UNIVERSIDAD DE ANTIOQUIA
1803
Facultad de Ciencias Exactas y Naturales

Quadrature Hybrid Optimization for ALMA Bands 2 and 3

Particle Swarm algorithm for millimeter and sub-millimeter microwave devices optimization

Master's thesis in Physics

Jorge Hernán Cárdenas Lopera

Faculty of Exact and Natural Sciences - Physics Institute

UNIVERSIDAD DE ANTIOQUIA
Medellín, Colombia 2022
www.udea.edu.co

Master's thesis

Quadrature Hybrid Optimization for ALMA Bands 2 and 3

Particle Swarm algorithm for millimeter and sub-millimeter microwave devices optimization

JORGE HERNÁN CÁRDENAS LOPERA

UNIVERSIDAD DE ANTIOQUIA
1803
Facultad de Ciencias Exactas y Naturales

Faculty of Exact and Natural Sciences
Physics Institute
Computational Physics and Astrophysics (FACom)
Universidad de Antioquia
Medellín, Colombia 2022

Quadrature Hybrid Optimization for ALMA Bands 2 and 3
Particle Swarm algorithm for millimeter and sub-millimeter microwave devices optimization

Candidate:
Jorge Hernán Cárdenas Lopera
jorge.cardenas3@udea.edu.co
Universidad de Antioquia.

Supervisor:
PhD Germán Chaparro.
Associate Professor, Institute of Physics
german.chaparro@udea.edu.co
Universidad de Antioquia.

Co-supervisor:
PhD(c) Óscar Restrepo.
Assisstant Professor
orestrepog@ecci.edu.co
Universidad ECCI.

Approved by:

Examiner	Date
Examiner	Date

Master's Thesis 2022
Faculty of Exact and Natural Sciences
Physics Institute
Computational Physics and Astrophysics (FACom)
Universidad de Antioquia
Medellín, Colombia 2022

Abstract

We introduce a novel optimization method based on the evolutionary algorithm Particle Swarm Optimization (PSO) to enhance the electromagnetic performance of quadrature hybrid designs. Optimization and simulations batches were conducted using a fully tuned and validated version of the algorithm, for the design of quadrature hybrids intended to operate ALMA (Atacama Large Millimeter Array) Band 2 (67-90 GHz), Band 3 (84-116 GHz), and Band 2+3 (67-116 GHz).

Thus, we present quadrature hybrid designs which are optimized to operate in ALMA Band 3 (84-116 GHz) with respect to their operational requirements for the scattering parameters and amplitude imbalance. Furthermore, the resulting designs take into account machining constraints related to cost and feasibility requirements.

Finally, this work provides a method that can be easily extended to optimize other microwave devices and waveguides for radio astronomy applications, with the benefit of speeding up the design process as well as reducing the computational costs.

Keywords: radio astronomy, ALMA, quadrature hybrid, directional coupler, optimization algorithm, amplitude imbalance, millimeter and submillimeter device.

Acknowledgements

It has been a long journey from the moment I got the 'Yes' to embark in a Master's in Physics, and I know that the fulfillment of this research is the product of many hands, minds and hearts.

In the first place, I want to thank to my supervisor Germán. It is not just because of his shared knowledge and experience, but for the guiding and leadership that gives room to grow. I hope that many undergraduate and graduate students have the chance to experience learning from a bold mind while feeling respected and considered.

Without question, Oscar Restrepo has been a great mentor during this time. Oscar, as a co-supervisor, was always there, providing ideas and the raw materials to develop what finally became our implementation of this method.

Along this research path I received invaluable information, input and feedback from Patricio Mena (NRAO) and David Monasterio (University of Chile). Two of the most knowledgeable and experienced engineers in the field, who provided fundamental concepts to test our designs and to validate our findings.

The University of Antioquia has played a major role in this project, not just because its great exact sciences faculty, but by providing financial support for this project to be presented in the *32nd IEEE International Symposium on Space Terahertz Technology*.

To my wife Juliana, who has been so patient and loving, but more importantly, has encouraged me to move forward in times when my strength was failing. She inspires me and provides me with the motivation to keep working for our dreams. To my baby, Emma, who is already filling me with joy and expectancy about the future. I am waiting to hold you in my arms.

To my parents, my sister, my brother in law, and Maximiliano, my nephew. They have always been a driving force. They supported me during the most difficult times and, because of them, I am always trying to pursue the best version of me.

Finally, as a person of faith who considers life as a continuous research and reframe of ideas, I want to thank God. As far as I know, there is no further explanation to many universal unknowns, and I am really fine with this one.

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

ALMA	Atacama Large Millimeter Array
EA	Evolutionary Algorithms
FF	Fitness Function
GA	Genetic Algorithms
IF	Intermediate Frequency
IR	Image Rejection
QH	Quadrature Hybrid
LSB	Lower Side Band
PSO	Particle Swarm Optimization
RF	Radio Frequency Signal
SED	Spectral Energy Distribution
TE	Transverse Electric
TEM	Transverse-electromagnetic
TM	Transverse Magnetic
WCA	Warm Cartridge Assembly
USB	Upper Side Band

Contents

List of Figures

List of Tables

1

Introduction

1.1 Radio Astronomy Fundamentals

1.1.1 Radio Emissions

Radio waves have the longest wavelength in the spectrum, and can be created from a myriad of astrophysical phenomena. Most radio-loud sources, like quasars, are bright in the radio spectrum, not due to thermal emission, but due to its free-free and synchroton emission.

One of the first radio detections was performed by Karl Jansky in 1930, and ended up identifying our own galaxy as an intense radio source. Synchrotron radiation was later introduced as one of the models to explain high radio brightness and that of the galactic radio background. This mechanism produces electromagnetic radiation as a consequence of a relativistic electron passing through a region with a magnetic field. The spectrum of this emission results from adding the emission spectra of individual electrons, which finally yields a total flux $F \sim \nu^{\alpha}$, where α is called the spectral index. Some of the most significant sources of this type of emission are pulsars, supernova remnants and active galactic nuclei (AGN).

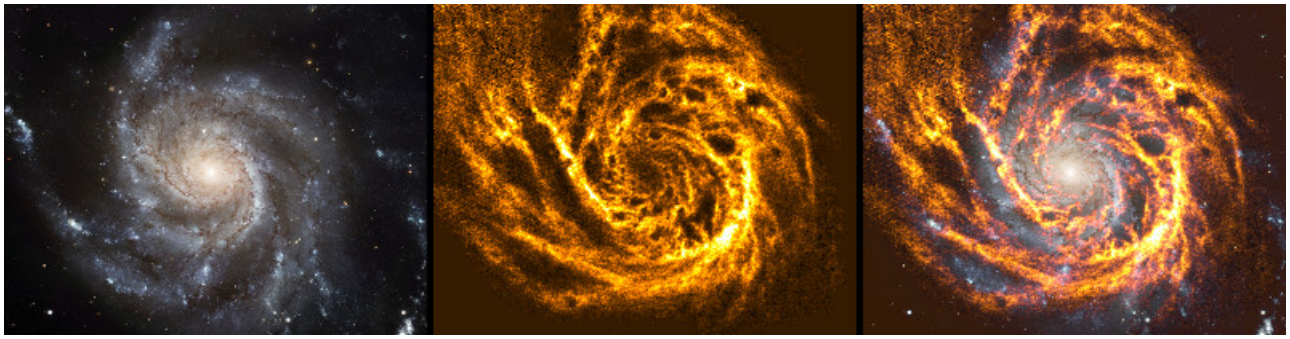


Figure 1.1: NGC 5457 in the visible spectrum taken by the Hubble Space Telescope and the atomic hydrogen detected using the VLA (Very Large Array). **Source:** http://galaxymap.org/drupal/node/202

Synchrotron radiation is just one mechanism among many others of the so-called free-particle emissions. *Brehmsstrahlung* radiation fits in this category resulting from the interaction of unbound free particles and ions. The electrostatic interaction determines the energy of the emitted photon. Moreover, to obtain radio emissions, the interaction should be relatively gentle collisions of electrons with ions.

