CHALMERS
UNIVERSITY OF TECHNOLOGY

Design and Analysis of a Narrowband Stepped Digital Elliptic Filter at 5 GHz

Master's thesis in Wireless, Photonics and Space Engineering

HANS SVANBERG

Department of Microtechnology and Nanoscience
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024

Master's thesis 2024

Design and Analysis of a Narrowband Stepped Digital Elliptic Filter at 5 GHz

HANS SVANBERG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Microtechnology and Nanoscience
Terahertz and Millimetre Wave Laboratory
Chalmers University of Technology
Gothenburg, Sweden 2024

Design and Analysis of a Narrowband Stepped Digital Elliptic Filter at 5 GHz

HANS SVANBERG

Supervisor: Jan Stake, Microtechnology and Nanoscience
Advisor: Anatoli Deleniv and Patrícia Pereira Bouça, Ericsson AB
Examiner: Jan Stake, Microtechnology and Nanoscience

Master's Thesis 2024
Department of Microtechnology and Nanoscience
Terahertz and Millimetre Wave Laboratory
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 (0)31-772 10 00

Typeset in LaTeX
Gothenburg, Sweden 2024

Design and Analysis of a Narrowband Stepped Digital Elliptic Filter at 5 GHz

HANS SVANBERG
Department of Microtechnology and Nanoscience
Chalmers University of Technology

Abstract

The increasing frequency and complexity of modern radio frequency (RF) systems makes more compact components essential, particularly in antenna arrays, where numerous electronic components must fit within the small spaces between antenna elements. Filters are a crucial component to every RF system, making the development of more compact filter solutions necessary. Furthermore, new and precise microfabrication methods are emerging where the cost of production is influenced by the physical size of the components.

In 1969, a narrowband bandpass filter structure, the stepped digital elliptic filter, was introduced. It has received little to no attention in subsequent research due to difficulties with tuning and fabrication. To the best of the author's knowledge, this is the most compact higher-order elliptic function bandpass filter structure consisting solely of distributed elements capable of achieving $n - 1$ transmission zeros. The unique structure also allows the resonators to be suspended without the use of dielectrics, enabling high Q values. Considering advances in microfabrication technology and the need for more compact solutions, it is now an opportune time to revisit this filter structure.

Starting from network theory and progressing to electromagnetic simulations, a filter prototype with a fractional bandwidth of 11 %, a ripple level of 10 %, and a minimum stopband attenuation of 40 dB was designed at 5 GHz, demonstrating close agreement between the simulated and theoretical results, with a return loss above 15 dB in the passband. The filter prototype was shown to be sensitive to fabrication errors, with tolerances greater than 5 µm severely degrading the return loss. Two different designs are proposed suitable for fabrication with the mentioned specifications.

In addition, filter characteristics and capacitance matrices (relating directly to the physical dimensions of the filter) for stopband attenuation levels of 30 dB, 40 dB and 60 dB, return losses of 28 dB and 14 dB, and fractional bandwidths (FBW) between 1 % and 30 % are presented, as well as the structure's dependence on scaling of the capacitance matrices.

Finally, the filter structure has been shown to be suitable for design using microfabrication, with almost any size and filter characteristics with FBWs below 30 % possible with sufficiently precise manufacturing tolerances.

Keywords: Filters, bandpass, narrowband, stepped digital elliptic filter, transmission zeros.

Acknowledgements

Thank you Ericsson, for giving me the opportunity to carry out my master's thesis at your company. Huge thanks to my supervisors at Ericsson, Anatoli Deleniv and Patrícia Pereira Bouça, for invaluable advice and guidance along the way. This thesis would not have been possible without you.

Thank you also to everyone at MC2, especially the people in the Terahertz and Millimetre Wave Laboratory, for creating such a supportive and inspiring environment during both the master's programme and the master's thesis. Your friendship, guidance, and teaching have been invaluable throughout this journey. In particular, I would like to express my deepest gratitude to my professor, Jan Stake, for his insightful feedback and support during the course of this thesis.

Hans Svanberg, Gothenburg, 2024-09-18

Nomenclature

Physics Constants

c_0	Speed of light in vacuum	$299\,792\,458\,\mathrm{m/s}$
ε_0	Permittivity in vacuum	$8.854\,187\,817\,620\,391 \times 10^{-12}\,\mathrm{F/m}$
μ_0	Permeability in vacuum	$1.256\,637\,061\,435\,917 \times 10^{-6}\,\mathrm{H/m}$

Other Symbols

C	Capacitance	[F]
L	Inductance	[H]
R	Resistance	[Ω]
G	Conductance	[S]
V	Voltage	[V]
I	Current	[A]
Z	Impedance	[Ω]
Y	Admittance	[S]
ϵ	Ripple	
RL	Return loss	
$A_{\min}$	Minimum stopband attenuation	
FBW	Fractional bandwidth	
n	Port index or total number of ports in a network	
n_{diag}	Diagonal scaling constant	
n_k	Nodal scaling constant	
f	Frequency	[Hz]
ω	Angular frequency	[rad/s]
λ	Wavelength	[1/m]
η	Intrinsic impedance	[Ω]

S_{11}	Complex reflection coefficient	
S_{21}	Complex transmission coefficient	
w	Width	[m]
b	Ground plate spacing	[m]
t	Conductor thickness	[m]
d	Length (spacing) between conductors	[m]
x	Reference length in width dimension	[m]
NA	Nodal alignment	[m]
C_{var}	Capacitance variation	[F]
C_{diff}	Mean nodal capacitance difference	[F]
l	Conductor length	[m]
$\mathbf{C}$	Capacitance matrix	[F]
$\mathbf{Y}$	Admittance matrix	[S]

Contents

List of Figures

List of Tables

1

Introduction

The electromagnetic spectrum has become increasingly congested as the demand for wireless communication and other technologies continues to grow. This surge in usage has intensified the competition for available frequencies, leading to challenges in managing and allocating this vital resource. Already in 1985 it was identified as a limited critical natural resource [1], subject to both legal and political problems, and much has changed since then. One way of solving this issue is to increase the frequency of operation to a less congested part of the electromagnetic spectrum. However, this approach alone is insufficient due to the high number of users and the limited availability of usable spectrum. Higher frequencies also bring new challenges, such as reduced signal range and increased difficulty in manufacturing components due to their smaller size. Another way of solving the issue is to conduct further research on narrowband filters with a sharp rejection in close proximity to the passband, allowing the spectrum to be used more effectively. Of all filter functions, the elliptical function has the sharpest rejection.

There are already many narrowband elliptical filter solutions available, but the emergence of precise microfabrication methods, where cost is often directly proportional to the physical size of the component, has increased the demand for more compact designs. This demand is further driven by radio frequency (RF) systems, particularly antenna arrays, where numerous electronic components must fit within the small $\lambda/2$-distance (half wavelength) between antenna elements. However, filters attempting to solve this problem [2] are often very complex and expensive to fabricate, which is generally undesirable.

Surface acoustic wave (SAW) filters, which are widely used in cellular phones due to their suitability for high volume production [3], utilise piezoelectric materials such as quartz to generate waves under electric fields. Their compact size, a result of their unique structure where thickness scales with frequency, makes them appealing for high-frequency applications. However, this also introduces fabrication challenges at higher frequencies, often limiting practical use to below 3 GHz. Although operation has been demonstrated up to 10 GHz [4], SAW filters remain costly unless mass-produced, limiting their viability for low volume or prototype-stage filters.

One compact narrowband elliptical bandpass filter realisation that seems to have been abandoned [5] due to difficulties with both construction and tuning is "The Stepped Digital Elliptic Filter" [6], originally introduced by Rhodes in 1969. Figure 1.1 shows a representation of the filter's structure, and Figure 1.2 illustrates its

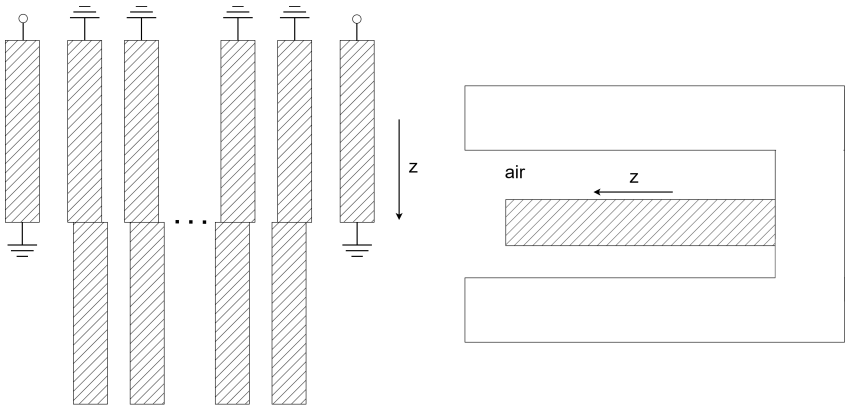


Figure 1.1: The stepped digital elliptic filter.

lumped element equivalent for a third-order filter. The filter has promising qualities; it can realise $n-1$ transmission zeros, is very compact, and consists only of distributed elements suspended in the air without the use of dielectrics, which allows for high Q values. Despite its promising compactness - achieved by parallel coupling of $\lambda/8$-long interdigital networks - the author could not find any further publications on this filter structure in the existing literature. This lack of subsequent research may be attributed to the challenges posed by its unique structure, particularly the precise alignment required for the parallel-coupled transmission lines. The gap in the literature, combined with advancements in modern microfabrication technology, makes this filter structure appropriate for re-evaluation and development.

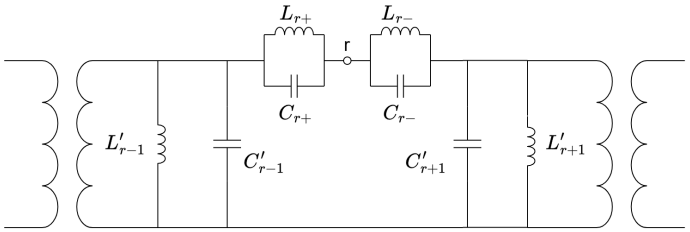


Figure 1.2: Lumped element equivalent of a third-order stepped digital elliptic filter.

Thus, the aim of the thesis is to fill a gap in the reported literature about this filter. By analysing and evaluating the specific filter structure under different potential filter and design requirements and thereby discovering and assessing potential problems, this research seeks to provide a deeper understanding of its performance and limitations. The findings will hopefully contribute to the development of future designs that can overcome existing challenges, ultimately paving the way for practical implementation in modern RF systems.

However, this thesis does not address the specific manufacturing processes of any filter prototypes. Additionally, the proposed filter design is not intended for microfabrication due to several factors: the high costs involved, the challenges in fabrication, specific supplier requirements, and the increased difficulty of evaluating the filter at high frequencies. The filter design will only be evaluated at 5 GHz for simplicity. Further research is needed to develop a design that can be fully optimised for microfabrication, addressing challenges such as material compatibility, fabrication tolerances, and cost efficiency.

The thesis will be structured in such a way that the relevant and general theory about filters, transmission lines, and manufacturing technologies is first treated, as well as a literature study on filters, eventually leading to the theory and method of design of the filter. Thereafter, analysis of the filter function and the physical structure of the filter along with simulated results will be presented. These findings will then be discussed and accompanied by suitable work for future researchers on the topic. Finally, the results are summarised, highlighting the most significant conclusions of this work.

2

Theory

In this chapter, the general theory relevant to filter design is presented.

2.1 Network analysis

Microwave components, such as filters, can be represented by n-port networks, as shown in Fig. 2.1, with the most common network being the two-port network. In these networks, V_1, V_2, V_n, I_1, I_2, and I_n represent the voltages and currents at ports 1 through and n, respectively. To analyse these components, various network pa-

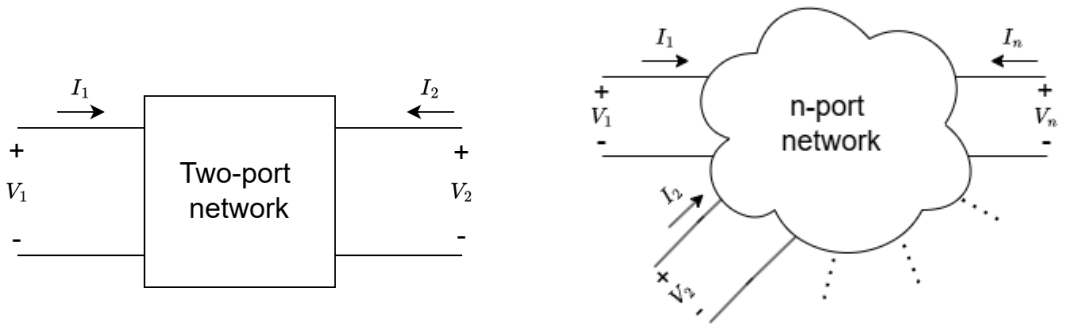


Figure 2.1: Two-port and n-port networks with voltages and currents.

rameters, such as impedance, admittance, and scattering parameters (S-parameters), are employed. These parameters are especially important at microwave frequencies, where conventional measurement tools, such as voltmeters and ammeters, are not suitable because of the high frequencies involved. Instead, transmission line theory and the concept of scattering waves are applied to characterise the behaviour of microwave networks and components.

2.1.1 Transmission line theory

The characteristic impedance of a transmission line is defined as

$$Z_0 = \frac{V_+}{I_+} \tag{2.1}$$

and the input impedance a length l away from a load impedance Z_L of a lossless transmission line is defined as

$$Z_{\text{in}}(l) = Z_0 \frac{Z_L + Z_0 \tan(\beta l)}{Z_0 + Z_L \tan(\beta l)} \tag{2.2}$$

where $\beta = 2\pi/\lambda$ is the propagating phase constant. If the line is shorted ($Z_L = 0$) or the line is left open ($Z_L = \infty$), eq. (2.2) simplifies to

$$Z_{\text{in,short}}(l) = jZ_0 \tan(\beta l) \tag{2.3}$$

and

$$Z_{\text{in,open}}(l) = -jZ_0 \cot(\beta l) \tag{2.4}$$

which allows for the realisation of impedances without adding load impedances.

2.1.2 Scattering matrix

The scattering matrix, containing the S-parameters, is often used to describe microwave input and output characteristics. In a two-port network, the S-parameters relate the incident and reflected normalised voltages at the respective ports as

$$\begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix} \tag{2.5}$$

where S_{11} and S_{22} are reflections at the corresponding ports, and S_{12} and S_{21} represent transmissions between the corresponding ports. The reflections are often referred to as return loss (RL), and the transmissions as insertion loss (IL) and are defined as

$$\text{RL} = -20 \log_{10} |S_{11}| \tag{2.6}$$

and

$$\text{IL} = -20 \log_{10} |S_{21}| \tag{2.7}$$

respectively where both are in units of dB.

The scattering matrix can be extended to multiport networks with n-ports

$$\begin{bmatrix} V_1^- \\ V_2^- \\ \vdots \\ V_n^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{11} & \dots & S_{1n} \\ S_{21} & S_{22} & \dots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \dots & S_{nn} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ \vdots \\ V_n^+ \end{bmatrix} \tag{2.8}$$

which allows for analysis of more complex networks.

2.1.3 Admittance and impedance matrices

In a two-port network, the admittance and impedance matrices are defined as

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{2.9}$$

and

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{2.10}$$

where the two matrices are related through

$$\mathbf{Z} = \mathbf{Y}^{-1}. \tag{2.11}$$

These parameters can be related to the S-parameters through any of the following equations [7]:

$$\mathbf{S} = \sqrt{\mathbf{Y_0}}(\mathbf{Z} - \mathbf{Z_0})(\mathbf{Z} + \mathbf{Z_0})^{-1}\sqrt{\mathbf{Z_0}} \tag{2.12a}$$
$$\mathbf{S} = \sqrt{\mathbf{Z_0}}(\mathbf{Y_0} - \mathbf{Y})(\mathbf{Y_0} + \mathbf{Y})^{-1}\sqrt{\mathbf{Y_0}} \tag{2.12b}$$
$$\mathbf{Z} = \sqrt{\mathbf{Z_0}}(\mathbf{U} + \mathbf{S})(\mathbf{U} - \mathbf{S})^{-1}\sqrt{\mathbf{Z_0}} \tag{2.12c}$$
$$\mathbf{Y} = \sqrt{\mathbf{Y_0}}(\mathbf{U} - \mathbf{S})(\mathbf{U} + \mathbf{S})^{-1}\sqrt{\mathbf{Y_0}} \tag{2.12d}$$

where the matrices

$$\mathbf{Z_0} = \begin{bmatrix} Z_{01} & 0 \\ 0 & Z_{02} \end{bmatrix} \tag{2.13a}$$
$$\mathbf{Y_0} = \begin{bmatrix} Y_{01} & 0 \\ 0 & Y_{02} \end{bmatrix} \tag{2.13b}$$

are the port impedances for a two-port network and $\mathbf{U}$ is the unity matrix.

The admittance and impedance matrices can also be extended to multiport networks in the same way as was shown for the scattering matrix in (2.12).

2.2 Transfer function

A systems response in the frequency domain is defined through its transfer function

$$H(s) = \frac{\text{response function}}{\text{driving function}} \tag{2.14}$$

which has the general shape

$$H(s) = \frac{b_n s^n + b_{n-1}s^{n-1} + ... + b_1 s + b_0}{a_m s^m + a_{m-1}s^{m-1} + ... + a_1 s + a_0} = \frac{N(s)}{D(s)} \tag{2.15}$$

where a and b are constants and N and D stands for numerator and denominator respectively. This function can be derived using time-domain analysis of electrical circuits.

A stable filter requires that the zeros of $D(s)$ (that is, the poles of $H(s)$) must lie in the left half of the complex plane. Therefore, $D(s)$ is a Hurwitz polynomial: all its roots and zeros lie either in the left half-plane or on the imaginary axis. If this is not the case, the filter will have an increasing magnitude over time, which is unphysical for a passive filter [8]. The poles and zeros of a filter will be different for every

filter network, and it is thus obvious that the pole and zero placements will directly change the filter's amplitude response over frequency.

Specifically for two-port lossless passive filter networks the transfer function is generally defined as

$$H(s) = |S_{21}(j\omega)|^2 = \frac{1}{1 + \epsilon^2 F_n^2(\omega)} \tag{2.16}$$

where ϵ is a ripple constant, F_n is the characteristic function, and ω is the frequency variable [9]. The characteristic function F_n defines the filtering function and the ripple constant ϵ defines the allowable variation of the function.

2.3 Filter functions and responses

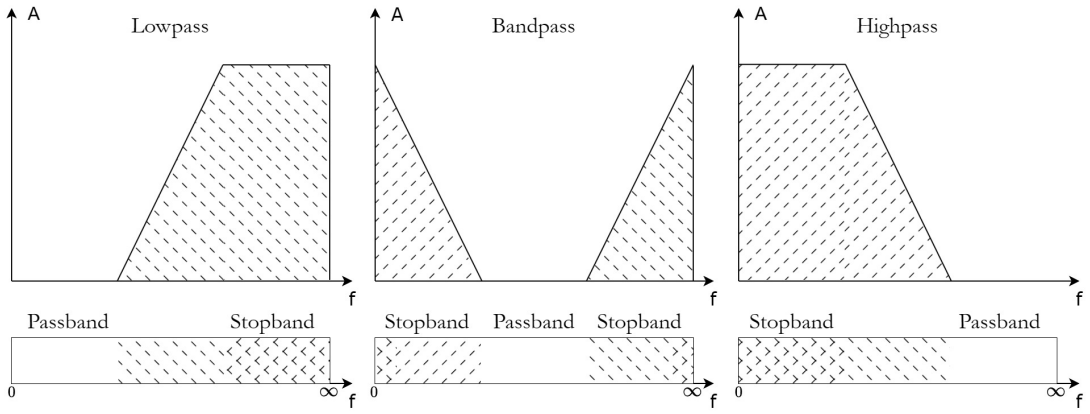


Figure 2.2: Attenuation over frequency for different filter responses.

Three different filter responses are shown in figure 2.2, namely, lowpass, bandpass and highpass. These filters are commonly used for a variety of applications. Another common one is the bandstop filter response, which is just the reciprocal of the bandpass filter. The point where the filter response starts to change from a passband to stopband, or vice versa, is usually referred to as the cutoff frequency.

Ideally, the change between passband and stopband should be instantaneous, with, for example, the stopband of the lowpass filter response beginning directly at the cutoff frequency. Achieving this ideal cutoff is physically impossible, but there are mathematical functions which aim to reproduce this ideal response as closely as possible, each with different advantages and disadvantages. These mathematical functions, when used in a filter context, are referred to as filter functions. Filter synthesis requires a filter function that can both satisfy the required specifications of a filter and be exactly realisable by a practical circuit.

Due to causality, a filter function cannot [10]:

- be zero over a finite band of frequencies,
- fall faster in amplitude than an exponential order function.

A filter function can thus only be zero for discrete frequencies, the so-called transmission zeros (TZ). The Bode-Fano criterion [11], [12] also shows that this is impossible since the passband reflection coefficient cannot be equal to zero unless the bandwidth is also equal to zero [13]. This leads to the immediate conclusion that the filter function can only be zero for finite, discrete frequencies, which proves that exactly reproducing an ideal lowpass response is impossible.

Three types of filter functions will be introduced in this section: maximally flat (also known as Butterworth), Chebyshev, and elliptical function filters.

2.3.1 Maximally flat

Maximally flat filters functions are designed to have a minimum amount of ripple and a linear phase response in the passband. However, this comes at the cost of a slower increase in attenuation, which requires a higher degree of the filter to achieve potential filter requirements.

The transfer function of such a filter is given by

$$H(s) = \frac{1}{1 + \omega^{2n}} \tag{2.17}$$

with n being the filter degree. The maximally flat filter does not have any transmission zeros at finite frequencies because all zeros in the transfer function are placed at infinity, and all poles are distributed equally on the unit circle in the left half-plane [9].

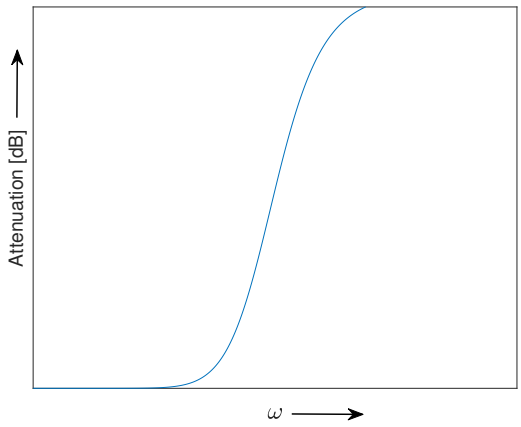


Figure 2.3: Maximally flat filter function lowpass response.

2.3.2 Chebyshev function

A Chebyshev filter function has an equiripple passband and a maximally flat stopband, with a transfer function given by

$$H(s) = \frac{1}{1 + \epsilon^2 T_n^2(\omega)} \tag{2.18}$$

where T_n is a chebyshev function of order n given by

$$T_n(\omega) = \begin{cases} \cos\left(n\cos^{-1}\omega\right) & \text{for } |\omega| \leq 1 \\ \cosh\left(n\cosh^{-1}\omega\right) & \text{for } |\omega| \geq 1 \end{cases}, \tag{2.19}$$

hence its name [9]. The filter response in Fig. 2.4 shows a slightly sharper roll-off

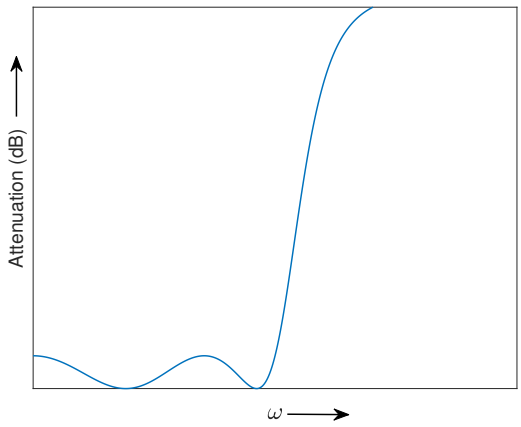


Figure 2.4: Chebyshev filter function lowpass response.

but at the expense of ripple in the passband. All transmission zeros of the transfer function are located at infinity, just as in the maximally flat case, but the poles are instead distributed on an ellipse in the left half-plane.

2.3.3 Elliptical function

The most important aspect of the elliptical filter functions and what it is intended to accomplish is the steepest possible roll-off between the passband and the stopband. This means that the pass- and stopbands of an elliptic filter can be closer in frequency than for Chebyshev and maximally flat filters. This is due to the location of the transmission zeros; the elliptic function has transmission zeros at finite frequencies, unlike the Chebyshev and maximally flat functions.

Another advantage of the elliptic filter function is that it requires the lowest degree of its characteristic function compared to all other characteristic functions, which makes it the most cost-effective (lowest degree) of the functions to obtain an as ideal response as possible [14]. A lower degree filter is also inherently more compact, due to the fewer number of resonant sections required. However, this comes at the cost of having the least linear phase in the passband [15], a more complex lumped element configuration, and ripple in both bands.

The transfer function for an elliptical filter function is given by

$$H(s) = \frac{1}{1 + \epsilon^2 F_n^2(\omega)} \tag{2.20}$$

where F_n is defined as in equations (2.21) and (2.22) for even and odd order respectively [16].

$$F_n = C_{\text{even}} \prod_{\nu=1}^{n/2} \frac{\omega^2 - \omega_\nu^2}{\omega^2 - (\omega_s/\omega_\nu)^2} \qquad F_n = C_{\text{odd}} \prod_{\nu=1}^{(n-1)/2} \frac{\omega^2 - \omega_\nu^2}{\omega^2 - (\omega_s/\omega_\nu)^2}$$

$$C_{\text{even}} = \prod_{\nu=1}^{n/2} \frac{1 - (\omega_s/\omega_\nu)^2}{1 - \omega_\nu^2} \quad (2.21) \qquad C_{\text{odd}} = \prod_{\nu=1}^{(n-1)/2} \frac{1 - (\omega_s/\omega_\nu)^2}{1 - \omega_\nu^2} \quad (2.22)$$

$$\omega_\nu = \text{sn}\left(\frac{(2\nu - 1)K(k^2)}{n}, k^2\right) \qquad \omega_\nu = \text{sn}\left(\frac{(2\nu)K(k^2)}{n}, k^2\right)$$

In equations (2.21) and (2.22), K is the complete elliptic integral of the first kind defined as

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{1 - k^2 \sin^2 \theta} \tag{2.23}$$

where the modulus k is defined as

$$k = \sin\theta \tag{2.24}$$

and where sn is the elliptic sine defined as

$$\text{sn}(u, k^2) = \sin\theta. \tag{2.25}$$

The modular angle θ defines the sharpness of the roll-off of the filter function, with a higher angle relating to a faster roll-off, but also a higher minimum stopband attenuation A_{min}.

