University of Alberta

EXACT AND APPROXIMATION ALGORITHMS FOR TWO COMBINATORIAL OPTIMIZATION PROBLEMS

by

Zhong Li

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

Spring 2011
Edmonton, Alberta

Examining Committee

Guohui Lin, Computing Science

Randy Goebel, Computing Science

Yu Hu, Electrical and Computer Engineering

Abstract

In this thesis, we present our work on two combinatorial optimization problems. The first problem is the Bandpass problem, and we designed a linear time exact algorithm for the 3-column case. The other work is on the Complementary Maximal Strip Recovery problem, for which we designed a 3-approximation algorithm.

The Bandpass problem arises in designing optimal communication networks which aims to minimize the communication cost by packing data flows into groups. In the mathematical definition of this problem, a network is represented by a binary matrix, and we let a bandpass stand for a data group. Given a binary matrix A and a positive integer B, a bandpass is a sequence of B consecutive 1's in a column. Our goal is to maximize the non-overlapping bandpasses in A by doing row permutations. The general Bandpass problem is NP-hard and was claimed to be NP-hard when the number of columns is three. Previously, a Row-Stacking algorithm for the 3-column case was proposed to produce a solution that is at most one less than the optimum. We show that for any given matrix A of three columns with a bandpass number $B \geq 2$, our Remainder-Driven algorithm can achieve an optimal solution in linear time.

The Complementary Maximal Strip Recovery (CMSR) problem is formulated from research on genome comparison. In this problem, given two sequences G_1 and G_2 of n gene markers, in which each marker occurs exactly once, we aim to partition G_1 and G_2 into a set of common substrings of length at least 2 after deleting a minimum number of markers. This problem has been shown NP-hard and APX-complete, and there is no constant ratio approximation algorithm. We designed a 3-approximation algorithm for the CMSR problem with a performance ratio analysis done through a novel inverse sequential amortization.

Acknowledgements

I would like to thank Dr. Lin, my supervisor. His valuable guidance not only inspired me in this research but also accelerated my goal to become an excellent researcher.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter, we introduce the background of the two combinatorial optimization problems that we worked on. The Bandpass problem arises in designing optimal communication networks which aims to minimize the communication cost. In this problem, we process the given binary matrix representing a network by doing row permutations to find an optimal placement. The CMSR problem is formulated from the research on genome comparison. Given two sequences of gene markers, we aim to map all common substrings of length at least two between the two given sequences by deleting a minimum number of markers.

1.1 The Bandpass Problem

The Bandpass problem was first formulated and presented in the Annual INFORMS meeting, October 2004, USA [2, 15]. Given an $m \times n$ matrix A of binary elements {0,1} and a positive integer B, a bandpass is a sequence of non-zero entries of length B in a column of A. The goal of this combinatorial optimization problem is to find an optimal row permutation of the matrix A to maximize the number of bandpasses in it, in which no two bandpasses share any entries.

This problem arises in designing optimal communication networks which aims to minimize the communication cost by packing information flows into groups. In a communication network, a sending point has m information packages to be sent to n different destination points. We can use a matrix A of dimension $m \times n$ to represent the sending point, where $A = \{a_{ij}\}$, $i = 1, \cdots m$, $j = 1, \cdots n$ and $a_{ij} = 1$ if the information package i is not destined for the destination point j, otherwise $a_{ij} = 0$. Then, for a given positive integer B, the cable of the sending point could reduce the communication cost by merging the information packages to the same destination using Dense Wavelength Division Multiplexing (DWDM) technology [2]. More details of the application can be referred to [2].

The general Bandpass problem is proven to be NP-hard when $B \geq 2$ [2, 15]. Dr. Lin proposed the Row-Stacking algorithm which produces a solution that is at most one bandpass less than the optimum when the given matrix has three columns [15]. Based on the Row-Stacking scheme, we designed a Remainder-Driven algorithm that can generate the optimal solution in linear time.

1.2 The CMSR Problem

In comparative genomics [4], one of the first steps is to decompose two given genomes into synthetic blocks — segments of chromosomes that are deemed homologous in the two input genomes. Many decomposition methods have been proposed [21, 20, 10], but they are vulnerable to ambiguities and errors. A few years back, the Maximal Strip Recovery (MSR) problem was formulated for eliminating noise and ambiguities in genomic maps [23], which are isolated points that do not co-exist with other points [9, 22]. In the more precise formulation, we are given two genomic maps G_1 and G_2 each of n distinct gene markers, and we want to retain the maximum number of markers in both G_1 and G_2 such that the resultant subsequences, denoted as G_1^* and G_2^*, can be partitioned into the same set of maximal substrings of length greater than or equal to two. Each retained marker thus belongs to exactly one of these substrings, which can appear in the reversed and negated form and are taken as nontrivial chromosomal segments. The deleted markers are regarded as noise or errors.

The MSR problem, and its several close variants, have been shown NP-hard [19, 5, 8]. More recently, it is shown to be APX-complete [5, 12], admitting a 4-approximation algorithm [8, 18]. This approximation algorithm is a modification of an earlier heuristics for computing a maximum clique (and its complement, a maximum independent set) [5, 22, 16], to convert the MSR problem to computing the maximum independent set in t-interval graphs [6], which admits a $2t$-approximation [8, 17]. In our work, we investigate the complementary optimization goal to minimize the number of deleted markers — the complementary MSR problem, or CMSR for short. CMSR is certainly NP-hard, and was proven to be APX-hard recently [13]. Nevertheless, there is no known constant ratio approximation algorithm. We present here a 3-approximation algorithm which is the first constant ratio approximation algorithm.

Chapter 2

A Linear Time Exact Algorithm for the Bandpass Problem [1]

In this chapter, we introduce the exact algorithm for the 3-column Bandpass problem in two sections. In Section 2.1, we give several important definitions which are frequently referred to in the next section. Then in Section 2.2, we consider a complete set of subcases for a given general 3-column Bandpass instance. Through case by case analysis, we present the solution for every case and prove its optimality.

2.1 Preliminaries

Definition 1 (Bandpass [2]) *B consecutive non-zero entries in the same column of the given $m \times n$ binary matrix A form a bandpass.*

Definition 2 (Bandpass Problem [2, 3, 15]) *Given an $m \times n$ matrix A of binary elements $\{0,1\}$ and a positive integer B, find a row permutation of A with the maximum number of non-overlapping bandpasses.*

In our work, we deal with a matrix of three columns. There are at most eight types of rows: (0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0) and (1, 1, 1) in a three-column binary matrix. For convenience, we use m_1 to represent the rows of type (0, 0, 0), so that $|m_1|$ stands for the number of (0, 0, 0) rows in the given matrix. Furthermore, we let $|m_1| = q_1 B + r_1$, where q_1 and r_1 are the quotient and remainder respectively of dividing $|m_1|$ by B. The symbols for every row type are shown in Table 2.1. Since m_1 does not contribute to any bandpass, we can ignore them hereafter.

Definition 3 (Maximum) *MAX represents the maximum possible number of bandpasses we can achieve. It is the upper bound of the optimal solution. We compute MAX using Equation (2.1).*

$$MAX = \left\lfloor \frac{\#\ of\ 1's\ in\ column1}{B} \right\rfloor + \left\lfloor \frac{\#\ of\ 1's\ in\ column2}{B} \right\rfloor + \left\lfloor \frac{\#\ of\ 1's\ in\ column3}{B} \right\rfloor \quad (2.1)$$

[1] The main result in this chapter appears as "Z. Li, G. Lin. The three column Bandpass problem is solvable in linear time. *Theoretical Computer Science*. 412:281–299, 2011." [14]

Symbol	Row Type	Quantity Representation
m_1	0 0 0	$\|m_1\| = q_1$B+r_1
m_2	0 0 1	$\|m_2\| = q_2$B+r_2
m_3	0 1 1	$\|m_3\| = q_3$B+r_3
m_4	0 1 0	$\|m_4\| = q_4$B+r_4
m_5	1 1 0	$\|m_5\| = q_5$B+r_5
m_6	1 1 1	$\|m_6\| = q_6$B+r_6
m_7	1 0 1	$\|m_7\| = q_7$B+r_7
m_8	1 0 0	$\|m_8\| = q_8$B+r_8

Table 2.1: Symbols for row types.

P_1			P_2			P_3		
m_2	0 0 1	$\|m_2\|$	m_2	0 0 1	$\|m_2\|$	m_4	0 1 0	$\|m_4\|$
m_3	0 1 1	$\|m_3\|$	m_7	1 0 1	$\|m_7\|$	m_3	0 1 1	$\|m_3\|$
m_4	0 1 0	$\|m_4\|$	m_8	1 0 0	$\|m_8\|$	m_2	0 0 1	$\|m_2\|$
m_5	1 1 0	$\|m_5\|$	m_5	1 1 0	$\|m_5\|$	m_7	1 0 1	$\|m_7\|$
m_6	1 1 1	$\|m_6\|$	m_6	1 1 1	$\|m_6\|$	m_6	1 1 1	$\|m_6\|$
m_7	1 0 1	$\|m_7\|$	m_3	0 1 1	$\|m_3\|$	m_5	1 1 0	$\|m_5\|$
m_8	1 0 0	$\|m_8\|$	m_4	0 1 0	$\|m_4\|$	m_8	1 0 0	$\|m_8\|$
P_4			P_5			P_6		
m_4	0 1 0	$\|m_4\|$	m_8	1 0 0	$\|m_8\|$	m_8	1 0 0	$\|m_8\|$
m_5	1 1 0	$\|m_5\|$	m_7	1 0 1	$\|m_7\|$	m_5	1 1 0	$\|m_5\|$
m_8	1 0 0	$\|m_8\|$	m_2	0 0 1	$\|m_2\|$	m_4	0 1 0	$\|m_4\|$
m_7	1 0 1	$\|m_7\|$	m_3	0 1 1	$\|m_3\|$	m_3	0 1 1	$\|m_3\|$
m_6	1 1 1	$\|m_6\|$	m_6	1 1 1	$\|m_6\|$	m_6	1 1 1	$\|m_6\|$
m_3	0 1 1	$\|m_3\|$	m_5	1 1 0	$\|m_5\|$	m_7	1 0 1	$\|m_7\|$
m_2	0 0 1	$\|m_2\|$	m_4	0 1 0	$\|m_4\|$	m_2	0 0 1	$\|m_2\|$

Table 2.2: Six alternative solutions of the Row-Stacking Algorithm.

Though we might not always be able to find a row permutation with MAX bandpasses for a given matrix, the Row-Stacking algorithm can construct a placement with at least $MAX - 1$ bandpasses. The **Row-Stacking Algorithm** [15] is that, given an $m \times 3$ matrix A of binary elements $\{0,1\}$ and a positive integer B, place the rows according to one of the six schemes shown in Table 2.2. We can get at least $MAX - 1$ bandpasses in each of the six placements.

Ideally, for a given matrix, if we can make the 1's in every column consecutive, then we get MAX bandpasses. Actually, we can not make it most of the time, for example when we have all types of rows or we simply have m_3, m_5 and m_7 in the given matrix. The key point of the Row-Stacking algorithm is that, always make the 1's in two columns consecutive, and then the 1's in the other column may be consecutive or separated into two bands.

Take the placement P_1 in Table 2.2 for instance, apparently, we can acquire the maximum number of bandpasses in the first two columns in which the 1's are consecutive. Then whether we can get MAX bandpasses depends on the third column. If the 1's in it are consecutive, then it is an optimal placement. Otherwise, the 1's in the third column are broken into two bands of 1's, then we group the 1's into bandpasses successively, until the remainders are less than B. This procedure is

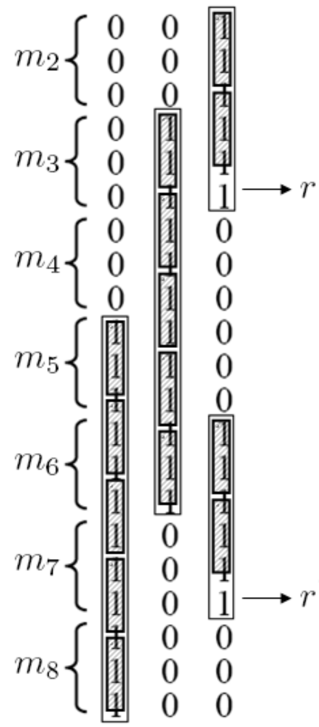

Figure 2.1: Extracting bandpasses in the Placement 1.

shown in Figure 2.1. Let the two remainders be r and $r^{'}$, respectively. If $r + r^{'} < B$, the placement P_1 is an optimal solution, because even there is a way to connect the two bands together, we can not get more bandpasses. Else, we have $r + r^{'} \geq B$ which means we get $MAX - 1$ bandpasses.

There are six alternative placements in the Row-Stacking scheme, the corresponding values of r and $r^{'}$ are shown in Table 2.3. Therefore, we can get MAX bandpasses as long as there exists one of the six placements with the corresponding $r + r^{'} < B$. Otherwise, we have $r + r^{'} \geq B$ for all the six placements.

Definition 4 (OPT) *We let OPT represent the number of bandpasses that we can get in the optimal placement. Thus, we have $OPT \leq MAX$.*

Based on the Row-Stacking algorithm, we further proposed the **Remainder-Driven Algorithm** which can get the optimal solution in linear time. For a given instance (A, B), we first use the Row-Stacking algorithm. If it can not achieve MAX bandpasses, then the Remainder-Driven algorithm will further process the given matrix to reach the optimum. Specifically, it first tries to construct an optimal placement by using the matrix consists of the remainder part of each type of rows, and put the integral part of each type of rows aside, since the matrix consists of the integral parts certainly can produce a maximum number of bandpasses by placing the rows of the same type together. If it can not achieve MAX bandpasses, then it will make use of the integral part if there is any. Otherwise, we will prove that $OPT = MAX - 1$.

Placement	r	r'
P_1	$(r_2 + r_3)\%\text{B}$	$(r_6 + r_7)\%\text{B}$
P_2	$(r_2 + r_7)\%\text{B}$	$(r_6 + r_3)\%\text{B}$
P_3	$(r_4 + r_3)\%\text{B}$	$(r_6 + r_5)\%\text{B}$
P_4	$(r_4 + r_5)\%\text{B}$	$(r_6 + r_3)\%\text{B}$
P_5	$(r_8 + r_7)\%\text{B}$	$(r_6 + r_5)\%\text{B}$
P_6	$(r_8 + r_5)\%\text{B}$	$(r_6 + r_7)\%\text{B}$

Table 2.3: Remaider values of the six placements.

Placement	Inequation
P_1	$(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$
P_2	$(r_2 + r_7)\%B + (r_6 + r_3)\%B \geq B$
P_3	$(r_4 + r_3)\%B + (r_6 + r_5)\%B \geq B$
P_4	$(r_4 + r_5)\%B + (r_6 + r_3)\%B \geq B$
P_5	$(r_8 + r_7)\%B + (r_6 + r_5)\%B \geq B$
P_6	$(r_8 + r_5)\%B + (r_6 + r_7)\%B \geq B$

Table 2.4: The conditions when the Row-Stacking algorithm can not achieve MAX bandpasses.

2.2 The Algorithm and Proofs of Optimality

In this section, we introduce how the Remainder-Driven algorithm works based on the Row-Stacking scheme. When the Row-Stacking algorithm can not achieve MAX bandpasses, we have $r + r' \geq B$ for all six placements. It is shown in Table 2.4 according to Table 2.3.

Generally speaking, for a given case which can not be optimized by the Row-Stacking scheme, the Remainder-Driven algorithm will first try to find a permutation with MAX bandpasses, but when it can not make it, we will prove that the solution with $MAX - 1$ bandpasses is already optimal. Based on this idea, we consider a complete set of four subcases for a given 3-column Bandpass instance, which is shown in Figure 2.2.

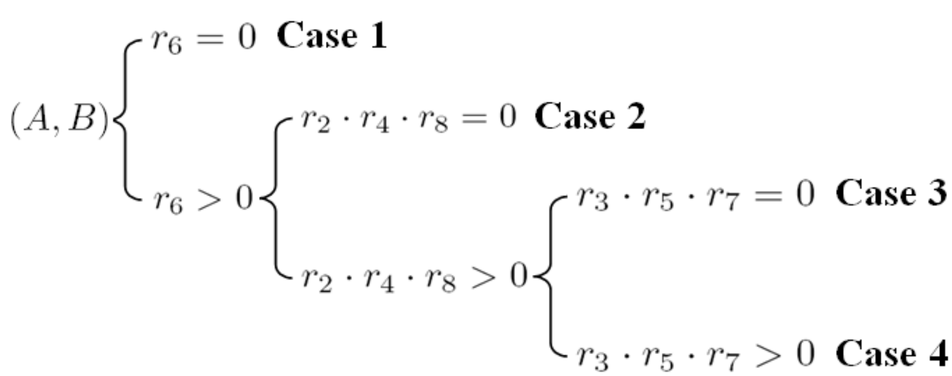

Figure 2.2: The skeleton of the Remaider-Driven algorithm.

2.2.1 Two Base Cases

Before stepping into the four subcases, we introduce two base cases first. Base case 1: we only have m_3, m_5, m_7 in the given matrix; Base case 2: the complement of Base case 1, where we only have m_2, m_4, m_6, m_8 in the given matrix. We present three lemmas for these two base cases which will be referenced frequently by later proofs of optimality.

Lemma 1 *If* $\forall m_i \in \{m_2, m_4, m_6, m_8\}$*, s.t.* $|m_i| = 0$*, and* $r_3 + r_5 \geq B$*,* $r_3 + r_7 \geq B$*,* $r_5 + r_7 \geq B$*,* $r_3 + r_5 + r_7 < 2B$*, then* $OPT = MAX - 1$*.*

PROOF. We first show that if one of q_3, q_5, q_7 is zero, then $OPT = MAX - 1$. Without loss of generality, assume $q_7 = 0$. From the definitions, we have $MAX = 2q_3 + 2q_5 + 3$, and if there is an optimal row placement P^* achieving MAX bandpasses, then there are $q_5 + 1$, $q_3 + q_5 + 1$, $q_3 + 1$ bandpasses in each column of P^*.