Radio signals related to thermal processes are comparatively dim, and are described by the Planck's law for blackbody emission,

$$B_{\nu}(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \quad . \tag{1.1}$$

Thus can be expressed as the Rayleigh-Jeans approximation, describing the thermal emission of a blackbody in the frequency band corresponding to radio waves, namely in the limit where $h\nu << kT$ (Fig. 1.2),

$$B_\nu(\nu, T) = \frac{2kT\nu^2}{c^2} \quad . \tag{1.2}$$

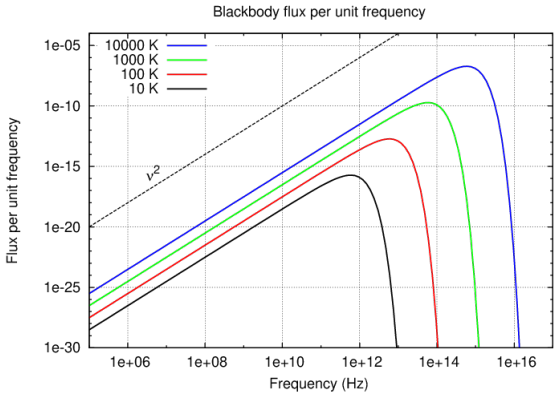


Figure 1.2: Black body flux per unit frequency. The Rayleigh-Jeans approximation dominates in the radio part of the spectrum. **Source:** http://spiff.rit.edu/classes/ast613/lectures/radio_i/radio_i.html

On the other hand, at the atomic level, emissions can result from state transitions, for example, *Radio Recombination lines* found in HII regions where electron recombination generates photon emission. Another atomic line come from the hyper-fine transition of neutral hydrogen, which has a natural frequency of 1.420 MHz. This is commonly known as 21-cm emission. By studying this line we can learn about the velocity of neutral hydrogen clouds, in our galaxy and other galaxies, helping measure the galactic dark matter haloes.

In addition to these radiation mechanisms, another source of radio line emission is due to molecular transitions. If we consider molecules as CO, O_2, or H_2, they can experiment longitudinal or transversal oscillations as well as rotational movements around different axes. This rotational behavior is quantized in terms of the total angular momentum J. The energy is thus defined as,

$$E = \left(\frac{h}{2\pi}\right)^2 \frac{J(J+1)}{2I} \quad . \tag{1.3}$$

Here, $I = \mu r^2$ is the molecular momentum of inertia. Therefore, for a CO molecule, in which a transition occurs between $J = 1$ to $J = 0$, a spectral line results at frequency of 115.2712 GHz. This molecule is frequently used as a tracer of clouds of molecular hydrogen as well as an indication of a cold environment in which there is a low level of ionizing radiation. Also, the interstellar medium is filled with both gas and small solid particles called *dust grains*, which play an important role in the scattering of photons.

1.1.2 ALMA Observatory

In a very broad definition, ALMA is an aperture synthesis telescope consisting of 66 antennas, operating in a millimiter and submillimeter regime in a broad frequency range. The idea behind aperture synthesis comes from *synthesizing* a pair of telescopes into a "new" one with a wider aperture. To achieve this, the source is observed by every telescope in the array, and then combining their outputs through mathematical methods. The resulting image resolution is equivalent to the one of a telescope with a diameter equal to the largest distance between antennas in the array.

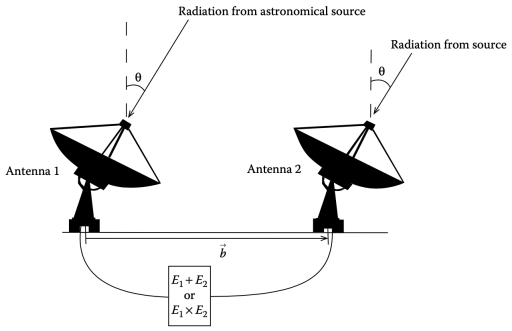


Figure 1.3: Two-antenna interferometer. **Source:** Marr et al., 2016

ALMA is composed of 66 12-meter antennas to reach an initial observation window covering from 84 to 950 GHz and, through a continuous upgrade plan, it got extended to 35-950 GHz. To appropriately operate in this frequency regime, the observatory had to be located in one of the driest places on earth, the plain of Chajnantor, in order to elude the effect of atmospheric water vapor that could prevent the electromagnetic wave to reach the instruments on the ground.

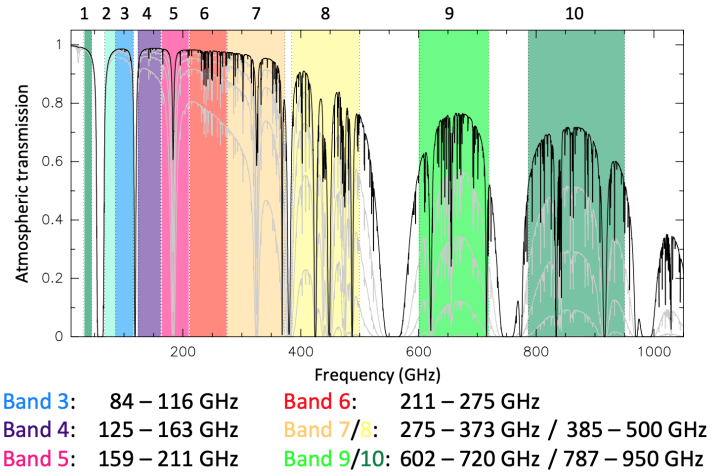


Figure 1.4: ALMA frequency bands. **Source:** https://astro.uni-bonn.de/ARC/events/proposalprep2022/tutorials/ALMA_Prop_Prep2022_LifeOfAnAlmaProject_1.pdf

Each antenna is loaded with a front-end including a cryostat which contains up to ten cartridges, each of them covering one frequency band (see Fig. 1.4), but only one band can be observed at a given moment. To observe in different bands or atmospheric windows, a rapid switching between bands is possible by mechanically adjusting the secondary.

1.1.2.1 Science in ALMA Bands 2 and 3

The scope of our research is set for the study of receiver-microwave devices intended to operate in ALMA Band 2 (67-90 GHz) and 3 (84-116 GHz), pushing the limits for microwave device optimization that could be used in the ongoing effort to develop and deploy a single receiver to fully cover ALMA Band 2+3 (67-116 GHz) (Yagoubov et al., 2020). The operation in these bands can benefit from improved sensitivities due to the implementation of optimized microwave devices.

Three level-one scientific goals have been defined for ALMA, all of which involve Band 2 and 3, namely a) to detect spectral line emission from CO or C+ in a normal galaxy like the Milky Way at a redshift of $z = 3$, in less than 24 hours of observation, b) ALMA should possess the ability to image the gas kinematics in a solar-mass protostellar/protoplanetary disk at a distance of 150 pc (roughly, the distance of the star-forming clouds in Ophiuchus or Corona Australis), enabling the study of the physical, chemical, and magnetic field structure of the disk and to detect the tidal gaps created by planets in formation, and c) The ability to provide precise images at an angular resolution of 0.1" (Schieven, 2022).

1.1.2.1.1 High redshift studies Band 2 and 3 possesses a low atmospheric opacity that is available for >87.5% of ALMA observing conditions (Yagoubov et al., 2020), allowing the detection of deuterated molecules in cold, dense gas and the study of redshifted CO and other species emission from galaxies. One of the essentials aims in ALMA relates to the understanding of the distribution and evolution of galaxies through time. To 'see' farther is cosmic time, ALMA looks for high redshift measurements of molecular and atomic lines as CO, CI, CII and H_2O (Thompson et al., 2001).