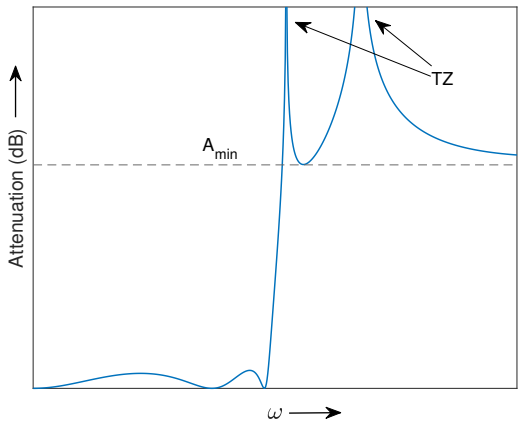


Figure 2.5: Elliptic filter function lowpass response with two transmission zeros.

The necessary filter degree n needed to achieve wanted passband and stopband attenuations α_p and α_s can be calculated using the following equation [14]:

$$n \geq \frac{\cosh^{-1}(1/k_1)}{\cosh^{-1}(1/k)} \tag{2.26}$$

where

$$k_1 = \left[\frac{10^{\alpha_p/10} - 1}{10^{\alpha_s/10} - 1}\right]^{1/2}$$

and n being rounded up to the closest integer.

2.4 Lowpass prototype filters

The building blocks for creating any filter function are circuit equivalents of lowpass, highpass, bandpass, and bandstop filters. These circuit equivalents are derived from lowpass prototypes, which are then transformed to the desired frequency response.

A lowpass prototype filter is, essentially, a filter template of lumped elements in the form of a lowpass filter that can later be transformed into other filter response types. The phase and amplitude response of the lowpass prototype is directly proportional to the practical realisation created from it; a bandpass filter's amplitude response at the centre and edge frequencies is exactly the same as the lowpass prototypes response at normalised 0 and 1 respectively [17]. A general lowpass prototype filter circuit is shown in 2.6.

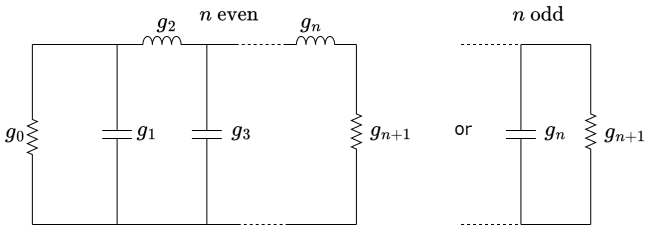


Figure 2.6: Prototype filter.

2.4.1 Elliptical function prototype

In order to realise the transmission zeros of an elliptical function filter, the lowpass prototype requires resonating LC-circuits. An alternative form of this lowpass prototype can also be derived, where the inductor and capacitor are instead directly connected to ground. These two configurations are presented in figure 2.7.

2.4.2 Element value determination

Once the correct lowpass prototype filter has been chosen, the element values for the R, L, C-components must be found. Unfortunately, unlike for example the Butterworth or Chebyshev filters, there is no formula that can simply derive the elliptical function filter networks element values [18]. However, there are extensive tables for both equally terminated, i.e. $g_0 = g_1 = 1$, and not equally terminated prototypes [19], [20], as well as software, for example Ansys Nuhertz FilterSolutions, which can calculate the element values for any desired frequency and attenuation.

2.5 Frequency and impedance transformations

As mentioned in section 2.4, frequency transformations are used to transform the normalised lowpass prototype filter into the frequency band of interest, while also choosing the filter type. In addition, since the lowpass prototype terminations are

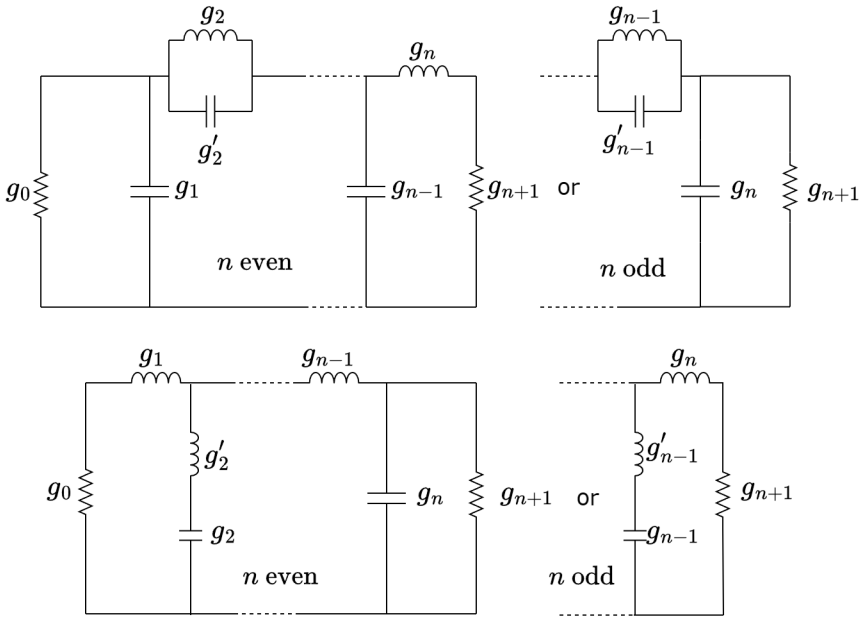


Figure 2.7: Lowpass prototype filters for elliptic function filter

generally normalised to a termination of $1\,\Omega$, which is not feasible in most applications, the impedance must also be transformed.

Table 2.1: Frequency transformations for different filter responses [6], [19].

Lowpass	$\omega' \rightarrow \dfrac{\tan\omega}{\tan\omega_c}$
Highpass	$\omega' \rightarrow \dfrac{\tan\omega_c}{\tan\omega}$
Bandpass	$\omega' \rightarrow a\left(\tan\omega - \dfrac{1}{\tan\omega}\right)$
Bandstop	$\omega' \rightarrow \dfrac{1}{a\left(\tan\omega - \frac{1}{\tan\omega}\right)}$

Using the different frequency transformations found in table 2.1, the normalised lowpass prototype can be transformed into any of the listed responses with a one-to-one relationship of the amplitude and phase responses.

In table 2.1, ω' is the lowpass prototype frequency variable, ω is the frequency variable, ω_c is the cutoff frequency and

$$a = \frac{\tan 2\omega_c}{2} \tag{2.27}$$

is the bandwidth scaling constant. The normalised cutoff frequency can be found through

$$\omega_c = \omega_n\left(1 - \frac{\mathrm{FBW}}{2}\right) \tag{2.28}$$

where ω_n is the normalisation frequency and FBW is the fractional bandwidth

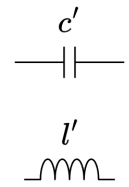
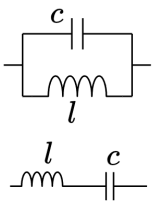
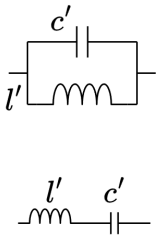
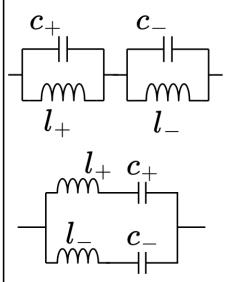
(FBW) defined as

$$\mathrm{FBW} = \frac{\omega_2 - \omega_1}{\omega_0} \tag{2.29}$$

where ω_1 and ω_2 are the lower and upper band edge frequencies and ω_0 is the centre frequency of the passband.

These frequency transformations also change the network itself, as shown for a bandpass transformation in table 2.2.

Table 2.2: Lowpass prototype to bandpass circuit transformation. The parameters a and λ are defined in equations (2.27) and (2.30) respectively.

Lowpass	Bandpass	Element values
		$c = \frac{1}{l} = ac'$ $l = \frac{1}{c} = al'$
		$c_+ = \frac{1}{l_-} = ac'(1+\lambda_{r-}^2)$ $c_- = \frac{1}{l_+} = ac'(1+\lambda_{r+}^2)$ $l_+ = \frac{1}{c_-} = al'(1+\lambda_{r-}^2)$ $l_- = \frac{1}{c_+} = al'(1+\lambda_{r+}^2)$

To obtain the correct locations of the attenuation poles of the bandpass filter network, two new variables need to be defined. First,

$$\lambda_{r\pm} = \sqrt{\left(\frac{\Omega_r}{2a}\right) + 1} \pm \frac{\Omega_r}{2a} \tag{2.30}$$

are the frequencies where the bandpass attenuation poles are located, and second Ω_r is the frequency where the lowpass prototype attenuation pole is located, as illustrated in Fig. 2.8. Intuitively, the doubling of the elements moving from the lowpass prototype to the bandpass circuit can be understood as transforming the lowpass response to a lowpass below and a highpass above the passband, thus requiring the double amount of attenuation poles.

2.6 Asymmetric response filters

All lowpass prototypes discussed so far have been symmetric. This is inherent due to the lossless inductors and capacitors terminated with a matching resistor, causing

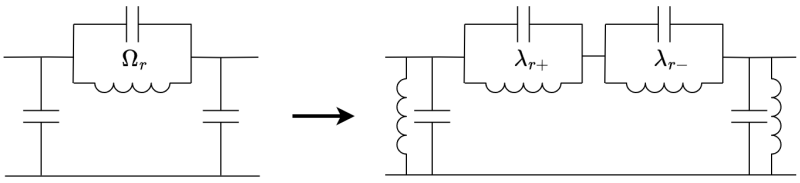


Figure 2.8: Attenuation pole transformation from lowpass to bandpass circuit.

the network to be symmetric around zero frequency [21]. Furthermore, all frequency transformations create filter responses that are symmetric around the passband centre. Due to these two properties, asymmetric filter responses are not directly realisable using ordinary lowpass prototype methods.

There is however a way of overcoming this specifically for the bandpass realisation. By adding hypothetical frequency-invariant reactive (FIR) elements to the lowpass prototype, which disappears during the synthesis process, an asymmetric bandpass response can be realised [22].

The advantages of asymmetric response filters are several, with the main advantage being that the transmission zeros can be placed more freely. This is especially useful when there are requirements in two bands, as shown in [23] and [24]. Using asymmetric filters one could potentially require only one filter instead of two cascaded filters to fulfill multiband requirements.

2.7 Distributed element equivalents

A common approach when designing a filter realised through distributed elements is to first find its approximate lumped element equivalent and then transform it into a distributed element realisation through Richard's transformation.

Starting with equations (2.3) and (2.4) for the shorted and open transmission lines and choosing $l = \lambda/8$ one arrives at

$$Z_{\text{in,shorted}}\left(\frac{\lambda}{8}\right) = jZ_0 \tag{2.31}$$

$$Z_{\text{in,open}}\left(\frac{\lambda}{8}\right) = -jZ_0 \tag{2.32}$$

since $\tan(\beta\lambda/8) = 1$. Equating these expressions to the known expressions for the impedances of lumped inductors and capacitors as seen in equations (2.33) and (2.34),

$$j\omega L = -jZ_{0,\text{short}} \Rightarrow Z_{0,\text{short}} = \omega L \tag{2.33}$$

$$\frac{1}{j\omega C} = -jZ_{0,\text{open}} \Rightarrow Z_{0,\text{open}} = \frac{1}{\omega C} \tag{2.34}$$

the Richard's transformations of Fig. 2.9 are found. Using the transformations in Fig. 2.9 a lumped element circuit consisting of capacitors and inductors can easily be translated into a distributed transmission line at a specific frequency relating to the length of $\lambda/8$.

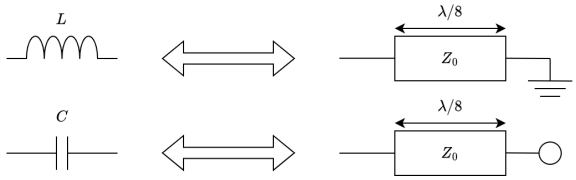


Figure 2.9: Physical conversion of elements using Richard's transformations.

2.8 Coupled line theory

Unshielded transmission lines relatively close to each other can experience coupling due to the electromagnetic fields surrounding the conductors, something that can be a problem or a resource depending on the situation. This coupling allows for more freedom on how a lumped element circuit can be realised in transmission line technology other than just the classical Richard's transformation shown in figure 2.9.

The characteristic impedance Z_0 of a lossless uniform transmission line operating in the TEM mode is related to its shunt capacitance by A lossless uniform TEM mode transmission lines shunt capacitance is connected to its characteristic impedance Z_0 through the relationship

$$Z_0 = \frac{\eta_0}{\sqrt{\varepsilon_r}} \frac{1}{(C/\varepsilon)} \tag{2.35}$$

where C/ε is the static capacitance ratio per unit length between conductors over the dielectric's permittivity.

One major contributor to these equivalent capacitive networks of coupled lines was Getsinger [25], who provided equations for both even- and odd-mode excitations, as well as graphs that simplified the filter design process for coupled line filters. This network relates to two parallel rectangular bars between parallel plates, as can be seen in figure 2.10, where the circles represent the conductors.

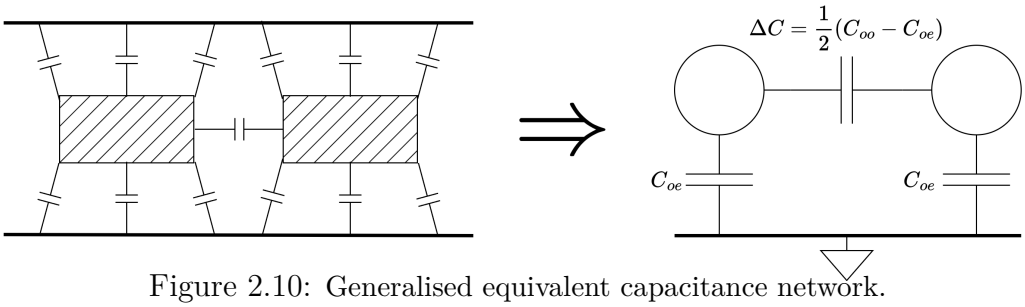


Figure 2.10: Generalised equivalent capacitance network.

C_{oe} and C_{oo} are the even- and odd-mode capacitances to ground when the two conductors are at the same or opposite potential respectively. This simple model can be used to model more complex structures where more conductors are added, which will be shown in section 2.9.

The characteristic impedances for the even- and odd-modes can then be calculated,

using the knowledge that these are TEM lines, as

$$Z_{oe} = \frac{1}{v_{\mathrm{p}} C_{oe}} \text{ and } Z_{oo} = \frac{1}{v_{\mathrm{p}} C_{oo}} \tag{2.36}$$

where v_{p} is the phase velocity of on the line.

2.8.1 Capacitance matrices

Using the lumped element filter networks of capacitors and inductors, assuming that they can be realised by parallel coupled lines, a capacitance matrix can be formed containing all the relevant information about the system. Assuming that there is no coupling between non-neighbouring lines for a n-node network, the general capacitance matrix that follows is

$$\mathbf{C}_{n\times n} = \frac{\eta}{Z_0} \begin{bmatrix} c_1 + c_{1,2} & -c_{1,2} & \dots & 0 \\ -c_{1,2} & c_2 + c_{1,2} + c_{2,3} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_n + c_{n-1,n} \end{bmatrix} \tag{2.37}$$

where c_n are the capacitances between the lines and ground, $c_{i,j}$ are the capacitances due to coupling, η is the wave impedance and Z_0 is the desired system impedance. The assumption that there is no coupling between non-neighbouring lines is what causes all matrix elements which are not directly next to the diagonal to be zero. Without this assumption, the analysis would become more complicated without a significant gain in precision [26]. With the help of electromagnetic solvers that exist today, this assumption can be easily evaluated for any type of coupled structure.

In certain cases, the impedance values related to the capacitance matrix become impossible to physically realise, or unwanted due to size or fabrication concerns of the wanted structure. The first case is generally solved by using a suitable frequency transformation, while the latter is solved by scaling of the capacitance network to achieve the desired impedance value [6](see sec.2.8.2).

2.8.2 Capacitance matrix scaling

As previously mentioned the scaling of the capacitance matrix relates to its physical realisability, and there are two ways to scale the matrix which relates to different physical properties:

i) scaling of an interior row and column
ii) scaling of a beginning or end row and column.

The main difference between these two scalings are that i) affects only the impedance levels of the lines, while ii) also changes the two-port characteristics of the network [27]. This means that by interior scaling of a row and column, any number of equivalent networks can be obtained without changing the input and output of the network. The only restriction on this scaling is that the sum of the elements of any row or column must be greater than or equal to zero to guarantee realisability.

If a unit element (UE) is added to the input and output of the network, the capacitance matrix is expanded with one more row and column. The UE capacitance matrix is as follows

$$\mathbf{C}_{\mathrm{UE}} = \frac{\eta}{Z_0}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

and when added to the end lines of the normal capacitance matrix (2.37) (ignoring impedance scaling)

$$\begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 1+C_1 & -c_{1,2} & \dots & 0 & 0 \\ 0 & -c_{1,2} & C_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & \dots & 1+C_n & -1 \\ 0 & 0 & \dots & \dots & -1 & 1 \end{bmatrix} \tag{2.38}$$

where C_n is the sum of the capacitance elements of the n-th line as in (2.37). The benefit of adding UEs is obvious, since now any scaling on the filter network itself abides the scaling type i), and thus does not change the two-port characteristics of the network.

The scaling procedure is shown in (2.39)

$$\begin{array}{c} \begin{matrix} & \times n_1 & \times n_2 & \dots & \times n_n & \end{matrix} \\ \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 1+C_1 & -c_{1,2} & \dots & 0 & 0 \\ 0 & -c_{1,2} & C_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 1+C_n & -1 \\ 0 & 0 & \dots & \dots & -1 & 1 \end{bmatrix} \end{array} \begin{matrix} \\ \times n_1 \\ \times n_2 \\ \vdots \\ \times n_n \end{matrix} \tag{2.39}$$

where $n_1, \dots, n_n$ are scaling constants for each of the separate lines excluding the UEs [27]. It can be seen that the UEs will have the impedance system impedance.

As previously mentioned, any number of equivalent networks can be found by scaling the interior elements of the matrix as shown in (2.39). However, the issue is finding one such network that is suitable for practical realisation in a coupled structure. One trait which is helpful because of symmetry is equal spacing between lines. This has been shown to be inherently satisfied for filters which are realised using multiple resonating line sections [6]. Because of this, the scaling factors should be chosen such that the variation in ground and coupling capacitances is minimised [6]. This ensures that the spacing between lines is generally equal.

For an interdigital realisation, the scaling factors are simply chosen such that the capacitances to ground are equal for every line. For stepped digital realisations, scaling factors with this property are not as easily found, requiring intermediate steps including scaling of the diagonal elements. The following diagonal scaling procedure with $n_{\mathrm{diag}} = 1.2$, as shown in (2.40) (excluding UEs), has been found

to allow scaling factors that give an optimum, or close to optimum, relationship between the ground and coupling capacitances [6].

$$\begin{bmatrix} C_1 & \dots & \dots & \dots & 0 \\ \vdots & n_{\text{diag}} \times C_1 & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \dots & n_{\text{diag}} \times C_n & \dots \\ 0 & 0 & \dots & \dots & C_n \end{bmatrix} \tag{2.40}$$

Essentially, the diagonal elements are just equaled to n_{diag} times the lower or upper diagonal value. If the matrix is uneven, the centre diagonal element can simply be chosen as n_{diag} times the lower or upper diagonal value.

2.8.3 Multiconductor transmission lines

The coupled line theory can be extended to multiconductor TEM transmission line cases, where the relationship between the vectors of line voltages $\hat{V}$ and line currents $\hat{I}$ along the n-conductor lines, are defined by the eigenvalue equations

$$[\mathbf{G} + j\omega\mathbf{C}][\mathbf{R} + j\omega\mathbf{L}]\hat{I} = \mathbf{YZ} = \gamma^2\hat{I} \tag{2.41a}$$

$$[\mathbf{R} + j\omega\mathbf{L}][\mathbf{G} + j\omega\mathbf{C}]\hat{V} = \mathbf{ZY} = \gamma^2\hat{V} \tag{2.41b}$$

where G, R, L and C are all symmetric matrices relating to the physical properties of the line. The eigenvalues γ are thus the same for $\hat{I}$ and $\hat{V}$, but the eigenvectors are not equal [28].

If the lines are assumed lossless the following relations hold

$$\mathbf{LC}\,\hat{I} = \frac{1}{v_p^2}\hat{I} \tag{2.42a}$$

$$\mathbf{CL}\,\hat{V} = \frac{1}{v_p^2}\hat{V} \tag{2.42b}$$

where v_p is the phase velocity on the line. Using these relations (assuming that the capacitance is known from lumped element prototype), the characteristic admittance and impedance matrices of the n-conductor lines can be calculated using

$$\mathbf{Y_0} = v_p\mathbf{C} \tag{2.43a}$$

$$\mathbf{Z_0} = \mathbf{Y_0}^{-1}. \tag{2.43b}$$

Using equations (2.41) the admittance and impedances can be diagonalised by

$$\mathbf{M_V}\,\mathbf{Z}\,\mathbf{Y}\,\mathbf{M_V}^{-1} = \boldsymbol{\gamma}^2 \tag{2.44}$$

which gives the characteristic admittance and impedance through

$$\mathbf{Y_0} = \mathbf{M_V}\gamma^{-1}\mathbf{M_V}^{-1}\mathbf{Y} = \mathbf{Z_0}^{-1} \tag{2.45}$$

where M_V and M_I ($M_I = M_V^{-T}$) denotes the voltage and current eigenvectors given by solving the characteristic equation.

The characteristic admittance matrix contains all information about the physical properties along with any coupling towards neighbouring lines of the filter network.

2.8.4 2*n*-port

Together with the voltage and current eigenvectors, a $2n$-port model which characterises the system is defined by

$$\mathbf{Y} = \sum_{m=1}^{n} \mathbf{Y_m} \tag{2.46}$$

with

$$[Y_m]_{2n\times 2n} = \begin{bmatrix} \coth(\gamma_m l)Y_0^m & -\mathrm{csch}(\gamma_m l)Y_0^m \\ -\mathrm{csch}(\gamma_m l)Y_0^m & \coth(\gamma_m l)Y_0^m \end{bmatrix}$$

where l is the length of the uniform coupled lines and m denotes the mth mode of propagation [29]. The 2n-port model is visualised in figure 2.11.

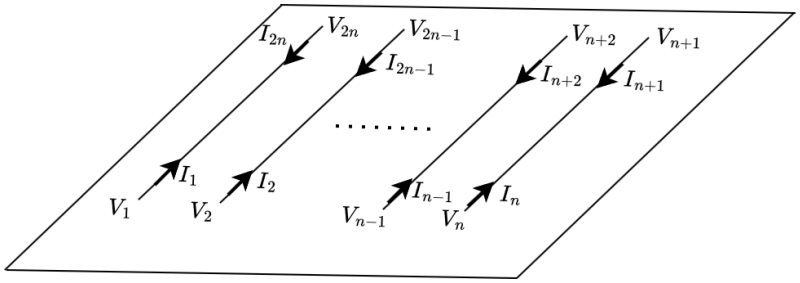


Figure 2.11: 2n-port model.

For the TEM case with a homogenous dielectric there is only one mode, but otherwise Y_0^m is found through

$$\mathbf{Y_0}^m = (\mathbf{Y}_{\mathrm{LM}} * \mathbf{M_V})\mathbf{D_m}\mathbf{M_I}^T$$

where Y_{LM} is the line-mode admittance matrices and D_m is a diagonal matrix given by

$$\mathbf{D_m} = \begin{cases} D(j,j) = 0 & \text{for } j \neq m \\ D(j,j) = 1 & \text{for } j = m \end{cases}$$

as defined in [29].

2.8.5 *n*-port characteristics extraction

Using the $2n$-port found from equation (2.46), the admittance matrix can be transformed to an S-parameter matrix using equation (2.12b). Once the S-parameter matrix is found, the n-port characteristics are identified as the S-parameters at the specific ports (corresponding to specific rows and columns of the matrix) of interest.

2.9 Filter types

In this section, a few relevant filters realised through distributed transmission line elements will be introduced and discussed.

2.9.1 Stripline filters

Stripline bandpass filters consists of several coupled half wavelength long strips, either end-coupled or parallel-coupled, with the difference shown in Fig. 2.12. The end-coupled configuration is inherently long, making the parallel-coupled configuration generally favourable. These filters are easily created through the use of

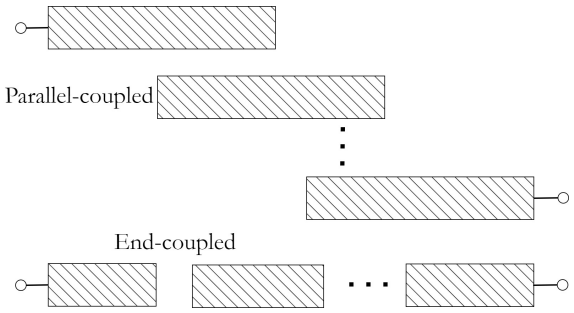
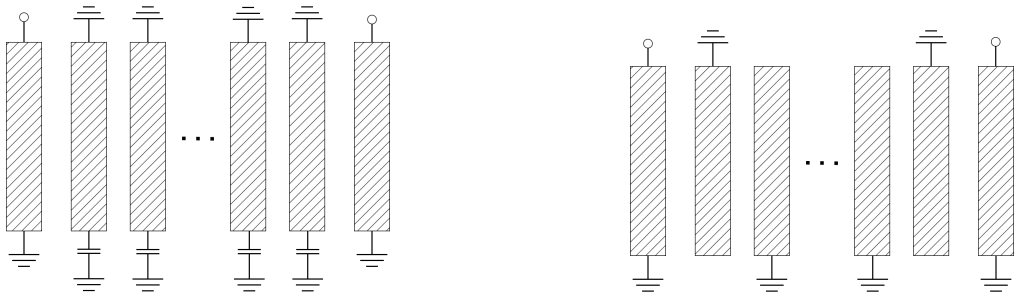


Figure 2.12: Two different stripline filters.

tabulated values and graphs for any frequency suitable for the stripline technology. Examples of these stripline filters can be found in [13].

