Scalable parallel rank order filters

Citation for published version (APA):

Document status and date:
Published: 01/01/2007

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 15. Oct. 2019
Scalable Parallel Rank Order Filters

M.G. van der Horst       R.H. Mak

February 26, 2007
Contents

1 Introduction 3

2 Problem Description 6
   2.1 Mathematical Definition 6
   2.2 Complexity measure 9

3 Problem Approach 10
   3.1 Divide and Conquer 10
   3.2 Divide, Re-use and Conquer 12
   3.3 Calculation Graph 13
   3.4 Cost of Merging Networks 14
   3.5 Mirroring 17

4 Generator Graphs 19
   4.1 Preliminaries 19
   4.2 Canonical generator graphs 21
   4.3 Cost calculation 23
   4.4 Constructing generator graphs 24

5 Heuristic Algorithms 27
   5.1 Preliminaries 27
   5.2 Divide and Conquer 28
   5.3 Recursive Tiling 29
   5.4 Bit-pattern Tiling 30
   5.5 Split Tiling 32
   5.6 Overlap Recursive Mirroring 35
   5.7 Multiple Dimensions 41
   5.8 Maximum reuse 42

6 Results 43
   6.1 Filter classes 43
   6.2 Heuristic algorithm evaluation 44
   6.3 Comparison with known filters 51
      6.3.1 Window sorters 51
      6.3.2 Median filter 52
      6.3.3 Minimum filter 53
   6.4 Summary 53
## 7 Parallel ROF Implementations

7.1 Sequential Implementation ........................................ 55
7.2 Parallel Implementation ............................................. 58
   7.2.1 VLSI Circuit .................................................. 58
   7.2.2 SIMD Processor ............................................... 60
7.3 Finite input streams .................................................. 62
7.4 Philips EVP16 .......................................................... 63

## 8 Conclusion ................................................................. 65

## Bibliography ................................................................. 66

## A Algorithm Results ....................................................... 68

A.1 One Dimensional Problems ........................................... 68
   A.1.1 Window Sorting .................................................. 69
   A.1.2 Median Filtering ............................................... 71
   A.1.3 Minimum Filtering ............................................. 74

A.2 Two Dimensional Problems ........................................... 77
   A.2.1 Window Sorting .................................................. 77
   A.2.2 Median Filtering ............................................... 77
   A.2.3 Minimum Filtering ............................................. 77

## B Sample solutions .......................................................... 79

B.1 Representation of the ROFs .......................................... 79
B.2 One Dimensional Problems .......................................... 79
   B.2.1 Window Sorting .................................................. 79
   B.2.2 Median Filtering ............................................... 81
   B.2.3 Minimum Filtering ............................................. 82

B.3 Two Dimensional Problems .......................................... 84
   B.3.1 Window Sorting .................................................. 84
   B.3.2 Median Filtering ............................................... 85
   B.3.3 Minimum Filtering ............................................. 87
Chapter 1

Introduction

Rank order filters are non-linear filters used in a wide range of applications. The median filter is perhaps the most common of the rank order filters. It can be used to remove noise from a signal while preserving edge information (see Figure 1.1), which is something that linear filters cannot do. This makes the median filter a popular filter in speech and image processing applications.

![Figure 1.1: Median filtering](image)

(a) Original SAN logo  
(b) Noisy logo  
(c) Noisy logo after a $7 \times 7$ median filter

Rank order filters work by sorting a window of inputs and selecting the input with a certain rank as the output. In the case of the median filter the element with the middle rank is selected, hence the name (although it is called a speckle filter in some applications, after the type of noise it removes). Other common rank order filters are the minimum and maximum filter, selecting respectively the input with the lowest and highest rank from the window. In image processing applications these minimum and maximum filters are sometimes also called dilation and erosion filters, because they make dark regions in an image respectively larger or smaller, as shown in Figure 1.2.

Rank order filters take a lot of processing power to implement, a lot more than the well-known linear filters. This is why rank order filters are rarely used in real-time signal processing applications. However, just like linear filters, rank order filters can benefit from parallelism. If enough parallelism is available any
CHAPTER 1. INTRODUCTION

(a) Logo after a $7 \times 7$ dilation/minimum filter  
(b) Original logo  
(c) Logo after a $7 \times 7$ erosion/maximum filter

Figure 1.2: Minimum and maximum filtering

desirable throughput can be achieved.

In this report we take a look at rank order filters and implement them in an efficient, parallel fashion such that the throughput of the implementation scales linearly with the amount of parallelism. What constitutes as “the amount of parallelism” depends on the implementation method used. When implementing the filter in hardware, as a VLSI circuit, the amount of parallelism would be the number of components used. When implementing the filter in software on a SIMD processor, the amount of parallelism is determined by the number of processing elements of the SIMD processor.

Implementing the filters in such a manner that the throughput scales linearly with the amount of parallelism is quite trivial. The calculations for the outputs of the filter have constant complexity and are independent of each other, therefore unlimited parallelism can be achieved. However, more efficient implementations are possible; the calculations of successive outputs have enough in common to allow the reuse of sub-calculations, thereby increasing efficiency. By exploiting this reuse we are able to find the most efficient implementations to date.

The implementation of rank order filters clearly requires some form of sorting. To make practical implementation possible we resort to so-called oblivious sorting algorithms. Oblivious sorting algorithms always perform the same number of operations in the same order, regardless of the data that has to be sorted. The well-known quick-sort algorithm, for example, is not an oblivious algorithm, whereas another well-known sorting algorithm, namely merge sort, is oblivious. We choose oblivious sorting algorithms because they are especially suited both for implementation on SIMD processors, since there is no data-dependency in the execution path, and for hardware implementation, since there is no overhead in the form of control logic.

In the literature there are already a number of parallel designs available for both hardware [2, 3, 9] and SIMD [8] implementations, but we improve upon all of them. The most important contributions described in this report, however, are the optimizations that can be applied to existing filters and the mathematical framework in which the design problem, and its solutions, can be described.
We introduce these contributions the following manner. First we formally define the problem of designing the implementation of a rank order filter in Chapter 2. In Chapter 3 we introduce the so-called calculation graph as an abstract representation of the calculations performed by a rank order filter based on oblivious sorting networks. Since a filter calculates (theoretically) infinitely many outputs, the associated calculation graph is also infinite. However, a calculation graph has regularities and in Chapter 4 we introduce generator graphs which are finite, yet still describe an entire calculation graph. The design of the filter implementation can be read from such a generator graph, but a suitable generator graph has to be found first, which is the subject of Chapter 5. In Chapter 6 we compare our designs with those found in the literature. The practical issues that arise when a design is turned into an actual implementation are discussed in Chapter 7. And finally, there is the conclusion in Chapter 8.
Chapter 2

Problem Description

The problem we address in this report is that of constructing efficient, scalable parallel implementations of rank order filters. We will not limit ourselves to filters that determine a single rank, but allow filters that produce a range of ranks as output.

For implementing rank order filters we use oblivious sorting methods, since the absence of data dependencies facilitates the implementation on SIMD machines and in VLSI circuits.

The outline of this chapter is as follows. The next section introduces a rigorous mathematical formulation of the problem and Section 2.2 gives the complexity measure by which we can judge the various algorithms used to solve the problem.

2.1 Mathematical Definition

A common property of filter applications is that the data to be filtered is organized in some underlying structure. For each location of this structure a filtered value has to be determined based on the original data item at that location and on the values of data items in some finite neighborhood. The precise way in which such a filter operates is given by its kernel, which describes both the structure of the neighborhood and the contribution of the individual values in that neighborhood to the resulting value.

In this report we restrict ourselves to very simple underlying structures, viz. to grid topologies. So, for $D > 0$, let $\mathbb{Z}^D$ be a $D$-dimensional grid with integral coordinates. The notion of a neighborhood is captured by a tile which is nothing but a finite set of grid locations. The size of a tile $T \in \mathcal{P}(\mathbb{Z}^D)$ is the number of grid locations in $T$ and is denoted by $|T|$. We say that two tiles $S$ and $T$ have an equivalent shape when there exists a translation vector $t \in \mathbb{Z}^D$ such that $T = S + t$ with tile $S + t$ defined as $\{s + t \mid s \in S\}$. Note that tiles that differ by a rotation are not considered equivalently shaped. In the sequel we will be interested in sets of tiles with equivalent shapes, laid out in repetitive patterns. So we define the function $\text{tile} : (\mathcal{P}(\mathbb{Z}^D) \times \mathbb{Z}^D \times \mathbb{Z}^D) \rightarrow \mathcal{P}(\mathcal{P}(\mathbb{Z}^D))$ given by:

$$\text{tile}(S, f, p) = \{S + f + g \cdot p \mid g \in \mathbb{Z}^D\} \quad (2.1)$$

In this definition the dot-operator “$\cdot$” stands for coordinate-wise vector multi-
CHAPTER 2. PROBLEM DESCRIPTION

Figure 2.1: The tiles generated by tile({(0, 0), (1, 0), (0, 1), (1, 2), (4, 3)})
CHAPTER 2. PROBLEM DESCRIPTION

Given an arbitrary function \( \text{data} : \mathbb{Z}^D \rightarrow \mathbb{D} \), the rank order filter determines for each \( W \in \mathcal{W} \) the sorted list \( W[l..u] \), that is the elements of \( W \) with ranks ranging from \( l \) up to and including \( u \). □

Note that most rank order filters are designed for a specific value or range of values for the parameters \( D, \mathcal{W}, l \) and \( u \). An image processing application, for example, only needs 2-dimensional rank order filters (\( D = 2 \)). Also, in most cases we have \( l = u \), i.e. rank order filters that select one specific rank. However, the definition above also includes so-called window sorters (\( l = 0 \) and \( u = |S_W| \)) and other intermediate problems which will be useful in designing rank order filters.

Some special cases of rank order filters are:

The 1-D median filter with window size \( 2K + 1 \):
- The number of dimensions is \( D = 1 \)
- The input signal is represented by the \text{data} function
- The shape of the windows is \( S_W = \text{hypercube}(K, D) \)
- The windows themselves are given by \( \mathcal{W} = \text{tile}(S_W, 1, 0) \)
- We are interested in the item with rank \( K \), hence \( l = u = K \)

The 2-D median filter with kernel size \( (2K + 1) \times (2K + 1) \):
- The number of dimensions is \( D = 2 \)
- The input image is represented by the \text{data} function
- The shape of the windows is \( S_W = \text{hypercube}(K, D) \)
- The windows themselves are given by \( \mathcal{W} = \text{tile}(S_W, (1, 1), (0, 0)) \)
- We are interested in the item with rank \( 2K^2 + 2K \), hence \( l = u = 2K^2 + K \)

Note that the function \text{tile} produces an infinite set. In practice the set of windows for a rank order filter is finite. However, for now we will assume the set of windows to be infinite, and save this aspect of practical implementation for Section 7.3.

Using the definition of a rank-order filter we can define what we consider to be a rank-order filter design problem.

Definition 2.1.2 (Rank order filter design problem (ROFDP)). Assume that we are given
- a number \( D \in \mathbb{N}^+ \) that specifies the dimension of the filter,
- a set of similar shaped tiles, called windows \( \mathcal{W} = \text{tile}(S_W, f_W, p_W) \),
- two numbers \( l \) and \( u \) such that \( 0 \leq l \leq u < |S_W| \).

Design a rank order filter for these parameters. □

To solve this problem and to find an algorithm that calculates the rank \([l..u]\) elements of every window in an oblivious manner, we will need some form of sorting network. However, the optimum networks for sorting lists of arbitrary size are unknown. We only know good instances when it comes to minimum-comparison sorting networks [7]. Therefore we are content to resort to heuristic methods to find the implementations of rank order filters.
CHAPTER 2. PROBLEM DESCRIPTION

2.2 Complexity measure

The elementary operation of oblivious sorting is a so-called compare and swap (C&S) operation. This operation effectively sorts a list of two inputs by comparing them and swapping them if necessary. The same result can be obtained by using a minimum and a maximum operator (the so-called extrema operations). Sorting is then accomplished by making the first element of the resulting list the minimum of the two inputs and the last element of the resulting list the maximum of the two inputs.

In this report we will use the number of extrema operations to measure the complexity of a ROF. We do so because the extrema count allows for a more fine-grained determination of the cost. If, for example, we only need one of the outputs of a C&S operation, we still need a single C&S operation. However, whenever two extrema operations are used to implement the C&S operation we can eliminate the one that produces the unneeded output.

Since there is no consensus in literature whether to use extrema or C&S operation counts, one has to be careful when comparing implementations. Even when the difference is taken into account there is no quick translation of costs. An C&S operation consists of two extrema operations, but it is possible that the design discussed in the literature can be optimized by converting all the C&S operations into extrema operations and subsequently removing the unnecessary operations.

However, whether C&S or extrema operations are used to implement a rank order filter, it is beneficial to minimize the operation count. We could have used an other complexity measure like the throughput (i.e. the number of results produced per unit of time). This measure, however, is implementation specific. Moreover, the operation count signifies the amount of work that needs to be done, therefore minimizing the amount of work will speed up any implementation, regardless what particular method of implementation is used.

Therefore the problem we investigate in this report is: What is the minimum number of extrema operations needed for a rank-order filter?
Chapter 3

Problem Approach

In this chapter we analyze the problem, derive some properties and introduce an abstract representation for our solutions, which we call the calculation graph. This graph captures the structure of the calculation performed by a rank order filter. To find this structure we first examine the basic operation of an oblivious sorting algorithm (namely the merge operation) in Section 3.1. In Section 3.2 we show that the problem of calculating the filter result provides the opportunity to reuse intermediate results. This leads to the introduction of the calculation graph in Section 3.3 which specifies which merge operations have to take place, provides a (partial) order for them and specifies where the operation results are (re)used. A cost function is needed to calculate the cost (in extrema operations) of an algorithm represented by a calculation graph. One such cost function, based on pruned Batcher’s merging networks, is introduced in Section 3.4. In Section 3.5 we show how a simple optimization leads to better merging networks and thus a different cost function.

3.1 Divide and Conquer

A familiar and efficient way of solving problems is the divide and conquer strategy. When this strategy is applied to a sorting problem the result is the merge sort algorithm. The basic operation of a merge sort algorithm is the merging of two sorted lists into a (larger) sorted list denoted by the function:

\[ T = \text{merge}(T_1, T_2) \] (3.1)

where \( \{T_1, T_2\} \) is a partition of \( T \), i.e. \( T_1 \cup T_2 = T \) and \( T_1 \cap T_2 = \emptyset \).

As a first step in designing a rank order filter we look at the possibility of dividing the problem of calculating a single window into smaller subproblems in a similar manner. As proper generalization of (3.1) we introduce

\[ T[l_0..u_0] = \text{merge}(T[l_1..u_1], T[l_2..u_2])[l_3..u_3] \] (3.2)

However, this equality only holds if the parameters of the merge function adhere to the following properties:

1. (a) \( 0 \leq l_i \leq ((l_0 - |T| + |T_i|) \uparrow 0) \) for \( i = 1, 2 \)
   
   (b) \( (u_0 \downarrow (|T_i| - 1)) \leq u_i < |T_i| \) for \( i = 1, 2 \)
2. (a) \( l_3 = l_0 - (l_1 + l_2) \)
    \( u_3 = u_0 - (l_1 + l_2) \)

A merge operation that adheres to these properties is called a valid merge operation. An example of such an operation is depicted in Figure 3.1.

![Figure 3.1: A merge operation according to (3.2)](image)

To explain property 1 we introduce the function “rnk” to denote the rank of \( \text{data}(g) \) in \( T \). More formally:

\[
\text{rnk}(g, T) = (\downarrow i : 0 \leq i < |T| \land T(i) = \text{data}(g) : i) \quad (3.3)
\]

Observe that for any element \( g \in T_i \) (for \( i = 1, 2 \)) it holds that \( \text{rnk}(g, T_i) \leq \text{rnk}(g, T) \). In other words, adding elements to a set can only increase the rank of an element already in the set, but it cannot decrease it.

Because \( T \) contains \( |T| - |T_i| \) elements more than \( T_i \), we therefore know that \( \text{rnk}(g, T_i) \leq \text{rnk}(g, T) \leq \text{rnk}(g, T_i) + (|T| - |T_i|) \). So the interval \([l_i .. u_i]\) must be such that it contains (at least) all elements \( g \in T_i \) with \( \text{rnk}(g, T_i) \leq u_0 \) and \( \text{rnk}(g, T_i) + (|T| - |T_i|) \geq l_0 \) (i.e. at least all the elements \( g \) of \( T_i \) for which \( l_0 \leq \text{rnk}(g, T) \leq u_0 \) might hold). This is exactly what property 1 expresses.

Property 2 can be derived by noting that the merge operation results in a list that is a sublist of \( T \). The missing elements are those that, as a consequence of property 1 have a rank in \( T \) that is not within the interval \([l_0 .. u_0]\). Of the removed elements there are \( l_1 + l_2 \) elements with a rank lower than \( l \). The removal of these elements means that an index \( r \) in the merged list corresponds to a rank \( r + l_1 + l_2 \) in the set \( T \) for \( l_0 \leq r + l_1 + l_2 \leq u_0 \).

From these properties we can draw some interesting conclusions. First of all, let us return to our original problem of calculating for a single window \( W \in W \) a range of ranks \([l .. u]\). When we write such a calculation as a series of nested merge operations on tiles \( T_j \) for some \( j \), we can use property 1 to calculate the ranges for the lower and upper bounds for each of the tiles \( T_j \). It turns out that these bounds depend only on \( |T_j| \) and \( |S_W| \). So introduce functions that calculate the highest valid lower bound, and the lowest valid upper bound:

\[
\text{low}(T) = (l - |S_W| + |T|) \uparrow 0
\]
\[
\text{up}(T) = u \downarrow (|T| - 1) \quad (3.4)
\]
Using these functions guarantees validity of the merge operation, w.r.t. property 1 and at the same time minimizes the size of the interval $[\text{low}(T) .. \text{up}(T)]$. A small interval means little information to calculate and keep track of, which benefits performance. So, once $T_1$ and $T_2$ have been chosen, the best choice for $l_1, u_1, l_2$ and $u_2$ is given by (3.4). Therefore we will be using these bounds throughout this report, and we introduce a special notation for it:

$$[T] = [T_1 \text{low}(T) .. \text{up}(T)]$$  \hfill (3.5)

Also, since property 2 uniquely determines $l_3$ and $u_3$, we can omit them in our notation of any valid merge operation. Therefore we introduce the operator $\bowtie$ so that we can write (3.2) as:

$$[T] = [T_1] \bowtie [T_2]$$ \hfill (3.6)

Note that the operator we introduce is associative and commutative, as can be expected from a generalization of the standard merge operation and therefore it is functionally correct to write:

$$([T_1] \bowtie [T_2]) \bowtie [T_3] = ([T_2] \bowtie [T_3]) \bowtie [T_1]$$  \hfill (3.7)

However, the costs of implementing the left and right hand side of this expression may differ. Consider the following example: $l = 0, u = |S_W| - 1, |S_W| > 4, |T_1| = 1, |T_2| = 1$ and $|T_3| = 2$. In this example the size of the tiles is the same as the size of the interval of ranks required from them. So, the left hand side of equation (3.7) tells us to merge two tiles of size 1 and the resulting tile (of size 2) with a tile of size 2. The right hand side, on the other hand, tells us to merge a tile of size 1 and size 2 and subsequently merge the resulting tile (of size 3) with a tile of size 1. These are all quite different merging operations and it is not surprising that they differ in cost.