The detection of CO in high redshift objects has an impact in the study of galaxies and stars formation. In particular, provides evidence of significant amounts of C and O being produced in the early universe. While some regions are not accessible in the optical spectrum, the detection of CO offers a tool to study them and to use its kinematics to obtain physical conditions as temperature and densities (Guilloteau, 2001). On the other hand, galaxies contain large molecular clouds in which new stars are forming. Under these clouds many atomic and molecular transitions take place in the radio ranges. For instance the $^{12}C^{18}O$ molecule transition $J = 1 \rightarrow 0$ emits at 109.782 GHz, or the SiO $J = 2 \rightarrow 1$ emits at 88.632 GHz, furthermore, the CII line at 1.90054 THz will be doppler shifted into the radio window for redshift >2.

Another interesting case is the study of high-redshift absorption lines. It takes place as we observe gas against a bright background source, and provides information on the ISM of the foreground galaxy (Schieven, 2022). Moreover, in ALMA cycle 9 observations, a Band 3 spectral survey is prepared to probe a spiral galaxy at $z \approx 0.9$ which lies in front of a bright background quasar at at $z \approx 2.5$. ALMA capability of very high redshift detections, enable the study of key elements from the first stellar objects ($z > 10$) by detecting the cooling lines.

1.1.2.1.2 Star- and planet-forming gas The study of rotational transitions are useful to characterize cold galactic sources such as pre-stellar cores, and other sources relevant to astrobiology like protoplanetary disks. There is on-going research on the evolutionary state of protostellar disks in the vicinity of the Ophiuchus molecular cloud. Dust around a protostar is expected to evolve as dust grains concentrate in the mid-plane of the disk. The evolution can be traced by determining the Spectral Energy Distribution (SED) to infer dust properties which can be obtained by measuring the dust temperature, density distribution and optical properties (Schieven, 2022). For a given dust grain composition, the slope of the emission flux vs the wavelength (see Fig. 1.7) traces the dust grain size distribution, thereby measuring its variation from the interstellar medium to protoplanetary disks allows for tracing of the evolution from dust, then to grains and, finally, to planets.

ALMA Band 2 allows the measurement of cool molecular gas in high-redshift galaxies, but there is an additional need to probe dense gas ($\gtrapprox 10^5 cm^{-3}$) which is associated with star-formation. This is achieved by observing molecules with higher dipole moments such as HCN or HCO^+ (Fuller et al., 2020).

1.1.2.1.3 Band 2+3 Regarding the extended ALMA Band 2+3, since the early ALMA design, technology and manufacture capabilities have advanced enough to expect a feasible production of a single wideband receiver to fully cover the range spanned by Bands 2 and 3 (Fuller et al., 2018). However, this is still to be proven, as the highly-demanding wideband constrains to fully cover the Band 2+3 set rigorous constructive

limits. Furthermore, despite the on-going efforts to produce a receiver that fully complies with ALMA requirements, there is no evidence to support an operational device which includes quadrature hybrids capable of fully covering this operational band, while fully complying with the required ΔA and $\Delta\phi$.

Apart from the technical challenges, the study of high redshift galaxies is highly favored by extending the capabilities of Band 3 by joining Band 2 in a single detector. Calculations suggest that the spectral survey would be 2.5 times faster by using a single receiver, and additionally, there will be an increase in redshift identification efficiency (Yagoubov et al., 2020). In Fig. 1.5, the CO ladder is presented for the different ALMA bands. As we consider Bands 2 and 3, we can see how distinct CO transition emissions can be detected in a wide range of redshift thus allowing for a continuous measurement of the CO by using a single receiver spanning Band 2+3.

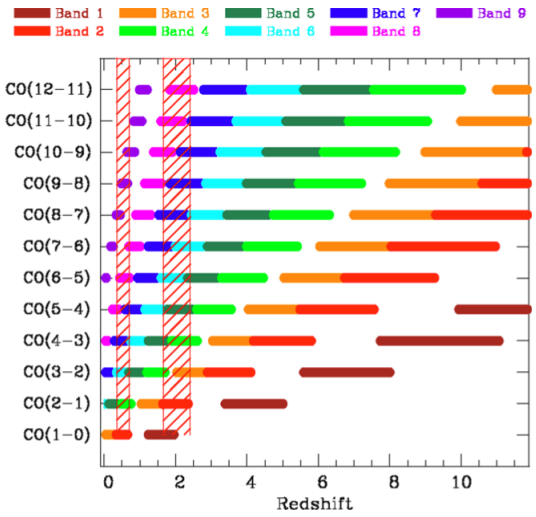


Figure 1.5: CO coverage by ALMA receivers. **Source:**Fuller et al., 2020

Now, regarding the dense gas tracers, the use of an extended band opens a the possibility to detect sources up to $z \approx 0.3$, which is a great improvement as compared to the $z \approx 0.05$ achieved by using single Band 3 observations. The benefit relies on the increase in the number of possible targets in which to trace in the search for gas parameters. The Fig. 1.6, shows the different species that could be detected to study dense star forming gas in the whole range spanned by the Band 2+3 at different redshifts.

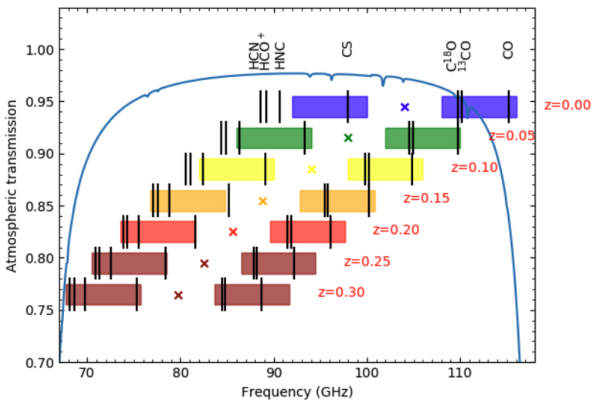


Figure 1.6: Dense gas tracers as a function of redshift in Band 2+3. **Source:**Fuller et al., 2020

While studying the planet formation process, as motioned above, there is great interest in tracing dust evolution. In the Fig. 1.7, we can see in the orange box, the corresponding Band 2+3. Previous measurements of the grain growth required observations from different telescopes with the setback of requiring calibration to ensure the accuracy needed to derive the showed slope and thus to obtain the dust grain size distribution. There is an accuracy gain by using an extended band and, consequently, there is improvement of the signal to noise ratio hence allowing a more precise measurement of the grain size variation within the disks.

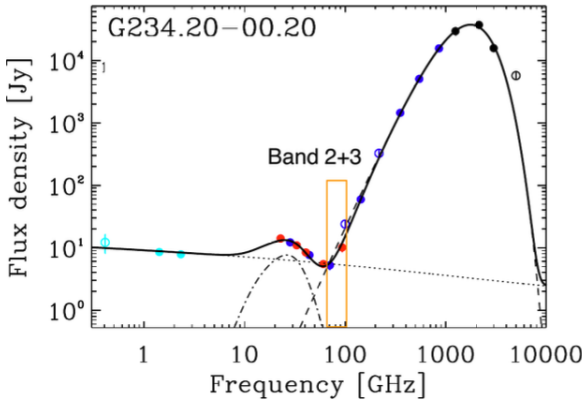


Figure 1.7: Spectral Energy Distribution of spinning dust. **Source:**Yagoubov et al., 2020

1.1.3 Radio Detectors

Radio telescope are completely passive devices intended to convert electromagnetic waves into electric current. It has two functional sections namely the front-end and the back-end. The former provides amplification and frequency down conversion of the collected waves. As it is usually composed of a set of waveguides hence it establishes an operational limit to the receiver. The latter has the function of further processing the signal by applying analog to digital conversion and finally storing the data.