2.9.2 Combline and interdigital filters

Even though the parallel-coupled stripline filter is shorter than the end-coupled one, more compact solutions are still favourable. By instead placing the resonant lines side by side, more compact structures such as the combline and interdigital filters can be created.



(a) Combline filter. (b) Interdigital filter.

Figure 2.13: Combline and interdigital filter realisations.

Combline filters consist of resonating lines coupled side by side that are less than a quarter-wavelength long at centre frequency, with two impedance transformers

as input and output lines. One end of the lines are coupled to ground through lumped element capacitors while the other is shorted. This allows for a very compact structure, however, since they require lumped elements they are not suitable for all purposes.

If the lumped elements of Fig. 2.13a are removed, and the lines alternate between being left short- and open circuited at each end, the interdigital filter seen in Fig. 2.13b is found. Although combline filters are slightly more compact than their interdigital equivalents, the interdigital filter has some advantages: they can achieve higher Q-values, have strong stopbands, consists solely of distributed elements and have quite large coupling tolerances [20].

The coupling between the resonating lines of the interdigital filter can be modeled as capacitances between adjacent lines and between lines and ground plates as shown in figure 2.14. This capacitance model is very convenient since the capacitances

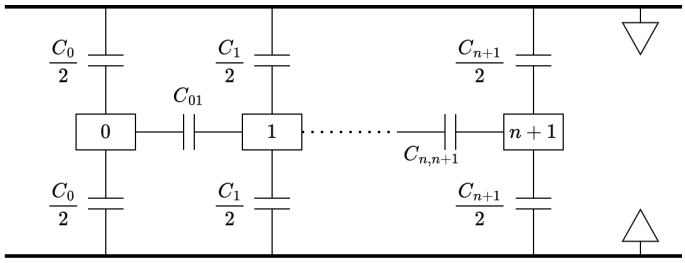


Figure 2.14: Capacitance model of an interdigital filter structure.

relate directly to the filters physical dimensions.

2.9.3 Elliptic filters

The previously mentioned filter realisations are not able to directly realise elliptic function filters. The realisation of elliptic filter functions is generally more complex and they are both difficult to construct and tune [5]. Three ways of realising elliptic filter functions will be covered in this section: the digital elliptic filter [30], the stepped digital elliptic filter [6], and the elliptic combline filter [31]. The digital elliptic and the elliptic combline filters will only be covered briefly, and the stepped digital will be covered more extensively. All three filters share the same coupled line characteristics and can be modelled with the capacitance model of the interdigital filter structure in Fig. 2.14.

2.9.3.1 The digital elliptic filter

The digital elliptic filter was introduced in 1967, and is called digital due to its finger-like appearance. This filter realisation is quite complex (but also uses the capacitance model of Fig. 2.14) and can be seen in Fig. 2.15. This realisation is for wideband purposes, but does not seem to occur in other publications after its introduction; probably due to the difficulties in constructing it. However, it is very important for historical reasons, as it was the first of these types of elliptic filters to be published [5].

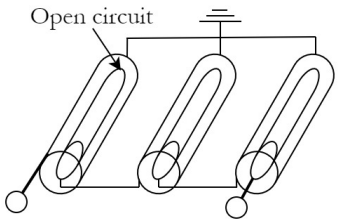


Figure 2.15: The digital elliptic filter.

2.9.3.2 The stepped digital elliptic filter

The stepped digital filter is directly inspired by the digital elliptic filter but instead aims at creating a narrowband solution. A representation of the stepped digital elliptic filter can be seen in Figure 2.16. This filter is very similar to the interdigital

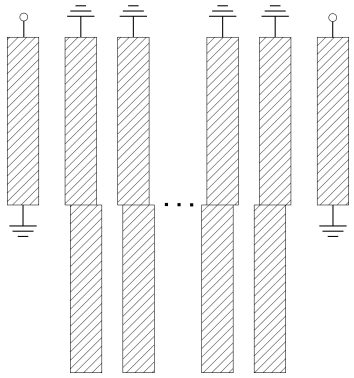


Figure 2.16: The stepped digital elliptic filter.

filter, but it consists of two parallel interdigital networks, each being an eight of a wavelength long, with UEs at the ends. The elements of network 1 are shorted, and the elements of network 2 being left open circuited. This approach allows for all shorted elements to be placed on the same side, while also allowing the elliptical function to be realised. This filter realisation has been found to be suitable for fractional bandwidths between 2 and 30%, with even lower bandwidths being difficult but possible to realise using individual pole tuning [6].

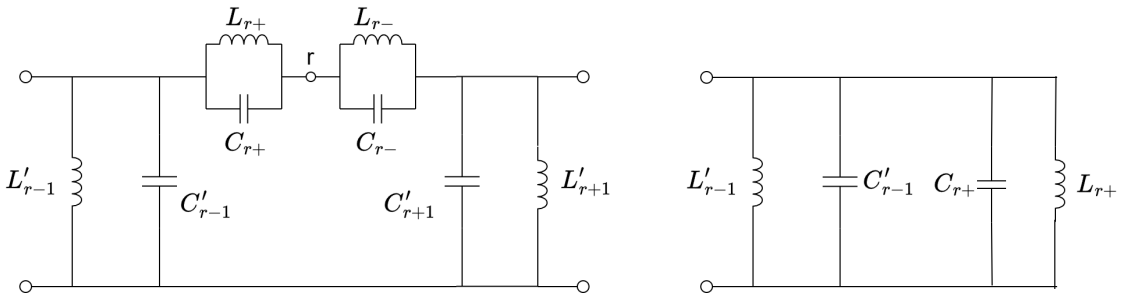
In order to create a narrow-bandpass elliptic function filter, the realisation of the poles using resonant sections must not change much depending on the bandwidth scaling factor [6]. If they do, then the realisation of the filter will be inherently wideband, like the digital elliptic filter. The solution to this problem was to apply the conventional bandpass transformation

$$\omega' \rightarrow a\left(\tan\omega - \frac{1}{\tan\omega}\right) \tag{2.47}$$

to the odd lowpass prototype in Fig. 2.7 with the band centre and lower and upper band edge frequencies being normalised to $\omega = \pi/4$, $\omega = \omega_c$, and $\omega = \pi/2 - \omega_c$ (for $\omega_c < \pi/4$) respectively [6]. For narrowband filters the bandwidth scaling factor $a \gg 1$, which causes the bandpass attenuation poles found through eq.

(2.30) to be in the order of unity, and thus the admittance levels to be approximately proportional to a. By coupling into the filter network through UEs the resonating sections will not be sensitive to scaling of the bandwidth [6], and thus a narrowband filter is possible to realise.

The coupled line realisation shown in Fig. 2.16 can be found by the following method. When the bandpass transformation is applied to the lowpass prototype in Fig. 2.7 the lumped elements are transformed according to table 2.2 into the lumped element resonating section of Fig. 2.17a.



(a) Resonating π-section. (b) Node analysis of node $r-1$.

Figure 2.17: Lumped element derivation of the stepped digital elliptic filter.

The resonating π-section of Fig. 2.17a is equal to an elliptic filter of degree 3. Under the assumption that there is no coupling between nonadjacent elements as suggested in [26], using node analysis at node $r-1$, all elements to the right of node r can be ignored, and the result is Fig. 2.17b.

Using Figs. 2.17b and 2.9, it can be clearly seen that the inductors can be realised as shorted $\lambda/8$-lines and that the capacitors can be realised as $\lambda/8$-lines left open at one end. If these lines are then set in parallel, the distributed element realisation of one of the resonating sections in figure 2.16 has been found. Repeating this node analysis for the other nodes then leads to the realisation of the remaining sections. The left- and rightmost UEs are then added as transformers to serve as input and output elements without changing the filter characteristics and to keep the filter sections at realisable impedance levels. This method can be extended to any number of odd nodes above 3 by extending the resonating π-section of Fig. 2.17a with additional nodes.

An experimental stepped digital elliptic filter was fabricated and measured by the author in 1969. A visualisation of the physical structure along with the measured filter characteristics, including a more in-depth derivation, can be found in [6].

2.9.3.3 The combline elliptic filter

The combline elliptic filter is another elliptic filter realisation which was introduced 1971 [31]. According to the author, superseded the stepped digital elliptic filter as a narrowband elliptic filter function solution, due to it being easily tuned with high-Q lumped element capacitors [5]. However, just as with the normal combline filter, this

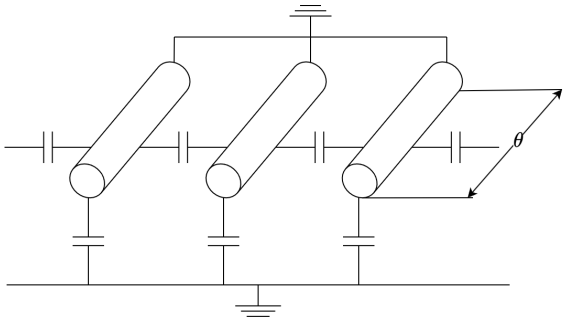


Figure 2.18: Combline elliptic filter, $\theta \leq 60°$ [5].

requires lumped elements. This filter type also does not seem to have been reported further in the literature. The realisation is visualised in Fig. 2.18.

2.10 Microfabrication techniques for filters

There are a great variety of techniques that can be employed for the manufacturing of high frequency components, all with their own strengths and weaknesses. In this section, a brief overview of different manufacturing methods is presented along with their corresponding strengths, weaknesses, and suitability for the specific filter structure.

2.10.1 RF MEMS

Though microelectromechanical (MEMS) technologies are widely used in several commercial products relating to large markets, such as airbag sensors and mobile handsets, they are yet to be readily available for RF components outside of academia [32]. RF MEMS switches are the first reported use of this technology for RF purposes [33], but have since been demonstrated for various components and purposes, including filters [34]. RF MEMS can substantially reduce the size, weight, and cost of RF components [32].

2.10.1.1 Polystrata process

The Polystrata^{TM} process can realise high Q coaxial lines. This is done through a thick-film metal micromaching process with integrated dielectric supports, that have been implemented for filters between 5 to 90 GHz. It has been shown that the size can be 10 to 100 times smaller than conventional combline or waveguide filters, but with comparable performance [35].

Even though the technology seems superior, it has drawbacks due it being a patented process. The process is very costly, making it unfeasible for most purposes and filter designers. It is also hard for larger companies to depend on a single manufacturer.

2.10.1.2 EFAB

EFAB^{TM} (electromechanical fabrication) is a technology similar to the Polystrata process, but instead suspends the resonator with a metal support [36]. This technology has been around for some time and is well established. The technology has also been shown to be suitable for compact low loss filter design [37].

However, since it requires a metal support, this introduces an inductance which affects the filter performance negatively and makes some filter characteristics impossible to realise.

2.10.1.3 Etching

Due to the fact that the resonating parts of the stepped digital elliptic filter are all equally thick, the spacings between the resonators could be etched (potentially in one layer) revealing the structure of the two networks, which can be seen in Fig. 2.19.

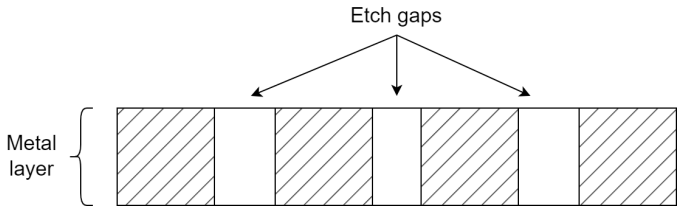


Figure 2.19: Possible etching of gaps to form resonators.

Normally, a process like Polystrata is required for these types of filter because the conductors are suspended. However, since the conductors are shorted at one end, they are already in a way suspended to the grounded wall, and do not need additional support. This is visualised in Fig. 2.20, from which it is obvious that this specific filter realisation does not require support in the form of dielectric or other materials.

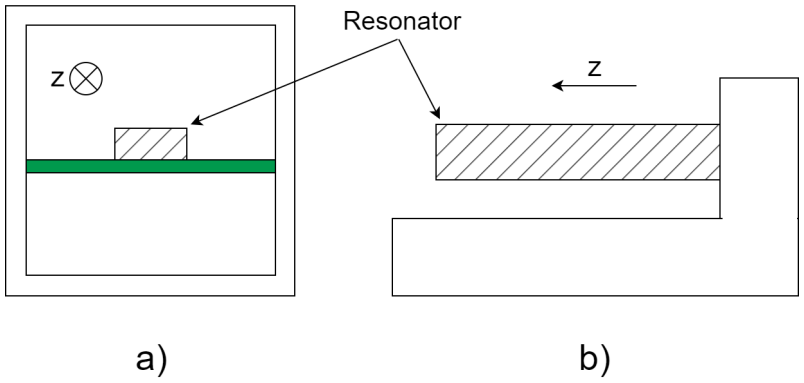


Figure 2.20: Comparison between a) simplified Polystrata process resonator resting on a dielectric support, and b) this filter structure.

3

Filter design method

In this chapter, the general filter design is explained, with methods based on the theory introduced in the previous chapter. In addition, specific design choices will be explained, along with a description of the simulations and optimisations and how they were conducted.

3.1 Design methodology

First, the theoretical steps of this filter realisation will be retraced. This step is essential; thorough understanding of the filter function and what affects it is crucial in order to be able to improve it. Thus, a script calculating the per unit-length capacitance matrices, the accompanying filter function, and a rough estimation of its physical realisation was implemented in MATLAB, where all possible variables could be changed and, in turn, optimised. This allowed for all possible variables to be analysed without the use of time-consuming three-dimensional electromagnetic simulations.

Secondly, this rough estimation of the physical realisation needed to be double checked to validate its accuracy and consistency. Again, in order to avoid the aforementioned time-consuming simulations, the per unit-length coupling matrices of the model were simulated in MAXWELL2D. The physical dimensions could then be adjusted until the coupling matrices were more or less identical to the theoretical ones derived in MATLAB. This adjustment procedure was automated through the built-in optimisation included in the MAXWELL2D software.

Third, the adjusted model was then simulated in three dimensions in HFSS to validate that the per unit-length model was accurate and to see how well the filter function of the physical structure aligns with the theoretical one. Once validated, all practical considerations needed to be taken into account. The physical structure, and certain elements needed in order to fabricate it, affects the filter function, and therefore, the structure had to be adjusted to counteract these changes. These changes were first theoretically estimated to find a good guess, and then optimised in HFSS. A tolerance analysis was performed to see how far the physical dimensions can be deviated due to, for example, fabrication errors.

3.2 Filter requirements

Since the importance of this thesis lies not in creating a filter for a specific application, but more as a study on the realisation of this filter, there is not a specific set of requirements.

However, as a general rule, the return loss in the passband should ideally be below $-20\,\mathrm{dB}$ for a good filter. The design must also be fabricatable, meaning that the capacitance matrices must correspond to physical dimensions of the filter that are not unphysical.

3.3 Filter network synthesis

As derived in section 2.17a, resonating π-sections had to be connected to one another depending on the degree of the filter. The required values of the lumped elements are already calculated in [20] for many different filter characteristics and degrees. The wanted filter characteristics that oblige to certain filter requirements could therefore be chosen simply and directly from these tables - which come with related lumped element values - from which the related capacitance matrices are found.

3.3.1 Finding the capacitance matrices

In order to obtain the theoretical filter function from the tabulated values, one has to start with the capacitance model and the related resonating π-section defined in figures 2.14 and 2.17a. From the tabulated capacitance values $\tilde{c_1}$ to $\tilde{c_n}$ and transmission zero locations Ω_1 to Ω_{n-1}, the transformed capacitance and inductance values in the resonating section are found using the transformations in table 2.2. This provides two capacitance matrices (3.2a) and (3.2b), one for each network, where

$$c_k = a\tilde{c_k} \tag{3.1a}$$

$$l_k = \frac{1}{c_k} \tag{3.1b}$$

$$c_{k,k+1} = a\tilde{c_k}(1+\lambda_{k+}^2),\ \text{if } k \text{ even} \tag{3.1c}$$

$$c_{k,k+1} = a\tilde{c_k}(1+\lambda_{k-}^2),\ \text{if } k \text{ odd} \tag{3.1d}$$

with $\lambda_\pm$ as defined in eq. (2.30).

$$\mathbf{C}^A_{n\times n} = \begin{bmatrix} c_1^A + c_{1,2}^A & -c_{1,2}^A & \dots & 0 \\ -c_{1,2}^A & c_2^A + c_{1,2}^A + c_{2,3}^A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_n^A + c_{n-1,n}^A \end{bmatrix} \tag{3.2a}$$

$$\mathbf{C}^B_{n\times n} = \begin{bmatrix} c_1^B + c_{1,2}^B & -c_{1,2}^B & \dots & 0 \\ -c_{1,2}^B & c_2^B + c_{1,2}^B + c_{2,3}^B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_n^B + c_{n-1,n}^B \end{bmatrix} \tag{3.2b}$$

Unit elements are now introduced to network 1 as shown in (2.39), and every row and every column of both matrices were then scaled by

$$\sqrt{\frac{1}{a'}} = \sqrt{\frac{\eta}{aZ_0}} \tag{3.3}$$

except for the rows and columns belonging to the unit elements. This means that the entirety of the $\mathbf{C}_{n\times n}$ matrices are scaled by $1/a'$, while the off diagonal unit elements are scaled by $1/\sqrt{a'}$ and the diagonal unit elements are not scaled at all.

As explained in section 2.8.1, scaling of the capacitance matrix does not affect the filter characteristics, only the physical realisation. Therefore, the only purpose of the following scaling is to arrive at a physically realisable structure. Knowing this, the diagonal is scaled by n_{diag} using the method shown in 2.40.

Once the diagonals had been scaled, the initial scaling constants $n'_1, \ldots, n'_n$ could be found through

$$n'_k = \frac{\sqrt{C^*_{k,k}}}{C_{k,k}} \tag{3.4}$$

where the asterisk $(*)$ denotes the diagonally scaled matrix. These scaling constants are now applied to every off-diagonal element (since the diagonal has already been scaled) by the method shown in 2.39. These initial scaling constants n'_k are not to be confused with scaling constants n_k, which can be freely decided to scale the capacitance matrix as needed at a later stage.

3.3.2 Extracting the filter function

The two initial capacitance matrices containing all the relevant information about the filter are now found. However, the capacitance matrices have not yet been connected to any frequency.

In order to extract the filter function, the methods introduced in 2.8.3 need to be implemented. The admittance and impedance matrices for the specified centre frequency ω_c are found through

$$\mathbf{Y} = j\omega_c \varepsilon \mathbf{C} \tag{3.5a}$$

$$\mathbf{Z} = j\omega_c / \varepsilon v_p^2 \mathbf{C} \tag{3.5b}$$

where the ε is needed since the capacitance matrix was derived per unit-length. These matrices now require diagonalisation, which is achieved by solving the characteristic equation for the eigenvalues and eigenvectors through equation (2.44). The characteristic admittance matrix $\mathbf{Y_0}$ is then found through equation (2.45).

In order to connect the two networks and to set port conditions, $2n$-port models of the two networks are required. The method of converting the characteristic admittance matrix into a $2n$-port is described in 2.8.4. Once the two matrices have been converted, they now need to be connected and to have their port conditions set. This is visualised in figure 3.1.

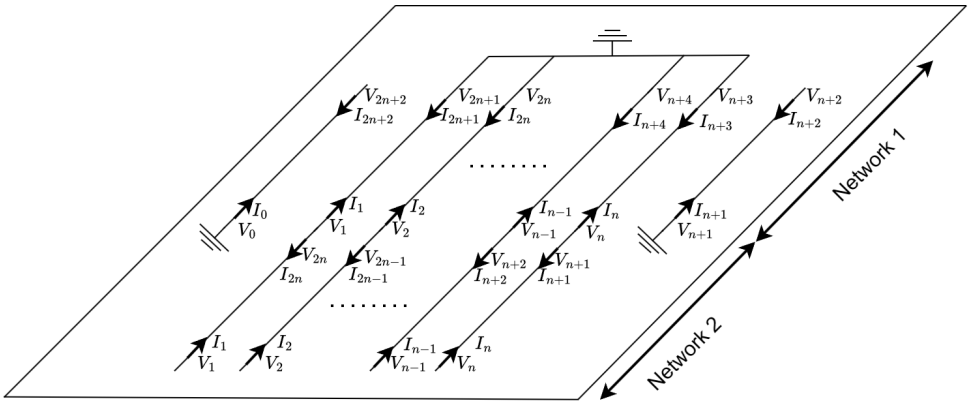


Figure 3.1: Connecting the two $2n$-port models and introducing port conditions. The numbering of the nodes is repeated due to the two individual networks.

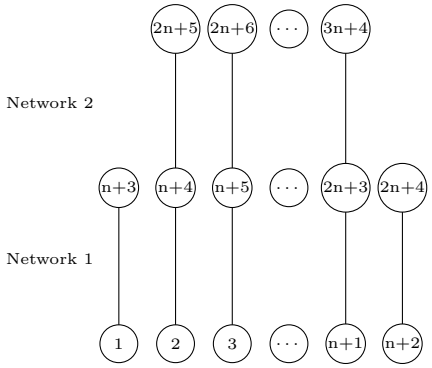


Figure 3.2: Node restructuring of the model in figure 3.1.

Connecting the two $2n$-port models as in figure 3.1 one combined characteristic admittance matrix, $\mathbf{Y}^{\text{combined}}_{3n+4\times 3n+4}$, is created which combines the admittances at all common nodes. In order to emulate the short circuits (i.e $I = 0$) at some of the ports belonging to network 1, the original admittance value at these ports are replaced by a very large value.

To simplify figure 3.1 and to visualise the combined matrix, all ports are renumbered to shared common numbers as in figure 3.2, which also simplifies the analysis.

Looking at figure 3.2 a $(3n + 4) \times (3n + 4)$ matrix can be constructed, where the diagonal elements are the nodal admittances and the off-diagonal elements are couplings between adjacent elements. Comparing with figure 3.1 the port conditions are applied to the corresponding nodes, and the input and output nodes are identified as nodes 1 and $n + 2$. Since these input and output nodes should have the overall system impedance level, these nodes are set to the desired system impedance level Z_0, which is normally $50\,\Omega$.

The 2-port S-parameters can thus be extracted from these nodes using equation (2.12), and the theoretical filter function is now derived.

3.4 Relating capacitance to a physical structure

Once the matrix $\mathbf{Y}^{\text{combined}}_{3n+4\times 3n+4}$ with its applied port conditions and its corresponding filter function has been verified, the capacitance matrices must now be related to a physical structure. One common method is to use Getsinger's graphs [25], which approximately connect the widths and spacings of the rectangular conductor to the capacitance values corresponding to the given conductor thickness-to-ground plate spacing ratios. This was also the way Rhodes found the physical dimensions in his paper [6]. One major flaw with this method is that the accuracy depends on the degree to which the graphs can be read, which also becomes harder at extreme thickness to ground plate ratios. Thus, a more general method is needed that also minimises the number of calculations, allowing for optimisation of the structure at a later stage.

3.4.1 Width and spacing approximations

The general method of approximating self-capacitance and coupling capacitance relating to the widths and spacings, respectively, is through conformal mapping techniques, namely the Schwarz-Christoffel transformation, which is also the way in which Getsinger derived his graphs [25] (see the Appendix of the referenced paper). These techniques are used in order to obtain analytical formulas for zero-thickness conductors, and approximate analytical formulas for conductors with a finite thickness.

For a strip of zero thickness between two infinite ground planes as in Fig. 3.3, the capacitance can be found through [38]

$$\frac{C_g}{\varepsilon} = 4\frac{K\left(\tanh\left(\frac{\pi}{2}\frac{w}{b}\right)\right)}{K\left(\operatorname{sech}\left(\frac{\pi}{2}\frac{w}{b}\right)\right)}, \tag{3.6}$$

where the subscript g stands for ground.

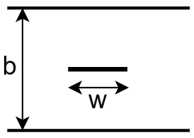


Figure 3.3: Zero-thickness conductor.

Since the conductors used will not be of zero-thickness, but rather relatively thick, this approximation will not be enough. If it is assumed that there is no interaction between fields below and above the strip, and assuming that there are no fields between the sides of the conductor to the ground plates, the formula can be adjusted to

$$\frac{C_g}{\varepsilon} = 4\frac{K\left(\tanh\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)}{K\left(\operatorname{sech}\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)} \tag{3.7}$$

where b' is the effective ground plate spacing by thickening the conductor, i.e.

$$b' = b - t. \tag{3.8}$$

The assumption that there are no fields between the sides of the conductor to the ground plates is clearly approximate, unless the ground plate spacing is very large compared to the conductor thickness (also making any capacitance negligible). A "thickness correction"-term thus needs to be added. One such approximate term has been derived in [39] and can be simply added as

$$\frac{C_g}{\varepsilon} = 4\frac{K\left(\tanh\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)}{K\left(\operatorname{sech}\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)} + \frac{2t}{b} \tag{3.9}$$

which gives the approximation for the capacitance between conductor and ground per unit-length used for finding the physical size of the conductors.

The capacitance between the lines is simply calculated as the parallel plate capacitance between the lines given by

$$\frac{C_{k,k+1}}{\varepsilon} = \frac{t}{d_{k,k+1}} \tag{3.10}$$

where d is the spacing between two conductors. This equation assumes that there is no coupling between the conductors anywhere else than between the sides of the conductors, which should be approximately true as long as the effective ground plate spacing b' is not too large.

3.4.1.1 Evaluating the width and spacing approximations

The spacing approximation is easily evaluated for a given capacitance matrix and conductor thickness through

$$d_{k,k+1} = \frac{t}{C_{k,k+1}/\varepsilon} \tag{3.11}$$

which is plotted against simulated values in Fig. 3.4. This simple case was simulated in MAXWELL2D as described in the figure.