### 3.2 Divide, Re-use and Conquer

Divide and conquer suffices when we have to "rank-filter" a single window, but our problem statement is concerned with multiple windows that overlap. This overlapping means that there could be common subexpressions that can be reused. This is illustrated in Figure 3.2. The figure shows that we can calculate the two windows $W_i$ and $W_{i+1}$ as $[W_i] = [i - 3] \bowtie [T_1]$ and...
\[ [W_{i+1}] = [T_i] \bowtie [i + 4] \]. This allows us to use \([T_i]\) twice, and thus the cost for calculating the intermediate result \([T_i]\) is amortized over two window calculations.

Dynamic programming is a standard technique for dealing with such situations, but it can only be applied if a problem exhibits two properties: optimal substructures and overlapping subproblems ([4] Chapter 15). We have already shown that the problem has overlapping subproblems, so we only have to show that it exhibits optimal substructures. Unfortunately this is not possible, since the subproblems are not independent; they share resources. Each subproblem chooses the tiles used in the construction of its windows, these tiles represent the resources and the cost for these resources is shared between the subproblems if reuse takes place. A good, every-day, analogy is car-pooling; perhaps the cheapest way to travel for each individual is by public transport, but travel by car becomes cheaper if you travel in a group and the costs can be amortized.

So, applying dynamic programming to our design problem is not possible. However, the variant of dynamic programming known as “memoization” can still be used in the implementation of the filter ([4] Chapter 15). Using memoization amounts to an implementation that stores all tiles in a table and uses a table-lookup whenever possible. This minimizes the number of extrema operations, at the cost of table lookup operations, and makes sure all previously calculated results are reused whenever possible. In Chapter [7] we show that the regularity of our designs can be used to implemented this table and the lookup operations efficiently with a small number of buffers.

3.3 Calculation Graph

In this section we introduce the notion of a calculation graph \( G = (V, E) \) to capture the computation performed by a ROF. The nodes of this graph, given by set \( V \subseteq P(\mathbb{Z}^D) \), represent all the (intermediate) results produced during the calculation, i.e. the tiles that will be put in the look-up table. The set of nodes must contain at least:

- all the windows \( W \subseteq V \), and
- all the necessary elements \( (\forall W, i : W \in W \land i \in W : \{i\} \in V) \)

The set of edges \( E \subseteq V \times V \) denotes how the tiles are used to construct other tiles. An edge from \( T_1 \) to \( T \) means that tile \( T_1 \) is used to construct tile \( T \) with a merge operation. The following constraints must hold for this set:

1. \((\forall T : T \in V \land |T| > 1 : |T| = (\bowtie T_1 : (T_1, T) \in E : |T_1|)) \)

2. \((\forall T : T \in V \land |T| > 1 : |\{(T_1, T) \mid (T_1, T) \in E\}| = 2) \)

Property [4] expresses that each tile that does not consist of a single data element is formed by merging the tiles specified by the set \( E \). Note that this implies that the constraints for a valid merge operation, as derived in Section [5], must also hold: i.e. together all the tiles used in the construction of a tile \( T \) must form a partition of that tile \( T \). This immediately implies that the graph must be acyclic, since merging tiles can only result in larger tiles. Furthermore, it implies that for each window there must exist paths that start in the tiles representing the window’s components and end in the tile representing the window.
The calculation graph effectively describes the calculation of the windows during the implementation. Each node in the graph represents the result of a calculation and the node’s incoming edges represent the recursive calculations needed to for the sub-tiles. However, we need another property besides 1 to define the calculation done by the implementation unambiguously.

Property 2 enforces that each merge operation merges only two tiles. This removes any ambiguity that arises when more than two tiles are merged. As mentioned in Section 3.1 the costs may differ depending on the order in which tiles are merged. By limiting the calculation graph to an in-degree of 2, we enforce the specification of the evaluation order of the merge operations, and thus remove the ambiguity.

In summary, any solution for a ROFDP can be described as a calculation graph \( G \). The nodes of this graph \( V \) are called tiles and represent the (intermediate) results of the calculation. The edges \( E \) of the graph represent the way in which tiles are combined, using merging operations, to form larger tiles.

The cost of the solution can be expressed as \( \mathcal{C}(G) \):

\[
\mathcal{C}(G) = \lim_{k \to \infty} \left( \sum_{T, T_1, T_2 : T \in (V \cap \text{hypercube}(k, D))} MC(T_1, T_2, T) \right)
\]

(3.8)

where \( MC(T_1, T_2, T) \) is the cost function that determines the cost, in extrema operations, for a merge operation \([T] = [T_1] \bowtie [T_2]\). Note that we have to take the limit here because \( W, V \) and \( E \) are infinite sets. So \( \mathcal{C}(G) \) effectively gives us the cost in extrema operations per window. From this point on we abbreviate this unit of cost as “epw”.

The Rank Order Filter Design Problem can now be restated as follows: given a specification find the calculation graph \( G \) that adheres to the properties above and minimizes \( \mathcal{C}(G) \) for a certain cost function \( MC \). Before we solve this problem we first examine the cost function.

### 3.4 Cost of Merging Networks

In principle the calculation graph is indifferent to the implementation of the merge function. In order to associate a cost with a calculation graph we have to choose a particular implementation. In this report we choose to implement each merging operation using an instance of Batcher’s merging network, because Batcher’s merging network is the most efficient, known, network for merging two sorted lists. Also, there is a simple formula to calculate the number of extrema operations used in such a network. Note that this approach differs from most of the literature, where a combination of the optimum sorting networks and Batcher’s merging networks is used. However, in Chapter 6 we show that we improve upon the results existing in the literature, despite the fact that we do not use the optimum sorting networks.

The formula for the number of compare and swap operations in the network can be found in [7]. However, recall that we measure the cost of our calculation in extrema operations. The cost function in [7] measures cost in C&S operations, so we had to adapt our cost function to produce its result in extrema
simplify the relation’s notation. The extra range of values introduced for which takes sorted lists, say extrema operations required by a pruned version of Batcher’s merging network the first cost function for such networks.

Pruning is a relatively simple optimization that can be done automatically. In the case of a software implementation it is called ”dead-code elimination” and in circuit design it is called ”unused logic elimination”, which is performed automatically by the compiler or synthesizer respectively. So it is by no means a new optimization, but in this section we introduce (to the best of our knowledge) the first cost function for such networks.

The function \( \text{PB}(s_1, s_2, l_0, u_0) \) for \( 0 \leq l_0 \leq u_0 < s_1 + s_2 \) returns the number of extrema operations required by a pruned version of Batcher’s merging network which takes sorted lists, say \( \overline{T}_1 \) and \( \overline{T}_2 \), of size \( s_1 = |\overline{T}_1| \) and \( s_2 = |\overline{T}_2| \) respectively and produces the interval \([l_0..u_0]\) of the sorted list \( \overline{T}_1 \cup \overline{T}_2 \). This function \( \text{PB} \) satisfies the following recurrence relation:

\[
\text{PB}(s_1, s_2, l_0, u_0) =
\begin{cases}
\begin{align*}
0 & \quad (s_1 = 0) \lor (s_2 = 0) \\
(u_0 \downarrow 1) - (l_0 \uparrow 0) + 1 & \quad (s_1 = -1) \lor (l_0 = s_1 + s_2) \\
\text{PB}(\lceil \frac{s_1}{2} \rceil, \lceil \frac{s_2}{2} \rceil, \lceil \frac{l_0}{2} \rceil, \lceil \frac{u_0}{2} \rceil) + \\
\text{PB}(\lceil \frac{s_1}{2} \rceil, \lceil \frac{s_2}{2} \rceil, \lceil \frac{l_0}{2} \rceil - 1, \lceil \frac{u_0}{2} \rceil - 1) + & \quad \text{otherwise} \\
(u_0 \downarrow 2\lceil \frac{s_1 + s_2 - 1}{2} \rceil) - (l_0 \uparrow 1) + 1 & \quad (u_0 \neq -1) \land (l_0 \neq s_1 + s_2)
\end{align*}
\end{cases}
\tag{3.9}
\]

Note that this recurrence relation has a larger domain than \( 0 \leq l_0 \leq u_0 < s_1 + s_2 \). We expanded the domain to \( l_0 \leq u_0 \land -1 \leq u_0 \land l_0 \leq s_1 + s_2 \) to simplify the relation’s notation. The extra range of values introduced for \( l_0 \) and \( u_0 \) does not really affect the outcome of the function, since we have that \( \text{PB}(s_1, s_2, l_0, u_0) = \text{PB}(s_1, s_2, l_0 \downarrow 0, u_0 \downarrow (s_1 + s_2 - 1)) \).

The recurrence relation is a modification of the cost function for Batcher’s merging networks without pruning given in \([2]\). The first case of \((3.9)\) are simple; first of all if \( s_1 = 0, s_2 = 0, u_0 = -1 \) or \( l_0 = s_1 + s_2 \), then there is no merging necessary, since either one of the lists is empty, or we need an empty interval from the resulting list.

The second case occurs when two lists of length 1 are merged. In this case the cost can also be easily calculated. If only one member of the resulting list is needed \( l_0 = u_0 \) then we need 1 extrema operation to determine it. If both members of the resulting list are needed \( l_0 < u_0 \) then we need 2 extrema operations.

The third case is the most complicated, but the example in Figure \([3.3]\) might help in understanding. In this figure a merging network is depicted by a box. The numbers in the upper left \((s_1)\) and upper right corner \((s_2)\) denote the size of the lists to be merged. The box contains a merging network that accomplishes that operation. This merging network consists of three parts, the first two are recursively merging the even and odd sequences, which have length \( \lceil \frac{s_1}{2} \rceil + \lceil \frac{s_2}{2} \rceil \) and \( \lceil \frac{s_1}{2} \rceil + \lceil \frac{s_2}{2} \rceil \) respectively. In the figure we see this recursion as the nesting of two boxes in the main box, in equation \((3.9)\) we see this as the two recursive calls.
The third part of the merging network consist of a network that merges the results of the other two (odd and even) parts. In the figure this part has been marked by a dashed box. Its structure is relatively simple; it pairwise compares and swaps elements from the even and the odd list, starting from element 1 from the even list and element 0 from the odd list. This is illustrated in the picture by the small grey numbers, which indicate the indices of the elements in the sorted lists.

This pair-merging network allows us to derive the lower and upper bounds of the ranks for the recursive calls; from the even sequence we need $\lceil \frac{l_0}{2} \rceil$ and $\lceil \frac{u_0}{2} \rceil$ as lower and upper bound respectively. This, and the structure of the pair-merging network implies that we need $\lceil \frac{l_0}{2} \rceil - 1$ through $\lceil \frac{u_0}{2} \rceil - 1$ from the odd sequence.

The cost of the pair-merging network is expressed by the last term of the third case of (3.9). This term looks daunting, but we can understand it by noting that every element passing through the pair-merging network takes part in one extrema operation, except for the element with index 0 and possibly the element with the highest index. So, we simply count the number of elements that pass through the pair-merging network to get the number of extrema operations.

The first element that passes through the pair-merging network and participates in an extrema operation is $l_0 \uparrow 1$. The last element that participates is
When \( u_0 < s_1 + s + 2 \). Otherwise the last element is \( s_1 + s_2 \) if \( s_1 + s_2 \) is odd, or \( s_1 + s_2 - 1 \) if \( s_1 + s_2 \) is even. Hence when pruning takes place the last element that participates in an extrema operation is the element with index \( u_0 \downarrow 2 \lfloor \frac{s_1 + s_2 - 1}{2} \rfloor \).

### 3.5 Mirroring

Since Batcher’s merging networks are the best merging networks known, we might expect that their pruned versions are also the best. This is, however, not the case.

This is caused by the fact that when the range of desired ranks is not symmetric around the center of the window the amount of pruning depends on the order in which the extrema operators are arranged in the network. As an example let \( T_1 \) be a list of size \(|T_1| = 1\) and let \( T_2 \) be a list of size \(|T_2| = 2\) and let \( T = T_1 \cup T_2 \) be a list of size 3. When \( T[0..2] \), i.e. all ranks, is required Batcher’s merging network requires 4 extrema operations (see Figure 3.4(a)). When \( T[1] \), i.e. the median, is required a pruned version of the network containing 2 extrema operations suffices (see Figure 3.4(b)). When \( T[2] \), i.e. the maximum, is required a pruned version of the network also containing 2 extrema operations suffices (see Figure 3.4(c)). However, when \( T[0] \), i.e. the minimum, is required a pruned version of the network containing only 1 extrema operation suffices (see Figure 3.4(d)).

![Figure 3.4: Batcher’s merging network for a 1-2 merge](image)
This example suggests that in case of pruning we can come up with another automatic optimization which involves varying the layout of the network by varying the sorting order. We either sort in the standard ascending order, or we switch to a descending order if that is less costly. Changing the sorting order amounts to taking the mirror image of a sorting network. Hence we obtain the following cost function:

\[
\text{MPB}(s_1, s_2, l_0, u_0) = \begin{cases} 
0 & (s_1 = 0) \lor (s_2 = 0) \\
(u_0 \downarrow 1) - (l_0 \uparrow 0) + 1 & (u_0 = -1) \lor (l_0 = s_1 + s_2) \\
\text{MPB}(\lceil \frac{s_1}{2} \rceil, \lceil \frac{s_2}{2} \rceil, \lceil \frac{l_0}{2} \rceil) + \text{MPB}(\lfloor \frac{s_1}{2} \rfloor, \lfloor \frac{s_2}{2} \rfloor, \lfloor \frac{l_0}{2} \rfloor - 1, \lfloor \frac{u_0}{2} \rfloor - 1) & (u_0 \neq -1) \land (l_0 \neq s_1 + s_2) \\
(u_0 \downarrow 2 \lfloor \frac{s_1 + s_2 - 1}{2} \rfloor) - (l_0 \uparrow 1) + 1 & \text{otherwise}
\end{cases}
\]

(3.10)

This function is almost identical to (3.9), but at each point in the recursion we now consider whether the pruned version of the original, or the pruned version of the mirror network, minimizes the cost.

This new optimization method can also be applied to networks other than pruned versions of Batcher’s merging networks. Optimal sorting networks that are pruned can also be mirrored prior to the pruning operation to see if this leads to a decrease in cost. In at least one case (namely in the solution presented in [8]) this is the case.

We are now in a position to define the cost function \( \text{MC} \) from (3.8). Using function \( \text{MPB} \) (3.10) and functions \( \text{low} \) and \( \text{up} \) as defined in (3.4) we define \( \text{MC} \) as:

\[
\text{MC}(T_1, T_2, T) = \text{MPB}(\text{up}(T_1) - \text{low}(T_1) + 1, \text{up}(T_2) - \text{low}(T_2) + 1, \text{low}(T) - \text{low}(T_1) - \text{low}(T_2), \text{up}(T) - \text{low}(T_1) - \text{low}(T_2))
\]

(3.11)
Chapter 4

Generator Graphs

A calculation graph is an infinite graph that represents the infinite calculation performed by a filter. The filter itself, however, being an algorithm implemented either in hardware or software, is a finite object. The gap between infinite calculation graphs and finite filter implementations is bridged by so-called generator graphs. A generator graph is a finite representation of a filter that exploits the regularities present in the calculations performed by such a filter. On the one hand the generator graph, as its name suggests, can be used to generate a calculation graph. On the other hand, a generator graph can serve as the starting point for a concrete filter implementation.

In this chapter generator graphs are defined and it is shown how calculation graphs can be generated from them. Finding a generator graph for a specific filter is the subject of Chapter 5. However, various heuristics presented in that chapter share a common approach, which we explain in this chapter as well. Deriving parallel implementations from generator graphs is the subject of Chapter 7.

4.1 Preliminaries

Although a calculation graph contains infinitely many tiles there are only a finite number of ways in which a tile of a certain shape is constructed. This suggests that one can condense the calculation graph such that only the different construction methods are retained. Consider all tiles of a specific shape in the calculation graph. We group these tiles according to their construction method and represent this group by a single node in the generator graph.

Recall that we have already introduced the function \( \text{tile} \) as a way of generating regular sets of tiles. The argument of this function is a triple consisting of a shape, a phase and a period:

\[
T = (S_T, f_T, p_T)
\]

We will use these triples as the nodes of our generator graph \( G_g = (V_g, E_g) \). Hence it follows that a node represents a tile-set \( T = \text{tile}(S_T, f_T, p_T) = \text{tile}(T) \). Note that the phase and periods can be used to identify the various construction methods for tiles of the same shape.
So, the node set of a generator graph $V_g$ is a subset of $(P(Z^D) \times Z^D \times Z^D)$, with each node representing a tile-set. Since distinct triples can describe the same tile set we introduce the function $\text{norm}(T)$ that rewrites these triples to a unique canonical form that satisfies:

\[
\text{tile}(\text{norm}(T)) = \text{tile}(T) \quad (4.1)
\]

\[
0_D \leq f_{\text{norm}}(T) \leq p_{\text{norm}}(T) \quad (4.2)
\]

\[
(\forall g : g \in S_{\text{norm}}(T) : g \geq 0_D) \quad (4.3)
\]

\[
(\forall t : t \in Z^D \land t \neq 0_D : (\exists g : g \in (S_{\text{norm}}(T) - t) : g < 0_D)) \quad (4.4)
\]

The canonical form describes the same tile set (4.1), but its phase lies between zero and the period (4.2). Also, the set of grid locations forming its shape are positioned in the positive hyper-quadrant (4.3) in such a way that there is no tile with the same shape in the positive hyper-quadrant that lies closer to the origin (4.4). All the nodes in the generator graph adhere to this canonical form.