Since the total number of rows is $|m_3| + |m_5| + |m_7| < (q_3 + q_5 + 2)B$, we conclude that in P^* there must be some bandpasses in the first column overlap (that is, share rows) with bandpasses in the third column. But none of the bandpasses in the first column would overlap with two bandpasses in the third column due to the non-existence of (1, 1, 1)-rows. Equivalently, there are pairs of overlapping bandpasses, one in the first column and the other in the third column. These overlapping regions, consisting of solely (1, 0, 1)-rows, separate the rows of P^* into chunks. For every bandpass (in the first or the third column) participating in the overlapping pairs, if a part of it belongs to a chunk, then the bandpass is said to belong to that chunk. Because there are $q_3 + q_5 + 1$ bandpasses in the second column of P^*, we conclude that there is (at least) one chunk in which the number of bandpasses in the second column is strictly less than the total number of bandpasses in the first and the third columns. Recall that inside a chunk, no bandpass in the first column would overlap with any bandpass in the third column. It follows that in this chunk strictly greater than $B - r_7$ 1's in the second column are not involved in any bandpasses. Nevertheless, in order to achieve MAX bandpasses, at most $r_3 + r_5 - B$ 1's in the second column of P^* can sit outside of generated bandpasses. This is a contradiction since $r_3 + r_5 - B < B - r_7$. This implies that $OPT = MAX - 1$.

When all q_3, q_5, q_7 are positive, we assume that $OPT = MAX = 2q_3 + 2q_5 + 2q_7 + 3$ is achieved by a row placement P^*. Then we examine where the topmost bandpass is in P^*. Assume without loss of generality that it occurs in the first column, then the second topmost bandpass should not occur in the first column, for otherwise at least B 1's would not be involved in any generated bandpasses in P^*. Again assume without loss of generality that the second topmost bandpass occurs in the second column. These two bandpasses must overlap for the same reason above. Due to the non-existence of (1, 1, 1)-rows, the third topmost bandpass does not overlap with the topmost bandpass. Suppose there are l (1, 0, 1)-rows in the topmost bandpass. If we take away the B rows in the topmost bandpass from the instance, the resultant new instance I' contains $|m_3'| = |m_3|$ (0, 1, 1)-rows, $|m_5'| = (m_5 - B + l)$ (1, 1, 0)-rows, and $|m_7'| = (|m_7| - l)$ (1, 0, 1)-rows. Apparently

$l \leq r_3 + r_7 - B$, implying that $r_7^{'} = r_7 - l \geq B - r_3 > 0$, $r_5^{'} = r_5 + l \leq r_3 + r_5 + r_7 - B < B$, $r_3^{'} + r_5^{'} = r_3 + r_5 + l \geq B$, $r_5^{'} + r_7^{'} = r_5 + r_7 \geq B$, $r_3^{'} + r_7^{'} = r_3 + r_7 - l \geq B$, and $r_3^{'} + r_5^{'} + r_7^{'} = r_3 + r_5 + r_7 \leq 2B$. This new instance $I^{'}$ satisfies the premises in the lemma, with B fewer rows than the original instance and again with $OPT(I^{'}) = MAX(I^{'})$.

It follows that if we were to apply the same reduction procedure, we will eventually end up with an instance that satisfies the premises in the lemma and with $OPT = MAX$, but one of q_3, q_5, q_7 is zero. This is a contradiction to the fact proven in the first half. Therefore, for all instances satisfying the premises, their optimal row placement contains only $MAX - 1$ bandpasses, suggesting that the Row-Stacking solutions are already optimal. This proves the lemma. □

Corollary 2 *If* $\forall m_i \in \{m_2, m_4, m_8\}$, *s.t.* $|m_i| = 0$, *and* $|m_6| = r_6$, $r_3 + r_6 < B$, $r_5 + r_6 < B$, $r_7 + r_6 < B$, $r_3 + r_5 + r_6 \geq B$, $r_3 + r_7 + r_6 \geq B$, $r_5 + r_7 + r_6 \geq B$, $r_3 + r_5 + 2r_6 + r_7 < 2B$, *then* $OPT = MAX - 1$.

PROOF. For a given instance I, assume there is a placement P^* with maximum bandpasses. Now we want to construct a new instance $I^{'}$ from I by changing m_6 rows to m_3,m_5,m_7 without loss of any bandpass in I. For each (1, 1, 1)-row, if it participated in three bandpasses, we change it to the corresponding two of $\{m_3, m_5, m_7\}$; else if it participated in one or two bandpasses, we change it to the corresponding m_3, m_5 or m_7; else, remove this row since it didn't participate in any bandpass. Suppose we changed $r_6^{'}$ rows to m_3, $r_6^{''}$ rows to m_5 and $r_6^{'''}$ rows to m_7. Then in $I^{'}$, $r_3^{'} = r_3 + r_6^{'} < B$, $r_5^{'} = r_5 + r_6^{''} < B$, $r_7^{'} = r_7 + r_6^{'''} < B$, where $r_6^{'} + r_6^{''} \geq B - (r_3 + r_5)$, $r_6^{'} + r_6^{'''} \geq B - (r_3 + r_7)$, $r_6^{''} + r_6^{'''} \geq B - (r_5 + r_7)$ and $r_6^{'} + r_6^{''} + r_6^{'''} \leq 2r_6$. Thus $r_3^{'} + r_5^{'} + r_7^{'} \leq r_3 + r_5 + 2r_6 + r_7 < 2B$, $r_3^{'} + r_5^{'} \geq B$, $r_3^{'} + r_7^{'} \geq B$, $r_5^{'} + r_7^{'} \geq B$. By Lemma 1, we have $OPT(I^{'}) = MAX(I^{'}) - 1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I) = MAX - 1$. □

Lemma 3 *When* $\forall m_i \in \{m_3, m_5, m_7\}$, *s.t.* $|m_i| = 0$, *and* $r_2 + r_6 \geq B$, $r_4 + r_6 \geq B$, $r_8 + r_6 \geq B$, $r_2 + r_4 + r_6 < 2B$, $r_4 + r_8 + r_6 < 2B$, $r_2 + r_8 + r_6 < 2B$, *if* $r_2 + r_4 + r_8 + 2r_6 < 3B$ *or* $q_2, q_4, q_8 = 0$, *then* $OPT = MAX - 1$.

PROOF. From the lemma premises and Eq. (2.1), we have $MAX = q_2 + q_4 + q_8 + 3q_6 + 3$, and if there were an optimal row placement P^* achieving MAX bandpasses, then there are $q_8 + q_6 + 1$, $q_4 + q_6 + 1$, $q_2 + q_6 + 1$ bandpasses in the first, second, third columns of P^*, respectively.

Since (0, 0, 1)-rows are not involved in any bandpasses formed in the first and the second columns, these bandpasses must overlap at least $(q_8 + q_6 + 1 + q_4 + q_6 + 1)B - (|m_4| + |m_6| + |m_8|) = q_6 B + 2B - r_4 - r_6 - r_8$ rows. These rows have 1 in both the first and the second columns, and thus must be (1, 1, 1)-rows. If one of these rows is involved in a bandpass generated in the third column, that is, there are three bandpasses, one from each column, overlapping at a (1, 1, 1)-row, then there are B consecutive rows of type m_6 in the optimal placement (which includes the shared

Placement	r	$r^{'}$
P_1	$(r_2+r_3)\%B$	r_7
P_2	$(r_2+r_7)\%B$	r_3
P_3	$(r_4+r_3)\%B$	r_5
P_4	$(r_4+r_5)\%B$	r_3
P_5	$(r_8+r_7)\%B$	r_5
P_6	$(r_8+r_5)\%B$	r_7

Table 2.5: Remaider values of the six placements when $r_6 = 0$.

(1, 1, 1)-row). Removing these B consecutive (1, 1, 1)-rows, on one hand we obtain a reduced instance $I^{'}$ for which all the premises hold except that q_6 decreases by 1; on the other hand, we obtain a row placement for $I^{'}$ achieving $MAX(I^{'}) = MAX(I) - 3$ bandpasses. It follows that by repeatedly reducing the instance whenever possible, we may assume without loss of generality that none of the $q_6B + 2B - r_4 - r_6 - r_8$ (1, 1, 1)-rows is involved in any bandpasses in the third column. Consequently, the maximum possible number of bandpasses in the third column becomes

$$\left\lfloor \frac{|m_2| + |m_6| - (q_6B + 2B - r_4 - r_6 - r_8)}{B} \right\rfloor = \left\lfloor q_2 + \frac{r_2 + r_4 + r_8 + 2r_6 - 2B}{B} \right\rfloor \quad (2.2)$$

Therefore, if $r_2 + r_4 + r_8 + 2r_6 < 3B$, this maximum possible number is q_2, a contradiction to $q_2 + q_6 + 1 = q_2 + 1$.

Note that $r_2+r_4+r_8+2r_6 < 4B$. Therefore, if $q_2, q_4, q_8 = 0$, this maximum possible number is $1 \le q_6+1$ and the equality holds only when $q_6 = 0$. In such a case, the bandpass in the third column may overlap with at most one of the bandpass in the first column and the bandpass in the second column, a contradiction to the fact that these three bandpasses must pairwise overlap. Hence, for all instances satisfying the premises, their optimal row placement contains only $MAX - 1$ bandpasses. This proves the lemma. □

2.2.2 Case 1 ($\mathbf{r_6 = 0}$)

We separate this case into two disjoint subcases according to whether $q_6 = 0$. One can verify that since $r_6 = 0$, Table 2.3 reduces to Table 2.5. If there exists a permutation satisfying $r + r^{'} < B$, then it is the optimal solution. Otherwise we have $(r_2+r_3)\%B+r_7 \ge B$, $(r_2+r_7)\%B+r_3 \ge B$, $(r_4+r_3)\%B+r_5 \ge B$, $(r_4+r_5)\%B+r_3 \ge B$, $(r_8+r_5)\%B+r_7 \ge B$ and $(r_8+r_7)\%B+r_5 \ge B$ simultaneously.

- **Case 1.1** $\mathbf{q_6 = 0}$

  In this case, since both $q_6 = 0$ and $r_6 = 0$, we have no (1, 1, 1)-row in the given matrix. We separate Case 1.1 into the following subcases.

  - If $\exists m_i \in \{m_2, m_4, m_8\}$, s.t. $|m_i| \ge B$.

m_2	0	0	1	$\|m_2\|$		$m_2^{'}$	0	0	1	$\|m_2\| - (r_2 + r_3)\%B$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						$m_2^{''}$	0	0	1	$(r_2 + r_3)\%B$

Table 2.6: Placement 1.1.01

m_2	0	0	1	r_2		$m_2^{'}$	0	0	1	$B - r_3$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						$m_2^{''}$	0	0	1	$r_2 + r_3 - B$

Table 2.7: Placement 1.1.02

Since m_2, m_4, m_8 can be symmetrically discussed, without loss of generality, here we assume $|m_2| \geq B$. We can get MAX bandpasses by using Placement 1.1.01 shown in Table 2.6, in which we take $(r_2 + r_3)\%B$ (0, 0, 1)-rows down to the bottom of the matrix to make the first band in the third column of length a multiple of B, together with that the 1's in the first and second columns are consecutive, thus the resultant matrix is optimal. **(In Placement 1.1.01, $\mathbf{m_2}$, $\mathbf{m_2^{'}}$ and $\mathbf{m_2^{''}}$ represent the same type of rows. We are indicating that the set of $\mathbf{m_2}$ rows in the left matrix is splitted into two sets, $\mathbf{m_2^{'}}$ and $\mathbf{m_2^{''}}$, in the resulting matrix to the right. We adopt this way of representation in every placement in the sequel.)**

– Else if $r_2 + r_3 \geq B$ (We can symmetrically consider the cases when $r_4 + r_5 \geq B$ or $r_8 + r_7 \geq B$).

  In this case, we can use a similar way as what we did in Placement 1.1.01 to get MAX bandpasses, and for the same reason, the resultant matrix in Placement 1.1.02 is optimal, which is shown in Table 2.7.

– Else if $r_2 + r_3 + r_5 + r_7 + r_8 \geq 2B$ (We can symmetrically consider the cases when $r_2 + r_3 + r_5 + r_7 + r_4 \geq 2B$ or $r_4 + r_3 + r_5 + r_7 + r_8 \geq 2B$).

  We can get MAX bandpasses by using Placement 1.1.03 shown in Table 2.8. Since we have $r_2 + r_3 < B$, by taking $B - (r_2 + r_3)$ (1, 0, 1)-rows up, we can make the first band in the third column of length a multiple of B, but meanwhile, these $B - (r_2 + r_3)$ (1, 0, 1)-rows are not involved in any bandpass in the first column. Since we have

						m_7''	1	0	1	$B-(r_2+r_3)$
m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4	$\Rightarrow$	m_4	0	1	0	r_4
m_5	1	1	0	r_5		m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7'	1	0	1	$r_7+r_2+r_3-B$

Table 2.8: Placement 1.1.03

$r_2 + r_3 + r_5 + r_7 + r_8 \geq 2B$, which means we can get one bandpass that we are supposed to get by using the remaining 1's in the first column. Thus, Placement 1.1.03 is optimal.

– Else if $r_2 + r_4 + r_8 + r_3 + r_5 + r_7 < 2B$.

  Here we claim that $OPT = MAX - 1$. We can reduce the current instance I to I' by changing m_2 ($|m_2| = r_2$) to m_3, m_4 ($|m_4| = r_4$) to m_5 and m_8 ($|m_8| = r_8$) to m_7. Then in I', $r_3' = r_2 + r_3 < B$, $r_5' = r_4 + r_5 < B$, $r_7' = r_7 + r_8 < B$, $r_3' + r_5' \geq B$, $r_3' + r_7' \geq B$, $r_5' + r_7' \geq B$ and $r_3' + r_5' + r_7' < 2B$. By Lemma 1, we can conclude that $OPT(I') = MAX(I') - 1$. Since $MAX(I') = MAX(I)$ and $OPT(I') \geq OPT(I)$, we have $OPT(I) = MAX(I) - 1$.

– Else, $r_2 + r_4 + r_8 + r_3 + r_5 + r_7 \geq 2B$.

  In this case, if $\exists m_i \in \{m_3, m_5, m_7\}$ s.t. $|m_i| \geq B$, we can get MAX bandpasses. Since m_3, m_5, m_7 can be symmetrically discussed, without loss of generality, here we assume $|m_3| \geq B$. Because $r_7 + r_8 < B$, together with the premise of this case, we will have $r_2 + r_3 + r_4 + r_5 > B$. Then if $r_2 + r_3 + r_4 < B$, by making use of extra B (0, 1, 1)-rows, we can get MAX bandpasses by using Placement 1.1.04 shown in Table 2.9. In this case, we are supposed to get one bandpass in the first column, two in the second column and two in the third column, so Placement 1.1.04 is optimal. Else we will have $r_2 + r_3 + r_4 \geq B$, similar to Placement 1.1.04, we use m_4 to adjust the matrix, which is illustrated by Placement 1.1.05 in Table 2.10, and for the same reason, Placement 1.1.05 is optimal. Else, $|m_3|, |m_5|, |m_7| < B$, then we are supposed to get one band in each column. Because $r_2 + r_3 + r_5 + r_7 + r_4 < 2B$, $r_2 + r_3 + r_5 + r_7 + r_8 < 2B$ and $r_4 + r_3 + r_5 + r_7 + r_8 < 2B$, which means the three bandpasses we are supposed to get are pairwise overlapping. This is impossible because they will form a circle. Thus we can get $MAX - 1$ bandpasses.

- **Case 1.2** $\quad \mathbf{q_6 > 0}$

In this case, we have at least B (1, 1, 1)-rows. First, we try to get an optimal placement without using m_6 in the way which was introduced in **Case 1.1**. If we can not make it, then

m_2	0	0	1	r_2	$\Rightarrow$	m_2	0	0	1	r_2
m_3	0	1	1	r_3+B		$m_3^{'}$	0	1	1	$B-r_2$
m_4	0	1	0	r_4		$m_5^{'}$	1	1	0	$r_2+r_3+r_4+r_5-B$
m_5	1	1	0	r_5		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_3^{''}$	0	1	1	r_2+r_3
						m_4	0	1	0	r_4
						$m_5^{''}$	1	1	0	$B-(r_2+r_3+r_4)$

Table 2.9: Placement 1.1.04

m_2	0	0	1	r_2	$\Rightarrow$	m_2	0	0	1	r_2
m_3	0	1	1	r_3+B		$m_3^{'}$	0	1	1	$B-r_2$
m_4	0	1	0	r_4		$m_4^{''}$	0	1	0	$r_2+r_3+r_4-B$
m_5	1	1	0	r_5		m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						$m_3^{''}$	0	1	1	r_2+r_3
						$m_4^{'}$	0	1	0	$B-(r_2+r_3)$

Table 2.10: Placement 1.1.05

we will have $|m_2| = r_2$, $|m_4| = r_4$, $|m_8| = r_8$, $r_2 + r_3 < B$, $r_4 + r_5 < B$, $r_8 + r_7 < B$, $r_2+r_3+r_5+r_7+r_8 < 2B$, $r_2+r_3+r_5+r_7+r_4 < 2B$ and $r_4+r_3+r_5+r_7+r_8 < 2B$. Now we make use of extra B m_6 to get MAX bandpasses, and we separate this case into the following subcases.

- If $\exists r_i \in \{r_2, r_4, r_8\}$, s.t. $r_i = 0$.