The width approximation (3.9) must be numerically evaluated, due to the elliptic and hyperbolic functions. The simple case of one conductor between two ground plates was simulated in MAXWELL2D, and the elliptic integrals of equation (3.9) were evaluated using MATLAB's built-in command **ellipke**. As is evident from Fig. 3.5a, the approximation of the width is wrong by some constant at larger width values, and the curve deviates at smaller width values. The constant was identified as approximately another thickness correction term, thus the corrected width approximation is found to be

$$\frac{C_g}{\varepsilon} = 4\frac{K\left(\tanh\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)}{K\left(\operatorname{sech}\left(\frac{\pi}{2}\frac{w}{b'}\right)\right)} + \frac{4t}{b}. \tag{3.12}$$

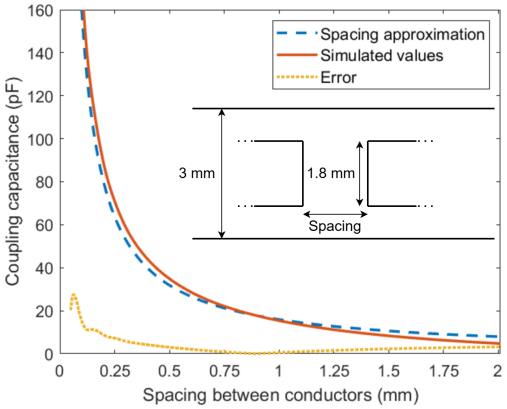
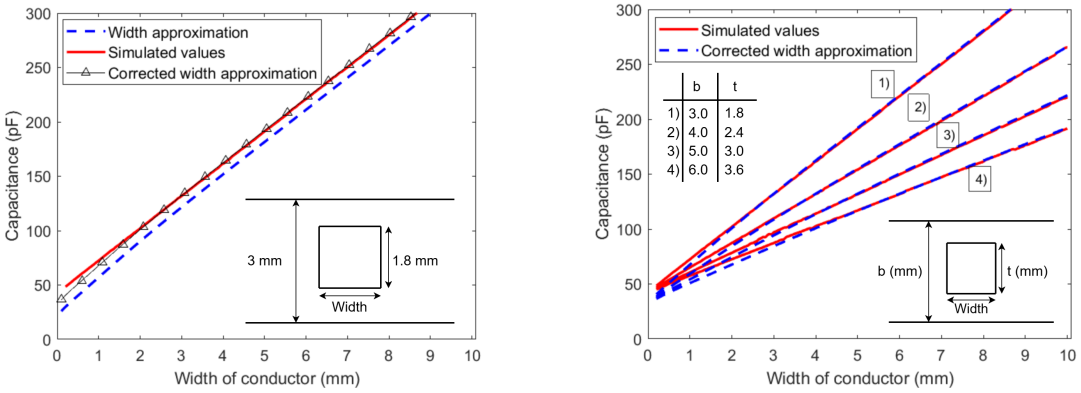


Figure 3.4: Evaluation of the spacing approximation (3.11) for the given dimensions.



(a) Correcting the approximation.

(b) Simulated and approximated values.

Figure 3.5: Evaluation of width approximation and corrected width approximation in comparison to simulated values.

This corrected width approximation is, as seen in Fig. 3.5b, accurate for wider conductors for different ground plate spacings and conductor thicknesses. At very small widths, some unknown effect occurs that is not accounted for. However, it is likely that this is due to field interactions between the different sides of the conductors.

Although the corrected approximation of equation (3.12) is more accurate for this simple case with only one conductor, in reality, the coupling between the sides and the ground plates will compete with the coupling between the sides of the conductors. For most cases the spacings are small enough to make this matter, causing the approximation in equation (3.9) to be more correct than the approximation in (3.12) for all sizes studied and was thus used instead of the corrected approximation.

3.4.2 From capacitance matrix to physical structure

As mentioned in 2.8.1, the diagonal values of the capacitance matrix contain both the capacitance to the ground plates as well as the couplings to the adjacent lines. Thus, the capacitance C_g in equation (3.12) is found from the capacitance matrix

as

$$C^A_{k,k} = \begin{cases} C_{0,0} - |C_{1,0}| & \text{if k=0} \\ C_{k,k} - |C_{k-1,k}| - |C_{k+1,k}| & \text{if k}\neq\text{0 or n+1} \\ C_{n+1,n+1} - |C_{n,n+1}| & \text{if k=n+1} \end{cases} \tag{3.13a}$$

$$C^B_{k,k} = \begin{cases} C_{1,1} - |C_{2,1}| & \text{if k=1} \\ C_{k,k} - |C_{k-1,k}| - |C_{k+1,k}| & \text{if k}\neq\text{1 or n} \\ C_{n,n} - |C_{n-1,n}| & \text{if k=n} \end{cases} \tag{3.13b}$$

for networks 1 and 2 respectively. The difference in numbering between the two matrices comes from the added unit elements at node 0 and 6 for network 1.

Given that both the thickness of the conductors and the ground plate spacing are known, all physical dimensions can now be calculated through equations (3.11) and (3.12) and an initial estimate of the filter prototype can be found.

If the ground capacitance is found to be negative from these expressions, the capacitance matrix is unphysical (negative conductor width) and thus cannot be realised as a physical structure.

3.5 Network alignment

Since the stepped digital elliptic filter consists of two separate networks, computed individually, great care has to be taken into proper alignment of the two networks. Because the physical dimensions of the filter depend on the capacitance matrix, correct scaling is of utmost importance (see 2.8.2). The synthesis is limited to realisations using rectangular conductors.

In order to create a homogeneous structure, i.e. a structure where the two networks align properly, two different approaches will be evaluated:

i) Node alignment optimisation,

ii) Capacitance matrix optimisation.

Both of these approaches boil down to finding a number that describes the homogenuity and the alignment of the structure, with the node alignment approach focusing solely on the physical dimensions and the capacitance matrix approach focusing on finding the optimal capacitance matrix that should then later transform into a homogeneous and well-aligned structure.

3.5.1 Node alignment optimisation

The node alignment approach consists of trying to find good approximations for the widths of and spacings between the conductors relating to the capacitance matrix defined in 2.8.1. Using the capacitance model for the interdigital filter (see figure 2.14), and explicitly defining the physical elements as in Fig. 3.6, a figure of merit (FoM) can be defined.

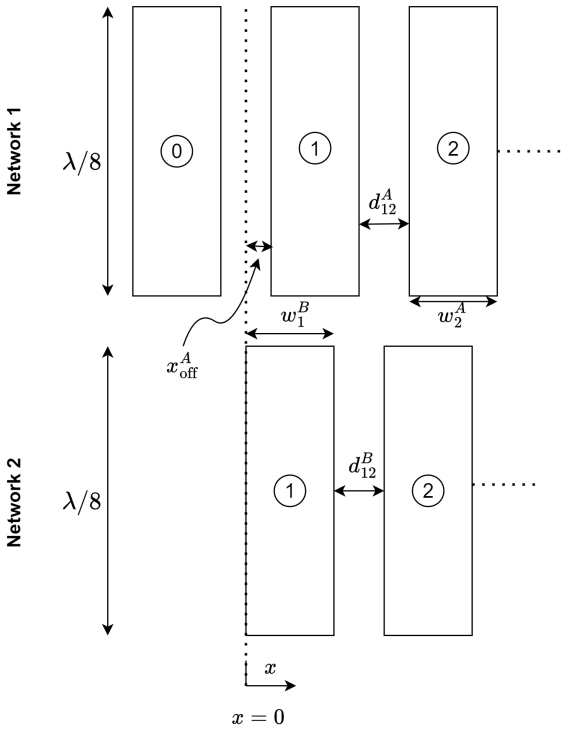


Figure 3.6: Definitions of variables used in node alignment calculation. Compare with figures 2.14 and 2.16 for reference.

As seen in Figs. 2.14 and 3.6, w_k is the width of conductor k, and $d_{k,k+1}$ is the spacing between conductors k and $k+1$. Additionally, with $x = 0$ being the reference beginning of both networks (excluding the leftmost UE as it does not require alignment), x_{off} is an offset from this point of reference, either to the right or left (i.e., along the width dimension). The subscripts A and B denote networks 1 and 2 respectively.

Thus, the node alignment NA for node k is given by

$$\text{NA}_k = x_{\text{off}}^A - x_{\text{off}}^B + \frac{w_k^A}{2} - \frac{w_k^B}{2} + \sum_{i=1}^{k-1}(w_i^A - w_i^B) + \sum_{i=1}^{k-1}(d_{i,i+1}^A - d_{i,i+1}^B) \tag{3.14}$$

where NA_k is defined as the difference between the middles of the two conductors at node k, noting that the summations are zero for $k = 1$. By rewriting all differences as $x_{\text{off}}^A - x_{\text{off}}^B = \Delta x_{\text{off}}$ and so on the node alignment expression can be given as

$$\text{NA}_k = \Delta x_{\text{off}} + \frac{\Delta w_k}{2} + \sum_{i=1}^{k-1}\Delta w_i + \sum_{i=1}^{k-1}\Delta d_{i,i+1} \tag{3.15}$$

where Δ represents the difference. A positive NA_k means that node k of network 1 is aligned slightly right off of the center of node k of network 2, and a negative NA_k means the opposite.

The figure of merit is then calculated as

$$\text{FoM} = \sum_{i=1}^{k}|\text{NA}_i| \tag{3.16}$$

which essentially means that the nodal misalignments of all conductors are summed. Therefore, if the FoM is minimised, the similarity of the two networks should be optimal.

One obvious disadvantage of this merit figure is that the widths of the different conductors could differ quite a lot between two different nodes, which might not be optimal. It should, however, find the solution where the two networks are, on average, most aligned, and thus the conductor widths should converge appropriately.

3.5.2 Capacitance matrix optimisation

The second approach instead targets the capacitance matrices directly without connecting them to physical dimensions. This is done by investigating how the scaling constants affect the capacitance matrices, not just finding the optimal ones as in section 3.5.1.

To evaluate which scaling constants are generally preferred, ground capacitance and coupling capacitance variations both within and between the networks were analysed. The capacitance variation is defined as

$$C_{\text{var}} = \max(C_g) - \min(C_g) \tag{3.17}$$

between all internal nodes. This measure analyses how large the capacitance variations and therefore widths are within one network. A large value of C_{var} would indicate a non-homogeneous structure.

While the variation in one network might be important, nodal alignment between the networks is more important. Thus, a mean nodal capacitance difference is defined as

$$C_{\text{diff}} = \frac{1}{n}\sum_{k=1}^{n}\left|C_{g,k}^{A} - C_{g,k}^{B}\right| \tag{3.18}$$

where n is the degree of the filter. If C_{diff} is minimised, the nodal alignment should be optimal. Variations and mean nodal differences are defined analogously for the coupling capacitance.

3.6 Fabrication considerations

In the previous sections, the filter structure analysed has been ideal. In order to make the filter fabricatable, the filter structure requires to be suspended. Suspending the filter requires a metal housing to which the conductors, and eventually, some way of connecting a cable to the filter, can be attached.

3.6.1 Correction for capacitances to metal housing

Since this type of filter realisation requires some metal housing, there will be unwanted couplings unless the space between the conductor and the metal wall is very large; and, therefore, not very compact. Instead, the length of every conductor in

network 2 has to be shortened by the amount of capacitance added by metal housing, and the first and last conductors of both networks will be made thinner due to coupling along the sides. This coupling capacitance is visualised in figure 3.7, together with the definition of l_g, which is the length between the ends of network 2 and the metal housing.

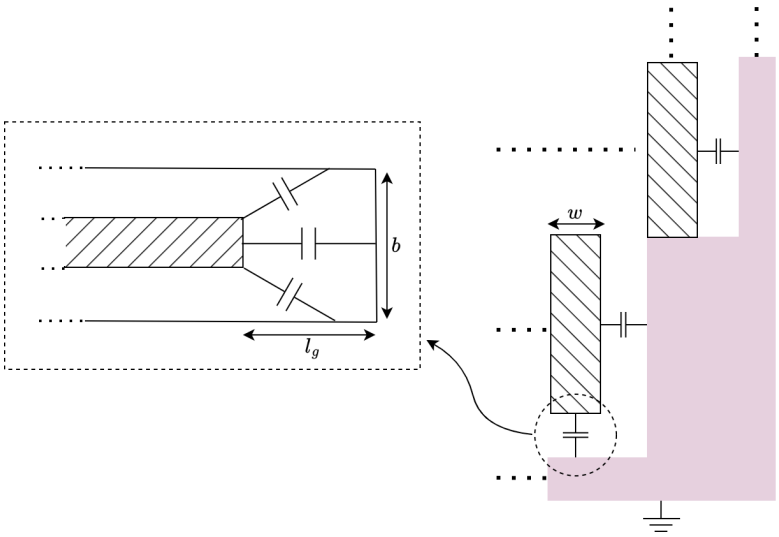


Figure 3.7: End wall capacitances for an example metal housing structure, where the unit element of network 1 is shorted by direct connection with the metal housing. Compare with figure 2.16 for reference.

The capacitances added by the side walls for network 2 are

$$\frac{C_{\text{side}}}{\varepsilon} = \frac{t}{d_{\text{side}}} \tag{3.19}$$

where d_{side} is the distance between the side walls and the closest conductor. With this particular housing structure shown in figure 3.7, the distances d_{side} are given by

$$d_{\text{side}} = \begin{cases} d_{01}^{A} + x_{\text{off}}^{B} - x_{\text{off}}^{A} & \text{(3.20a)} \\ \sum_{1}^{N} d_{i,i+1}^{A} + \sum_{2}^{N-1} + x_{\text{off}}^{A} - x_{\text{off}}^{B} - \sum_{1}^{N} d_{i,i+1}^{B} - \sum_{1}^{N} w_i^{B} & \text{(3.20b)} \\ N = \text{filter degree} \end{cases}$$

where eq. (3.20a) is the distance between the first conductor of network 2 and the side wall, and eq. (3.20b) is the distance between the conductor N of network 2 and the side wall. These lengths are easily derived through inspection of figures 3.6 and 3.7.

The length l_{g} can be chosen arbitrarily; but a more compact design is favourable and therefore the metal housing will have an effect by increasing the overall capacitance of the line, while not affecting the coupling between the lines. To compensate for the extra capacitance C_{end}, the length of the individual lines will need to be adjusted to reduce the added capacitance. The capacitance added to the ends of network 2 are

$$\frac{C_{\text{end}}}{\varepsilon} = \frac{t \cdot w}{l_{\text{g}}} \tag{3.21}$$

where the capacitance is now multiplied with the width because the capacitance is along the length axis. The thickness times the width is the area of the end of the open-ended line.

To approximate the amount by which individual lines are to be shortened, variables C_{line}, C'_{line}, l', and Δl must be defined. Here, they represent the capacitance of the original length line, the capacitance of the line of adjusted length, the adjusted line length, and the difference in original and adjusted length, respectively. See also figure 3.8. It is thus obvious that if the following relationship holds,

$$C_{\text{line}} l = C'_{\text{line}} l' + C_{\text{end}}, \tag{3.22}$$

this compensation is achieved.

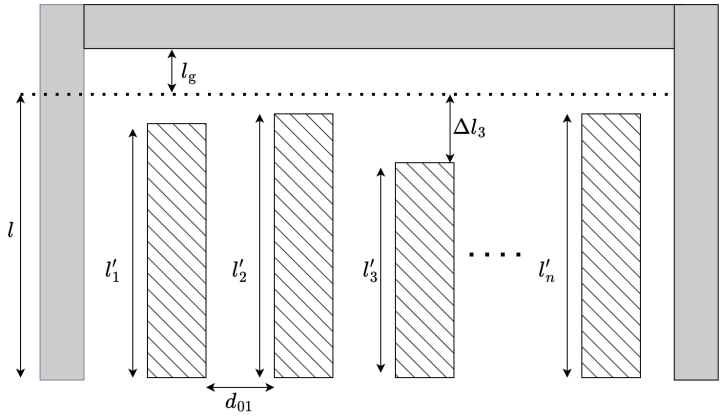


Figure 3.8: Individually adjusted lengths at the end of network 2.

With adjusted lines lengths of network 2, an adjusted variant of equation (3.21) is defined to be

$$\frac{C'_{\text{end}}}{\varepsilon} = \frac{t \cdot w}{l_{\text{g}} + \Delta l}, \tag{3.23}$$

noting that the length adjustment $\Delta l = l - l'$. Using (3.22) and (3.23) the resulting expression becomes

$$\frac{C_{\text{line}}}{\varepsilon} = \frac{t \cdot w}{\Delta l (l_{\text{g}} + \Delta l)} \tag{3.24}$$

which has the solution

$$\Delta l = \frac{1}{2}\left(\sqrt{l_{\text{g}}^2 + 4\frac{t \cdot w}{C_{\text{line}}/\varepsilon}} - l_{\text{g}}\right) \tag{3.25}$$

for a given l_{g}. Crucially, equation (3.24) demonstrates that due to the capacitance's dependence on width, individual adjustments to line lengths will be necessary based on their respective widths. Values of the length adjustment Δl for different end length walls l_{g} over capacitance is plotted in Fig. 3.9.

Intuitively, this equation should also include the spacings between conductors, since a reduced length will reduce the coupling between lines and, therefore, also the total

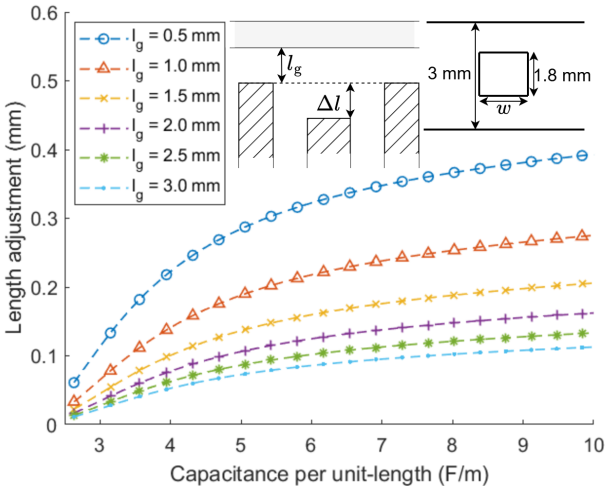


Figure 3.9: Length adjustment Δl for different end wall lengths l_g using the dimensions specified in the figure. The capacitances are related to different widths using either eq. (3.12) or Fig. 3.5b.

capacitance of the line. However, given that couplings between the ends of the lines have not been integrated into the theory, it will be assumed that this coupling gain and loss balance each other out.

It also does not include the edge areas coupling to the top and bottom ground plates, which is probably the largest error contribution if the end length l_g is longer than the distance between the distance from the centre of the line to the ground plate, i.e. $b/2$. Thus, if l_g is relatively large compared to b', the fringing capacitances of an isolated rectangular bar C'_f [25] can be incorporated into equation (3.21) as

$$C_{\text{end}} = \frac{t \cdot w}{l_g} + 2C'_f \tag{3.26}$$

to more accurately estimate Δl for these cases.

3.6.2 Excitation options

In order to be able to measure this filter, it needs to be able to be connected to standard measurement cables like 50 Ω coaxial cables. Two different options for excitation of the filter are briefly introduced and discussed.

3.6.2.1 Replacing UE with end-launch SMA pin

By replacing the UEs with the pin of an end-launch SMA connector as shown in Fig. 3.10, the desired coupling could be achieved directly without the need for transition design. By electrically connecting the pin to the wall, by for example solder, and having the pin be $\lambda/8$ long inside the structure, the pin acts just as the rectangular UE. By changing the spacing between the pin and the rectangular connector at node 1, the same coupling capacitance can be achieved with this input. Furthermore, by adding a metal housing wall to the left of the circular pin in Fig. 3.10, the ground

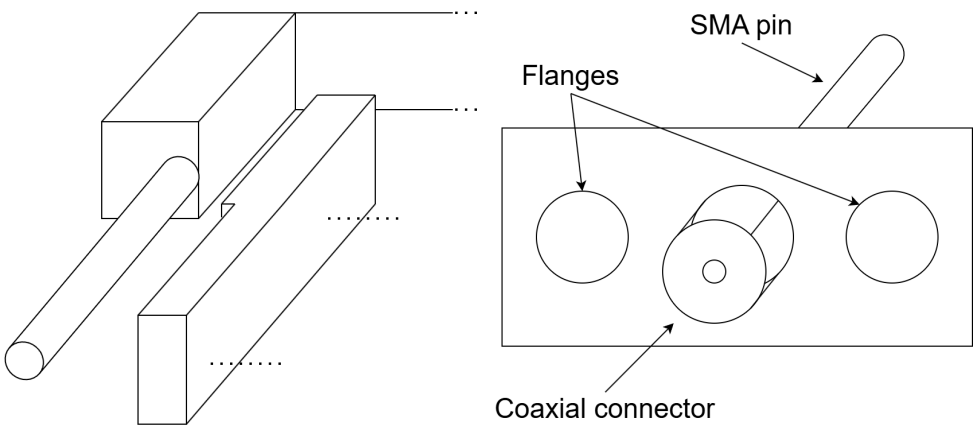


Figure 3.10: UE replaced by SMA end-launch pin.

capacitance C_g can also be matched to the desired capacitance matrix. This coupling can be seen in the top right of Fig. 3.7. The same procedure can then be applied to replace the second UE.

This design works as long as the ground plate spacing is larger than the diameter of the pin, which is normally around 1.3 mm in diameter. It is also dependent on frequency, the pin has to be long enough to reach the $\lambda/8$ required length inside the structure. However, since there are several different commercially available SMA-connectors, finding a suitable connector should be feasible for most frequencies.

3.6.2.2 Suspended stripline to microstrip to SMA transition

Since the conductors in the structure are essentially equal to thick suspended striplines (SSP), a transition between an SSP and the rectangular UE could be a viable option. SSPs also have the problem that they are not easily connected to a common end-launch SMA connector. However, a microstrip fits perfectly for this purpose, and SSP-to-microstrip transitions have been reported in the literature [40].

3.7 Simulations and computer aided optimisation

In order to be able to test different types of structures, filter characteristics, etc., the whole procedure discussed in section 3.3 was programmed in MATLAB. Thus, the theoretical aspects of the filter design and an initial physical structure could be evaluated without the use of CAD tools. In order to validate and correct the calculations in MATLAB, a 2D model was created for each network in MAXWELL2D. Lastly, the 3D model of the filter was created, tested, and optimised in HFFS in which the final filter prototype was designed. The workflow is shown in figure 3.11 and will be more thoroughly described in the following sections.

3.7.1 MATLAB simulations and optimisations

To avoid time-consuming electromagnetic simulations, as much of the design process as possible was programmed in MATLAB. This not only sped up the design process

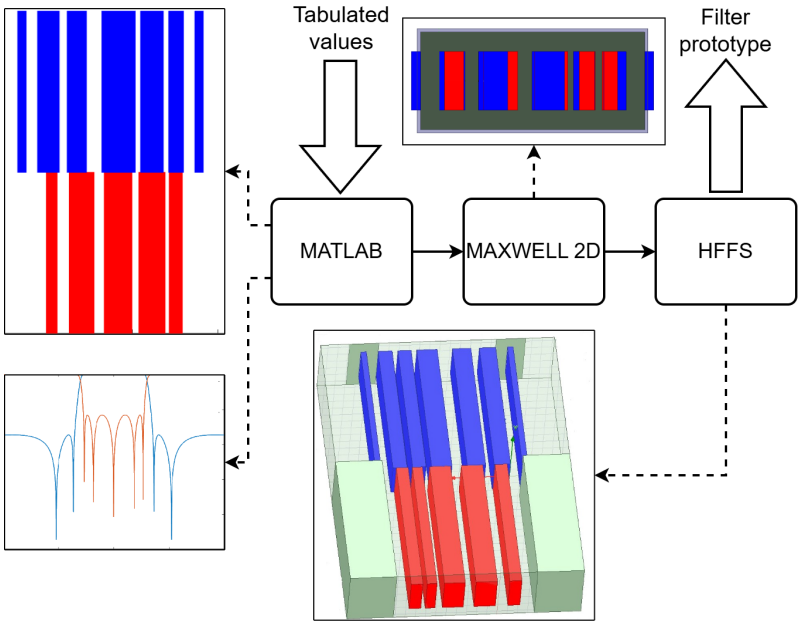


Figure 3.11: Workflow from theoretical tabulated values as input to a filter prototype as output. The blue conductors represent network 1 and the red network 2.

but also allowed for larger control over the simulation process. Parts of the MATLAB code are included in appendix A.

3.7.1.1 Initial estimate of filter prototype

Starting from tabulated values that relate to specific filter characteristics, the capacitance matrices are found as shown in section 3.3.1. Given these matrices, conductor thickness, and ground plate spacing, an initial physical structure can be realised by relating the widths and spacings to the corresponding capacitance values as described in section 3.4. If there is no width that satisfies this condition, then the capacitance matrix is deemed physically unrealisable and thus needs further optimisation as discussed in section 3.5. The initial physical structure is then passed through the FoM calculation of equation (3.16), through which a measure of the network alignment is given. To summarise the steps used to derive this estimate:

1. Choose filter characteristics
2. Find corresponding tabulated values in [20]
3. Calculate the two capacitance matrices $\mathbf{C}^A_{n+2}$ and $\mathbf{C}^B_n$
4. Calculate the admittance matrix $\mathbf{Y}^{\text{combined}}_{3n+4\times 3n+4}$ to verify S-parameters
5. Relate diagonal values of $\mathbf{C}^A_{n+2}$ and $\mathbf{C}^B_n$ to conductor widths (eq. (3.9))
6. Relate off-diagonal values of $\mathbf{C}^A_{n+2}$ and $\mathbf{C}^B_n$ to conductor spacings (eq. (3.10))
7. Apply corrections to lengths of network 2
8. Measure network alignment using equation (3.16)

If two conductors corresponding to the same common node in Fig. 3.2 are completely disattached, as is the case in Fig. 4.7a, the filter is physically unrealisable and electrically not equivalent to the resonating sections of figure 2.17a. To fix this,

one could either manually change the scaling of the capacitance matrices or let an optimiser calculate the optimum scaling constants. However, since the effect of changing scaling constants on larger matrices is hard to anticipate, optimisation is sometimes required.

