Selection of Severe Transport Situations in the High-Pressure Gas Network of GTS

Introduction of a Similarity Measure based on Optimal Flow Patterns

S. Maring

TU Delft

gasunie transport services

Selection of Severe Transport Situations in the High-Pressure Gas Network of GTS

Introduction of a Similarity Measure based on Optimal Flow Patterns

by

S. Maring

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday January 26, 2018 at 11:00 AM.

Student number:	4226518	
Project duration:	March 1, 2017 – January 26, 2018	
Thesis committee:	Prof. dr. ir. C. Vuik,	TU Delft, supervisor
	Dr. J.W. van der Woude,	TU Delft, supervisor
	Drs. H. Dijkhuis,	Gasunie Transport Services, supervisor
	Dr. J.J. Steringa,	Gasunie Transport Services, supervisor

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TUDelft gasunie transport services

Abstract

Gasunie Transport Services (GTS) is the operator of the natural gas transmission network in the Netherlands. This network is an entry-exit system where entry and exit capacities can be booked by parties called shippers. Variations in usage lead to an enormous number of possible entry-exit combinations. GTS must be able to accommodate all realistically possible gas transport scenarios that result from these entry-exit combinations. In principle, hydraulic testing of all these scenarios is required to see if the current network is optimally sized for this task. However, calculating all scenarios is extremely time consuming, so a smaller set of severe scenarios that also covers the less severe ones, is needed to represent the complete set.

Reduction of the complete set of transport scenarios is a mathematically challenging task. The current method to reduce the number of these scenarios is workable and probably meets the requirements of day-to-day planning calculations at GTS. This method makes use of the so-called end point representation, which describes transport scenarios by their capacities on entry and exit points only, while e.g. the flow pattern is unknown. The distance function for the current reduction method measures the difference between scenarios by comparing the capacities on end points, where the end points are correlated by their respective transport distances.

A new representation is introduced in this report: the flow representation. This representation makes use of the flow patterns that emerge from balanced combinations of entry and exit capacities. A flow pattern follows from an entry-exit combination by determining its minimum associated transport load. Compared to the end point representation, the flow representation is more intricate to obtain (more calculations are required to get flow patterns), but the result is a more accurate representation in terms of the transport physics of the network. The pay-off is that the distance function for the flow representation can be a lot simpler, e.g. a weighted norm which approximates the transport effort. It also turns out to be more adaptable. For example the diameter, pressure drop and other network information can easily be included in this weighted norm.

In this report both representations are compared. However, from the experiments no final conclusion can be given for which representation has a better performance. Future studies, e.g. involving hydraulic calculations, are recommended to conclude this matter.

Preface

This report is written for my graduation project to fulfil the requirements on the Master degree of Applied Mathematics at the Delft University of Technology. Gasunie Transport Services is in need of a better method for planning gas flows. The employees of the planning department want a good motivation for their method and a faster program to make calculations for the planning. In this report, it is seen that there are interesting theories for the planning of gas and there are much more areas of research which can be explored.

I want to acknowledge some people who have supported, guided and helped me to do this research project and write this report.

First of all, I want to thank my supervisors Kees Vuik and Jacob van der Woude for their support from the TU Delft. They gave me good feedback during my research process and on my report. Despite the distance, they were involved in my project by meeting in person and by having a Skype meeting.

I also would like to express my gratitude to my dialy supervisors Harry Dijkhuis and Jarig Steringa at Gasunie. They were very committed to my project which made me enthusiastic for this research. Our lively discussions on the methods which are explored gave me inspiration for the mathematical analysis on the subjects.

Moreover, I want to thank the whole planning department for letting me feel I was one of the team. I received knowledge about the Gasunie business by attending meetings. Furthermore, I enjoyed the corporate retreats like dinners and drinks. The trip to the pressure regulation station and the trip to the control room were unique and very interesting. In time of sickness, kind messages and interest were very welcome and appreciated. I want to thank Jan Albert Laverman and Patrick Tel especially. They were my roommates at the Gasunie building. We have had nice conversations and they gave me other information of Gasunie, the gas world and other energy subjects. From other departments of Gasunie I got good responses too. I want to thank Marco Hoogwerf for his explanation of the MCA program and its methods and listening to my ideas.

Lastly, I want to thank Davie and my family for their support. In the time of this project I experienced again that I can always count on their help.

Finally, I hope this report will arouse interest in the gas planning and new research will arise from this report.

Sandra Maring
Groningen, January 2018

Contents

1
Introduction

1.1. Gasunie Transport Services

Gasunie Transport Services (GTS) is the gas Transmission System Operator (TSO) of the Netherlands. GTS is a regulated part of the holding Gasunie, to which also a TSO in Germany belongs. GTS supplies entry and exit capacity to the Dutch gas market. This is done through contracts with parties that supply or demand gas at the entry and exit points. These parties are called shippers, which are companies from the Netherlands, or other European countries, like Germany, the United Kingdom and Russia [16, 19]. Gasunie is responsible for a network of 15.000 kilometer of gas pipes and numerous stations [10], through which 1236 Tera Watt hour was transported in 2016 [12].

The entry and exit system of Gasunie differs in planning and optimisation from a system where capacity can be controlled by the operator [19]. The main planning issue is to check whether every realistic situation is feasible by the current network. First from the many possible entry and exit combinations, the most challenging transport situations must be found. Second, if a planned situation cannot be accommodated, projects arise to make the situation transportable. Possible solutions to such a problem are to add new assets to the system or to debottleneck the existing system.

The optimisation of this transport is a complex problem. Transport of capacity involves the presence of gas of parties at the entry and exit points. Knowledge of scope, policy and behaviour of all the parties is needed. Dealing with this problem can be dealt with by mathematical theories such as statistics, linear algebra and numerical analysis [19].

Gas transmission systems operate on an hourly timescale. Every shipper is required to notify their need of capacity within contractual limits. This notification is sent a couple of hours in advance for every hour of the day. Each year has 8760 hours and for each hour, billions of different transport situations exist. This implies that the planning of the gas transmission involves many possible scenarios. It is not desirable to analyse all possibilities, because it requires many computations, which is extremely time consuming. Fortunately, there are ways to reduce the amount of transport situations that have to be evaluated.

At Gasunie, reduction of the set of transport situations is done by calculating an approximation of the gas flow through the network and only considering the most severe situations. After this selection, the set of severe situations is further reduced by taking out similar ones. To define whether situations are similar, a measure is needed to compare the situations. Researchers of the planning department at Gasunie decided to use a certain parametrisation of a quadratic form distance for this comparison.

The goal of this project is to find the most suitable comparison method to obtain the minimal amount of tests to be checked to ensure the feasibility of the gas network. This calls for an analysis of different measures and different approaches.

1.2. Dutch gas network

In this report, only the High Pressure Grid (HTL) of the Dutch gas network is considered. The planning of the Intermediate Pressure Grid (RTL), involves a different method because of the lower pressures and smaller transport distances. There are three different types of natural gas in the Dutch gas network. H-gas has a high calorific value and is usually imported. L-gas is gas with a lower caloric value and is exported to customers in

Germany and Belgium. The lowest quality is called G-gas, or Groningen gas. This is the gas quality commonly used for the gas stoves and the CV boiler of Dutch households.

The Dutch gas network consist not only on entry points, exit points and pipelines, but also of compressor stations, blending stations (to blend different types of natural gas), pressure regulation-stations and nitrogen injections [4]. Compressor stations increase the pressure in the pipe such that there is enough pressure to transport the gas over the required distance. Blending stations are used for the blending of different gas qualities to G-gas. Nitrogen injection supports the process of converting H-gas into G-gas with nitrogen. Pressure regulation-stations are used for e.g. the transition from the HTL-network to the RTL-network; odorisation takes place there and it is made sure that the pressure is at the right level to safely transport gas into the smaller pipes. The odorisation is needed to detect natural gas (without it the gas is odourless).

An overview of the HTL-gas infrastructure of the Netherlands is given on page 4. The yellow lines are the pipelines which transport H-gas, the grey lines are the pipelines which transport G-gas. Furthermore, the stations described above can be seen in the figure, along with gas storage facilities.

1.3. Research project

The research assignment of this project is to find improvements to an already good and efficient method which computes whether the gas network will suffice for every scenario which can occur scenario-based planning methodology. There are many factors which play a role in finding and evaluating scenarios. Much time and memory is needed to compute all scenarios. The questions related to this problem are stated below.

1. What are important *properties* of the Dutch gas network to characterise the load of gas transport situations?

2. What is the current method for *generating* stress tests? Are there better methods for generating these stress tests? What are the (dis)similarities and what are the (dis)advantages of the methods?

3. What is the current method for *reducing* the generated set of stress tests? Are there better methods for reducing the set of stress tests? What are the (dis)similarities and what are the (dis)advantages of the methods?

4. What are the criteria for *similarity* of stress tests?

The third research question covers the main part of the research presented in this report.

In this report, example networks are used to clarify various methods, techniques and problems. The five example networks are depicted in Figures 1.2, 1.3, 1.4 and 1.5 (pages 5 and 6). Entry points are denoted by an N and the exit points by an X. At every entry and exit point bounds of the capacity are given: one lower bound (lb) and one upper bound (ub) are given by $[lb, ub]$. The pipelines have a length shown in blue.

In Figure 1.2 the *one pipeline network* is given. This is a simple network with only one pipeline. The network, is special in that is has an entry and an exit point at the same location. In Figure 1.3a an H-shaped network is shown: the *simple H-network*. This network is often used to show the effect of the two directions in which gas can flow through a pipe (pipeline between E and F). This network includes many symmetrical components. To avoid degeneration, the *advanced H-network* is introduced (Figure 1.3b), where extra exits and asymmetry is added. The fourth network is the *triangular network*, seen in Figure 1.4. This network has some of the characteristics of the previous networks, but with an addition of a loop structure. An approximation of the HTL part of the Dutch gas network is the *shopping cart network* in Figure 1.5 . This network has three loops, one of which is not really a loop. Besides loops, a new type of component is visible: gas storages. For the simple networks there are no bounds on the pipelines, but there exists lower and upper bounds on the flows through the pipelines in the shopping cart network.

1.4. Outline report

The following chapters answer the research questions as in the outline given below.

Chapter 2: Gas transport

Safety and supply of natural gas is ensured by GTS. Transport situations are being calculated while making specific assumptions. The effort to transport gas can be defined as the power to transport gas through pipeline segments. This quantity in turn can be approximated by the transport moment. By maximisation of this transport moment severe situations are be generated and these situations should be checked whether they fail for the network to be examined.

Chapter 3: Generation of stress tests

Severe situations (stress tests) can be found under certain conditions and optimisations. In Section 3.1 the current method of generating stress tests is explained and illustrated with an example. As an alternative, another approach of representing the stress tests is introduced and how to find them. This is explained in Section 3.2. The two stress test generation methods are discussed of the example networks given in previous section (Dutch gas network, Figures 1.2 - 1.5).

Chapter 4: Reduction of the generated set of stress tests

The set of stress tests has to be reduced to make the number of computations manageable. First, in Section 4.1 the current reduction method is considered. A distance function with a specific parametrisation (the quadratic form distance) is used to reduce the set. The distance function is illustrated by an example and it is investigated whether this function is a metric distance function. Second, the set of stress tests will be shown in a so-called “flow-representation”. The set of stress tests with the other representation should also be reduced. Because these stress tests have been established differently, other norms can be used for this problem. These norms are discussed in Section 4.2. An adapted $\mathscr{L}_p$-norm is found which fulfil the requirements for the distance between scenarios. Some additional reduction criteria are considered in Section 4.3 (e.g. pressure drop and diameter of a pipeline). The further reduction process once the best norms are found is given in Section 4.4.

Chapter 5: Results of the reduction methods

The reduction methods of previous chapters will be tested and the different techniques will be compared in this chapter. Similarities and differences between the methods are examined. First, the method of comparison will be explained. The criteria of similarity will be shown in Section 5.1. In Section 5.2 the comparison between the reduction techniques is done. The impact of other properties will be discussed in Section 5.3.

Chapter 6: Conclusions

The results of the research questions are concluded in this chapter. Also, some discussion on the research and future work is mentioned.

Figure 1.1: Dutch HTL-gas network in 2014

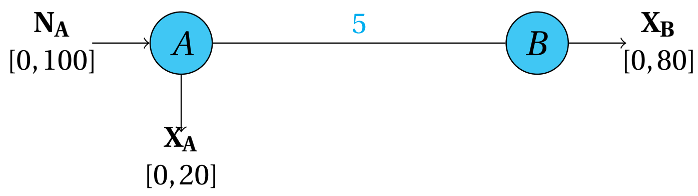
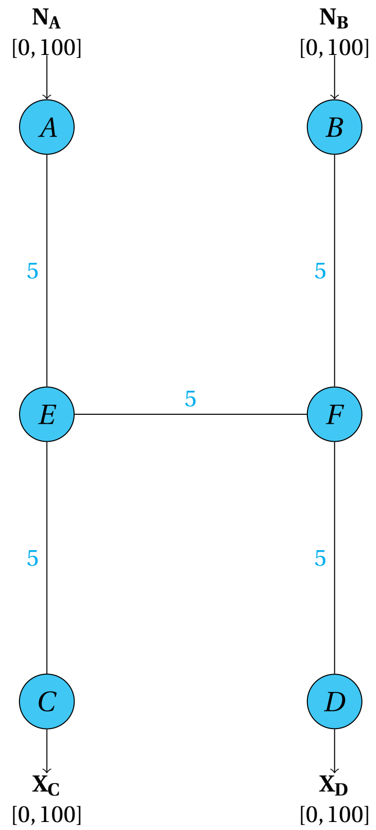
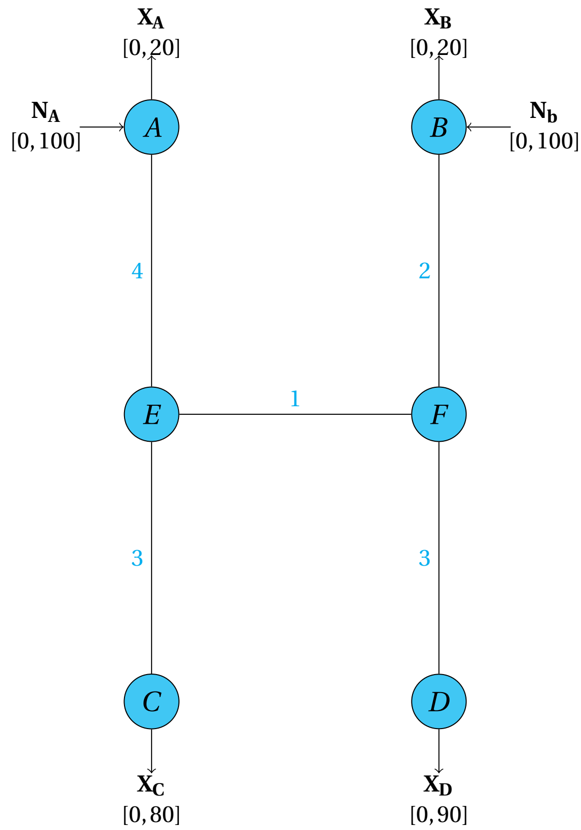


Figure 1.2: One pipeline network



(a) Simple H-network



(b) Advanced H-network

Figure 1.3: Two H-networks

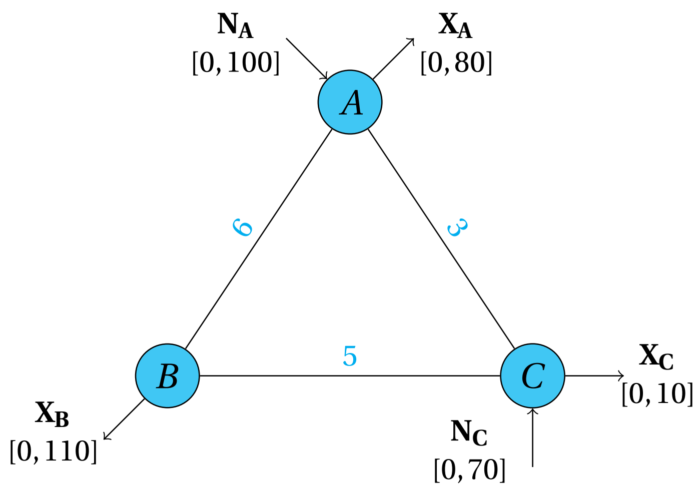


Figure 1.4: Triangular network

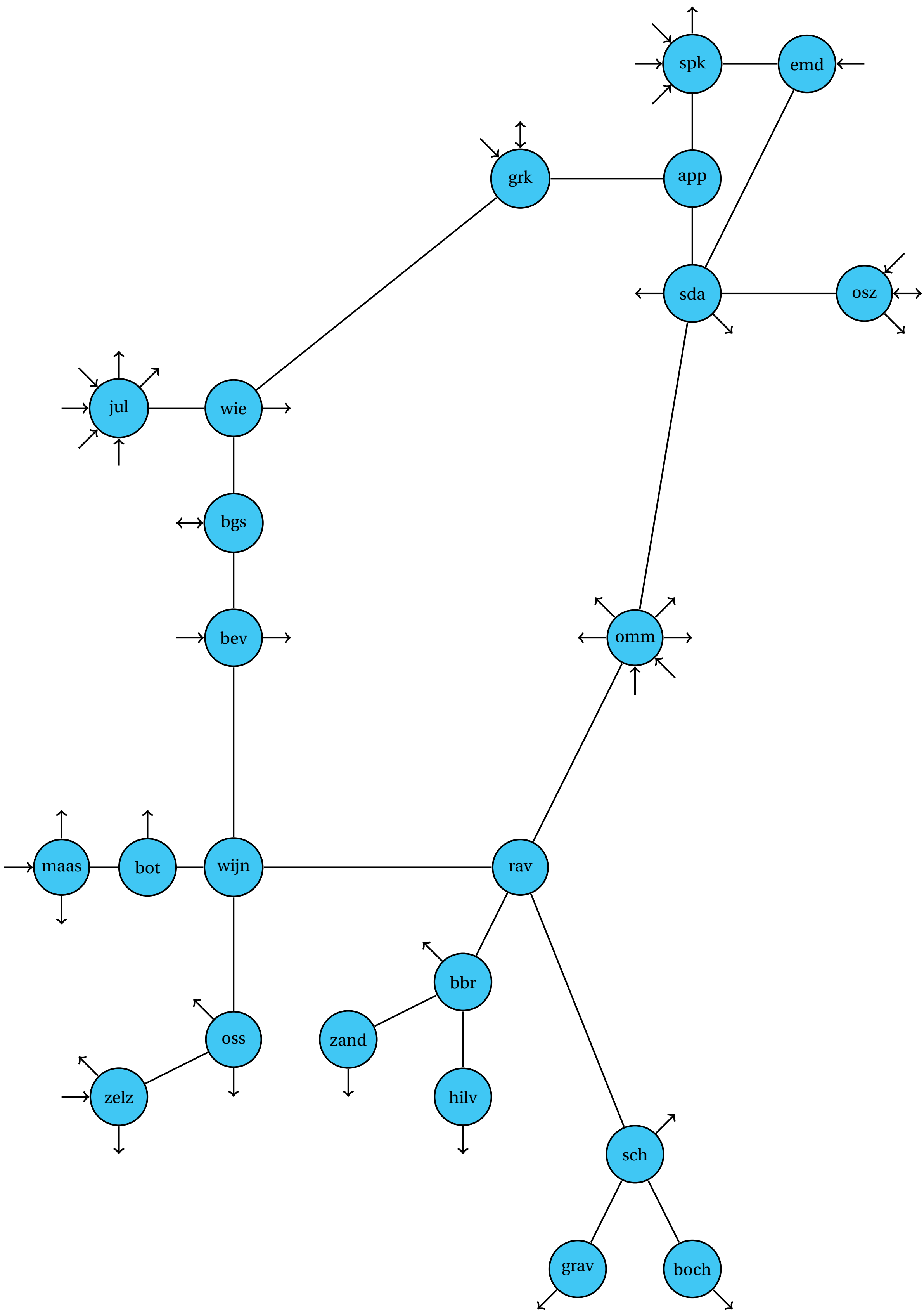


Figure 1.5: Shopping cart network

2

Gas transport

Gasunie is required by law to ensure safety and supply of natural gas in the HTL and RTL networks of the Netherlands. Therefore calculations must be made to ensure these regulations. This is done at the planning department of GTS. These calculations answer questions like: is the current gas network sufficient for its task, or is it necessary that pipelines have to be added or updated. Realistic situations in a gas transmission system where the entry and exit capacities are balanced are called transport situations. Balance means that the sum of the entry capacities equals the sum of the exit capacities, which is also known as the flow conservation law. The transportation of gas depends on pipe lengths and diameter, amount of flow, pressure requirements, compressor stations, temperature and more.

An example of a transport situation is shown in Figure 2.1 on page 8. This network, the shopping cart network, is an approximation of the H-gas part of the HTL gas network in the Netherlands. The flows in this example are fictional. It is seen in this transport situation that the incoming flow equals the outgoing flow. There is a flow of 1000 going from North to South and there is a flow from the North to the West of 750.

Realistic situations are situations that can happen in practice. An assumption underlying these situations is that shippers always adhere to their contracts with Gasunie. If a shipper has a contract for entry point A, a maximum and a minimum are agreed on the capacity of the gas transported into the network at A. These bounds of entry and exit points are known to Gasunie from contractual information and can be further restricted to a certain extent. A list of all the assumptions made in this report can be found in Appendix A on page 49.

It is sufficient to check the network for only the most severe situations. If these situations are feasible for the network, the non-severe situations will also be feasible [18]. The severity of the situation is given by a quantity called transport moment, which approximates the transport load. The transport load is the power needed to get a gas quantity from one location to another through a pipeline. The equation for this power is given by Equation (2.1):

$$P = Q \cdot \Delta p \tag{2.1}$$

P is the power in Watt, Q is the flow in m^3/s and Δp is the pressure drop in N/m^2. The pressure drop is not linear in a long pipeline, but in parts of the pipeline the pressure drop can be approximated by pipeline length. It is assumed that transport distance and amount of flow through the network are the most important quantities for the transport moment.

In the next chapter, two equations are given which approximate the transport moment from different perspectives.