The nodes are used to group tiles that have the same construction method and the edges of the generator graph are used to capture these construction methods. To this end the edges of the generator graph mimic those of the calculation graph. Since a tile containing more than one grid location is constructed from two other tiles the generator graph also has this property:

\[
(\forall T : T \in V_g \land |S_T| > 1 : |(U, T)|((U, T) \in E_g) = 2)
\]

However, the tiles $T_1$ and $T_2$, used in the construction of a tile $T$, may be contained in the same tile-set and thus the same node of the generator graph. This means that the generator graph is a multi-graph and thus the set of edges is a multi-set. From our definition of the nodes of the generator graph it then follows:

\[
(\forall T, T_1, T_2 : (T_1, T) \in E_g \land (T_2, T) \in E_g \land T_1 \neq T_2 :
\]

\[
\text{tile}(T) \subseteq ([T_1] \bowtie [T_2] \mid T_1 \in \text{tile}(T_1) \land T_2 \in \text{tile}(T_2)) \quad (4.5)
\]

\[
(\forall T, T_1 : (T_1, T) \in E_g \land (\forall T_2 : (T_2, T) \in E_g : T_1 = T_2) :
\]

\[
\text{tile}(T) \subseteq ([T_1] \bowtie [T_2] \mid T_1, T_2 \in \text{tile}(T_1) \land T_1 \neq T_2) \quad (4.6)
\]

These restrictions express that for every pair of edges describing the construction of a tile set $T$, there are two other tile-sets $T_1$ and $T_2$ in which the necessary tiles can be found. Equation (4.5) expresses this for cases where the composing tiles come from distinct tile-sets and (4.6) expresses it for cases where the composing tiles come from the same tile-set.

Consider the example generator graph in Figure 4.1. For ease of reading we have chosen to denote shapes in a graphical manner. Each shape is a set of grid locations, and in the one dimensional case we can represent this set as a linear sequence of squares. When a square is black (■) this means that that specific grid location is in the set, when the square is blank (□) that specific grid location is not in the set. We start the numbering of grid locations at 0 because of the normalization requirements. So sequence "■■■■" denotes the set \{0,1,2,3\}, and "□□□□" denotes the set \{0,3\}.

Figure 4.1 shows us an example of why the generator graph is a directed multi-graph, whereas the calculation graph is not a multi-graph. A tile $T$ from the tile-set $\text{tile}(■■■■, 1, 2)$ is formed by taking two tiles $T_1$ and $T_2$ from the
same tile-set tile((■, 1, 2)). In the generator graph nodes represent tile-sets and we might need two tiles from the same tile-set to generate a tile, hence we need the generator graph to be a multi-graph. In the calculation graph each tile is represented by a single node and hence we only need at most one edge between nodes.

4.2 Canonical generator graphs

Using the preliminaries from the previous section we can now define the relationship between the nodes of the calculation graph \( V \) and the generator graph \( V_g \) more formally:

\[
V = (\cup T : T \in V_g : \text{tile}(S_T, f_T, p_T))
\]

The set of edges of the calculation graph describe the construction method for the tiles/nodes of the calculation graph. These edges can also be derived from the generator graph. Using restriction (4.5) and (4.6) we specify:

\[
(\forall T, T_1 : T \in \text{tile}(T) \land (T_1, T) \in E_g : (\exists T_1 : T_1 \in \text{tile}(T) : (T_1, T) \in E))
\]

Note that since \( E_g \) is a multi-set, a duplicate edge in \( E_g \) means that there are two (distinct) edges in \( E \).

This, however, does not describe exactly which tiles \( T_1 \in T_1 \) and \( T_2 \in T_2 \) should be used in the construction of a tile \( T \in T \), only that a construction
CHAPTER 4. GENERA TOR GRAPHS

method exists for some $T_1$ and $T_2$. However, there are at most two combinations of tiles $T_1$ and $T_2$ that partition the tile $T$. This can be explained as follows: take a grid location $g \in T$. We now have either $g \in T_1$ or $g \in T_2$, i.e. the tile containing $g$ comes either from tile-set $T_1$ or $T_2$, hence there are two options. One might argue that there could be multiple tiles $T_1 \in T_1$ such that $g \in T_1$, however there is only one $T_1$ capable of forming $T$ with one other tile $T_2$ from the other tile-set because the shape of the tiles is fixed for $T_1$, $T_2$ and $T$. The same reasoning can, of course, be used to show that there is only one suitable $T_2 \in T_2$.

When multiple options exist a choice must be made. This choice does not influence the cost of the resulting calculation graph. The cost of merging, after all, depends only on the size of the tiles, which is not influenced by the choice. The choice, however, must be made in a consistent manner. This ensures that always the same calculation graph is generated from a generator graph. This, in turn, ensures that the implementation based on the generator graph will correspond to the calculation graph.

To make such a consistent choice we introduce an ordering on the grid locations and the tile-sets. The grid locations are ordered by sorting them in ascending order on their first coordinate, then on their second coordinate, etc. The tile-sets are ordered first on the size of the shape then on their phase and subsequently on their period. Whenever a choice has to be made we make it such that the tile containing the grid location with the lowest rank of the resulting tile ($g \in T$) comes from the tile-set with the lowest rank. In most cases it is possible to make a consistent choice, but the example of Section 4.1 shows a case where this is not possible: tile($\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}$, 0, 1). Consider the lowest coordinates of the tiles from tile($\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}$, 0, 1) if this coordinate is even the tile can only be formed by taking a tile containing this coordinate from tile($\begin{array}{c}
  \begin{array}{c}
  | \
  |
\end{array}
\end{array}$, 0, 1), but when this coordinate is odd the tile can only be formed by taking a tile containing this coordinate from tile($\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}$, 1, 2).

To solve this problem we introduce a canonical form such that all choices during the generation of a calculation graph from generator graphs in canonical form can be made in a consistent manner. The canonical form entails nothing more than that all the nodes in the generator graph should have the same period and that this period should be as small as possible. We call this period the meta-period of the graph and it can be calculated by taking the least common multiple of the periods of all the nodes of a generator graph (that has minimum periods for all its tile-sets). This meta-period also plays an important role in actually implementing the design in Chapter 7.

Note that it is always possible to get the canonical form of a generator graph since nodes that do not have the required period can simply be duplicated a number of times. We have done this for the generator graph from Figure 4.1 and the result is shown in Figure 4.2.

A major benefit of this canonical form is that the the choices in generating the calculation graph can now be made consistently. For example the tiles from node (\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}, 0, 2) in Figure 4.2 are always constructed by a tile from (\begin{array}{c}
  \begin{array}{c}
  | \n  |
\end{array}
\end{array}, 0, 2) followed by a tile from (\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}, 1, 2). Conversely the tiles from node (\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}, 1, 2) are always constructed by a tile from (\begin{array}{ccc}
  & & \\
  & & \\
  & & \\
\end{array}, 1, 2) followed by a tile from (\begin{array}{c}
  \begin{array}{c}
  | \n  |
\end{array}
\end{array}, 1, 2). In contrast the tiles from the original graph in Figure 4.1 had to alternate between these construction methods and thus did not allow a consistent choice.
4.3 Cost calculation

We can calculate the cost of a calculation graph generated by a generator graph without actually generating the calculation graph itself. Take Figure 4.2 for example; the fact that node (■, 1, 2) has incoming edges from (■, 0, 2) and (■, 1, 2) means that it represents the merging of two tiles of size 1 and that this merging occurs with a period of 2. To calculate the cost of merging two tiles of size 1 we need to know the range of ranks needed from the resulting tile. This range of ranks can be calculated using (3.4) and, for the sake of simplicity, we look at the window sorting problem here, where all ranks are needed. Then the merging of two tiles of size 1 requires 2 extrema operations. Because the period of the windows is 1 and the period of these mergings is 2 we know that this cost of the merging operation can be amortized over 2 windows. Hence the contribution of this merging operation to the total is 1 epw. In a similar manner we can calculate the contribution of all the nodes in the generator graph to the total cost. For Figure 4.2 this has been done in Table 4.1.

Note that, besides not having to actually generate the calculation graph to calculate the cost, the generator graph has another advantage; we are able to avoid taking limits as required by definition (3.8).
### 4.4 Constructing generator graphs

Chapter 5 describes, in detail, several heuristic algorithms for constructing generator graphs. In this section we discuss the framework used by some of the algorithms from that chapter and indicate how other heuristic algorithms can be designed.

The algorithms for constructing a generator graph start with an empty generator graph \((\emptyset, \emptyset)\) and add nodes and edges until a generator graph is formed such that the calculation graph derived from it is valid.

To add a node and edges to the generator graph the \texttt{add} function is used, which returns the updated graph:

\[
\text{add}(G_g, T, T_1, T_2) = (V_g \cup \{T\}, E_g \cup \{(T_1, T), (T_2, T)\})
\]  

(4.7)

This adds the node \(T\) to the set of nodes and adds incoming edges (from \(T_1\) and \(T_2\)) to the graph. Note that for the resulting graph to be valid the construction of the tiles in \(T\) from the tiles in \(T_1\) and \(T_2\) must be possible, that \(T_1\) and \(T_2\) must be in the set of nodes, and that all the triples have to be in canonical form. Moreover, recall that \(E_g\) is a multi-set, so in case \(T_1 = T_2\) two edges are added.

Some algorithms create a generator graph in a top down fashion. These algorithms start by defining the construction of the windows (specified by the ROFDP) and then recursively define the construction method of the window’s components.

The top down approach is based on two functions. The first one is a function that specifies how tiles from an existing tile-set should be split into two sub-tiles from (possibly different) tile-sets and which (possibly other) function specifies the construction method of these new tile-sets. This function forms the core of the heuristic algorithm and we therefore call it the heuristic function. The second function used by this approach is the function that recursively calls the heuristic function and adds the results to the generator graph.

The method of construction of a tile set is defined by the heuristic function, say \(f\). The function differs for each algorithm, but all functions return the same kind of tuple: \(f(T) = ((T_1, f_1), (T_2, f_2))\), where \(T_1 = \text{norm}(T_1)\) and \(T_2 = \text{norm}(T_2)\). The function effectively specifies that the tiles in tile set \(\text{tile}(T)\) should be constructed from tiles in \(\text{tile}(T_1)\) and \(\text{tile}(T_2)\). It also specifies that the tile-sets \(\text{tile}(T_1)\) and \(\text{tile}(T_2)\) should be constructed in the manner

<table>
<thead>
<tr>
<th>Node</th>
<th>Cost per operation</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>(■■■■■, 0, 2)</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>(■■■■■, 1, 2)</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>(■■■■■, 0, 2)</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>(■■■■■, 1, 2)</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>(■■, 1, 2)</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>(■, 0, 2)</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(■, 1, 2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total cost:</strong> (in epw)</td>
<td><strong>22</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Cost of the generator graph in Figure 4.2
defined by the heuristic functions \( f_1 \) and \( f_2 \) respectively. For basic algorithms we have \( f = f_1 = f_2 \), but more sophisticated algorithms switch heuristics at a certain point.

Although this first function differs per algorithm, there are enough similarities between algorithms to introduce the so-called split function. This function is used to split a tile at a specific point into two sub-tiles:

\[
\text{split}(T, c, p_1, f_1, p_2, f_2) = ((\text{cut}(T, c, p_1), f_1), (\text{rem}(T, c, p_2), f_2))
\]

The result returned by the split function adheres to the format for the heuristic function and is often used as a basis for such a function.

To split a tile into two sub-tiles the split function uses the cut and rem functions which respectively return the part of a tile closest to the origin and the remainder of the tile. More formally:

\[
\begin{align*}
\text{cut}(T, c, p_{\text{new}}) &= \text{norm}(\left(S_T \cap \text{hyperbar}(c), f_T, p_{\text{new}}\right)) \quad (4.9) \\
\text{rem}(T, c, p_{\text{new}}) &= \text{norm}(\left(S_T \setminus \text{hyperbar}(c), f_T, p_{\text{new}}\right)) \quad (4.10)
\end{align*}
\]

The functions \( \text{cut}(T, c, p_{\text{new}}) \) and \( \text{rem}(T, c, p_{\text{new}}) \) therefore return normalized triples based on \( T \). The difference between the two functions is in the shape of the returned triple. The shape of the original triple is effectively cut in two pieces based on the cut-off point described by vector \( c \). The function cut returns the part closest to the origin, while the function rem returns the remainder. The triples returned by these functions have a period indicated by the vector \( p_{\text{new}} \).

The hyperbar function is used by the cut and rem functions to partition a tile. As it name implies the function does nothing more than returning a tile in the shape of a hyper-bar with dimensions indicated by vector \( s \):

\[
\text{hyperbar}(s) = \{ g | 0 \leq g < s \}
\]

Let us examine \( \text{split}((\square\square\square, 0, 1), 2, 1, f_1, 2, f_2) \) as an example. This formula specifies that the tiles from \( (\square\square\square, 0, 2) \) have to be formed from tiles of size 2 and 4, with periods 1 and 2 respectively:

\[
\begin{align*}
\text{split}((\square\square\square, 0, 2), 2, 1, f_1, 2, f_2) &= \text{Definition of split} \\
&= ((\text{cut}((\square\square\square, 0, 2), 2, 1), f_1), (\text{rem}((\square\square\square, 0, 2), 2, 2), f_2)) \\
&= \text{Definition of cut, rem and hyperbar} \\
&= ((\text{norm}((\square\square\square \cap \square, 0, 1)), f_1), (\text{norm}((\square\square\square \setminus \square, 0, 2), f_2)) \\
&= \text{Set arithmetic} \\
&= ((\text{norm}((\square, 0, 1)), f_1), (\text{norm}((\square\square\square, 0, 2), f_2)) \\
&= \text{Rewrite to canonical form} \\
&= ((\square, 0, 1), f_1), ((\square\square\square, 2, 2), f_2))
\end{align*}
\]

Note that the split function does nothing to ensure correctness. We could also have used the example:
split((■■■■■■, 0, 1), 2, 3, f_1, 2, f_2) = (((■■, 0, 3), f_1), ((■■■■, 2, 2), f_2)),
which would result in a generator graph that is in violation of (4.5). So split is only a notational shorthand for our algorithms. That an algorithms provides us with a correct generator graph has to be shown for each algorithm individually.

The other major part of the frame work for top-down algorithms is the function that recursively calls the heuristic function until the construction method of every tile has been specified. This function is called TD:

\[
TD(G_g, T, f) =
\begin{cases}
G_g & T \in V_g \\
(V_g \cup \{T\}, E_g) & T \notin V_g \land |S_T| = 1 \\
TD(TD(add(G_g, T, T_1, T_2), T_1, f_1), T_2, f_2) & \text{otherwise}
\end{cases}
\]

where \(((T_1, f_1), (T_2, f_2)) = f(T)\) (4.14)

This function returns an updated generator graph by taking a generator graph \(G_g\) by adding the triple \(T\) to it (we assume that \(T = \text{norm}(T)\)). This explains the first two cases in (4.14): in the first case the triple \(T\) is already part of the generator graph and thus nothing has to be done. In the second case the triple \(T\) is not yet part of the generator graph, so it has to be added, and since it is of size 1 it requires no incoming edges to describe its construction. The last case is more complicated, not only must the triple \(T\) be added as a node, also two edges need to be added that describe how the tiles in the tile set \((T)\) must be constructed. The particular manner in which this construction takes place is described by the heuristic function \(f\), which (for most algorithms) is based on the \textit{split} function described earlier.
Chapter 5

Heuristic Algorithms

In the previous chapter the problem of designing an efficient ROF has been reduced to the problem of finding a generator graph with low cost. In this chapter we introduce several heuristic algorithms capable of finding such low-cost generator graphs. Finding optimal ROF’s of small size can be done by a brute force algorithm, just as optimal sorting networks have been found for small sorting problems [7], but for larger problems the number of possible generator graphs explodes. Therefore we resort to heuristic algorithms.

In the next few sections we develop several heuristic algorithms for (mainly one dimensional) ROFDPs using the method sketched in Chapter 4. This means that we introduce various ways of splitting tile-sets. We also show how an algorithm designed for a one dimensional ROFDP can be generalized to multiple dimensions.

5.1 Preliminaries

In the following sections several heuristic algorithms to construct generator graphs are presented. The way these algorithms work is illustrated using two running examples. The first example concerns the design of a 1-dimensional window sorter for window size 6, i.e. window-set \( W = \text{tile}(\text{□□□□□□}, 0, 1) \) and \( l = 0 \) and \( u = 5 \). The second example concerns a 1-dimensional window sorter for window size 7, i.e. window-set \( W = \text{tile}(\text{□□□□□□□}, 0, 1) \) and \( l = 0 \) and \( u = 6 \). Furthermore we introduce three functions that are frequently used in the definition of the heuristics. They are:

\[
\begin{align*}
pow2(a) & = (\exists i : i \in \mathbb{N} : 2^i = a) \quad (5.1) \\
msb_1(a) & = (\uparrow i : 2^i < a : 2^i) \quad (5.2) \\
lsb_1(a) & = (\uparrow i : 2^i | a : 2^i) \quad (5.3)
\end{align*}
\]

These functions respectively return: a boolean value that indicates whether \( a \) is a power of two and the power of two represented by the most-significant 1-bit of \( a \) and the least-significant 1-bit of \( a \).

An operation which is used often by the algorithms is selecting the triple describing tiles with the largest or smallest shape from a set. To indicate these...
we use the following notation:

\[ \downarrow Q \] for an element of \( Q \) such that \( |S_{\downarrow Q}| = (\downarrow q : q \in Q : |S_q|) \)

\[ \uparrow Q \] for an element of \( Q \) such that \( |S_{\uparrow Q}| = (\uparrow q : q \in Q : |S_q|) \)  

\[ (5.4) \]

### 5.2 Divide and Conquer

This algorithm is based on the observation that, when using Batcher’s merging networks, the cheapest way to obtain a sorted list of size \( s \) is by merging sorted lists of size: \( \lceil \frac{s}{2} \rceil \) and \( \lfloor \frac{s}{2} \rfloor \). This is a simple divide and conquer approach leading to the construction:

\[ G_g = TD((\emptyset, \emptyset), \text{norm}(W), \text{dac}) \]

where the heuristic function \( \text{dac} \) is specified by:

\[ G_g = TD((\emptyset, \emptyset), \text{norm}(W), \text{dac}) \]

\[ \text{dac}(T) = \text{split}(T, \lceil \frac{|S_T|}{2} \rceil, p_T, \text{dac}, p_T, \text{dac}) \]

The generator graphs produced by this algorithm for our running examples are shown in Figure 5.1. The resulting rank order filters require 18 epw (extrema operations per window) and 28 epw for sorting windows of size of 6 and 7 respectively. Since each tile-set has a period equal to the period of the windows the \( \text{dac} \)-heuristic does not reuse any tiles.
Figure 5.2: Example of overlapping for $T = (■■■■■■■, 0, 1)$

### 5.3 Recursive Tiling

The divide and conquer algorithm from the previous section provides generator graphs in which no reuse takes place. Before we introduce an algorithm that generates generator graphs with reuse, let us examine a possible way in which reuse can take place.

A solid tile $T \in T$ can be formed by combining two tiles $T_1$ and $T_2$ with $(\forall g_1, g_2 : g_1 \in T_1 \land g_2 \in T_2 : g_1 < g_2)$. When $|T_1|$ is a multiple of $p_T$ then $T_2$ can be reused in the construction of $T + |T_1|$ like so: $[T + |T_1|] = [T_2] \bowtie [T_1 + |T|]$. Two examples of this are shown in Figure 5.2.