  Here we can get MAX bandpasses. For m_2, m_4, m_8 can be symmetrically discussed, without loss of generality, assume $r_4 = 0$. Thus $|m_4| = r_4 = 0$, there is no (0, 1, 0)-row in the matrix. In Placement 1.2.01 which is shown in Table 2.11, we use m_6 to adjust the matrix, then the 1's in the first two columns are consecutive, and the second band in the third column is of length exactly B. Thus, Placement 1.2.01 is optimal.

m_2	0	0	1	r_2	$\Rightarrow$	m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_5	1	1	0	r_5		$m_6^{''}$	1	1	1	r_7
m_6	1	1	1	B		m_5	1	1	0	r_5
m_7	1	0	1	r_7		$m_6^{'}$	1	1	1	$B-r_7$
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_8	1	0	0	r_8

Table 2.11: Placement 1.2.01

m_2	0	0	1	r_2	$\Rightarrow$	m_8	1	0	0	r_8
m_3	0	1	1	$r_3 + B$		m_7	1	0	1	r_7
m_4	0	1	0	r_4		m_6''	1	1	1	$B - (r_7 + r_8)$
m_5	1	1	0	r_5		m_3''	0	1	1	r_8
m_6	1	1	1	B		m_4	0	1	0	r_4
m_7	1	0	1	r_7		m_5'	1	1	0	$r_5 + r_7 + r_8 - B$
m_8	1	0	0	r_8		m_3'	0	1	1	$r_3 + B - r_8$
						m_2	0	0	1	r_2
						m_6'	1	1	1	$r_7 + r_8$
						m_5''	1	1	0	$B - (r_7 + r_8)$

Table 2.12: Placement 1.2.02

m_2	0	0	1	r_2	$\Rightarrow$	m_8	1	0	0	r_8
m_3	0	1	1	r_3		m_7	1	0	1	r_7
m_4	0	1	0	r_4		m_6''	1	1	1	$B - (r_7 + r_8)$
m_5	1	1	0	r_5		m_3''	0	1	1	r_8
m_6	1	1	1	B		m_4	0	1	0	r_4
m_7	1	0	1	r_7		m_5'	1	1	0	$r_5 + r_7 + r_8 - B$
m_8	1	0	0	r_8		m_3'	0	1	1	$r_3 - r_8$
						m_2	0	0	1	r_2
						m_6'	1	1	1	$r_7 + r_8$
						m_5''	1	1	0	$B - (r_7 + r_8)$

Table 2.13: Placement 1.2.03

- Else if $\exists m_i \in \{m_3, m_5, m_7\}$, s.t. $|m_i| > B$.

  In this case, we can get MAX bandpasses. Since m_3, m_5, m_7 can be symmetrically discussed, without loss of generality, we assume $|m_3| > B$. Then we can make use of extra B (0, 1, 1)-rows to get an optimal solution, which is shown in Placement 1.2.02. We are supposed to get two bandpasses in the first column, three in the second and third columns, so Placement 1.2.02 is optimal.

- Else if $r_3 \geq r_8$ (We can symmetrically consider the cases when $r_5 \geq r_2$ or $r_7 \geq r_4$).

  Similar to Placement 1.2.02, we can get MAX bandpasses. The optimal solution is shown in Placement 1.2.03.

- Else if $r_3 + r_5 + r_7 \geq B$.

  In this case, we can get MAX bandpasses. Since we have $r_5 \geq B - (r_3 + r_7)$, in Placement 1.2.04, we make use of both m_5 and m_6 to adjust the matrix, then we can get two bandpasses in each column that we are supposed to get. Thus, Placement 1.2.04 which is shown in Table 2.14 is optimal.

- Else, $r_3 + r_5 + r_7 < B$.

$$
\begin{array}{ccccc}
m_2 & 0 & 0 & 1 & r_2 \\
m_3 & 0 & 1 & 1 & r_3 \\
m_4 & 0 & 1 & 0 & r_4 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_6 & 1 & 1 & 1 & B \\
m_7 & 1 & 0 & 1 & r_7 \\
m_8 & 1 & 0 & 0 & r_8
\end{array}
\quad \Rightarrow \quad
\begin{array}{ccccc}
m_2 & 0 & 0 & 1 & r_2 \\
m_3 & 0 & 1 & 1 & r_3 \\
m_6^{'} & 1 & 1 & 1 & r_7 \\
m_5^{'} & 1 & 1 & 0 & B-(r_3+r_7) \\
m_8 & 1 & 0 & 0 & r_8 \\
m_5^{''} & 1 & 1 & 0 & r_3+r_5+r_7-B \\
m_4 & 0 & 1 & 0 & r_4 \\
m_6^{''} & 1 & 1 & 1 & B-r_7 \\
m_7 & 1 & 0 & 1 & r_7
\end{array}
$$

Table 2.14: Placement 1.2.04

Placement	r	$r^{'}$
P_1	$(r_2+r_3)\%B$	$(r_6+r_7)\%B$
P_2	$(r_2+r_7)\%B$	$(r_6+r_3)\%B$
P_3	r_3	$(r_6+r_5)\%B$
P_4	r_5	$(r_6+r_3)\%B$
P_5	$(r_8+r_7)\%B$	$(r_6+r_5)\%B$
P_6	$(r_8+r_5)\%B$	$(r_6+r_7)\%B$

Table 2.15: Remaider values of the six placements when $r_4 = 0$.

In this case, $OPT = MAX - 1$. We can reduce the original instance I to $I^{'}$ by changing m_3, m_5, m_7 to m_6, then in $I^{'}$, we have $r_6^{'} = r_3 + r_5 + r_7$, $|m_2| = r_2$, $|m_4| = r_4$, $|m_8| = r_8$, $r_2 + r_6^{'} + r_4 < 2B$, $r_2 + r_6^{'} + r_8 < 2B$ and $r_4 + r_6^{'} + r_8 < 2B$. By Lemma 3, we can conclude that $OPT(I^{'}) = MAX(I^{'}) - 1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, we have $OPT(I) = MAX(I) - 1$.

2.2.3 Case 2 ($\mathbf{r_6 > 0}$ and $\mathbf{r_2 \cdot r_4 \cdot r_8 = 0}$)

In this case, since we have $\exists r_i \in \{r_2, r_4, r_8\}$, s.t. $r_i = 0$, without loss of generality, we assume $r_4 = 0$ hereafter, and then Table 2.3 will reduce to Table 2.15. If there exists a permutation satisfying $r + r^{'} < B$, it is the optimal solution. Otherwise, we have $(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$, $(r_2 + r_7)\%B + (r_6 + r_3)\%B \geq B$, $r_3 + (r_6 + r_5)\%B \geq B$, $r_5 + (r_6 + r_3)\%B \geq B$, $(r_8 + r_5)\%B + (r_6 + r_7)\%B \geq B$ and $(r_8 + r_7)\%B + (r_6 + r_5)\%B \geq B$ simultaneously. In the following, we separate this case into two disjoint subcases according to whether $q_4 = 0$.

- **Case 2.1** $\mathbf{q_4 = 0}$

  In this case, since both $q_4 = 0$ and $r_4 = 0$, it means that we have no (0, 1, 0)-row in the given matrix. We separate this case into the following subcases.

  - If $|m_2| + r_7 \geq B$ (We can symmetrically consider the case when $|m_8| + m_3 \geq B$).

    In Placement 2.1.01, we can make the 1's in the first two columns consecutive and one of the two bands in the third column of length exactly B. Thus Placement 2.1.01 is

$$
\begin{array}{lcccl}
m_2 & 0 & 0 & 1 & r_2 + B \\
m_3 & 0 & 1 & 1 & r_3 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_6 & 1 & 1 & 1 & r_6 \\
m_7 & 1 & 0 & 1 & r_7 \\
m_8 & 1 & 0 & 0 & r_8
\end{array}
\quad \Rightarrow \quad
\begin{array}{lcccl}
m_2 & 0 & 0 & 1 & |m_2| + r_7 - B \\
m_3 & 0 & 1 & 1 & r_3 \\
m_6 & 1 & 1 & 1 & r_6 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_8 & 1 & 0 & 0 & r_8 \\
m_7 & 1 & 0 & 1 & r_7 \\
m_2^{''} & 0 & 0 & 1 & B - r_7
\end{array}
$$

Table 2.16: Placement 2.1.01

$$
\begin{array}{lcccl}
m_2 & 0 & 0 & 1 & r_2 \\
m_3 & 0 & 1 & 1 & r_3 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_6 & 1 & 1 & 1 & |m_6| \\
m_7 & 1 & 0 & 1 & r_7 \\
m_8 & 1 & 0 & 0 & r_8
\end{array}
\quad \Rightarrow \quad
\begin{array}{lcccl}
m_2 & 0 & 0 & 1 & r_2 \\
m_3 & 0 & 1 & 1 & r_3 \\
m_6^{''} & 1 & 1 & 1 & B - r_3 \\
m_7 & 1 & 0 & 1 & r_7 \\
m_6^{'} & 1 & 1 & 1 & |m_6| + r_3 - B \\
m_5 & 1 & 1 & 0 & r_5 \\
m_8 & 1 & 0 & 0 & r_8
\end{array}
$$

Table 2.17: Placement 2.1.02

optimal.

– Else if $|m_6| + r_3 \geq B$ (We can symmetrically consider the cases when $|m_6| + r_5 \geq B$ or $|m_6| + r_7 \geq B$).

  In this case, we use m_6 to adjust the matrix which is shown in Placement 2.1.02. The 1's in the first and third columns are consecutive, and in the second column the first band is of length B. Thus we can get MAX bandpasses in Placement 2.1.02.

– Else if $r_2 + r_3 + r_6 \geq B$.

  For $|m_6| = r_6$ and $r_6 + r_3 < B$, together with the premise of this subcase, we have $r_2 \geq B - (r_6 + r_3)$. In Placement 2.1.03, the 1's in the first two columns are consecutive, and one band in the third column is of length exactly B. Thus Placement 2.1.03 is optimal.

– Else if $r_5 + r_6 + r_8 \geq B$.

$$
\begin{array}{lcccl}
m_2 & 0 & 0 & 1 & r_2 \\
m_3 & 0 & 1 & 1 & r_3 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_6 & 1 & 1 & 1 & r_6 \\
m_7 & 1 & 0 & 1 & r_7 \\
m_8 & 1 & 0 & 0 & r_8
\end{array}
\quad \Rightarrow \quad
\begin{array}{lcccl}
m_2^{'} & 0 & 0 & 1 & B - (r_6 + r_3) \\
m_3 & 0 & 1 & 1 & r_3 \\
m_6 & 1 & 1 & 1 & r_6 \\
m_5 & 1 & 1 & 0 & r_5 \\
m_8 & 1 & 0 & 0 & r_8 \\
m_7 & 1 & 0 & 1 & r_7 \\
m_2^{''} & 0 & 0 & 1 & r_2 + r_3 + r_6 - B
\end{array}
$$

Table 2.18: Placement 2.1.03

m_2	0	0	1	r_2		m_8''	1	0	0	$B-(r_6+r_5)$
m_3	0	1	1	r_3		m_5	1	1	0	r_5
m_5	1	1	0	r_5		m_6	1	1	1	r_6
m_6	1	1	1	r_6	$\Rightarrow$	m_3	0	1	1	r_3
m_7	1	0	1	r_7		m_2	0	0	1	r_2
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_8	1	0	0	$r_8+r_6+r_5-B$

Table 2.19: Placement 2.1.04

m_2	0	0	1	r_2		m_7''	1	0	1	$B-(r_2+r_3+r_6)$
m_3	0	1	1	r_3		m_2	0	0	1	r_2
m_5	1	1	0	r_5		m_3	0	1	1	r_3
m_6	1	1	1	r_6	$\Rightarrow$	m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		m_8	1	0	0	r_8
						m_7'	1	0	1	$r_2+r_3+r_6+r_7-B$

Table 2.20: Placement 2.1.05

Similar to the previous subcase, Placement 2.1.04 shown in Table 2.19 is optimal.

– Else if $r_2+r_3+r_5+2r_6+r_7+r_8 \geq 2B$

  In Placement 2.1.03, if $r_2+r_3+r_6 < B$, we can take some (1, 0, 1)-rows up to make the first band in the third column of length B, and the corresponding 1's are not involved in any bandpass in the first column. For the premise of this subcase, we can get one bandpass that we are supposed to get in the first column. Thus Placement 2.1.05 shown in Table 2.20 is optimal.

– Else, we have $OPT = MAX - 1$.

  Now $|m_2| = r_2$ and $r_2+r_3+r_6 < B$; $|m_6| = r_6$ and $r_6+r_3 < B$, $r_6+r_5 < B$, $r_6+r_7 < B$; $|m_8| = r_8$ and $r_5+r_6+r_8 < B$, $r_2+r_3+r_5+2r_6+r_7+r_8 < 2B$. From the current instance I, we change m_2 to m_3, m_8 to m_5. Then in the new instance I', we have $r_3' = r_2+r_3$, $r_5' = r_5+r_8$, $r_7' = r_7$, satisfying that $r_3'+r_5'+r_7'+2r_6 < 2B$, and $r_3'+r_7'+r_6 = r_2+r_3+r_6+r_7 \geq B$, similarly $r_3'+r_5'+r_6 \geq B$, $r_5'+r_7'+r_6 \geq B$. By Corollary 2, we have $OPT(I') = MAX(I') - 1$. Since $MAX(I') = MAX(I)$ and $OPT(I') \geq OPT(I)$, then $OPT(I) = MAX(I) - 1$.

- **Case 2.2** $\quad \mathbf{q_4 > 0}$

In this case, we can make use of B (0, 1, 0)-rows. First, we try to get an optimal placement without using m_4 in the way which was introduced in **Case 2.1**. If we can not make it, then we will have $|m_2| = r_2$, $|m_8| = r_8$, $|m_6| = r_6$, $r_2+r_3 < B$, $r_2+r_7 < B$, $r_8+r_5 < B$,

m_2	0	0	1	r_2		$m_4^{'}$	0	1	0	r_5
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	B		m_6	1	1	1	r_6
m_5	1	1	0	r_5	$\Rightarrow$	m_7	1	0	1	r_7
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		$m_4^{''}$	0	1	0	$B - r_5$

Table 2.21: Placement 2.2.01

						$m_4^{'}$	0	1	0	r_5
m_2	0	0	1	r_2		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_6	1	1	1	r_6
m_4	0	1	0	B		m_2	0	0	1	r_2
m_5	1	1	0	r_5	$\Rightarrow$	m_7	1	0	1	r_7
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		$m_4^{''}$	0	1	0	$B - r_5$

Table 2.22: Placement 2.2.02

$r_8 + r_7 < B$, $r_6 + r_3 < B$, $r_6 + r_5 < B$, $r_6 + r_7 < B$. Now we are supposed to get one bandpass in each column. We separate this case into the following subcases.

- If $r_3 + r_6 + r_7 \geq B$ (We can symmetrically consider the case when $r_5 + r_6 + r_7 \geq B$). Here we can get one bandpass that we are supposed to get without using r_2 (0, 0, 1)-rows in the third column. In Placement 2.2.01, by making use of B (0, 1, 0)-rows, we get one bandpass in the first and third columns respectively, and two in the second column. Thus Placement 2.2.01 is optimal.

- Else if $r_5 + r_7 + r_8 \geq B$.

  We are supposed to get one bandpass in the first column, and we can get it without using m_6. Placement 2.2.02 is similar to Placement 2.2.01, the only difference is that by making use of r_2 (0, 0, 1)-rows, the third column can achieve one bandpass while the first column wasted all the m_6 1's. The first column can achieve one bandpass for the premise of this case. Thus Placement 2.2.02 is optimal.

- Else if $r_2 + r_3 + r_5 + r_6 + 2r_7 + r_8 \geq 2B$.

  Now we have $r_5 + r_7 + r_8 < B$ and $r_5 + r_6 + r_7 + r_8 \geq B$, so $r_6 \geq B - (r_5 + r_7 + r_8)$. In Placement 2.2.03, the second band in the second column is exactly one bandpass, and the first column can achieve one bandpass. For the premise of this case, we can get one bandpass in the third column. Thus Placement 2.2.03 is optimal.

- Else if $r_2 + 2r_3 + 2r_5 + 2r_6 + 2r_7 + r_8 \geq 3B$.

						$m_4^{'}$	0	1	0	$B-(r_7+r_8)$
m_2	0	0	1	r_2		$m_6^{'}$	1	1	1	$r_5+r_6+r_7+r_8-B$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	B		m_2	0	0	1	r_2
m_5	1	1	0	r_5	$\Rightarrow$	m_7	1	0	1	r_7
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		$m_6^{''}$	1	1	1	$B-(r_5+r_7+r_8)$
						$m_4^{''}$	0	1	0	r_7+r_8

Table 2.23: Placement 2.2.03

						$m_4^{'}$	0	1	0	$r_2+r_3+r_5+r_6+r_7-B$
						$m_6^{'}$	1	1	1	$r_5+r_6+r_7+r_8-B$
m_2	0	0	1	r_2		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_2	0	0	1	r_2
m_4	0	1	0	B		m_7	1	0	1	r_7
m_5	1	1	0	r_5	$\Rightarrow$	$m_6^{'''}$	1	1	1	$2B-(r_2+r_3+r_5+r_6+2r_7+r_8)$
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		$m_6^{''}$	1	1	1	$r_2+r_3+r_6+r_7-B$
						$m_4^{''}$	0	1	0	$2B-(r_2+r_3+r_5+r_6+r_7)$

Table 2.24: Placement 2.2.04

In Placement 2.2.03, since now $r_2+r_3+r_5+r_6+2r_7+r_8 < 2B$, we can not get one bandpass in the third column. We take $2B-(r_2+r_3+r_5+r_6+2r_7+r_8)$ (1, 1, 1)-rows up to make the first column able to achieve one bandpass, but these rows are not involved in any bandpass in the second column. By using m_4, we can make the second band in the second column of length exactly B, and then the first band is long enough to produce one bandpass for the premise of the case. Thus Placement 2.2.04 is optimal.

– Else if $|m_3| \geq B$ (We can symmetrically consider the case when $|m_5| \geq B$ or $|m_7| \geq B$).

  We can make use of extra B (0, 1, 1)-rows. In Placement 2.2.02, the first band in the third column is not long enough to achieve one bandpass, but now we have enough (0, 1, 1)-rows to make it exactly one bandpass, and the second band is long enough to achieve one bandpass for $r_2+r_3+r_6+r_7 \geq B$. In the second column, the second band is exactly two bandpasses. The first column can achieve one bandpass. Thus we can get all bandpasses that we are supposed to get in Placement 2.2.04, which is an optimal placement.