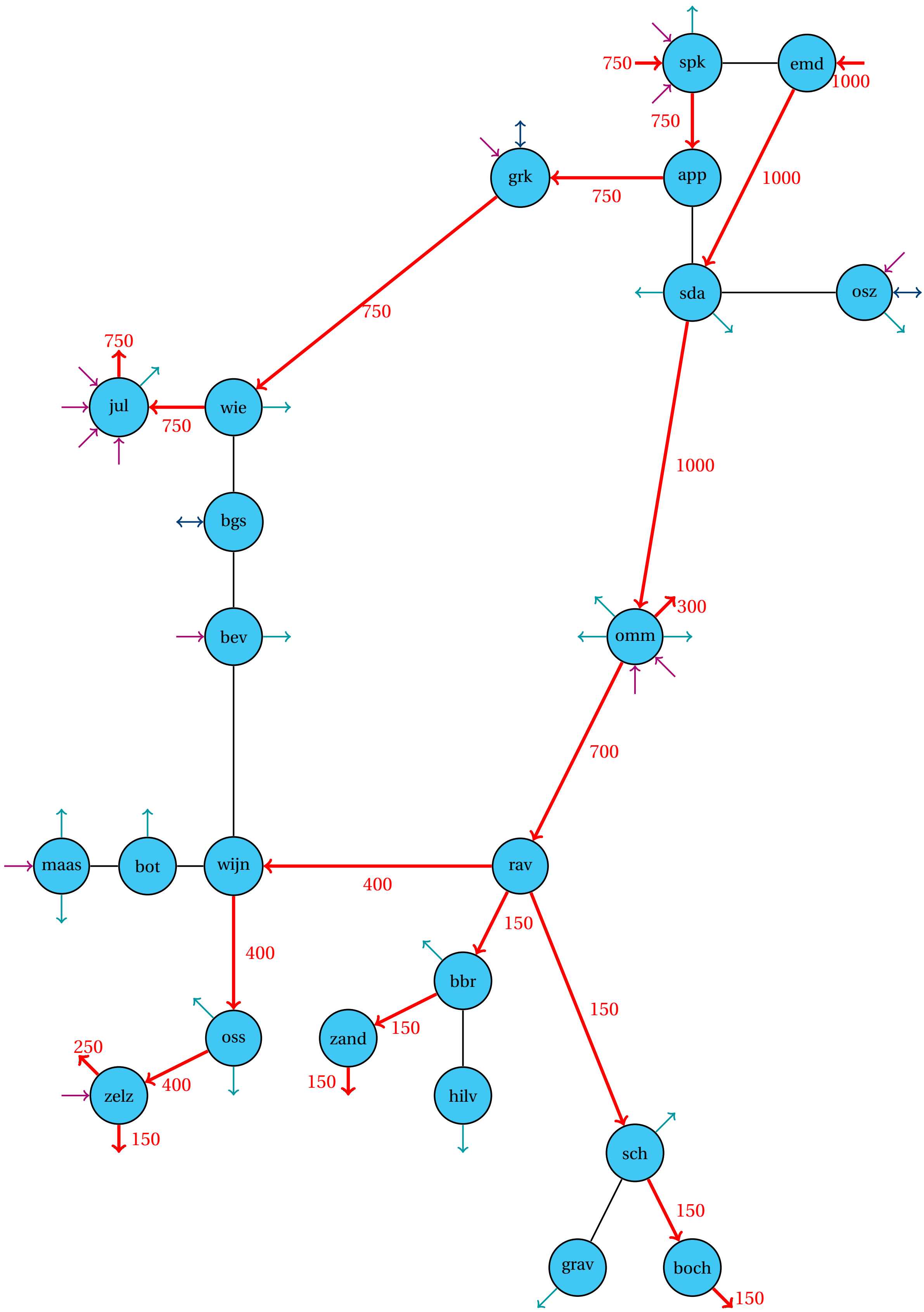


Figure 2.1: Example of a transport situation with fictional flows

3

Generation of stress tests

To ensure the safety of the gas network, hydraulic calculations on transport situations have to be made. However, these calculations are expensive. Since it is too expensive to compute all hydraulic calculations, only the situations that are more likely to fail are used. These balanced situations are severe transport situations, which are called stress tests [19]. Stress tests are the most severe realistic transport situations. To define the severeness of a transport situation, the transport moment is used as is described in previous chapter. It is assumed that the transport distance in relation to the gas flow over that distance are the most important quantities for the transport load. When using these two quantities, the transport moment is approximated. This approximation is needed, because the real hydraulic calculations are too expensive as well. Thus the equations for the transport moment are defined by the transport distance and the gas flow.

Transport moment based on capacities through end points

Two equations are used to define the approximated transport moment. The first equation is from the shippers point of view: the capacities on entry and exit points (end points). Shippers are only interested in the entry and exit capacities, because they book their capacities on those points and GTS is responsible for the transport in between. Equation (3.1) is the corresponding equation which is not based on the resulting flows, but on the capacities of the end points and the mutual distances. The exact flow pattern is not needed for the calculations of this end point representation. This transport moment is relative to a certain anchor point (τ). It is only necessary to take an entry or exit points for τ in practice, because this results in the most severe transport situations: the stress tests [18].

Let I be the amount of exit points and J is the amount of entry points. Exit point $i \in [1, I]$ is defined by X_i and entry point $j \in [1, J]$ is N_j, the capacity on the end points is given by $c(\cdot)$ and the distance between end points is given by $d(\cdot,\cdot)$. The distance function d is the shortest path from one point to another along the gas pipelines.

$$T(\tau) = \sum_{i=1}^{I} c(X_i) d(X_i, \tau) - \sum_{j=1}^{J} c(N_j) d(N_j, \tau) \tag{3.1}$$

Transport moment based on flows on pipelines

The second equation is an explicit equation which relates the pipeline length directly to the flow through that pipeline. The total transport moment T is calculated by the sum of the transport moment per pipeline segment. This implies if the transport distance is twice as long, the transport moment will be twice as large, and the same holds for the amount of flow. Therefore the transport moments are calculated by the product of the flow through the pipeline segment (f) and the length of that segment (L). The total transport moment formula is seen in Equation (3.2), where E is the set of pipeline segments. This equation approaches the physical reality more than the previous definition of the transport moment, because this equation is the linearisation of the power equation, which is shown in the previous chapter.

$$T_{\text{total}} = \sum_{e \in E} T(e) = \sum_{e \in E} f(e) \cdot L(e) \tag{3.2}$$

$f(e)$ is the flow through pipe $e \in E$ and $L(e)$ is the length of pipe $e \in E$.

3.1. Stress tests based on end points

The current method to generate stress tests is based on the end points. The severe situations are found by choosing the capacities on the entry and exit points such that the transport moment is maximised under the conditions of the conservation law and the capacity boundaries [18]. It is unknown at this moment what happens inside the network for the approximated transport moment, so it is only necessary to have information of the capacity of the end points and the distances between end points for this approach. All situations are covered by successively taking all entry and exit points as anchor point τ for the approximation of the transport moment. In this way all directions of the flow will be taken into account.

Figure 3.1 shows a network with a feasible transport situation. Here, the entry and exit points are given and the internal network is like a black box, because its flow pattern is unknown for finding the stress tests. Assuming this is a realistic and severe situation for this network, the vector of this transport situation stated below is a stress test.

$$\text{The capacities on the points} \begin{pmatrix} N_A \\ N_B \\ N_D \\ N_E \\ X_B \\ X_C \\ X_D \\ X_F \\ X_G \\ X_H \end{pmatrix} \text{are} \begin{pmatrix} 400 \\ 200 \\ 100 \\ 300 \\ -50 \\ -150 \\ -100 \\ -200 \\ -300 \\ -200 \end{pmatrix}, \text{where the exits are given by a minus sign.}$$

The flows at the exit points are negative, such that the flow direction is preserved in the vector. It can be seen that this stress test is balanced, because the sum of the entry capacities equals the sum of the exit capacities.

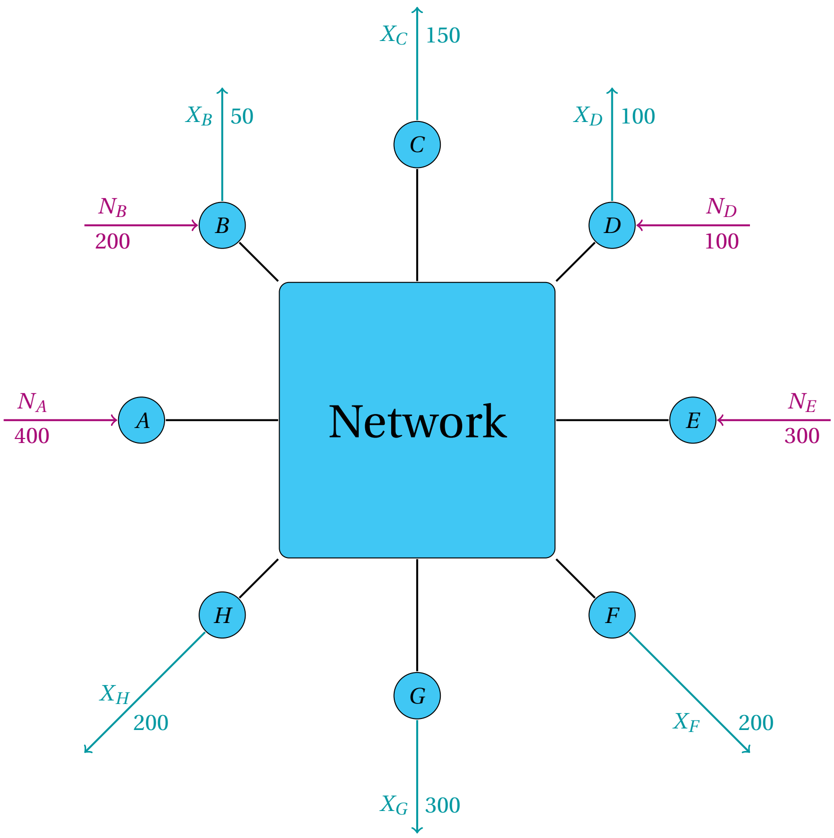


Figure 3.1: Example of a feasible transport situation

3.1.1. Stress test algorithm

The algorithm to find the stress tests is given in Algorithm 3.1, which is the algorithm stated in the article by Steringa, et al [18]. The algorithm maximises the transport moment of Equation (3.1) by iterating over all anchor points and adding capacity at the closest entry point to the anchor point. The same is done for the exit capacity furthest away from the anchor point such that balance is preserved. By putting the exit capacity furthest away, the distance is maximised and by adding maximal capacity, the transport moment is maximised. Adding more capacity to the network when the maximum is already reached, the transport moment will decrease and the algorithm terminates. An example for finding the stress tests by this algorithm is given in the next paragraph.

Algorithm 3.1 The Stress Test algorithm

```
for all anchor points do
    while transport moment does not decrease and maintaining balance do
        Add entry capacity at or, in case of maximum capacity used, near the anchor point
        Add exit capacity furthest away from the anchor point at which the bounds will not exceed
    end while
    Store resulting entry-exit combination as the stress test for the chosen transport direction
end for
```

3.1.2. Example: Advanced H-network

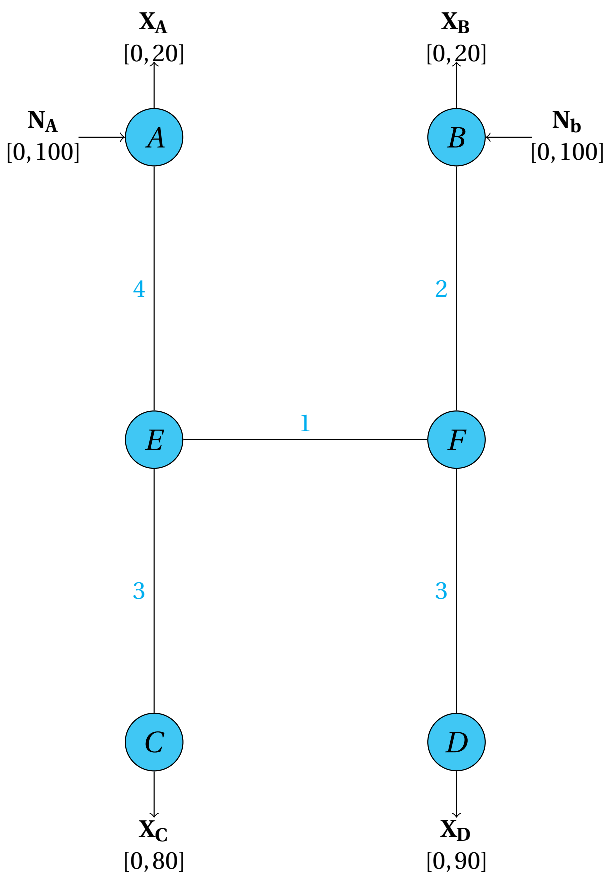


Figure 3.2: Example generating stress tests

The network considered for this example is the advanced H-network, is illustrated in Figure 3.2. There are two entry points (N_A, N_B), four exit points (X_A, X_B, X_C, X_D) and five pipelines. The length of each pipeline is given in the figure next to the relevant pipeline and the gas can flow in both directions. The lower bound and the upper bound of each entry and exit is given at these points by $[lb, ub]$, where lb is the lower bound and

ub is the upper bound.

Often, the simple H-network is used for illustrating the algorithm. The simple network has an almost identical structure as the advanced version, but it has no exits at points A and B, all pipe segments have the same length and the bounds on the entry and exit points are symmetric. In this example, the pipe lengths are not chosen symmetric, so that the calculations between different anchor points are more clear. The exit points at A and B are put into the network to show that some scenarios are similar, but not equal. The iterations of the algorithm are described below. The iterations of the first anchor point are explained in more detail than the other anchor points. If a choice has to be made between entries/ exits, the first entry/ exit in alphabetical order is chosen.

Anchor point N_A

The illustration of the iterations for this anchor point is given in Figure 3.3. Although the flows through the pipelines are unknown in the algorithm, in this network they can be seen easily and it gives a better view on the transport situation when flow is indicated. This flow is visualised in the figure in red.

- The entry closest to anchor point N_A is N_A itself and exit furthest away to the anchor point is X_D. The minimum of the upper bounds on N_A and X_D that can be used is $\min\{100, 90\} = 90$. Therefore the transport moment on the capacities becomes:

$$\begin{aligned} T(N_A) &= \sum_{j=\{A,B,C,D\}} c(X_j)d(N_A, X_j) - \sum_{i=\{A,B\}} c(N_i)d(N_A, N_i) \\ &= 90 \cdot 8 - 90 \cdot 0 \\ &= 720 \end{aligned}$$

  The resulting flow addition is seen in Figure 3.3a.

- More capacity at the entry and exit points can be added until the transport moment is decreasing. The entry point closest with capacity left is N_A and the exit point furthest away with capacity is X_B. X_C can be chosen as well, but there is made a choice to have alphabetical priority. The minimum of the maximal capacity that can be chosen on the entry and exit is 10. If this capacity is added, the transport moment becomes:

$$\begin{aligned} T(N_A) &= \sum_{j=\{A,B,C,D\}} c(X_j)d(N_A, X_j) - \sum_{i=\{A,B\}} c(N_i)d(N_A, N_i) \\ &= 10 \cdot 7 + 90 \cdot 8 - 100 \cdot 0 \\ &= 790 \end{aligned}$$

  The scenario has become more severe, which can be seen in Figure 3.3b.

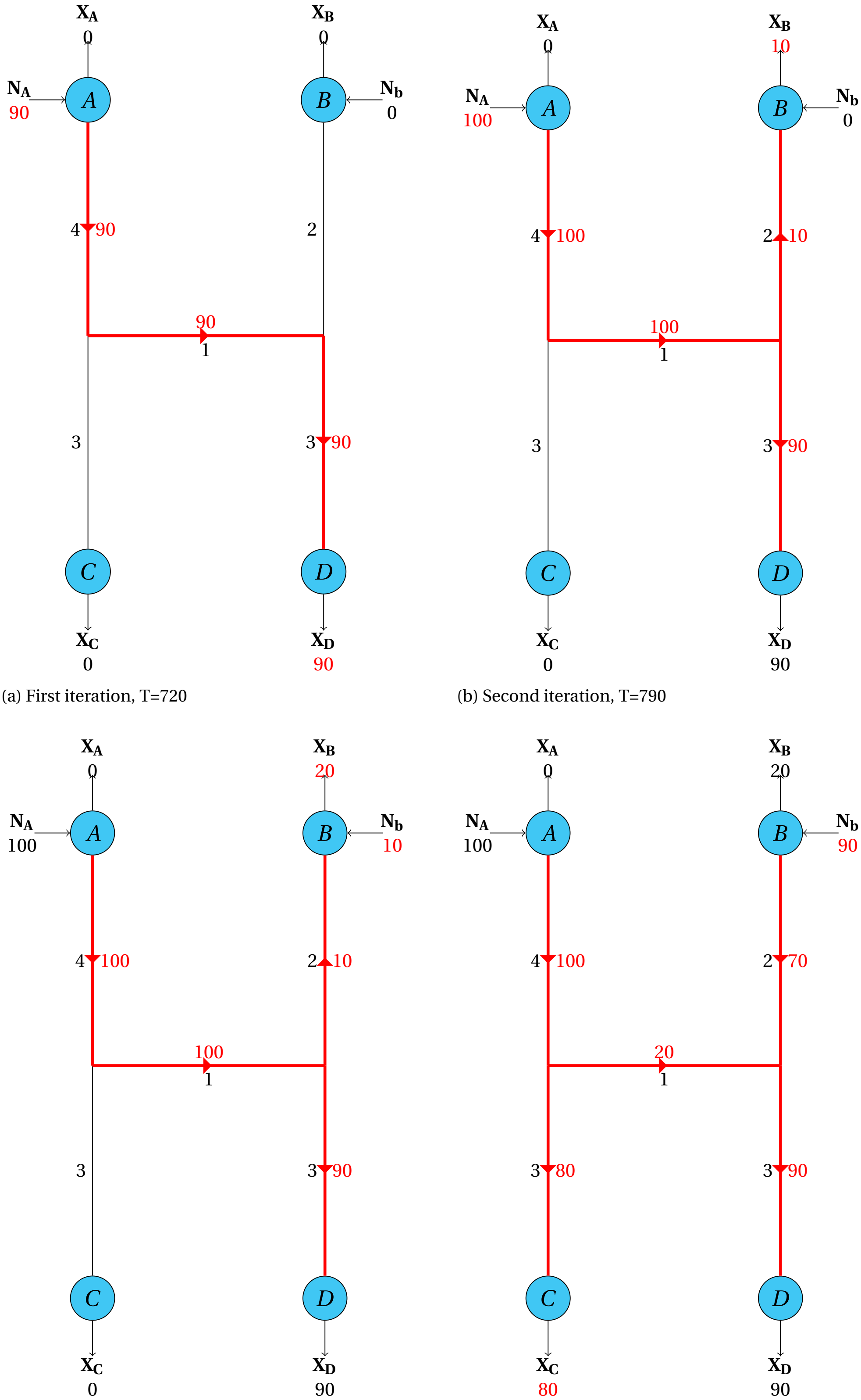
- The next closest entry point is N_B and exit point X_B has some capacity left. The capacity added to these points is 10. By adding this capacity, the transport moment is:

$$\begin{aligned} T(N_A) &= \sum_{j=\{A,B,C,D\}} c(X_j)d(N_A, X_j) - \sum_{i=\{A,B\}} c(N_i)d(N_A, N_i) \\ &= 20 \cdot 7 + 90 \cdot 8 - 100 \cdot 0 - 10 \cdot 7 \\ &= 790 \end{aligned}$$

  As is seen in Figure 3.3c, the transport moment does not increase, because the transport load does not change.

- At exit point X_B all possible capacity is being used, so the next exit, which is far away from the anchor point and has capacity left is X_C. The entry point is N_B. The maximal capacity which can be added on these points is 80. See Figure 3.3d for the result of this iteration.

$$\begin{aligned} T(N_A) &= \sum_{j=\{A,B,C,D\}} c(X_j)d(N_A, X_j) - \sum_{i=\{A,B\}} c(N_i)d(N_A, N_i) \\ &= 20 \cdot 7 + 90 \cdot 8 + 80 \cdot 7 - 100 \cdot 0 - 90 \cdot 7 \\ &= 790 \end{aligned}$$



(a) First iteration, T=720

(b) Second iteration, T=790

(c) Third iteration, T=790

(d) Fourth iteration, T=790

Figure 3.3: Finding stress test for anchor point N_A

- This fifth iteration will be the last iteration.

$$\begin{aligned} T(N_A) &= \sum_{j=\{A,B,C,D\}} c(X_j)d(N_A,X_j) - \sum_{i=\{A,B\}} c(N_i)d(N_A,N_i) \\ &= 10\cdot 0 + 20\cdot 7 + 90\cdot 8 + 80\cdot 7 - 100\cdot 0 - 100\cdot 7 \\ &= 720 \end{aligned}$$

This transport moment is decreased with respect to the previous transport moment. So the addition of flow on exit X_A and entry N_B will stop the while-loop and the stress test in Equation (3.3) is found on the anchor point at N_A. In this simple example, the flow through the network can easily be generated when knowing the capacities on the entry and exit points. The fifth iteration is seen in Figure 3.4 and it is seen that this is less severe than the situation at iteration 4. Iterations two, three and four produce the same transport moment. This is not likely to happen in more complex networks and it is indifferent which one is chosen. For algorithmic simplicity, the last scenario before the transport moment is decreased, is stored. When another order of entry and exit points are chosen, an other order of the iterations is found, so another stress test with the same transport moment is found.

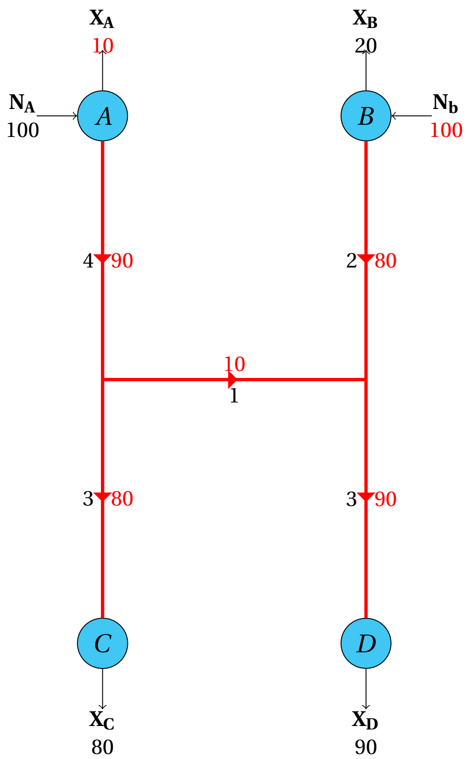


Figure 3.4: Fifth iteration, T=720

$$\begin{pmatrix} 100 \\ 90 \\ -0 \\ -20 \\ -80 \\ -90 \end{pmatrix} \tag{3.3}$$

This stress test represents the capacity on the entry and exit points. It is seen that the stress test is balanced, because the sum of the capacities on the entries is 190 and the same capacity sum is found for the exit points.

The calculations of the other anchor points are given in the next few paragraphs. In each iteration, the closest entry, the furthest exit, the added capacity is given. Then, in each iteration the calculations to the transport moment is given.

Anchor point N_B

- Entry N_B, exit X_A, capacity that can be added is 20.

$$\begin{aligned} T(N_B) &= 20 \cdot 7 - 20 \cdot 0 \\ &= 140 \end{aligned}$$

- Entry N_B, exit X_C, capacity that can be added is 80.

$$\begin{aligned} T(N_B) &= 20 \cdot 7 + 80 \cdot 6 - 100 \cdot 0 \\ &= 620 \end{aligned}$$

- Entry N_A, exit X_D, capacity that can be added is 90.

$$\begin{aligned} T(N_B) &= 20 \cdot 7 + 80 \cdot 6 + 90 \cdot 5 - 90 \cdot 7 - 100 \cdot 0 \\ &= 440 \end{aligned}$$

This addition of capacity causes a smaller transport moment, therefore the last addition is not included for the stress test. It is seen that 620 is the maximum transport moment and therefore the most severe scenario, because the addition in the third iteration of 90 causes a lower transport moment. Then $T(N_B) = 440$. In the third iteration, the additional flow can also be 1, but that will also decrease the transport moment, because then the transport moment becomes: $T(N_B) = 618$. The final stress test for this anchor point is:

$$\begin{pmatrix} 0 \\ 100 \\ -20 \\ -0 \\ -80 \\ -0 \end{pmatrix}$$

Anchor points X_A and X_B

The stress tests for the exit points X_A and X_B is the same as for N_A and N_B respectively, because they are located at the same point and therefore the same exit and entry points will be chosen. So for every point where the distance is zero, these cases can be considered as the same. Algorithm 3.1 will again do the same iterations as in N_A and N_B.