Note that the part that could be reused in the divide and conquer algorithm is relatively small; for any $T$ the reusable part would have size $\lceil \frac{|T|}{2} \rceil$. So, instead of adopting the divide and conquer algorithm so that it exhibits reuse, we immediately turn to an algorithm that maximizes the size of the reused parts.

This algorithm simply splits a tile in a remainder and in an overlapping part. In the first step the remainder will be of size one and the overlapping part of size $|S_T| - 1$. This overlapping part then has a period of 2, so this remainder is split into a part of size 2 and a part of size $|S_T| - 3$. More formally this algorithm is described by the heuristic function $rts$:

$$rts(c, m)(T) = \begin{cases} 
\text{split}(T, c, p_T, d \text{ac}, (2p_T) \downarrow m, rts(2c, m)) & c < |S_T| \\
\text{d\text{ac}}(T) & c = |S_T| \\
\text{r\text{tm}}(T) & c > |S_T| 
\end{cases}$$

$$rtsm(T) = \text{split}(T, \text{msb}_1(|S_T|), p_T, d \text{ac}, p_T, rtsm)$$

The $rts$ function has two parameters: $c$ and $m$. The parameter $c$ equals the size of the tile that is split of. The algorithm starts with $c = 1$ and, since $c$ can
only be doubled, this means that \( c \) is always a power of two. Conceptually the period of the overlapping part will double in each recursive step. However, to make sure that the generator graph produced by the algorithm is valid we need to limit the period at a certain value. This limit is indicated by the parameter \( m \). Since a higher period allows for more reuse this parameter will be set as high as possible, but low enough to ensure that a valid generator graph is produced. This leads to \( m = 2 \lfloor \log_2 |S_W| + 1 \rfloor \).

The \( \text{rts} \) function recursively calls itself to split the tiles, but at a certain stage \( c \) will be larger or equal to the size of the tile that is under construction. In these cases we have to switch heuristics. The \( \text{dac} \) heuristic is a possible choice, but while using the \( \text{rts} \) heuristic a lot of tiles with sizes that are a power of two have been generated. We can reuse these tiles by switching to the correct heuristic. When \( c \), which is a power of two, equals the tile size the \( \text{dac} \) heuristic will reuse them automatically since all the tiles it generates will be powers of two. In the case where \( c \) is larger than the tile size we switch to the \( \text{rtm} \) heuristic to make sure that we construct the tile from tiles whose size are a power of two.

This heuristic performs well for odd window sizes, however for even window sizes we have \( m = 2 \lfloor \log_2 |S_W| + 1 \rfloor = 2 \), which severely limits reuse. We therefore take a special step in the case of even window sizes, defined by the \( \text{rtf} \) heuristic:

\[
\text{rtf}(T) = \begin{cases} 
\text{split}(T, 1, \text{pr} \ldots \text{pr}, \text{rts}(1, 2 \lfloor \log_2 |S_T| \rfloor)) & |S_T| \mod 2 = 0 \\
\text{rts}(1, 2 \lfloor \log_2 (|S_T| + 1) \rfloor)(T) & |S_T| \mod 2 = 1
\end{cases}
\]

The generator graph produced by the recursive tiling algorithm for one dimensional problems is therefore defined by:

\[
G_g = \text{TD}((\emptyset, \emptyset), (\text{norm}(W), \text{rtf}))
\]

The result of this algorithm for our running examples can be seen in figures 4.2 and 5.3. The cost for both solutions is respectively 22 and 20.5 epw for window sizes 6 and 7. So for windows of size 7 this algorithm is better than the divide and conquer algorithm. In fact, for most problems this algorithm performs better than the divide and conquer algorithm. The reason for this is the reuse; for the problem with window size 7 the tiles of size 2, 4 and 6 can be reused, with the tile of size 4 even being used for 4 different windows. It is this amount of reuse that allows the recursive tiling algorithm to improve upon the divide and conquer algorithm. Unfortunately, only the tiles of size 2 and 4 can be reused for the problem with window size 6 and this reuse is not enough to improve upon the divide and conquer algorithm.

### 5.4 Bit-pattern Tiling

Our experiments with the heuristic from the previous section shows that power of two tiles (i.e. tiles with a size that is a power of two) can be created efficiently using the \( \text{dac} \)-heuristic. In this section we therefore present an algorithm that produces windows based on such power of two tiles. Since the type of power of two tiles depends on the pattern of bits representing the window size, we have named this heuristic “Bit-pattern tiling”.

This algorithm comes in two variants; one that splits off the least significant power of two from an existing tile, and one that splits of the most significant
power of two. In effect we either iterate through the bit-pattern from left to right or from right to lift.

The first variant of the algorithm is called "Least Significant Bit Tiling" or "LSB Tiling" and is defined as follows:

\[
G_g = TD((\emptyset, \emptyset), \text{norm}(W), \text{lsbt})
\]

where the heuristic function \( \text{lsbt} \) is specified by:

\[
\text{lsbt}(T) =
\begin{cases}
  \text{dac}(T) & \text{pow2}(|S_T|) \\
  \text{split}(T, \text{lsb}_1(|S_T|), p_T, \text{lsbt}, 2p_T, \text{lsbt}) & \neg\text{pow2}(|S_T|) \land p_T = \text{lsb}_1(|S_T|) \\
  \text{split}(T, \text{lsb}_1(|S_T|), p_T, \text{lsbt}, p_T, \text{lsbt}) & \text{otherwise}
\end{cases}
\]

The second variant is called "Most Significant Bit Tiling" or "MSB Tiling":

\[
G_g = TD((\emptyset, \emptyset), \text{norm}(W), \text{msbt})
\]

where the heuristic function \( \text{msbt} \) is specified by:

\[
\text{msbt}(T) =
\begin{cases}
  \text{dac}(T) & \text{pow2}(|S_T|) \\
  \text{split}(T, \text{msb}_1(|S_T|), p_T, \text{msbt}, 2p_T, \text{msbt}) & \neg\text{pow2}(|S_T|) \land p_T = \text{msb}_1(|S_T|) \\
  \text{split}(T, \text{msb}_1(|S_T|), p_T, \text{msbt}, p_T, \text{msbt}) & \text{otherwise}
\end{cases}
\]
Both algorithms consist of three cases. The first case applies only to tiles that have a size that is a power of two. Those tiles are constructed using the dac-heuristic. The other two cases are quite similar; a piece of the tile is split off and the size of this piece is the least- and most-significant bit of the tile’s size for \( \text{lsbt} \) and \( \text{msbt} \) respectively. The only difference between the two cases is the period of the part that is split off. If possible, i.e. if we still end up with a valid construction, this period is double the period of the original tile since it can be reused in that case.

Both algorithms produce the same generator graph for our running example with window size 6 (see Figure 5.4, which costs 20 epw. However, there is very little reuse in this solution, so although the new algorithms improve upon the recursive tiling algorithm, they do not improve upon the divide and conquer algorithm. The LSB tiling algorithm produces, for windows of size 7, the same graph as the recursive tiling algorithm (shown in Figure 5.3) and thus has the same cost (20.5 epw). The MSB tiling algorithm also produces an already familiar graph for windows of size 7, namely the same one produced by the divide an conquer algorithm (Figure 5.1(b)) which has a cost of 28 epw.

5.5 Split Tiling

The fact that power of two tiles can be formed relatively cheaply is exploited in the heuristic from the previous section, but not to its full extent. In this section we propose a heuristic that will form each window from a set of power of two tiles, just like the Bit-pattern tiling heuristic from the previous section, but there will be more (possible) orders in which these power of two tiles are combined. Most notably we make use of the fact that the merge operation is
CHAPTER 5. HEURISTIC ALGORITHMS

commutative (see Section 3.1), which allows us to construct tiles which are split by "holes". Hence we call this heuristic "Split Tiling".

This heuristic assumes that the period of the windows is one ($p_w = 1$) and that the shape of the tile is solid to begin with, i.e. $S_w = \{0, \ldots, |S_w| - 1\}$.

The algorithm works by first creating all tiles with a size that is a power of two and that have a period equal to their size. This can be done at a very low cost.

$$G = TD((\emptyset, \emptyset), ((0, \ldots, msb(|S_w|) - 1), 0, msb(|S_w|), sth))$$

$$sth(T) = split(T, |S_T|/2, p_T/2, sth, p_T/2, sth)$$

This only generates part of the final generator graph. The second step is to use the tiles in this partial generator graph to form the window. Since the meta-period of the graph created so far is $msb(|S_w|)$, there are $msb(|S_w|)$ different combinations of tiles that form a window; one combination for each phase in the period $msb(|S_w|)$. The combination of tiles that form the window at phase $i$ is expressed by $comb(i)$:

$$comb(f) = cr(|S_w|, f)$$

$$cr(s, f) = ct(gt(msb_1(s), f)) \cup cr(s - gt(msb_1(s), f), f + gt(msb_1(s), f))$$

$$ct(s, f) = \begin{cases} 
    s & f \mod s = 0 \\
    gt(s/2, f) & f \mod s \neq 0 
\end{cases}$$

$$gt(s, f) = \begin{cases} 
    s & f \mod s = 0 \\
    gt(s/2, f) & f \mod s \neq 0 
\end{cases}$$

The function $gt$ returns the size of the largest tile smaller than or equal to $s$ that starts at phase $f$. This function is then used by the recursive function $cr$ to find the set of triples describing the tiles that will form the window. The function $cr$ returns the set of tiles that form a tile of size $s$ at phase $f$. The function $ct$ is used to construct the actual triple, based on its size. Note that this triple is not normalized. This is done to make sure that the set returned by $cr$ contains one triple for each tile used in the construction of the window. For example: constructing the tiles from $tile(\square\square\square\square\square, 1, 4)$ requires tiles from $tile(\square, 1, 1)$, $tile(\square\square, 2, 2)$ and $tile(\square\square, 4, 2)$. If we would normalize the triples we would end up with only two distinct triples, and the implicit information in the phase would be destroyed too.

Now, the order in which those tiles are combined can be varied, therefore we have made three variants of this heuristic. These variants are named "Split Tiling 1", "Split Tiling 2" and "Split Tiling 3" respectively. Each of the split tiling algorithm variants generates a partial generator graph $G_i$ based on $comb(i)$ for $0 \leq i < msb(|S_w|)$. We then combine all the partial generator graphs to create the final generator graph like so:

$$G_g = G_i \cup (\cup i : 0 \leq i < msb(|S_w|) : G_i)$$

To generate the partial graphs $G_i$ all the variants of this heuristic have to merge the tiles in $comb(i)$ until the entire window is formed. This bottom up
approach can be expressed as follows:

\[
\begin{align*}
\text{ST}(G_i, S, f) &= \begin{cases} 
G_i, & |S| = 1 \\
\text{ST}(\text{ma}(G_i, T_1, T_2), \text{sup}(S, T_1, T_2), f_1) & \text{otherwise}
\end{cases} \\
\text{ma}(G_i, T_1, T_2) &= \text{add}(G_i, \text{norm}(\text{mt}(T_1, T_2)), \text{norm}(T_1), \text{norm}(T_2)) \\
\text{mt}(T_1, T_2) &= (\text{ms}(T_1, T_2), f_1, \text{msb}_{1}(|S_W|)) \\
\text{ms}(T_1, T_2) &= S_{T_1} \cup (S_{T_2} + f_{T_2} - f_{T_1}) \\
\text{sup}(S, T_1, T_2) &= (S \setminus \{T_1, T_2\}) \cup \{\text{mt}(T_1, T_2)\}
\end{align*}
\]

The ST function generates the partial generator graph \(G_i\) by repeatedly taking two tiles from the set \(S\), merging them and adding them to \(G_i\). This will eventually lead to a single tile in the set \(S\) (the window at phase \(i\)), which indicates that the partial graph is finished. If the recursion is not yet finished the ST function uses the function ma to merge two tiles and add the new tile and its construction method to \(G_i\). It does this by using mt which merges the tile, which in turn uses ms to merge the shape of the tiles. The function \(\text{sup}\) is used to update the set of tiles that still have to be merged.

The selection of the tiles from the set \(S\) is determined by the selection function \(f\), which differs per variant of the split tiling algorithm. The function \(f(S) = (T_1, T_2, f_1)\) returns a triple with \(T_1 \in S\) and \(T_2 \in S\) indicating the tiles to be merged and \(f_1\) indicating a (possibly) new selection function for the next step in the algorithm.

So, in contrast to the algorithms discussed so far the "Split Tiling" approach is bottom-up; given some tiles it constructs the windows. Whereas the other algorithms are top-down approaches; given the window they decide in which way it should be split.

The differences between the variants of the "Split Tiling" algorithm lie in the selection function and for "Split Tiling 1" we have \(G_i = \text{ST}((\emptyset, \emptyset), \text{comb}(i), \text{st1})\):

\[
\text{st1}(S) = (\downarrow S, (S \setminus \{\downarrow S\}), \text{st1})
\]

So "Split Tiling 1" simply keeps on merging the smallest tiles in the set, until the entire window has been constructed.

For "Split Tiling 2" we have \(G_i = \text{ST}((\emptyset, \emptyset), \text{comb}(i), \text{st2}(i))\):

\[
\text{st2}(i)(S) = \begin{cases} 
\text{st2a}(i)(S), & (\exists T_1, T_2 : T_1, T_2 \in (S \cap \text{comb}(i)) \land T_1 \neq T_2 : |S_{T_1}| = |S_{T_2}|) \\
\text{st1}(S), & \text{otherwise}
\end{cases}
\]

\[
\text{st2a}(i)(S) = (T_1, T_2, \text{st2}) \text{ such that:}
\]

\[
|S_{T_1}| = |S_{T_2}|
\]

\[
T_1, T_2 \in (S \cap \text{comb}(i))
\]

\[
T_1 \neq T_2
\]

\[
|S_{T_1}| = (\downarrow T_3, T_4 : T_3, T_4 \in (S \cap \text{comb}(i)) \land T_3 \neq T_4 \land |S_{T_3}| = |S_{T_4}| : |S_{T_3}|)
\]

This variant first merges all the initial tiles of the same size, and only after that it starts merging the smallest tiles until the entire window is formed. The idea
is that Batcher’s merging networks are more efficient when lists of the same size are merged.

For "Split Tiling 3" we have $G_i = ST((\emptyset, \emptyset), \text{comb}(i), \text{st3}):$

$$st3(S) = \begin{cases} st3a(S) & (\exists T_1, T_2 : T_1, T_2 \in S \land T_1 \neq T_2 : |S_{T_1}| = |S_{T_2}|) \\ st3b(S) & \text{otherwise} \end{cases}$$

$$st3a(S) = (T_1, T_2, \text{st3}) \text{ such that } |S_{T_1}| = |S_{T_2}| | T_1, T_2 \in S \land T_1 \neq T_2$$

$$st3b(S) = ((\downarrow S, \downarrow (S \setminus \{\uparrow S\})), \text{st3b})$$

This variant first merges all the tiles of the same size in $S$. If all the tiles have different sizes it simply starts merging the largest two tiles.

For windows of size 6 all three variants of the algorithm produce the same generator graph (depicted in Figure 5.5), with a cost of 17 epw. This is an improvement over all the algorithms we have seen so far. This is not surprising since this is the first design for this problem in which tiles of size 4 are reused.

For windows of size 7 the first variant produces the generator graph in Figure 5.6 with a cost of 22.5 epw and the other two variants produce the same graph as the recursive tiling algorithm (Figure 5.3) with a cost of 20.5 extrema operations.

### 5.6 Overlap Recursive Mirroring

So far our algorithms have attempted to maximize reuse by following some algorithm based on the shape of the windows. However, the cost of the generator
CHAPTER 5. HEURISTIC ALGORITHMS

Figure 5.6: $G_g$ for the Split Tiling 1 algorithm for window size 7 (22.5 epw)

The graph also depends on $l$ and $u$ and we can obtain better results by taking a look at the possible ways of reuse and selecting the best one based on the cost function. This is what the "Overlap Recursive Mirroring" algorithm does.

For every tile it has to construct the heuristic checks whether that tile is a hyper-bar and whether it overlaps with the next (periodic) version of itself. If this is the case the overlapping part is a good candidate for an intermediate tile. Of course, a tile may overlap with multiple versions of itself. The overlapping part of multiple versions is smaller than the overlapping part of two single versions, but the former can be reused multiple times. This is a difficult trade-off and the "ORM" algorithm makes this trade-off by simply trying a number of possibilities and selecting the one with the lowest cost.

To keep the calculation time under control we introduce the parameter $R$ which indicates the maximum number of times an overlap is re-used. So, for $R = 0$ the algorithm only considers constructing the tile without making use of any overlap. For $R = 1$ it also considers constructing a tile by re-using the overlap between two consecutive occurrences of a tile. For $R = 2$ it considers all the possibilities of $R = 1$ and the possibility of re-using the overlap between three consecutive occurrences of a tile. And so forth for larger values of $R$.

The function $\text{ovl}(T)$ returns the set of possible manners in which tiles from $\text{tile}(T)$ can be overlapped with periodic versions of itself, given that the tile is a hyper-bar.

$$\text{ovl}(R, T) = \{ g | g \in \text{hyperbar}(R \cdot 1_D) \land (S_T \cap (g \cdot p_T + S_T)) \neq \emptyset \}$$
So each element of the returned set is a vector. This vector no longer represents a coordinate on the grid, but a way in which the tiles overlap with other tiles from the same tile-set. So (0,0) means no overlapping, (1,0) means that a tile overlaps with the next tile in the first dimension, but that there is no overlapping in the second dimension. This function is used to generate the set of overlaps considered by the algorithm. Note that the parameter $R$ is used here to keep the number of possibilities in check.

Note that we only apply the $\text{ovl}$ function when $S_T$ is a hyper bar. We do this because this makes it possible to implement the $\text{ovl}$ function in a fast manner in practice. Whenever the tile is not a hyper-bar we only consider the overlaps represented by the set $\{0_D\}$.

Other functions that deal with the overlapping part are:

\[
\begin{align*}
\text{os}(T,o) &= (\bigcap g : g \in \text{hyperbar}(o + 1_D) : S_T + g \cdot p_T) \\
\text{op}(T,o) &= (o + 1_D) \cdot p_T \\
\text{ot}(T,o) &= \text{norm}(\text{os}(T,o), f_r, \text{op}(T,o))
\end{align*}
\]

In these functions the parameter $o$ is an element of the set produced by the $\text{ovl}$ function that indicates the manner of overlap under consideration. The function $\text{os}$ returns the shape of the overlapping part of the tiles, $\text{op}$ the period and $\text{ot}$ returns the normalized triple representing the tile-set of overlapping parts.