– Else, we have the following instance I satisfying that $OPT = MAX - 1$.

					$m_4^{'}$	0	1	0	$r_2+r_3+r_5+r_6+r_7-B$
					$m_3^{'}$	0	1	1	$2B-(r_2+r_5+r_6+2r_7+r_8)$
m_2	0	0	1	r_2	$m_6^{'}$	1	1	1	$r_5+r_6+r_7+r_8-B$
m_3	0	1	1	r_3+B	m_2	0	0	1	r_2
m_4	0	1	0	B	m_7	1	0	1	r_7
m_5	1	1	0	r_5	m_8	1	0	0	r_8
m_6	1	1	1	r_6	m_5	1	1	0	r_5
m_7	1	0	1	r_7	$m_6^{''}$	1	1	1	$B-(r_5+r_7+r_8)$
m_8	1	0	0	r_8	$m_3^{''}$	0	1	1	$r_2+r_3+r_5+r_6+2r_7+r_8-B$
					$m_4^{''}$	0	1	0	$2B-(r_2+r_3+r_5+r_6+r_7)$

Table 2.25: Placement 2.2.05

m_2	0	0	1	r_2		
m_3	0	1	1	r_3		$r_3+r_6+r_7<B$,
m_4	0	1	0	q_4B		$r_5+r_6+r_7<B$,
m_5	1	1	0	r_5	satisfying	$r_2+r_3+r_5+2r_6+r_7+r_8<2B$,
m_6	1	1	1	r_6		$r_2+r_3+r_5+r_6+2r_7+r_8<2B$,
m_7	1	0	1	r_7		$r_2+2r_3+2r_5+2r_6+2r_7+r_8<3B$.
m_8	1	0	0	r_8		

In this case, we are supposed to get one bandpass in each of the first and third columns and q_4+1 bandpasses in the second column. Since $r_2+r_3+r_5+r_6+r_7+r_8<2B$, the bandpass that we are supposed to get in the first column and that in the third column must overlap at least $2B-(r_2+r_3+r_5+r_6+r_7+r_8)$ rows, which consist of m_6 and m_7. The bandpasses in the second column can not make use of this overlapping area. Otherwise, we can get one bandpass in the first column without using m_8 or get one bandpass in the third column without using m_2, which contradicts to that $r_5+r_6+r_7<B$ and $r_3+r_6+r_7<B$.

Suppose we can get q_4+1 bandpasses in the second column, then the total number of rows should be at least

$$T=(q_4+1)B+2B-(r_2+r_3+r_5+r_6+r_7+r_8)+r_2+r_8.$$

While the actual total number of rows is

$$T^{'}=q_4B+r_2+r_3+r_5+r_6+r_7+r_8.$$

Because $T-T^{'}=3B-(r_2+2r_3+2r_5+2r_6+2r_7+r_8)>0$. This is a contradiction. So we can get at most $MAX-1$ bandpasses.

2.2.4 Case 3 ($\mathbf{r_6,r_2,r_4,r_8>0}$ and $\mathbf{r_3\cdot r_5\cdot r_7=0}$)

In this case, we have $\exists r_i\in\{r_3,r_5,r_7\}$, s.t. $r_i=0$, without loss of generality, we assume $r_5=0$ hereafter, and then Table 2.3 reduces to Table 2.26. If there exists a permutation satisfying $r+r^{'}<$

Placement	r	$r^{'}$
P_1	$(r_2+r_3)\%B$	$(r_6+r_7)\%B$
P_2	$(r_2+r_7)\%B$	$(r_6+r_3)\%B$
P_3	$(r_3+r_4)\%B$	r_6
P_4	r_4	$(r_6+r_3)\%B$
P_5	$(r_8+r_7)\%B$	r_6
P_6	$r_8\%B$	$(r_6+r_7)\%B$

Table 2.26: Remaider values of the six placements when $r_5 = 0$.

$$\begin{array}{lllll} m_2 & 0 & 0 & 1 & r_2 \\ m_3 & 0 & 1 & 1 & r_3+B \\ m_4 & 0 & 1 & 0 & r_4 \\ m_6 & 1 & 1 & 1 & r_6 \\ m_7 & 1 & 0 & 1 & r_7 \\ m_8 & 1 & 0 & 0 & r_8 \end{array} \Rightarrow \begin{array}{lllll} m_2 & 0 & 0 & 1 & r_2 \\ m_3^{'} & 0 & 1 & 1 & B-r_2 \\ m_4 & 0 & 1 & 0 & r_4 \\ m_3^{''} & 0 & 1 & 1 & r_2+r_3 \\ m_6 & 1 & 1 & 1 & r_6 \\ m_7 & 1 & 0 & 1 & r_7 \\ m_8 & 1 & 0 & 0 & r_8 \end{array}$$

Table 2.27: Placement 3.1.01

B, it is the optimal solution. Otherwise we have $(r_2+r_3)\%B+(r_6+r_7)\%B \geq B$, $(r_2+r_7)\%B+(r_6+r_3)\%B \geq B$, $(r_3+r_4)\%B+r_6 \geq B$, $r_4+(r_6+r_3)\%B \geq B$, $r_8+(r_6+r_7)\%B \geq B$ and $(r_8+r_7)\%B+r_6 \geq B$ simultaneously. In the following, we separate this case into two disjoint subcases according to whether $q_5 = 0$.

- **Case 3.1** $\mathbf{q_5 = 0}$

  In this case, since both $q_5 = 0$ and $r_5 = 0$, it means we have no (1, 1, 0)-row in the given matrix. We separate this case into the following subcases.

  - If $|m_3| \geq B$ (We can symmetrically consider the case when $|m_7| \geq B$).

    In Placement 3.1.01, the 1's in the first two columns are consecutive, and the first band in the third column is of length B. Thus Placement 3.1.01 is optimal.

  - Else if $r_2+r_3 > B$ (We can symmetrically consider the case when $r_2+r_7 > B$).

    Similar to Placement 3.1.01, we can get MAX bandpasses in Placement 3.1.02.

  - Else if $r_3+r_4 > B$ (We can symmetrically consider the case when $r_7+r_8 > B$).

    In the premises of **Case 3**, we have $(r_3+r_4)\%B+r_6 \geq B$. If $r_3+r_4 > B$, then we have $r_3+r_4+r_6 \geq 2B$, and $r_4 < B$, so $r_3+r_6 > B$. In a similar way, we can show that $r_3+r_6 > B$ leads to $r_3+r_4 > B$. Therefore, $r_3+r_4 > B$ iff $r_3+r_6 > B$ (Similarly $r_7+r_8 > B$ iff $r_7+r_6 > B$). In this case, we assume $r_3+r_4 > B$. In Placement 3.1.03, the 1's in the first and third columns are consecutive, and the first band in the second column forms exactly one bandpass. Thus Placement 3.1.03 is optimal.

m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		$m_3^{'}$	0	1	1	$B - r_2$
m_4	0	1	0	r_4	$\Rightarrow$	m_4	0	1	0	r_4
m_6	1	1	1	r_6		$m_3^{''}$	0	1	1	$r_2 + r_3 - B$
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_8	1	0	0	r_8

Table 2.28: Placement 3.1.02

m_2	0	0	1	r_2		m_4	0	1	0	r_4
m_3	0	1	1	r_3		$m_3^{''}$	0	1	1	$B - r_4$
m_4	0	1	0	r_4	$\Rightarrow$	m_2	0	0	1	r_2
m_6	1	1	1	r_6		$m_3^{'}$	0	1	1	$r_3 + r_4 - B$
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_8	1	0	0	r_8

Table 2.29: Placement 3.1.03

- Else if $r_3 + r_6 + r_7 \geq B$.

  Here we can get one bandpass without using m_2 in the third column. Thus Placement 3.1.04 is optimal since the 1's are consecutive in each column.

- Else if $r_2 + r_3 + r_6 + r_7 + r_4 \geq 2B$ (We can symmetrically consider the cases when $r_2 + r_3 + r_6 + r_7 + r_8 \geq 2B$ or $r_4 + r_3 + r_6 + r_7 + r_8 \geq 2B$).

  In this case, since $r_3 + r_4 < B$, then $r_2 + r_6 + r_7 > B$. This together with $r_2 + r_7 < B$ implies $r_6 > B - (r_2 + r_7)$. In Placement 3.1.05, we can get one bandpass in each column, so this placement is optimal.

- Else, we have $r_3 + r_6 + r_7 < B$, $r_2 + r_3 + r_6 + r_7 + r_4 < 2B$, $r_2 + r_3 + r_6 + r_7 + r_8 < 2B$ and $r_4 + r_3 + r_6 + r_7 + r_8 < 2B$ simultaneously.

  In this case, if $\exists m_i \in \{m_2, m_4, m_8\}$, s.t. $|m_i| > B$, without loss of generality, we assume $|m_2| > B$. At the same time, if $r_2 + r_4 + r_8 + 2(r_3 + r_6 + r_7) \geq 3B$, we can get

m_2	0	0	1	r_2						
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_3	0	1	1	r_3
m_6	1	1	1	r_6	$\Rightarrow$	m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		m_8	1	0	0	r_8

Table 2.30: Placement 3.1.04

m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_7	1	0	1	r_7
m_4	0	1	0	r_4		$m_6^{'}$	1	1	1	$B-(r_2+r_7)$
m_6	1	1	1	r_6	$\Rightarrow$	m_8	1	0	0	r_8
m_7	1	0	1	r_7		$m_6^{''}$	1	1	1	$r_2+r_6+r_7-B$
m_8	1	0	0	r_8		m_3	0	1	1	r_3
						m_4	0	1	0	r_4

Table 2.31: Placement 3.1.05

m_2	0	0	1	r_2+B		$m_2^{'}$	0	0	1	$r_2+r_3+r_4+r_6+r_7-B$
m_3	0	1	1	r_3		$m_6^{''}$	1	1	1	$r_6+r_7+r_8-B$
m_4	0	1	0	r_4		m_3	0	1	1	r_3
m_6	1	1	1	r_6		m_4	0	1	0	r_4
m_7	1	0	1	r_7	$\Rightarrow$	$m_6^{'''}$	1	1	1	$2B-(r_3+r_4+r_6+r_7+r_8)$
m_8	1	0	0	r_8		m_8	1	0	0	r_8
						m_7	1	0	1	r_7
						m_6	1	1	1	$r_3+r_4+r_6-B$
						$m_2^{''}$	0	0	1	$2B-(r_3+r_4+r_6+r_7)$

Table 2.32: Placement 3.1.06

MAX bandpasses by Placement 3.1.06 shown in Table 2.32. For the premises of this case, we can get one bandpass in the first column, one in the second column, and two in the third column, because $r_2 + r_4 + r_8 + 2(r_3 + r_6 + r_7) \geq 3B$.

Otherwise, we have (1) $\exists m_i \in \{m_2, m_4, m_8\}$, s.t. $|m_i| > B$ and $r_2 + r_4 + r_8 + 2(r_3 + r_6 + r_7) < 3B$; or (2) $\forall m_i \in \{m_2, m_4, m_8\}$, s.t. $|m_i| < B$; In both two cases, we have $OPT = MAX - 1$. In the first case, assume the current instance is I. By changing $|m_3| = r_3$, $|m_7| = r_7$ to m_6, we get a new instance $I^{'}$, in which there is no m_3, m_5, m_7 rows. For the premises of this case, we have $r_6^{'} = r_3 + r_6 + r_7 < B$, $r_2+r_6^{'}+r_4 < 2B$, $r_2+r_6^{'}+r_8 < 2B$, $r_4+r_6^{'}+r_8 < 2B$ and $r_2+r_4+r_8+2r_6^{'} < 3B$. By Lemma 3, we have $OPT(I^{'}) = MAX(I^{'}) - 1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I) = MAX(I) - 1$. In the second case, we have $q_2, q_4, q_8 = 0$. By changing $|m_3| = r_3$, $|m_7| = r_7$ to m_6, we can get a new instance $I^{'}$ in which $r_6^{'} = r_3+r_6+r_7 < B$. By Lemma 3, we have $OPT(I^{'}) = MAX(I^{'})-1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I) = MAX(I) - 1$.

- **Case 3.2** $\mathbf{q_5 = 0}$

In this case, we can get MAX bandpasses by using Placement 3.2.01 shown in Table 2.33. It is necessary to mention that, now we have $r_3 + r_6 + r_7 < B$, and in the premises of Case 3, we have $r_6 + r_7 + r_8 > B$. So $r_8 > r_3$. In Placement 3.2.01, we can get two bandpasses in

						m_4	0	1	0	r_4
m_2	0	0	1	r_2		$m_5^{'}$	1	1	0	$r_3 + r_6$
m_3	0	1	1	r_3		$m_8^{'}$	1	0	0	$r_8 - r_3$
m_4	0	1	0	r_4		m_7	1	0	1	r_7
m_5	1	1	0	B	$\Rightarrow$	m_2	0	0	1	r_2
m_6	1	1	1	r_6		m_3	0	1	1	r_3
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		$m_5^{''}$	1	1	0	$B - (r_3 + r_6)$
						$m_8^{''}$	1	0	0	r_3

Table 2.33: Placement 3.2.01

each of the first two columns and one in the third column. Thus Placement 3.2.01 is optimal.

2.2.5 Case 4 ($\mathbf{r_6, r_2, r_4, r_8, r_3, r_5, r_7 > 0}$)

In this case, we have to refer to Table 2.3, which can not be reduced, to see the remainder makeup for each of the six placements. Same to the other three cases, we first use the Row-Stacking algorithm, if it can not produce MAX bandpasses, then we have the following conditions simultaneously.

$$(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$$

$$(r_2 + r_7)\%B + (r_6 + r_3)\%B \geq B$$

$$(r_4 + r_3)\%B + (r_6 + r_5)\%B \geq B$$

$$(r_4 + r_5)\%B + (r_6 + r_3)\%B \geq B$$

$$(r_8 + r_5)\%B + (r_6 + r_7)\%B \geq B$$

$$(r_8 + r_7)\%B + (r_6 + r_5)\%B \geq B$$

In the following, we separate this case into two subcases. In Case 4.1, we have $\exists r_i \in \{r_3, r_5, r_7\}$, s.t. $r_6 + r_i \geq B$. Then in Case 4.2, we have $r_6 + r_3 < B$, $r_6 + r_5 < B$ and $r_6 + r_7 < B$.

Case 4.1 $\quad \exists \mathbf{r_i} \in \{\mathbf{r_3}, \mathbf{r_5}, \mathbf{r_7}\}$, s.t. $\mathbf{r_6} + \mathbf{r_i} \geq \mathbf{B}$

In this case, we assume $r_6 + r_7 \geq B$ (We can symmetrically consider the cases when $r_6 + r_3 \geq B$ or $r_6 + r_5 \geq B$). We separate this case into the following subcases.

- If $|m_2| + r_3 \geq B$ (We can symmetrically consider the case when $|m_8| + r_5 \geq B$).

  Here we have $r_7 \geq B - r_6$ and $|m_2| \geq B - r_3$. In Placement 4.1.01, the 1's in the first two columns are consecutive, and there are three bands in the third column, and two of them are of length exactly B. Thus we can get MAX bandpasses in Placement 4.1.01.

						m_2'	0	0	1	$B - r_3$
m_2	0	0	1	$\|m_2\|$		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	m_6	1	1	1	r_6
m_6	1	1	1	r_6		m_7'	1	0	1	$B - r_6$
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7''	1	0	1	$r_6 + r_7 - B$
						m_2''	0	0	1	$\|m_2\| - (B - r_3)$

Table 2.34: Placement 4.1.01

						m_4''	0	1	0	$\|m_4\| - (B - r_3)$
m_2	0	0	1	r_2		m_5	1	1	0	r_5
m_3	0	1	1	r_3		m_6	1	1	1	r_6
m_4	0	1	0	$\|m_4\|$		m_7'	1	0	1	$r_2 + r_3 + r_7 - B$
m_5	1	1	0	r_5	$\Rightarrow$	m_8	1	0	0	r_8
m_6	1	1	1	r_6		m_7''	1	0	1	$B - (r_2 + r_3)$
m_7	1	0	1	r_7		m_2	0	0	1	r_2
m_8	1	0	0	r_8		m_3	0	1	1	r_3
						m_4'	0	1	0	$B - r_3$

Table 2.35: Placement 4.1.02

- Else if $|m_4| + r_3 \geq B$ (We can symmetrically consider the case when $|m_4| + r_5 \geq B$).

  Since $r_2 + r_3 < B$, $r_6 + r_7 \geq B$ and $(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$, we have $r_2 + r_3 + r_6 + r_7 \geq 2B$. For $r_6 < B$, then $r_2 + r_3 + r_7 > B$. Therefore, we have $r_7 > B - (r_2 + r_3)$ and $|m_4| \geq B - r_3$. In Placement 4.1.02, the 1's in the first column are consecutive, and in each of the second and third columns, there are two bands of 1's, one of which is exactly one bandpass. Thus Placement 4.1.02 is optimal.

- Else if $|m_3| \geq r_8$ (We can symmetrically consider the case when $|m_5| \geq r_2$).

  This case is similar to last case. Since $r_5 + r_8 < B$, $r_6 + r_7 \geq B$ and $(r_5 + r_8)\%B + (r_6 + r_7)\%B \geq B$, we can get $r_5 + r_6 + r_8 > B$. In Placement 4.1.03, the 1's in the third column are consecutive, and in each of the first and second columns, there are two bands, one of which is of length B. Thus Placement 4.1.03 is optimal.

- Else if $r_6 + r_3 < B$ (We can symmetrically consider the case when $r_6 + r_5 < B$).