3.7.1.2 Network alignment optimisation tools and prerequisites

To employ the network alignment optimisation schemes introduced in section 3.5, the **Global Optimization Toolbox** was used in MATLAB [41]. This toolbox allows for global optimisations, which is needed since there are an infinite number of physical structures that realise this filter function, and therefore also an infinite number of structures where the networks align well. To ensure that the optimiser finds the global optimum (or close to), the option **Multistart** within this toolbox is used. This allows the optimiser to start from different points within the given boundaries, thus not locking itself in at values close to the initial guess. However, even though a global optimiser is used, well-estimated parameter boundaries are still preferred due to the high computational cost associated with these optimisations.

In order to visually inspect the node alignment and to give sense to the FoM value, a visualisation function was developed in MATLAB. An example of this visualisation function is visible in the upper left of Figure 3.11. Using this figure to represent the FoM value simplified the analysis of whether a specific FoM number related to a realisable structure or not.

3.7.1.3 Node alignment optimisation

In order to optimise the nodal alignment given the initial estimate of the filter prototype, the FoM equation (3.16) is used as an objective function with the Sequential Quadratic Programming (SQP) optimisation algorithm. SQP algorithms are a group of constrained non-linear optimisation algorithms with Schittkowski [42] having published the most efficient method [43]. A variant of this method is included in the aforementioned toolbox. The structure is aligned through optimisation with

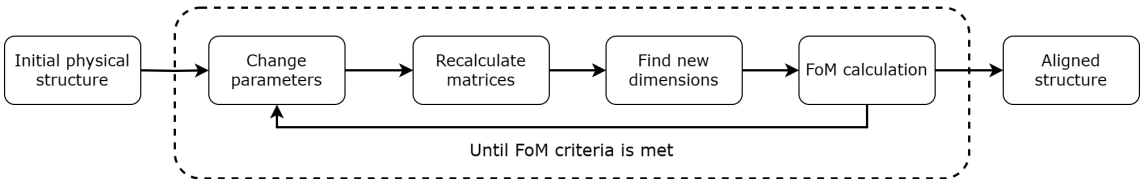


Figure 3.12: A visualisation of the optimisation method.

the steps shown in figure 3.12. The parameters are generally the scaling constants n_k and n_{diag}, but other parameters could also be changed.

Through this optimisation procedure, it could quickly be evaluated if certain filter characteristics are realisable or not.

3.7.2 MAXWELL2D simulations

Once a physical structure is found that satisfies the FoM criteria, a 2D version (zero length) of the structure was simulated in MAXWELL2D with an example shown in Fig. 3.11. This software is capable of calculating the capacitance matrices per unit-length of 2D-structures. This allowed the approximate dimensions found through MATLAB to be double-checked and corrected for each individual network.

3.7.2.1 2D-correction of the initial structure

Two models were created, one for each network, which also included the side walls of the metal housing introduced in section 3.6.1. Each conductor was then excited by an equal, arbitrary voltage source, and the ground plates and the side walls were set as ground. By relating each of these excitations to a matrix, the capacitance matrix of the network is automatically found after analysis. Then, by comparison, the dimensions in MAXWELL2D could be adjusted until the simulated and theoretically derived matrices were more or less identical.

3.7.3 HFSS simulations and optimisations

Once a sufficient 2D model of the filter prototype has been identified, a 3D filter prototype was created in HFSS with an example shown in Fig. 3.11. All simulations assumed perfect electrical conductors (PEC), which was later compared with a worst-case scenario using aluminium conductors.

3.7.3.1 Optimisation

Since there are always electromagnetic effects that have not been accounted for, or have been only approximately accounted for, an optimisation is required.

For a filter of degree n, there are $5n + 2$ parameters directly related to the widths, spacings, lengths, and alignment of the structure. This quickly becomes cumbersome even for fairly low-degree filters. Therefore, the most important parameters needed to be identified so that these parameters could be optimised first. Once the most important parameters were optimised, the next most important parameters were optimised and so on.

Two different built-in optimisation algorithms were used, the **Genetic Algorithm** (GA) and the **Mixed-Integer SQP** (MISQP). The GA is extremely time-consuming. However, it performs very well even when far from the optimal solution and with many different parameters. It, however, rarely meets the wanted optimisation criteria, or requires too many iterations to do so. Therefore, a combination between the two algorithms seems to be ideal; first, finding good initial values from the GA, and then fine-tuning optimisation using MISQP.

3.8 Tolerance analysis

In order to verify the design's robustness toward fabrication errors, a tolerance analysis will be carried out. The return loss of the filter seems to be generally more sensitive to fabrication imperfections as opposed to the insertion loss, and thus, the tolerance analysis focused on how errors in the capacitance matrices affected it.

Three different approaches will be tested:

i) error only on the most sensitive element of the capacitance matrix,
ii) errors normally distributed on lower triangular part of matrix,
iii) percentage error of all matrix elements.

It is clear from equation 3.10 that a higher coupling capacitance in the capacitance matrix indicates an increased sensitivity to differences in physical size. Therefore, it can be assumed that changes on the most sensitive parts of the coupling matrix due to fabrication imperfections will be responsible for a majority of the deterioration of the filter function. This element is identified as the maximum off-diagonal value and is substituted with a value calculated by (3.10) for a distance $d_{k,k+1} \pm \Delta d_{k,k+1}$, where the error is the tolerance of interest.

To evaluate ii), it is assumed that the errors follow a normal distribution. Thus, the error introduced to each element will be normally distributed where $x \pm \Delta x$ are assumed to be the limits of the 95 % confidence interval, where Δx would be the precision specified by a manufacturer. In this case, x represents either the width or the spacing between conductors depending on the element of the matrix. Once the errors have been distributed on the lower part of the matrix, the symmetry of the matrix is restored by applying the new values with errors to the upper triangular matrix.

Approach iii) is simply evaluated by increasing or decreasing the value of every element by a certain percentage.

The conversion between capacitance values and physical lengths is found using equations (3.9) and (3.10). However, because the spacing was identified to be much more sensitive than the width, the width error can be assumed to be zero without changing the results meaningfully.

4

Results

In this chapter, theoretical filter function analysis of filter prototypes with a few different characteristics, and the results of the simulated filter prototype along with a proposed design for fabrication will be presented. In addition, a tolerance analysis will also be presented that examines how errors in fabrication affect filter performance.

Three different filter prototypes with different stopband attenuation levels were chosen to be examined theoretically: 30 dB, 40 dB and 60 dB. These attenuation levels were chosen to investigate how the transmission zero locations in proximity to the passband affected both the capacitance matrices as well as the realisation of the physical structure. To be able to compare with the original stepped digital elliptic filter created by Rhodes, the same return loss and fractional bandwidth levels of 20 dB and 11 % as well as filter degree of 5 were chosen [6]. The centre frequency at which the filters were designed was 5 GHz which is higher than the centre frequency of 2.15 GHz chosen in [6]. The tabulated values for the three different filter

Table 4.1: Elliptic filter function prototype values for filters with RL = 20 dB for different A_{min}. The included and more values can be found in [20].

A_{min}	c_1	c_2	c_3	c_4	c_5	Ω_2	Ω_4
30 dB	0.792	0.245	1.352	0.827	0.468	1.910	1.328
40 dB	0.867	0.137	1.512	0.407	0.669	2.438	1.617
60 dB	0.926	0.059	1.666	0.161	0.836	3.612	2.304

characteristics mentioned are presented in table 4.1.

Furthermore, to also be able to evaluate the filter for different return loss and FBW levels, the characteristics of the filter for return losses of 14 dB and 28 dB, FBWs of 2 %, 5 % and 8 % with a A_{min} of 30 dB were examined.

Table 4.2: Elliptic filter function prototype values for filters with $A_{\text{min}} = 30$ dB for different RL. The included and more values can be found in [20].

RL	c_1	c_2	c_3	c_4	c_5	Ω_2	Ω_4
14 dB	1.057	0.328	1.474	1.093	0.667	1.706	1.225
28 dB	0.586	0.177	1.257	0.603	0.324	2.267	1.521

The tabulated values for these filter prototypes are shown in table 4.2.

4.1 Analysis of the filter function

Using the tabulated values from table 4.1, the initial capacitance matrices C^A and C^B, the combined admittance matrix $\mathbf{Y}_{19\times19}^{\text{combined}}$ and related theoretical filter function of figure 4.1 were found. By comparison of the filter functions with equal charac-

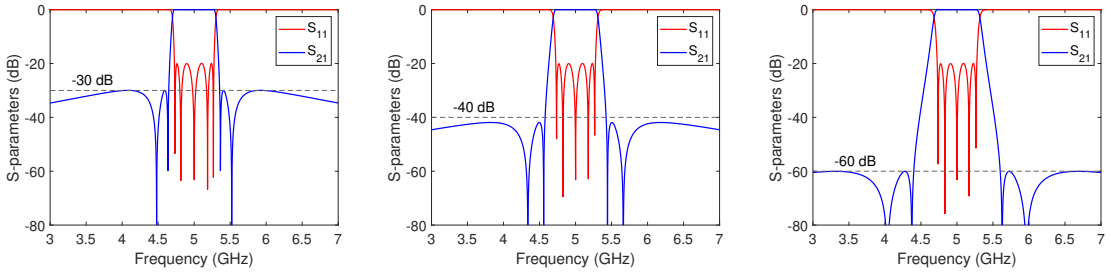


Figure 4.1: Filter functions for FBW=11%, RL=20 dB and, A_{min}= 30 dB, 40 dB and 60 dB.

teristics except for the minimum stopband attenuation in Fig. 4.1 it is clear that a lower A_{min} causes the transmission zero locations to be closer to the passband, causing a sharper rejection. This means that choosing which tabulated values to use will be a trade-off between rejection sharpness and minimum stopband attenuation depending on the filter requirements. All other general return and insertion loss characteristics remain the same between the different filter functions.

The filter functions in Fig. 4.1 are, as explained in section 2.8.2, independent of the scaling of the capacitance matrices. Thus, in order to compare the capacitance matrices between the three filter functions, they all need to be derived using the same scaling constants. The capacitance matrices for the three filter functions are derived using an arbitrary diagonal scaling constant n_{diag} of 1.2 and nodal scaling constants n_k all equal to 1, and are displayed without the UEs (since they are equal when the scaling constants are equal) in table 4.3.

The results show that a lower A_{min} causes all capacitance values to be larger. This causes problems, for example, at node 4 in both networks for $A_{\text{min}} = 30\,\text{dB}$, where the ground capacitance defined in equation (3.13) is negative for this capacitance matrix. A negative capacitance is obviously unphysical, and so the capacitance matrices belonging to $A_{\text{min}} = 30\,\text{dB}$ are practically unrealisable with this arbitrary scaling.

Next, the filter function will be evaluated depending on different FBW while keeping return loss and minimum stopband attenuation constant. From table 4.3 it is clear that the fractional bandwidth changes the matrices. The ground, coupling, and mean nodal differences as defined in section 4.2.2 are plotted in Fig. 4.3. The mean nodal difference keeps decreasing together with the FBW, showing that the lower FBW, the more nodally aligned are network 1 and 2. This is true for any A_{min}. Also, the capacitance variations within the networks are minimised around a FBW of 7%.

Table 4.3: Capacitance per unit-length matrices excluding UEs belonging to the filter functions in Fig. 4.1 with $n_{\text{diag}} = 1.2$, FBW = 11 %, RL = 20 dB, and n_k=1 for all nodes.

	Network 1	Network 2
$A_{\text{min}} = 30\,\text{dB}$	$\begin{bmatrix} 11.70 & -6.01 & 0 & 0 & 0 \\ -6.01 & 14.04 & -3.62 & 0 & 0 \\ 0 & -3.62 & 19.23 & -10.28 & 0 \\ 0 & 0 & -10.28 & 19.23 & -9.76 \\ 0 & 0 & 0 & -9.76 & 16.02 \end{bmatrix}$	$\begin{bmatrix} 9.15 & -4.32 & 0 & 0 & 0 \\ -4.32 & 14.04 & -5.03 & 0 & 0 \\ 0 & -5.03 & 18.07 & -8.17 & 0 \\ 0 & 0 & -8.17 & 19.23 & -12.28 \\ 0 & 0 & 0 & -12.28 & 17.60 \end{bmatrix}$
$A_{\text{min}} = 40\,\text{dB}$	$\begin{bmatrix} 10.46 & -4.45 & 0 & 0 & 0 \\ -4.45 & 12.55 & -2.44 & 0 & 0 \\ 0 & -2.44 & 14.08 & -6.3 & 0 \\ 0 & 0 & -6.3 & 14.08 & -5.71 \\ 0 & 0 & 0 & -5.71 & 11.74 \end{bmatrix}$	$\begin{bmatrix} 8.26 & -2.92 & 0 & 0 & 0 \\ -2.92 & 12.55 & -3.71 & 0 & 0 \\ 0 & -3.71 & 13.50 & -4.77 & 0 \\ 0 & 0 & -4.77 & 14.08 & -7.56 \\ 0 & 0 & 0 & -7.56 & 12.16 \end{bmatrix}$
$A_{\text{min}} = 60\,\text{dB}$	$\begin{bmatrix} 9.55 & -3.06 & 0 & 0 & 0 \\ -3.06 & 11.46 & -1.40 & 0 & 0 \\ 0 & -1.40 & 11.57 & -3.85 & 0 \\ 0 & 0 & -3.85 & 11.57 & -3.07 \\ 0 & 0 & 0 & -3.07 & 9.64 \end{bmatrix}$	$\begin{bmatrix} 7.67 & -1.65 & 0 & 0 & 0 \\ -1.65 & 11.46 & -2.58 & 0 & 0 \\ 0 & -2.58 & 11.28 & -2.59 & 0 \\ 0 & 0 & -2.59 & 11.57 & -4.56 \\ 0 & 0 & 0 & -4.56 & 9.32 \end{bmatrix}$
	(1) (2) (3) (4) (5)	(1) (2) (3) (4) (5)

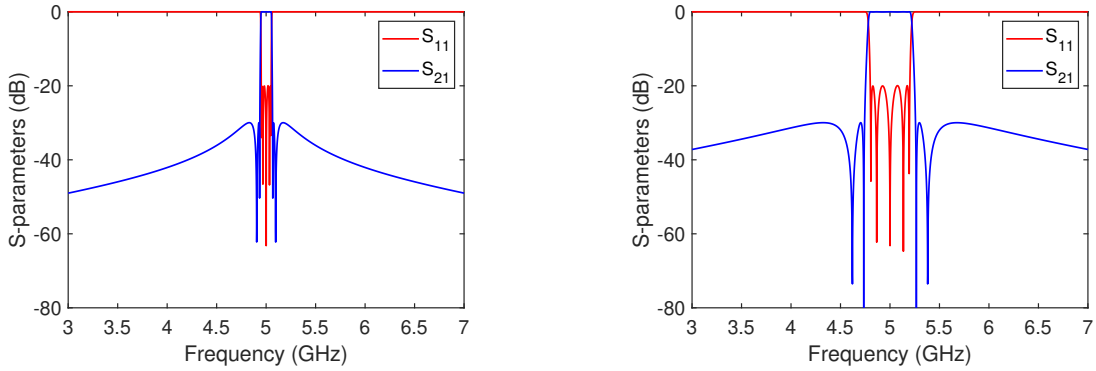


Figure 4.2: Filter functions for $A_{\text{min}} = 30\,\text{dB}$ and FBW=2% and 8%.

Table 4.4: Capacitance per unit-length matrices excluding UEs belonging to the filter functions in Fig. 4.2 with $A_{\text{min}} = 30\,\text{dB}$, $n_{\text{diag}} = 1.2$ and n_k=1 for all nodes.

	Network 1	Network 2
FBW = 2%	$\begin{bmatrix} 10.02 & -1.34 & 0 & 0 & 0 \\ -1.34 & 12.03 & -3.89 & 0 & 0 \\ 0 & -3.89 & 19.17 & -9.49 & 0 \\ 0 & 0 & -9.49 & 19.17 & -10.71 \\ 0 & 0 & 0 & -10.71 & 15.98 \end{bmatrix}$	$\begin{bmatrix} 9.56 & -4.32 & 0 & 0 & 0 \\ -4.32 & 12.03 & -4.13 & 0 & 0 \\ 0 & -4.13 & 18.96 & -9.10 & 0 \\ 0 & 0 & -9.10 & 19.17 & -11.16 \\ 0 & 0 & 0 & -11.16 & 16.26 \end{bmatrix}$
FBW = 8%	$\begin{bmatrix} 11.12 & -5.60 & 0 & 0 & 0 \\ -5.60 & 13.34 & -3.71 & 0 & 0 \\ 0 & -3.71 & 19.18 & -10.00 & 0 \\ 0 & 0 & -10.00 & 19.18 & -10.06 \\ 0 & 0 & 0 & -10.06 & 15.99 \end{bmatrix}$	$\begin{bmatrix} 9.28 & -4.41 & 0 & 0 & 0 \\ -4.41 & 13.34 & -4.71 & 0 & 0 \\ 0 & -4.71 & 18.34 & -8.46 & 0 \\ 0 & 0 & -8.46 & 19.18 & -11.89 \\ 0 & 0 & 0 & -11.89 & 17.13 \end{bmatrix}$
	(1) (2) (3) (4) (5)	(1) (2) (3) (4) (5)

The capacitance values are scaled by the reflection level as shown in table 4.5. The mean nodal differences of the capacitances for the three different return loss levels 14 dB, 20 dB, and 28 dB are presented in table 4.6, with the result that a lower reflection is connected to a more nodally aligned network, although the difference is very small.

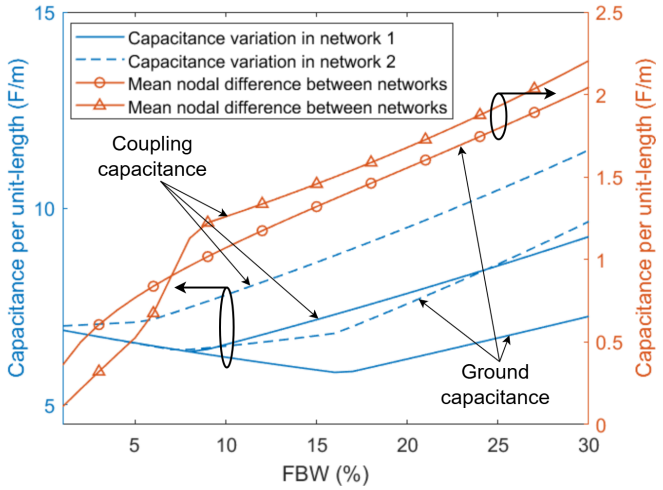


Figure 4.3: Capacitance variation over FBW for $A_{\min} = 30\,\text{dB}$.

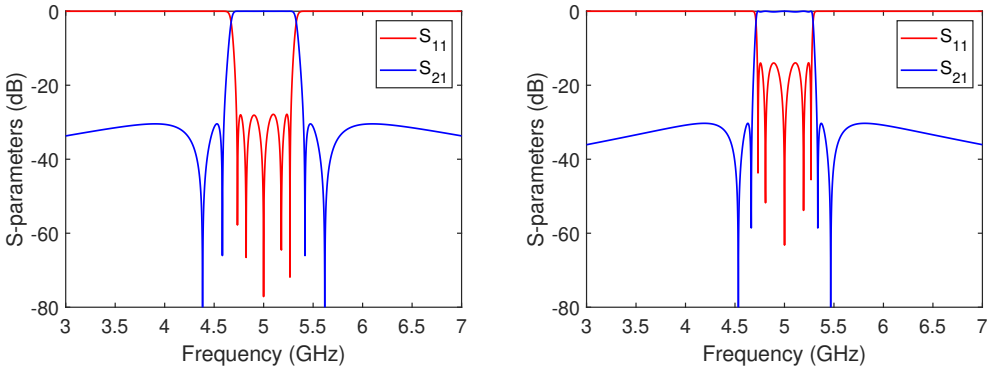


Figure 4.4: Filter functions for $A_{\min} = 30\,\text{dB}$, FBW=11% and RL=28 dB and 14 dB.

Table 4.5: Capacitance per unit-length matrices excluding UEs belonging to the filter functions in Fig. 4.4 with $A_{\min} = 30\,\text{dB}$, $n_{\text{diag}} = 1.2$ and n_k=1 for all nodes.

	Network 1	Network 2
RL = 28 dB	$\begin{bmatrix} 9.04 & -4.63 & 0 & 0 & 0 \\ -4.63 & 10.84 & -2.51 & 0 & 0 \\ 0 & -2.51 & 14.16 & -7.32 & 0 \\ 0 & 0 & -7.32 & 14.16 & -7.04 \\ 0 & 0 & 0 & -7.04 & 11.80 \end{bmatrix}$	$\begin{bmatrix} 6.66 & -3.13 & 0 & 0 & 0 \\ -3.13 & 10.84 & -3.70 & 0 & 0 \\ 0 & -3.70 & 13.29 & -5.63 & 0 \\ 0 & 0 & -5.63 & 14.16 & -9.15 \\ 0 & 0 & 0 & -9.15 & 12.91 \end{bmatrix}$
RL = 14 dB	$\begin{bmatrix} 15.07 & -7.75 & 0 & 0 & 0 \\ -7.75 & 18.08 & -5.01 & 0 & 0 \\ 0 & -5.01 & 25.50 & -13.99 & 0 \\ 0 & 0 & -13.99 & 25.50 & -13.04 \\ 0 & 0 & 0 & -13.04 & 21.25 \end{bmatrix}$	$\begin{bmatrix} 12.28 & -5.77 & 0 & 0 & 0 \\ -5.77 & 18.08 & -6.72 & 0 & 0 \\ 0 & -6.72 & 23.96 & -11.32 & 0 \\ 0 & 0 & -11.32 & 25.50 & -16.12 \\ 0 & 0 & 0 & -16.12 & 23.46 \end{bmatrix}$
	(1) (2) (3) (4) (5)	(1) (2) (3) (4) (5)

Table 4.6: Capacitance per unit-length difference between networks for different return losses RL with $A_{\min} = 30\,\text{dB}$ and FBW=11 %.

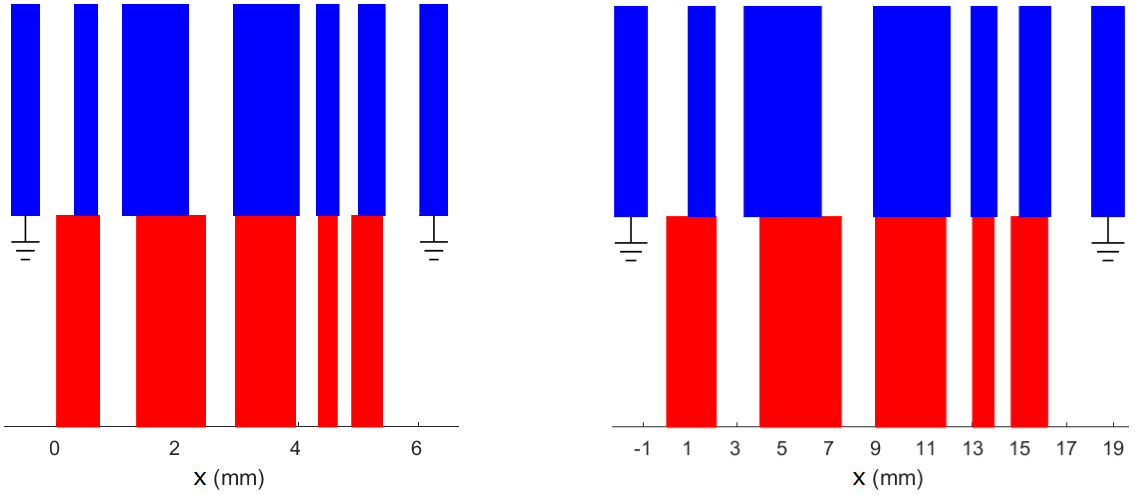
	RL = 28 dB	RL = 20 dB	RL = 14 dB
Mean nodal difference of ground capacitance	1.10	1.12	1.17
Mean nodal difference of coupling capacitance	1.20	1.30	1.37

4.2 Analysis of the physical structure

In this section, the results of the analysis of the physical structure depending on conductor thickness, ground plate spacing and scaling constants are presented.

4.2.1 Structure dependence on physical dimensions

Using the methods introduced in section 3.4, the process of relating a capacitance matrix to widths of and spacings between conductors is rather straightforward. However, in order to use equations (3.9) and (3.11) both the thickness t and the ground plate spacing b need to be known. Any value of these dimensions is theoretically possible, as long as the ground plate spacing is larger than the conductor thickness, but not all values create desired physical structures. Optimal conductor thickness and ground plate spacing should allow for both a compact structure and well-aligned nodes, whilst still not being impossible to fabricate.

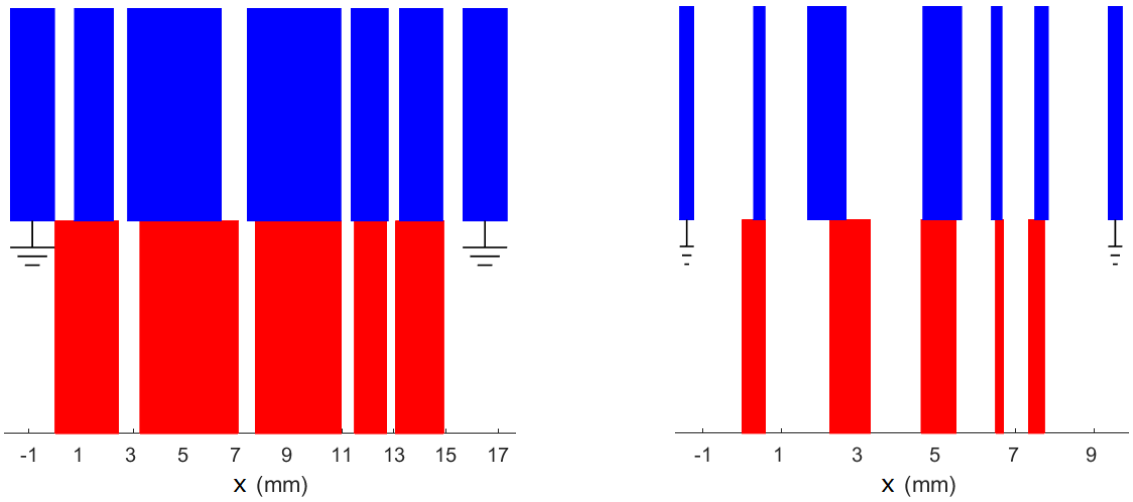


(a) b=3 mm, t=1.8 mm (b) b=9 mm, t=5.4 mm

Figure 4.5: Nodal alignments for thickness-to-ground plate ratios t/b of 0.6. Blue conductors represents network 1 and red conductors network 2.