The generator graph created by the function $\text{orm}$ heuristic:

\[
G_g = \text{orm}(R, (\emptyset, \emptyset), \text{norm}(W))
\]

\[
\text{orm}(R, G_g, T) \in \{\text{ov}(G_g, T, o) | o \in \text{ovl}(R, T)\}
\]

such that $\text{C}(\text{orm}(R, G_g, T)) = (\downarrow o : o \in \text{ovl}(R, T) : \text{C}(\text{ovl}(G_g, T, o)))$

Note that we abuse the notation here slightly: the cost function $\text{C}$ should receive a calculation graph instead of a generator graph as its parameter. We allow this, however, since the cost can be calculated from the generator graph directly (see Section 4.3).

The function $\text{orm}$ simply selects the best generator graph generated by the $\text{ov}$ function. This $\text{ov}$ function generates a generator graph by using the given overlapping pattern $o$ for the tile $T$.

\[
\text{ov}(G_g, T, o) = \begin{cases} 
\text{nov}(G_g, T) & \text{o} = 0_D \\
\text{ovt}(G_g, T, o) & \text{otherwise}
\end{cases}
\]

The generator graph in which no overlap for tiles in $T$ is used is generated by $\text{nov}$ and the generator graph in which overlap is used is generated by $\text{ovt}$.

An example: The tiles defined by the triple $T = (■■■■, 0, 1)$ overlap each other. For $R = 0$ the algorithm constructs the tile without exploiting any overlap in the tiles. This method is explained later on in this section. For $R = 1$ the algorithm also considers constructing $(■■■■, 1, 2)$ as an intermediate result, since these tiles are the overlapping part of two consecutive tiles from $T$. For $R = 2$ the algorithm considers all the previous possibilities and also the construction of $(■■■■, 2, 3)$, since those tiles are the overlapping part of three consecutive tiles from $T$. 

\[\text{orm}\]
CHAPTER 5. HEURISTIC ALGORITHMS

So, that is where the "Overlap" part of the algorithm’s name comes from. The "Recursive" part is because this algorithm is applied recursively to the resulting intermediate tiles. And the "Mirroring" part is because the shape of tiles that have to be constructed to form the original tiles in combination with the overlap are mirror symmetric. In the last case from our example we construct \((\square, 2, 3)\) as the overlap and we would need \((\square, 0, 3), (\square \square \square, 1, 3)\) and \((\square, 4, 3)\) to be able to construct \(T\).

This is all done by the \(\text{ovt}\) function:

\[
\text{ovt}(G_g, T, o) = \text{pq}_1(G_g, T, \text{ot}(T, o), \text{enq}(T, o), \{\text{ot}(T, o)\})
\]

\[
\text{rt}(T, o, g) = \text{norm}((S_T + g \cdot p_T) \setminus \text{os}(T, o), f_T + g \cdot p_T, \text{op}(T, o))
\]

\[
\text{enq}(T, o) = \{\text{rt}(T, o, g) | g \in \text{hyperbar}(o + 1D)\}
\]

The function \(\text{rt}\) calculates the triple representing the tile-sets describing the remaining (non-overlapping) parts at phase \(g\) for an overlapping pattern \(o\).

The function \(\text{enq}\) uses this function in turn to create a set the remaining parts for each phase. This set, or queue, is then processed by the following function to add the construction method (overlapping part and the remaining part) for the tiles of \(T\) to the generator graph for each of the phases:

\[
\text{pq}_1(G_g, T, O, Q_1, Q_2) = \begin{cases} 
\text{pq}_2(G_g, Q_2) & Q_1 = \emptyset \\
\text{pq}_1(\text{add}(G_g, (S_T, p_Q, f_Q), T, O, Q), O, Q_1 \setminus \{Q\}, Q_2 \cup \{Q\}) & \text{with } Q \in Q_1 \text{ otherwise}
\end{cases}
\]

The \(\text{pq}_2\) function is then used to make sure that the tiles used in the construction are actually constructed themselves using the \(\text{orm}\) heuristic:

\[
\text{pq}_2(G_g, Q) = \begin{cases} 
G_g & Q = \emptyset \\
\text{pq}_2(\text{orm}(G_g, [Q], Q \setminus \{Q\}) & \text{otherwise}
\end{cases}
\]

This is how the \(\text{ovt}\) function is used to construct tiles with overlap. The \(\text{nov}\) is used when the algorithm decides to consider the possibility of constructing a tile without overlap (either as one of the possibilities, or because the tile simply does not overlap). The function searches in the tiles generated so far for any tile that overlaps with the tile that is to be constructed. At the very least it finds \((\square, 0, 1)\), but perhaps it can find something bigger. If the tile it finds is of "reasonable" size, it is used in the construction. Otherwise the tile to be constructed is simply created by splitting it into two parts. This splitting is done in one of two ways: if the tile is not a hyper-bar (i.e. it contains gaps) then the largest possible hyper-bar is found and split from the tile. If the tile is a hyper-bar we take the dimension in which it is largest (say size \(x\)) and split the tile along that dimension in a tile of size \(1s_2(x)\) and \(x - 1s_1(x)\).

The algorithm decides what a "reasonable" size is by estimating the amount of work that still has to be done after the available tile has been reused. It does this by taking the size of the tile and subtracting the size of the available tile. It then estimates the amount of work that has to be done if the tile is split. It does this by taking the size of the largest of the two components in which the tile should be split. The idea here is that the smaller of the two parts is reused in the larger part, so that no extra costs are incurred for the smaller part. These
two estimates are compared and the one that is estimated to lead to the lowest cost is used.

So, the construction of tiles without reusing the overlap is done as follows:

\[
\text{nov}(G_\text{g}, T) = \text{orm}(\text{orm}(\text{add}(G_\text{g}, T, T_1, T_2), T_1), T_2)
\]

where \((T_1, T_2) = \text{ch}(V_\text{g}, T)\)

The function \text{nov} adds the construction method specified by the function \text{ch} to the generator graph and calls \text{orm} to add the composing tile sets to the generator graph. The function \text{ch} makes the choice between either by reusing available tiles, or by splitting the tile without reuse:

\[
\text{ch}(V_\text{g}, T) = \begin{cases} 
\text{reu}(V_\text{g}, T) & (|S_T| - |S_{T_2}|) \leq (|S_{T_1}| \uparrow |S_{T_4}|) \\
\text{spl}(T) & \text{otherwise}
\end{cases}
\]

where \((T_1, T_2) = \text{reu}(V_\text{g}, T) \wedge (T_3, T_4) = \text{spl}(T)\)

Splitting depends on whether the shape is a hyper-bar or not:

\[
\text{spl}(T) = \begin{cases} 
\text{spl}_1(T) & S_T \text{ is a hyperbar} \\
\text{spl}_2(T) & \text{otherwise}
\end{cases}
\]

If the shape happens to be a hyper-bar we split it along the largest dimension:

\[
\text{spl}_1(T) = (\text{norm}(S_{T_1}, f_T, p_T), \text{norm}(S_{T_2}, f_T, p_T)) \text{ such that }
\]
\[
S_T = \text{hyperbar}(q)
\]
\[
S_{T_1} = \text{hyperbar}(r)
\]
\[
S_{T_2} = S_T \setminus S_{T_1}
\]
\[
r(e) = \text{lsb}_1(q(e))
\]
\[
(\forall d: 0 \leq d < D \land d \neq e : r(d) = q(d))
\]
\[
(\forall d: 0 \leq d < D : q(d) \leq q(e))
\]

Otherwise we just extract the largest hyper-bar from the shape:

\[
\text{spl}_2(T) = (\text{norm}(S_{T_1}, f_T, p_T), \text{norm}(S_{T_2}, f_T, p_T)) \text{ such that }
\]
\[
S_{T_2} = S_T \setminus S_{T_1}
\]
\[
S_{T_1} = \text{the largest hyper-bar in } S_T
\]

The \text{reu} function returns, given an original tile, a triple containing an already existing tile that can be reused and the rest of the original tile.

\[
\text{reu}(V_\text{g}, T) = (T_1, \text{norm}((S_T + f_T) \setminus (S_{T_1} + g \cdot p_T + f_T), 0, p_T)) \text{ such that }
\]
\[
g \in \mathbb{Z}^D
\]
\[
T_1 = \{T_2 \mid T_2 \in V_\text{g} \land (p_T \mod p_T = 0) \land
\]
\[
(S_{T_2} + g \cdot p_{T_2} + f_{T_2}) \subseteq (S_T + f_T)\}
\]

Where the set from which \(T_1\) is selected is the set containing all the triples describing tile set of which all the tiles are a subset of a tile described by \(T\).
Furthermore, this set is constructed in such a way that the tiles occur at a regular position in $T$, which is expressed by the variable $g$.

Note that an increase in $R$ generally results in lower cost generator graphs, since the algorithm considers more generator graphs. However, it does not necessarily mean that it considers a superset. Therefore it is not guaranteed that an increase in $R$ leads to a reduction in cost. The two dimensional median filtering is a good example of this (see Appendix A.2.2). Consider the following example; the function $\text{orm}$ selects the best generator graphs generated by the $\text{ov}$ function. For $R = 1$ the best graph it encounters is $G_1$. For $R = 2$ it also encounters $G_1$, but also the slightly better $G_2$. So the $R = 1$ algorithm selects $G_1$ and the $R = 2$ algorithm selects $G_2$. However, if this happens to be an intermediate step in the generation of the final graph, then there are still some tiles that have to be added to the selected graph. The cost of adding tiles depends on the chosen graph, since tiles already in the chosen graph might be reused. It is possible that some of the tiles which are to be added can be constructed with more reuse in $G_1$ than in $G_2$. The net effect of this is that the final calculation graph produced by the $R = 1$ algorithm can be less costly than the $R = 2$ algorithm, because the $R = 1$ algorithm based its solution on the (at first sight slightly more costly) $G_1$ graph.

![Figure 5.7: $G_g$ for the ORM 1 algorithm for window size 6 (17.5 epw)](image)

The generator graph produced by this algorithm for our running example with window sizes 6 for $R = 1$ is shown in Figure 5.7. This generator graph has a cost of 17.5 epw, which is slightly more expensive than our best solution so far. But when the parameter $R$ is increased to 2 we get the generator graph
CHAPTER 5. HEURISTIC ALGORITHMS

Figure 5.8: A generator graph $G_g(V_g, E_g)$ for $W = \text{tile}(\Box\Box\Box\Box\Box\Box, 0, 1) (16 \text{ epw})$

from Figure 5.8 The reuse of the tile of size 4 gives this solution its efficiency and results in a cost of 16 epw, which is the lowest cost found by our algorithms. For the running example with window sizes 7 the ORM algorithm produces the same calculation graphs as the recursive tiling algorithm (depicted in Figure 5.3), with a cost of 20.5 extrema operations, which is also the lowest cost found by our algorithms.

5.7 Multiple Dimensions

Most of our algorithms (except for Overlap Recursive Mirroring) are designed for one dimensional problems. However, they can be generalized to algorithms for multi-dimensional problems using the method described in this section.

This method is only applicable for a certain class of problems, namely the class of problems whose windows have the shape of a hyper-bar.

$$(\exists s : s \in \mathbb{Z}^D : S_W = \text{hyperbar}(s)) \quad (5.5)$$

The generalized algorithm then functions as follows: We first use the one-dimensional algorithm to produce a solution for windows of size $s(0)$. This solution shows us how we can construct those windows from tiles of size 1. We then take this solution and extend all the generated tiles into another dimension. All the tiles get a size of $s(1)$ in this dimension. The construction methods we have for all the tiles are still valid, except for the tiles of size $1 \times s(1)$. We never generated a construction method for them because they used to be of size 1 in the one-dimensional problem.

So now we execute the algorithm again, but take the tiles with size $1 \times s(1)$ as the new windows and generate a way to constructing these tiles. Since these tiles only have a size larger than one in the second dimension, they can be considered
one dimensional in practice and the algorithm can be executed normally on them.

We repeat this process as often as necessary to get a solution in all the dimensions.

5.8 Maximum reuse

Upon reviewing this report prior to publication we were able to devise a new heuristic which we call the "Maximum Reuse" heuristic. This heuristic is for one dimensional windows, but can be generalized for multi dimensional windows as described in the previous section.

The algorithm works like the ORM 1 algorithm, but instead of choosing between no overlap and reusing the overlap between two successive windows it always chooses for the reuse. Hence the name of this heuristic, since this method results in that the largest possible tile being reused. But only the overlapping tile is fed back into the maximum reuse heuristic, all other tiles, and tiles with no overlap are created using the msbt heuristic.

This new heuristic provides solutions that are close (in cost) to those of the ORM 1 algorithm. Since it is not an improvement upon the ORM algorithm we have not updated the tables and graphs in this report. However, the new heuristic deserves to be mentioned here because its running time is much lower than that of any of the ORM algorithms.
Chapter 6

Results

In the previous chapters we have presented several heuristic algorithms that find generator graphs for specific ROFs. Although we have not yet discussed how an abstract representation in the format of a generator graph can be turned into a concrete implementation, we can already determine the performance of such an abstract filter in terms of extrema operations per window. In this chapter we use this performance measure for two purposes. First we compare the heuristic algorithms from Chapter 5. Second we compare our best abstract designs to designs found in the literature.

6.1 Filter classes

In this section we define the three filter classes that we use to evaluate the heuristic algorithms. These three classes are chosen such that they are representative for filters commonly found in literature and provide sufficient variety to illustrate the differences between the heuristic algorithms.

Although our implementations could easily generate calculation graphs for ROFDPs with an arbitrary number of dimensions, we restrict our attention to one- and two-dimensional cases, where in the latter case we only consider square windows. Although for each window size there are many rank order filters that select a single rank and many more that select an interval of ranks, we restricted our attention to the following three classes:

**Window sorters** Since in a window sorter all the ranks are obtained, this class gives us an upper-bound for any filter which has to select one or more ranks.

**Median filters** This rank order filter is interesting because the median is the most costly single rank that we could wish to obtain.

**Minimum filters** This rank order filter is interesting because the minimum is the least costly single rank that we could wish to obtain. Note that the minimum filter, on grounds of symmetry, cost the same as the maximum filter.
6.2 Heuristic algorithm evaluation

In this section we compare the heuristic algorithms from Chapter 5 by studying their performance on the various filter classes defined in the previous section.

The results of the heuristic algorithms for the classes are plotted in Figures 6.1, 6.2, 6.3, 6.4, 6.5, and 6.6. In all figures the vertical axis shows the cost of the resulting calculation graph in extrema operations per window. The horizontal axis contains the window size. The results of the different heuristic algorithms have been plotted using distinct symbols and colors.

The graphs (sub-figures (a)) show us that the same algorithm can exhibit a quite different trend when it is applied to a different class of filters. All algorithms, for example, display roughly the same, approximately linear, trend when applied to window sorting (except for LSB tiling). However, when those same algorithms are applied to the median filtering algorithm they show different trends; the divide and conquer algorithm shows a linear trend, while the ORM algorithms seem to indicate a logarithmic curve. This difference also holds for the minimum filtering problem, where the divide and conquer algorithm appears to be logarithmic, but the ORM algorithms seem to have an almost constant upper-bound.

Despite the difference in trends the ORM algorithm produces the best filter in each of the three classes. Specifically it is the ORM 4 algorithm (displayed by the small dots) that provides the design with the lowest cost, compared to the other algorithms.

Since we are mainly interested in these best designs, we limit our examination of the trends to the trend of the ORM algorithm. To this end the figures also contain a graph (sub-figures (b)) that contains the results of the ORM algorithms for a specific $R$, with a curve fitted to the data points. For the 1-dimensional cases these graphs contain the results for $R = 4$. However, for the 2-dimensional cases the original graphs contain too few data points to examine the trend effectively. This was solved by taking larger window sizes and generating additional designs. However, the running time of the ORM algorithm for larger window sizes is impractical for larger $R$ and therefore we opted to use ORM 1 algorithm instead.

The curve fitting for both the window sorting and median filtering for 1-dimensional problems is quite satisfactory. The window sorting seems to follow a linear curve, while the median filtering problem exhibits a logarithmic curve. The 2-dimensional variants are fitted with the same curve as the 1-dimensional variants. Even though the curve fits well the predictive quality of the 2-dimensional curves is questionable.

It is possible that the curves would fit better if we had fitted them to the results of the ORM algorithm with a larger parameter $R$. After all, we see that the ORM algorithms produce approximately the same results for smaller window sizes, but for larger windows the ORM algorithms with a higher $R$ produce better results.

The minimum filtering problem seems to be in another category entirely. Fitting a curve to it proves difficult and we have not been able to find an appropriate one. It seems that the one dimensional variant requires between three and four extrema operations for all the window sizes we examined. The 2-dimensional variant seems to require between four and eight extrema operations.

The cost of the minimum filter increases with the size of the window, but this
CHAPTER 6. RESULTS

(a) All algorithms

(b) ORM 4

(c) Legend

Figure 6.1: 1-Dimensional Window Sorting
CHAPTER 6. RESULTS

(a) All algorithms

(b) ORM 1

(c) Legend

Figure 6.2: 2-Dimensional Window Sorting
CHAPTER 6. RESULTS

Figure 6.3: 1-Dimensional Median Filter

(a) All algorithms

(b) ORM 4

(c) Legend

Figure 6.3: 1-Dimensional Median Filter
CHAPTER 6. RESULTS

(a) All algorithms

(b) ORM 1

(c) Legend

Figure 6.4: 2-Dimensional Median Filter
CHAPTER 6. RESULTS

(a) All algorithms

(b) ORM 4

(c) Legend

Figure 6.5: 1-Dimensional Minimum Filter
CHAPTER 6. RESULTS

50

(a) All algorithms

(b) ORM 1

(c) Legend

Figure 6.6: 2-Dimensional Minimum Filter
increase is extremely small. So it is safe to say that implementing a minimum (or maximum) filter requires very few extrema operations, regardless of the number of dimensions and the window size. This is interesting since the best solutions for minimum filter are actually the best solutions for any problem where the cost function always returns a cost of 1 operation. This is discussed in more detail in Chapter 8.

6.3 Comparison with known filters

In the previous section we compared the heuristic algorithms. In this section we compare the results found by our best algorithm with filters taken from the literature. Whenever a publication contains multiple filters for the same ROFDP we have compared our solution with the best one in the publication. These comparisons are based on the cost of the filter in epw. Note that for some publications this measure requires the conversion of the cost in C&S operations (as presented in the publication) to the cost in extrema operations. We do this using the rule that 1 C&S operation costs 2 extrema operations. This conversion is not entirely fair; it might be possible to prune some of those extrema operations, but this can only be done when the design is known. However, most publications only mention their design method and the cost of the resulting filters, but not the filters themselves. Therefore we mark all the costs that have been converted with "+", to indicate that the cost might be lower if extrema operations are pruned.

6.3.1 Window sorters

In this section we survey the first class of problems: the window sorters. We first examine the one-dimensional cases, followed by the two-dimensional cases.