  In the premises of **Case 4**, $r_6 + r_7 \geq B$, then $r_3 + r_6 + r_7 \geq B$, so $r_7 \geq B - (r_6 + r_3)$. In Placement 4.1.04, the 1's in the second column are consecutive, in each of the first and third columns, there are two bands, one of which is of length exactly B. Thus Placement 4.1.04 is optimal.

						m_8	1	0	0	r_8
m_2	0	0	1	r_2		m_5	1	1	0	r_5
m_3	0	1	1	$\|m_3\|$		$m_6^{'}$	1	1	1	$B-(r_5+r_8)$
m_4	0	1	0	r_4		m_3	0	1	1	r_8
m_5	1	1	0	r_5	$\Rightarrow$	m_2	0	0	1	r_2
m_6	1	1	1	r_6		m_7	1	0	1	r_7
m_7	1	0	1	r_7		$m_6^{''}$	1	1	1	$r_5+r_6+r_8-B$
m_8	1	0	0	r_8		m_3	0	1	1	$\|m_3\|-r_8$
						m_4	0	1	0	r_4

Table 2.36: Placement 4.1.03

						m_2	0	0	1	r_2
m_2	0	0	1	r_2		$m_7^{''}$	1	0	1	$r_3+r_6+r_7-B$
m_3	0	1	1	r_3		$m_8^{''}$	1	0	0	r_8-r_3
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	m_4	0	1	0	r_4
m_6	1	1	1	r_6		m_3	0	1	1	r_3
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		$m_7^{'}$	1	0	1	$B-(r_6+r_3)$
						$m_8^{'}$	1	0	0	r_3

Table 2.37: Placement 4.1.04

- Else if $|m_7| \geq r_4$.

  Similar to Placement 4.1.03, we can get MAX bandpasses in Placement 4.1.05.

- Else if $r_3 + r_5 \geq B$.

  Now we have $r_2 > r_5$ and $r_6 + r_7 \geq B$. In Placement 4.1.06, there are two bands of 1's in each column, one of which is of length B. Thus Placement 4.1.06 is optimal.

- Else if $r_3 + r_5 + r_6 + r_7 \geq 2B$.

						m_4	0	1	0	r_4
m_2	0	0	1	r_2		m_5	1	1	0	r_5
m_3	0	1	1	r_3		$m_6^{'}$	1	1	1	$B-(r_4+r_5)$
m_4	0	1	0	r_4		$m_7^{'}$	1	0	1	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_2	0	0	1	r_2
m_6	1	1	1	r_6		m_3	0	1	1	r_3
m_7	1	0	1	$\|m_7\|$		$m_6^{''}$	1	1	1	$r_4+r_5+r_6-B$
m_8	1	0	0	r_8		$m_7^{''}$	1	0	1	$\|m_7\|-r_4$
						m_8	1	0	0	r_8

Table 2.38: Placement 4.1.05

						m_2'	0	0	1	r_5
						m_3'	0	1	1	$B-r_5$
m_2	0	0	1	r_2		m_5	1	1	0	r_5
m_3	0	1	1	r_3		m_8	1	0	0	r_8
m_4	0	1	0	r_4		m_7'	1	0	1	r_6+r_7-B
m_5	1	1	0	r_5	$\Rightarrow$	m_2''	0	0	1	r_2-r_5
m_6	1	1	1	r_6		m_7''	1	0	1	$B-r_6$
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		m_3''	0	1	1	r_3+r_5-B
						m_4	0	1	0	r_4

Table 2.39: Placement 4.1.06

m_2	0	0	1	r_2		m_2'	0	0	1	r_5
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_6	1	1	1	$B-(r_3+r_5)$
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_8	1	0	0	r_3
m_7	1	0	1	r_7		m_4	0	1	0	r_4
m_8	1	0	0	r_8		m_6	1	1	1	$r_3+r_5+r_6-B$
						m_7	1	0	1	r_7

Table 2.40: Placement 4.1.07

For $r_7 < B$, then $r_3+r_5+r_6 > B$. In Placement 4.1.07, there are two bands of 1's in each column, one of which is of length B. Thus Placement 4.1.07 is optimal.

- Else if $r_2+r_3+r_5+r_6+r_7+r_4 \geq 3B$ (We can symmetrically consider the cases when $r_2+r_3+r_5+r_6+r_7+r_8 \geq 3B$ or $r_4+r_3+r_5+r_6+r_7+r_8 \geq 3B$).

  For $r_2+r_3 < B$ and $r_7 < B$, then $r_4+r_5+r_6 > B$. Also for $r_4+r_5 < B$, then $r_6 > B-(r_4+r_5)$. In Placement 4.1.08, the 1's in the first column are consecutive, there are two bands of 1's in the second column, one of which is exactly one bandpass. In the third column, since $r_2+r_3+r_5+r_6+r_7+r_4 \geq 3B$, we can get two bandpasses that we are supposed to get. Thus Placement 4.1.08 is optimal.

- Else, we have $OPT = MAX - 1$.

  From the current instance I, we can get a new instance I' by changing $m_3(|m_3| = r_3)$, $m_5(|m_5| = r_5)$ and $m_7(|m_7| = r_7)$ to m_6. Then in I', we only have m_2, m_4, m_6, m_8, and $|m_2| = r_2$, $|m_4| = r_4$, $|m_8| = r_8$, $r_6' = r_3+r_5+r_6+r_7-B$, satisfying that

$$r_2+r_6'+r_4 = r_2+r_3+r_5+r_6+r_7+r_4-B < 2B,$$

$$r_2+r_6'+r_8 = r_2+r_3+r_5+r_6+r_7+r_8-B < 2B,$$

						m_4	0	1	0	r_4
m_2	0	0	1	r_2		m_5	1	1	0	r_5
m_3	0	1	1	r_3		$m_6^{'}$	1	1	1	$B-(r_4+r_5)$
m_4	0	1	0	r_4		m_8	1	0	0	r_3
m_5	1	1	0	r_5	$\Rightarrow$	m_7	1	0	1	r_7
m_6	1	1	1	r_6		$m_6^{''}$	1	1	1	$r_4+r_5+r_6-B$
m_7	1	0	1	r_7		m_3	0	1	1	r_3
m_8	1	0	0	r_8		m_2	0	0	1	r_2

Table 2.41: Placement 4.1.08

$$r_4 + r_6^{'} + r_8 = r_4 + r_3 + r_5 + r_6 + r_7 + r_8 - B < 2B.$$

By Lemma 3, we have $OPT(I^{'}) = MAX(I^{'}) - 1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I) = MAX(I) - 1$.

Case 4.2 $\quad \forall \mathbf{r_i} \in \{\mathbf{r_3}, \mathbf{r_5}, \mathbf{r_7}\}$**, s.t.** $\mathbf{r_6} + \mathbf{r_i} < \mathbf{B}$

In this case, we have $\forall r_i \in \{r_3, r_5, r_7\}$, s.t. $r_6 + r_i < B$, then we can induct the following three conditions:

$$(1) \quad r_2 + r_3 \geq B \text{ iff } r_2 + r_7 \geq B$$

$$(2) \quad r_4 + r_3 \geq B \text{ iff } r_4 + r_5 \geq B$$

$$(3) \quad r_8 + r_5 \geq B \text{ iff } r_8 + r_7 \geq B$$

Since they can be symmetrically discussed, without loss of generality, we only prove (1). For $r_6 + r_7 < B$, if $r_2 + r_3 \geq B$, together with the premise of **Case 4**: $(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$, then we will have $r_2 + r_3 + r_6 + r_7 \geq 2B$. Since $r_6 + r_3 < B$, then $r_2 + r_7 \geq B$. Thus $r_2 + r_3 \geq B$ leads to $r_2 + r_7 \geq B$. Similarly, we can prove that $r_2 + r_7 \geq B$ leads to $r_2 + r_3 \geq B$. In the following, we separate this case into subcases according to how many of the three conditions above are satisfied.

- At least two of (1), (2), (3) are satisfied. Since (1), (2), (3) can be symmetrically discussed, without loss of generality, here we assume that (1) and (3) are satisfied simultaneously. Then in Placement 4.2.01, the 1's in the second column are consecutive, and there are two bands in each of the first and third columns, one of them is exactly one bandpass. Thus Placement 4.2.01 is optimal.
- Exactly one of (1), (2), (3) is satisfied, without loss of generality, we assume (1) is satisfied here. We separate this case into the following subcases.
  - If $|m_6| > B$.

    We can make use of extra B (1, 1, 1)-rows. Since $r_2 + r_3 \geq B$ and $r_6 + r_7 < B$, together with the premise of **Case 4** that $(r_2 + r_3)\%B + (r_6 + r_7)\%B \geq B$, then we

						$m_2^{'}$	0	0	1	$B - r_3$
m_2	0	0	1	r_2		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	m_6	1	1	1	r_6
m_6	1	1	1	r_6		$m_7^{''}$	1	0	1	$r_7 + r_8 - B$
m_7	1	0	1	r_7		$m_2^{''}$	0	0	1	$r_2 + r_3 - B$
m_8	1	0	0	r_8		$m_7^{'}$	1	0	1	$B - r_8$
						m_8	1	0	0	r_8

Table 2.42: Placement 4.2.01

						m_2	0	0	1	r_2
m_2	0	0	1	r_2		$m_3^{''}$	0	1	1	$r_3 + r_6 + r_7 - B$
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		$m_3^{'}$	0	1	1	$B - (r_6 + r_7)$
m_5	1	1	0	r_5	$\Rightarrow$	$m_6^{''}$	1	1	1	$r_6 + r_7$
m_6	1	1	1	$r_6 + B$		m_5	1	1	0	r_5
m_7	1	0	1	r_7		$m_6^{'}$	1	1	1	$B - r_7$
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_8	1	0	0	r_8

Table 2.43: Placement 4.2.02

have $r_2+r_3+r_6+r_7 \geq 2B$. For $r_2 < B$, then $r_3+r_6+r_7 > B$. Also for $r_6+r_7 < B$, then $r_3 > B - (r_6 + r_7)$. In Placement 4.2.02, the 1's in the first two columns are consecutive, and there are three bands in the third column, two of them are of length exactly B. Thus Placement 4.2.02 is optimal.

– Else if $|m_4| > B$ (We can symmetrically consider the case when $|m_8| > B$).

  In Placement 4.2.03, the 1's in the first column are consecutive, and there are two bands in each of the second and third columns, one of which is exactly one bandpass. Thus Placement 4.2.03 is optimal.

– Else if $|m_3| > B$ (We can symmetrically consider the cases when $|m_5| > B$ or $|m_7| > B$).

  In Placement 4.2.04, the 1's in the first two columns are consecutive, and there are three bands in the third column, two of them are of length exactly B. Thus Placement 4.2.04 is optimal.

– Else if $r_3 + r_5 + r_6 \geq B$.

  In this case, we can get one bandpass in the second column without using m_4. In Placement 4.2.05, the 1's in the first two columns are consecutive, and there are two bands in the third column, one of them is exactly one bandpass. Thus Placement 4.2.05 is

						m_2	0	0	1	r_2
m_2	0	0	1	r_2		$m_3^{'}$	0	1	1	$B - r_2$
m_3	0	1	1	r_3		$m_4^{''}$	0	1	0	$r_4 + r_5$
m_4	0	1	0	$r_4 + B$		$m_3^{''}$	0	1	1	$r_2 + r_3 - B$
m_5	1	1	0	r_5	$\Rightarrow$	m_6	1	1	1	r_6
m_6	1	1	1	r_6		m_7	1	0	1	r_7
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_5	1	1	0	r_5
						$m_4^{'}$	0	1	0	$B - r_5$

Table 2.44: Placement 4.2.03

						$m_2^{'}$	0	0	1	$B - r_7$
m_2	0	0	1	r_2		m_7	1	0	1	r_7
m_3	0	1	1	$r_3 + B$		m_8	1	0	0	r_8
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	m_6	1	1	1	r_6
m_6	1	1	1	r_6		$m_3^{'}$	0	1	1	$B - r_6$
m_7	1	0	1	r_7		m_4	0	1	0	r_4
m_8	1	0	0	r_8		$m_3^{''}$	0	1	1	$r_3 + r_6$
						$m_2^{''}$	0	0	1	$r_2 + r_7 - B$

Table 2.45: Placement 4.2.04

optimal.

– Else if $r_5 + r_6 + r_7 \geq B$.

  Similar to the previous case, we can get one bandpass without using m_8 in the first column. For the same reason which was discussed in Placement 4.2.05, in this case, Placement 4.2.06 is optimal.

– Else if $r_3 + r_4 + r_6 \geq B$.

  We are supposed to get one bandpass in the second column. By the premise of this case, we can get the bandpass in the second column without using m_5. For the same reason which was discussed in Placement 4.2.05, in this case, Placement 4.2.07 is optimal.

m_2	0	0	1	r_2		$m_2^{'}$	0	0	1	$B - r_7$
m_3	0	1	1	r_3		m_7	1	0	1	r_7
m_4	0	1	0	r_4		m_8	1	0	0	r_8
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_3	0	1	1	r_3
m_8	1	0	0	r_8		$m_2^{''}$	0	0	1	$r_2 + r_7 - B$

Table 2.46: Placement 4.2.05

m_2	0	0	1	r_2		$m_2^{'}$	0	0	1	$B - r_3$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_2^{''}$	0	0	1	$r_2 + r_3 - B$

Table 2.47: Placement 4.2.06

						m_2	0	0	1	r_2
m_2	0	0	1	r_2		$m_7^{'}$	1	0	1	$B - r_2$
m_3	0	1	1	r_3		m_8	1	0	0	r_8
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	$m_7^{''}$	1	0	1	$r_2 + r_7 - B$
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_3	0	1	1	r_3
m_8	1	0	0	r_8		m_4	0	1	0	r_4

Table 2.48: Placement 4.2.07

– Else if $r_2 + r_3 + r_7 \geq 2B$.

  Here we can get two bandpasses without using m_6 in the third column. For the same reason which was discussed in Placement 4.2.05, in this case, Placement 4.2.08 is optimal.

– Else if $r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$.

  Since $r_3 + r_4 + r_6 < B$ and $r_3 + r_4 + r_5 + r_6 \geq B$, then $r_5 \geq B - (r_3 + r_4 + r_6)$. From $r_2 + r_3 + r_6 + r_7 \geq 2B$, $r_2 < B$ and $r_6 + r_7 < B$, we have $r_3 > B - (r_6 + r_7)$. In Placement 4.2.09, we can get one band in the first column from the premise of this case, one in the second column and two in the third column. Thus Placement 4.2.09 is optimal.

						$m_2^{'}$	0	0	1	$B - r_3$
m_2	0	0	1	r_2		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_6	1	1	1	r_6
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_2^{''}$	0	0	1	$r_2 + r_3 - B$

Table 2.49: Placement 4.2.08

						m_2	0	0	1	r_2
m_2	0	0	1	r_2		m_3''	0	1	1	$r_3+r_6+r_7-B$
m_3	0	1	1	r_3		m_5'	1	1	0	$B-(r_3+r_4+r_6)$
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_3'	0	1	1	$B-(r_6+r_7)$
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		m_8	1	0	0	r_8
						m_5''	1	1	0	$r_3+r_4+r_5+r_6-B$

Table 2.50: Placement 4.2.09

						m_2''	0	0	1	r_2+r_7-B
m_2	0	0	1	r_2		m_6'	1	1	1	$2B-(r_2+r_3+r_7)$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_6''	1	1	1	$r_2+r_3+r_6+r_7-2B$
m_6	1	1	1	r_6		m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						m_2'	0	0	1	$B-r_7$

Table 2.51: Placement 4.2.10

- Else if $r_2+r_3+r_5+r_6+2r_7+r_8 \geq 3B$.

  Since $r_2+r_3+r_7 < 2B$ and $r_2+r_3+r_6+r_7 \geq 2B$, then $r_6 \geq 2B-(r_2+r_3+r_7)$. In Placement 4.2.10, we can get one bandpass in the first column for $r_2+r_3+r_5+r_6+2r_7+r_8 \geq 3B$, one in the second column and two in the third column. Thus Placement 4.2.10 is optimal.

- Else if $r_2+r_4+r_8+2r_3+2r_5+2r_6+2r_7 \geq 4B$.

  Since $r_3+r_4 < B$ and $r_5+r_6+r_7 < B$, then $r_2+r_3+r_5+r_6+r_7+r_8 > 2B$. This together with $r_7+r_8 < B$ and $r_3+r_5+r_6 < B$ implies $r_2 > 2B-(r_3+r_5+r_6+r_7+r_8)$. Now we have $r_2+r_3+r_5+r_6+2r_7+r_8 < 3B$, $r_2+r_3+r_6+r_7 \geq 2B$ and $r_5+r_6+r_7+r_8 \geq B$, then $r_6 > 3B-(r_2+r_3+r_5+r_6+2r_7+r_8)$. In Placement 4.2.11, we can get one bandpass in the first column, one in the second column and two in the third column. Thus Placement 4.2.11 is optimal.

- Else, we have $|m_3| = r_3$, $|m_4| = r_4$, $|m_5| = r_5$, $|m_6| = r_6$, $|m_7| = r_7$, $|m_8| = r_8$, then $OPT = MAX - 1$.

  Suppose we can get MAX bandpasses. Since $r_3+r_4+r_5+r_6+r_7+r_8 < 2B$, the bandpasses in the first column and the second column must overlap at least $2B - r_3 - r_4 - r_5 - r_6 - r_7 - r_8$ rows, which consist of m_5 and m_6. The bandpasses

m_2	0	0	1	r_2		$m_2^{''}$	0	0	1	$2B-(r_3+r_5+r_6+r_7+r_8)$
m_3	0	1	1	r_3		$m_6^{'}$	1	1	1	$r_5+r_6+r_7+r_8-B$
m_4	0	1	0	r_4		m_3	0	1	1	r_3
m_5	1	1	0	r_5	$\Rightarrow$	m_4	0	1	0	r_4
m_6	1	1	1	r_6		$m_6^{''}$	1	1	1	$r_2+r_3+r_6+r_7-2B$
m_7	1	0	1	r_7		m_5	1	1	0	r_5
m_8	1	0	0	r_8		m_8	1	0	0	r_8
						$m_6^{'''}$	1	1	1	$3B-(r_2+r_3+r_5+r_6+2r_7+r_8)$
						m_7	1	0	1	r_7
						$m_2^{'}$	0	0	1	$r_2+r_3+r_5+r_6+r_7+r_8-2B$

Table 2.52: Placement 4.2.11

in the third column can not make use of the overlapping area of the first and second columns. Otherwise, we will have $r_3+r_5+r_6 \geq B$ or $r_5+r_6+r_7 \geq B$ or $r_3+r_4+r_5+2r_6+r_7+r_8 \geq 2B$, which is a contradiction. Assume we can get q_2+2 bandpasses in the third column, then the total number of rows should be at least $T = (q_2+2)B + (2B-r_3-r_4-r_5-r_6-r_7-r_8)+r_4+r_8$. It is greater than the actual number of rows which is $T^{'} = q_2B+r_2+r_3+r_4+r_5+r_6+r_7+r_8$, because $T-T^{'} = 4B-(r_2+r_4+r_8+2r_3+2r_5+2r_6+2r_7) > 0$. This is a contradiction. Thus $OPT = MAX - 1$.