Anchor point X_C

- Entry N_B, exit X_A, capacity that can be added is 20.

$$\begin{aligned} T(X_C) &= 20 \cdot 7 - 20 \cdot 6 \\ &= 20 \end{aligned}$$

- Entry N_B, exit X_D, capacity that can be added is 80.

$$\begin{aligned} T(X_C) &= 20 \cdot 7 + 80 \cdot 7 - 100 \cdot 6 \\ &= 100 \end{aligned}$$

- Entry N_A, exit X_D, capacity that can be added is 10.

$$\begin{aligned} T(X_C) &= 20 \cdot 7 + 90 \cdot 7 - 10 \cdot 7 - 100 \cdot 6 \\ &= 100 \end{aligned}$$

- Entry N_A, exit X_B, capacity that can be added is 20.

$$\begin{aligned} T(X_C) &= 20 \cdot 7 + 20 \cdot 6 + 90 \cdot 7 - 30 \cdot 7 - 100 \cdot 6 \\ &= 80 \end{aligned}$$

This capacity addition causes a smaller transport moment, therefore the last addition is not included for the stress test.

$$\begin{pmatrix} 10 \\ 100 \\ -20 \\ -0 \\ -0 \\ -90 \end{pmatrix}$$

Anchor point X_D

- Entry N_B, exit X_A, capacity that can be added is 20.

$$\begin{aligned} T(X_D) &= 20 \cdot 8 - 20 \cdot 5 \\ &= 60 \end{aligned}$$

- Entry N_B, exit X_C, capacity that can be added is 80.

$$\begin{aligned} T(X_D) &= 20 \cdot 8 + 80 \cdot 7 - 100 \cdot 5 \\ &= 220 \end{aligned}$$

- Entry N_A, exit X_B, capacity that can be added is 20.

$$\begin{aligned} T(X_D) &= 20 \cdot 8 + 20 \cdot 5 + 80 \cdot 7 - 20 \cdot 8 - 100 \cdot 5 \\ &= 160 \end{aligned}$$

This capacity addition causes a smaller transport moment, therefore the last addition is not included for the stress test.

$$\begin{pmatrix} 0 \\ 100 \\ -20 \\ -0 \\ -80 \\ -0 \end{pmatrix}$$

Finally, different stress tests found, which are all balanced by definition of the algorithm, see Equation (3.4). The stress tests of anchor point B and D are the same, because the order of nearest entry is the same at each anchor point. And the order of exit points farthest away is almost the same, it holds for the first two exit points: X_A and X_C. After using these two exit points in the algorithm, the transport moment of the capacities is decreased. Therefore, the stress tests of the anchor points B and D are the same.

$$s_A = \begin{pmatrix} 100 \\ 90 \\ -0 \\ -20 \\ -80 \\ -90 \end{pmatrix}, \quad s_B = \begin{pmatrix} 0 \\ 100 \\ -20 \\ -0 \\ -80 \\ -0 \end{pmatrix}, \quad s_C = \begin{pmatrix} 10 \\ 100 \\ -20 \\ -0 \\ -0 \\ -90 \end{pmatrix}, \quad s_D = \begin{pmatrix} 0 \\ 100 \\ -20 \\ -0 \\ -80 \\ -0 \end{pmatrix} \tag{3.4}$$

A comparison of the two different approximations of the transport moment is made by the stress tests in Table 3.1. The equations of the two transport moments are given in Equation (3.1) and in Equation (3.2) on page 9. The flow of each unique stress test is illustrated in Figure 3.6 on page 18.

Anchor point	T based on capacities end points	T based on flows on pipelines
N_A, X_A	790	1070
N_B, X_B	620	620
X_C	100	520
X_D	220	620

Table 3.1: Two approximations of the transport moment of stress tests

It is seen in the values of the table above and Figure 3.5 on page 17 that there is no clear correlation in these two transport moments. There are two stress tests (of anchor points B and D) which have the same transport moment based on the flow, but have a different transport moment based on capacities. That is because the transport moment based on capacities on end points is dependent on the corresponding anchor point.

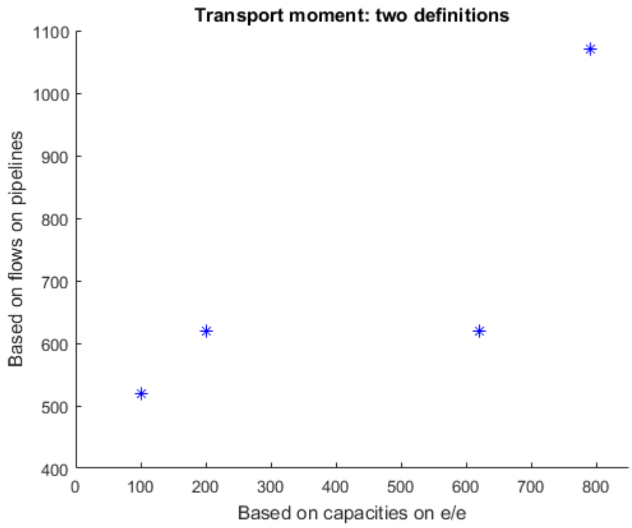


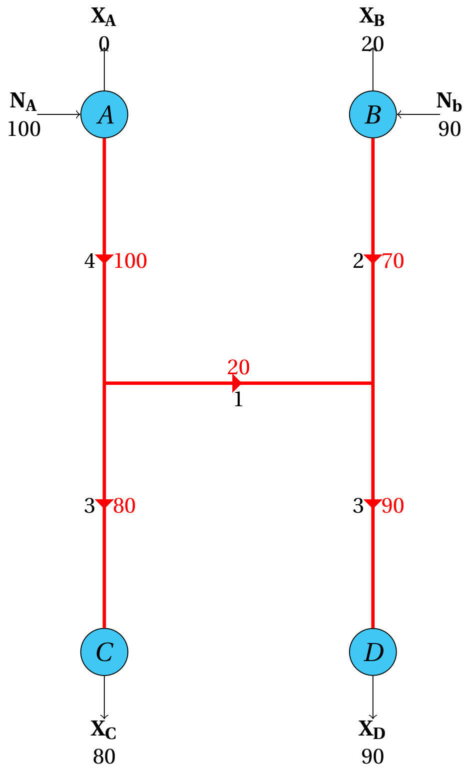
Figure 3.5: Two approximations of the transport moment on different stress tests

3.2. Stress tests based on flows through pipelines

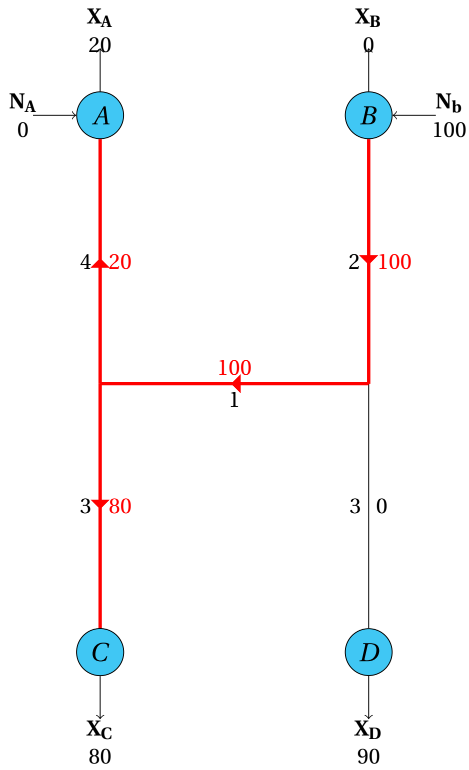
In the previous section, the stress tests are based on the capacities on the entry and exit points. In that case, the representation includes only a few components so these are simple calculations. However, the load of the gas transport may alternatively be based on the flow through the pipelines. When describing the problem of generating stress tests based on flow through pipelines, the amount of physical information in the stress tests is increased. The contractual bounds on the entries and exits are still needed for the computations and the structure in the network is added to the known data. Apart from exit and entry points, there are points in between as well which are connected to the network and the directions of flow are defined. The flow which is generated is an approximation of the real flow, because the real flow requires calculations with more parameters such as the pressure in the pipeline. In this chapter two methods of generating stress tests based on flow through pipelines are given. The first method generates these stress tests directly from the network information and the second one obtains the stress tests by transforming the stress tests based on end points. Both methods are based on pipelines, so the transport moment of Equation (3.2) on page 9 is used for both approaches.

3.2.1. Generation directly from the network to stress tests

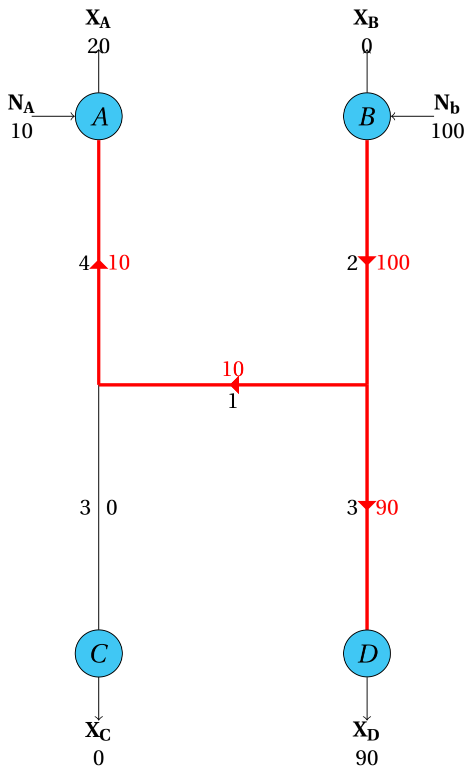
Using the network information, severe transport situations are obtained by solving the maximisation problem of the transport moment with certain conditions. These conditions are bounds on the entry and exit points, flow conservation on the end points (what comes in, goes out) and the flow conservation on the inner network points. In mathematics, this problem can be displayed into a LP, because it has a linear optimisation function and linear constraints [7]. The standard form of a linear program is given in Equation (3.5). $c \in \mathbb{R}^m$ is the cost vector of the minimisation function, $x \in \mathbb{R}^m$ is the vector which is to be optimised. The conditions of $Ax \leq b$ are the conditions with an inequality, where $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$, where n is the amount of conditions with an inequality. The same holds for A_{eq} and b_{eq}, which represent the equality conditions. Furthermore, lb, ub $\in \mathbb{R}^m$ are the lower bound and the upper bound respectively.



(a) Stress test, anchor point A



(b) Stress test, anchor point B



(c) Stress test, anchor point C

Figure 3.6: Flow on the stress tests

$$\begin{aligned} \min \quad & c^T x && \text{(3.5a)} \\ \text{subject to} \quad & Ax \leq b && \text{(3.5b)} \\ & A_{eq} x = b_{eq} && \text{(3.5c)} \\ & lb \leq x \leq ub && \text{(3.5d)} \end{aligned}$$

The set of solutions to the linear constraints is called a polyhedron [8]. If the polyhedron is bounded, then it is called a polytope. Thus a polyhedron $P \subseteq \mathbb{R}^n$ is a polytope if there exists lower and upper bounds $lb, ub \in \mathbb{R}^n$ such that $lb \leq x \leq ub$ for all $x \in P$.

Optimisation problem displayed as graph

The linear program leads to a problem which is similar as the maximal flow through a graph. This looks like the gas network, but some adjustments have to be done. A graph from an LP is constructed by a start node and an final node with internal nodes. The flow goes from the start node to the final node along the network edges and internal nodes. However, in the original gas networks, there are multiple start and final nodes which are the entry and exit points. Therefore the end points are replaced by a start and a final node.

So there are two nodes added to the problem which are the representatives of the entry and exit points to have one node from which all gas flows (N) and one node to which all gas gas flows (X). The flow from N (representing all entry points) must equal the incoming flow at X (representing all exit points). Below is a full description of this representation. Afterwards the transformations of the example networks are given.

1. Take the nodes from the network and make edges between nodes which are connected.
2. The weight on each edge is the distance between the corresponding nodes.
3. The bounds of the flow is taken $[lb, ub] = [-\infty, \infty]$, because now it assumed that through each pipeline the flow can go both ways and there are no restrictions on the amount of flow (this bound will not be shown in the illustrations).
4. Connect each node which has an entry with N and put the capacity bounds of the corresponding entry point on the edge. These are the bounds of the flow through those edges.
5. Connect each node which has an exit point the same way as for entry points at X.
6. The edges connected to N and X have weight zero (this is not shown in the illustrations).

The mathematical definitions describing the network are given below.

- $V = \{N, X, v_1, \ldots, v_k\}$ is the set of nodes in the network.
- $E = \{(v_1, v_2)$: there is a direct pipeline between node v_1 and $v_2\}$ is the set of edges in the network.
- $G = (V, E)$ is the graph that describes the network by nodes and edges.
- $f : E \to \mathbb{R}$ is the function that describes the amount of flow through an edge or a set of edges.
- $lb, ub : E \to \mathbb{R}$ are the lower and upper bounds (respectively) on the amount of flow through the pipelines.
- $L : E \to \mathbb{R}$ are the weights on edges. In the gas network problem, the weights are the corresponding lengths of the pipelines.
- For $A \subseteq V$ is $\delta^{in}(A) = \{e \in E$: e has only one end in A where the flow towards A is positive$\}$ [5]
- For $A \subseteq V$ is $\delta^{out}(A) = \{e \in E$: e has only one end in A where the flow leaving A is positive$\}$

With these definitions describing the network, the following constraints are found to model the gas network:

- $f(\delta^{out}(N)) = f(\delta^{in}(X))$. This means that the total amount of flow from the entries is equal to the total amount of flow to the exits. This ensures the flow conservation in the network.

- $f(\delta^{out}(v)) = f(\delta^{in}(v))$ for every $v \in V \backslash \{N, X\}$. The incoming flow must equal the outgoing flow of an internal node $(v_1, \ldots, v_k)$.
- $f(\delta^{in}(N)) = 0$. There is no flow towards entry points in the network.
- $f(\delta^{out}(X)) = 0$. There is no flow from exit points in the network.

Finding the severe situations by the description from above, a maximisation problem is constructed by several corresponding constraints. If these constraints and the maximisation function are linear, then problems like this can be handled by linear programming.

Finding severe transport situations

Now that the description of the new network representation is given, the severe transport situations should be found. In Equations (3.6), the desired optimisation problem is defined. This is not a linear program, because the maximisation function is $\text{fun}(f) = L^T|f|$, so a NLP algorithm is used. Yet, this non-linearity does not have a big impact on the running time, because this function is only piecewise linear. From the first constraint follows the flow conservation at the entries and exits.

$$\begin{aligned} &\text{maximise} && L \cdot |f| && (3.6a)\\ &\text{subject to} && f(\delta^{out}(v)) - f(\delta^{in}(v)) = 0, \qquad v \in V\ \{N, X\} && (3.6b)\\ &&& f \in [lb, ub] && (3.6c) \end{aligned}$$

In the system above f is the flow vector for each edge in the new representation. The directions of positive flow are stored in the adjacency matrix A_{adj}, with the amount of internal nodes as rows and the amount of columns are the number of edges. The entries of this matrix are defined as:

$$a_{ij} = \begin{cases} -1 & \text{edge } j \text{ is defined in direction from node } i \\ +1 & \text{edge } j \text{ is defined in direction towards node } i \\ 0 & \text{edge } j \text{ has no direct connection to node } i. \end{cases}$$

The columns of the edges which are connected to the entry node N have one +1 in that column and for the exit node X there is one −1. There are a '−1' and a '−1' in the column for internal edges, because both adjacent nodes are internal nodes (not N or X). The constraint $A_{adj} \cdot f = 0$ ensures the flow conservation in each node. That is because on row i of A there is a −1 for each outgoing edge and a +1 for each incoming edge for node i. So the summation of each row multiplied with the flow on each edge must be zero. Then there is equally incoming flow and outgoing flow in each internal node.

An example is given below. This is the network matrix belonging to the one pipeline network. The graph is given in Figure 3.7b, which is explained in the paragraph above this figure. The network matrix of this network is given in Equation (3.7). If there is a flow is of 20 from entry A to exit A and a flow from entry A to exit B of 80, there is balance (see Equation (3.8)). But when there is more incoming flow than outgoing flow, there is no balance and the result of the constraint will not be zero (see Equation (3.9)).

$$A_{adj} = \begin{matrix} \\ A \\ B \end{matrix} \begin{matrix} NA & AX & BX & AB \\ \left(\begin{matrix} +1 \\ 0 \end{matrix} \right. & \begin{matrix} -1 \\ 0 \end{matrix} & \begin{matrix} 0 \\ -1 \end{matrix} & \left. \begin{matrix} -1 \\ +1 \end{matrix} \right) \end{matrix} \tag{3.7}$$

$$\begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 100 \\ 20 \\ 80 \\ 80 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3.8}$$

$$\begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 100 \\ 20 \\ 60 \\ 60 \end{pmatrix} = \begin{pmatrix} 20 \\ 0 \end{pmatrix} \tag{3.9}$$

The flow conservation of the entries and exits is a linear combination of these constraints, so no extra condition is needed. If all internal nodes satisfy the flow conservation constraint, the incoming flow must equal the outgoing flow.

Matlab has a function called `fmincon` which minimises a variable or vector f by a function fun, with a given initial value x_0 [14]. This function will return a scalar. This Matlab function use the same variables as in Equation (3.5), but the first line (3.5a) is $\min \text{fun}(x)$. Optionally, non-linear constraints can be added. The general inputs for `fmincon` are stated below.

```
[x, fval] = fmincon(fun, x0, A, b, Aeq, beq, lb, ub)
```

Here `x` is the optimised vector and `fval` is the function value with the optimised vector.

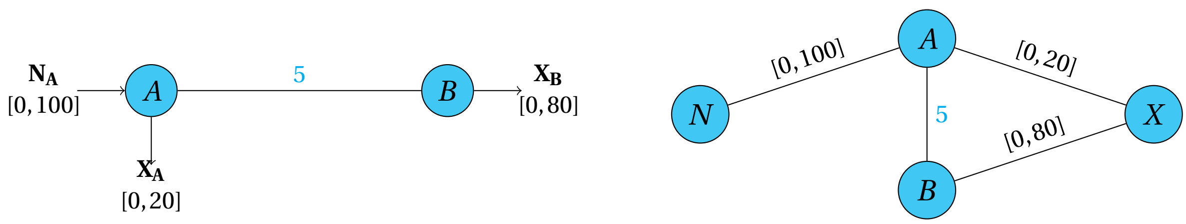
The problem of the maximisation of the transport moment on a gas network, the initial value for the entry and exit edges is chosen to be between the lower bound and the upper bound of these entries and exits and the internal edges is chosen to be zero. Generally, this initial value will not comply with some flow conservation constraints, but the function will adjust this initial value to a feasible optimisation variable vector. In this case, there are no inequality constraints (apart from the lower and upper bound), so the following Matlab code will give the most severe transport situation, based on flows on the pipelines and their lengths. `TMfun` is the transport moment function, which is negative, because it has to be maximised and the standard Matlab function minimises the input function. `flow` is the flow vector and `TM` is the value of the transport moment.

```
[flow, TM] = fmincon(-TMfun, f0, [], [], A, zeros(1,V), lb, ub)
```

Result for each example network

The algorithm, which is explained in the previous paragraphs, is applied to the first four example networks (page 5). The new representation of those networks is illustrated next to the original representation of the networks. Then the most severe situation is calculated. That situation is illustrated in the original network. The stress tests found for the new network is given by v_{flow}. The stress test with the largest transport moment of the capacity representation is also given by v_{cap}.

One pipeline network The original one pipeline network is given in Figure 3.7a. From the flow representation of Figure 3.7b the transport situation is calculated with maximal transport moment for this network. The transport moment is 400 and the stress test based on capacities on entry and exit points and the stress test based on flows through the pipelines corresponding to this moment is given in Equation (3.10). The amount of flow going from the entry at A to the exit at A does not contribute to the severeness of the network. At this point the optimisation function, in this case the transport moment, does not improve by adding more capacity to the network, so the Matlab function has stopped. Moreover, for the pipeline representation is only the value of AB required. The other values of the vector are only needed for the optimisation and finding the stress test.



(a) Original representation (b) Flow representation

Figure 3.7: One pipeline network

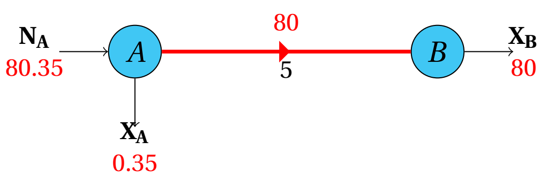
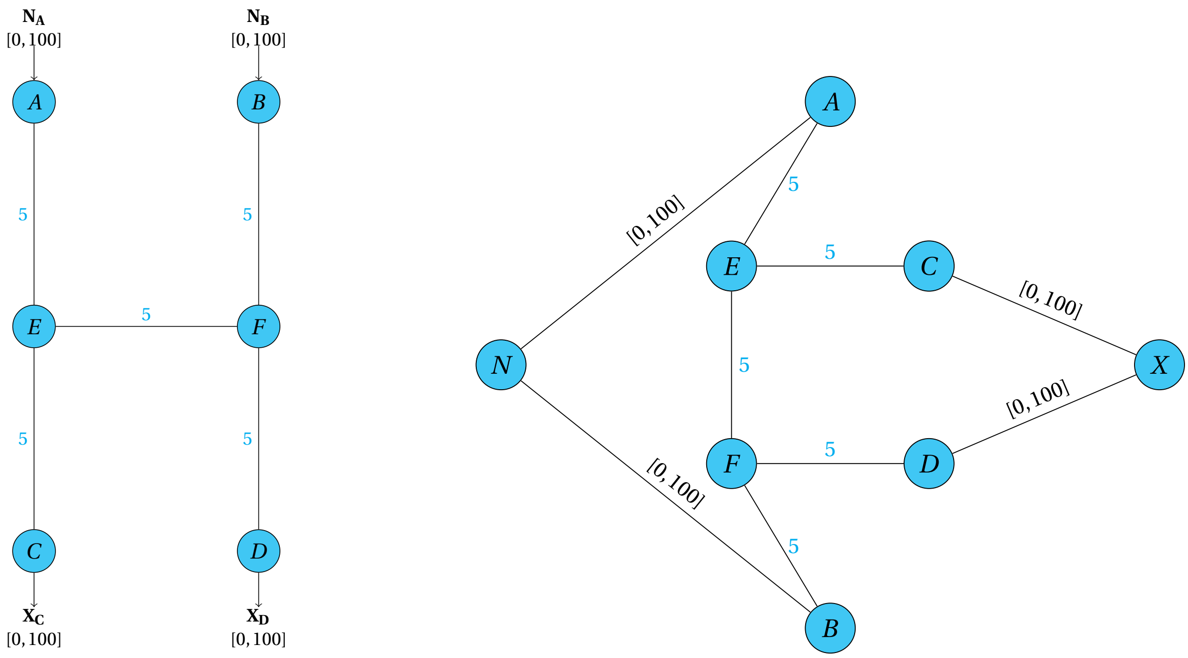


Figure 3.8: One pipeline network: maximal transport moment

$$v_{\text{cap}} = c\begin{pmatrix} N_A \\ X_A \\ X_B \end{pmatrix} = \begin{pmatrix} 100 \\ -20 \\ -80 \end{pmatrix} \tag{3.10}$$

$$v_{\text{flow}} = f\begin{pmatrix} NA \\ AX \\ BX \\ AB \end{pmatrix} = \begin{pmatrix} 80.35 \\ 0.35 \\ 80 \\ 80 \end{pmatrix} \tag{3.11}$$

Simple H-network The original and flow representation of this network is given in Figure 3.9. The transport moment found with this network is $5 \cdot 4 \cdot 100 = 2000$. The situation is shown in Figure 3.10. This transport situation is the one which is expected in terms of severeness. The most severe stress test based on end points is the same as the stress tests based on pipelines as can be seen in Equation (3.12).