The physical dimensions and nodal alignments of the capacitance matrix related to the tabulated values for $A_{\min} = 40\,\text{dB}$ in table 4.1 were visualised for different thickness-to-ground plate spacing ratios in Fig. 4.5. It was identified that if the thickness-to-ground plate spacing ratio, t/b, is kept constant, the nodal alignment of the conductors will also remain the same. The blue conductors represent network 1 and the red conductors represent network 2.

From Fig. 4.6 the general conclusion is that a higher thickness-to-ground ratio means a more compact structure, but a worse nodal alignment due to the conductors becoming very thin. The structure relating to the t/b-ratio of 0.8 is probably very difficult, if not impossible, to fabricate due to the misalignments of the nodes. On the other hand, the structure relating to the 0.4-ratio is probably also very difficult to fabricate, not due to misalignment, but instead due to the small spacings between the conductors.

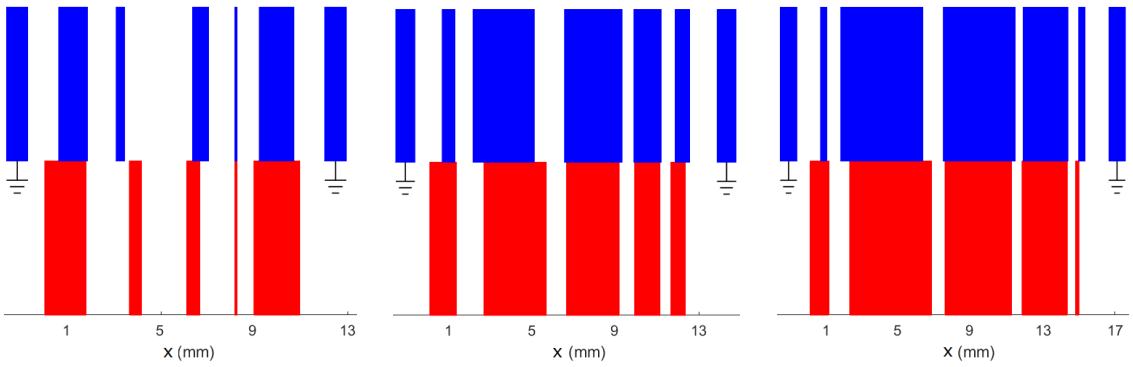


(a) b=6 mm, $t/b = 0.4$ (b) b=6 mm, $t/b = 0.8$

Figure 4.6: Nodal alignments for different thickness-to-ground plate ratios t/b.

4.2.2 Structure dependence on scaling constants

In the capacitance matrices of [6], a diagonal scaling constant n_{diag} of 1.2 was used because it was identified as an optimum relationship between the coupling and the ground capacitances of both networks, supposedly resulting in a minimum variation between the capacitances. Changing this scaling constant also changes the physical structure, with a smaller diagonal constant causing thinner widths of the inner elements because of a smaller capacitance and a larger diagonal constant having the opposite effect; see Fig. 4.7 and the related capacitance matrices in table 4.7.



(a) $n_{\text{diag}} = 0.7$ (b) $n_{\text{diag}} = 1.4$ (c) $n_{\text{diag}} = 1.8$

Figure 4.7: Nodal alignments for different n_{diag} values with b=6 mm and t=3.6 mm.

To evaluate which diagonal scaling constant is generally preferred, ground capacitances and coupling capacitances were plotted for different diagonal scaling constants in Figs. 4.8 and 4.9, respectively. It does not seem correct, or another metric for calculating capacitance variations was used, that 1.2 would be found to be the optimal value for n_{diag}.

Instead, knowing that the nodal scaling constants n_k can scale individual nodes and thus counteract capacitance variations, a minimum nodal difference between the two networks is the desired result of optimal diagonal scaling constants. The value of

Table 4.7: Capacitance per unit-length matrices excluding UEs belonging to the physical structures in Fig. 4.1 with n_k=1 for all nodes.

	Network 1					Network 2				
$n_{\text{diag}} = 0.7$	10.46	−3.40	0	0	0	8.26	−4.32	0	0	0
	−3.40	7.32	−1.42	0	0	−4.32	7.32	−2.16	0	0
	0	−1.42	8.22	−3.68	0	0	−2.16	7.87	−2.78	0
	0	0	−3.68	8.22	−4.36	0	0	−2.78	8.22	−5.77
	0	0	0	−4.36	11.74	0	0	0	−5.77	12.16
$n_{\text{diag}} = 1.4$	10.46	−4.80	0	0	0	8.26	−3.16	0	0	0
	−4.80	14.64	−2.85	0	0	−3.16	14.64	−4.33	0	0
	0	−2.85	16.43	−7.35	0	0	−4.33	15.75	−5.56	0
	0	0	−7.35	16.43	−6.17	0	0	−5.56	16.43	−8.16
	0	0	0	−6.17	11.74	0	0	0	−8.16	12.16
$n_{\text{diag}} = 1.8$	10.46	−5.45	0	0	0	8.26	−3.58	0	0	0
	−5.45	18.82	−3.66	0	0	−3.58	18.82	−5.57	0	0
	0	−3.66	21.12	−9.45	0	0	−5.57	20.25	−7.15	0
	0	0	−9.45	21.12	−7.00	0	0	−7.15	21.12	−9.25
	0	0	0	−7.00	11.74	0	0	0	−9.25	12.16
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)

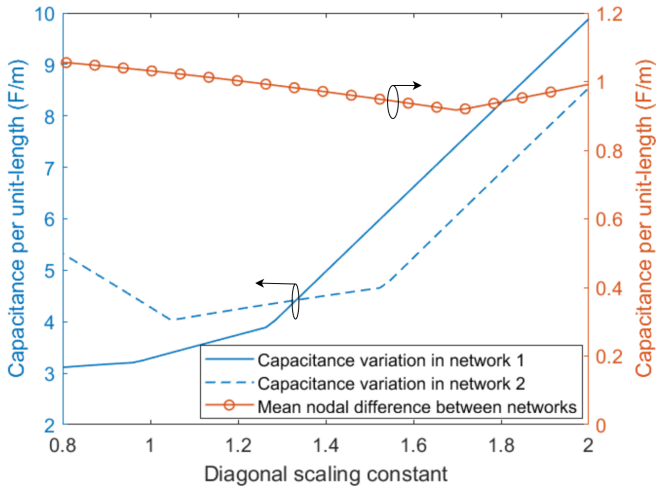


Figure 4.8: Variation of ground capacitances C_g over n_{diag} with $A_{\text{min}} = 30\,\text{dB}$.

n_{diag} that corresponds to this is 1.71. Since the nodal scaling constants scale both networks, there will be no effect on the mean capacitance difference between the networks.

However, the capacitance variation within the two networks will change. Choosing the optimal diagonal constant n_{diag} and the nodal constants n_k will therefore be a trade-off between the compactness of the structure, the variations in capacitance within each network, and the nodal alignment between the networks.

The value of n_{diag} for optimal nodal alignment is approximately true for all A_{min} and all realisable FBWs, but changes slightly with different return losses. For example, the capacitance matrix with RL = 28 dB in table 4.5 had an optimal n_{diag} value of 1.84.

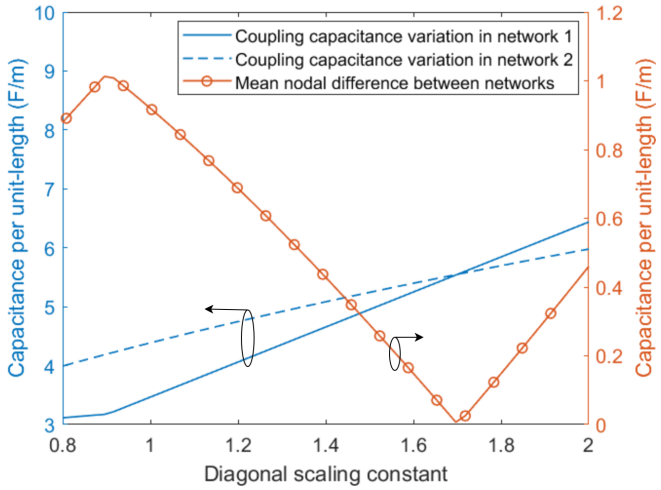


Figure 4.9: Variation of coupling capacitances $C_{k,k+1}$ over n_{diag} with $A_{\text{min}} = 30\,\text{dB}$.

4.2.3 Nodal alignment optimisation procedure evaluation

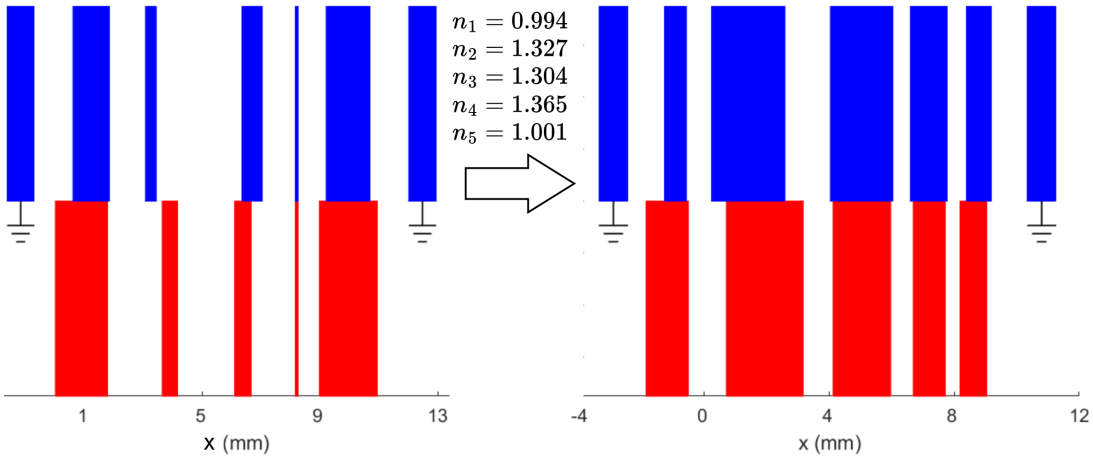


Figure 4.10: Scaling constants found by optimisation procedure of the physical structure in Fig. 4.7a as well as the original (left) and optimised (right) structure. For $A_{\text{min}} = 30\,\text{dB}$, RL = 20 % and FBW=11 %

In Fig. 4.10 it is shown how the unrealisable structure in Fig. 4.7a is made realisable by applying the nodal scaling constants n_k found through optimisation.

Table 4.8: Optimised capacitance per unit-length matrices excluding UEs belonging to the physical structure to the right in Fig. 4.10.

Network 1

$$\begin{bmatrix} 10.33 & -4.48 & 0 & 0 & 0 \\ -4.48 & 12.89 & -2.46 & 0 & 0 \\ 0 & -2.46 & 13.97 & -6.54 & 0 \\ 0 & 0 & -6.54 & 15.31 & -5.97 \\ 0 & 0 & 0 & -5.97 & 11.77 \end{bmatrix}$$

(1) (2) (3) (4) (5)

Network 2

$$\begin{bmatrix} 8.16 & -2.95 & 0 & 0 & 0 \\ -2.95 & 12.89 & -3.74 & 0 & 0 \\ 0 & -3.74 & 13.39 & -4.95 & 0 \\ 0 & 0 & -4.95 & 15.31 & -7.89 \\ 0 & 0 & 0 & -7.89 & 12.20 \end{bmatrix}$$

(1) (2) (3) (4) (5)

The optimised capacitance matrices are presented in table 4.8.

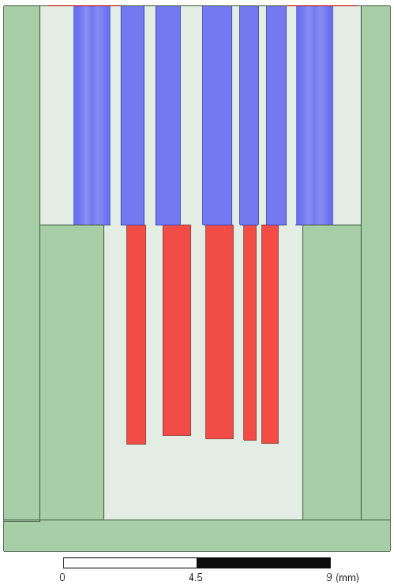
4.3 Simulated prototype filter results

A filter prototype was designed for a filter with a minimum stopband attenuation of 40 dB, a return loss of 20 dB and a fractional bandwidth of 11 %. The prototype values can be found in table 4.1. The initial capacitance matrices of network 1 and 2 are presented in (4.1a) and (4.1b) respectively.

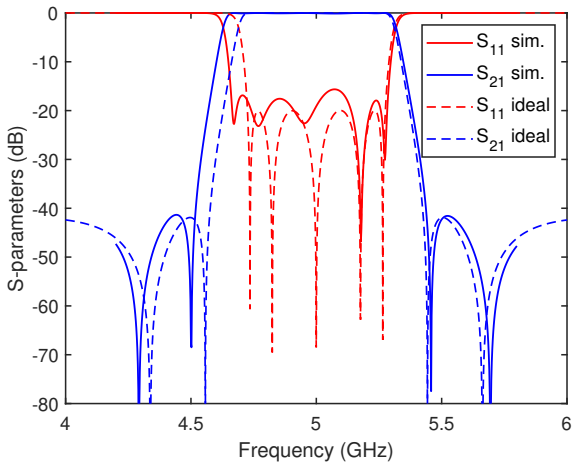
$$\begin{bmatrix} 7.54 & -3.14 & 0 & 0 & 0 & 0 & 0 \\ -3.14 & 10.46 & -4.45 & 0 & 0 & 0 & 0 \\ 0 & -4.45 & 12.55 & -2.44 & 0 & 0 & 0 \\ 0 & 0 & -2.44 & 14.08 & -6.3 & 0 & 0 \\ 0 & 0 & 0 & -6.3 & 14.08 & -5.71 & 0 \\ 0 & 0 & 0 & 0 & -5.71 & 11.74 & -3.14 \\ 0 & 0 & 0 & 0 & 0 & -3.14 & 7.54 \end{bmatrix} \tag{4.1a}$$

$$\begin{bmatrix} 8.26 & -2.92 & 0 & 0 & 0 \\ -2.92 & 12.55 & -3.71 & 0 & 0 \\ 0 & -3.71 & 13.50 & -4.77 & 0 \\ 0 & 0 & -4.77 & 14.08 & -7.56 \\ 0 & 0 & 0 & -7.56 & 12.16 \end{bmatrix} \tag{4.1b}$$

Through nodal alignment optimisation of the capacitance matrices, followed by MAXWELL2D corrections and then S-parameter optimisation of the related physical structure in HFFS, the filter structure visible in Fig. 4.11a was designed. A comparison between the ideal and the simulated filter characteristics is presented in Fig. 4.11b. The structure was designed with circular coaxial pins as unit elements, and all metal surfaces were assumed to be perfect electrical conductors (PEC).



(a) Filter prototype structure.



(b) Simulated and ideal S-parameters of the filter.

Figure 4.11: Physical dimensions and S-parameters of simulated filter with $A_{\min} = 40\,\text{dB}$, RL = 20 dB and FBW = 11 %. Values are presented in table 4.9.

The physical dimensions of the filter prototype are presented in table 4.9, where node 0 is the leftmost element in Fig. 4.11a.

Table 4.9: Physical parameters relating to capacitance matrices (4.1) in mm.

Node	Width	Spacing	Length
0	⌀1.3		7.365
		0.596	
1	0.762		7.365
		0.397	
2	0.831		7.365
		0.755	
3	0.976		7.365
		0.271	
4	0.628		7.365
		0.263	
5	0.676		7.365
		0.562	
6	⌀1.3		7.365

(a) Network 1

Node	Width	Spacing	Length
1	0.651		7.388
		0.585	
2	0.928		7.104
		0.501	
3	0.938		7.202
		0.352	
4	0.416		7.260
		0.192	
5	0.547		7.377

(b) Network 2

	x_{off}	$d_{\text{side,left}}$	$d_{\text{side,right}}$	b	t	l_g
Network 1	-0.052	1.285	1.021	3	1.8	
Network 2	0.123	0.977	1.030	3	1.8	2.576

The propagation of the wave in the structure is shown in Fig. B.4. More figures of wave propagation at different phases are found in the Appendix.

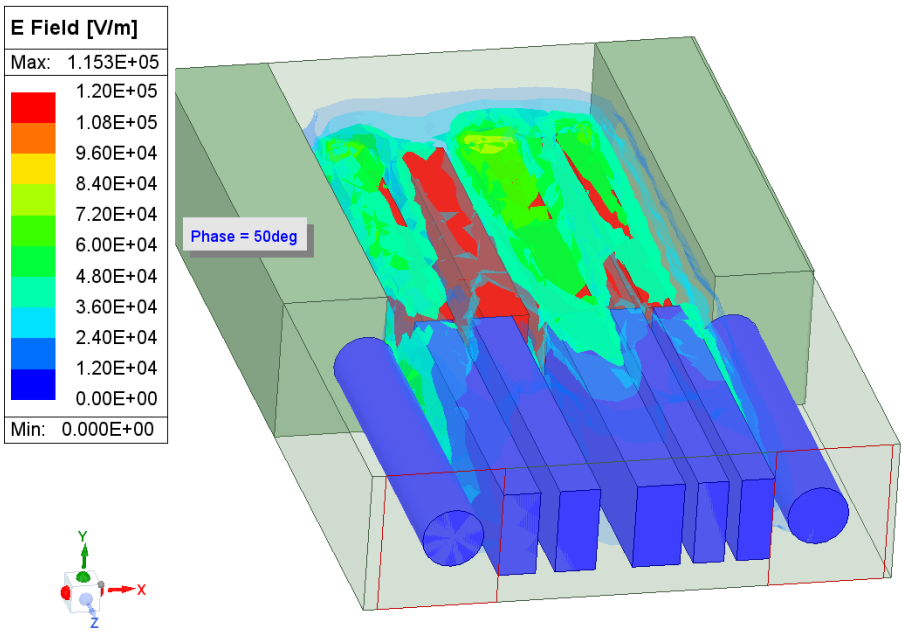


Figure 4.12: Wave propagation in filter at 5 GHz.

The group delay response of the filter is also presented in Fig. 4.13.

The filter was also tested for a worst-case scenario with all metal surfaces chosen as aluminium. A comparison of the filter characteristics using either aluminium or PEC surfaces is presented in Fig. 4.14.

It is evident from Fig. 4.14 that the return loss is fairly unaffected between the two different cases, while the insertion loss differs considerably within the passband.

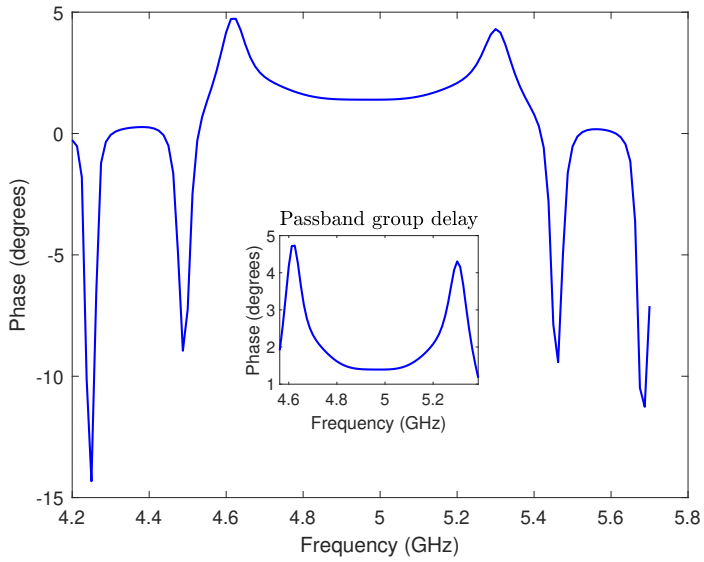


Figure 4.13: Group delay response of filter.

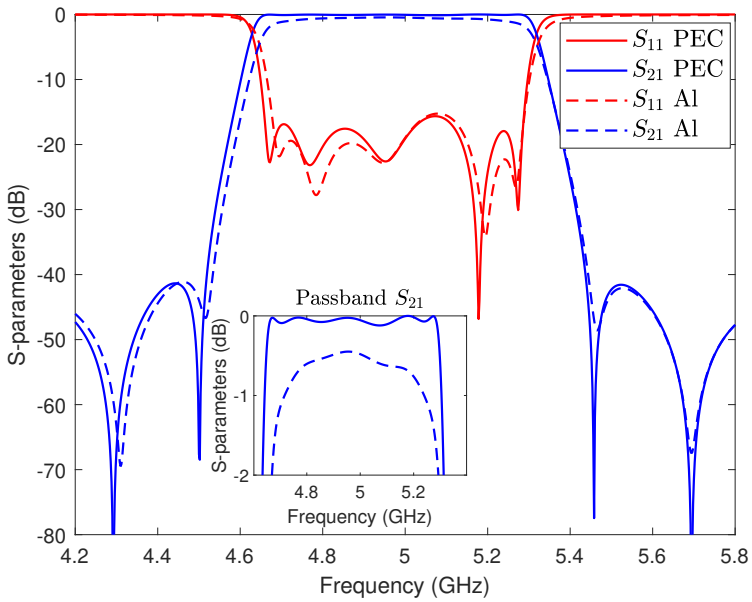
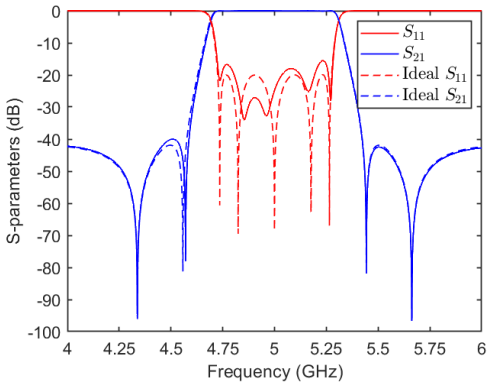


Figure 4.14: Filter characteristics using either PEC or aluminium as metal surfaces.

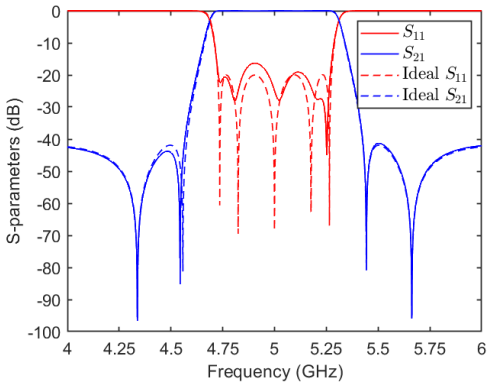
This effect should be even more pronounced as the frequency increases as a result of decreased skin depth.

4.3.1 Tolerance analysis

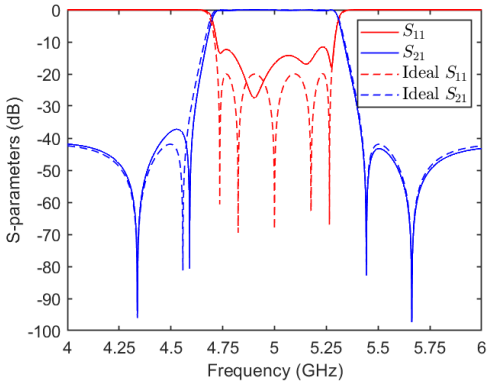
The results of the tolerance analysis for the three different approaches, i), ii) and iii) discussed in section 3.8 show that the filter characteristics, particularly the return loss, are sensitive to small changes for this filter structure. All graphs presented in the tolerance analysis are for the capacitance matrices in 4.1.



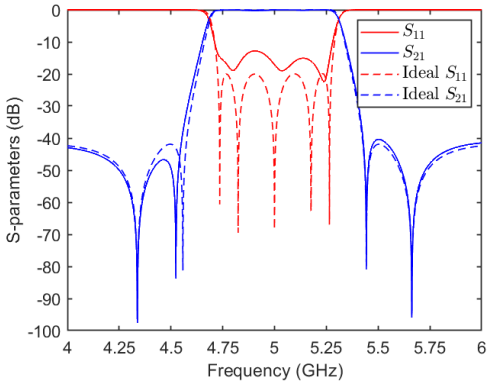
(a) $\Delta d = 2\,\mu\text{m}$



(b) $\Delta d = -2\,\mu\text{m}$

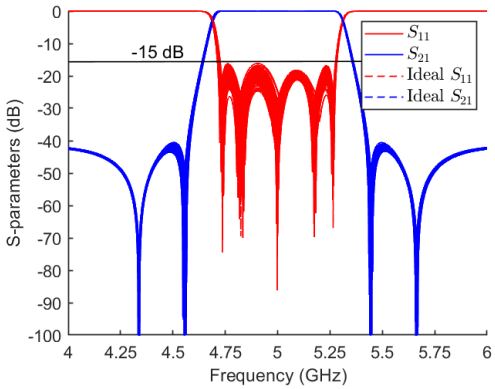


(c) $\Delta d = 5\,\mu\text{m}$

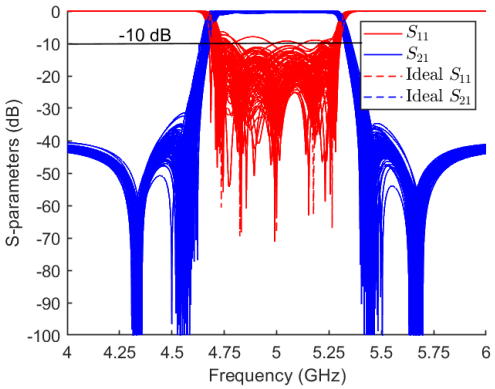


(d) $\Delta d = -5\,\mu\text{m}$

Figure 4.15: Approach i): Filter function as a result of change of most sensitive element for the capacitance matrices in table 4.1.