- None of (1), (2), (3) is satisfied. Then we have $r_2+r_3 < B$, $r_2+r_7 < B$, $r_4+r_3 < B$, $r_4+r_5 < B$, $r_8+r_5 < B$, $r_8+r_7 < B$ simultaneously. We separate this case into two subcases. In the first one, $\exists m_i \in \{m_3, m_5, m_7\}$, s.t. $|m_i| > B$; in the second one, we have $\forall m_i \in \{m_3, m_5, m_7\}$, s.t. $|m_i| < B$.

  – If $\exists m_i \in \{m_3, m_5, m_7\}$, s.t. $|m_i| > B$.

    Since m_3, m_5, m_7 can be symmetrically discussed, without loss of generality, we assume $|m_7| > B$ in the sequel. It means we can make use of extra B rows of m_7. We separate this case into the following subcases.

    * If $|m_6| > B$.

      We can make use of extra B rows of m_6. In Placement 4.2.12, the 1's in the first column are consecutive. In each of the second and third columns, there are two bands, one of which is exactly one bandpass. Thus Placement 4.2.12 is optimal.

    * Else if $|m_2| > B$ (We can symmetrically consider the cases when $|m_4| > B$ or $|m_8| > B$).

      Here we can make use of extra B rows of m_2. In Placement 4.2.13, the 1's in the second column are consecutive. In each of the first and third columns, there are two bands, one of which is exactly one bandpass. Thus Placement 4.2.13 is optimal.

m_2	0	0	1	r_2
m_3	0	1	1	r_3
m_4	0	1	0	r_4
m_5	1	1	0	r_5
m_6	1	1	1	$r_6 + B$
m_7	1	0	1	$r_7 + B$
m_8	1	0	0	r_8

$\Rightarrow$

m_4	0	1	0	r_4
m_3	0	1	1	r_3
$m_6^{'}$	1	1	1	$B - (r_3 + r_4)$
$m_7^{'}$	1	0	1	r_4
m_8	1	0	0	r_8
m_5	1	1	0	r_5
$m_6^{''}$	1	1	1	$r_3 + r_4 + r_6$
$m_7^{''}$	1	0	1	$B + r_7 - r_4$
m_2	0	0	1	r_2

Table 2.53: Placement 4.2.12

m_2	0	0	1	$r_2 + B$
m_3	0	1	1	r_3
m_4	0	1	0	r_4
m_5	1	1	0	r_5
m_6	1	1	1	r_6
m_7	1	0	1	$r_7 + B$
m_8	1	0	0	r_8

$\Rightarrow$

$m_2^{'}$	0	0	1	$B - r_3$
m_3	0	1	1	r_3
m_4	0	1	0	r_4
m_5	1	1	0	r_5
m_6	1	1	1	r_6
$m_7^{''}$	1	0	1	$r_7 + r_8$
$m_2^{''}$	0	0	1	$r_2 + r_3$
$m_7^{'}$	1	0	1	$B - r_8$
m_8	1	0	0	r_8

Table 2.54: Placement 4.2.13

$*$ Else if $r_4 + r_5 + r_6 \geq B$.

In this case, we can get one bandpass without using m_3 in the second column. In Placement 4.2.14, we can get two bandpasses in each of the first and third columns, one in the second column. Thus Placement 4.2.14 is optimal.

$*$ Else if $r_6 + r_7 + r_8 \geq B$.

Here we can get one bandpass without using m_5 in the first column. In Placement 4.2.15, the 1's in the second column are consecutive, and we can get two bandpasses in each of the first and third columns. Thus Placement 4.2.15 is optimal.

m_2	0	0	1	r_2
m_3	0	1	1	r_3
m_4	0	1	0	r_4
m_5	1	1	0	r_5
m_6	1	1	1	r_6
m_7	1	0	1	$r_7 + B$
m_8	1	0	0	r_8

$\Rightarrow$

m_8	1	0	0	r_8
$m_7^{'}$	1	0	1	$B - r_8$
m_2	0	0	1	r_2
m_3	0	1	1	r_3
$m_7^{''}$	1	0	1	$r_7 + r_8$
m_6	1	1	1	r_6
m_5	1	1	0	r_5
m_4	0	1	0	r_4

Table 2.55: Placement 4.2.14

m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		$m_7^{'}$	1	0	1	$B - r_2$
m_4	0	1	0	r_4		m_8	1	0	0	r_8
m_5	1	1	0	r_5	$\Rightarrow$	$m_7^{''}$	1	0	1	$r_2 + r_7$
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	$r_7 + B$		m_3	0	1	1	r_3
m_8	1	0	0	r_8		m_4	0	1	0	r_4
						m_5	1	1	0	r_5

Table 2.56: Placement 4.2.15

						m_8	1	0	0	r_8
m_2	0	0	1	r_2		$m_7^{'}$	1	0	1	$B - r_8$
m_3	0	1	1	r_3		$m_2^{''}$	0	0	1	$r_2 + r_6 + r_7 + r_8 - B$
m_4	0	1	0	r_4		m_3	0	1	1	r_3
m_5	1	1	0	r_5	$\Rightarrow$	m_4	0	1	0	r_4
m_6	1	1	1	r_6		m_5	1	1	0	r_5
m_7	1	0	1	$r_7 + B$		m_6	1	1	1	r_6
m_8	1	0	0	r_8		$m_7^{''}$	1	0	1	$r_7 + r_8$
						$m_2^{'}$	0	0	1	$B - (r_6 + r_7 + r_8)$

Table 2.57: Placement 4.2.16

* Else if $r_2 + r_6 + r_7 + r_8 \geq B$.

  Since $r_6 + r_7 + r_8 < B$, then $r_2 > B - (r_6 + r_7 + r_8)$. In Placement 4.2.16, the 1's in the second columns are consecutive. In each of the first and second columns, there are two bands, one of which is exactly one bandpass. Thus Placement 4.2.16 is optimal.

* Else if $r_2 + r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$.

  Since $r_2 + r_6 + r_7 + r_8 < B$ and $r_2 + r_3 + r_6 + r_7 + r_8 > B$, then $r_3 > B - (r_2 + r_6 + r_7 + r_8)$. In Placement 4.2.17, we can get one bandpass in the second column, two in each of the other two columns. Thus Placement 4.2.17 is optimal.

* Else, we have $OPT = MAX - 1$.

  From the current instance I, we can construct a new instance $I^{'}$ by changing m_2 ($|m_2| = r_2$) and m_8 ($|m_8| = r_8$) to m_7, m_4 ($|m_4| = r_4$) to m_5. Then in $I^{'}$, we only have m_3, m_5, m_6 and m_7, satisfying that $r_3^{'} = r_3$, $r_3^{'} + r_6 < B$, $r_5^{'} = r_4 + r_5$, $r_5^{'} + r_6 < B$, $r_7^{'} = r_2 + r_8 + r_7$, $r_7^{'} + r_6 < B$ and $r_3^{'} + r_5^{'} + 2r_6 + r_7^{'} < 2B$. By Corollary 2, we have $OPT(I^{'}) = MAX(I^{'}) - 1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I) = MAX(I) - 1$.

– Else $|m_3| = r_3$, $|m_5| = r_5$, $|m_7| = r_7$.

						m_8	1	0	0	r_8
m_2	0	0	1	r_2		m_7'	1	0	1	$B - r_8$
m_3	0	1	1	r_3		m_3''	0	1	1	$r_2 + r_3 + r_6 + r_7 + r_8 - B$
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	$r_7 + B$		m_7''	1	0	1	$r_7 + r_8$
m_8	1	0	0	r_8		m_3'	0	1	1	$B - (r_2 + r_6 + r_7 + r_8)$
						m_2	0	0	1	r_2

Table 2.58: Placement 4.2.17

						m_7''	1	0	1	y
m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_6''	1	1	1	x
m_5	1	1	0	r_5	$\Rightarrow$	m_4	0	1	0	r_4
m_6	1	1	1	r_6		m_5	1	1	0	r_5
m_7	1	0	1	r_7		m_6'	1	1	1	$r_6 - x$
m_8	1	0	0	r_8		m_7'	1	0	1	$r_7 - y$
						m_8	1	0	0	r_8

Table 2.59: Placement 4.2.18

Here we are supposed to get one bandpass in each column in the remainder matrix. We separate this case into the following subcases.

* If $r_2 + r_3 + r_5 + r_6 + r_7 + r_8 \geq 2B$ (We can symmetrically consider the cases when $r_2 + r_3 + r_5 + r_6 + r_7 + r_4 \geq 2B$ or $r_4 + r_3 + r_5 + r_6 + r_7 + r_8 \geq 2B$).
  In this case, keeping the 1's in the second column consecutive, we take some m_6 up to between m_3 and m_4 and/or some m_7 up to the top lines of the matrix, in order to make the first band of 1's in the third column is of length exactly B. To reach this goal, we have to take $B - (r_2 + r_3)$ this many m_6 and/or m_7 up, which are inevitably not involved in any bandpass in the first column. To make the remaining band of 1's in the first column is of length at least B, we must have $r_5 + r_6 + r_7 + r_8 - (B - r_2 - r_3) \geq B$, which is right the premise of the case, thus we have enough 1's in the first column to form a bandpass. This is shown in Placement 4.2.18.
* Else if $\exists m_i \in \{m_2, m_4, m_8\}$, s.t. $|m_i| > B$.
  Without loss of generality, here we assume $|m_2| > B$ and separate this case into the following subcases.
  + If $r_3 + r_4 + r_5 \geq B$ (We can symmetrically consider the case when $r_5 + r_7 + r_8 \geq B$).

m_2	0	0	1	r_2+B		$m_2^{'}$	0	0	1	$B-r_3$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_6	1	1	1	r_6
m_8	1	0	0	r_8		m_7	1	0	1	r_7
						$m_2^{''}$	0	0	1	r_2+r_3

Table 2.60: Placement 4.2.19

m_2	0	0	1	r_2+B		$m_2^{''}$	0	0	1	r_2+r_7
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_6	1	1	1	r_6
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_2^{'}$	0	0	1	$B-r_7$

Table 2.61: Placement 4.2.20

It means we can get one bandpass without using the r_6 (1, 1, 1)-rows in the second column. Then by Placement 4.2.19, we have room for adjustment by making use of extra B rows of m_2. We can get four bandpasses which we are supposed to get. Thus Placement 4.2.19 is optimal.

+ Else if $r_3 + r_5 + r_6 \geq B$ (We can symmetrically consider the case when $r_5 + r_6 + r_7 \geq B$).

  It is similar to the previous case, here we can get a bandpass without using the r_4 (0, 1, 0)-rows in the second column. In Placement 4.2.20, after dropping those rows, we have room for m_2 to adjust the matrix. Then we can get four bandpasses which we are supposed to get. Thus Placement 4.2.20 is optimal.

+ Else if $r_3 + r_4 + 2r_5 + r_6 + r_7 + r_8 \geq 2B$.

  Since $r_5+r_7+r_8 < B$ and $r_5+r_6+r_7+r_8 > B$, then $r_6 > B-(r_5+r_7+r_8)$. In Placement 4.2.21, we can get four bandpasses. Thus Placement 4.2.21 is optimal.

+ Else if $r_2 + r_4 + r_8 + 2r_3 + 2r_5 + 2r_6 + 2r_7 \geq 3B$.

  Since the previous case failed to get MAX bandpasses, now we have $r_3 + r_4 + 2r_5 + r_6 + r_7 + r_8 < 2B$. It means the band of 1's in the second column is not enough to form a bandpass, because too many (1, 1, 1)-rows are not involved in any bandpass in the second column. Thus we move exactly $2B - (r_3 + r_4 + 2r_5 + r_6 + r_7 + r_8)$ rows of m_6 up to between m_5 and m_8 to make sure there is

						$m_2^{'}$	0	0	1	$B + r_2 - (r_5 + r_8)$
m_2	0	0	1	$r_2 + B$		$m_6^{'}$	1	1	1	$r_5 + r_6 + r_7 + r_8 - B$
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_6^{''}$	1	1	1	$B - (r_5 + r_7 + r_8)$
						$m_2^{''}$	0	0	1	$r_5 + r_8$

Table 2.62: Placement 4.2.21

						$m_2^{'}$	0	0	1	$r_2 + r_3 + r_4 + r_5 + r_6 + r_7 - B$
						$m_6^{'}$	1	1	1	$r_5 + r_6 + r_7 + r_8 - B$
m_2	0	0	1	$r_2 + B$		m_3	0	1	1	r_3
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_5	1	1	0	r_5
m_5	1	1	0	r_5	$\Rightarrow$	$m_6^{''}$	1	1	1	$2B - (r_3 + r_4 + 2r_5 + r_6 + r_7 + r_8)$
m_6	1	1	1	r_6		m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_6^{'''}$	1	1	1	$r_3 + r_4 + r_5 + r_6 - B$
						$m_2^{''}$	0	0	1	$2B - (r_3 + r_4 + r_5 + r_6 + r_7)$

Table 2.63: Placement 4.2.22

exactly one bandpass in the second column. At the same time the 1's in the rows we just moved are wasted in the third column. By keeping the second band in the third column a bandpass, we can get another bandpass for the premise of this case. Thus Placement 4.2.22 is optimal.

+ Else if $|m_6| > B$.

  In this case, we can make use of extra B rows of m_6. If $r_5 \geq r_2$, then in Placement 4.2.22, we can make the 1's in the first column consecutive, and there are two bands in each of the second and third columns, one of them is exactly one bandpass. Thus Placement 4.2.22 is optimal. Else if $r_3 + r_5 + r_6 \geq B$. In Placement 4.2.23, since now we have $r_5 < r_2$, the premise of this case is $r_3 + r_5 + r_6 \geq B$, then we have $r_2 + r_3 + r_6 \geq B$. We can get three bandpasses, so Placement 4.2.23 is optimal. Else if $r_3 + r_5 + r_6 + r_7 \geq B$, since $r_3 + r_5 + r_6 < B$, then $r_7 > B - (r_3 + r_5 + r_6)$. In Placement 4.2.24, we can get two bandpasses in each column. Thus Placement 4.2.24 is optimal. Else, we have $OPT = MAX - 1$. Suppose the current instance is I, then we can get another instance $I^{'}$ by changing m_3, m_5, m_7 to m_6. Apparently, $OPT(I^{'}) \geq OPT(I)$. In $I^{'}$, we only have m_2, m_4, m_6, m_8, satisfying that

m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		$m_6^{'}$	1	1	1	$B-(r_2+r_3)$
m_5	1	1	0	r_5		$m_5^{'}$	1	1	0	r_2
m_6	1	1	1	r_6+B	$\Rightarrow$	m_8	1	0	0	r_8
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	r_8		$m_6^{''}$	1	1	1	$r_6+(r_2+r_3)$
						$m_5^{''}$	1	1	0	r_5-r_2
						m_4	0	1	0	r_4

Table 2.64: Placement 4.2.23

m_2	0	0	1	r_2		m_2	0	0	1	r_2
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		m_6	1	1	1	r_6
m_5	1	1	0	r_5	$\Rightarrow$	m_5	1	1	0	r_5
m_6	1	1	1	r_6		m_7	1	0	1	r_7
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8						

Table 2.65: Placement 4.2.24

$r_2^{'}=r_2, r_4^{'}=r_4, r_8^{'}=r_8, r_6^{'}=r_3+r_5+r_6+r_7<B$, and

$$r_2^{'}+r_6^{'}+r_4^{'}=r_2+r_3+r_5+r_6+r_7+r_4<2B,$$

$$r_2^{'}+r_6^{'}+r_8^{'}=r_2+r_3+r_5+r_6+r_7+r_8<2B,$$

$$r_4^{'}+r_6^{'}+r_8^{'}=r_4+r_3+r_5+r_6+r_7+r_8<2B,$$

$$r_2^{'}+r_4^{'}+r_8^{'}+r_6^{'}=r_2+r_4+r_8+2r_3+2r_5+2r_6+2r_7<3B.$$

By Lemma 3, we have $OPT(I^{'}) = MAX(I^{'})-1$. Since $MAX(I^{'}) = MAX(I)$ and $OPT(I^{'}) \geq OPT(I)$, then $OPT(I)=MAX(I)-1$.

m_2	0	0	1	r_2		m_4	0	1	0	r_4
m_3	0	1	1	r_3		m_3	0	1	1	r_3
m_4	0	1	0	r_4		$m_6^{'}$	1	1	1	r_5+r_6
m_5	1	1	0	r_5		$m_7^{'}$	1	0	1	$B-(r_3+r_5+r_6)$
m_6	1	1	1	r_6+B	$\Rightarrow$	m_8	1	0	0	r_8
m_7	1	0	1	r_7		$m_7^{''}$	1	0	1	$r_3+r_5+r_6+r_7-B$
m_8	1	0	0	r_8		m_2	0	0	1	r_2
						$m_6^{''}$	1	1	1	$B-r_5$
						m_5	1	1	0	r_5

Table 2.66: Placement 4.2.25

m_2	0	0	1	r_2		$m_8^{'}$	1	0	0	$B - r_5$
m_3	0	1	1	r_3		m_5	1	1	0	r_5
m_4	0	1	0	r_4		m_4	0	1	0	r_4
m_5	1	1	0	r_5	$\Rightarrow$	m_3	0	1	1	r_3
m_6	1	1	1	r_6		m_6	1	1	1	r_6
m_7	1	0	1	r_7		m_7	1	0	1	r_7
m_8	1	0	0	$r_8 + B$		$m_8^{''}$	1	0	0	$r_5 + r_8$

Table 2.67: Placement 4.2.26

+ Else if $|m_8| > B$ (We can symmetrically consider the case when $|m_4| > B$).
  In this case, we can make use of extra B (1, 0, 0)-rows. If $r_3 + r_6 + r_7 \geq B$, then we can get one bandpass without using the r_2 (0, 0, 1)-rows in the third column. In Placement 4.2.25, we can get four bandpasses. Thus Placement 4.2.25 is optimal. Else, we have $OPT = MAX - 1$. Suppose we can get MAX bandpasses, then we should get $q_8 + 1$ bands in the first column, $q_4 + 1$ bands in the second column, $q_2 + 1$ bands in the third column and the total number of rows is

  $$T = (q_2 + q_4 + q_8)B + r_2 + r_3 + r_4 + r_5 + r_6 + r_7 + r_8.$$

  Since $(q_4+q_8)B+r_3+r_4+r_5+r_6+r_7+r_8 < (q_4+q_8+2)B$, the bandpasses in the first column must overlap with that in second column. The bandpasses in third column can not make use of the overlapping area of the first and second columns. Otherwise, we will have $r_3 + r_5 + r_6 \geq B$ or $r_5 + r_6 + r_7 \geq B$ or $r_3 + r_6 + r_7 \geq B$, which is a contradiction. If we can get $q_2 + 1$ bandpasses in the third column, then the total number of rows should be

  $$T^{'} = (q_2+1)B + q_4B + r_4 + q_8B + r_8 + 2B - (r_3 + r_4 + r_5 + r_6 + r_7 + r_8).$$

  Because $T^{'} - T = 3B - (r_2 + r_4 + r_8 + 2r_3 + 2r_5 + 2r_6 + 2r_7) > 0$, which is a contradiction. Thus $OPT = MAX - 1$.