(a) Original representation (b) Flow representation

Figure 3.9: Simple H-network

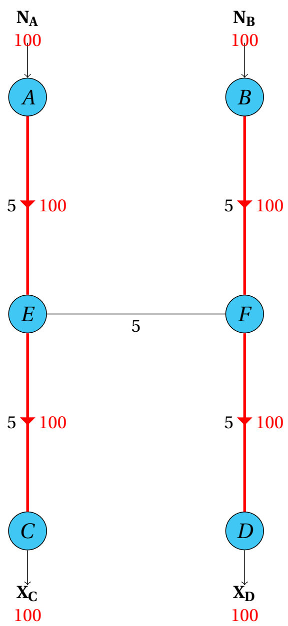
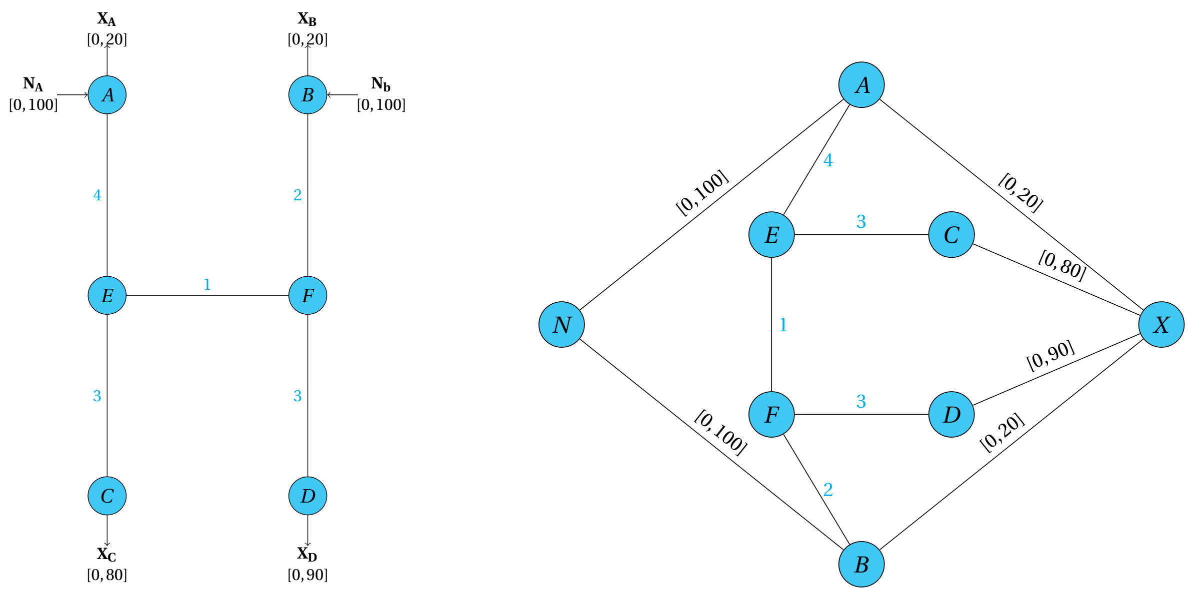


Figure 3.10: Simple H-network: maximal transport moment

$$v_{\text{cap}} = c \begin{pmatrix} N_A \\ N_B \\ X_C \\ X_D \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ -100 \\ -100 \end{pmatrix} \tag{3.12}$$

$$v_{\text{flow}} = f \begin{pmatrix} NA \\ NB \\ CX \\ DX \\ AE \\ BF \\ EC \\ FD \\ EF \end{pmatrix} = \begin{pmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 0 \end{pmatrix} \tag{3.13}$$

Advanced H-network The two network representations are found in Figure 3.11. The transport moment of the optimisation problem is 1070. The stress test vectors are found in Equation (3.15) and they are illustrated in Figure 3.12. This is almost the same situation as the situation that is found in previous section (see Equation (3.14)). In the previous section there was a flow of 20 from N_A to X_B, but in the new representation this flow does not contribute to the severeness of the scenario. Therefore the stress test is accepted as long as the difference between NB and BX is 70 and the LP conditions are satisfied. The entry and exit edges in the new representation do not have to be saved, only the internal edges are needed and then the difference disappears. Just as the one pipeline network the algorithm stops if the optimisation function does not improve.



(a) Original representation

(b) Flow representation

Figure 3.11: Advanced H-network

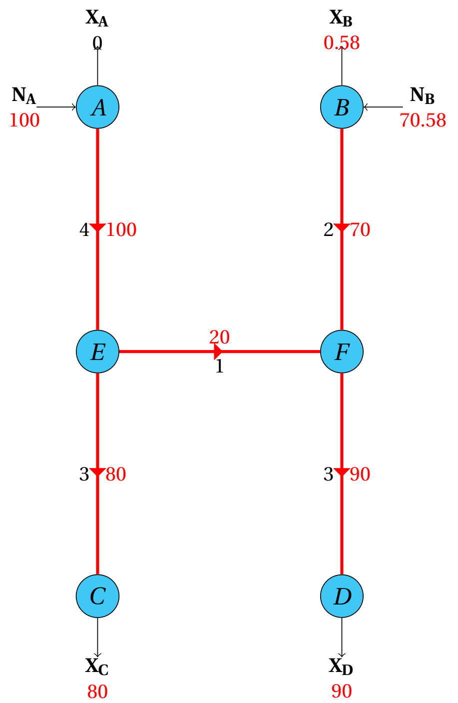
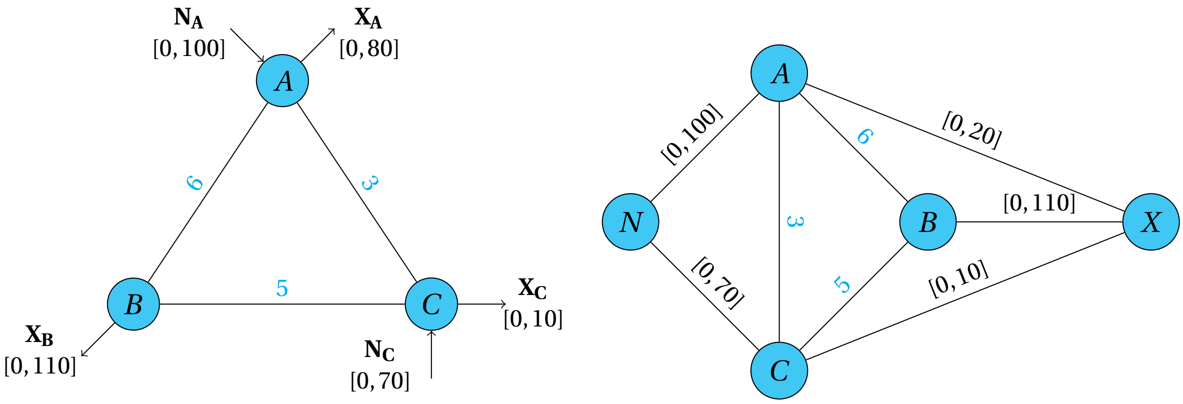


Figure 3.12: Advanced H-network: maximal transport moment

$$v_{\text{cap}} = c\begin{pmatrix} N_A \\ N_B \\ X_A \\ X_B \\ X_C \\ X_D \end{pmatrix} = \begin{pmatrix} 100 \\ 90 \\ 0 \\ -20 \\ -80 \\ -90 \end{pmatrix} \tag{3.14}$$

$$v_{\text{flow}} = f\begin{pmatrix} NA \\ NB \\ AX \\ BX \\ CX \\ DX \\ AE \\ BF \\ EC \\ FD \\ EF \end{pmatrix} = \begin{pmatrix} 100 \\ 70.58 \\ 0 \\ 0.58 \\ 80 \\ 90 \\ 100 \\ 70 \\ 80 \\ 90 \\ 20 \end{pmatrix} \tag{3.15}$$

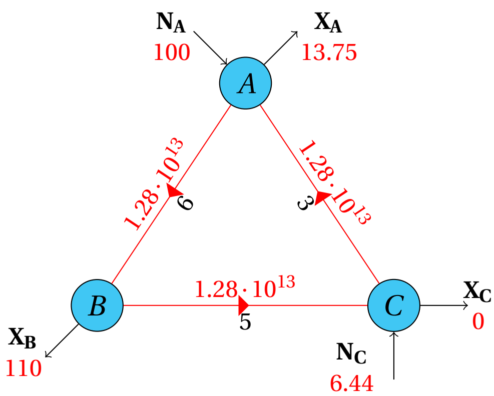
Triangular network This network is a simple network, but with a loop. In Figure 3.13 the transformation of this network is seen. When applying the algorithm to this network, a problem arises. The conditions still hold, but much gas is flowing through the loop. Because of the large amount of flow, the capacities on the entry and exit points are not significant for the optimisation function. Therefore the scenarios of v_{cap} and v_{flow} are not the same. The algorithm has stopped because the maximal number of iterations is exceeded.



(a) Original representation

(b) Flow representation

Figure 3.13: Triangular network



$$v_{\text{cap}} = c\begin{pmatrix} N_A \\ N_C \\ X_A \\ X_B \\ X_C \end{pmatrix} = \begin{pmatrix} 100 \\ 20 \\ 0 \\ -110 \\ -10 \end{pmatrix} \tag{3.16}$$

$$v_{\text{flow}} = f\begin{pmatrix} NA \\ NC \\ AX \\ BX \\ CX \\ AB \\ AC \\ BC \end{pmatrix} = \begin{pmatrix} 100 \\ 1.82 \\ 0 \\ 101.82 \\ 0 \\ 1.28 \cdot 10^{13} \\ -1.28 \cdot 10^{13} \\ 1.28 \cdot 10^{13} \end{pmatrix} \tag{3.17}$$

Figure 3.14: Triangular network: maximal transport moment

The solution is not sufficient when running this function for the LP given in Equation (3.6) for a network with a loop. It was seen in Figure 3.14 and also the Figure B.1 on page 52 for the shopping cart network. The constraints to the problem are satisfied, but the flow in the loop is too high and this situation is not realistic. Moreover, this situation is probably too severe in reality for the gas network. This is not a problem of using the non-linear program, but the problem is with the definitions within the model. There are some solutions to this problem:

- When reducing flow in the loop of the found solution until one pipeline has no flow, the problem of too much flow in the loop is solved. Although this situation is more sensible than the previous one, it is still not realistic and moreover this solution does not make use of the shortest path method. Mostly the gas will flow along the shortest paths. The result is seen in Figure B.2 on page 53.
- The second possible solution is to put bounds on the pipelines. There exist bounds on the flow through pipelines in the program MCA. These bounds can be used to suppress the excessive flow. The result of this solution is seen in Figure B.3 on page 51. Despite the realistic scenario, the bounds on the pipelines are chosen manually and this scenario is not necessary the most severe scenario. Again, the shortest path is not used here.
- All gas networks have visible loops, so these loops are known. Therefore, the third solution is to add a non-linear constraint to the problem which prevents that flow through the pipelines in a loop do not

have the same direction. Yet, the function does not find a feasible solution when running Matlab with this addition. Unfortunately `fmincon` is protected in Matlab and therefore there is not enough information how this function optimises the non-linear program. Besides the optimisation of maximising the transport moment, the flow pattern must be found such that the path is minimised. So the maximisation of the transport moment should be done at the same time as the minimisation of the transport path.

Previous solutions did not use the shortest path for the flow. For the stress test based on end points, this method was not required, because the internal flow pattern was not needed to find the severe scenarios. Moreover, this method (discussed in Section 3.1) finds a set of severe scenarios, but the methods with non-linear programming only finds the scenario with the maximal transport moment. A set with stress tests is preferred, because this set is representative for all the transport situations. Just one stress test can be the only representative for all transport situations if the network is very simple. More stress tests can be found by finding the vertices of the polytope. The vertices are the intersections of the linear conditions of the NLP. However, this only holds when the polytope is convex, but the non-linear conditions of the problem can cause a non-convex polytope. Another solution is to find scenarios in half planes. This is a more general approach than the vertices. The anchor points in the stress test algorithm of the end points take direction of the network into account, so such method may be considered to find the stress tests based on pipelines.

Finding a method more similar to the anchor point method of the stress test algorithm is to allow a particular directions in the network. This takes out an amount of freedom for generating the stress tests, but after taking different directions every time the algorithm has run, a set is generated with many stress tests. It is questionable whether this method gives the right set of scenarios. Restricting some directions may not give the most severe scenarios.

Finding the stress tests based on capacities of end points, MCA also calculates a flow pattern based on these stress tests. This lead to the same amount of stress tests as the generation method of end points. How MCA converts the stress tests will be explained and discussed in the next section.

3.2.2. Conversion from the end point representation of stress tests

The following method is implemented in MCA to find the flow vector from the capacity vector. The (capacity) stress tests found by the stress test algorithm (page 11) are considered to be known. There are also bounds on the pipelines given: an upper and a lower bound. The conversion from the known stress test to the pipeline stress test is done by adding flow on the pipelines such that the capacity stress test is satisfied. The optimal flow pattern is found by minimising the following transport moment: $T = L^T f$ within the predefined bounds for pipelines. Transport moment T is here defined by the vector of pipeline lengths L and flow vector f. It is possible that no flow pattern is found if the pipeline stress test have to satisfy the known transport situation and the bounds on the pipelines. However, the bounds on the pipelines are not strict, but undesirable. Therefore, if a flow on a pipeline must exceed its bound, the cost function transport moment is replaced by another function $\overline{L}(f)$ which has higher costs than the transport moment.

Figure 3.15 is an illustration to the cost function for one pipeline. When the amount of flow through the pipeline is below the upper bound (ub), then the cost function is equal to the transport moment. If the flow exceeds the upper bound, another cost function is used to give a penalty to the solution and crossing the bound will be punished. Trivially, the same holds for the lower bound, which represents the opposite flow direction.

When the stress vectors for the flow are derived for each scenario, this set should be reduced just as in the previous case to reduce the amount of computations. The analysis for the reduction of the set of stress tests is given in the next chapter.

3.3. Conclusions

There are two types of representations of the transport situations discussed in this chapter. The first one is described by the capacities on the end points. These stress tests for a certain network are found by applying the stress test algorithm. The contractual bounds of capacity on the end points and their mutual distances are input variables for this algorithm. The algorithm makes use of the more implicit transport moment in which directions are taken into account such that a set of stress tests is generated. This representation is taken from the shippers perspective, because they are only interested in the end points of the gas network.

The other representation is by flow through pipelines. This representation is taken from the modeller's perspective and it can be used for a more explicit definition of the transport situation. To generate the stress

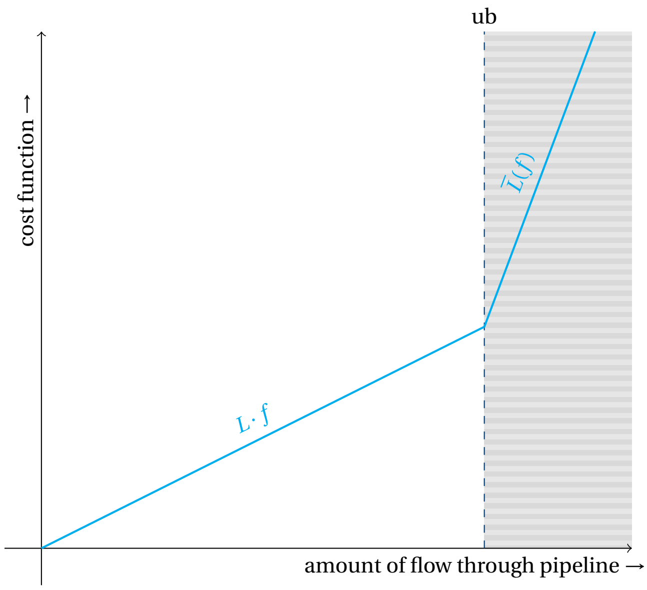


Figure 3.15: Cost function of a pipeline for finding the flow pattern

tests for the scenarios based on pipelines, more calculations are needed. In the previous representation calculations for the flow pattern was not required, because the network was like a black box in that representation. So in pipeline one, more calculations and information are required, but also more information is gained by knowing the flow pattern. This pattern gives information which is valuable for the continuation of the planning process.

No sufficient generation method directly from the network information for the stress tests based on pipelines is found in this network. An optimisation function (transport moment) and conditions belonging to the problem are known, so (non-)linear programming is used to solve the optimisation problem. Non-linearity appears in the optimisation problem, because the flow vector must be positive in the optimisation function, but the flow vector must include the direction by negative and positive values for the conditions in the optimisation problem. Therefore, the optimisation function includes the absolute function which is not completely linear, but piecewise linear. When applying only the bounds of capacity and flow conservation conditions, the networks with loops get unrealistic solutions because much gas is flowing around in the loops. Therefore a non-linear condition must be added to the Non-Linear Programming (NLP) to ensure that no unusual amount of flow is in the loops. The Matlab function `fmincon` does not find a feasible solution to the NLP after this addition. Even if the function finds the solution, there is only one stress test found; the most severe one according to the transport moment. Moreover, in the NLP the solution for the flow pattern finds a pattern with the shortest path. There are some possible solutions given in this chapter which are not carried out in this report.

In the continuation of this report, the representation of the pipelines is received from the program MCA of Gasunie. In MCA, a conversion is made from the stress tests based on end points to pipelines. This conversion leads to the same scenarios as in the end point representation but in a different form, i.e. flows through pipelines. The flow pattern found for a scenario takes the shortest path along the pipelines into account, which is realistic.

4

Reduction of the generated set of stress tests

In the previous chapter, stress tests have been generated. Generally, a large number of stress tests are found. The number of stress tests is equal to the number of exit points (around 1100 in HTL-network) plus the number of entry points (around 50 in HTL-network) [15]. To reduce the number of stress tests, similarities between the stress tests can be used. To make this possible, a distance has to be defined to measure the difference between the stress tests and determine to the similarity of stress tests.

4.1. Current reducing method

The first step in the current reducing method is to have the generated set reduced by keeping only unique stress tests. In the example of Section 3.1 it is seen that equal stress tests are found for different anchor points. However, this unique set of stress test is still too large for the final computations. Therefore, the quadratic form distance is used to measure the "difference" between the stress tests and a reduction criterion can be given. The $\mathscr{L}_p$-norm (see Equation (4.1)) is an often used distance for vectors. However, this distance is not adequate in this case, because the severeness of gas transport is dependent on both flow and transport distance.

$$\mathscr{L}_p(\mathbf{x}) = \left(\sum_i |x_i|^p \right)^{1/p} \tag{4.1}$$

4.1.1. Quadratic form distance

The comparison of numerical objects is not only done for scenarios in gas networks, but also in other fields. In the article of Skopal, et al [17] images are compared to each other and with the QFD (Quadratic Form Distance) a rate of similarity is given. There are some techniques for comparing images by making histograms of the colours at every location of the image and comparing them with other images. However, if the image has some noise or is scaled or rotated, those techniques do not give a good similarity measure between the distorted image and the original. The QFD takes distortion of the image into account by increasing the number of dimensions where more dependencies of the image are considered. An example of the dimensions of an image is the amount of red, green and blue of each pixel, so for m pixels, the dimension of the image is $3m$, where only colour is included. Texture for example is also a quantity which can be added to be a dimension.

The quadratic form distance can also be used for comparing stress tests, because besides the Euclidean distance (or other $\mathscr{L}_p$-distances) between the vectors of stress tests, the geographical distance has also to be taken into account for comparing stress tests. When two entry points N_1 and N_2 are close to each other in the network, the transport load from N_1 to exit point X is assumed to be similar to the transport load from N_2 to X.

The aim is that the quadratic form distance will give a good similarity rate of stress tests. The stress tests can be denoted as n-dimensional vectors where n is the total number of entry and exit points. The capacity on the entry points is given in the vector with a plus sign and the capacity on the exit points is given in the vector with a minus sign, just as in the previous chapter.

The quadratic form distance of the vectors x and y is given by

$$QFD_{\mathbf{A}}(x,y) = \sqrt{(x-y)^T\,\mathbf{A}(x-y)} \tag{4.2}$$

where **A** is an $n \times n$ symmetric semi-positive definite matrix [13] and x and y are the n-dimensional vectors which represent two stress tests. The transport moment represents the severeness of the situation, so the distance of the transport must be included in the matrix **A**.

In the article of Skopal et al the matrix **A** is defined by

$$a_{ij} = 1 - \frac{d_{ij}}{d_{\max}}, \quad \text{with } d_{\max} = \max_{i,j} d_{ij}, \quad i,j = 1,2,\ldots,n, \tag{4.3}$$

where a_{ij} are the entries of matrix **A** and d_{ij} is the Euclidean distance between representatives of colours i and j.

For gas transport situations, the matrix **A** is defined as in Equation (4.3), but with another distance d_{ij} than in the colour comparison, see the master thesis of K. Lindenberg [13]. In the gas network the distance is chosen to be the shortest path along the pipelines between point i and point j. These points are entry and exit points.

This quadratic form distance is more preferable than the $\mathscr{L}_p$-distance, because the QFD gives also the correlation between different dimensions (capacities and shortest paths), while the $\mathscr{L}_p$-distances give the combination of the distances of each dimension independently and only the difference of capacities.

4.1.2. Example QFD

Illustrating the previous theory, the parametrisation of the one pipeline network is given in this paragraph. This network was given in the introduction in Figure 1.2 on page 5 and matrix **A** is derived by Equation (4.3) with **D** as the distance matrix of the shortest paths via pipelines. The considered network has only one pipeline, so the length of the pipeline can be arbitrary chosen as long as it is greater than zero. Call this length $\ell > 0$.

The distance matrix is for the network above is equal for every bounds on the capacity:

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & \ell \\ 0 & 0 & \ell \\ \ell & \ell & 0 \end{pmatrix} \tag{4.4}$$

The matrix **A** can be computed from **D**, which is given in Equation (4.5). Matrix **A** is not a symmetric positive matrix, which can be seen with the definition in next subsection.

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} - \frac{1}{\ell}\begin{pmatrix} 0 & 0 & \ell \\ 0 & 0 & \ell \\ \ell & \ell & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

From the stress test algorithm (Algorithm 3.1), the most severe situation is found by adding some capacity to the network until the bounds are exceeded. Some of the severe situations can easily be seen in this example. Like the situation that there is 80 injected at point A, no capacity leaves point A directly through the exit point of A and 80 will leave at the exit point at B. This situation is represented in the following stress test vector $(80, 0, -80)^T$.

Another severe situation is $(100, -20, -80)^T$, so here 100 is injected to point A and immediately 20 is leaving point A and 80 leaves point B at its exit point. In this example it is easily seen that these two situations give the same transport moment, because $T = f \cdot L = 80\ell$ for both situations.

In fact, all situations of the form $(80 + p, -p, -80)^T$ with $p \in [0, 20]$ will give the same transport moment. The QFD will show that they are similar, because the distance is zero. This is proven by taking two arbitrary stress tests of the following form $x_p = (80 + p, -p, -80)^T$ and $x_q = (80 + q, -q, -80)^T$ and comparing them with the QFD.

$$
\begin{aligned}
QFD_{\mathbf{A}}(x_p, x_q) &= \sqrt{(x_p - x_q)^T\,\mathbf{A}(x_p - x_q)} \\
&= \sqrt{\begin{pmatrix} p-q & q-p & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p-q \\ q-p \\ 0 \end{pmatrix}} \\
&= \sqrt{\begin{pmatrix} p-q & q-p & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}} \\
&= 0
\end{aligned}
$$

So for the one pipeline network all situations which correspond to $(80+p, -p, -80)^T$ with $p \in [0,20]$ are equally severe and most severe for the network with these bounds on the entry and exit points.

Using this relatively simple network the situations can be seen and be proven rather simply. For more complex networks, the computer should do the work. The computations for stress tests of the advanced H-network is given in Appendix C, Section C.1. There were three unique stress tests are found, so three comparisons have been done.

4.1.3. Metric distance

The quadratic form distance looks like a well defined distance, but is it a metric distance as required by mathematicians? In this section it is shown that the QFD with the defined parametrisation is not for every gas network a full metric distance. However, a less strict property of the distance is sufficient for checking the similarity between stress tests, because when the distance between two stress tests is zero, then these two does not always have to be the exact same situations. This can occur if an entry and an exit point are at the same location (see previous section). The situation when there is no flow in the network has distance zero to the situation when there is flow from the entry point to the exit point on the same location. It is necessary that $(x-y)^T\,\mathbf{A}(x-y) \geq 0$ for every $x, y \in \mathbb{R}^n$, because then $\text{QFD}_{\mathbf{A}}(x,y) \in \mathbb{R}$.

The matrix $\mathbf{A}$ is always symmetric if the entries are defined by $a_{ij} = 1 - \frac{d_{ij}}{d_{\max}}$, because $d_{ij} = d_{ji}$, thus $a_{ij} = a_{ji}$.