(a) $\Delta x = 1\,\mu\text{m}$



(b) $\Delta x = 5\,\mu\text{m}$

Figure 4.16: Approach ii): Filter function as a result of a normally distributed change in all elements of both networks for the capacitance matrices in table 4.1.

Approach i), see Fig. 4.15, shows that the most sensitive element has a significant impact on the return loss of the filter. However, if the manufacturing precision is

less than 2 µm, the degradation is fairly small.

Approach ii), see Fig. 4.16, shows that in the case when $\Delta x = \pm 5\,\mu\text{m}$ the filter structure can still produce satisfactory results although with more degradation.

Approach iii), an approximate relationship between S_{11} and the percentage of error of the capacitance matrix can be found by averaging the results for many fixed frequency points within the passband for gradually increasing error percentages; see Fig. 4.17.

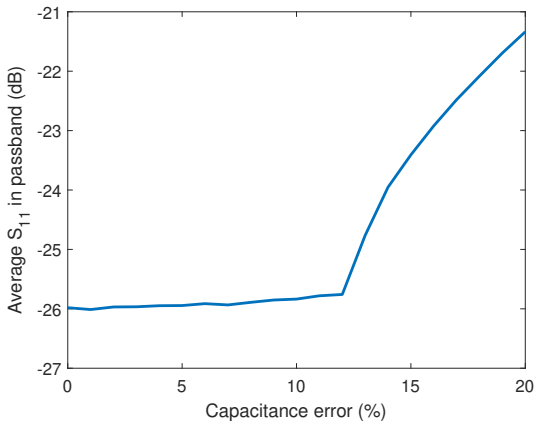
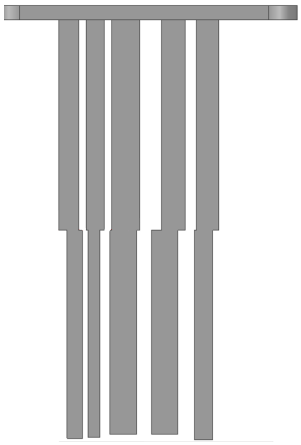


Figure 4.17: Approach iii): Average S_{11} in passband over percentage error in capacitance matrix.

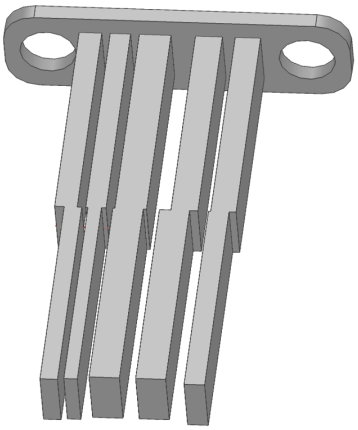
4.4 Proposed prototype design for fabrication

Drawings of a proposed filter prototype suitable for fabrication excited by an end-launch SMA pin (see section 3.6.2.1) are presented. The proposed filter prototype consists of two parts, the metal insert part (see Fig. 4.18) which is the realisation of networks 1 and 2, together with the grounding back wall of network 1, and the housing (see Figs. 4.19 and 4.20) which was designed in two different versions.

Figures 4.19 and 4.20 also show how the insert is supposed to be placed within the housing to create the final structure. The two holes in the insert are for the SMA pin, which needs to be connected to the metal housing to ensure proper grounding.

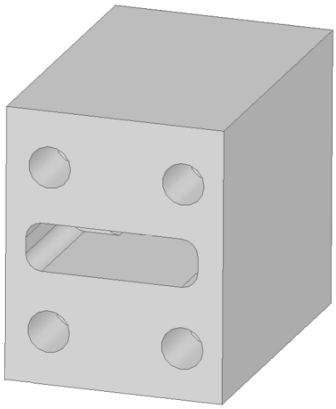


(a) Network 1 and 2 insert in 2D.

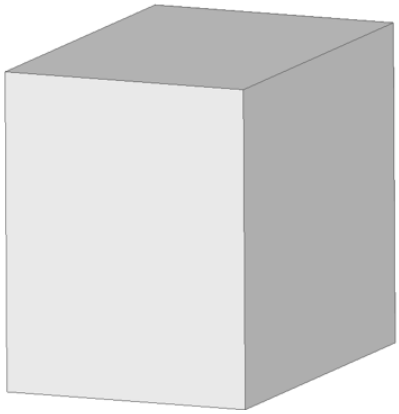


(b) Network 1 and 2 insert in 3D.

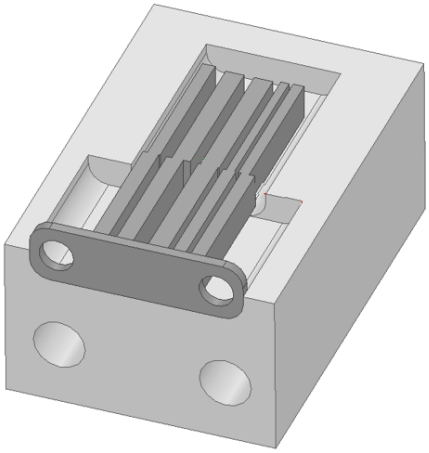
Figure 4.18: Filter structure metal insert part design.



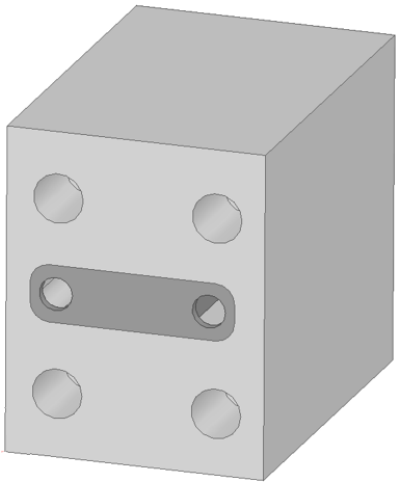
(a) Front, holes for flange screws and insert.



(b) Back.

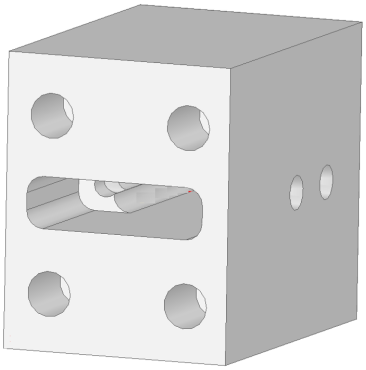


(c) With insert, cut horizontally.

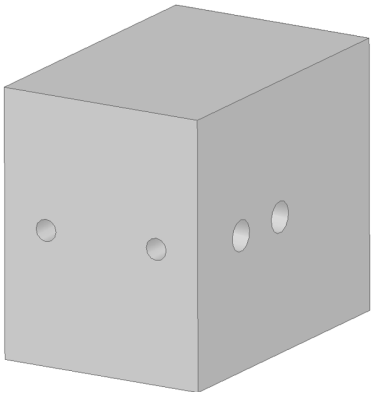


(d) Final structure.

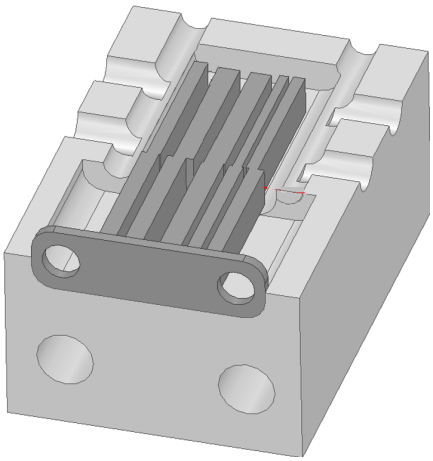
Figure 4.19: Proposed design, option 1.



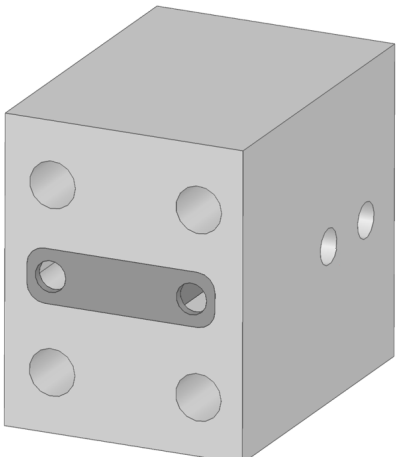
(a) Front, holes for flange screws and insert.



(b) Back, holes for soldering.



(c) With insert, cut horizontally.



(d) Final structure.

Figure 4.20: Proposed design, option 2.

5

Discussion and future work

5.1 Optimal diagonal scaling constant

In section 4.2.2 it was argued that a mean nodal difference between the two networks was the desired result when choosing the diagonal scaling constant n_{diag}. In order to achieve the most aligned individual nodes this is generally the case, but for other purposes it might not always be optimal. It is also highly debatable of what the best number for the diagonal scaling constant actually is: as Rhodes suggested a value of 1.2 and this work suggests a value of 1.71 due to the possibility of scaling each node individually, which was not done in Rhodes' paper [6].

The best diagonal scaling constant will probably differ slightly with each design, but with the use of Figs. 4.8 and 4.9 the general behaviour, alignment and homogeneity of the filter structure related to the scaling constant are visualised. This should simplify the filter design process.

5.2 Width and spacing approximations

Evidently, with the dimensions of the final filter in mind, the spacing and width approximations of equations (3.10) and (3.12) are not accurate enough for filters with dimensions of the conductors below roughly 1 mm. The optimisation scheme developed in MATLAB could therefore be significantly more accurate with the use of another way to find the spacings and widths. Other similar ways to compute the width approximation are through the impedance calculations of a thick strip as published by Wheeler [44], and for the spacing approximation there are expressions for the odd- and even-mode impedances for coupled striplines published by Cohn [45]. At least the spacing approximation could benefit from a more accurate expression.

There are, however, also entirely different methods in which this could be done. One way is through eigenmode analysis [46], using integration of electric and magnetic fields which should be able to give more accurate results even when the conductors are of finite thickness; another way to do this would be to implement a method of moments (MoM) approach as described in [47]. The downside of a MoM approach is that it is generally time-consuming, but with good enough accuracy and theoretical modelling of parasitic capacitances, the simulations in HFFS could potentially be used only to verify the final design.

5.3 Varying ground plate distance

By varying the distance b' between the conductors and the ground plate, the ground capacitance C_g can be individually adjusted, adding another opportunity for individual tuning of the resonators. By increasing the distance for a specific conductor, the width of that conductor alone will increase, while the spacing between conductors remains the same (assuming that there is no coupling between the sides of the conductors and the ground plate). This process is visualised in Fig. 5.1 for the case when one conductor is disproportionately wide, and thus the ground plate distance is made smaller to make the conductors equally wide.

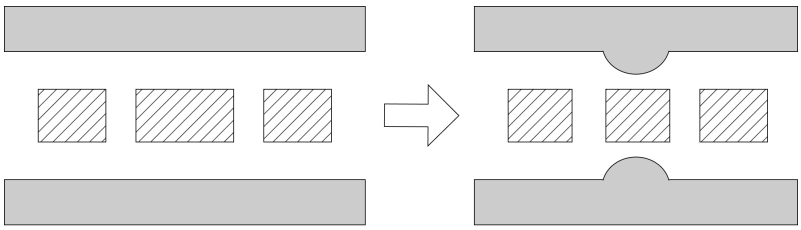


Figure 5.1: Example of varying ground plate distances to change conductor width.

This approach proves to work quite well in theory since it adds another dimensions to the scaling of the matrices, where the width decrease or increase is predicted by equation (3.9). However, practically, there are three main concerns which makes this hard to realise:

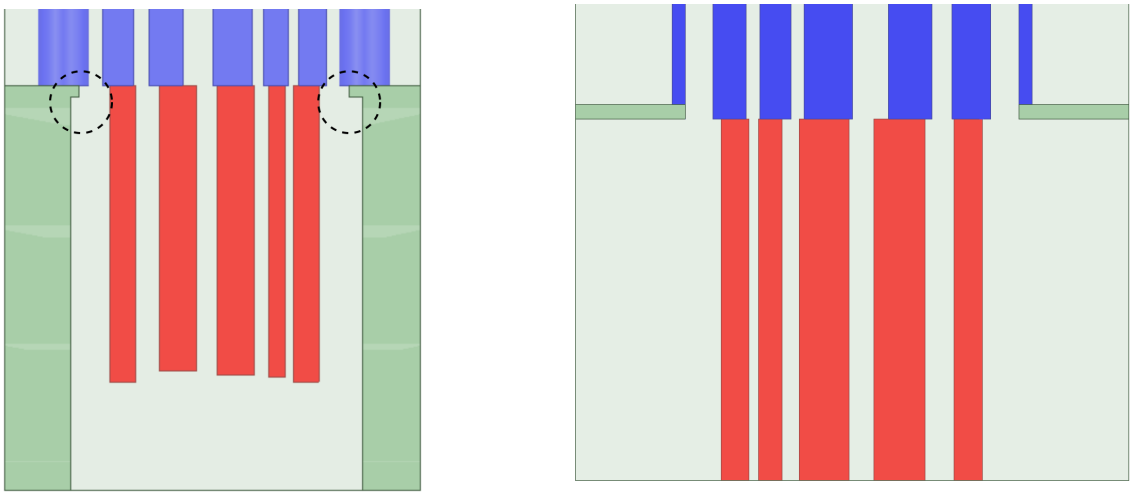
i) the ground plate change does not only affect one conductor,
ii) the spacing is more sensitive than the width,
iii) and the fabrication becomes more difficult.

It was evident from simulations that i) could severely degrade the filter characteristics and was deemed too hard to realise. However, with for example a MoM approach as discussed in the previous section, the effect of changing the ground plate distance can be better predicted. With a more precise approximation of change, and if it is not too hard to fabricate, this approach could be useful for certain scalings or sizes of the structure which causes uneven conductor dimensions.

5.4 Decreasing the impact of side walls

One problem with the current shorting of the UEs by connection to the end walls is that the conductors corresponding to nodes 1 and 5 of network 2 are significantly affected by their distance from these end walls. This was identified to be one of the main issues that could negatively impact the overall design and is also the reason why some problematic areas as described in Fig. 5.3 appeared during the design process. Two suggested ways to fix this are proposed:

i) Removing a portion of the end walls (Fig. 5.2a),
ii) Different design of the UE shorting (Fig. 5.2b).



(a) Removal of portion of end walls.

(b) Different UE short design.

Figure 5.2: Possible design changes to decrease the impact of side walls.

The design changes simplify and give more freedom in the design, but might make the filter harder to manufacture through methods such as milling. However, with microfabrication these design changes should not be any more difficult to manufacture.

5.5 Overlapping conductors

Even though the nodes of the filter might align relatively well, certain problematic areas can arise, which might make the filter difficult to fabricate. One way to overcome this issue is to let the lengths of individual conductors extend to the other network, making the conductors overlap slightly.

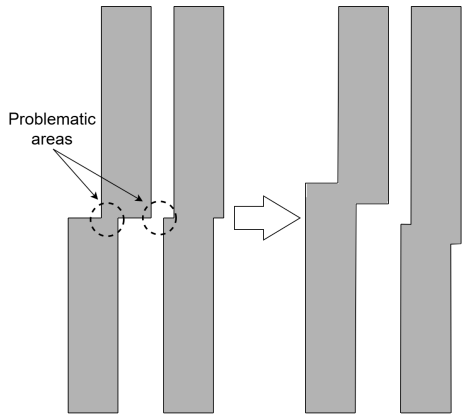


Figure 5.3: Problematic areas fixed by overlapping of conductors.

In Fig. 5.3, two problematic areas for a hypothetical node alignment are identified and corrected through the overlapping procedure. If the capacitance variation between networks is fairly low, this procedure should only slightly alter the capacitance

matrices.

5.6 HFFS simulation and optimisation improvements

If the S-parameters of each of the non-common ports, corresponding to the ends of the lines of network 1 and 2, in Fig. 3.2 could be extracted, the capacitance matrix to S-parameter method described in section 2.8.1 could be reversed, and thus the capacitance matrix could also be found directly for the 3D structure. The advantages of being able to inspect the capacitance matrix are several. The return and insertion loss characteristics of the filter can be significantly degraded by a single node deviating from the desired capacitance value, even if the other values match the theoretical matrices. If the capacitance matrix of the 3D structure were known, this issue could be easily corrected. Another advantage is that the optimisation procedure could be performed in MATLAB, with the goal of matching the theoretical capacitance matrix as closely as possible, rather than simply achieving sufficient filter characteristics. This approach should be faster and require fewer iterations, as only the incorrect nodes would be adjusted.

This capacitance matrix extraction was not included in this work, thus making it impossible to compare the theoretical capacitance matrix with the capacitance matrix of the simulated 3D-structure. Implementing this is therefore strongly recommended in any future work.

5.7 MEMS technology fabrication

By increasing the frequency, the design could potentially be suitable for MEMS fabrication as the size decreases with frequency. Also, with the increased precision that comes with, for example, Polystrata, scaling constants that create very compact filter realisations could potentially be utilised, as they would now be fabricatable.

5.8 Asymmetric filter realisation

By introducing hypothetical FIR elements during the filter synthesis process, it has been found that an asymmetric response filter with a similar physical structure can be achieved. However, the asymmetric structure has not been further analysed. Using the methods developed in this thesis, the asymmetrical realisation could be analysed and compared to the symmetrical structure. Since the asymmetrical structure allows for more flexible pass- and stopbands, the effect of the pole placement could be investigated in relation to the physical structure.

6

Conclusion

This work has revolved around the stepped digital elliptic filter and its possible realisations. A comprehensive design process from theory to fabrication and analysis of the filter has been presented for a few different bandpass filter characteristics.

A physical realisation of a 5th degree stepped digital elliptical filter's dependence on scaling of its related capacitance matrices was investigated with the focus on finding the optimal scaling relating to an optimal physical realisation. An optimal scaling was not clearly identified; however, graphs and figures on the dependence on scaling of the capacitance matrix and its related physical realisation are presented. These findings should help future filter designers when choosing the best scaling for their purposes.

It was also found that the optimisation based on the nodal alignment approach introduced in section 3.5.1 and applied in MATLAB as described in section 3.7.1.3 could successfully find well-aligned physical realisations of a 5th degree stepped digital elliptic filter in many cases. The same method can be extended to any odd-order stepped digital filter above 3. However, further work is needed to improve or replace the approximate relations between capacitance and physical dimensions with more precise methods.

The filter characteristics of the proposed design have been shown to align well with the theoretical filter characteristics in simulations, although with slightly worse performance. The slight difference in performance is probably caused by errors in the capacitance matrix of the simulated realisation. Relatively small errors in the conductor spacing (around 5 µm) have been shown to degrade the return loss of the filter, whereas the insertion loss is relatively unaffected. Furthermore, a worst-case comparison between perfect electrical conductors and aluminium as conductor materials shows that while the return loss remains essentially unchanged, the insertion loss in the passband, particularly at the band edges, is affected, with a difference of roughly 2 dB at the band edges between the aluminium and PEC cases.

The current design proposed is not suitable for MEMS fabrication. The reasons are that the frequency is too low (5 GHz) and that the overall design is too large, also mainly due to frequency. By increasing the frequency to the n258 band (around 26 GHz) or perhaps even higher, this filter realisation could be suitable for MEMS technology, as the size of the length dimension, which is always equal to $\lambda/4$, decreases from 15 mm to 2.9 mm.

Bibliography

[1] C. Herter, "The electromagnetic spectrum: A critical natural resource," *Nat. Res. 651*, vol. 25, no. 3, 1985.

[2] T. G. Williamson, T. R. Harris, G. Che, and N. Ghalichechian, "Novel Compact Tunable Filter Prototype Realized with Vanadium Dioxide," in *2024 IEEE International Microwave Filter Workshop (IMFW)*, 2024, pp. 15–17, DOI: 10.1109/IMFW59690.2024.10477139.

[3] C. C. W. Ruppel, "Acoustic wave filter technologya review," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, no. 9, pp. 1390–1400, 2017, DOI: 10.1109/TUFFC.2017.2690905.

[4] H. Nakahata, A. Hachigo, K. Itakura, and S. Shikata, "Fabrication of high frequency saw filters from 5 to 10 ghz using sio/sub 2//zno/diamond structure," in *2000 IEEE Ultrasonics Symposium. Proceedings. An International Symposium (Cat. No.00CH37121)*, vol. 1, 2000, 349–352 vol.1, DOI: 10.1109/ULTSYM.2000.922570.

[5] H. Sobol and K. Tomiyasu, "Milestones of microwaves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 594–611, 2002, DOI: 10.1109/22.989945.

[6] J. D. Rhodes, "The Stepped Digital Elliptic Filter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 17, no. 4, pp. 178–184, 1969, DOI: 10.1109/TMTT.1969.112692.

[7] D. A. Frickey, "Conversions between S, Z, Y, H, ABCD, and T parameters which are valid for complex source and load impedances," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 2, pp. 205–211, Feb. 1994, DOI: 10.1109/22.275248.

[8] J.-S. Hong, *Microstrip Filters for RF/Microwave Applications*. Wiley, 2011, ch. 3, p. 30, DOI: 10.1002/9780470937297.ch3.

[9] P. Bhartia and P. Pramanick, *Modern RF and Microwave Filter Design*. Artech House, 2016.

[10] R. J. Cameron, C. M. Kudsia, and R. R. Mansour, *Microwave Filters for Communication Systems*. New Jersey: Wiley, 2007, ch. 3, p. 98.

[11] R. M. Fano, "Theoretical Limitations on the Broad-Band Matching of Arbitrary Impedances," *Journal of the Franklin Institute*, vol. 249, pp. 139–154, 1950.

[12] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. New York: Van Nostrand, 1945.

[13] D. M. Pozar, *Microwave Engineering*. Wiley, 2011, vol. 4th edition, ch. 5, p. 267.

[14] G. C. Temes and S. Mitra, *Modern Filter Theory and Design*. New York: Wiley, 1973, ch. 2, pp. 52–53.

[15] S. J. Orfanidis. "Lecture notes on elliptic filter design [PDF document]." (2006), [Online]. Available: https://eceweb1.rutgers.edu/~orfanidi/ece521/.

[16] P. Jarry and J. Beneat, *Advanced Design Techniques and Realizations of Microwave and RF Filters*. IEEE, 2008, ch. 2, p. 27, DOI: 10.1002/9780470294178.

[17] R. J. Cameron, C. M. Kudsia, and R. R. Mansour, *Microwave Filters for Communication Systems*. New Jersey: Wiley, 2007, ch. 3, p. 127.

[18] J.-S. Hong, *Microstrip Filters for RF/Microwave Applications*. Wiley, 2011, ch. 3, p. 44.

[19] R. Saal and E. Ulbrich, "On the design of filters by synthesis," *IRE Transactions on Circuit Theory*, vol. 5, no. 4, pp. 284–327, 1958, DOI: 10.1109/TCT.1958.1086481.

[20] A. I. Zverev, *Handbook of filter synthesis*. Wiley, 1967, ch. 1, p. 29, ISBN: 978-0-471-74942-4.

[21] R. J. Cameron, C. M. Kudsia, and R. R. Mansour, *Microwave Filters for Communication Systems*. New Jersey: Wiley, 2007, ch. 3, pp. 136–137.

[22] R. Baum, "Design of unsymmetrical band-pass filters," *IRE Transactions on Circuit Theory*, vol. 4, no. 2, pp. 33–40, 1957, DOI: 10.1109/TCT.1957.1086352.

[23] P. Lenoir, S. Bila, F. Seyfert, D. Baillargeat, and S. Verdeyme, "Synthesis and design of asymmetrical dual-band bandpass filters based on equivalent network simplification," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 7, pp. 3090–3097, 2006, DOI: 10.1109/TMTT.2006.877037.

[24] S. D'Oro and G. Macchiarella, "Design of asymmetric filters with requirements in two bands of finite extension," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 6, pp. 1045–1049, 2001, DOI: 10.1109/22.925489.

[25] W. J. Getsinger, "Coupled rectangular bars between parallel plates," *IRE Transactions on Microwave Theory and Techniques*, vol. 10, no. 1, pp. 65–72, Jan. 1962, DOI: 10.1109/TMTT.1962.1125447.

[26] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave filters, impedance-matching networks, and coupling structures*. Norwood, MA: Artech House, 1980, pp. 195–196.

[27] R. J. Wenzel, "Exact theory of interdigital band-pass filters and related coupled band-pass structures," *IEEE Transactions on Microwave Theory and Techniques*, vol. 13, no. 5, pp. 559–575, Sep. 1965, DOI: 10.1109/TMTT.1965.1126051.

[28] R. F. Harrington, "Losses on multiconductor transmission lines in multilayered dielectric media," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 7, pp. 705–710, Jul. 1984, DOI: 10.1109/TMTT.1984.1132757.

[29] A. Tripathi and V. K. Tripathi, "A configuration-oriented spice model for multiconductor transmission lines in an inhomogeneous medium," *IEEE Transac-*

tions on Microwave Theory and Techniques, vol. 46, no. 12, pp. 1997–2005, Dec. 1998, DOI: 10.1109/22.739272.

[30] M. C. Horton and W. R J, "The digital elliptic filter - a compact sharp-cutoff design for wide bandstop or bandpass requirements," *IEEE Transactions on Microwave Theory and Techniques*, vol. 15, no. 5, pp. 307–314, May 1967, DOI: 10.1109/TMTT.1967.1126457.

[31] R. Levy and J. Rhodes, "A comb-line elliptic filter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 19, no. 1, pp. 26–29, 1971, DOI: 10.1109/TMTT.1971.1127441.

[32] S. Lucyszyn, *Advanced RF MEMS*. Cambridge University Press, 2010, ch. 1, p. 1, DOI: 10.1017/CBO9780511781995.

[33] K. E. Peterson, "Micromechanical membrane switches on silicon," *IBM Journal of Research and Development*, vol. 23, no. 4, pp. 376–385, 1979, DOI: 10.1147/rd.234.0376.