+ Else if $r_3 + r_6 + r_7 \geq B$ and $r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$.
  It means we can get one bandpass without using m_2 in the third column. For $r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$, we can get two bandpasses in each of the first two columns by adjusting the matrix with m_5, which is shown in Placement 4.2.26. Thus Placement 4.2.26 is optimal.

+ Else, we have $OPT = MAX - 1$.
  Suppose we can get MAX bandpasses, then we should get one bandpass in the first column, one bandpass in the second column, $q_2 + 1$ bandpasses in the third

m_2	0	0	1	r_2		$m_5^{'}$	1	1	0	x
m_3	0	1	1	r_3		m_4	0	1	0	r_4
m_4	0	1	0	r_4		m_3	0	1	1	r_3
m_5	1	1	0	r_5	$\Rightarrow$	m_6	1	1	1	r_6
m_6	1	1	1	r_6		m_7	1	0	1	r_7
m_7	1	0	1	r_7		m_8	1	0	0	r_8
m_8	1	0	0	r_8		$m_5^{''}$	1	1	0	$r_5 - x$

Table 2.68: Placement 4.2.27

column and the total number of rows is

$$T = q_2B + r_2 + r_3 + r_4 + r_5 + r_6 + r_7 + r_8.$$

Since $r_3+r_4+r_5+r_6+r_7+r_8 < 2B$, the bandpasses in the first column must overlap with that in the second column. The bandpasses in the third column can not make use of the overlapping area of the first and third columns. Otherwise, we will have $r_3 + r_5 + r_6 \geq B$ or $r_5 + r_6 + r_7 \geq B$ or $r_3 + r_6 + r_7 \geq B$ and $r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$, which is a contradiction. If we can get $q_2 + 1$ bandpasses in the third column, then the total number of rows should be

$$T^{'} = (q_2 + 1)B + r_4 + r_8 + r_5 + 2B - (r_3 + r_4 + 2r_5 + r_6 + r_7 + r_8).$$

Because $T^{'} - T = 3B - (r_2 + r_4 + r_8 + 2r_3 + 2r_5 + 2r_6 + 2r_7) > 0$, which is a contradiction. Thus, $OPT = MAX - 1$.

* Else if $|m_6| > B$.

  This case is the same as the subcase ($|m_6| > B$) in the previous case.

* Else if $r_3+r_4+r_5 \geq B$ and $r_2+r_3+r_5+2r_6+r_7+r_8 \geq 2B$ (We can symmetrically consider the cases when $r_5 + r_6 + r_7 \geq B$ and $r_2 + r_3 + r_4 + r_5 + 2r_6 + r_7 \geq 2B$ or $r_3 + r_6 + r_7 \geq B$ and $r_3 + r_4 + r_5 + 2r_6 + r_7 + r_8 \geq 2B$), we can get MAX bandpasses, the solution is similar to Placement 4.2.26.

* Else, we have $OPT = MAX - 1$.

  Now we have $\forall m_i \in \{m_2, m_3, m_4, m_5, m_6, m_7, m_8\}$, s.t. $|m_i| = r_i$, we are supposed to get one bandpasses in each column, and they pairwise overlap, which is impossible. Thus, we can get $MAX - 1$ bandpasses at most.

Through **Case 1** to **Case 4**, we have considered a complete set of subcases for a given instance of the 3-column Bandpass problem. For each subcase, we have presented the optimal placement and proved its optimality. Because the complete set of subcases is finite, the 3-column Bandpass problem can be solved in linear time. This conclusion is proved by Thereome 12 in Chapter 4.

Chapter 3

A 3-approximation Algorithm for the CMSR Problem [2]

In this chapter, we present our 3-approximation algorithm for the CMSR problem in three sections. In Section 3.1, we introduce this problem through an example, and define some terms which are frequently used in later sections. In Section 3.2, we present some structural properties of this problem, based on which we designed the 3-approximation algorithm. At last, the algorithm and the proof of its approximation ratio are presented in Section 3.3.

3.1 Preliminaries

In the sequel, we use a lower case letter to denote a gene marker. A negation sign together with the succeeding gene indicate that the gene is in its reversal and negated form. We reserve the "•" symbol for connection use. For example, $a \bullet b$ means gene b comes directly after gene a. When a common substring (also called strip, or synthetic block) of G_1 and G_2 is identified, we will (often) label it using a capital letter. We abuse this capital letter a bit to also denote the set of genes in the substring.

We first look at an example instance of the CMSR problem (which is also an instance of the MSR problem), in which

$$G_1 =< a, b, c, d, e, f, g, h, i, j, k, l >,$$

$$G_2 =< -i, -d, -g, -f, h, a, c, b, -l, -k, -j, -e > .$$

(we use commas to separate the gene markers for easier reading). By deleting markers c, d, e, h from both G_1 and G_2, the resultant subsequences are

$$G_1^* =< a, b, f, g, i, j, k, l >,$$

$$G_2^* =< -i, -g, -f, a, b, -l, -k, -j > .$$

[2]The main result in this chapter appears as "H. Jiang, Z. Li, G. Lin, L. Wang, B. Zhu. Exact and approximation algorithms for the complementary maximal strip recovery problem. *Journal of Combinatorial Optimization*, (Nov 2010)." [11]

These two resultant subsequences can be decomposed into three maximal substrings $S_1 = a \bullet b$, $S_2 = f \bullet g \bullet i$ (appearing in the reversal and negated form in G_2^*), and $S_3 = j \bullet k \bullet l$ (appearing in the reversal and negated form in G_2^*). For this small instance, one can prove that the optimal solution to the MSR problem has size 8, and (consequently) the optimal solution to the CMSR problem has size 4.

We use OPT to denote an optimal solution to the instance of the CMSR problem. That is, OPT is a minimum-size subset of letters that, deleting them from G_1 and G_2 gives the remainder sequences G_1^* and G_2^*, respectively, which can be partitioned into maximal common substrings.

Given any CMSR instance, in at most quadratic time, we can determine all maximal common substrings of length at least two in G_1 and G_2 and the remaining are isolated letters. Note that the quadratic time could be improved to a linear time, with proper data structure such as suffix-tree. We use unit to refer to a maximal common substring or an isolated letter. A unit and its reversed negated form are considered identical. The above determined units form a common partition of G_1 and G_2, i.e. every letter in G_1 occurs in exactly one of these substrings. For ease of presentation, the maximal common substrings are called type-0 substrings; the isolated letters are called isolates.

In our algorithm Approx-CMSR, all type-0 substrings are kept in the resultant sequences and our goal is to eliminate the isolates, by deleting them to "merge" some letters into substrings. Here "merge" refers to either appending an isolate to some existing substring, or merging two isolates into a novel common substring.

3.2 Structural Properties

Lemma 4 *For any CMSR instance, there exists an optimal solution OPT such that*

(1) for each type-0 substring S, either $S \subset OPT$ or $S \cap OPT = \emptyset$;

(2) if $|S| \geq 4$, then $S \cap OPT = \emptyset$.

PROOF. For a type-0 substring S, assume to the contrary that some but not all of its letters are in OPT. We know that the letters of $S - OPT$ appear consecutively in both G_1^* and G_2^*, and they form or participate in a single maximal substring, denoted as T. We may put letters of $S \cap OPT$ back to G_1^* and G_2^* according to their positions in G_1 and G_2, respectively. These letters do not break but participate in the maximal substring T. This contradicts the optimality of OPT. Therefore, either $S \subset OPT$, or $S \cap OPT = \emptyset$.

If S has length of 4 or greater and $S \subset OPT$, we again put the letters of S back to G_1^* and G_2^* according to their positions in G_1 and G_2, respectively. This added S, as a consecutive segment, might break into maximal substrings of G_1^* and G_2^* to give rise to at most 4 distinct letters that no longer belong to any maximal substrings. Since S becomes a (or part of a) maximal common substring, we can delete the (at least 4) letters of S from OPT while adding to OPT the (at most 4) letters that fall out of maximal substrings. The added letters certainly do not belong to any type-0 substrings. Therefore, this letter-swapping process gives another optimal solution that contains one less type-0 substring of length at least 4. Repeating the same argument if necessary, at the end we will achieve an optimal solution that does not contain any type-0 substring of length at least 4. □

The above Lemma 4 tells that for every type-0 substring, either all its letters are kept in OPT or none of them is in OPT. Thus we can partition OPT into a subset O_3 of length-3 type-0 substrings, a subset O_2 of length-2 type-0 substrings and a set O_1 of isolates, then we have $OPT = O_3 \cup O_2 \cup O_1$. It follows that the number of letters in OPT is

$$|OPT| = 3|O_3| + 2|O_2| + |O_1| \tag{3.1}$$

3.3 The Approximation Algorithm

By Lemma 4, all type-0 substrings of length 4 and greater are retained in our approximation algorithm to be presented next. The output of our algorithm will be compared against an optimal solution OPT which also retains all these substrings. In the following, we only deal with length-3 and length-2 type-0 substrings, and isolates.

Here we use an example to illustrate our 3-step greedy algorithm. Given

$$G_1 =< a, x, b, u, e, f, g, i, j, h, y, k, c, z, d, v >,$$

$$G_2 =< e, z, f, g, v, x, h, k, a, b, c, d, u, i, y, j > .$$

In the first step, our algorithm maps all maximal common substrings and retains all type-0 substrings. This can be shown by Figure 3.1, in which we reserved $\{fg\}$ in Step 1. In the second step, our algorithm recursively removes a target isolate, denoted as u; such a removed isolate has to satisfy the condition (C) listed in the following, with the goal that removing it from (the current) G_1 and G_2 gives rise to (at least) a new common substring of length 2. This procedure may consist of several iterations. We continue use the example. In the first iteration, our algorithm deleted x and merged a and b into a substring ab. This can be shown by Figure 3.2. Then in the second iteration, it deleted y and reserved hk, ij. In the third iteration, it deleted z to merge c, d and attach e to an existing substring fg. These two iterations can be illustrated by Figure 3.3 and 3.4 separately. Each of these new generated common substrings is not a common substring to the original G_1 and G_2, thus is called a type-1 substring for distinction purpose. Note that after such isolate removal, some units (type-0 and/or type-1 substrings, and/or isolates) might be able to be merged into longer maximal common substrings. For consistency we do not merge two existing substrings; but we will append isolates to existing substrings (type-0 or type-1) whenever possible, since our goal is to get rid of isolates. These appended isolates are no longer isolates, and the extended substrings keep their type (type-0 or type-1). When none of the isolates satisfying condition (C) can be identified, the algorithm enters the last step to remove all the remaining isolates, if any. This can be shown by Figure 3.5.

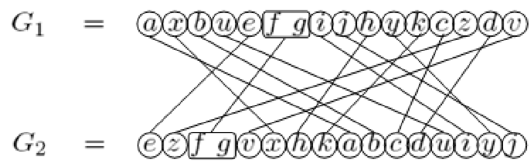

Figure 3.1: Step 1: Mapping maximal common substrings.

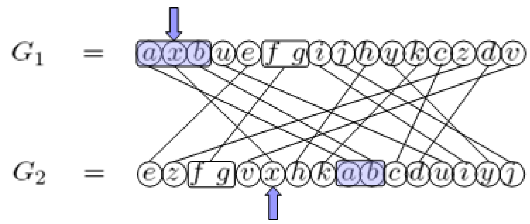

Figure 3.2: Step 2: Deleting x and reserve a,b.

Definition 5 *Condition (C): In either G_1 or G_2, two neighboring units of u are also isolates; and after removing u, they form into a type-1 common substring of length* 2*.*

It could be the case that in both G_1 and G_2, the two neighboring units of u form into a type-1 common substring of length 2 after deleting u; our algorithm will identify the case and subsequently all these isolates become no longer isolates. There is another (disjoint) case in which, besides forming the type-1 common substring of length 2, another neighboring isolate of u in the different sequence can be appended to an existing, or the newly formed, substring; our algorithm will identify this case too and subsequently the appended isolate becomes no longer an isolate. Intuitively, removing isolate u saves (i.e., retains) at least two other isolates, and can save one or two more isolates.

For ease of discussion, let $U = \{u_1, u_2, ..., u_m\}$ denote the set of isolates located in sequential order by our algorithm, which are all removed. Associated with each u_j , let V_j denote the set of neighboring isolates of u_j in the current G_1 and G_2 that become no longer isolates after removing u_j. We have $|V_j| \geq 2$, for $j = 1, 2, ..., m$. In particular, the two neighboring isolates of u_j that form a type-1 substring after deleting u_j are denoted as a_j and b_j (where there are two such pairs, a_j and b_j refer to an arbitrary one of them). Let R denote the set of remaining isolates at the time

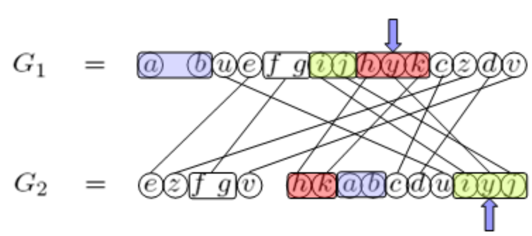

Figure 3.3: Step 2: Deleting y and reserve h,i,j,k.

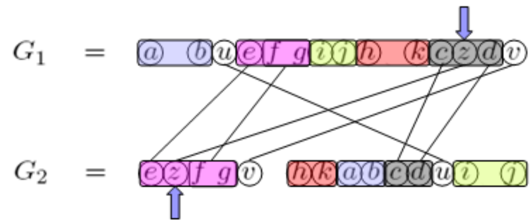

Figure 3.4: Step 2: Deleting z and reserve c,d,e.

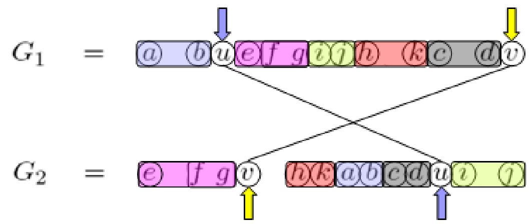

Figure 3.5: Step 3: Deleting all remaining isolates.

the algorithm finds no isolates satisfying condition (C); that is, R is the set of isolates deleted by our algorithm at the last step. The following two lemmas state some preliminary observations.

Lemma 5 *The set of all isolates I is the union of the disjoint sets $U, V_1, V_2, ..., V_m$, and R, that is, $I = U \cup (\bigcup_{j=1}^{m} V_j) \cup R$; moreover, the algorithm deletes all isolates of $U \cup R$, but no others.*

Lemma 6 *In the original input sequences G_1 and G_2, the letters in between a_j and b_j all belong to $\{u_1, u_2, ..., u_{j-1}, u_j\}$; moreover, u_j is in between a_j and b_j in exactly one of G_1 and G_2.*

Recall that we use in the discussion an optimal solution OPT which satisfies the two properties listed in Lemma 4. Consider the inverse process of deleting units of OPT from G_1 and G_2 to obtain the final sequences G_1^* and G_2^* . In this inverse process, we add the units of OPT back to G_1^* and G_2^* using their original positions in G_1 and G_2 to re-construct G_1 and G_2. At the beginning of this process, there are no isolated letters in G_1^* or G_2^*; all the isolates of I are thus either units of $I \cap O_1$, or generated by inserting units of OPT back, which break the maximal common substrings into fragments of which some are single letters. At any time of the process, inserting one unit of OPT back to the current G_1 and G_2 can generate at most four fragments of single letters, since in the worst case two current length-2 substrings can be broken into four such fragments. Some of these single letters might not be the isolates of $U \cup R$; those that are in $U \cup R$, as well as the inserted unit when it belongs to $(U \cup R) \cap O_1$, are said to be associated with the inserted unit of OPT.We firstly insert units of O_3 and O_2, one by one; each of them is associated with at most four isolates of $U \cup R$ (Lemma 7); the resultant sequences are denoted as G_1^0 and G_2^0.