Assume matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is SPD (Symmetric Positive Definite), i.e. $v^T\,\mathbf{A}v > 0$ for every non-trivial $v \in \mathbb{R}^n$. Then $QFD_{\mathbf{A}}$ is a metric distance, because the next four requirements hold of the definition of a metric distance function [9]. The detailed proof of these requirements are found in Appendix C, Section C.2. The first two requirements are easy to see, but the others are more difficult. For the last requirement, the triangular inequality, the theorem of Cauchy Schwarz is used for this special case.

Definition (Metric distance)**.** $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is a metric distance function if the following requirements hold:*

- $d(x,y) \geq 0$
- $d(x,y) = d(y,x)$
- $d(x,y) = 0 \iff x = y$
- $d(x,y) \leq d(x,z) + d(y,z)$

In the simple example of previous subsection (4.1.2), the matrix $\mathbf{A}$ is not positive definite, but this is not a necessarily desired property for measuring the similarity between stress tests. The QFD is a square root of a vector-matrix-vector multiplication, so it is desired that the vector-matrix-vector multiplication is non-negative. Therefore, the matrix $\mathbf{A}$ is required to be SSPD (Symmetric Semi-Positive Definite), which is a less strict definition than SPD. If the matrix $\mathbf{A}$ is SSPD, the QFD is a semi-metric distance [9]. The definition of a SSPD-matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is $v^T\,\mathbf{A}v \geq 0$ for every non-trivial $v \in \mathbb{R}^n$. The definition of a semi-metric distance is given below and the proofs are given in Appendix C in Section C.2.

Definition (Semi-metric distance)**.** $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is a metric distance function if the following requirements hold:*

- $d(x,y) \geq 0$

- $d(x,y) = d(y,x)$
- $d(x,x) = 0$
- $d(x,y) \leq d(x,z) + d(y,z)$

If **A** is a (semi-)negative definite matrix or an indefinite matrix, the $QFD_{\mathbf{A}}$ is not a distance which can be used for the similarity of stress tests.

4.1.4. Other definitions of the matrix

K. Lindenberg [13] tested some definitions of **A** on symmetric semi-positive definiteness, however she did not find a definition such that it is SSPD for every transport network. In these definitions relations to transport distance, diameter and a combination of the two are considered in different forms. The conclusion of Lindenberg was that the original form of the $\text{QFD}_{\mathbf{A}}$, with $a_{ij} = 1 - \frac{d_{ij}}{d_{\max}}$ is the best found parametrisation of the QFD for the gas transport situations.

At Gasunie, $a_{ij} = 1 - \frac{d_{ij}}{d_{\max}}$ is used as the definition of **A** without any problems. It is possible that some properties of the gas network are not included, which leads to a symmetric semi-positive definite matrix. More research is needed for to discover these properties.

4.2. Reduction of stress tests based on pipelines

The flow representation of the stress tests give a more explicit description of the transport situation. Therefore, the quadratic form distance is unnecessary for this type of vectors. To use the quadratic form distance on the flow vectors, a diagonal matrix is found for the parametrisation matrix **A**, where the diagonal elements are the pipeline lengths. Then rewriting the QFD for the flow vector f a weighted $\mathscr{L}_p$-norm is found, which is shown in Equation (4.6) with weight vector w^2.

$$\begin{aligned} QFD(f) &= \left(f^T \begin{pmatrix} w_1^2 & & \\ & \ddots & \\ & & w_n^2 \end{pmatrix} f \right)^{1/2} \\ &= \left(\sum_i (w_i f_i)^2 \right)^{1/2} \\ &= \mathscr{L}_2^w(f) \end{aligned} \tag{4.6}$$

The distance between two scenarios must depend on the flow or capacities and the dependency of the transport distance in the network will be investigated in this paragraph.

The difference between two transport scenarios must depend on the relation of the transport distance and the amount of flow or capacity in the network. Therefore the distances to be investigated have to be dependent on the transport distance by adding a weight to the distance. This is already implemented at the quadratic form distance by the parametrisation matrix **A**. The standard $\mathscr{L}_p$-distance is metric which is given in Equation (4.7a) [9]. This norm is often used as the distance function for vectors, because the function is simple and it works in a multidimensional space $\mathbb{C}^n$. The weighted $\mathscr{L}_p$-distance $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with weight vector $w \in \mathbb{R}^n_{\geq 0}$ becomes as in Equation (4.7b).

$$d(x,y) = \mathscr{L}_p(x-y) = \left(\sum_i (|x_i - y_i|)^p \right)^{\frac{1}{p}} \tag{4.7a}$$

$$d_w(x,y) = \mathscr{L}_p^w(x-y) = \left(\sum_i (w_i |x_i - y_i|)^p \right)^{\frac{1}{p}} \tag{4.7b}$$

The weighted $\mathscr{L}_p$-norm is a metric distance. The proof of this statement is given in Appendix C, Section C.3.

The adjustment of the weight w to the $\mathscr{L}_p$-norm ensures that the entries of the compared vectors will not be distributed equally to the distance function. This is preferred for the stress tests, because transporting over a longer transport distance with a certain flow will have a larger transport load than the same flow over a shorter distance. This assumption that the weighted norm is preferred over the standard norm is checked in the next paragraph. There are requirements for the difference measure between transport scenarios. These requirements are tested in the next paragraph.

4.2.1. Test distances

The (semi-)metrics used for comparing the stress tests $x, y \in \mathbb{R}^n$ based on the capacities on entry and exit points are the following (Equation (4.8)).

$$d(x,y) = \sum_i |x_i - y_i| \qquad (\mathscr{L}_1\text{-norm})$$

$$d(x,y) = \sqrt{\sum_i (x_i - y_i)^2} \qquad (\mathscr{L}_2\text{-norm})$$

$$d_{\mathbf{A}}(x,y) = \sqrt{(x-y)^T \mathbf{A}(x-y)} \qquad (\text{QFD}_A)$$

$$\text{with } a_{ij} = 1 - d_{ij}/d_{\max} \qquad (4.8)$$

For the quadratic form distance on the capacities of the end points, the distance matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ is used, where the element d_{ij} is the transport distance between endpoint i and j. The weight of the flows is a vector which represents the pipeline lengths. The distances used for the flow representation of the stress tests are given in Equation (4.9).

$$d(x,y) = \sum_i |x_i - y_i| \qquad (\mathscr{L}_1\text{-norm})$$

$$d(x,y) = \sqrt{\sum_i (x_i - y_i)^2} \qquad (\mathscr{L}_2\text{-norm})$$

$$d_L(x,y) = \sum_i L_i |x_i - y_i| \qquad (\text{Weighted } \mathscr{L}_1\text{-norm})$$

$$d_L(x,y) = \sqrt{\sum_i L_i^2 (x_i - y_i)^2} \qquad (\text{Weighted } \mathscr{L}_2\text{-norm})$$

$$(4.9)$$

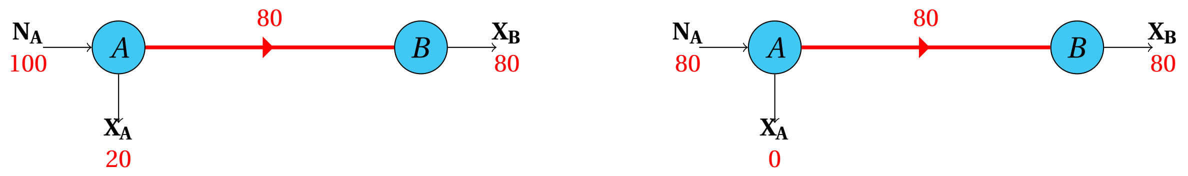
The distances above are for stress tests $x, y \in \mathbb{R}^m$ described by the flows on pipelines. The first two norms are only defined by the flows through each pipeline i of the network. The weighted norms are defined by these flows as well, but also in combination to the pipeline length vector $L \in \mathbb{R}^m$, which is called the weight vector of these norms.

The first check if the distances suffices the requirements of the gas transport network, the simple networks *one pipeline network*, *simple H-network* and a temporary network (network 2) are used. The requirements are drawn up to get the conditions for the network physics. The scenarios which are used to check the distances are illustrated in Figure 4.1 (page 32), Figure 4.2 (page 32) and Figure 4.3 on page 33. The requirements are listed below.

- $d(1A, 1B) = 0$, because these scenarios are in terms of gas transport similarly severe.
- $d(2A, 2B) = d(2B, 2C)$, because the scenarios of A and C should have the same distance to scenario B.
- $d(2A, 2B) > d(2A, 2C)$, the pipeline of scenario B is ten times larger than the other pipelines and therefore is the difference between scenario A and scenario B larger than the difference between scenario A and scenario C.
- $d(3A, 3C) = d(3B, 3C)$, scenarios A and B have both three different pipelines with the same amount of flow than scenario C.
- $d(3C, 3D) < d(3C, 3E)$, because there is more flow difference in scenarios C and E than in C and D.

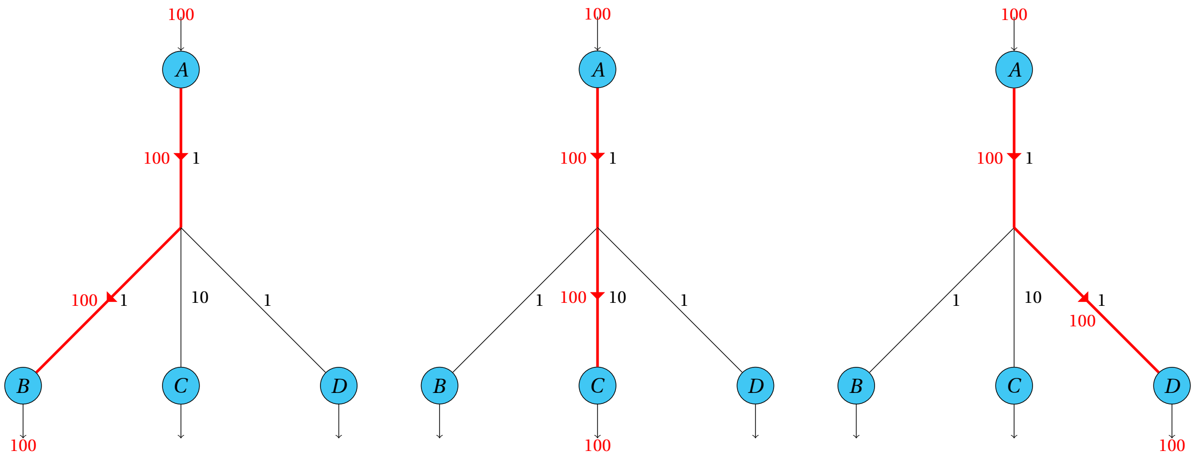
The main conclusion which is drawn from Table 4.1 is that the distance measure for gas transportation scenarios has to be weighted with network distances (path lengths and pipeline lengths). For the distances concerning the flow through pipelines both weighted $\mathscr{L}_1$-norm and weighted $\mathscr{L}_2$-norm satisfy the requirements and the QFD for the stress tests based on end points is the only sufficient distance which is tested in this report.

The requirement for the first network holds almost for every distance function. For network 2, there are two requirements. The first one is that d(2A,2B)=d(2B,2C) holds for every distance function in the table.



(a) Scenario 1A (b) Scenario 1B

Figure 4.1: Scenarios for the one pipeline network



(a) Scenario 2A (b) Scenario 2B (c) Scenario 2C

Figure 4.2: Scenarios for network 2

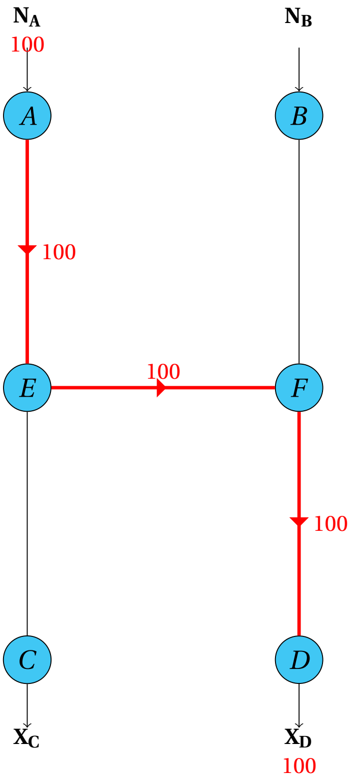
However, the second requirement that the difference between 2A and 2B is strict larger than the difference between 2A and 2C does not hold for every distance, it only holds for the quadratic form distance and the two weighted $\mathscr{L}_p$-norms.

The difference between 3A and 3C must have the same value as the difference between 3B and 3C. This requirement is satisfied by all the distance functions which are investigated. The difference between scenarios 3C and 3E should be bigger than the difference between scenarios 3C and 3D. Actually, $d(3C,3D) = 2 \cdot d(3C,3E)$ for every definition of the distance in the table. This is caused by the flow that is chosen and that in each scenario the same flow pattern is used. On every pipeline (except for pipeline EF) of the three scenarios 3C, 3D and 3E is the flow 100, 80, 60 respectively. Therefore, $d(3C,3D) = 2 \cdot d(3C,3E)$ is reasonable, but the same result is in other networks not necessary.

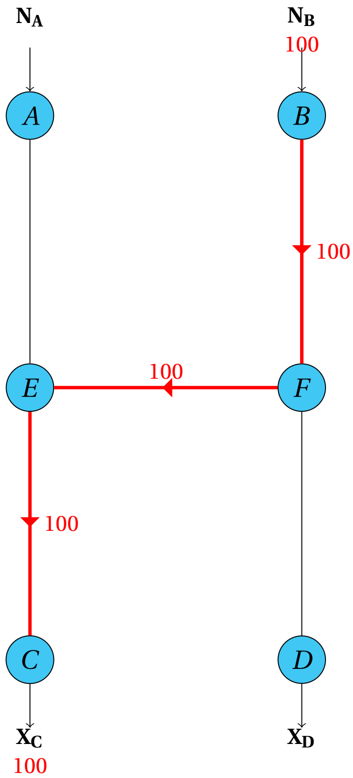
In Appendix B.2 on page 55 distances of the networks simple H-network, advanced H-network and the shopping cart network are plotted. For each of the networks, the quadratic form distance between stress tests based on capacities on end points is plotted on the x-axis. On the y-axis there are the two norms plotted:

Distance	Capacities on end points			Flows on pipelines			
	$\mathscr{L}_1$-norm	$\mathscr{L}_2$-norm	QFD_A	$\mathscr{L}_1$-norm	$\mathscr{L}_2$-norm	$\mathscr{L}_1^L$-norm	$\mathscr{L}_2^L$-norm
d(1A,1B)	40	28.28	0	0	0	0	0
d(2A,2B)	200	141.42	141.42	200	141.42	1100	331.66
d(2A,2C)	200	141.42	60.30	200	141.42	200	141.42
d(2B,2C)	200	141.42	141.42	200	141.42	1100	331.66
d(3A,3C)	200	141.42	141.42	300	173.21	1500	387.30
d(3B,3C)	200	141.42	141.42	300	173.21	1500	387.30
d(3C,3D)	80	40.00	32.66	80	40.00	400	89.44
d(3C,3E)	160	80.00	65.32	160	80.00	800	178.89

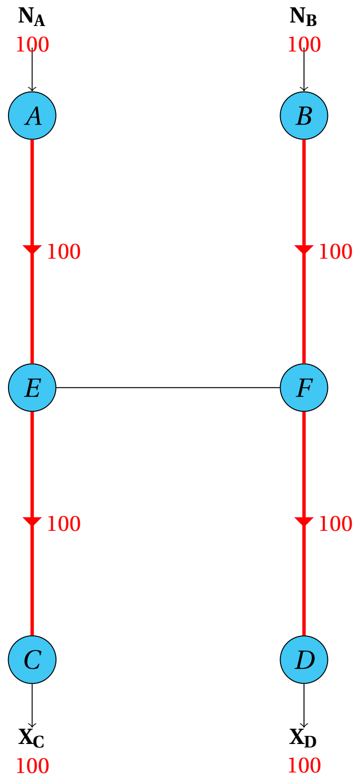
Table 4.1: Verifying requirements of the distance function for transport situations



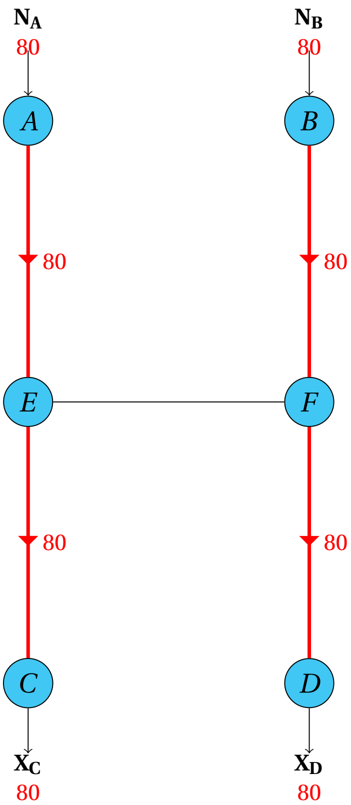
(a) Scenario 3A



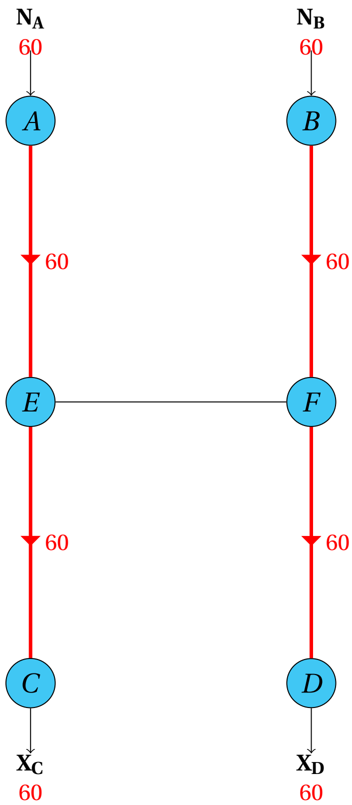
(b) Scenario 3B



(c) Scenario 3C



(d) Scenario 3D



(e) Scenario 3E

Figure 4.3: Scenarios for the simple H-network

the weighted $\mathscr{L}_1$-norm and the weighted $\mathscr{L}_2$-norm of the difference of stress tests based on flow through pipelines. The values of the distances are rounded off such that the density can be seen, which is represented by the different colours.

The figures show that the more complex the network the wider the relative spread is between the two representations of the stress tests. This is expected, because when the inner network gets more complex, this leads to more divergence of the two representations of the stress tests.

Both weighted norms meet the requirements of the difference measure for the stress tests based on flows through pipelines, so the $\mathscr{L}_1^L$-norm is used in this report as the distance function for the scenarios based on pipelines.

4.3. Other reduction criteria

Not only the flow through gas pipelines and the length of these pipelines have influence on the severeness of the transport situation, but other network aspects as well. For example the diameter and the pressure drop also have an effect on the severeness of the scenario. Instead of only taking the length of each pipeline as the weight for the weighted $\mathscr{L}_1$-norm, the diameter can also taken into account. To get the right combination between the length and the diameter, the physical relations must be kept in mind. The larger the diameter of a pipeline, the less severe the situation with the same amount of flow is. This is also related to the pressure drop formula. This formula is shown in Equation (4.10) [21]. p_{in} is the pressure at the start of the pipeline, the p_{out} is the pressure at the end of the pipeline and $\Delta p = p_{in} - p_{out}$ is the pressure drop. Q is the flow from the start to the end of the pipeline. The term in front of the flow is interesting for the weight of the new distance function.

$$p_{in}^2 - p_{out}^2 = k \cdot \frac{L}{D^5} \cdot |Q| \cdot Q \tag{4.10}$$

The units of the quantities in Equation (4.10) are given in Table 4.2 and in Appendix E [21].

Description	Quantity	Unit
Pressure at the start of the pipe	p_{in}	bar
Pressure at the end of the pipe	p_{out}	bar
Constant	k	60^4 kg/m^3
Length of the pipe	L	km
Diameter	D	m
Volumetric flow	Q	dam^3/h

Table 4.2: Quantities and units of the equation of pressure drop (Equation (4.10))

The units of satisfy Equation (4.10), which is found below. The unit bar can be rewritten as 1 bar = 10^5 Pa = 10^5 J/m^3 = 10^5 kg/(ms^2). So the left hand side has the following unit:

$$\begin{aligned} [\text{bar}] - [\text{bar}] &= \left[10^5 \frac{kg}{ms^2}\right] - \left[10^5 \frac{kg}{ms^2}\right] \qquad \text{(to SI base units)} \\ &= 10^5 \frac{[kg]}{[m][s]^2} \\ &= 10^5 [kg][m]^{-1}[s]^{-2} \end{aligned}$$

And the right hand side has the same unit:

$$\begin{aligned} &[60^4 kg/m^3] \cdot \frac{[km]}{[m]^5} \cdot |[dam^3/h]| \cdot [dam^3/h] \\ &= [60^4 kg/m^3] \cdot 10^3 \frac{[m]}{[m]^5} \cdot \left| \frac{10}{3600} [m^3/s] \right| \cdot \frac{10}{3600} [m^3/s] \qquad \text{(to SI base units)} \\ &= 60^4 \frac{[kg]}{[m]^3} \cdot 10^3 \frac{[m]}{[m]^5} \cdot \frac{10^2}{60^4} \frac{[m]^6}{[s]^2} \\ &= 10^5 [kg][m]^{-3} \cdot [m]^{-4} \cdot [m]^6 \cdot [s]^{-2} \\ &= 10^5 [kg][m]^{-1}[s]^{-2} \end{aligned}$$

So the units of the pressure drop equation is correct. Adding the term of the length and diameter of a pipeline, the new weighted $\mathscr{L}_1$-norm is as in Equation (4.11). k is assumed to be a fixed constant and therefore irrelevant for the new distance function.

$$\mathscr{L}_1^{LD}(x-y) = \sum_i \frac{L_i}{D_i^5}|x_i - y_i| \tag{4.11}$$

The reduction distance appear even more like the pressure drop equation (Equation 4.10) if the following distance is used:

$$\mathscr{L}^{\Delta p}(x-y) = \sum_i \frac{L_i}{D_i^5}|x_i - y_i|(x_i - y_i) \tag{4.12}$$

Despite the appearance of the pressure drop, the distance above is not actually a distance function, because it is not always larger than zero. There is a version of the pressure drop equation which does not take the transport direction into account. In this equation the flow is squared and becomes a mathematical distance function.

$$\mathscr{L}^{\Delta p}(x-y) = \sum_i \frac{L_i}{D_i^5}(x_i - y_i)^2 \tag{4.13}$$

In the next chapter the results in comparison with the original weighted $\mathscr{L}_1$-norm is given.

4.4. Reduction process

When the distance (described in Section 4.2) between stress tests is less than a certain number ε, then the stress tests are considered to be similar. So considering one stress test v, all stress tests in the ball with radius ε around this stress test, $B(v,\varepsilon)$, are considered to be similar. The set of all these similar vectors can be reduced to one stress vector.

Actually, the stress test set can be reduced even further. The scenarios which are less severe than the ones in the ball can be removed from the set, because of the assumption that a scenario is feasible if a more severe one is feasible. An illustration in $\mathbb{R}^2$ is given in Figure 4.4a. In this figure the cone of clustering is shown in blue. All vectors in this cone can be reduced to the vector v. All the stress tests within this cone are in the same cluster. However, the clustering technique is not unique, see Figure 4.4b. When vector v_3 in this figure is considered first, v_2 will be in the cluster of v_3, but if v_1 is first considered, vector v_2 is in the cluster of v_1. For a bigger set, the clustering of stress tests is a more complex problem. Gasunie has chosen to consider the stress test with the largest corresponding distance first. Thus the most severe situations have priority of being the main vector in a cluster. The goal of the clustering is to get less stress test vectors, so for this purpose it is indifferent to which cluster vector v_2 belongs.