<table>
<thead>
<tr>
<th>Window</th>
<th>ORM 4</th>
<th>Lucke &amp; Parhi [9]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5.00</td>
<td>5.00*</td>
</tr>
<tr>
<td>5</td>
<td>12.00</td>
<td>12.00*</td>
</tr>
<tr>
<td>6</td>
<td>16.00</td>
<td>15.00*</td>
</tr>
<tr>
<td>7</td>
<td>20.50</td>
<td>20.88*</td>
</tr>
<tr>
<td>9</td>
<td>29.00</td>
<td>29.00*</td>
</tr>
</tbody>
</table>

Table 6.1: Cost of known solutions for 1-dimensional window sorting

Table 6.1 contains the costs of one-dimensional window sorters. Lucke and Parhi examine such window sorters, which they call merge sorters (because of the merging used), in the first part of their paper [9]. Their method is quite similar to our ORM algorithm, in that it exploits overlap between successive windows. However, instead of working with infinite input streams they optimize for a fixed number of overlapping windows, which they call the block size. For all block sizes considered the results shown in Table 6.1 are the best results reported by them.

In one specific instance in Table 6.1 indeed in anything we found in the literature, there is a better solution than the one found by the ORM algorithm. The 1-dimensional window sorter for $|S_W| = 6$ by Lucke and Parhi [9] requires
only 15 epw, compared to the 16 required by the solution of the ORM algorithm. Unfortunately we cannot explain this anomaly since we have been unable to reconstruct the design of Lucke and Parhi using the algorithm described in their article.

<table>
<thead>
<tr>
<th>Window</th>
<th>ORM 4</th>
<th>Kolte et al. [8]</th>
<th>Lucke &amp; Parhi [9]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 × 3</td>
<td>35.00</td>
<td>50.00</td>
<td>37.50*</td>
</tr>
<tr>
<td>4 × 4</td>
<td>76.00</td>
<td>128.00</td>
<td></td>
</tr>
<tr>
<td>5 × 5</td>
<td>152.00</td>
<td>288.00</td>
<td>180.44*</td>
</tr>
<tr>
<td>6 × 6</td>
<td>254.50</td>
<td>520.00</td>
<td></td>
</tr>
<tr>
<td>7 × 7</td>
<td>379.00</td>
<td>820.00</td>
<td></td>
</tr>
<tr>
<td>8 × 8</td>
<td>526.67</td>
<td>1,232.00</td>
<td></td>
</tr>
<tr>
<td>9 × 9</td>
<td>711.00</td>
<td>1,804.00</td>
<td></td>
</tr>
<tr>
<td>10 × 10</td>
<td>940.71</td>
<td>2,522.00</td>
<td></td>
</tr>
<tr>
<td>11 × 11</td>
<td>1,169.13</td>
<td>3,346.00</td>
<td></td>
</tr>
<tr>
<td>12 × 12</td>
<td>1,379.83</td>
<td>4,400.00</td>
<td></td>
</tr>
<tr>
<td>13 × 13</td>
<td>1,688.85</td>
<td>5,572.00</td>
<td></td>
</tr>
<tr>
<td>14 × 14</td>
<td>2,017.88</td>
<td>7,050.00</td>
<td></td>
</tr>
<tr>
<td>15 × 15</td>
<td>2,339.65</td>
<td>8,494.00</td>
<td></td>
</tr>
<tr>
<td>16 × 16</td>
<td>2,627.45</td>
<td>10,348.00</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.2: Known solutions for 2-dimensional window sorting

Lucke and Parhi also examined 2-dimensional window sorters [9] and so did Kolte, Smith and Su in their paper [8]. The results from these papers are shown in Table 6.2 and compared to our ORM 4 algorithm. Again the ORM 4 algorithm performs best, but the results of Lucke and Parhi come very close. Although the focus of Kolte et al. lies on finding implementations that are easy to implement on a SIMD processor, they nevertheless manage to get good results.

6.3.2 Median filter

In this section we survey the second class of problems: the median filters. We first examine the one-dimensional cases, followed by the two-dimensional cases.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.00</td>
<td>5.00*</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5.00</td>
<td>11.20*</td>
<td>31.00*</td>
</tr>
<tr>
<td>7</td>
<td>8.50</td>
<td>19.66*</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10.00</td>
<td>32.00*</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.3: Known solutions for 1-dimensional median filtering

In the article [9] Lucke and Parhi also make an excursion to the median filter. Their approach is to specify the median filter as a stack filter in a max-min structure and then applying their method for implementing stack filters. However, as shown in Table 6.3 the resulting solution is not very efficient. In
fact, it is less costly to implement a window sorter according to their own method (see Table 6.1) than it is to implement a median filter (for $|S_W| = 5$).

Chakrabarti focusses entirely on 1-dimensional median filters in his paper [2] and achieves better results. In a later publication with Wang [3] he also considers 2-dimensional median filtering and again obtains good results (see Table 6.4). Still, the ORM 4 algorithm produces even better results.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3 × 3</td>
<td>16.00</td>
<td>26.00*</td>
<td>17.00</td>
</tr>
<tr>
<td>5 × 5</td>
<td>53.00</td>
<td>61.00*</td>
<td>107.00</td>
</tr>
<tr>
<td>7 × 7</td>
<td>105.75</td>
<td></td>
<td>317.00</td>
</tr>
<tr>
<td>9 × 9</td>
<td>171.00</td>
<td></td>
<td>685.00</td>
</tr>
<tr>
<td>11 × 11</td>
<td>232.75</td>
<td></td>
<td>1254.00</td>
</tr>
<tr>
<td>13 × 13</td>
<td>321.25</td>
<td></td>
<td>2036.00</td>
</tr>
<tr>
<td>15 × 15</td>
<td>404.38</td>
<td></td>
<td>3018.00</td>
</tr>
</tbody>
</table>

Table 6.4: Known solutions for 2-dimensional median filtering

6.3.3 Minimum filter

In this section we examine the third and final class of filters, namely the minimum filters. There has been little attention for the minimum filter in the literature, but Harrington [5] has examined the problem and has patented a novel method for both 1 and 2-dimensional minimum filters. The 1-dimensional method proposed by Harrington requires just under 3 epw for any window size $|S_W|$. In most cases this is better than the solutions found by our algorithms. Our algorithms only obtain solutions with such a similar low cost for $|S_W| = 2^i - 1$ for $i \in \mathbb{Z}$.

The two dimensional approach specified in [5], however, is less efficient. The proposed method for 2-dimensional rectangular windows requires just under 11 epw, which is more expensive than the solutions found by our algorithms. Furthermore, unlike our approach and the 1-dimensional approach from [5], the two dimensional approach in [5] is not suited for every associative operation. The method presented in the patent relies on combining overlapping tiles and thus it is only suited for idempotent operators like minimum and maximum.

6.4 Summary

In the previous chapter we presented several heuristic algorithms that generate a generator graph and thus a design for a ROF. The so-called ”Overlap Recursive Mirroring” or ORM algorithm is the best algorithm we designed. The only method we found in the literature that is capable of obtaining better results than our ORM algorithm is by Harrington [5], but only for 1-dimensional minimum filtering. The approach by Lucke and Parhi [9] is capable of obtaining similar results, but only for the window sorting problem. For the median filtering problem our ORM algorithm is the best so far.
The ORM algorithm has a parameter $R$ that gives an indication of the search space covered by the algorithm. As $R$ increases more and more possible designs are considered, and better designs are found. For $R = 4$ the ORM algorithm finds the better solutions than for $R = 3$, however in the majority of cases up to $|S_W| \leq 250$ the ORM algorithm with $R = 3$ already provides the best solution. The solutions provided by $R = 4$ are only better in a handful of cases.

An overview of the costs of the designs generated by our algorithms can be found in Appendix A. Furthermore the detailed designs of several of the smaller solutions for window sorting and median filtering can be found in Appendix B.
Chapter 7

Parallel ROF Implementations

In this chapter we present two kinds of parallel ROF implementations, viz. implementations in the form of VLSI-circuits and implementations in the form of parallel programs suited for SIMD machines like the EVP$^{16}$. We discuss the scalability of the resulting implementations and for some specific 2-dimensional filters we provide the number of clock cycles per output required by the EVP$^{16}$ solution. As an intermediate step towards a parallel implementation we first show how to obtain a sequential program and in passing we also show how to handle finite input streams.

7.1 Sequential Implementation

So far we have used calculation graphs as abstract representations of ROFs. Recall that a calculation graph is an acyclic graph whose nodes represent merging operations and whose edges represent the dependencies those operations. By sorting the nodes topologically we obtain an order in which to perform the merge operations and thereby a sequential program.

Unfortunately, because the calculation graph is an infinite graph the program is infinite as well. We can overcome this problem by introducing loops that capture the repetitive nature of the calculation graph.

This repetitive behavior is captured by the generator graph. Note that any two tiles (or windows) from the same tile-set in the generator graph are calculated in the same way and that these tiles are separated by a meta-period. Therefore we can obtain a loop in our program that handles one meta-period per iteration.

The number of windows in a single meta-period is $I = \frac{\text{meta-period}}{\text{proc.}}$, and this is therefore the number of windows which are calculated in a single iteration of the sequential program. This differs from common filtering programs which usually calculate only one result per iteration of the loop.

Because a single iteration of our program calculates $I$ new windows it follows that all tiles needed for the calculation of those windows also need to be “obtained”. We use the word obtain and not calculate, since there are two ways
in which the tile can be obtained. On the one hand it can, of course, be calculated by merging the tiles on which it depends. These tiles, in turn, have to be obtained and the recursion ends in the tiles of size one, which form the actual inputs. On the other hand the memoization technique discussed in Section 3.3 can be used. Whenever a tile has been computed that is needed in a future iteration that tile is stored in memory, so that in future iterations the tile only needs to be retrieved, not calculated.

Note that tiles need not be stored indefinitely; by examining the calculation graph we can determine when tiles are no longer needed and remove them from memory. In fact the regularity of the calculation graph allows us to simply introduce a fixed number of buffers which are read and written in a regular pattern.

Whereas the number of windows per iteration is fixed at $I$, the phasing of the windows within the meta-period can be selected freely. This phase determines which of the actual windows are calculated in a single iteration. Therefore this choice affects the implementation, since it determines which tiles need to be buffered and therefore the size of the buffer.

As an illustration consider the ROFDP with $W = \text{tile}(\bullet\bullet\bullet, 0, 1)$ for which a solution is given in Figure 7.1. This figure shows the generator graph with meta-period 2 and some tiles of the three tile sets \( \text{tile}(\bullet\bullet\bullet, 0, 1) \), \( \text{tile}(\bullet\bullet, 1, 2) \) and \( \text{tile}(\bullet, 0, 1) \). The construction of these individual tiles can be derived from the generator graph:

\[
B_{2i+1} = A_{2i+1} \otimes A_{2i+2} \\
W_{2i} = A_{2i} \otimes B_{2i+1} \\
W_{2i+1} = B_{2i+1} \otimes A_{2i+3}
\]

If we choose phase 0, then $W_{2i}$ and $W_{2i+1}$ (for $i \in \mathbb{Z}$) would fall into the same meta-period. So the iteration that calculates $W_0$ and $W_1$ needs to obtain tiles $A_0$, $B_1$, $A_2$, and, recursively, $A_1$ and $A_2$. Of these tiles $A_0$ and $A_1$ can be read from memory, because they were also needed in the previous iteration, and
tiles $A_2$ and $A_4$ need to be stored in memory because they will be needed in the next iteration. Choosing phase of 0 therefore requires the buffering of two tiles of size 1 between each iteration.

Next, choose the phase to be 1, then $W_{2i+1}$ and $W_{2i+2}$ would fall into the same meta-period. So the iteration that calculates $W_1$ and $W_2$ needs to obtain tiles $B_1$, $A_3$, $A_2$, $B_3$ and, recursively, $A_1$, $A_2$, $A_3$ and $A_4$. Note that the tiles $A_2$ and $A_3$ have to be obtained twice in this iteration. Of the tiles to be obtained the tiles $B_1$ and $A_2$ can be read from memory, because they were needed in the previous iteration, and tiles $B_3$ and $A_4$ need to be stored in memory because they will be needed in the next iteration (the one that calculates $W_3$ and $W_4$). Choosing a phasing of 1 therefore requires the buffering of one tile of size 1 and one tile of size 2 between each iteration.

It seems that a phase of 1, for our example, requires the largest amount of buffering. However, keep in mind that the size of a tile and the size of the interval of ranks required from a tile are not necessarily the same (see Section 3.1). If Figure 7.1 is a solution for a minimum or maximum filter, for example, then we only need one item from the tile of size 2. So the buffer cost for both phases would be 2 data items, in that case, therefore, the phase would not matter.

Finding the best phase can be done with a simple exhaustive search, since the meta-period for a dimension is an upper bound for the number of phases in that dimension. Note that choosing the correct phase can diminish the amount of buffering, but that it does not adversely affect the number of extrema operations that have to be performed.

\[
\text{do } \text{true } \rightarrow \\
\{ \text{buf}0 = A_{2i} \} \{ \text{buf}1 = A_{2i+1} \} \\
\text{a2} := \text{read}(A_{2i+2}); \\
\{ \text{a2} = A_{2i+2} \} \\
b := \text{buf}1 \bowtie \text{a2}; \\
\{ \text{b} = B_{2i+1} = A_{2i+1} \bowtie A_{2i+2} \} \\
W_{2i} := \text{buf}0 \bowtie \text{b}; \\
\{ W_{2i} = A_{2i} \bowtie B_{2i+1} \} \\
\text{buf}0 := \text{a2}; \\
\{ \text{buf}0 = A_{2i+2} \} \\
a3 := \text{read}(A_{2i+3}); \\
\{ \text{a3} = A_{2i+3} \} \\
W_{2i+1} := \text{b} \bowtie \text{a3}; \\
\{ W_{2i+1} = B_{2i+1} \bowtie A_{2i+3} \} \\
\text{buf}1 := \text{a3}; \\
\{ \text{buf}1 = A_{2i+3} \} \\
i := i + 1; \\
\text{od}
\]

Figure 7.2: Pseudo-code for the example in Figure 7.1

Once the phase has been chosen the sequential program can be written. The program executes an infinite loop that calculates $I$ windows per iteration and
between each iteration the program buffers several data items, in accordance with the explanation above. Effectively, we have chosen a part of the calculation graph and implemented it as a number of statements. Because the calculation graph can be formed by replication of the chosen part we can implement the filter by repeating those statements in a loop.

A phase of zero for our example leads to the pseudo-code in Figure 7.2. Note that the pseudo-code resembles the generator graph. The number of merge operations equals the number of tile-sets with tiles larger than one and the number of read operations equals the number of tile-sets with tiles of size one in the generator graph. It is, in fact, possible to generate the program code for a rank order filter automatically from the generator graph.

In Section 7.3 we will show how to deal with an input stream that is finite in one or more dimensions. This will effectively limit the number of iterations for the loops that are related to those finite dimensions. Using this we can generalize our approach to multi-dimensional problems where the program will consist of several nested loops, one loop for each dimension. The approach will work as long as there is at most one dimension in which the inputs are infinite; the loop related to that dimension will become the outermost loop.

7.2 Parallel Implementation

The parallel implementations in this section are based on one iteration of the sequential program. The VLSI implementation, for example, will be based on components that implement one such iteration. Furthermore, although there are data dependencies between consecutive iterations, it is also possible to implement each iteration on a processing element of a SIMD machine since these dependencies can be resolved.

7.2.1 VLSI Circuit

Implementing the rank order filter in parallel in a VLSI circuit is quite straightforward.

First we create a component that performs the same operation as a single iteration of the sequential program. This component has two sets of inputs and two sets of outputs. One set of inputs is for the tiles of size one, the actual inputs of the filter. The other set of inputs is for the tiles the component needs to read from a buffer from a previous iteration. Similarly the outputs are divided into two sets; those for the complete windows and those for the tiles that need to be buffered for a next iteration.

This component is depicted in Figure 7.3. Note that we used boxes with the symbol "⊲ ⊳" to denote the merging networks. One such component can be used to get an implementation that calculates $I$ outputs in parallel. To do this we simply connect the components buffer outputs to a buffer and connect the buffer to the components buffer inputs, as is depicted in Figure 7.4(a).

To obtain a circuit that calculates $L = IP$ windows in parallel we simply use $P$ components and connect them in a similar manner. For $P = 4$ this has been done in Figure 7.4(b). In this way any number of components may be connected to increase the number of results calculated by the circuit.
CHAPTER 7. PARALLEL ROF IMPLEMENTATIONS

(a) Generic

(b) For the running example

Figure 7.3: Basic component for VLSI implementation

(a) $P = 1$

(b) $P = 4$

Figure 7.4: Parallel rank order filtering circuit
The actual cycle time of the circuit is independent of $P$. There is a dependency between components since tiles that are calculated in one component may be needed (also) in the others, but it is the depth of the calculation graph (from tiles of size 1 to the windows) and the calculation time of the merging networks that determine the cycle time. Neither of these depend on $P$, therefore the cycle time is independent of the amount of parallelism, and an arbitrary throughput can be obtained. This throughput is $O(L)$ and therefore it scales linearly with the size of the circuit.

The circuit may even be pipelined to obtain a smaller cycle time, since it is essentially a feed forward network. The calculation graph is an directed acyclic graph and this is reflected in the circuit; smaller tiles are fed into the rest of the circuit to calculate larger tiles. Although there appears to be a cycle in the circuit (see Figure 7.4) a detailed examination of the circuit will show that there are no cycles, only some feed forward paths that contain one or (possibly) more buffers.

A note on optimization: a component that performs a compare and swap operation is cheaper (in terms of the number of gates) than two components that perform an extrema operation. Therefore a good optimization step when implementing the merging networks is to replace all the pairs of extrema components that actually perform a C&S operation by a C&S component.

### 7.2.2 SIMD Processor

In an SIMD processor with $P$ processing elements (PEs) there are more restrictions to implementing an algorithm than in hardware. However we can calculate $P$ iterations of the sequential program in one iteration of the SIMD program. We can also look at it from a different angle and consider one iteration of the SIMD program as the simulation of one cycle of the VLSI implementation (with $P$ components).

The problem in vectorizing the sequential program is that there is a data dependency between the iterations; the tiles stored in the buffers. The reading from and writing to these buffers needs to be synchronized such that a read from a buffer is immediately followed by a write to that buffer. In this way the communication of the tile stored in the buffer can be implemented by a shift operation. During the shift operation each of the PEs receives a tile from its neighbor that is working on the preceding iteration (the read from the buffer) and sends a tile to its neighbor that is working on the next iteration (the write to the buffer).

The calculation of a tile that has to be written to a buffer will never depend on the contents of that same buffer (tiles can only depend on smaller tiles), it is therefore always possible to rearrange the statements in such a way that the read from and write to a buffer follow each other immediately.