Radio telescope antennas are primarily receiving devices whose most salient properties are,

- sensitivity: the light gathering power. This property is highly dependent on the collecting area thus larger dishes are capable of more light collecting. As global parameter, the sensitivity is determined by the noise added in every step of the gathering and processing of the observed signal.
- dish irregularities: the surface can be unpolished conversely to the optical counterpart. The requirement is that irregularities must be smaller than the wavelength of the radiation. This is the reason why at long wavelengths the dish can be a mesh with holes.
- resolution: for radio telescopes the angular resolution $\theta \propto \frac{\lambda}{D}$, is the ability to see fine details in the sky. It depends on the wavelength and the diameter D, thus, the larger the dish the better the angular resolution.
- directivity: accounts for the ratio of the maximum to the average radiation intensity. It is a measure of how concentrated is the radiation in a given direction.
- gain: the gain is more generally defined as the relation between the output and the input power. As for the astronomical observations, we relate the antenna temperature to flux density for a point source, thus we measure how much antenna temperature results for each Jansky of flux density from a point source. In radio astronomy, a large gain means that the telescope is very sensitive thus it has a strong response to the incident radiation. Moreover, a large gain implies that even a weak source with a small flux density will produce a measurable antenna temperature.

- radiation pattern: describes the response of the antenna as a function of direction. This pattern is presented in the form of lobes depicting the gain in polar coordinates.
- efficiency: due to the losses associated to material's conductivity or dielectric losses, the collected energy is not fully converted into electric power.
- feed: its main goal is to confine electromagnetic waves in transmission lines carrying the signal to be further processed in the receivers. As its main purposes is to treat the faint signal collected during previous stages by reducing its frequency, and efficiently amplifying it, the receiver must be as close as possible to the feed to conserve the signal by avoiding the power loss due to wave transport.

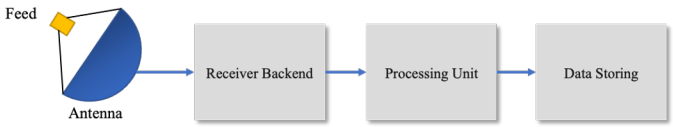


Figure 1.8: Simplified Radio Telescope model. **Source:** Marr et al., 2016

ALMA receivers consist of a cold and a warm cartridge assembly (WCA) (Fig. 1.9). The signal collected by the dish is concentrated using a set of mirrors at the top of the front-end (WCA), to be guided into the receiver. The receiver has a three-stage cooling *cryostat* (80 K, 15 K and 4 K), housing a down-converting system. Is the 4K stage where the input signal is collected by the feedhorn, as the cryogenic temperatures enable the use of superconducting materials and improves the sensitivity by reducing the thermal noise of the instrument. The observed signal has a power of the order of $10^{-15} - 10^{-20}$ watt, meaning that the power received from the background can be many orders of magnitude higher hence the use of different techniques to gain sensitivity and reduce the noise coming from different sources.

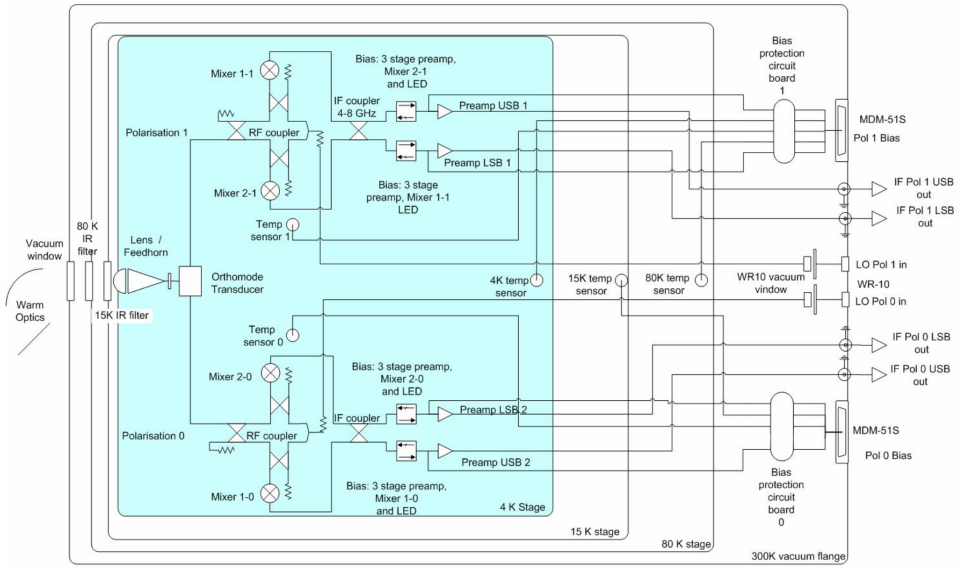


Figure 1.9: Schematic diagram of an ALMA receiver. **Source:** Claude et al., 2008

The following stage is the *down-conversion*, by means of mixers, and the *amplification* which uses low-noise amplifiers (LNA). This particular stage of major interest, and which we will study in the coming sections due to addition of a key component for the image-rejection and down-conversion operation: the *Quadrature Hybrid* (QH). Optimizing the design of this component is the main goal of this work.

1.1.3.1 The Receiver

The *homodyne*, *heterodyne* and *superheterodyne* are the most popular type of receivers in radio astronomy. Despite the architecture of choice, the central function of the front-end is to convert the observed signal into a lower frequency. This operation reduces the losses typical at higher frequencies and, simultaneously, reduces the fabrication difficulty of the waveguides involved in further processing of the signal. Furthermore, at lower frequencies the amplification can take place with higher stability.

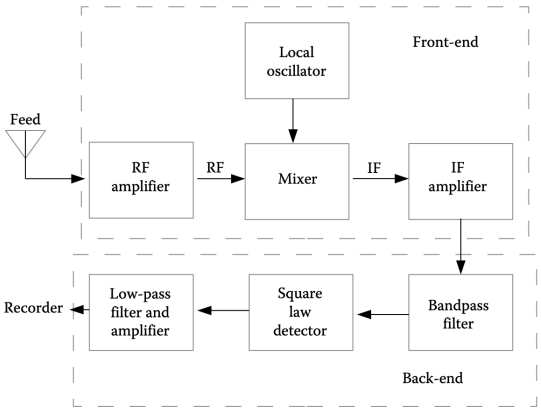


Figure 1.10: Components of a radio telescope's receiver. **Source:** Marr et al., 2016.

While the homodyne receiver directly converts the RF signal to a DC level without any use of intermediate frequencies, heterodyne receivers convert the received signal to a lower intermediate frequency (IF). The term heterodyne implies the combination or mixing of two signals with different frequencies to achieve a frequency shift. In this type of receivers, the *mixer* is the fundamental device to achieve the frequency conversion by multiplying our incoming signal (RF) with a well-characterized signal from a *local oscillator* (LO).

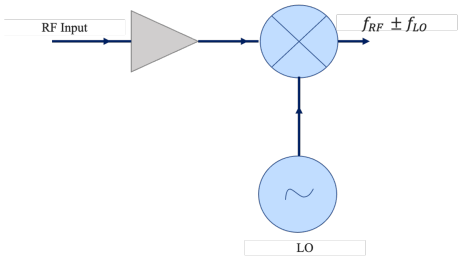


Figure 1.11: Mixing process to produce down-converted signals.

If we assume our RF signal as,

$$v_{RF} = v_r \cos \omega_r t \quad , \tag{1.4}$$

and the LO signal as,

$$v_{LO} = v_0 \cos \omega_0 t \quad , \tag{1.5}$$

The combining operation (Fig. 1.11) can be made by a simple T-junction or a directional coupler (Pozar, 1998). Thus, the result of operating both signals is a intermediate frequency (IF) , $F_{IF} = F_{RF} - F_{LO}$, and a much higher frequency $F_{image} = F_{RF} + F_{LO}$,

$$v_{out} = v_{RF} \times v_{LO} \tag{1.6}$$

$$v_{out} = v_r v_0 \cos\omega_r t \cos\omega_0 t \tag{1.7}$$

$$v_{out} = v_r v_0 \frac{1}{2}\left[\cos\left(\omega_r + \omega_0\right)t + \cos\left(\omega_r - \omega_0\right)t\right] \quad . \tag{1.8}$$

We see that the heterodyning operation yields two signals at two different bands, called the lower side band (LSB) and the upper side band (USB), the latter also known as the *image* (see Fig. 1.12).