Lemma 7 *The number of isolates of $U \cup R$ associated with each unit of $O_3 \cup O_2$ is at most four.*

Next, we insert isolates of $O_1 \cap (u_j \cup V_j)$ back into G_1^0 and G_2^0, for $j = 1, 2, ..., m$ sequentially. At the end of the inserting isolates of $O_1 \cap (u_j \cup V_j)$, the resultant sequences are denoted as G_1^j and G_2^j.We emphasize that this sequential order is very important, as we need it in the proofs of Lemmas 8 and 9. Lemma 9 counts the average number of isolates of $U \cup R$ associated with each isolate of $O_1 \cap (u_j \cup V_j)$.

Lemma 8 *For any j , u_j is an isolated letter in G_1^j and G_2^j.*

PROOF. We prove this lemma by (finite) induction. Firstly, we notice that a_1, b_1, and u_1 cannot co-exist in G_1^* and G_2^* , since otherwise u_1 would be the only letter in between a_1 and b_1 in exactly one of G_1^* and G_2^* , and thus an isolated letter. Therefore, $O_1 \cap (u_1 \cup V_1) \neq \emptyset$. After inserting isolates of $O_1 \cap (u_1 \cup V_1)$ back, a_1, b_1, and u_1 are all present in G_1^1 and G_2^1. For the same reason that u_1 is the only letter in between a_1 and b_1 in exactly one of G_1^1 and G_2^1, u_1 is an isolated letter. That is, the lemma holds for $j = 1$.

Assume the lemma holds for all $i = 1, 2, ..., j - 1$, that is, $u_1, u_2, ..., u_{j-1}$ are isolated letters in G_1^{j-1} and G_2^{j-1} , and thus they are all isolated letters in G_1^j and G_2^j. Due to the co-existence of a_j, b_j , and u_j in G_1^j and G_2^j, Lemma 6 tells that if u_j is not an isolated letter, then it can only pair with some letter of $\{u_1, u_2, ..., u_{j-1}\}$ to sit together in a substring. This is a contradiction to the inductive assumption. Therefore, u_j is an isolated letter in G_1^j and G_2^j. □

Lemma 9 *For any j , the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most* 2.5*. Moreover, by the end of this iteration of inserting process, u_j is associated to some unit of OPT.*

PROOF. Recall that we insert isolates of $O_1 \cap (u_j \cup V_j)$ back into G_1^0 and G_2^0 in sequential order of j. When we start to insert isolates of $O_1 \cap (u_j \cup V_j)$, all isolates of $O_1 \cap \left(\cup_{i=1}^{j-1} u_i \cup V_i\right)$ have been inserted and the resultant sequences are G_1^{j-1} and G_2^{j-1}.

Firstly, if $O_1 \cap (u_j \cup V_j) = \emptyset$, then the lemma is proved automatically. So we assume in the following that $O_1 \cap (u_j \cup V_j) \neq \emptyset$. Let a_j and b_j be the two neighboring isolates of u_j when the approximation algorithm located u_j, as in Lemma 6, such that by removing u_j, $a_j \cdot b_j$ became a type-1 length-2 substring. We consider the following two disjoint cases: $u_j \in O_1$ and $u_j \notin O_1$.

In the first case, $u_j \in O_1$. When $a_j, b_j \in O_1$ and a_j and b_j are separated by certain letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1 (G_2, respectively), inserting a_j and b_j into G_1^{j-1} (G_2^{j-1}, respectively) does not generate any new isolates of $U \cup R$; when $a_j, b_j \in O_1$ and a_j and b_j are separated by no letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1 (G_2, respectively), inserting a_j and b_j into G_1^{j-1} (G_2^{j-1}, respectively) can generate at most two isolates of $U \cup R$. When one and only one of a_j and b_j is in O_1, then inserting it into G_1^{j-1} and G_2^{j-1} does not generate any new isolates of $U \cup R$.

If $|V_j| = 4$, then the other two letters, c_j and d_j, have the same properties as a_j and b_j. When $|V_j \cap O_1| = 4$, that is, $a_j, b_j, c_j, d_j \in OPT$, inserting a_j, b_j and c_j, d_j can generate at most 8 new isolates of $U \cup R$; When $|V_j \cap O_1| = 3$, and assuming $a_j, b_j, c_j \in OPT$, inserting a_j, b_j can generate at most 4 new isolates of $U \cup R$, but inserting c_j generates no new isolates of $U \cup R$; When $|V_j \cap O_1| = 2$, and in the first scenario assuming $a_j, b_j \in OPT$, inserting a_j, b_j can generate at most 4 new isolates of $U \cup R$; in the second scenario assuming $a_j, c_j \in OPT$, inserting a_j, c_j generates no new isolates of $U \cup R$; When $|V_j \cap O_1| = 1$, and assuming $a_j \in OPT$, inserting a_j generates no new isolates of $U \cup R$. After inserting isolates of $O_1 \cap V_j$, if any, inserting u_j back into the current G_1^{j-1} and G_2^{j-1} does not generate any new isolates of $U \cup R$. In summary, for $|O_1 \cap V_j| = 4, 3, 2, 1,$ and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $8, 4, 4, 0,$ and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $8/5$.

If $|V_j| = 3$, then the third letter, c_j, was appended to an existing (type-0 or type-1) substring S when the approximation algorithm removed u_j. Similarly to the discussion on a_j and b_j, c_j and S can only be separated by letters of $\{u_1, u_2, \ldots, u_{j-1}\}$, besides u_j, in G_1 and G_2. Moreover, u_j is in between c_j and S in at most one of G_1 and G_2. Therefore, when $c_j \in O_1$, inserting it into G_1^{j-1} and G_2^{j-1} can generate at most one new isolate of $U \cup R$. After inserting isolates of $O_1 \cap V_j$, if any, inserting u_j back into the current G_1^{j-1} and G_2^{j-1} does not generate any new isolates of $U \cup R$. Therefore, for $|O_1 \cap V_j| = 3, 2, 1,$ and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $5, 4, 1,$ and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $4/3$.

If $|V_j| = 2$, after inserting isolates of $O_1 \cap V_j$, if any, inserting u_j back into the current G_1^{j-1} and G_2^{j-1} can generate at most two isolates of $U \cup R$. Therefore, for $|O_1 \cap V_j| = 2, 1,$ and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $6, 2,$ and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most 2.

In the second case, $u_j \notin O_1$. Assume without loss of generality that u_j is in between a_j and b_j in G_1 in Lemma 7. When $a_j \in O_1$ ($b_j \in O_1$, respectively) and a_j (b_j, respectively) and u_j are separated by certain letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1, inserting a_j (b_j, respectively) into G_1^{j-1} does not generate any new isolates of $U \cup R$. When $a_j \in O_1$ ($b_j \in O_1$, respectively) and a_j (b_j, respectively) and u_j are separated by no letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1, inserting a_j (b_j, respectively) into G_1^{j-1} can generate at most two isolates of $U \cup R$, including u_j. Nonetheless, when $a_j, b_j \in O_1$ and a_j and b_j are separated by no letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1, inserting a_j and b_j into G_1^{j-1} can generate at most three isolates of $U \cup R$, including u_j. Similarly, when $a_j, b_j \in O_1$ and a_j and b_j are separated by certain letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_2, inserting a_j and b_j into G_2^{j-1} does not generate any new isolates of $U \cup R$; when $a_j, b_j \in O_1$ and a_j and b_j are separated by no letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_2, inserting a_j and b_j into G_2^{j-1} can generate at most two isolates of $U \cup R$.

If $|V_j| = 4$, then the other two letters, c_j and d_j, have the same properties as a_j and b_j. Note that when inserting a_j and b_j into G_1^{j-1} generates new isolates of $U \cup R$, these isolates will be seen again when inserting c_j and d_j into G_2^{j-1}. Therefore, for $|O_1 \cap V_j| = 4, 3, 2, 1$, and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $7, 4, 2, 0$, and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $7/4$.

If $|V_j| = 3$, then the third letter, c_j, was appended to an existing (type-0 or type-1) substring S when the approximation algorithm removed u_j. Similarly to the discussion on a_j and b_j, c_j and S can only be separated by letters of $\{u_1, u_2, \ldots, u_{j-1}\}$ in G_1 and G_2, besides u_j in G_2. Therefore, when $c_j \in O_1$, inserting c_j into G_2^{j-1} can generate at most one new isolate of $U \cup R$, which will be seen when inserting b_j into G_1^{j-1}. Note that S might start with a_j or end with b_j. For $|O_1 \cap V_j| = 3, 2, 1$, and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $4, 2, 0$, and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $4/3$.

If $|V_j| = 2$, for $|O_1 \cap V_j| = 2, 1$, and 0, respectively, the total number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $5, 2$, and 0, respectively. It follows that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ is at most $5/2$.

From the above case analysis, we conclude that the average number of isolates of $U \cup R$ associated with isolates of $O_1 \cap (u_j \cup V_j)$ in the worst case is $5/2 = 2.5$. □

Lastly, we insert isolates of $O_1 \cap R$ back into G_1^m and G_2^m. At the end of this last inserting process, we achieve the input sequences G_1 and G_2.

Lemma 10 *The average number of isolates of $U \cup R$ associated with each isolate in $O_1 \cap R$ is at most* 3.

PROOF. The key fact used in the proof is that after locating isolate u_m, removing it from the current sequences, and making letters in V_m non-isolates, the approximation algorithm finds no more isolates to iterate the process. That is, for any two remaining isolates $r, s \in R$ that are not separated by any existing (type-0 or type-1) substring in both sequences (that is, r and s can potentially form into a substring, or participate together), there are at least two other isolates, duplications are separately counted, in between them, counting from both sequences.

In sequences G_1^m and G_2^m obtained after inserting units of $O_3 \cup O_2 \cup \left(O_1 \cap \left(U \cup \cup_{j=1}^m V_j\right)\right)$ into G_1^* and G_2^*, some units of R are already isolates, while the other reside in substrings (of length at least two). These units residing in substrings are to be singled out by inserting units of $O_1 \cap R$ into G_1^m and G_2^m; and it is these units that are associated with isolates of $O_1 \cap R$.

Let $S_1, S_2, \ldots, S_k$ denote the substrings in G_1^m and G_2^m that are made of isolates of R; and $T_1, T_2, \ldots, T_\ell$ denote the fragments of substrings in G_1^m and G_2^m, where the substrings are not purely made of isolates of R, but the fragments are. Note that $|S_i| \geq 2$ for every i. To single out all letters of $\left(\cup_{i=1}^k S_i\right) \cup \left(\cup_{j=1}^\ell T_j\right)$, we first need at least one isolate of $O_1 \cap R$ to chop each T_i off its host substring; Afterwards, the above argument states that for every two adjacent letters in S_i or T_j, there are at least two isolates of $O_1 \cap R$ in between them, counting from both sequences. This gives a lower bound on the minimum number of isolates of $O_1 \cap R$. Since each isolate of $O_1 \cap R$ can appear in two places, we have

$$2|O_1 \cap R| \geq \ell + \sum_{i=1}^{k} 2(|S_i| - 1) + \sum_{j=1}^{\ell} 2(|T_j| - 1) \geq \sum_{i=1}^{k} |S_i| + \sum_{j=1}^{\ell} |T_j|.$$

Therefore, the total number of isolates of $U \cup R$ (in this case, R only) that are associated with isolates of $O_1 \cap R$ is at most $\sum_{i=1}^{k} |S_i| + \sum_{j=1}^{\ell} |T_j| + |O_1 \cap R|$, which is less than or equal to $3|O_1 \cap R|$. This proves the lemma. □

Theorem 11 *The CMSR problem admits a* 3*-approximation algorithm.*

PROOF. To summarize, all isolates of $U \cup R$ are associated with units of OPT. From Lemmas 7, 9, and 10, we have

$$|U \cup R| \leq 4|O_3 \cup O_2| + 2.5|O_1 \cap \left(U \cup \left(\cup_{j=1}^{m} V_j\right)\right)| + 3|O_1 \cap R| \leq \frac{4}{3} \times 3|O_3| + 2 \times 2|O_2| + 3 \times |O_1| \leq 3|OPT|,$$

where $|OPT|$ denotes the number of letters in OPT and thus $|OPT| = 3|O_3| + 2|O_2| + |O_1|$. Note that the approximation algorithm deletes all isolates of $U \cup R$, but no others, and therefore it is a 3-approximation algorithm. □

Chapter 4

Conclusions and Future Work

In this chapter, we first summarize our results on the Bandpass and the CMSR problems, respectively. Then we introduce the future work for these two combinatorial optimization problems.

In Chapter 2, we presented our Remainder-Driven algorithm for the 3-column Bandpass problem through a case by case analysis style. Now we can conclude the following Theorem 12.

Theorem 12 *The three column Bandpass problem with any bandpass number $B \geq 2$ can be solved exactly in linear time.*

PROOF. In analysis from Case 1 to Case 4 in Chapter 2, we have shown that in most cases, the six solutions returned from the Row-Stacking algorithm include an optimum; all the exceptional cases are recognized and solved by the Remainder-Driven algorithm in Section 2.2. Because the complete set of subcases for a 3-column instance considered by the Remainder-Driven algorithm is finite, the 3-column Bandpass problem can be solved in linear time. □

In Chapter 3, we presented a 3-approximation algorithm for the CMSR problem. The key design technique is greedy, and the performance ratio is proven using a novel inverse amortized analysis, through which we can construct a mapping between our algorithm's solution and the optimal solution.

In the future, we will investigate the general Bandpass problem which is proven to be NP-hard. For a general matrix A and $B = 2$ instance, a 2-approximation algorithm [15] was proposed based on results of the maximum weighted set packing problem [7, 1]. But there is no constant ratio approximation algorithm for the given matrices with $B > 2$, and we may work on it. For the CMSR problem, we are currently working on improved approximation algorithms. Furthermore, it is more practical and challenging to deal with multiple genome sequences. Our greedy algorithm still works, but we have to examine if the approximation ratio is 3. Although still given two genome sequences, it is much more difficult to deal with if the makers in each sequence can occur more than once. In this case, we have to revise the weight function and design a new scheme for computing gains in our greedy algorithm. Also we need to check and prove its approximation ratio.

Bibliography

[1] E. Arkin and R. Hassin. On local search for weighted packing problems. *Mathematics of Operations Research*, 23:640–648, 1998.

[2] D. Babayev, G. Bell, and U. Nuriyev. The Bandpass problem: combinatorial optimization and library of problems. *Journal of Combinatorial Optimization*, 18:151–172, 2009.

[3] G. Bell and D. Babayev. Bandpass problem. *In Annual INFORMS meeting, 2004. October, 2004, Denver, CO, USA*.

[4] J. Bowers, C. Abbey, and S. Anderson. A high-density genetic recombination map of sequence tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics*, 165:367–386, 2003.

[5] L. Bulteau, G. Fertin, and I. Rusu. Maximal strip recovery problem with gaps: hardness and approximation algorithms. *In: Proceedings of the 20th Annual International Symposium on Algorithms and Computation (ISAAC'09)*, 5878:710–719, 2009.

[6] A. Butman, D. Hermelin, M. Lewenstein, and D. Rawitz. Optimization problems in multiple-interval graphs. *In: Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'07)*, pp 268–277, 2007.

[7] B. Chandra and M. Halldorsson. Greedy local improvement and weighted set packing approximation. *In: Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'99)*, pp 169–176, 1999.

[8] Z. Chen, B. Fu, M. Jiang, and B. Zhu. On recovering synthetic blocks from comparative maps. *Journal of Combinatorial Optimization*, 18:307–318, 2009.

[9] V. Choi, C. Cheng, Q. Zhu, and D. Sankoff. Algorithms for the extraction of synteny blocks from comparative maps. *In: Proceedings of the 7th International Workshop on Algorithms in Bioinformatics (WABI'07)*, pp 277–288, 2007.

[10] S. Hannenhalli and P. Pevzner. To cut or not to cut (applications of comparative physical maps in molecular evolution). *In: Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'96)*, pp 304–313, 1996.

[11] H. Jiang, Z. Li, G. Lin, L. Wang, and B. Zhu. Exact and approximation algorithms for the complementary maximal strip recovery problem. *Journal of Combinatorial Optimization*, 2010. DOI 10.1007/s10878-010-9366-y.

[12] M. Jiang. Inapproximability of maximal strip recovery. *In: Proceedings of the 20th Annual International Symposium on Algorithms and Computation (ISAAC'09)*, 5878:616–625, 2009.

[13] M. Jiang. Inapproximability of maximal strip recovery II. *In: Proceedings of the 4th Annual Frontiers of Algorithmics Workshop (FAW'10)*, 6213:53–64, 2010.

[14] Z. Li and G. Lin. The three column Bandpass problem is solvable in linear time. *Theoretical Computer Science*, 412:281–299, 2011.

[15] G. Lin. On the Bandpass problem. *Journal of Combinatorial Optimization*, 2009. DOI 10.1007/s10878-009-9273-2.

[16] P. Ostergard. A fast algorithm for the maximum clique problem. *Discrete Applied Mathematics*, 120:195–205, 2002.

[17] B. Reuven, M. Halldorsson, J. Naor, H. Shachnai, and I. Shapira. Scheduling split intervals. *SIAM Journal on Computing*, 36:1–15, 2006.

[18] S. Vialette. On the computational complexity of 2-interval pattern matching problems. *Theoretical Computer Science*, 312:223–249, 2004.

[19] L. Wang and B. Zhu. On the tractability of maximal strip recovery. *Journal of Computational Biology*, 17:907–914, 2010.

[20] G. Watterson, W. Ewens, T. Hall, and A. Morgan. The chromosome inversion problem. *Journal of Theoretical Biology*, 99:1–7, 1982.

[21] S. Yancopoulos, O. Attie, and R. Friedberg. Efficient sorting of genomics permutations by translocation, inversion and block interchange. *Bioinformatics*, 21:3340–3346, 2005.

[22] C. Zheng, Q. Zhu, and D. Sankoff. Removing noise and ambiguities from comparative maps in rearrangement analysis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 4:515–522, 2007.

[23] B. Zhu. Efficient exact and approximate algorithms for the complement of maximal strip recovery. *Algorithmic aspects in Information and Management*, 6124:325–333, 2010.