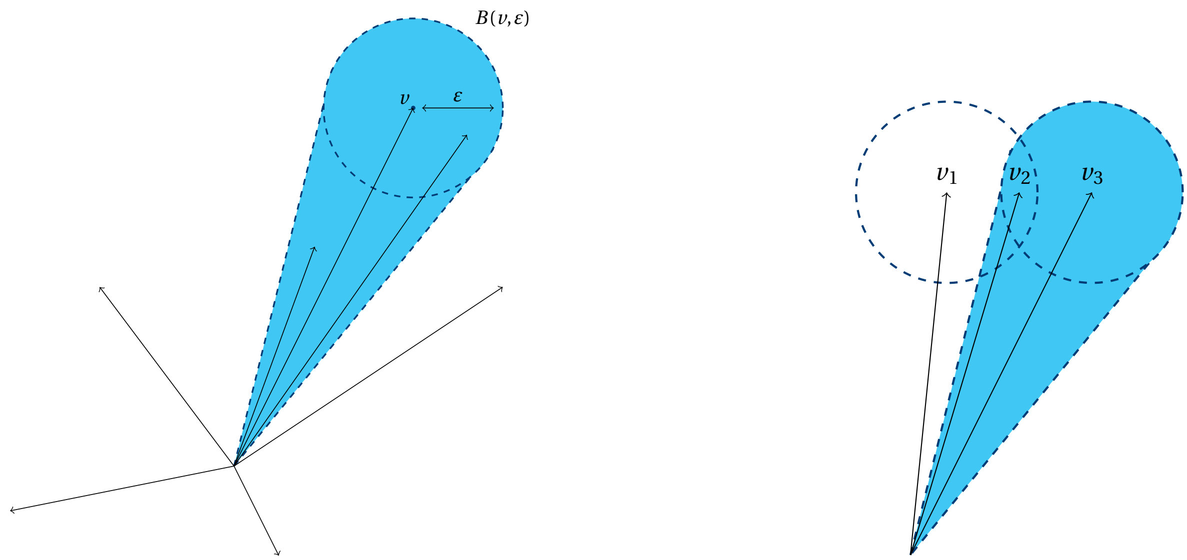
Despite this procedure of priority, the reduced set of stress tests will change easily for a small change in the network. It is preferred to have a stable set of stress tests for which the final hydraulic calculations are done.

Furthermore, the question of what ε should be is difficult. This value of ε is strongly dependent of the network. If the capacities of one network are larger than the capacities of the other network, the values of the stress tests will be bigger and therefore ε is different to achieve the similar clusters. In the next chapter, the value of ε is determined for each network and both types of stress tests and these values are analysed.

4.5. Conclusions

The amount of stress tests found by the generation methods discussed in Chapter 3 is too much in order to check these scenarios by hydraulic calculations. This set of stress tests can be reduced by leaving out the similar stress tests and the less severe ones. The stress tests within the ball with radius ε of a certain stress test vector v are considered to be similar. Moreover, all vectors within the cone of clustering can be reduced to v.

All relative distances have to take the properties of the network into account. For example, the transport distance is not equal to the Euclidean distance. In this chapter, there are distance functions tested on all the requirements for transport situations. The conclusion is that these functions must depend on the flow or capacity of gas and the transport distance. For the stress tests based on end points, the only sufficient function found is the quadratic form distance:



(a) Cone of clustering

(b) Clustering is not unique

Figure 4.4: Clustering method

$$QFD_{\mathbf{A}}(x,y) = \sqrt{(x-y)^T\,\mathbf{A}(x-y)} \qquad (\text{QFD}_A)$$

$$\text{with } a_{ij} = 1 - \frac{d_{ij}}{d_{\max}}, \quad \text{with } d_{\max} = \max_{i,j} d_{ij}$$

In this equation i, j are the end points and d_{ij} is the transport distance between end point i and j. $\mathbf{A}$ is the parametrisation matrix with entries a_{ij}.

The distance function for stress tests based on pipelines, two weighted norms are found: the weighted $\mathscr{L}_1$-norm and the weighted $\mathscr{L}_2$-norm. Both norms satisfy the requirements for scenarios, so there is chosen to use the weighted $\mathscr{L}_1$-norm in this report:

$$d_w(x,y) = \sum_i w_i |x_i - y_i| \qquad (\text{Weighted } \mathscr{L}_1\text{-norm})$$

First, the weight w_i for this norm is chosen to be the length L_i of the i^{th} pipeline. This choice of the weight vector leads directly to the transport moment for stress tests based on pipelines (see Equation 3.2 on page 9).

However, the wider the pipeline, the less is the pressure drop. So the addition of the diameter to the distance function should improve the comparison of the stress tests. The term $\frac{L_i}{D_i}$ in the pressure drop equation (Equation 4.10) relates the length L_i and diameter D_i of pipeline i, so this term should be sufficient as weight for the the weighted $\mathscr{L}_1$-norm.

In the next chapter, the problems and results of the reduction techniques are given.

5
Results of the reduction methods

Research question number four ("What are the criteria for similarity of stress tests?") is answered in this chapter. The answer will be discussed for different kind of stress tests and different networks. The comparison of the two representations of stress tests will be done with these criteria. The impact of the structure of the network will also be examined.

First, the criteria are determined by calculating the reduced set of stress tests by the program MCA as a reference. The ε is found by taking the highest number such that the distance (QFD_A or $\mathscr{L}_1^L$) between the stress tests is larger than that number. If the ε is larger than this number, some of the stress tests are in the cone of clustering of other stress tests. Then the set of stress test should be reduced even further and the reference from MCA does not suffice. If the value of ε is smaller, the set of stress tests is not a representative of all transport situations. This value for ε is found for the simple H-network, the advanced H-network, the triangular network and the shopping cart network. In total $4 \cdot 2 = 8$ criteria will be found. The hypothesis is that these criteria are different from each other, because for flow through pipelines another unit is used and the distance norm is different for both types of stress tests. Moreover, the networks have other bounds and the dimensions are also different for every network. For example, the stress test vector of capacities on end points of the simple H-network is in $\mathbb{R}^4$ while the one for the advanced H-network is in $\mathbb{R}^6$. Therefore, it is expected that the criteria of similarity are different from each other.

After this comparison, the differences between reducing methods of capacities on end points and flows through pipelines can be examined. This is done by generating large numbers of feasible transport situations by MCA. These situations are not necessarily the most severe scenarios. The transport situations are given in both representations: vectors of end points and vectors of pipelines. For every situation, the distance is measured to the stress tests of the reduced set, where the identifier of the nearest vector is saved. This is done for capacities on end points as well as for flows through the pipelines. The hypothesis here is that the classification of the situations to the stress tests is shows a higher degree of similarity for the simple and advanced H-network than for the triangular network and the shopping cart network. This hypothesis is made due to the fact that there are loops in the last two networks. Therefore, an optimisation step is performed to get the vectors of pipelines. The optimisation step involves the shortest path optimisation as explained on page 25.

5.1. Determining criteria of similarity

The criteria for each network and each representation of stress tests are found by Equation (5.1). S is the set of stress tests found by MCA and $d(\cdot,\cdot)$ is the distance function. The quadratic form distance and the weighted $\mathscr{L}_1$-norm are the distance functions for the end point stress tests and the pipeline stress tests respectively. It is assumed that the stress tests from MCA are the ideal representatives of all the transport situations, therefore, the minimum distance between the stress tests is the radius of the ball as explained in Section 4.4. If the radius is larger, another stress test should not be included in the final set of stress tests. If the radius is smaller, there is a possibility that a transport situation is not covered by the cones of clustering.

$$\varepsilon = \min_{s_i, s_j \in S} d(s_i, s_j) \tag{5.1}$$

Network	End points	Pipelines	Pipelines scaled
Simple H-net	1414	1535	307
Advanced H-net	1140	655	164
Triangular network	258	7170	24
Shopping cart network	2519	19248	160

Table 5.1: Criteria ε for every network and type of stress test

The criteria found for each network and each type of stress test is given in Table 5.1. The values are found by Equation (5.1) and rounded down to an integer. It is rounded down, because there should not be areas which coincide with two cones. In the last column, the criteria for stress test of pipelines are scaled by dividing the criteria with the maximum pipeline length. It is useful to know information of the criteria in order to get the criteria for a new network.

If the pipeline lengths of a network are multiplied with a factor α, the criteria of the end points do not change, because the quadratic form distance has a parametrisation in which all transport distances are divided by the maximal transport distance. The criteria of the pipelines are α times as large, because in the weighted $\mathscr{L}_1$-norm there is no division done. However, this does not hold if only one pipeline length is changed.

The values of possible relations to the criteria of similarity are given in Table D.1 on page 65. The relation found in the network properties is the Euler's Polyhedron Theorem [2]:

Theorem (Euler's Polyhedron Theorem)**.** *For a graph $G = (V, E)$ the following relation holds:*

$$V - E + F = 2 \tag{5.2}$$

V is the number of the nodes in graph G, E is the number of edges and F is the number of faces in the graph.

A face is the connected area of the complement of a graph [6]. Only one face is unbounded, which is called the external face. Every plane has at least the external face. There are loops in the network if there are more than one faces in the plane in which the graph lies.

The theorem can be applied for gas networks if the following definitions are given:

- the number of nodes V is the number of the internal points: all network points excluding the end points,
- the number of edges E are the number of the pipelines,
- and the number of faces F are the number of loops plus one.

This theorem is thus applicable for gas network, which is useful for finding the number of loops in the network.

The calculations to find other relations to the criteria are found in Table D.2 on page 66. However, no trends are found in the data of the network. For example, there is no relation between the two radii of the cones of clustering and neither a relation of the radius to the pipeline length, number of edges and number of nodes. Finally, there is no relation found to the total flow.

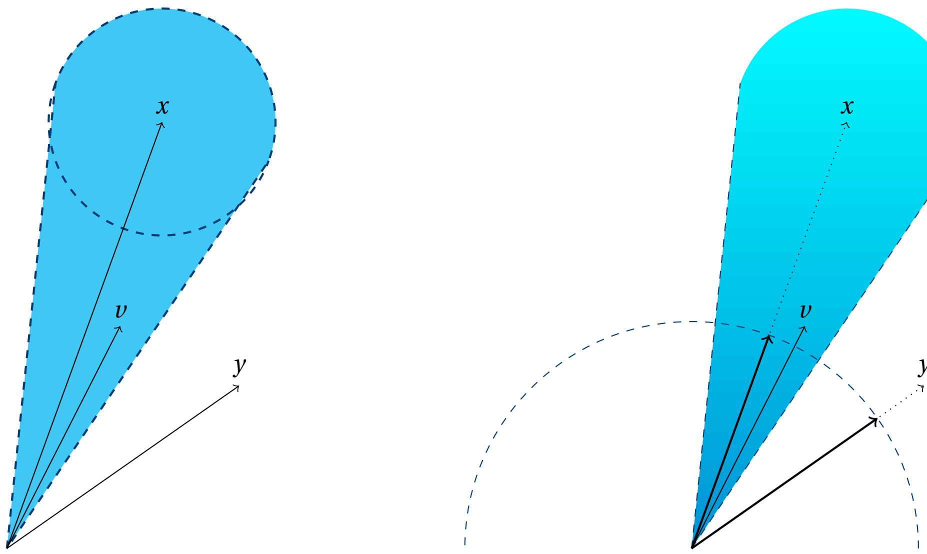
5.2. Classification of transport situations

To measure the similarities between the two distance functions, two methods are discussed below. Again, the stress tests found in MCA for each network is used. Then there are scenarios randomly generated with MCA. These scenarios are randomly drawn from the uniform distribution, but not all scenarios are feasible. For example in the simple H-net, at both entries the upper bounds are 100, but the exits have upper bound 80 and 90. So the transport situation where every end bound have maximal capacity is not feasible. Therefore the set of feasible draws is not uniform in the end.

For stress test vector x the normalisation in Equation (5.3a) is used for end points. Equation (5.3b) is used for pipelines. The normalisation for the stress tests is done such that the clustering is done properly. If the vectors are not normalised and then the distance is measured, problems arise as can be seen in Figure 5.1a. Scenario v is in the cone of stress test x, but this vector is closer to stress test y. If the stress tests are normalised, which is shown in Figure 5.1b, the distance from scenario vector v to x is smaller than the distance to y.

$$x_{\text{norm}} = \frac{1}{QFD_{\mathbf{A}}(x,x)}x \tag{5.3a}$$

$$x_{\text{norm}} = \frac{1}{\mathscr{L}_1^L(x,x)}x \tag{5.3b}$$



(a) Cone of clustering

(b) Problem of clustering without normalisation

Figure 5.1: Clustering transport situations

Network	Number of loops	Number similar of 500 cases	Similarity percentage	Switched cases
Simple H-net	0	419	83.8	81
Advanced H-net	0	405	81.0	49
Triangular network	1	306	72.0	41
Shopping cart network	2	98	19.6	0

Table 5.2: Similarity of clustering of the different types of stress tests for 500 scenarios

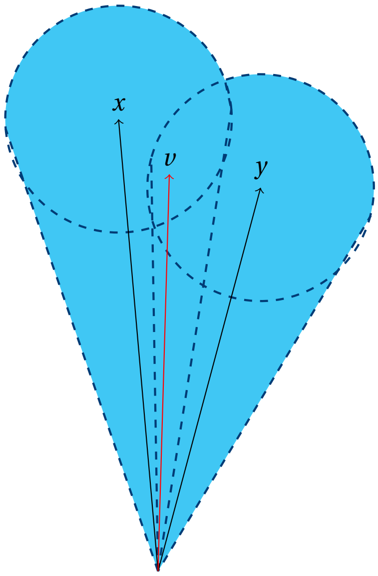
It is seen in Table 5.2 that if the number of loops increases and the network is more complex, the number of similar classifications decreases. So the more complex the network, the more the classifications of the scenarios diverge in the end point and flow representations. The last column 'switched cases' are the scenarios that for example are closest to stress test A and second closest to B for one representation and the other way around for the other representation.

In the table the results are given where the nearest stress test vector is used for the classification of a randomly generated scenario. However, there is a possibility that one scenario is in two cones, but is closer to one stress test than another. Then it is preferred to know in which cones this scenario lies. An illustration is given in Figure 5.2. In Figure 5.2a it is seen that vector v is closer to x than to y, but v is in both cones.

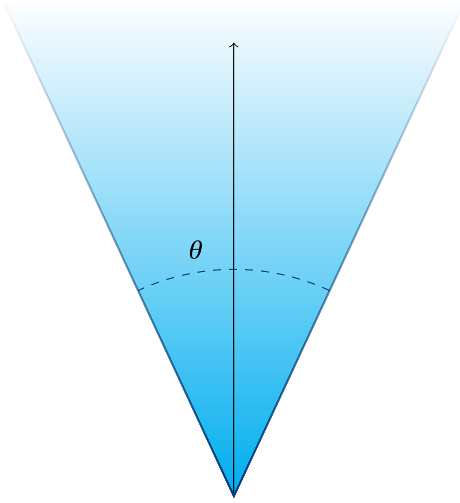
In Figure 5.2b, a stress test with its cone is shown. The angle of the cone can be calculated, because the radius of the cone ε and the length of the scenario vector are known (the vectors are normalised). Therefore the angle θ can be calculated as in Equation (5.4).

$$\tan\theta = \varepsilon \tag{5.4}$$

Now the angle θ of the cone of clustering is known, so the randomly drawn scenarios are in the cone of clustering of some stress test if the angle with that stress test is less than θ. The general equation for the angle



(a) Transport situation in two cones



(b) The angle of the cluster

Figure 5.2: Method for finding all cones for scenarios

between two vectors x, y is given in Equation 5.5. However, the $\mathscr{L}_1$-norm and the weighted $\mathscr{L}_1$-norm do not induce an inner product space. Actually, an inner product that induces a norm exists if and only if that norm satisfies the parallelogram law (see Equation (5.6)) [3]. A counter example for the (weighted) $\mathscr{L}_1$-norm and proofs of the $QFD_{\mathbf{A}}$ and the weighted $\mathscr{L}_2$-norm can be found in Section C.4 on page 63. The proofs are found by using the definition of an inner product (see below) [3].

$$\cos\varphi = \frac{\langle x, y\rangle}{\|x\|\|y\|} \tag{5.5}$$

Theorem (Parallelogram law)**.** *Let $\|\cdot\|$ be a norm. Two quantities x, y satisfy the parallelogram law if*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{5.6}$$

Definition (Inner product)**.** *For every $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ the following holds:*

1. *Non-negativity: $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0 \iff x = 0$*
2. *Linearity: $\langle z, \alpha x + \beta y\rangle = \alpha\langle z, x\rangle + \beta\langle z, y\rangle$*
3. *Symmetry: $\langle x, y\rangle = \langle y, x\rangle$*

To find the angle between the vectors x, y of the end point representation, the following equation is used:

$$\cos\theta = \frac{x^T \mathbf{A} y}{\sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y}} \tag{5.7}$$

The equation to find the angle between the vectors x, y with weight vector w is shown below.

$$\cos\theta = \frac{\sum_i w_i^2 x_i y_i}{\sqrt{\sum_i (w_i x_i)^2}\sqrt{\sum_i (w_i y_i)^2}} \tag{5.8}$$

In Table 5.3 results of the comparison of classification are given. Three stress tests are found for the Simple H-network, two for the Advanced H-network, four for the Triangular network and nine for the Shopping cart network. It appears that these stress tests do not represent all other scenarios or there is some error in the method. In Table 5.4 it is seen that for many of the scenarios no classification is found. In Table 5.3 the number of equal classes is seen for both representations. For the Simple H-network, all scenarios which are classified have exactly one equal class. There are no scenarios in this network that are in more than one cone, so every scenario which is clustered has the same classification in both representations. For the Advanced H-network there are 18 cases which have a completely different classification for the two methods. Most of the scenarios have at least one cone in common of both representations. It can be seen that the networks with a loop do not have many corresponding classes.

Network	# classes	0 equal class	1 equal class	2 equal class	>2 equal class
Simple H-network	3	0	3556	0	0
Advanced H-network	2	18	8464	1419	12
Triangular network	4	0	49	0	0
Shopping cart network	9	0	0	0	0

Table 5.3: Comparison of the classification for the two representations (10000 samples)

Network	End points	Pipelines
Simple H-network	4454	6444
Advanced H-network	68	5
Triangular network	9950	9914
Shopping cart network	10000	10000

Table 5.4: The number of scenarios that cannot be classified (10000 samples)

5.3. Addition of the pipeline diameter

A possible improvement of the weight of the norm is to add the diameter of pipelines. The term for this improvement is explained in Section 4.3. The diameter for the H-networks is for the middle $0.6m$ and the other pipelines have diameter $1.189m$. It is seen in Table 5.5 that different criteria are found. Mostly the criteria are increased, except for the Triangular network. In this network pipeline AC is $0.6m$ in diameter and the other pipelines have diameter $1.189m$.

Network	Criteria pipelines
Simple H-network	7012
Advanced H-network	713
Triangular network	1423
Shopping cart network	39445

Table 5.5: Criteria of reduction with $\mathscr{L}_1^{LD}$

More results should be found if the classification method is improved.

5.4. Conclusions

The radius of the cone of clustering, i.e. the criteria of similarity are different in each network and each representation of the scenarios. The relations of the network to the radius of the cone of clustering are useful to know such that a new radius can be calculated if the network changes. There are two relations found: Euler's polyhedron theorem and the relation of multiplying all pipeline lengths or transport distances with a certain value. The first relation has no correlation to the criterion of similarity, but only the relation between the number of nodes, edges and loops in the network. The second relation applies only if all transport distances are scaled with the same number. If this distance is changed with a parameter α, the criterion for the end point representation does not change, because the the parametrisation matrix of the QFD is established by dividing the lengths by the maximal transport distance. The criterion for the flow representation is scaled by α if all pipeline lengths are scaled by α.

Reducing the set of stress tests, the angle between stress tests needs to be calculated. The calculation of this angle an inner product is needed. However there exists no inner product for the weighted $\mathscr{L}_1$. Therefore, the weighted $\mathscr{L}_2$ is used instead of the $\mathscr{L}_1$-norm, because this norm induces an inner product space. Either this method is not correct or the data which is derived from MCA does not match for the classification method, because for the networks with loops no classification is found.

6

Conclusions

6.1. Summary of the results

Gasunie Transport Services is required to secure the supply of the natural gas in the Netherlands. The planning of the network involves testing to see if all severe scenarios are feasible. The calculations are time consuming, therefore the set of scenarios should be reduced to a set of scenarios which represent all transport scenarios. It is assumed that if the most severe scenarios, which are called stress tests, are feasible, the less severe ones are feasible as well. The effort to transport gas from one location to another through pipelines is given by the power equation:

$$P = Q\Delta p \qquad \text{(2.1, revisited)}$$

P is the power, Q is the flow and Δp is the pressure drop. The pressure drop equation contains non-linear terms, e.g. L/D^5 and $Q \cdot |Q|$. In the first term the parameter L is the pipeline length and parameter D is the diameter. The transport moment $L^T Q$ is an approximation of the effort and is used to measure the severeness of a scenario.

The current reduction method makes use of stress tests based on capacities on end points to represent the scenarios ('end points' is the name for entry, exit or storage points). In this representation, the bounds of the capacities on the end points are known as well as the transport distances between the end points. The network itself is treated as black box. The stress tests in this representation are obtained by a specific algorithm, which makes use of anchor points such that various transport directions are taken into account. The set of stress tests can be reduced by using the quadratic form distance to measure the difference between the stress tests. All scenarios within a specific radius or those which are less severe to certain stress tests can be grouped to that stress test. The QFD calculates the difference between scenario x and y as follows:

$$QFD_{\mathbf{A}}(x,y) = \sqrt{(x-y)^T \mathbf{A}(x-y)} \qquad \text{(QFD}_A\text{, revisited)}$$
$$\text{with } a_{ij} = 1 - d_{ij}/d_{\max}$$

Here d_{ij} is the transport distance in kilometres between end point i and end point j and $d_{\max} = \max_{ij} d_{ij}$. The entries of matrix $\mathbf{A}$ are defined by a_{ij}.

The current reduction method does not require many calculations, but the distance to measure the difference between scenarios is implicit, because of the 'black box' and the physical part of the matrix $\mathbf{A}$ is not easily adaptable.

To overcome these problems, a flow representation can be introduced. This representation defines flows through all pipeline segments in a unique way, i.e. by choosing the optimal flow pattern in terms of transport load. The transport moment for this representation is a linearisation of the power equation. The transport moment formula becomes as in Equation (3.2).

$$T = L^T Q \qquad \text{(3.2, revisited)}$$

Various norms can be used to compare scenarios in this flow representation. The weighted $\mathscr{L}_1$-norm and the weighted $\mathscr{L}_2$-norm seem the most promising:

$$\mathscr{L}_1^L(x,y) = \sum_i L_i |x_i - y_i| \qquad \text{(Weighted } \mathscr{L}_1\text{-norm, revisited)}$$

$$\mathscr{L}_2^L(x,y) = \sqrt{\sum_i L_i^2 (x_i - y_i)^2} \qquad \text{(Weighted } \mathscr{L}_2\text{-norm, revisited)}$$

Both norms satisfy the requirements for the comparison of stress tests based on flows through pipelines, but there is no inner product which induce the weighted $\mathscr{L}_1$-norm. The inner product is needed to find the angle between two stress tests. Therefore, the weighted $\mathscr{L}_2$-norm may be more suitable.

The comparison of the two representations, the current representation from the end point perspective and the method introduced in this research of the pipeline point of view is done by comparing their classification. The stress tests obtained from the program MCA are considered to be known. Randomly drawn realistic stress tests can be classified into the cones of clustering of the stress tests for both representations. This classification of both representations is compared. Unfortunately, no clear conclusions can be drawn from these results.