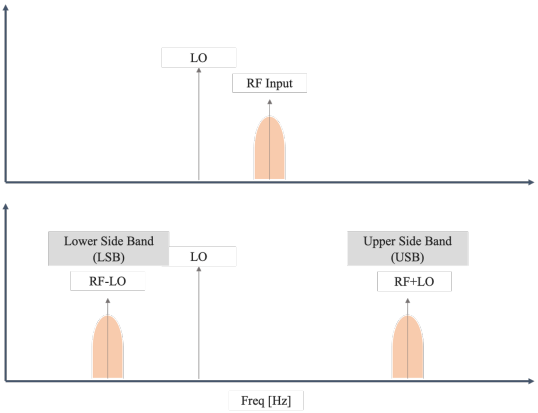


Figure 1.12: Resulting side bands after the mixing process.

From an observational perspective, if the image sideband is located in a part of the spectrum with a higher atmospheric absorption, noise addition to the observation takes place hence degrading the sensitivity (Khudchenko et al., 2017). Some types of mixers, called Double Sideband mixers, or DSB, use passband filters to get rid of the USB. As the down-conversion happens by means of this configuration, the resulting upper and lower side bands get combined at IF frequency hence they are indistinguishable. Moreover, as the upper side band tends to have a signal with a higher noise level associated to the atmospheric absorption at higher frequencies, our observations will degrade in quality (Mena et al., 2011; Finger, 2013).

However, ALMA uses another type of mixers called Sideband Separating (2SB) mixers, which have been shown to improve sensitivity (Hesper et al., 2009). This type of mixer yields pure LSB and USB signals (terminals I_1 and I_2 in Fig. 1.13) without the need of passband filters.

The procedure to achieve this goes as follows. A 90° hybrid (RF Quadrature hybrid) equally divides the incoming RF signal into two ports, one with a 0° phase while the other suffers a 90° phase shift (see Fig. 1.13). Each RF signal is mixed with an LO signal using M_1 and M_2. After the initial mixing, the result is a down-converted IF signal with in-phase and quadrature components. The next step is to further apply a hybrid (IF quadrature hybrid), in which a phase cancellation leads to two signals, a pure LSB and a pure USB signal at terminals I_1 and I_2, respectively. Thus, the good quality of the band separation is determined by that of the IF quadrature hybrid. We will come back to the requirements for this hybrid below.

Let us consider the ALMA Band 3 (84-116 GHz) as an example. One of the requirements for any ALMA receiver cartridge is that the total IF bandwidth is 8 GHz per polarization line (Claude et al., 2008). According to ALMA specifications, the IF bandwidth must be 4-12 GHz. The expected signal (RF) is in the 84-116 GHz band which is then down-converted to a 6 GHz IF (Vassilev et al., 2017).

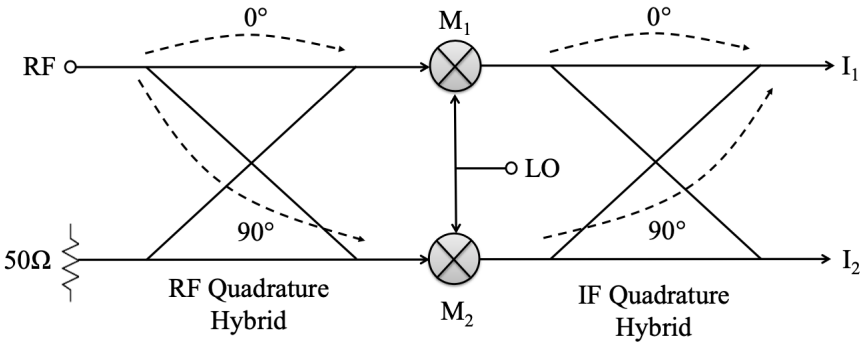


Figure 1.13: Sideband separating mixer approach in which the quadrature power division occurs on the incoming RF signal. **Source:** Finger, 2013

Therefore, the RF/IF quadrature hybrid needs to produce a phase shift of precisely 90° while keeping the output signals as equal in amplitude as possible, for the whole band of operation. This poses a technical challenge from a design and fabrication standpoint, especially when covering a very wide band, as both parameters are antagonist.

A measurement of the quality of this type of quadrature hybrid needs to be related to a very low amplitude and phase imbalance (see Fig 1.14). A proxy for this measurement is given by the Image Rejection, defined as:

$$IR = -10\log\left(\frac{\gamma^2 + 1 + 2\gamma\cos\Delta\phi}{\gamma^2 + 1 - 2\gamma\cos\Delta\phi}\right) \quad . \tag{1.9}$$

Here $\Delta\phi$ (phase imbalance) is the phase deviation from the quadrature, and $\gamma = \Delta A = \left|\frac{Amp_1}{Amp_2}\right|$ is the amplitude ratio, which implies that, as $\gamma = 1$ no amplitude imbalance exists between the resulting signals (Henderson and Cook, 2001). Moreover, a $\Delta\phi = 0°$ and $\gamma = 1$ yields an ideal infinite image rejection. Any deviation from those values reduces the image rejection. To have an IR > 25 dB, it requires that $\Delta A <$1 dB, even more, to further increase the image rejection, it is essential that $0° < \Delta\phi \lesssim 3°$.

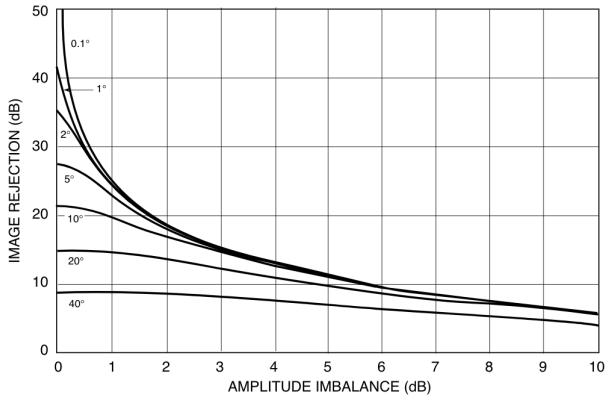


Figure 1.14: Image rejection vs amplitude and phase imbalance. **Source:** Henderson and Cook, 2001.

1.2 Problem statement

In order to attend the scientific goals set for ALMA in its 2030 development roadmap, the improvement in sensitivity is set as a mid-term objective. As mentioned, this can be achieved by getting a better image rejection as a consequence of tighter constraints in the amplitude and phase imbalances for the quadrature hybrids involved in the sideband separating operation.

The success of the band separation process is highly determined by the amplitude imbalance (ΔA) and phase imbalance ($\Delta\phi$), each of which is an antagonist to the other, meaning that improving one parameter, consequently, degrades the other. ALMA has set $\Delta A < 1$ dB and $\Delta\phi \pm 1°$ in order to reach an appropriate image rejection response hence the device must fully comply with these parameters along the full operational band

This project specifically focuses on finding an optimization method for these devices intended to work on ALMA Band 2 (67-90 GHz) and ALMA Band 3 (84-116 GHz) due to the scientific interest in these bands, as well as due to the growing interest in the development of a single receiver cartridge capable of fully covering the extended ALMA 2+3 band (67-116 GHz).

1.3 Objectives

Optimize the geometric design and the electromagnetic behavior for a Quadrature Hybrid (QH) through a machine learning algorithm, applied to the development of a receiver intended to operate in the band 2+3 (67-116 GHz) for ALMA Observatory.