[34] B. Pradhan and B. Gupta, "Novel tunable band reject filter using RF MEMS technology," in *Proceedings of the 2014 IEEE Students' Technology Symposium*, 2014, pp. 335–340, DOI: 10.1109/TechSym.2014.6808071.

[35] J. R. Reid, J. M. Oliver, K. Vanhille, and D. Sherrer, "Three dimensional metal micromachining: A disruptive technology for millimeter-wave filters," in *2012 IEEE 12th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2012, pp. 17–20, DOI: 10.1109/SiRF.2012.6160165.

[36] A. Cohen, G. Zhang, F.-G. Tseng, U. Frodis, F. Mansfeld, and P. Will, "EFAB: rapid, low-cost desktop micromachining of high aspect ratio true 3-D MEMS," in *Technical Digest. IEEE International MEMS 99 Conference. Twelfth IEEE International Conference on Micro Electro Mechanical Systems*, 1999, pp. 244–251, DOI: 10.1109/MEMSYS.1999.746824.

[37] R. Chen and E. Brown, "An ultra-compact low loss 30-ghz micromachined coaxial filter," in *2005 European Microwave Conference*, vol. 1, 2005, DOI: 10.1109/EUMC.2005.1608936.

[38] D. W. Kammler, "Calculation of characteristic admittances and coupling coefficients for strip transmission lines," *IEEE Transactions on Microwave Theory and Techniques*, vol. 16, no. 11, pp. 925–937, Nov. 1968, DOI: 10.1109/TMTT.1968.1126828.

[39] J. D. Horgan, "Coupled strip transmission lines with rectangular inner conductors," *IRE Transactions on Microwave Theory and Techniques*, vol. 5, no. 2, pp. 92–99, Apr. 1957, DOI: 10.1109/TMTT.1957.1125105.

[40] S. Lin and M. Eron, "Development of an ultra-wideband suspended stripline to shielded microstrip transition," *IEEE Microwave and Wireless Components Letters*, vol. 21, no. 9, pp. 474–476, 2011, DOI: 10.1109/LMWC.2011.2163153.

[41] Mathworks, *Global Optimization Toolbox - MATLAB & Simulink*, Accessed: 2024-05-13. [Online]. Available: `www.mathworks.com/help/gads`.

[42] K. Schittkowski, "NLQPL: A fortran-subroutine solving constrained nonlinear programming problems," *Annals of Operations Research*, vol. 5, pp. 485–500, 1986, DOI: 10.1007/BF02739235.

[43] Mathworks, *Constrained Nonlinear Optimization Algorithms - MATLAB & Simulink*, Accessed: 2024-05-13. [Online]. Available: `www.mathworks.com/`

help / optim / ug / constrained - nonlinear - optimization - algorithms . html.

[44] H. Wheeler, "Transmission-line properties of a strip line between parallel planes," *IEEE Transactions on Microwave Theory and Techniques*, vol. 26, no. 11, pp. 866–876, 1978, DOI: 10.1109/TMTT.1978.1129505.

[45] S. Cohn, "Shielded coupled-strip transmission line," *IRE Transactions on Microwave Theory and Techniques*, vol. 3, no. 5, pp. 29–38, 1955, DOI: 10.1109/TMTT.1955.1124973.

[46] T. Hiraoka, K. Ding, and J. P. Hsu, "Calculation of characteristic impedance for parallel coupled stripline based on eigenmode analysis and equivalent network," *IEEE MTT-S International Microwave Symposium Digest*, pp. 269–272, 2009, DOI: 10.1109/MWSYM.2009.5165685.

[47] M. Saih, H. Rouijaa, and A. Ghammaz, "Computation of multiconductor transmission line capacitance using method of moment," in *Proceedings of 2014 Mediterranean Microwave Symposium (MMS2014)*, 2014, pp. 1–6, DOI: 10.1109/MMS.2014.7088781.

A

MATLAB code

A.1 Calculation of filter function

A.1.1 Main code

```
%% Stepped digital filter
% INFO: Extracts S-parameters from capacitance coupling matrix
% So far this code can handle
%   - Homoegeneous dielectrics
%   - 5th order stepped digital filter
% Assuming:
%   - Stepped digital filter with two networks and two unit elements.
%   - The length of each network is lambda/8
%   - No coupling between non-adjacents elements
%   - Short circuit admittances: term_const=10e7, change if needed

%N1=Network 1
%N2=Network 2

%% Parameter setup and initial calculations
%Instructions:
%   1) Choose frequency / plotting parameters
%   2) Choose filter characteristics, find C and Omega from tables
clear
if isempty(findall(0,'Type','figure'))==0 %Clears old plots
   close all
end

par=struct;

%Frequency / plotting
par.f_1=3e9; %Lower sweep frequency
par.f_2=7e9; %Upper sweep frequency
par.f_c=5e9; %Center frequency of filter
par.resolution=3000; %How many frequency points
```

```
par.plot_window_S21=[par.f_1 par.f_2 -100 0]; %[Xmin Xmax Ymin(dB) Ymax(dB)]
par.plot_S_param='yes';

%Filter characteristics
par.order=5;        %Order of filter
par.frac_bw=0.11;   %Fractional bandwidth
par.Z0=50; %Scales the matrix to XX-Ohms

%Tabulated values
par.C=[1,1,1,1,1]; par.Omega_r=[1,1];

%Capacitance matrix scaling
par.align_center_element_impedance='yes'; %Aligns center element impedance level to other lines
par.diag_scaling_const=1.2; %changes impedances of node 2 to 4
par.n_mat=1*[1,1,1,1,1].'; %Choose scaling freely

%THE FOLLOWING SHOULD NOT BE CHANGED

%Set constant parameters
par.c=physconst('lightspeed');
par.epsilon_0=8.8541878128e-12;
par.mu_0=1.25663706212e-6;
par.lambda_c=par.c/par.f_c; %Wavelength at center freq
par.l=par.lambda_c/8; %As long as the length of the lines are lambda/8 this is constant
par.beta=2*pi/par.lambda_c; %Wavenumber
par.eta_0=120*pi;
par.vp=1/sqrt(par.epsilon_0*par.mu_0*par.epsilon_r); %Phase velocity
par.eps_eff=(par.c/par.vp)^2;

%Set filter parameters
par.norm_freq=pi/4;
par.norm_omega_c=par.norm_freq*(1-par.frac_bw/2);
par.a=tan(2*par.norm_omega_c)/2;
par.lambda_r_plus=sqrt(((par.Omega_r(1:end)./(2*par.a)).^2)+1)+par.Omega_r(1:end)./(2*par.a);
par.lambda_r_minus=sqrt(((par.Omega_r(1:end)./(2*par.a)).^2)+1)-par.Omega_r(1:end)./(2*par.a);
par.imp_scaling_constant=par.eta_0/(par.Z0*sqrt(par.epsilon_r));
par.size_N1=par.order+2;
par.size_N2=par.order;
par.M_size=par.order*3+4; %N1+N2+4 unit element nodes, size of final coupling matrix
par.term_const=10e6;
```

```
71
72
73  %% 5-th order filter
74  l=par.l;
75  a=par.a; order=par.order; size_N1=par.size_N1; size_N2=par.size_N2;
76  lambda_mat=[par.lambda_r_plus(1) par.lambda_r_minus(1) par.
        lambda_r_plus(2) par.lambda_r_minus(2)];
77  C1=par.C(1); C2=par.C(2); C3=par.C(3); C4=par.C(4); C5=par.C(5);
78
79  Y_aa=a*C1; Y_bb=0; Y_cc=a*C3; Y_dd=0; Y_ee=a*C5;
80  Y_ab=a*C2*(1+par.lambda_r_plus(1)^2);
81  Y_bc=a*C2*(1+par.lambda_r_minus(1)^2);
82  Y_cd=a*C4*(1+par.lambda_r_plus(2)^2);
83  Y_de=a*C4*(1+par.lambda_r_minus(2)^2);
84
85  Y=zeros(par.order);
86  %Network 1
87  Y(1,1)=Y_aa+Y_ab; Y(2,2)=Y_bb+Y_ab+Y_bc; Y(3,3)=Y_cc+Y_bc+Y_cd; Y
        (4,4)=Y_dd+Y_cd+Y_de; Y(5,5)=Y_ee+Y_de;
88  Y(2,1)=-Y_ab;
89  Y(3,2)=-Y_bc;
90  Y(4,3)=-Y_cd;
91  Y(5,4)=-Y_de;
92
93  Y1=Y+tril(Y,-1).';
94
95  %Network 2
96  [Y_ab, Y_bc] = deal(Y_bc,Y_ab); %Switches the values for the coming
        calculations
97  [Y_cd, Y_de] = deal(Y_de,Y_cd); %See Rhodes eq.(6) and (7) for
        details
98
99  Y(1,1)=Y_aa+Y_ab; Y(2,2)=Y_bb+Y_ab+Y_bc; Y(3,3)=Y_cc+Y_bc+Y_cd; Y
        (4,4)=Y_dd+Y_cd+Y_de; Y(5,5)=Y_ee+Y_de;
100 Y(2,1)=-Y_ab;
101 Y(3,2)=-Y_bc;
102 Y(4,3)=-Y_cd;
103 Y(5,4)=-Y_de;
104
105 Y2=Y+tril(Y,-1).';
106
107 %Transformer action on network 1, and scaling by 1/sqrt(a) and a
108 Y1_transformer = zeros(par.size_N1, par.size_N1);  %Creates N+2
        matrix
109 Y1_transformer(2:par.size_N1-1, 2:par.size_N1-1) = Y1/a;
110 Y2_scaled=Y2/a;
```

```
111
112 Y1_transformer(1:2, 1:2) =  [1 -1/sqrt(a); -1/sqrt(a) 1/a+
       Y1_transformer(2,2)]; %Adds in top left corner
113 Y1_transformer(par.size_N1-1:par.size_N1, par.size_N1-1:par.size_N1)
       = [1/a+Y1_transformer(par.size_N1-1,par.size_N1-1) -1/sqrt(a); -1/
       sqrt(a) 1]; %Adds in bottom right corner
114
115 Y1_scaled_except_diag=Y1_transformer*par.imp_scaling_constant;
116 Y2_scaled_except_diag=Y2_scaled*par.imp_scaling_constant;
117
118 %Scales matrix by finding scaling constants
119 [Y1_scaled, Y2_scaled, scaling_constants_Y1, scaling_constants_Y2]=
       scale_matrix(Y1_scaled_except_diag,Y2_scaled_except_diag,par.
       diag_scaling_const,par);
120
121 %To get correct units of the matrix
122 C1_mat=Y1_scaled*par.epsilon_0;
123 C2_mat=Y2_scaled*par.epsilon_0;
124
125 %Initialize loop
126 steps=par.resolution;
127 f_2=par.f_2; f_1=par.f_1;
128 step=abs(f_2-f_1)/steps;
129
130 M_size=par.M_size;
131
132 S11=zeros(steps,1);
133 S21=zeros(steps,1);
134 for i=1:steps
135     f_m=f_1+(i-1)*step;
136
137     Z_A1=1i*C1_mat^(-1)*2*pi*f_m/par.vp^2;
138     Y_A1=1i*C1_mat*2*pi*f_m;
139     Z_A2=1i*C2_mat^(-1)*2*pi*f_m/par.vp^2;
140     Y_A2=1i*C2_mat*2*pi*f_m;
141
142     %The following comes from Knockaert paper
143     ZY_1=(Z_A1*Y_A1);
144     YZ_1=(Y_A1*Z_A1);
145     ZY_2=(Z_A2*Y_A2);
146     YZ_2=(Y_A2*Z_A2);
147
148     %SVD is used since eig() gives numerical error for eigenvalue
149     %Calculates the eigenvectors
150     [Mv_1,~]=eig(YZ_1);   [Mv_2,~]=eig(YZ_2);
151     %Calculates the eigenvalues
```

```
152     eig_1=-diag(svd(YZ_1)); eig_2=-diag(svd(YZ_2));
153     Mi_2=eye(order) / Mv_2.';
154
155     gamma_mat_1=sqrt(eig_1);
156     gamma_mat_2=sqrt(eig_2);
157
158     ZT_1=Z_A1*Mv_1; Tgamma_1=Mv_1*gamma_mat_1;
159     Zc_1= ZT_1 / Tgamma_1;
160     Yc_1 = eye(order+2) / Zc_1;
161
162     ZT_2=Z_A2*Mv_2; Tgamma_2=Mv_2*gamma_mat_2;
163     Zc_2= ZT_2 / Tgamma_2;
164     Yc_2 = eye(order) / Zc_2;
165
166
167     %Only one mode since it is a TEM line
168     gamma_1=sqrt(eig_1(1,1));
169     gamma_2=sqrt(eig_2(1,1));
170
171     Ym_1=[coth(gamma_1*l)*Yc_1 -csch(gamma_1*l)*Yc_1; -csch(gamma_1*l
            )*Yc_1 coth(gamma_1*l)*Yc_1];
172     Ym_2=[coth(gamma_2*l)*Yc_2 -csch(gamma_2*l)*Yc_2; -csch(gamma_2*l
            )*Yc_2 coth(gamma_2*l)*Yc_2];
173
174     n2_start=order+4; %order+2(Network 1)+1(unit element)+1(get to
            next node)
175     n1_end=2*order+5; %order+2(N1)+1(UE)+1(next)+order(N2)+1(UE)+1(
            get to next)
176     combined_matrix=zeros(M_size);
177     combined_matrix(1:length(Ym_1),1:length(Ym_1))=Ym_1; %Adds
            network 1 to the combined matrix
178     combined_matrix(size_N1+2:2*order+3,size_N1+2:2*order+3)=
            combined_matrix(size_N1+2:2*order+3,size_N1+2:2*order+3)+Ym_2
            (1:order,1:order); %Adds common elements to the diagonal
179     combined_matrix(n1_end:length(combined_matrix),n1_end:length(
            combined_matrix))=Ym_2(order+1:end,order+1:end); %Adds most
            elements of Ym_2
180
181     combined_matrix(n2_start:2*order+3,n1_end:length(combined_matrix)
            )=Ym_2(1:order,order+1:2*order);
182     combined_matrix(n1_end:length(combined_matrix),n2_start:2*order
            +3)=Ym_2(order+1:2*order,1:order);
183
184     %Add port termination conditions, short circuit (network 1)
185     term_const=1i*par.term_const; %Short circuit = infinite
            admittance
```

```
186
187     combined_matrix(2:order+1,2:order+1)=combined_matrix(2:order+1,2:
            order+1)+diag(diag(eye(order)*term_const));
188     combined_matrix(order+3,order+3)=term_const;  %Shorted part of
            input transformer
189     combined_matrix((order+2)*2,(order+2)*2)=term_const;  %Shorted
            part of output transformer
190
191     Y0=zeros(length(combined_matrix));
192     Y0(1,1)=1/par.Z0;
193     Y0(7,7)=1/par.Z0;
194
195     S=(Y0-combined_matrix) / (Y0+combined_matrix);
196     two_port=[S(1,1) S(1,7); S(7,1) S(7,7)];
197     two_port_dB=20*log10(abs([S(1,1) S(1,7); S(7,1) S(7,7)]));
198     S21(i)=two_port_dB(2,1);
199     S11(i)=two_port_dB(1,1);
200
201     %Just to save center frequency values if wanted
202     if f_m==par.f_c
203         Yc_2_fc=Yc_2;
204         Yc_1_fc=Yc_1;
205         S_fc=S;
206         combined_matrix_fc=combined_matrix;
207
208         %Impedance of the lines
209         Zc_1_lines=1./real(diag(Yc_1_fc(2:6,2:6)));
210         Zc_2_lines=1./real(diag(Yc_2_fc));
211     end
212 end
213 if isequal(lower(par.plot_S_param),'yes')
214     linewidth=1.4;
215     freq=linspace(f_1*1e-9,f_1*1e-9+step*steps*1e-9,length(S21));
216     plot(freq,S11,'r','LineWidth',linewidth)
217     hold on
218     plot(freq,S21,'b','LineWidth',linewidth)
219     axis([par.plot_window])
220     xlabel('Frequency (GHz)','FontSize',15)
221     ylabel('S-parameters (dB)','FontSize',15)
222     legend('S_{11}','S_{21}','FontSize',15)
223 end
```

A.1.2 Function for scaling the matrices

```
1 function [Y1_scaled, Y2_scaled, scaling_constants_Y1,
     scaling_constants_Y2] = scale_matrix(Y1,Y2,diag_scaling_const,par)
```

```
% Takes in matrices Y1 and Y2 that are yet to be scaled
% Only works for 5th degree filters but is easily extended
% to higher order filters if needed, same method
% IMPORTANT! The matrices have already been scaled
  with /a and /sqrt(a) where needed
   Y1_org=Y1; Y2_org=Y2;

   %Scale the diagonals by scaling constant (N1)
   Y1(3,3)=Y1_org(2,2)*diag_scaling_const;
   Y1(5,5)=Y1_org(6,6)*diag_scaling_const;

   %Scale the diagonals by scaling constant (N2)
   Y2(2,2)=Y1(2,2)*diag_scaling_const;
   Y2(4,4)=Y1(6,6)*diag_scaling_const;

   %Not necesarry, makes the center elements align more closely in
      impedance level to the other lines
   if strcmp(par.align_center_element_impedance,'yes')
       Y1(4,4)=Y1(5,5); %Network 1 center element
       find_scaling_Y1=sqrt(diag(Y1(2:6,2:6))./diag(Y1_org(2:6,2:6))
          );
       Y2(3,3)=Y2(3,3)*find_scaling_Y1(3)^2; %Network 2 center
          element
   end

   find_scaling_Y1=sqrt(diag(Y1(2:6,2:6))./diag(Y1_org(2:6,2:6)));
   find_scaling_Y2=sqrt(diag(Y2)./diag(Y2_org));

   %Saves these diagonals for later
   Y1_diag=diag(Y1);
   Y2_diag=diag(Y2);

   %Symbolically adds all scaling constants to off-diagonal elements
   syms n1 n2 n3 n4 n5

   sym_Y1=sym(Y1);
   sym_Y2=sym(Y2);

   sym_Y1(2,:)=sym_Y1(2,:)*n1; sym_Y1(:,2)=sym_Y1(:,2)*n1;
   sym_Y1(3,:)=sym_Y1(3,:)*n2; sym_Y1(:,3)=sym_Y1(:,3)*n2;
   sym_Y1(4,:)=sym_Y1(4,:)*n3; sym_Y1(:,4)=sym_Y1(:,4)*n3;
   sym_Y1(5,:)=sym_Y1(5,:)*n4; sym_Y1(:,5)=sym_Y1(:,5)*n4;
   sym_Y1(6,:)=sym_Y1(6,:)*n5; sym_Y1(:,6)=sym_Y1(:,6)*n5;

   sym_Y2(1,:)=sym_Y2(1,:)*n1; sym_Y2(:,1)=sym_Y2(:,1)*n1;
   sym_Y2(2,:)=sym_Y2(2,:)*n2; sym_Y2(:,2)=sym_Y2(:,2)*n2;
```

```
45      sym_Y2(3,:)=sym_Y2(3,:)*n3; sym_Y2(:,3)=sym_Y2(:,3)*n3;
46      sym_Y2(4,:)=sym_Y2(4,:)*n4; sym_Y2(:,4)=sym_Y2(:,4)*n4;
47      sym_Y2(5,:)=sym_Y2(5,:)*n5; sym_Y2(:,5)=sym_Y2(:,5)*n5;
48
49      %Substitutes symbolic scaling constants to the previously found
            ones
50      Y1_new=double(subs(sym_Y1, [n1 n2 n3 n4 n5], find_scaling_Y1.'));
51      Y2_new=double(subs(sym_Y2, [n1 n2 n3 n4 n5], find_scaling_Y2.'));
52
53      %Creates the new matrices
54      Y1_scaled=Y1_new; Y2_scaled=Y2_new;
55
56      %The diagonal elements were also scaled in the symbolic
            calculations,
57      %so now the diagonals have to be reverted to the saved diagonals
58      for i=1:length(Y1)
59          Y1_scaled(i,i)=Y1_diag(i); %Final, scaled matrix
60      end
61
62      for i=1:length(Y2)
63          Y2_scaled(i,i)=Y2_diag(i); %Final, scaled matrix
64      end
65
66      scaling_constants_Y1=find_scaling_Y1;
67      scaling_constants_Y2=find_scaling_Y2;
68  end
```

B

Wave propagation figures

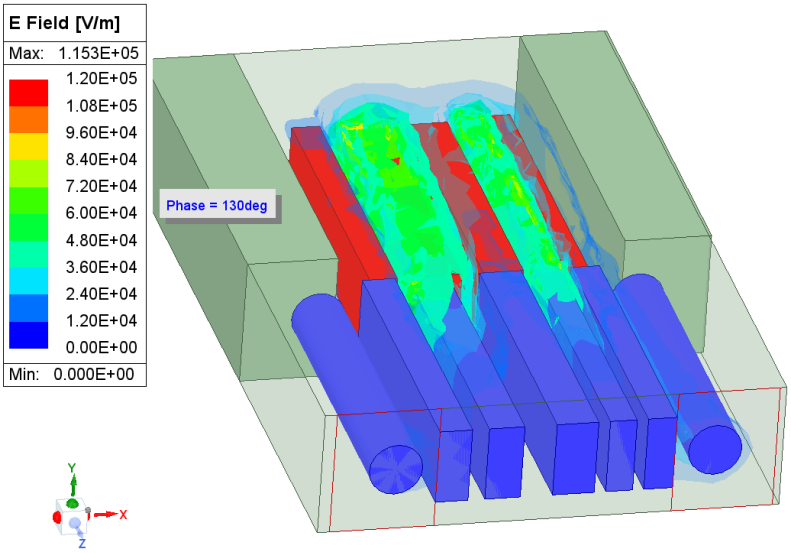


Figure B.1: Wave propagation in filter at 5 GHz.

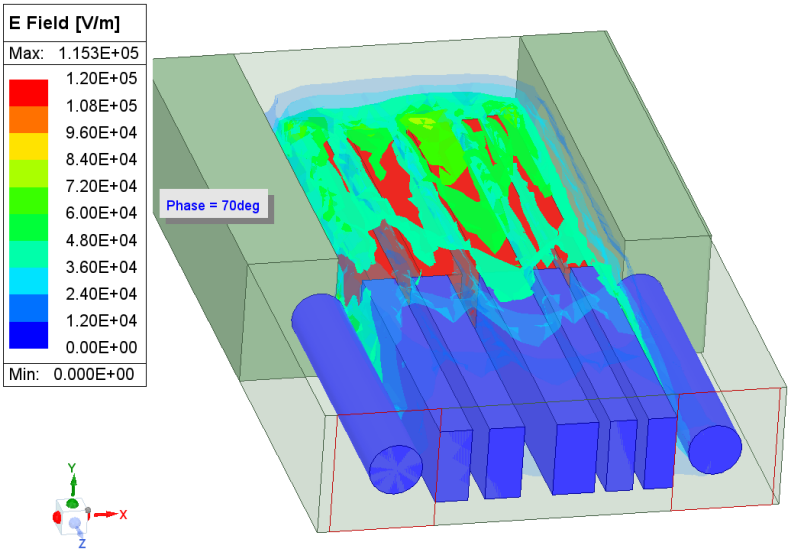


Figure B.2: Wave propagation in filter at 5 GHz.

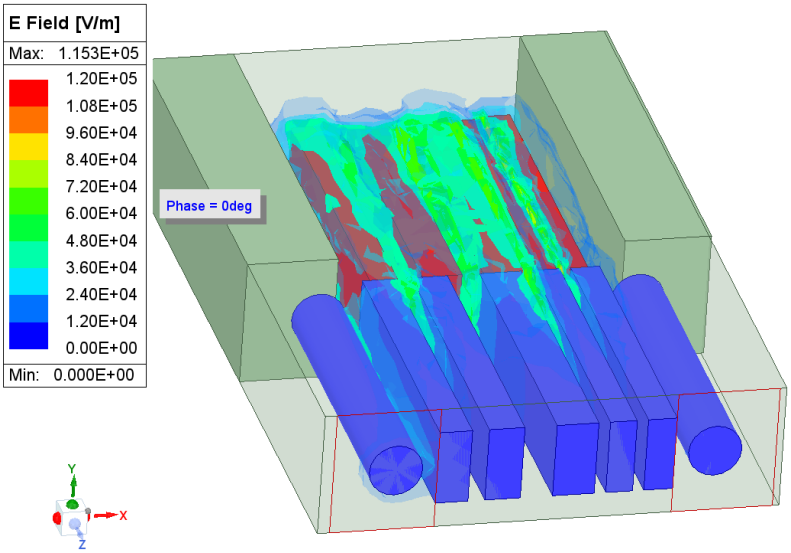


Figure B.3: Wave propagation in filter at 5 GHz.

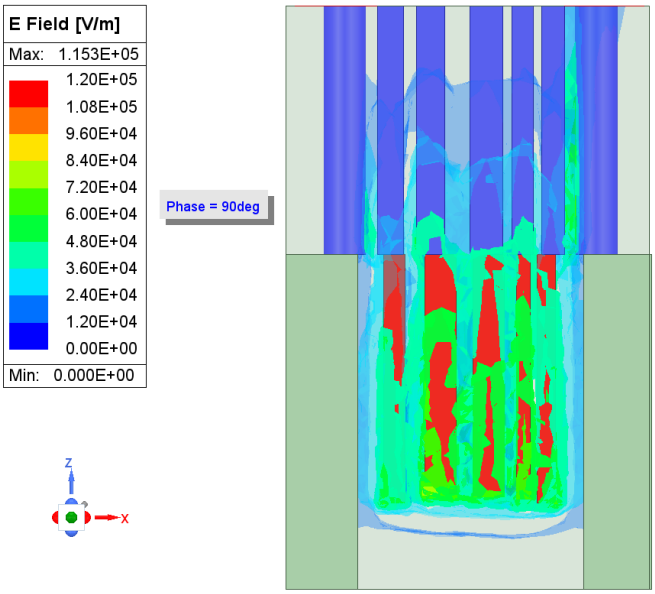


Figure B.4: Wave propagation in filter at 5 GHz.