Note that during a shift operation the tile received by PE 0 has to come from a buffer and the element send by PE $P - 1$ has to be placed in that buffer, and that the contents of this buffer need to be kept for the next iteration of the SIMD program. Most SIMD processor have some sort of functionality to handle the data items at the edge of the PE array during a shift operation, so this is not a problem.

There is one last problem that we have to tackle before we have our SIMD implementation. This concerns the reading of the inputs and the writing of the
outputs. Since all the tile-set in the generator graph have a period equal to the meta-period the implementation requires the inputs and produces the outputs with that same period. Consider our running example; the sequential program calculates first the next even and then the next odd window in an iteration. The SIMD program, however, first calculates the next $P$ even windows and than the next $P$ odd windows. The same reasoning holds for the reading of the inputs.

This is no problem if the inputs and outputs happen to be in memory in that particular order. However in most cases the inputs are in memory in consecutive order and the outputs are also desired in a consecutive order. Therefore we can use so-called strided memory access to implement our algorithm, as shown in Figure 7.5.

```
do true →
  {buf0 = A2;
   buf1 = A2+1}
a2_vec := strided_read (A2; 2); {a2_vec[j] = A2(i+j)+2 for 0 ≤ j < P}
a3_vec := strided_read (A2+3; 2); {a3_vec[j] = A2(i+j)+3 for 0 ≤ j < P}
t1_vec, buf1 := shift (buf1, a3_vec);
  {t1_vec[j] = A2(i+j)+1 for 0 ≤ j < P ∧ buf1 = A2(i+P)+1}
  b_vec := t1_vec ⊲ ⊳ a2_vec;
  {b_vec[j] = B2(i+j)+1 = A2(i+j)+1 ⊲ ⊳ A2(i+j)+2 for 0 ≤ j < P}
t0_vec, buf0 := shift (buf0, a2_vec);
  {t0_vec[j] = A2(i+j) for 0 ≤ j < P ∧ buf0 = A2(i+P)}
W2i := t0_vec ⊲ ⊳ b_vec;
  {W2(i+j) = A2(i+j) ⊲ ⊳ B2(i+j)+1 for 0 ≤ j < P}
W2i+1 := b_vec ⊲ ⊳ a3_vec;
  {W2(i+j)+1 = B2(i+j)+1 ⊲ ⊳ A2(i+j)+3 for 0 ≤ j < P}
i := i + P;
)
```

Figure 7.5: SIMD pseudo-code for the example in Figure 7.1

Not many SIMD processors support strided memory access, but another solution to this problem is possible if the processor is capable of shuffling the vectors. A shuffle operation is used to reorder the elements of a vector. So, after the inputs have been read in consecutive order we use several (masked) shuffling operations to distribute them to the correct PEs. This process is depicted in Figure 7.6 for an SIMD processor with 4 PEs. On the top the figure shows the inputs as they read from memory, at the bottom it shows how two vectors have been created, both containing the inputs with a period of two, but the left-most vector with a phase of 0, and the right-most vector with a phase of 1. The numbers at the arrows indicate the number of the shuffle operation in which the movement of the input to the destination vector is accomplished. The figure shows that 4 shuffle operations are needed, assuming that each shuffle operation can have only one source and one destination vector. A similar shuffling pattern can be used to get the outputs back into a consecutive order.
Figure 7.6: Shuffling consecutive inputs for $P = 4$ and a meta-period of 2

Note that the number of shuffling operations depends only on the meta-period, not on the number of processing elements ($P$). Therefore the running time of an iteration of the SIMD program is still independent of $P$ and we end up with an SIMD implementation of which the throughput ($O(L) = O(IP)$) scales linearly with the number of processing elements.

A note on optimization: a SIMD processor is usually capable of the required extrema operations. However should it be capable of the C&S operation then using that instruction wherever possible in the implementation would be a good idea. A C&S (or extrema) operation can even be simulated by comparing two values and using the result as a mask to some assignment operations, as shown by the pseudo-code in Table 7.1.

<table>
<thead>
<tr>
<th>Option 1:</th>
<th>Option 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>t := a;</td>
<td>mask := compare.gte(a, b);</td>
</tr>
<tr>
<td>a := min(a,b);</td>
<td>t := a;</td>
</tr>
<tr>
<td>b := max(t,b);</td>
<td>a := masked.move(b, mask);</td>
</tr>
<tr>
<td></td>
<td>b := masked.move(t, mask);</td>
</tr>
</tbody>
</table>

Table 7.1: Two options for implementing a C&S operation

Clearly option 1 requires less instructions, and by using dead code elimination it can be automatically removed from the program if necessary. However, option 2 might be faster than simple extrema operations if the processor also provides VLIW parallelism. Although option 2 clearly requires more operations, in a VLIW processor there might be multiple functional units capable of a move operation (addition of zero, shifting over a distance of zero, etc.), while there is usually only one functional unit capable of extrema operations. So option 2 implementation might require less clock cycles for some processors. For an example of such a processor, see Section 7.4.

### 7.3 Finite input streams

So far we have considered input streams that are infinite. In practice, however, this is not the case. How do our implementations cope with this?

The answer is relatively simple: the finite input streams are extended to contain a number of elements that is a multiple of $L$. In that way all of our
implementations can simply make an integer number of iterations. Extending the input stream with dummy inputs is something that happens routinely in image processing to prevent artifacts at the edge of a picture.

The only problem with this solution is that the implementation now has some overhead compared to the theoretical solution found by the calculation graph. However, this overhead is $O(L)$ for the entire input stream, so it becomes less and less important for the larger input streams.

This solution can be extended to multiple-dimensions. If there are $d \leq D$ dimensions in which the input stream is finite then this introduces an overhead of $O(L^d)$. Again this overhead becomes less and less important as the size of the input stream increases.

### 7.4 Philips EVP\textsubscript{16}

The Philips EVP\textsubscript{16} is an SIMD processor with $P = 16$ processing elements \[1\]. The EVP\textsubscript{16} is a processor that supports extrema operations and shuffle operations. So, although it does not support C&S operations or strided memory access, it is possible to implement our solution. We wrote a program that, given a generator graph, automatically produces the program code for the EVP\textsubscript{16}.

The EVP\textsubscript{16} also supports VLIW parallelism and masked operations. The extrema operations and compare operations are only supported on one unit\[1\], while assignment or move operations are supported on 4 units\[2\]. This means that the alternative implementation of the C&S operation, i.e. option 2 from Table 7.1, is possible and might even be an improvement.

Option 2 would also take two clock cycles (one for the compare and move operation and the other for the two remaining masked move operations). So there seems to be no gain. However, with this implementation we can start a new C&S operation in each clock cycle. Assuming, of course, that the second of the subsequent C&S operations does not require the outputs of the first as input. This is something to take into account.

Another thing we must take into account is that the compiler also needs to schedule some other operations, like the memory accesses and the shuffling of vectors. These operations can only be performed on the same units that are used by option 2. So producing the most efficient program code is an entirely new optimization problem.

We have not examined this problem extensively, because with (for example) $N$ C&S operations in a program, and two options for implementing each C&S operation, there would be $2^N$ configurations of the program in total. Needless to say, it is impractical to examine all these possibilities and select the best one.

To find a good program we use two steps. The first step is to find all the C&S operations of which one of the extrema operations can be pruned. For these C&S operations the best implementation is option 1, since that implementation reduces to a single extrema operation and thus requires only one clock cycle on the ALU unit.

After this first step there may still be C&S operations left for which we can choose between the two alternative implementations. So in the second step we use a simple heuristic algorithm to find good choices for these remaining C&S

---

1 VALU
2 VLSU, VMAC, VALU and VSHU
operations. The choices for these operations can be represented as a bit vector. We start by using the bit vector representing that all C&S operations are implemented as option 2. We then examine the neighborhood of the current configuration for a new configuration that is better. The neighborhood of a configuration consists of all configurations with a Hamming distance of one to the current configuration, i.e., all the configurations that have a different option for a single one of the C&S operations. The best configuration in this neighborhood is taken and we apply the same algorithm again and again, until there is no more improvement. The performance of the resulting implementations after such an optimization can be found in Table 7.2. However, solving this optimization problem with a more sophisticated algorithm may improve these results.

<table>
<thead>
<tr>
<th>Window:</th>
<th>3 × 3</th>
<th>5 × 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clock cycles per:</td>
<td>iteration</td>
<td>output</td>
</tr>
<tr>
<td>Using only option 1</td>
<td>84</td>
<td>1.31</td>
</tr>
<tr>
<td>Using only option 2</td>
<td>77</td>
<td>1.20</td>
</tr>
<tr>
<td>After optimization</td>
<td>71</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Table 7.2: Clock cycles for median filtering on the EVP$_{16}$

Table 7.2 presents the information on two 2-dimensional median filters implemented on the EVP$_{16}$. The implementations produce $L = IP$ outputs per iteration. The number of windows per iteration happens to be $I = 4$ for both problems and the processor has $P = 16$ PEs, therefore the number of clock cycles per output is simply the number of clock cycles per iteration divided by $L = 64$. The rows of the table indicate the number of clock cycles for different implementation strategies. The row marked "Using only option 1" and "Using only option 2" shows the results when step 2 of our optimization is skipped and all remaining C&S operations are implemented using option 1 or option 2 respectively. The row marked "After optimization" shows the results when the neighborhood search algorithm is used to optimize the implementation.

Note that the 3 × 3 and 5 × 5 median filters require 1.15 and 6.6 clock cycles per output respectively on the AltiVec, according to [8]. Our implementation on the EVP$_{16}$ with 1.11 and 3.77 clock cycles respectively is an improvement on those numbers, especially for the latter filter.

Table 7.2 also shows that median filters can be implemented very efficiently. No sequential algorithm is capable of producing outputs at the rate shown in the table. There are sequential algorithms that have a better complexity than our solutions, but the oblivious approach allows for easy parallelism. For practical window sizes this parallelism compensates small differences in complexity and provides better overall performance.
Chapter 8

Conclusion

In this report we have presented a mathematical framework for representing both the rank-order filter design problem (ROFDP) and the solutions to that problem, i.e. the rank-order filters (ROF) themselves. We have also introduced several heuristic algorithms for solving the ROFDP and have compared the filters found by those algorithms to filters found in the literature. Furthermore, we have shown how our designs can be used to obtain scalable, parallel implementations of ROFs as VLSI circuits or as programs on a SIMD machine like the EVP\textsubscript{16}.

A novelty of our designs is the mirroring of merging and/or sorting networks prior to the pruning (or "dead-code elimination"). We have not encountered this optimization in the literature so far, but it improves results significantly. Furthermore we present a cost function for pruned merging networks, both for the standard versions and for the versions with the mirroring optimization applied.

The major result of our research, however, we consider the mathematical framework that uses concepts like tiling, the calculation graph and the generator graph. This framework enables us to formalize the reuse that takes place in a ROF. In addition it allows us to make a clear distinction between the calculation to be performed by the ROF and the (parallel) implementation itself. This separation of concerns allows us to design the solution first and choose the implementation method and the amount of parallelism at a later stage. In all cases considered this approach has resulted in designs that require less extrema operations per window than the ones reported in the literature so far, except for 1-dimensional minimum filtering [5] and a 1-dimensional window sorter for windows of size 6 [9].

Since we have not been unable to provide an upper bound for the cost for arbitrary large window sizes. The asymptotic complexity of our designs is unknown, just like the designs presented in [2, 3, 8, 9]. Although the cost of our ROF designs is very low, we know that the complexity is not as low some of the best algorithms found in the literature [6, 10], due to the fact that we resort to oblivious sorting methods. Because our designs rely on oblivious sorting methods they have two major advantages over these algorithms. First of all they allow for scalable parallelism; the throughput of the filter scales linearly with the amount of parallelism. Secondly, the constant factor for our designs is very low, which means that our designs perform better for practical
The methods presented in this report are widely applicable. Any problem consisting of a sliding window and an aggregation operator that is both associative and commutative can benefit from our approach. The cost of an implementation using an aggregation function with a cost of one (like the addition, multiplication, minimum and maximum operators) is equal to the cost of the minimum filter. The only difference is that in the first two cases the cost is measured in addition operations per window and multiplications operations per window respectively, instead of extrema operations per window.

Other problems for which our methods can be adapted are weighted rank order filters. In these filters the set of coordinates forming the windows become multi-sets, and the weight of a coordinate is represented by the number of times it is in the multi-set.

Although the generator graphs found by our algorithms are already quite good there is still room for improvement. The largest improvement, however, can be achieved by improving on the implementation of our designs on SIMD processors. Generating efficient code that can be readily optimized by the compiler for such machines is a hard problem that we have not examined fully yet. All these improvements, however, are possible topics for future research.

In summary, our research has shown that the computational gap between linear filters and rank order filters is not as large as was previously believed. The improvements made in this report are sufficient to make the use of rank order filters in real-time systems attractive.
Bibliography


Appendix A

Algorithm Results

The next sections of this appendix present tables containing the results of the different algorithms described in Chapter 5. Each row of such a table corresponds to a specific window size $W$, and each column to a specific algorithm. For the sake of brevity, we have adopted the following numbering scheme to identify the algorithms:

0. Divide and Conquer (Section 5.2)
1. Recursive Tiling (Section 5.3)
2. LSB Tiling (Section 5.4)
3. MSB Tiling (Section 5.4)
4. Split Tiling 1 (Section 5.5)
5. Split Tiling 2 (Section 5.5)
6. Split Tiling 3 (Section 5.5)
7. Overlap Recursive Mirroring 1 (Section 5.6 with $R = 1$)
8. Overlap Recursive Mirroring 2 (Section 5.6 with $R = 2$)
9. Overlap Recursive Mirroring 3 (Section 5.6 with $R = 3$)
10. Overlap Recursive Mirroring 4 (Section 5.6 with $R = 4$)

The last column of these tables indicates which of the algorithms provided a solution with the minimum cost.

A.1 One Dimensional Problems

In this section we present the results of our algorithms for one-dimensional ROFDPs.

- $D = 1$
- $W = \text{tile} \left( \text{hyperbar}(W), 1, 0 \right)$

The following subsections consider various choices for $l$ and $u$. 
### A.1.1 Window Sorting

This section considers the problem of sorting the entire window. Hence $l = 0$ and $u = w + 1$.

<table>
<thead>
<tr>
<th>$w$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Best strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
</tbody>
</table>

**APPENDIX A. ALGORITHM RESULTS**
### A.1.2 Median Filtering

This section considers the problem of selecting the median of a window. Hence \( l = a = \lceil \frac{w}{2} \rceil \).

<table>
<thead>
<tr>
<th>w</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Best strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>3</td>
<td>4.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
<td>1, 2, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>4</td>
<td>11.00</td>
<td>6.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>5</td>
<td>18.00</td>
<td>9.50</td>
<td>11.00</td>
<td>8.50</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>6</td>
<td>24.00</td>
<td>12.00</td>
<td>13.50</td>
<td>10.80</td>
<td>8.50</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>7</td>
<td>30.00</td>
<td>15.50</td>
<td>16.50</td>
<td>13.30</td>
<td>11.00</td>
<td>9.50</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>9.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>8</td>
<td>35.00</td>
<td>18.00</td>
<td>18.50</td>
<td>15.30</td>
<td>12.50</td>
<td>11.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>9</td>
<td>40.00</td>
<td>20.50</td>
<td>20.50</td>
<td>17.30</td>
<td>14.50</td>
<td>13.00</td>
<td>12.00</td>
<td>12.00</td>
<td>12.00</td>
<td>12.00</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>10</td>
<td>45.00</td>
<td>23.00</td>
<td>23.00</td>
<td>19.80</td>
<td>17.00</td>
<td>15.50</td>
<td>14.50</td>
<td>14.50</td>
<td>14.50</td>
<td>14.50</td>
<td>1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
</tbody>
</table>

\[ a = \lceil \frac{w}{2} \rceil \]

\( a \) is the median of the window.

\[ l = a \]

\( l \) is the length of the window.

\[ w \] is the width of the window.

\[ \lfloor \] is the floor function.
APPENDIX A. ALGORITHM RESULTS

74

<table>
<thead>
<tr>
<th>n</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Best strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>2</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>3</td>
<td>2.0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>4</td>
<td>3.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>5</td>
<td>3.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>6</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>7</td>
<td>4.0</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>8</td>
<td>3.5</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>9</td>
<td>4.0</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>10</td>
<td>5.0</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>11</td>
<td>6.0</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>12</td>
<td>6.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>13</td>
<td>6.0</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>3.75</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>14</td>
<td>7.0</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>15</td>
<td>7.0</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>3.25</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>16</td>
<td>7.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>17</td>
<td>7.0</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>2.75</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>18</td>
<td>7.0</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>19</td>
<td>7.0</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>2.25</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>20</td>
<td>7.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
</tbody>
</table>

A.1.3 Minimum Filtering

This section considers the problem of selecting the minimum of a window. Hence

\[ l = u = 0. \]
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
<th>Column 7</th>
<th>Column 8</th>
<th>Column 9</th>
<th>Column 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.0</td>
<td>5.25</td>
<td>9.0</td>
<td>9.0</td>
<td>5.94</td>
<td>3.61</td>
<td>3.91</td>
<td>3.48</td>
<td>3.48</td>
<td>3.48</td>
</tr>
<tr>
<td>10.0</td>
<td>2.91</td>
<td>2.91</td>
<td>10.0</td>
<td>5.97</td>
<td>2.91</td>
<td>2.91</td>
<td>2.91</td>
<td>2.91</td>
<td>2.91</td>
</tr>
<tr>
<td>10.0</td>
<td>6.0</td>
<td>7.0</td>
<td>8.0</td>
<td>6.0</td>
<td>3.86</td>
<td>6.0</td>
<td>3.41</td>
<td>3.41</td>
<td>3.41</td>
</tr>
<tr>
<td>11.0</td>
<td>4.0</td>
<td>4.0</td>
<td>7.5</td>
<td>6.02</td>
<td>3.36</td>
<td>4.0</td>
<td>3.28</td>
<td>3.17</td>
<td>3.17</td>
</tr>
<tr>
<td>10.0</td>
<td>5.0</td>
<td>7.5</td>
<td>7.5</td>
<td>6.02</td>
<td>3.36</td>
<td>4.5</td>
<td>3.78</td>
<td>3.69</td>
<td>3.58</td>
</tr>
<tr>
<td>11.0</td>
<td>3.25</td>
<td>3.25</td>
<td>8.0</td>
<td>6.05</td>
<td>3.13</td>
<td>3.25</td>
<td>3.06</td>
<td>3.06</td>
<td>3.03</td>
</tr>
<tr>
<td>8.0</td>
<td>4.25</td>
<td>4.5</td>
<td>7.0</td>
<td>5.81</td>
<td>3.88</td>
<td>4.0</td>
<td>3.58</td>
<td>3.58</td>
<td>3.58</td>
</tr>
<tr>
<td>11.0</td>
<td>4.25</td>
<td>4.5</td>
<td>8.0</td>
<td>6.09</td>
<td>3.86</td>
<td>4.0</td>
<td>3.81</td>
<td>3.81</td>
<td>3.81</td>
</tr>
<tr>
<td>11.0</td>
<td>3.0</td>
<td>3.0</td>
<td>9.0</td>
<td>6.11</td>
<td>3.5</td>
<td>3.0</td>
<td>2.97</td>
<td>2.96</td>
<td>2.96</td>
</tr>
<tr>
<td>8.0</td>
<td>4.25</td>
<td>4.5</td>
<td>7.0</td>
<td>5.75</td>
<td>3.63</td>
<td>3.48</td>
<td>3.52</td>
<td>3.52</td>
<td>3.52</td>
</tr>
<tr>
<td>11.0</td>
<td>4.25</td>
<td>4.5</td>
<td>8.0</td>
<td>5.83</td>
<td>3.36</td>
<td>3.31</td>
<td>3.17</td>
<td>3.17</td>
<td>3.17</td>
</tr>
<tr>
<td>10.0</td>
<td>5.25</td>
<td>8.0</td>
<td>9.0</td>
<td>6.08</td>
<td>3.88</td>
<td>4.0</td>
<td>3.81</td>
<td>3.81</td>
<td>3.81</td>
</tr>
<tr>
<td>11.0</td>
<td>3.0</td>
<td>3.0</td>
<td>9.0</td>
<td>6.11</td>
<td>3.5</td>
<td>3.0</td>
<td>2.97</td>
<td>2.96</td>
<td>2.96</td>
</tr>
<tr>
<td>8.0</td>
<td>4.25</td>
<td>4.5</td>
<td>7.0</td>
<td>5.75</td>
<td>3.63</td>
<td>3.48</td>
<td>3.52</td>
<td>3.52</td>
<td>3.52</td>
</tr>
<tr>
<td>11.0</td>
<td>4.25</td>
<td>4.5</td>
<td>8.0</td>
<td>5.83</td>
<td>3.36</td>
<td>3.31</td>
<td>3.17</td>
<td>3.17</td>
<td>3.17</td>
</tr>
</tbody>
</table>

APPENDIX A. ALGORITHM RESULTS
APPENDIX A. ALGORITHM RESULTS

76
A.2 Two Dimensional Problems

In this section we present the results of our algorithms for two-dimensional ROFDPs with squares windows.