1.3.1 Specific objectives

- Define the most relevant electromagnetic parameters to be used in the optimization process, building the figures of merit and the appropriate cost functions.
- Analyze algorithms to intervene the different geometries and design parameters of the quadrature hybrid.
- Determine the geometric and electromagnetic parameters, including machining constraints, to ensure flat response of the Quadrature Hybrid through the full ALMA 2+3 Band, while reducing the phase and amplitude imbalance.
- Document the algorithm and the code to repeat the optimization process, either on different ALMA receivers or other millimeter/submillimeter observatories, contrasting the optimization and simulation results to those of previous researches.
- Define a Quadrature Hybrid fabrication and characterization plan.

1.4 Thesis Outline

Despite the vast amount of information regarding microwave devices design and transmission line theory, there is a lack of sources that clearly and concisely provide the fundamental knowledge to extensively study quadrature hybrids. Therefore, this thesis is intended to provide specific information on the subject of multi-branch directional couplers, more specifically the quadrature hybrid.

Thus, in the chapters 2 through 4, we develop the theoretical framework to study the multi-branch quadrature hybrid, its physics and the optimization requirements. In these chapters we compile all necessary information to

model waveguides, directional couplers, and multi-branch hybrids, which we expect to be useful for future works.

In order to have a clearer insight of the quadrature hybrid, in chapter 5 we introduce a simulation of a well-characterized classical quadrature hybrid to be used as a reference to state the optimization problem of our interest. Herein we define the optimization constraints involved in the operation of the Quadrature Hybrid, all of which are extensively used in the following chapters.

Thereafter, in chapters 6 and 7, we develop the method to get optimized quadrature hybrid designs that could improve the performance of such devices in the operational ALMA Bands 2 and 3. These chapters include the background information of a) types of optimization algorithms, b) heuristic and meta-heuristic strategies, and c) the Particle Swarm Optimization (PSO) algorithm, its architecture and how it was developed to fulfill the quadrature hybrid optimization needs. More importantly, we present a validation benchmark, consisting of an optimization batch applied to a fully-characterized Quadrature Hybrid designed to operate in ALMA Band 5.

Chapter 8 presents the most relevant results of the optimization process for the Quadrature Hybrid designs. To fully cover the operational bands while improving the amplitude and phase imbalances, a number of geometries were considered during the optimization and simulations namely, 8-,10 and 12-branch geometries. Furthermore, alternatives geometries, which added extra cavities to the main guides, were tested as well.

Finally, after applying our proposed optimization method, two fully-optimized and simulated geometries were selected. As we compare their electromagnetic performance with those of other authors in the field, significant improvements have been found regarding scattering parameters, amplitude and phase imbalances. Moreover, the resulting optimized designs provide feasible designs for ALMA W-Band and Band 3.

2

Physics of the Guiding Structures

- The guiding structures have the purpose of transferring electromagnetic waves in the most efficient manner between to points.
- They are usually characterized by the frequency or range of frequencies they are intended to transport.
- They are studied considering the amount of power they effectively carry or loss between the two points.
- We are looking for solutions of the Maxwell's equations to study the propagation of the electromagnetic wave along the structure.

2.1 Maxwell's Equations

Solving Maxwell's equations while considering boundary conditions is the foundation of the models that describe the different the electromagnetic behavior of the microwave structures and devices to be studied in this work.

$$\boldsymbol{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{2.1}$$

$$\boldsymbol{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \mathcal{J} \tag{2.2}$$

$$\boldsymbol{\nabla} \cdot \vec{D} = \rho \tag{2.3}$$

$$\boldsymbol{\nabla} \cdot \vec{B} = 0 \tag{2.4}$$

Equations (2.1, 2.4) describe the interaction between charges in four vectors: $\vec{B}$ (magnetic flux density), $\vec{E}$ (electric field intensity), $\vec{H}$ (magnetic field intensity) and $\vec{D}$ (electric flux density); $\mathcal{J}$ is the electric current density and ρ is the electric charge density. Though the goal of this section is not to provide a very detailed development and solution of the previous set of equations, some significant results can be highlighted for the sake of the forthcoming discussions.

In addition to the set of equations (2.1, 2.4), we need to consider that the pairs $\vec{E}$, $\vec{D}$ and $\vec{B}$, $\vec{H}$ are connected through the following relations,

$$\vec{D} = \epsilon \vec{E} \tag{2.5}$$

$$\vec{B} = \mu \vec{H} \tag{2.6}$$

where μ and ϵ are correspondingly the permeability and permitivity of the medium; such relations can be expressed in a more general fashion for anisotropic materials, in which case such quantities are dyads.

$$\begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \begin{bmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \tag{2.7}$$

If the medium is a conductor, there is an additional relation namely Ohm's Law,

$$\mathcal{J} = \sigma \vec{E} \tag{2.8}$$

where σ is the conductivity. Altogether, the properties of the medium are described by parameters independent of the fields: μ, ϵ and σ. Finally, there is a fundamental relation to be consider namely the conservation of charge,

$$\boldsymbol{\nabla} \cdot J + \frac{\partial \rho}{\partial t} = 0 \tag{2.9}$$

The preceding description of electromagnetic behavior is based on the assumption of free-space with no presence of particular materials. It is useful to include dielectric bodies to develop the idea of microwave components and to study how electromagnetic fields behave in the presence transitioning interfaces between two materials with different properties.

2.1.1 Fields in different media and boundary conditions

In the study of macroscopic electromagnetic fields, quantities as $\vec{D}$ and $\vec{H}$ have more complicated relations that include electric and magnetic dipoles and quadrupoles (Jackson, 1999) but, in most materials, the higher order terms are neglected, leaving only the terms of the dominant dipole hence producing electric and magnetic polarization, **P** and **M**. For instance, **P** informs us about the permittivity as a complex parameter of which its imaginary part accounts for the losses in the medium. $\vec{D}$ can be defined as,

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P} \tag{2.10}$$

$$\vec{P} = \epsilon_0 \chi_e \vec{E} \tag{2.11}$$

where χ_e is the complex quantity called *electric susceptibility*. By operating 2.10 and 2.11, we obtain an expression for the flux displacement,

$$\vec{D} = \epsilon_0 (1 + \chi_e) \vec{E} \tag{2.12}$$

If we use the same procedure, we arrive to a definition of $\vec{B}$ in terms of magnetic polarization $\mathcal{M}$,

$$\vec{B} = \mu \vec{H} + \mathcal{M} \tag{2.13}$$

$$\vec{B} = \mu_0 (1 + \chi_m) \vec{H} \tag{2.14}$$

This result is meaningful for the study of microwave devices including ferromagnetic materials. Moreover, the set of equations provide a framework to face problems in which materials, other than free-space, are involved of which dielectric properties must be accounted for, in order to understand their electric response in the presence of fields.

Consider now two different materials joint together and both having distinct physical properties hence creating what can be understood as a discontinuity in the medium. It is expected that the electromagnetic waves propagate differently as in the case of free-space thus exhibiting behaviors as transmission, reflection and resonances among others. Then the question arises: how to model such a vast number of possible behaviors while having multiple transitioning interfaces between materials?. Furthermore, considering different geometries and frequency regimes can make the problem extremely complex and almost impossible to be solved analytically.