A better comparison for these two representations may determine the set of stress test for these methods and make the final hydraulic calculations for these stress tests. If one method does not find the failing scenario and the other one does, the second method is 'better' than the first one. If both methods find all the possible failing scenarios for every gas network, the one with the least amount of stress tests may be preferable.

Furthermore, in this report methods have been examined to generate stress tests based on flows through pipelines. To find these stress tests directly from network information, non-linear programming is used. However, some problems arose. For example the solution found has much gas flowing around in loops, the shortest path is not used for the flow pattern and only one solution is found instead of a set of stress tests. In Section 6.3 possible solutions to this problem are given. In this report, the stress tests based on flows through pipelines are used from MCA, which calculates these stress tests from the stress tests based on capacities at end points.

Without knowing which method performs best according to the hydraulic calculations, a general conclusion can be given. First of all the method of end points has worked well in practice. There are no precalculations needed for generating the stress tests and there is not much information needed beforehand. However, the parametrisation matrix in the QFD cannot be changed easily and the vector space that is created is difficult to imagine. The method of flows through pipelines needs more network information and currently, pre-calculation of the stress test algorithm is required. Despite the data and calculations, this reduction method is closer to the physics: the transport moment is the approximation of the transportation effort where the pressure drop is linearised to the pipeline length. The weighted $\mathscr{L}_1$-norm and weighted $\mathscr{L}_2$-norm are easily adapted. For example, first the weight was taken as the pipeline length, but instead the combination of the pipeline length and the diameter can be taken as the weight. The combination of pipeline length and diameter (L/D^5) is closer to the pressure drop equation. So there is a great potential in the method with the flow representation.

6.2. Discussion

The transport moment is an approximation of the gas transport load. The amount of flow, transport distance and diameter are included in the reduction methods and some results are drawn in this report. A concept for including the pressure drop is written in Section 4.3, but other components are important as well. Components such as blending stations and different types of gas qualities also have an impact on the transport load. Nitrogen stations are often used to get a lower gas quality, but there is not a infinite amount of nitrogen available, so that troubles the process. Moreover, it is assumed that a large amount of gas has a linear relation to the transport moment: the higher the amount of gas the higher the transport moment. It is not investigated what the transport load for a scenario with a low amount of gas is.

The random scenarios which are drawn for the classification of scenarios to the stress tests are mostly non-severe scenarios. The capacity on end points are drawn from a uniform distribution, but to get a feasible transport situation, all capacities will probably not be uniformly chosen. The scenarios which have capacities close to the boundaries are important to evaluate as well.

6.3. Future work

This research project may have raised more questions than it has solved. Other projects may be initiated starting from the theory and results of this project. The subjects of the future work are listed below.

- Most importantly, the verification of the two methods of reduction must be done differently. When the hydraulic calculations are done, a better comparison of these methods can be done. The results of the clustering with these calculations can be used as reference. The representation which clustering is most similar to this reference clustering should be the best representation, because it is closest to the reality. Furthermore it is preferred to have a stable set of stress tests, so when a component of the network is changed, the set should not change much.
- There is potential to get an algorithm which computes stress tests based on pipelines directly from the network. Some solutions have been tested in this report. Knowledge of combinatorial optimisation is required to investigate this problem further. The vertices of the polytope or half spaces of the problem can be used as stress tests. In this report a Matlab function is used which not always finds a feasible solution. Other programs can be used to solve non-linear problems, but there was not enough time to try these for this thesis.
- For the reduction of the set of stress tests based on pipelines, the weight for the weighted $\mathscr{L}_1$-norm is chosen to be the pipeline length or a combination of the pipeline length and diameter. Other weights may be used as well. There are other parameters which contribute to the transport load, for example the pressure drop. At Gasunie, the models mostly run with pressure. In this project, it is assumed that there is no pressure for simplicity. In Section 4.3 a description is given for adding pressure drop to the model. When this parameter is added, Gasunie has a better view of the results.
- The effect of the loops in a network can be an interesting subject as well. In this report it is seen that the representations diverge more if loops are involved. The number of similarities of the classifications decrease as the number of loops are increased.

Bibliography

[1] Archipel. Gasuniegebouw, 1993. URL http://archipelvzw.be/media/img/original/9698_gasuniegebouw.jpg. File: FrontLow.png, Accessed 2-March-2017.

[2] Béla Bollobás. *Modern graph theory*, volume 184. Springer Science & Business Media, 2013. page 22.

[3] James V. Burke. Lecture notes hilbert spaces, November 2015.

[4] Emile J. L. Chappin. Aardgas en distributie, 2016. URL http://eduweb.eeni.tbm.tudelft.nl/TB141E/?aardgas-transport. Online; accessed 4-April-2017.

[5] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization*. Wiley Interscience, 1998. page 12.

[6] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization*. Wiley Interscience, 1998. page 96.

[7] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization*. Wiley Interscience, 1998. page 325-335.

[8] William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization*. Wiley Interscience, 1998. page 204-206.

[9] Michel Marie Deza and Elena Deza. Encyclopedia of distances. In *Encyclopedia of Distances*, pages 1-583. Springer, 2009.

[10] Gasunie. Netwerkkaart uk, 2016. Received on intranet Gasunie, accessed 4-April-2017.

[11] Gasunie. Blending station, 2017. Online; accessed 4-April-2017.

[12] Gasunie. Jaarverslag 2016, 2017. URL http://report2016.gasunie.nl/. Online; accessed 4-April-2017.

[13] Kimberley Lindenberg. Comparing severe gas transport situations through the network. Master's thesis, TU Delft, 2015.

[14] Mathworks. Find minimum of constrained nonlinear multivariable function, 2017. URL https://nl.mathworks.com/help/optim/ug/fmincon.html. Online; accessed on 20-July-2017.

[15] Gasunie Transport Services. Entry- & exitcapaciteit, 2016. URL https://www.gasunietransportservices.nl/over-gts/gastransport/entry-exitcapaciteit. Online; accessed 27-March-2017.

[16] Gasunie Transport Services. Shipper worden, 2017. URL https://www.gasunietransportservices.nl/shippers/klant-worden/shipper-worden. Online; accessed 24-March-2017.

[17] Tomáš Skopal, Tomáš Bartoš, and Jakub Lokoč. On (not) indexing quadratic form distance by metric access methods. In *Proceedings of the 14th International Conference on Extending Database Technology*, pages 249-258. ACM, 2011.

[18] Jarig J. Steringa and Marco Hoogwerf. Generating network stress tests for hydraulic evaluation. *SIMONE Congress*, 13, apr 2016.

[19] Jarig J. Steringa, Marco Hoogwerf, Harry Dijkhuis, et al. A systematic approach to transmission stress tests in entry-exit systems. In *PSIG Annual Meeting*. Pipeline Simulation Interest Group, 2015.

[20] Universiteitsfonds. Tu delft ewi, 2016. URL https://i0.wp.com/universiteitsfonds.tudelft.nl/wp-content/uploads/2016/06/DF9E665F114F45CCAFB2C73D78F8173B.jpg?fit=1600%2C1067. File: FrontLow.png, 2-March-2017.

[21] Tom van der Hoeven. *Math in Gas and the art of linearization*. Energy Delta Institute, 2004.

[22] Eric W. Weisstein. Positive definite matrix, 2017. URL http://mathworld.wolfram.com/PositiveDefiniteMatrix.html. Online; accessed 28-March-2017.

A

Assumptions

- A network has no negative transport distance.
- The bounds on entry and exit points are known.
- The contracts with the shippers are not violated, this means that
  - there is a balance of feed in and take off of gas in the network, and
  - the minimal and maximal feed in and take off of gas is known and will not be violated.
- The length of the pipe and the amount of gas in the transport network are the most important quantities for measuring the severity of the gas transportation (Chapter 2).
- The gas in a pipeline can flow into two directions and the quantities are the same in both directions. Thus the transport moment of the flow through a pipeline in one direction equals the transport moment of the flow in the other direction with the same flow through the same pipeline.
- When methods work for the example networks, it will work in any gas transport network. Excluding the following:
  - Compressor stations
  - Blending stations
  - Pressure regulation station
  - Nitrogen injection stations
- The more severe transport situations dominate the less severe transport situations. That means that if the severe situations are feasible, the less severe ones are also feasible. Moreover, this holds for the entire cone as described in Section 4.4.

B

Figures

B.1. Generation of stress tests based on pipelines

The shopping cart network with the generated scenarios are given on pages 52-54.

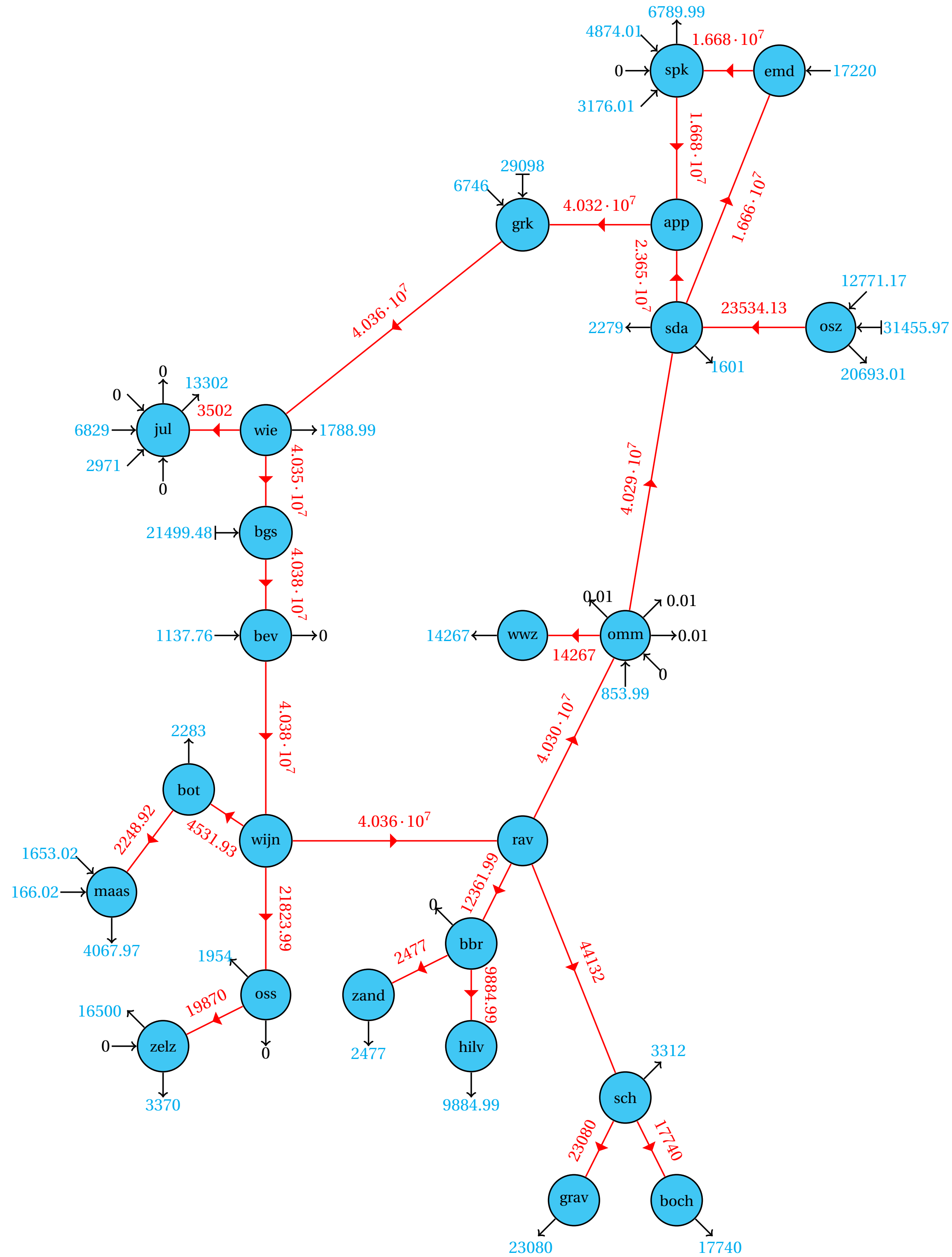


Figure B.1: Shopping cart network: scenario with maximal transport moment obtained by linear constraints

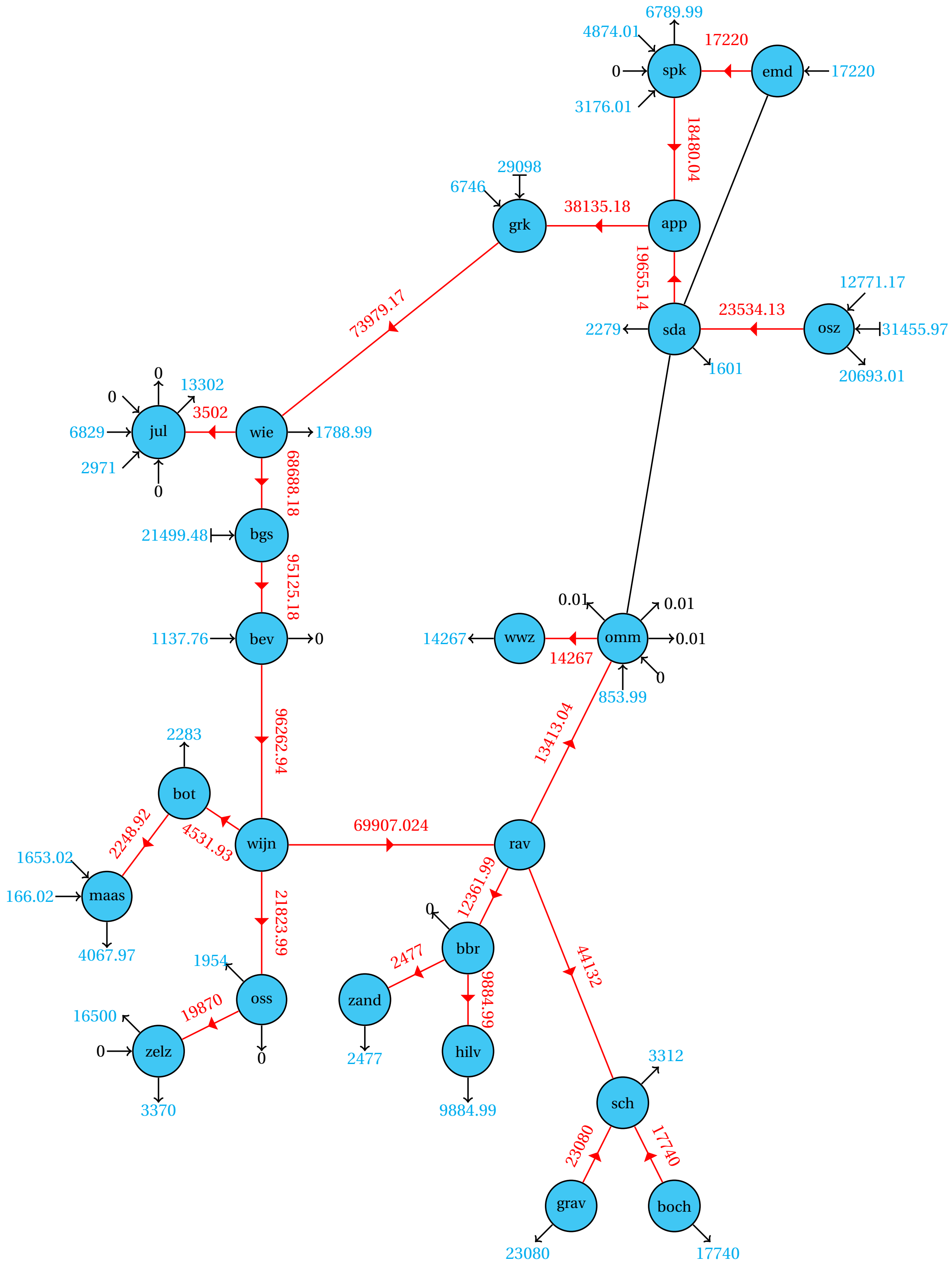


Figure B.2: Shopping cart network: manually reducing flow in loops

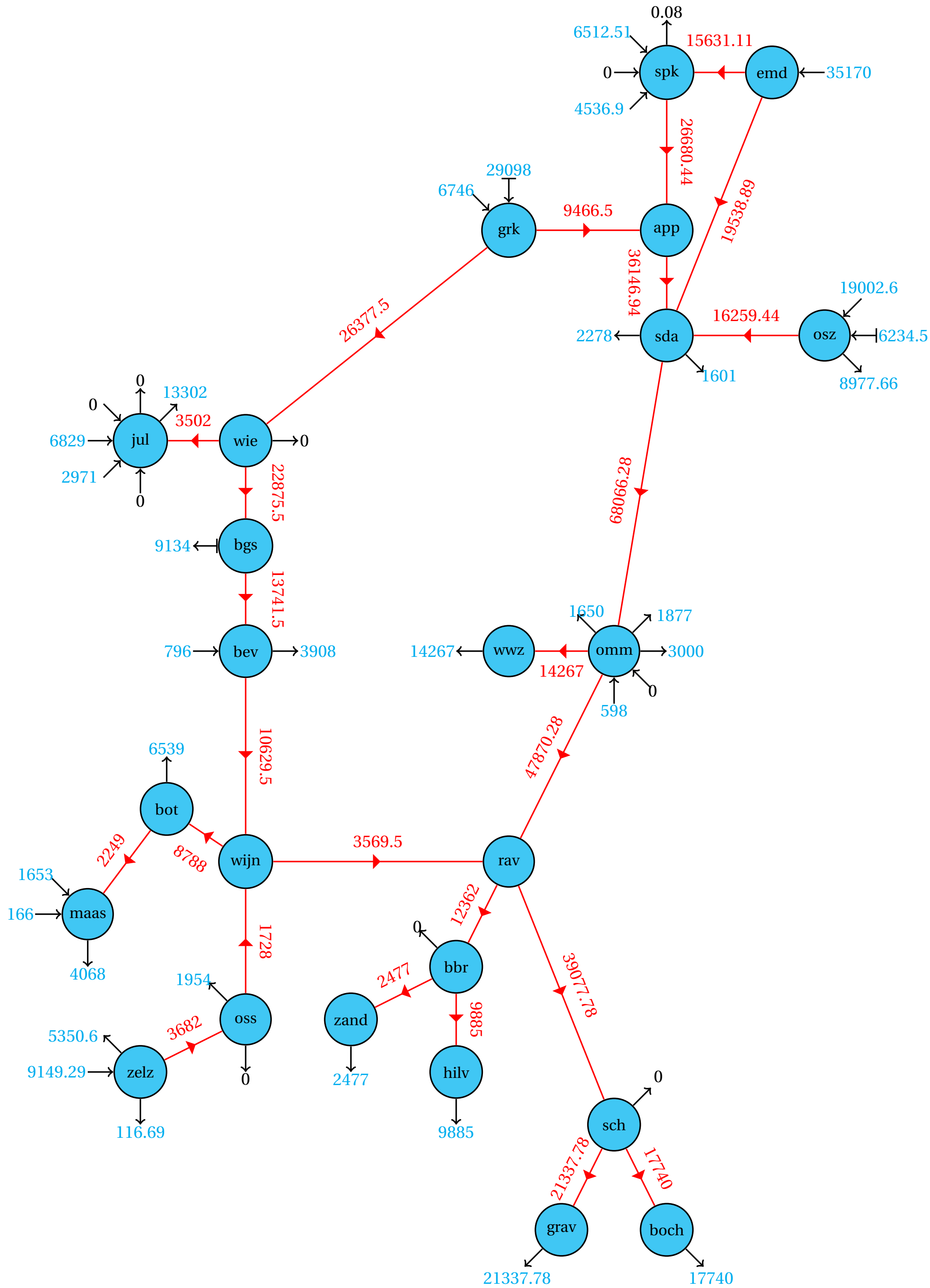


Figure B.3: Shopping cart network: scenario generated with bounds on flow through the pipelines

B.2. Comparison (semi-) metrics and stress tests

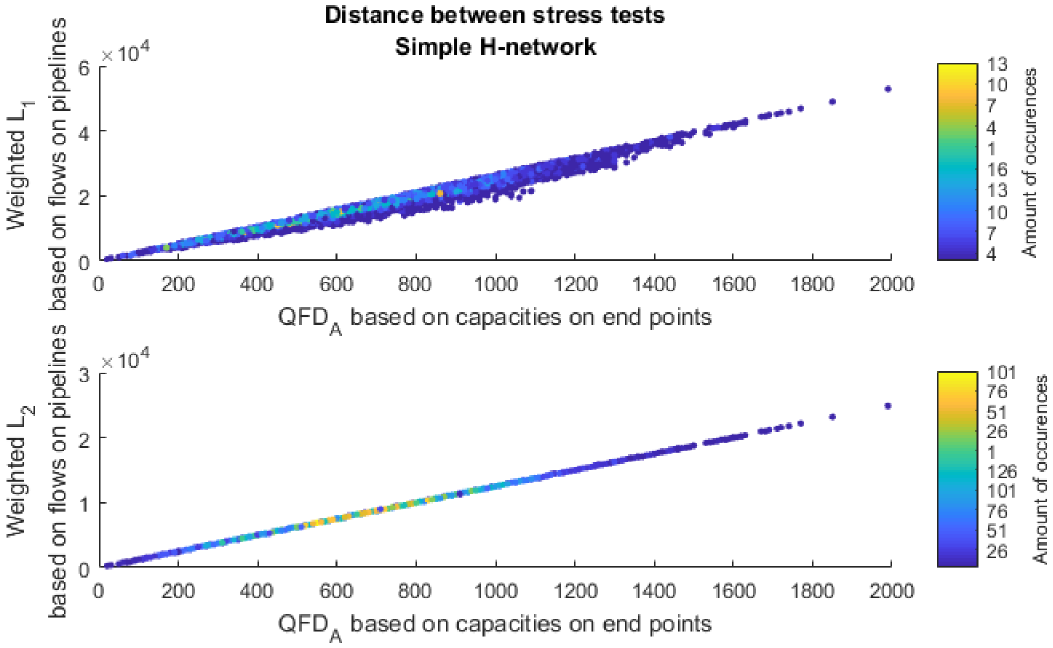


Figure B.4: Distance stress tests of the simple H-network

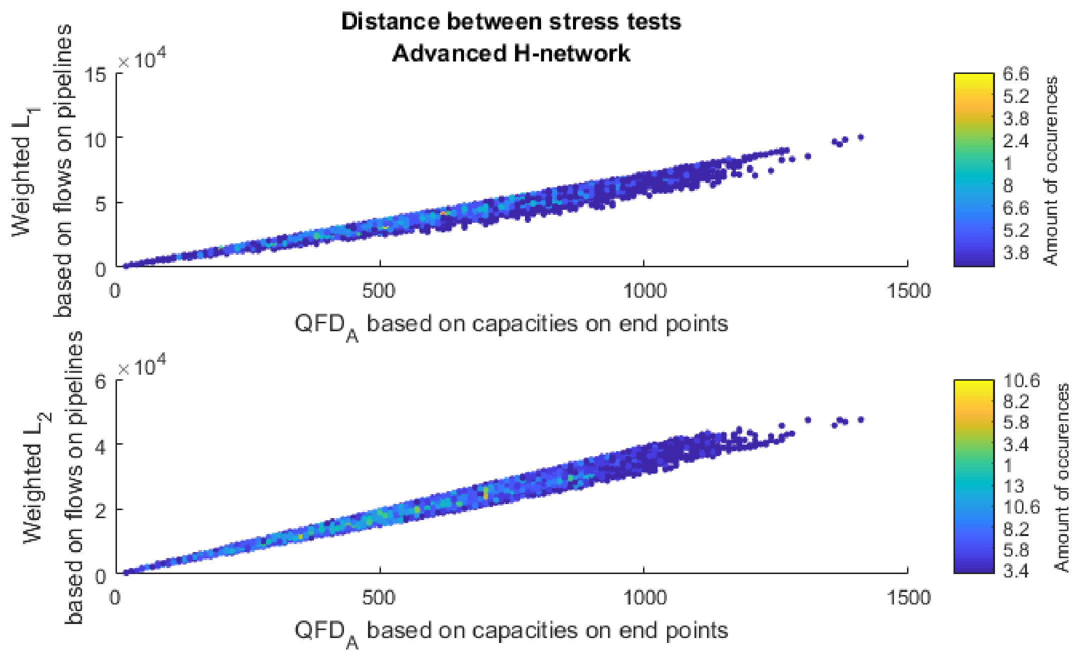


Figure B.5: Distance stress tests of the advanced H-network

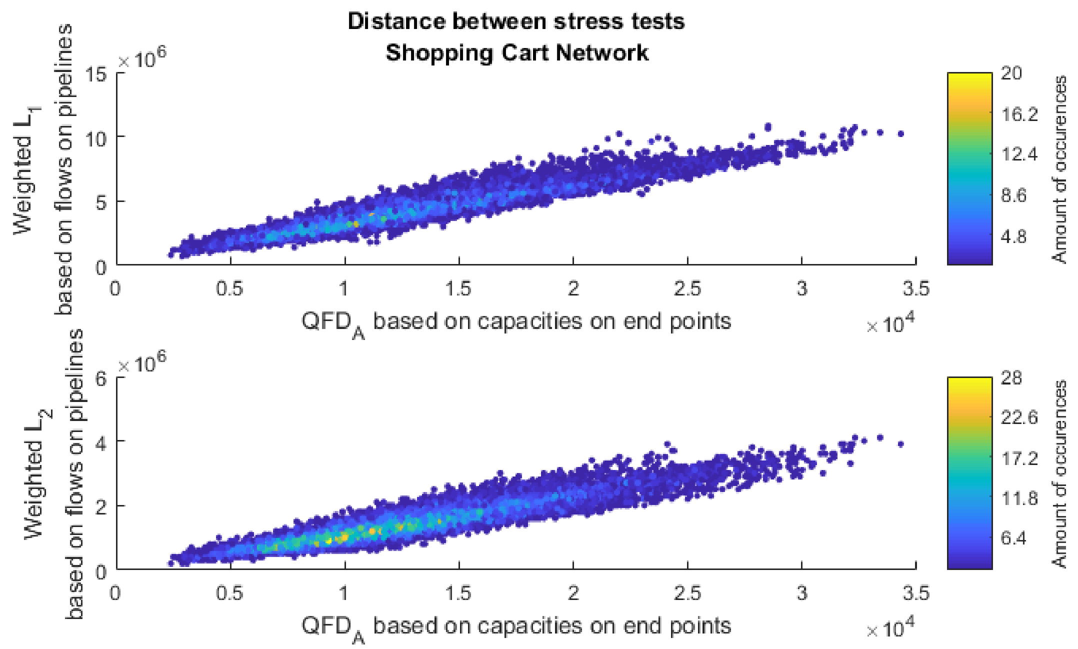


Figure B.6: Distance stress tests of the shopping cart network

C

Proofs and examples

C.1. Example QFD on advanced H-network

In this subsection, the distance of stress tests of the advanced H-network in Section 4.1.2 is calculated. The illustration of this network is given on page 5. The stress tests for this network are found in Section 3.1.2 and these are again given below in Equation (C.1). The order of capacities in the stress tests is N_A, N_B, X_A, X_B, X_C, X_D.