- \( D = 2 \)
- \( W = \text{tile}(\text{hyperbar}((W, W), 0, D, 1, D)) \)

The following subsections consider various choices for \( l \) and \( u \).

### A.2.1 Window Sorting

This section considers the problem of sorting the entire window. Hence \( l = 0 \) and \( u = W \times W \).

<table>
<thead>
<tr>
<th>( W )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Best strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10</td>
</tr>
</tbody>
</table>

### A.2.2 Median Filtering

This section considers the problem of selecting the median of a window. Hence \( l = 0 \) and \( u = W \times W \).

### A.2.3 Minimum Filtering

This section considers the problem of selecting the minimum of a window. Hence \( l = u = 0 \).
<table>
<thead>
<tr>
<th></th>
<th>8.00</th>
<th>7.50</th>
<th>8.00</th>
<th>8.00</th>
<th>6.50</th>
<th>6.10</th>
<th>6.00</th>
<th>6.00</th>
<th>6.00</th>
<th>5, 6, 7, 8, 9, 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>8.00</td>
<td>7.50</td>
<td>8.00</td>
<td>8.00</td>
<td>6.50</td>
<td>6.10</td>
<td>6.00</td>
<td>6.00</td>
<td>6.00</td>
<td>5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>11</td>
<td>10.00</td>
<td>5.00</td>
<td>5.00</td>
<td>10.00</td>
<td>6.75</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00, 1, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>12</td>
<td>8.00</td>
<td>7.50</td>
<td>8.00</td>
<td>8.00</td>
<td>6.50</td>
<td>6.50</td>
<td>6.50</td>
<td>6.00</td>
<td>5.83</td>
<td>5.83, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>13</td>
<td>12.00</td>
<td>5.50</td>
<td>6.00</td>
<td>10.00</td>
<td>7.00</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>14</td>
<td>15.00</td>
<td>8.25</td>
<td>10.00</td>
<td>10.00</td>
<td>7.50</td>
<td>6.50</td>
<td>6.50</td>
<td>6.50</td>
<td>6.50</td>
<td>6.50, 7, 8, 9, 10</td>
</tr>
<tr>
<td>15</td>
<td>12.00</td>
<td>5.38</td>
<td>5.25</td>
<td>12.00</td>
<td>7.75</td>
<td>5.25</td>
<td>5.25</td>
<td>5.25</td>
<td>5.25</td>
<td>5.25, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>16</td>
<td>8.00</td>
<td>8.13</td>
<td>8.00</td>
<td>8.00</td>
<td>8.00</td>
<td>7.00</td>
<td>7.00</td>
<td>6.25</td>
<td>6.25</td>
<td>6.25, 7, 8, 9, 10</td>
</tr>
<tr>
<td>17</td>
<td>14.00</td>
<td>6.00</td>
<td>6.00</td>
<td>10.00</td>
<td>8.13</td>
<td>6.00</td>
<td>6.00</td>
<td>5.75</td>
<td>5.75</td>
<td>5.75, 7, 8, 9, 10</td>
</tr>
<tr>
<td>18</td>
<td>10.00</td>
<td>9.00</td>
<td>10.00</td>
<td>10.00</td>
<td>8.25</td>
<td>7.00</td>
<td>7.00</td>
<td>6.75</td>
<td>6.75</td>
<td>6.75, 8, 9, 10</td>
</tr>
<tr>
<td>19</td>
<td>14.00</td>
<td>5.75</td>
<td>5.50</td>
<td>12.00</td>
<td>8.38</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50</td>
<td>5.50, 2, 5, 6, 7, 8, 9, 10</td>
</tr>
<tr>
<td>20</td>
<td>10.00</td>
<td>8.75</td>
<td>10.00</td>
<td>10.00</td>
<td>8.00</td>
<td>7.00</td>
<td>7.00</td>
<td>6.50</td>
<td>6.50</td>
<td>6.50, 8, 9, 10</td>
</tr>
</tbody>
</table>
Appendix B

Sample solutions

In this appendix we present the details of some of the smaller ROF we found.

B.1 Representation of the ROFs

The ROFs in the rest of this appendix are presented in a table with a specific format.

The first two rows of such a table describe the windows and the desired interval of ranks from those windows that have to be calculated by the filter algorithm. The filter itself is described by the generator graph. The next two rows give the meta-period of this graph and the number of extrema operations per window.

The remaining rows of the table are used to describe the generator graph itself. Each row describes a single node of the graph. The first column contains the triple associated with the node. The next two columns contain the nodes that have edges leading towards the node in question. The final column contains the contribution of this node to the final cost of the entire filter.

B.2 One Dimensional Problems

This section contains some solutions for one dimensional problems.

B.2.1 Window Sorting

The first problem is that of sorting the entire window.

<table>
<thead>
<tr>
<th>Window</th>
<th>[■■■, 0, 1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired Interval:</td>
<td>[0, 2]</td>
</tr>
<tr>
<td>Meta-period:</td>
<td>2</td>
</tr>
<tr>
<td>Total Cost:</td>
<td>5.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(v_g)</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[■, 0, 2]</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[■, 1, 2]</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[■■, 1, 2]</td>
<td>[■, 1, 2]</td>
<td>1.00</td>
</tr>
<tr>
<td>[■■■, 0, 2]</td>
<td>[■■, 1, 2]</td>
<td>2.00</td>
</tr>
<tr>
<td>[■■■, 1, 2]</td>
<td>[■, 1, 2]</td>
<td>2.00</td>
</tr>
</tbody>
</table>
### APPENDIX B. SAMPLE SOLUTIONS

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0..1)</td>
<td>[0..0]</td>
<td>1</td>
<td>8.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V_v</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 1)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0, 0)</td>
<td>[0, 0)</td>
<td>2.00</td>
</tr>
<tr>
<td>[0, 1)</td>
<td></td>
<td>6.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0..1)</td>
<td>[0..0)</td>
<td>1</td>
<td>12.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V_v</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 1)</td>
<td>[1, 1)</td>
<td>1.00</td>
</tr>
<tr>
<td>[1, 1)</td>
<td>[1, 1)</td>
<td>3.00</td>
</tr>
<tr>
<td>[1, 1)</td>
<td>[1, 1)</td>
<td>4.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0..1)</td>
<td>[0..0)</td>
<td>1</td>
<td>16.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V_v</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 0)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 3)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[2, 3)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0, 3)</td>
<td>[0, 1)</td>
<td>0.67</td>
</tr>
<tr>
<td>[1, 1)</td>
<td>[1, 1)</td>
<td>0.67</td>
</tr>
<tr>
<td>[2, 2)</td>
<td>[2, 2)</td>
<td>0.07</td>
</tr>
<tr>
<td>[3, 2)</td>
<td>[3, 2)</td>
<td>0.07</td>
</tr>
<tr>
<td>[0, 3)</td>
<td>[0, 0)</td>
<td>2.00</td>
</tr>
<tr>
<td>[0, 3)</td>
<td>[0, 0)</td>
<td>4.00</td>
</tr>
<tr>
<td>[0, 3)</td>
<td>[0, 0)</td>
<td>4.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0..1)</td>
<td>[0..0)</td>
<td>1</td>
<td>20.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V_v</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[2, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[1, 4)</td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 4)</td>
<td>[1, 4)</td>
<td>0.50</td>
</tr>
<tr>
<td>[2, 4)</td>
<td>[3, 4)</td>
<td>0.50</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[0, 4)</td>
<td>1.90</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[3, 4)</td>
<td>3.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[3, 4)</td>
<td>3.00</td>
</tr>
<tr>
<td>[0, 4)</td>
<td>[0, 0)</td>
<td>3.00</td>
</tr>
<tr>
<td>[2, 4)</td>
<td>[2, 4)</td>
<td>3.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[3, 4)</td>
<td>3.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[3, 4)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0..1)</td>
<td>[0..0)</td>
<td>6</td>
<td>25.33</td>
</tr>
</tbody>
</table>
## B.2.2 Median Filtering

This is the median filtering problem.
Window: [■■■■■■■, 0, 1)
Desired Interval: [3..3]
Meta-period: 4
Total Cost: 8.50

<table>
<thead>
<tr>
<th>(b_i)</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[■, 0, 4)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■, 1, 4)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■, 2, 4)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■, 3, 4)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■■, 1, 4)</td>
<td>[■, 1, 4)</td>
<td>[■, 2, 4)</td>
</tr>
<tr>
<td>[■■■, 1, 4)</td>
<td>[■■, 1, 4)</td>
<td>[■■, 2, 4)</td>
</tr>
<tr>
<td>[■■■■, 1, 4)</td>
<td>[■■■, 1, 4)</td>
<td>[■■■, 2, 4)</td>
</tr>
<tr>
<td>[■■■■, 3, 4)</td>
<td>[■■■, 3, 4)</td>
<td>[■■■, 4, 4)</td>
</tr>
<tr>
<td>[■■■■■, 3, 4)</td>
<td>[■■■■, 3, 4)</td>
<td>[■■■■, 4, 4)</td>
</tr>
<tr>
<td>[■■■■■■, 3, 4)</td>
<td>[■■■■, 3, 4)</td>
<td>[■■■■, 4, 4)</td>
</tr>
</tbody>
</table>

Window: [■■■■■■■■■, 0, 1)
Desired Interval: [4..4]
Meta-period: 2
Total Cost: 10.00

<table>
<thead>
<tr>
<th>(b_i)</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[■, 0, 2)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■, 1, 2)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■■, 1, 2)</td>
<td>[■, 1, 2)</td>
<td>[■, 2, 2)</td>
</tr>
<tr>
<td>[■■■, 1, 2)</td>
<td>[■■, 1, 2)</td>
<td>[■■, 2, 2)</td>
</tr>
<tr>
<td>[■■■■, 1, 2)</td>
<td>[■■■, 1, 2)</td>
<td>[■■■, 2, 2)</td>
</tr>
<tr>
<td>[■■■■■, 1, 2)</td>
<td>[■■■■, 1, 2)</td>
<td>[■■■■, 2, 2)</td>
</tr>
<tr>
<td>[■■■■■■, 1, 2)</td>
<td>[■■■■, 1, 2)</td>
<td>[■■■■, 2, 2)</td>
</tr>
</tbody>
</table>

Window: [■■■■■■■■■, 0, 1)
Desired Interval: [0..0]
Meta-period: 2
Total Cost: 1.50

<table>
<thead>
<tr>
<th>(b_i)</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[■, 0, 2)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■, 1, 2)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■■, 0, 2)</td>
<td>[■, 0, 2)</td>
<td>[■, 1, 2)</td>
</tr>
<tr>
<td>[■■, 1, 2)</td>
<td>[■, 1, 2)</td>
<td>[■, 2, 2)</td>
</tr>
</tbody>
</table>

Window: [■■■■■■■■■, 0, 1)
Desired Interval: [0..0]
Meta-period: 1
Total Cost: 2.00

<table>
<thead>
<tr>
<th>(b_i)</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[■, 0, 1)</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>[■■■, 0, 1)</td>
<td>[■, 0, 1)</td>
<td>[■, 1, 1)</td>
</tr>
<tr>
<td>[■■■■, 0, 1)</td>
<td>[■■■, 0, 1)</td>
<td>[■■■, 1, 1)</td>
</tr>
</tbody>
</table>

Window: [■■■■■■■■■, 0, 1)
Desired Interval: [0..0]
Meta-period: 2
Total Cost: 2.00

B.2.3 Minimum Filtering
This is the minimum filtering problem.
## APPENDIX B. SAMPLE SOLUTIONS

### Table: Sample Solutions

<table>
<thead>
<tr>
<th>$V_0$</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0.2, 0.2)</td>
<td>[1, 2)</td>
<td>0.50</td>
</tr>
<tr>
<td>[1.2, 1.2)</td>
<td>[1, 2)</td>
<td>0.50</td>
</tr>
<tr>
<td>[1, 2)</td>
<td>[0, 2)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

### Table: Sample Solutions (Continued)

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval: [0.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost: 2.50</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$V_0$</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0.2, 0.2)</td>
<td>[1, 2)</td>
<td>0.50</td>
</tr>
<tr>
<td>[1.2, 1.2)</td>
<td>[1, 2)</td>
<td>0.50</td>
</tr>
<tr>
<td>[1, 2)</td>
<td>[0, 2)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

### Table: Sample Solutions (Continued)

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval: [0.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost: 2.25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$V_0$</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[2, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0.4, 0.4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[1.4, 1.4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[1, 4)</td>
<td>[3, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
</tbody>
</table>

### Table: Sample Solutions (Continued)

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval: [0.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cost: 2.75</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$V_0$</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[1, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[2, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[3, 4)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[0.4, 0.4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[1.4, 1.4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[1, 4)</td>
<td>[3, 4)</td>
<td>0.25</td>
</tr>
<tr>
<td>[3, 4)</td>
<td>[1, 4)</td>
<td>0.25</td>
</tr>
</tbody>
</table>
APPENDIX B. SAMPLE SOLUTIONS

84

B.3 Two Dimensional Problems

In this section we present some solutions for two dimensional problems.

We extend our notation for a set of coordinates to two dimensions by drawing matrices of squares. The upper left square of such a matrix corresponds to the coordinate (0,0) the square to the right corresponds to (1,0), etc.

B.3.1 Window Sorting

The first problem is that of sorting the entire window.

<table>
<thead>
<tr>
<th>Window:</th>
<th></th>
<th></th>
<th>(0, 0), (1, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired Interval:</td>
<td>[0, 8]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meta-period:</td>
<td>(2, 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Cost:</td>
<td>35.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V_g</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ] (0, 0), (2, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[ ] (0, 1), (2, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[ ] (1, 0), (2, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[ ] (1, 1), (2, 2)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 0), (2, 2)</td>
</tr>
<tr>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
<td>[ ] (0, 0), (2, 2)</td>
</tr>
<tr>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
<td>[ ] (0, 1), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (1, 0), (2, 2)</td>
<td>[ ] (1, 0), (2, 2)</td>
</tr>
<tr>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
<td>[ ] (1, 1), (2, 2)</td>
</tr>
</tbody>
</table>
**B.3.2 Median Filtering**

This is the median filtering problem.
APPENDIX B. SAMPLE SOLUTIONS

Window: 

| Desired Interval: [4, 4] |
|------------------------|------------------|
| Meta-period: (2, 2)    | Total Cost: 16.00 |

<table>
<thead>
<tr>
<th>Vₖ</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>2.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>2.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.50</td>
</tr>
</tbody>
</table>

Window: 

| Desired Interval: [12, 12] |
|------------------------|------------------|
| Meta-period: (2, 2)    | Total Cost: 16.00 |

<table>
<thead>
<tr>
<th>Vₖ</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>0.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>1.50</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>2.00</td>
</tr>
<tr>
<td>[.0, 0, (2, 2)]</td>
<td>(.0, 0, (2, 2))</td>
<td>2.00</td>
</tr>
</tbody>
</table>
### APPENDIX B. SAMPLE SOLUTIONS

#### B.3.3 Minimum Filtering

This is the minimum filtering problem.

```
<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval</th>
<th>Meta-period</th>
<th>Total Cost</th>
<th>$V_x$</th>
<th>Incoming edges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0), (2, 2)</td>
<td>(0, 0)</td>
<td>(2, 2)</td>
<td>3.00</td>
<td>(0, 0), (2, 2)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(0, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 0), (2, 2)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(0, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 1), (2, 2)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>(1, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 1), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(1, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 1), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 1), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(1, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 1), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 1), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(1, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(1, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(0, 1), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>(1, 0), (2, 2)</td>
<td></td>
<td></td>
<td></td>
<td>(0, 0), (2, 2)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>
```
APPENDIX B. SAMPLE SOLUTIONS

<table>
<thead>
<tr>
<th>Window:</th>
<th>Desired Interval:</th>
<th>Meta-period:</th>
<th>Total Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.00</td>
</tr>
<tr>
<td>$(0, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.00</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.00</td>
</tr>
<tr>
<td>$(1, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.00</td>
</tr>
<tr>
<td>$(0, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(0, 0)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$(1, 1)$</td>
<td>$[0, 0]$</td>
<td>$(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>$[1,0]$, $(2, 2)$</td>
<td>$[1,0]$, $(2, 2)$</td>
<td>$[0,0]$, $(2, 2)$</td>
<td>0.25</td>
</tr>
<tr>
<td>$[1,1]$, $(2, 2)$</td>
<td>$[1,1]$, $(2, 2)$</td>
<td>$[0,1]$, $(2, 2)$</td>
<td>0.25</td>
</tr>
</tbody>
</table>