To answer that question, we can start by stating that vectors $\vec{D}$ and $\vec{E}$ are our objects of interest at the interfaces

between media. The more general case is one in which there is a difference between the fields in the material at each side of the interface. Figure 2.1 depicts how electric and magnetic fields are behaving in each material, for example, the electric field $\vec{E}$ has tangential component in every medium. An additional insight of the fields behavior comes from the Maxwell's equations in their integral form, which results from applying Stoke's and divergence theorems to the equations (2.1 - 2.4) (Pozar, 1998).

$$\oint_S \vec{D} \cdot d\vec{s} = \int_V \rho \, dV \tag{2.15}$$

$$\oint_C \vec{E} \cdot d\vec{l} = -\frac{\partial}{\partial t} \int_S \vec{B} \cdot \vec{ds} - \int_S \vec{M} \cdot \vec{ds} \tag{2.16}$$

Let us see equation 2.15 in the light of figure 2.1. The integration on the closed surfaces yields no contribution from tangential components and with two distinct $\vec{D}$ normal components. At the interface, the integral is reduced to,

$$\hat{n} \cdot (\vec{D_2} - \vec{D_1}) = \rho_s \tag{2.17}$$

meaning that there is a surface charge density associated with the difference between the flux displacement at the material interface. Using similar arguments we can find the tangential components for the electric field,

$$(\vec{E_2} - \vec{E_1}) \times \hat{n} = \vec{M_s} \tag{2.18}$$

$$\hat{n} \times (\vec{H_2} - \vec{H_2}) = \vec{J_s} \tag{2.19}$$

where $\vec{M_s}$ is the magnetic current density and $\vec{J_s}$ is the electric current at the surface.

The results (2.17 - 2.19) can be further simplified when the interface is composed by two lossless dielectrics. As there are no charge or current densities at the surfaces, we can state that the normal component of $\vec{D}$ must satisfy the boundary condition on either side of the interface, and the same applies to $\vec{E}$. On the other hand, as for the magnetic fields, the normal component of $\vec{B}$ and the tangential component of $\vec{H}$ must satisfy boundary conditions at the material interface (Jackson, 1999). Equation 2.20 describes, for instance, the boundary conditions for two materials with different magnetic permeabilities. A detailed development of boundary conditions can be found in Collin (1991).

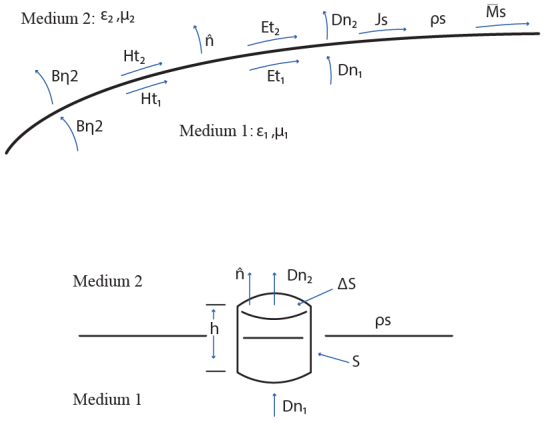


Figure 2.1: Graphical depiction of material interfaces and the conditions for fields at the boundaries.

$$\vec{B_2} \cdot \vec{n} = \vec{B_1} \cdot \vec{n} \quad , \quad \vec{B_2} \times \vec{n} = \frac{\mu_2}{\mu_1} \vec{B_1} \times \vec{n} \tag{2.20}$$

2.1.2 Poynting vector

After assessing the fields and their relations at the boundary, we can introduce the *Poyinting's theorem*. Consider that there is a distribution of charge and current hence we can calculate the rate of doing work by external fields in a volume V as,

$$\int_V \vec{J} \cdot \vec{E}\, d^3x \tag{2.21}$$

On the other hand, electric and magnetic energy densities can be obtained and added to further calculate the total energy density,

$$u = \frac{1}{2}(\vec{E} \cdot \vec{D} + \vec{B} \cdot \vec{H}) \tag{2.22}$$

Equations 2.21 and 2.22 produce the continuity equation (Jackson, 1999),

$$\frac{\partial u}{\partial t} + \nabla \cdot (\vec{E} \times \vec{H}) = -\vec{J} \times \vec{E} \tag{2.23}$$

The equation 2.23 can be seen as formulation of energy conservation in which the quantity $\vec{E} \times \vec{H}$ represents a flow of energy usually notated as **S** and called *Poynting vector*. An additional expression can be obtained as we consider periodic fields as complex quantities, in which case we obtain a complex Poynting vector,

$$\vec{S} = \frac{1}{2}\vec{E} \times \vec{H}^* \tag{2.24}$$

This vector can be understood as an average power flow through the surface, which can be obtained from the real part of the normal component of $\vec{S}$. If you are interested on a more lengthy and detailed derivation of the Poyting vector, you can find Jordan and Balmain (1968) very useful.

An important consequence stems from use of this vector: there is an energy flow that can be studied in the fields rather than in the currents, furthermore, it is thereby implied that there is a physical energy transfer mechanism or physical transmission line.

The question arises about how to study the amount of energy contained in an electromagnetic field. Though it is clear that there must be a relation between *that* energy and mechanical work, mainly due to the interaction and motion of charges, the discussion can be limited to think in terms of energy *stored* in the fields flowing at a given rate. This idea is of major relevance as it becomes the ground to introduce the analysis of fields in the form of voltages, currents and, finally as lumped circuits.

2.1.3 Maxwell's equations and guiding structures

To approach the phenomena of energy transport, the initial step must be that of solving the set of Maxwell's equations (2.1, 2.4). In the appendix A.1, the reader can find a detailed description of how to solve those equations by means of phasor forms and Hemlhotz equations. Some important results are the expressions for the E and H field in terms of traveling waves,

$$E_x = E^+ e^{-jkz} + E^- e^{jkz} \quad , \tag{2.25}$$

and

$$H_y = \frac{1}{\eta}[E^+ e^{-jkz} + E^- e^{jkz}] \quad . \tag{2.26}$$

where $\eta = \omega\mu/k = \sqrt{\mu/\epsilon}$, is the wave impedance or the *intrinsic impedance* of the medium. Likewise, we find that $\vec{E}$ and $\vec{H}$ are orthogonal to the direction of propagation and to each other (see Fig. 2.2).

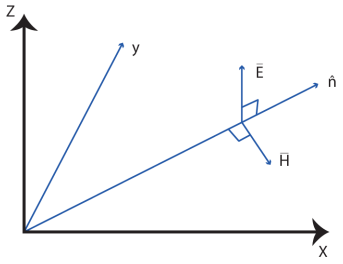


Figure 2.2: A generalized view of the direction of propagation and the orthogonal fields.

The notion of *transport of energy in electromagnetic fields in the form of waves*, can be extended by thinking about electromagnetic fields confined within a closed surface. Consider a perfect conductor namely a material with infinite σ thus all fields must be zero inside the conductor material. We can find that 2.18 reduces to,

$$\hat{n} \times \vec{E} = 0 \tag{2.27}$$

which means that the tangential components are zero or very small due to *skin depth* effect. Furthermore, 2.27 is accompanied with additional boundary conditions like $\hat{n} \times \vec{H} = \vec{J}$ and $\hat{n} \cdot \vec{D} = \rho_s$ (Pozar, 1998); the type of boundary complying with the mentioned conditions is frequently known as *electric wall*. Such boundaries enable the electromagnetic wave to travel inside a closed surface made of conducting material as the energy can only flow or "escape" through the walls in very small amounts, in an event more precisely know as a *loss*. Additionally, there are also *losses* related to the surface currents in the walls, which finally manifests itself as heat.

In order discuss to a greater extent the transport of energy through cavities made of conductors, let us first approach the types of propagation behaviors that an electromagnetic wave could have.

2.1.3.1 Wave Modes

We have seen that traveling electromagnetic waves, in the case of planar waves, have orthogonal components to the direction of propagation. A more broad depiction is that of waves having more components, not just in a perpendicular plane but in a arbitrary plane relative to the traveling direction, manifesting itself as polarization, reflections and even as constructive or destructive behaviors.