$$s_A = \begin{pmatrix} 100 \\ 90 \\ -0 \\ -20 \\ -80 \\ -90 \end{pmatrix}, \quad s_B = \begin{pmatrix} 0 \\ 100 \\ -20 \\ -0 \\ -80 \\ -0 \end{pmatrix}, \quad s_C = \begin{pmatrix} 10 \\ 100 \\ -20 \\ -0 \\ -0 \\ -90 \end{pmatrix} \tag{C.1}$$

Three distances can be computed from these stress tests. The matrices used for the QFD are:

$$D = \begin{pmatrix} 0 & 7 & 0 & 7 & 7 & 8 \\ 7 & 0 & 7 & 0 & 6 & 7 \\ 0 & 7 & 0 & 7 & 7 & 8 \\ 7 & 0 & 7 & 0 & 6 & 7 \\ 7 & 6 & 7 & 6 & 0 & 7 \\ 8 & 7 & 8 & 7 & 7 & 0 \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 1 & \frac{1}{8} & 1 & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{8} & 1 & \frac{1}{8} & 1 & \frac{1}{4} & \frac{1}{8} \\ 1 & \frac{1}{8} & 1 & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{8} & 1 & \frac{1}{8} & 1 & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & 1 & \frac{1}{8} \\ 0 & \frac{1}{8} & 0 & \frac{1}{8} & \frac{1}{8} & 1 \end{pmatrix}$$

Now, the distances are calculated:

$$\begin{aligned} QFD_{\mathbf{A}}(s_A, s_B) &= \sqrt{(s_A - s_B)^T \, \mathbf{A}(s_A - S_B)} \\ &= \sqrt{23175} \\ &= 15\sqrt{103} \\ &\approx 152.2334 \end{aligned}$$

$$\begin{aligned} QFD_{\mathbf{A}}(s_A, s_C) &= \sqrt{(s_A - s_C)^T \, \mathbf{A}(s_A - S_C)} \\ &= \sqrt{17575} \\ &= 5\sqrt{703} \\ &\approx 132.5707 \end{aligned}$$

$$\begin{aligned} QFD_{\mathbf{A}}(s_B, s_C) &= \sqrt{(s_B - s_C)^T\,\mathbf{A}(s_B - S_C)} \\ &= \sqrt{13000} \\ &= 10\sqrt{130} \\ &\approx 114.0175 \end{aligned}$$

It follows that $QFD_{\mathbf{A}}(s_A, s_B) > QFD_{\mathbf{A}}(s_A, s_C) > QFD_{\mathbf{A}}(s_B, s_C)$, so the scenario generated from anchor point A is more similar to the scenario generated from C than from stress test with anchor point B.

C.2. Proof QFD is semi-metric

$QFD_{\mathbf{A}}(u,v) \geq 0$

If **A** is a SSPD-matrix (or SPD-matrix):

$$
\begin{aligned}
(u-v)^T \mathbf{A}(u-v) &\geq 0 && (\mathbf{A} \text{ is SSPD}) \\
\sqrt{(u-v)^T \mathbf{A}(u-v)} &\geq 0 \\
QFD_{\mathbf{A}}(u,v) &\geq 0
\end{aligned}
$$

So non-negativity holds for the quadratic form distance.

$QFD_{\mathbf{A}}(u,v) = QFD_{\mathbf{A}}(v,u)$

If **A** is a SSPD-matrix (or SPD-matrix), then

$$
\begin{aligned}
QFD_{\mathbf{A}}(u,v) &= \sqrt{(u-v)^T \mathbf{A}(u-v)} \\
&= \sqrt{(v-u)^T \mathbf{A}(v-u)} \\
&= QFD_{\mathbf{A}}(v,u)
\end{aligned}
$$

So symmetry holds.

$QFD_{\mathbf{A}}(u,v) \leq QFD_{\mathbf{A}}(u,w) + QFD_{\mathbf{A}}(w,v)$

If **A** is a SSPD-matrix (or a SPD-matrix), then the theorem of Cauchy-Schwarz holds in this case (Equation (C.2)).

$$
x^T \mathbf{A} y \leq \sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y} \tag{C.2}
$$

The proof of this special case of the theorem is given below. Here it is used that a SSPD-matrix (or SPD) can be written as $\mathbf{A} = \mathbf{B}^T \mathbf{B}$ [22].

$$
\begin{aligned}
0 &\leq \left(\sqrt{x^T \mathbf{A} x}\,\mathbf{B} y - \sqrt{y^T \mathbf{A} y}\,\mathbf{B} x\right)^T \left(\sqrt{x^T \mathbf{A} x}\,\mathbf{B} y - \sqrt{y^T \mathbf{A} y}\,\mathbf{B} x\right) \\
&= x^T \mathbf{A} x (\mathbf{B} y)^T \mathbf{B} y - 2\sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y} (\mathbf{B} x)^T \mathbf{B} y + y^T \mathbf{A} y (\mathbf{B} x)^T \mathbf{B} x \\
&= x^T \mathbf{A} x y^T \mathbf{B}^T \mathbf{B} y - 2\sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y} x^T \mathbf{B}^T \mathbf{B} y + y^T \mathbf{A} y x^T \mathbf{B}^T \mathbf{B} x \\
&= x^T \mathbf{A} x y^T \mathbf{A} y - 2\sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y x^T \mathbf{A} y} + y^T \mathbf{A} y x^T \mathbf{A} x \\
&= 2 x^T \mathbf{A} x y^T \mathbf{A} y - 2\sqrt{x^T \mathbf{A} x}\sqrt{y^T \mathbf{A} y x^T \mathbf{A} y}
\end{aligned}
$$

This can be used for the triangular inequality $d(u,v) \leq d(u,w) + d(w,v)$:

$$
\begin{aligned}
(QFD_{\mathbf{A}}(u,v))^2 &= (u-v)^T \mathbf{A}(u-v) \\
&= (u-w+w-v)^T \mathbf{A}(u-w+w-v) \\
&= (u-w)^T \mathbf{A}(u-w) + 2(u-w)^T \mathbf{A}(w-v) + (w-v)^T \mathbf{A}(w-v) && (\mathbf{A} \text{ is symmetric}) \\
&\leq (u-w)^T \mathbf{A}(u-w) + 2\sqrt{(u-w)^T \mathbf{A}(u-w)}\sqrt{(w-v)^T \mathbf{A}(w-v)} + (w-v)^T \mathbf{A}(w-v) \\
& && (\text{Cauchy-Schwarz, see Equation (C.2)}) \\
&= \left(\sqrt{(u-w)^T \mathbf{A}(u-w)} + \sqrt{(w-v)^T \mathbf{A}(w-v)}\right)^2
\end{aligned}
$$

So the triangular inequality holds for the QFD with a symmetric (semi-)positive definite matrix.

$QFD_{\mathbf{A}}(u,v) = 0 \iff u = v$

If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix, then $QFD_{\mathbf{A}}(u,v) = 0$ if and only if $u = v$ for every $u, v \in \mathbb{R}^n$. The proof is given below.

Assume $u = v$:

$$\begin{aligned} QFD_{\mathbf{A}}(u,u) &= \sqrt{(u-u)^T \, \mathbf{A}(u-u)} \\ &= 0 \end{aligned}$$

Assume $QFD_{\mathbf{A}}(u,v) = 0$, then

$$\begin{aligned} QFD_{\mathbf{A}}(u,v) &= 0 \\ &\iff \\ \sqrt{(u-v)^T \, \mathbf{A}(u-v)} &= 0 \\ &\iff \\ (u-v)^T \, \mathbf{A}(u-v) &= 0 \\ &\iff \\ (u-v)^T \, \mathbf{A}(u-v) = 0 \; \forall u, v (u \neq v) \\ \text{or } u = v \end{aligned}$$

As **A** is positive definite, the first possibility does not hold. Therefore $d(u,v) = 0 \iff u = v$. If **A** is semi-positive definite, this reason does not hold. Therefore, if **A** is semi-positive definite, $QFD_{\mathbf{A}}$ is not metric. However, the equality $d(u,u) = 0$ holds for every u, so then $QFD_{\mathbf{A}}$ is a semi-metric distance.

C.3. Proof weighted $\mathscr{L}_p$ distance is metric

$d_w(x, y) \geq 0$

$$\begin{aligned} d_w(x, y) &= \left(\sum_i \left(w_i |x_i - y_i| \right)^p \right)^{\frac{1}{p}} \\ &\geq 0 && (w \geq 0, \text{ and } |x_i - y_i| \geq 0) \end{aligned}$$

$d_w(x, y) = d_w(y, x)$

$$\begin{aligned} d_w(x, y) &= \left(\sum_i \left(w_i |x_i - y_i| \right)^p \right)^{\frac{1}{p}} \\ &= \left(\sum_i \left(w_i |y_i - x_i| \right)^p \right)^{\frac{1}{p}} \\ &= d_w(y, x) \end{aligned}$$

$d_w(x, y) = 0 \iff x = y$

$$\begin{aligned} d_w(x, y) &= 0 \\ &\iff \\ \left(\sum_i \left(w_i |x_i - y_i| \right)^p \right)^{\frac{1}{p}} &= 0 \\ &\iff \\ \sum_i \left(w_i |x_i - y_i| \right)^p &= 0 \\ &\iff && \text{(All terms are positive)} \\ \left(w_i |x_i - y_i| \right)^p &= 0 && \forall i \\ &\iff \\ w_i |x_i - y_i| &= 0 && \forall i \\ &\iff && \text{(Non-trivial cases)} \\ |x_i - y_i| &= 0 && \forall i \\ &\iff \\ x_i &= y_i && \forall i \\ &\iff \\ x &= y \end{aligned}$$

$d_w(x,y) \leq d_w(x,z) + d_w(z,y)$

$$\begin{aligned}
(d_w(x,y))^p &= \sum_i w_i^p |x_i - y_i| \\
&= \sum_i \left(w_i|x_i - y_i|\right)\left(w_i|x_i - y_i|\right)^{p-1} \\
&\leq \sum_i \left(w_i|x_i - z_i| + w_i|z_i - y_i|\right)\left(w_i|x_i - y_i|\right)^{p-1} && \text{(Triangular inequality)} \\
&\leq \left(\left(\sum_i w_i^p|x_i - z_i|^p\right)^{\frac{1}{p}} + \left(\sum_i w_i^p|z_i - y_i|^p\right)^{\frac{1}{p}}\right)\left(\sum_i \left(w_i|x_i - y_i|\right)^p\right)^{\frac{p-1}{p}} && \text{(Hölder's inequality)} \\
&= (d_w(x,z) + d_w(z,y))(d_w(x,y))^{p-1} \\
\left(d_w(x,y)\right)^p &\leq \left(d_w(x,z) + d_w(z,y)\right)(d_w(x,y))^{p-1} \\
\Rightarrow d_w(x,y) &\leq d_w(x,z) + d_w(z,y)
\end{aligned}$$

C.4. Two proofs and a counterexample inner product spaces

For the weighted $\mathscr{L}_1$-norm a counterexample is found and the proofs are given for the weighted $\mathscr{L}_2$-norm and the $QFD_{\mathbf{A}}$ are given.

Take the following functions on the interval $[0,1]$ for the weighted $\mathscr{L}_1$-norm with general weight $w \in \mathbb{R}$ as the norm.

$$f(x) = \begin{cases} 0 & x \in [0, \frac{1}{2}] \\ 1 & x \in [\frac{1}{2}, 1] \end{cases}$$
$$g(x) = \begin{cases} 1 & x \in [0, \frac{1}{2}] \\ 0 & x \in [\frac{1}{2}, 1] \end{cases}$$

Then the right hand side of the parallelogram equation (page 40) becomes:

$$\begin{aligned} \|f(x)+g(x)\|^2_{\mathscr{L}_1^w} + \|f(x)-g(x)\|^2_{\mathscr{L}_1^w} &= w^2 + w^2 \\ &= 2w^2 \end{aligned}$$

The solution from above is not equal to the left hand side:

$$\begin{aligned} 2\|f(x)\|^2_{\mathscr{L}_1^w} + 2\|g(x)\|^2_{\mathscr{L}_1^w} &= 2\left(\frac{1}{2}w\right)^2 + 2\left(\frac{1}{2}w\right)^2 \\ &= w^2 \end{aligned}$$

Therefore, the weighted $\mathscr{L}_1$-norm is not induced by an inner product.

The three requirements for an inner product are given on page 40. The $QFD_{\mathbf{A}}$ satisfies these requirements for the inner product $\langle x, y\rangle = x^T \mathbf{A} y$ if $\mathbf{A}$ is symmetric positive definite (SPD).

1. Non-negativity, for every $x \in \mathbb{R}^n$:

$$\langle x, x\rangle = x^T \mathbf{A} x \geq 0 \qquad (\mathbf{A} \text{ is SPD})$$
$$x = 0 \implies \langle x, x\rangle = 0$$
$$\langle x, x\rangle = 0 \implies x = 0 \qquad (\mathbf{A} \text{ is SPD})$$

2. Linearity, for every $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta in \mathbb{R}$

$$\begin{aligned} \langle z, \alpha x + \beta y\rangle &= z^T \mathbf{A}(\alpha x + \beta y) \\ &= z^T \mathbf{A}\alpha x + z^T \mathbf{A}\beta y \\ &= \alpha\langle z, x\rangle + \beta\langle z, y\rangle \end{aligned}$$

3. Symmetry, for every $x, y \in \mathbb{R}^n$

$$\begin{aligned} \langle x, y\rangle &= x^T \mathbf{A} y \\ &= y^T \mathbf{A} x \\ &= \langle y, x\rangle \end{aligned}$$

The weighted $\mathscr{L}_2$-norm satisfies these requirements for the inner product $\langle x, y\rangle = \sum_i (w_i x_i)(w_i y_i)$.

1. Non-negativity, for every $x \in \mathbb{R}^n$:

$$\langle x, x\rangle = \sum_i (w_i x_i)^2 \geq 0$$
$$x = 0 \implies \langle x, x\rangle = 0$$
$$\langle x, x\rangle = 0 \implies x = 0 \qquad (w \text{ is non trivial})$$

2. Linearity, for every $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$

$$\begin{aligned}\langle z, \alpha x + \beta y\rangle &= \sum_i (w_i z_i)(w_i(\alpha x_i + \beta y_i)) \\ &= \sum_i \alpha(w_i z_i)(w_i x_i) + \beta(w_i z_i)(w_i y_i) \\ &= \alpha\langle z, x\rangle + \beta\langle z, y\rangle\end{aligned}$$

3. Symmetry, for every $x, y \in \mathbb{R}^n$

$$\begin{aligned}\langle x, y\rangle &= \sum_i (w_i x_i)(w_i y_i) \\ &= (w_i y_i)(w_i x_i) \\ &= \langle y, x\rangle\end{aligned}$$

D

Tables

D.1. Data networks and the relation to criteria of similarity

The criteria of end points is defined by the variable ε_{end} and the criteria of pipelines is given by $\varepsilon_{\text{pipe}}$, which are the radii of the cones of clustering. $L_{\max}$ is the maximal pipeline length in the network and L_{tot} is the total amount of length in the network. Define the two stress tests s_1 and s_2, which satisfy Equation (5.1), see page 37. Q^i_{tot} is the total amount of flow going from entries to exits of stress test s_i, where $i \in \{1,2\}$. Q_{totdiff} is the difference between the two Q_{tot}. N is the amount of entry points, X is the amount of exit points and S is the amount of storages. V is the amount of internal points and E is the total amount of (internal) pipelines.

	Simple H-net	Advanced H-net	Triangular network	Shopping cart network
ε_{end}	1414	1140	258	2519
$\varepsilon_{\text{pipe}}$	1535	655	7170	19248
$L_{\max}$	5	4	300	120
L_{tot}	25	13	700	1209
Q^1_{tot}	$1 \cdot 10^3$	$1 \cdot 10^5$	$1 \cdot 10^3$	$6.16 \cdot 10^6$
Q^2_{tot}	$1 \cdot 10^3$	$6.2 \cdot 10^5$	$1.2 \cdot 10^3$	$8.70 \cdot 10^6$
Q_{totdiff}	0	$5.2 \cdot 10^5$	$2 \cdot 10^2$	$2.54 \cdot 10^6$
N	2	2	2	16
X	2	4	3	24
S	0	0	0	3
V	2	6	3	23
E	5	5	3	24
F	1	1	2	3

Table D.1: Relations of the networks

	Simple H-net	Advanced H-net	Triangular network	Shopping cart network
$\varepsilon_{\text{end}}/V$	707.11	190.03	86.07	109.54
$\varepsilon_{\text{pipe}}/L_{\text{max}}$	307.08	163.78	23.90	160.41
$\left(\varepsilon_{\text{pipe}}/\varepsilon_{\text{end}}\right) \cdot L_{\text{max}}$	4.61	6.96	10.80	15.71
$(\varepsilon_{\text{end}} - \varepsilon_{\text{pipe}})/\varepsilon_{\text{end}}$	-0.08	0.43	-26.77	-6.64
$(\varepsilon_{\text{end}} - \varepsilon_{\text{pipe}})/\varepsilon_{\text{end}} \cdot L_{\text{max}}$	0.78	0.86	0.91	0.94
$(\varepsilon_{\text{end}} - \varepsilon_{\text{pipe}})/\varepsilon_{\text{end}} \cdot V$	-0.17	2.55	-80.31	-152.73
$\varepsilon_{\text{end}}/E$	282.84	228.04	86.07	104.98
$\varepsilon_{\text{pipe}}/E$	307.08	131.02	2390.00	802.06
$\varepsilon_{\text{pipe}}/E \cdot L_{\text{max}}$	61.42	32.76	7.97	6.68
$\varepsilon_{\text{end}}/\varepsilon_{\text{pipe}}$	0.92	1.74	0.04	0.13
$\varepsilon_{\text{end}}/\varepsilon_{\text{pipe}} \cdot L_{\text{max}}$	1.84	10.44	0.11	3.01
$\varepsilon_{\text{end}}/\varepsilon_{pipe} \cdot V$	4.61	8.70	0.11	3.14
$(\varepsilon_{\text{end}} - \varepsilon_{\text{pipe}})/Q_{\text{totdiff}}$	-	0.00	-34.56	-0.01
$(\varepsilon_{\text{end}} - \varepsilon_{\text{pipe}})/(Q^1_{\text{tot}}/Q^2_{\text{tot}})$	-121.19	3007.44	-8294.16	-23635.93

Table D.2: Relations of the criteria

E

Definitions, Abbreviations and Symbols

Glossary

Calorific value Energy value or volume of energy by gas (in MJ/m^3) [11]. 1

End point Collective name for entry point, exit point or storage point. 9

Face Simply connected regions divided by a graph in a plane. 38

Polyhedron Set of solutions to the linear constraints of linear programming. 19

Polytope Bounded polyhedron. 19, 25, 45

Shipper Party who is recognised by the network operator of the national grid and consequently has programme responsibility [16]. 1, 9

Stress test Transport situation which is severe for the transport of gas and is within the contractual bounds of the shippers. 9, 27

Transport moment Quantity of the transport load, mostly dependent on the amount of flow through the pipes and the length of the pipes. 7, 9

Transport situation Balanced combination of the quantities on the entry and exit points, the capacity is the most common used quantity. 7

Vertex Intersections of the linear conditions of the (non-)linear programming. 25

List of Abbreviations

G-gas Groningen gas (Wobbe index ≤ 44.4). 2

GTS Gasunie Transport Services. 1, 7, 9

H-gas Natural gas with a Wobbe index ≥ 49.0. 1, 7

HTL High Pressure Grid. 1, 2, 7

L-gas Natural gas with a Wobbe index between 44.4 and 47.2. 1

LP Linear Programming. 17

MCA Multiple Case Analysis. 24, 37, 44

NLP Non-Linear Programming. 20, 26

QFD Quadratic Form Distance. 27

RTL Intermediate Pressure Grid. 1, 7

SPD Symmetric Positive Definite. 29, 59

SSPD Symmetric Semi-Positive Definite. 29, 59

TSO Transmission System Operator. 1

List of Symbols

$\mathbf{A}$ Parametrisation matrix for quadratic form distance. 28

c Capacity on an end point (entry, exit or storage point). 9

d Distance function (between end points or vectors). 9

$\mathbf{D}$ Distance matrix of end points. 28

ε Similarity number of stress tests. If the outcome of a metric distance from two stress tests is smaller than this number, the stress tests are similar. 35

f Flow through pipelines. 9, 25

L Length of a pipeline. 9, 25

lb Lower bound. 2, 17

Δp Pressure drop. 34

p_{in} Pressure at the start of the pipe. 34

p_{out} Pressure at the end of the pipe. 34

Q Gas flow. 34

τ Anchor point in the stress test algorithm. 9, 10

ub Upper bound. 2, 17