Imposing a preferred direction of propagation for the equations A.11 and A.12, result in a solution to Maxwell's equations in which there is no component along the traveling axis, usually known as *Transverse-electromagnetic* or **TEM**. Two additional conditions occur regarding the components along the propagation axis . There is a case in which the electric field has component along the traveling axis while the magnetic fields does not; this situation is known as *Transverse Magnetic* or **TM**. A second case, known as *Transverse Electric* or **TE**, is the one in which the magnetic field has component along the traveling axis while the electric field remains in the transverse plane hence does not have a component along the propagation axis.

Maxwell's equations can be solved for TEM waves, offering further insight about the particular case. Let us first consider TEM waves as *uniform* given that there is no change in amplitude in the surface along the z axis; a second consideration is that E_z and H_z are zero given the definition of TEM, and consequently $\partial/\partial x$ and

By applying similar procedures we can derive mathematical descriptions for non-uniform TEM waves, but now the derivatives $\partial/\partial x$ and $\partial/\partial y$ are not zero (Montgomery et al., 1987). Authors like Jackson and Montgomery et al. provide a deeper study of geometries like coaxial conductor, parallel plates and spherical conductors, and how the wave modes are mathematically described hence different ways to calculate impedance and fields.

Let us now consider the TE waves namely E field is fully transverse or equivalently, $E_z = 0$. With that condition at hand and by expanding the rotational and divergence equations from phasor form of Maxwell's equations, we get,

$$\frac{\partial E_y}{\partial z} = j\omega\mu H_x \quad , \quad \frac{\partial E_x}{\partial z} = -j\omega\mu H_y \tag{2.37}$$

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = j\omega\epsilon E_x \quad , \quad \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = j\omega\epsilon E_y \tag{2.38}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu H_z \tag{2.39}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = 0 \tag{2.40}$$

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0 \tag{2.41}$$

$$\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} + \frac{\partial H_z}{\partial z} = 0 \tag{2.42}$$

This set of equations can be used to get a wave equation for every field component, in this case we are going to solve for H_z. If we assume that every field component varies with z like $e^{-\gamma z}$, we can plug solutions of the type $E_{coordinate}(z) = E_{coordinate}e^{-\gamma z}$ into 2.37 to get,

$$-\gamma E_y = j\omega\mu H_x \quad , \quad -\gamma E_x = j\omega\mu H_y \tag{2.43}$$

We can further substitute this results into 2.39 and 2.42 and solve for H_z,

$$\frac{\partial^2 H_z}{\partial x^2} + \frac{\partial^2 H_z}{\partial y^2} + \left(\gamma^2 + \omega^2\epsilon\mu\right) H_z = 0 \tag{2.44}$$

Once a solution is found for the equation 2.44 of the type A.14 and A.15, we can get the other components namely H_x, H_y, E_x and E_y, by using the whole set of equations 2.37 - 2.42. This means that all fields can be obtained from a single quantity, H_z in this case.

Alternatively to 2.44, we can solve for E_z hence we get a wave equation,

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \left(\gamma^2 + \omega^2\epsilon\mu\right) E_z = 0 \tag{2.45}$$

A major conclusion from this section is that a rectangular waveguide, which has all walls made of the same conducting material, cannot support the TEM propagation mode, TE and TM are the only allowed modes in such a structure. This is a consequence of the fully transverse layout of the fields, one in which none of them have a component in the direction of propagation, moreover, there are walls in the x and y direction, which necessarily implies due to the boundary conditions, that both fields must be zero at some point along the waveguide. If you are interested about diving deeper into other guiding structures, parallel plates for instance, and solving Maxwell's equations for TE and TM wave, please refer to Ellingson (2020) or Jordan and Balmain (1968).

2.1.3.2 Waveguides

In the beginning, the two-wire waveguides, parallel plates for instance, were considered the best way for power transfer until the development of closed hollow tubes and the prove of TM/TE modes propagating in them, furthermore, it was proved that they have higher power-carrying capacity and lower losses (Pozar, 1998). But, the problem remained on how to model the electromagnetic behavior of such devices.

It is important to keep in mind that solving Maxwell's equations and studying the types of traveling waves that result from those solutions, are important to provide a tool to evaluate how energy is transported along an enclosed surface containing the fields. In order to keep a simple but clear development of this work, we are limiting ourselves to the exploration of symmetric guiding structures with uniform cross-section, in particular the *rectangular waveguide*.

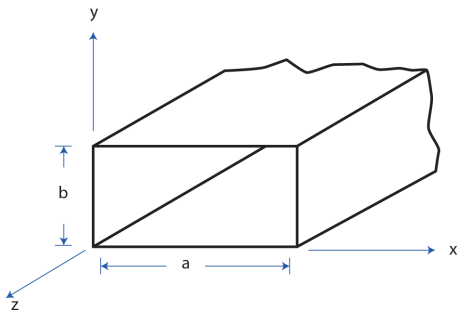


Figure 2.4: Rectangular cross-section for a waveguide.

Let us consider a guiding structure as showed in the figure 2.4, which has perfectly conducting walls. To study this type of waveguide we can use TE and TM waves modes. A comprehensive derivation of the equations can be found in Jackson and Montgomery et al.; here we will describe the process as well as the major findings necessary to analyze microwave devices.

To find the relations that govern the electromagnetic waves inside a rectangular waveguide, we have at hand two approaches. The first treats the problem by dividing the fields in two different components namely *tranverse* and *longitudinal*, while the second relies on solving the wave equation 2.44, depending on the type of solutions we are looking for, particularly TE or TM waves.

A particular case for TE waves can be found in A.1, where a solution to 2.44 is provided. In order to continue developing the theoretical framework to model guiding structures as transmission lines, we can highlight some major results.

In the first place, the parameter γ called the *phase propagation constant* is expressed in terms of the rectangular waveguide dimensions,

$$\gamma = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 - \omega^2\epsilon\mu} \quad . \tag{2.46}$$

Secondly, there is a critical value, usually known as **cut-off frequency**, which defines a real- or imaginary-valued γ hence determining the wave propagation (see A.1). This value is the one that satisfies,

$$\omega_c = \sqrt{\frac{1}{\epsilon\mu}\left[\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\right]} \quad . \tag{2.47}$$

This value ω depends on the geometric size of the cross-section of the waveguide and, simultaneously, depends

on the values **m** and **n**, thereby different cut-off frequencies can be found for "higher" modes, namely higher values of m and n.

Further meaningful relations were found, such as the *wavelength of the guide*,

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2}} \tag{2.48}$$

$$\lambda_g = \frac{\lambda_0}{\sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0} - \left(\frac{\lambda_0}{\lambda_c}\right)^2}} \tag{2.49}$$

As values are assigned to m and n, specific modes arise and expressions for ω_c and λ_c are found, particularly for TE_{mn} with m=1 and n=0 namely TE_{10}-Mode.

$$\omega_c = \frac{\pi}{\epsilon\mu a} \tag{2.50}$$

$$\lambda_c = 2a. \tag{2.51}$$

We can conclude that the cut-off frequency corresponds to the one in which half wavelength is equal to the width a of the waveguide; interesting enough is the fact that, as for TE_{10} mode, the dimension b has no relevance. For practical purposes, the last result in which we get expressions for TE_{10} mode, is of major significance for rectangular waveguides, as this mode is considered the dominant in such type of structures.

3

Transmission lines and Microwave circuits

Radio astronomy applications highly benefits from the developments herein explained in particular regarding the design of devices intended to work in different regimes with the highest standards in terms of sensitivity. To reach a transmission line formalism, a definition of impedance along a waveguide required as it is one of the key parameters allowing the development of transmission line theory. Each case needs to be addressed by its own, whether we are studying TEM, TM, or TE waves; expressions are needed for fields and impedance in each particular case such as parallel plates, cylindrical waveguides as well as for guides with conductors within namely coaxial lines, among others.

3.1 Impedance in waveguides and the Transmission Line formalism

After our study on how to solve Maxwell's equations and how waves propagate in a medium hence how electromagnetic waves behave in a guiding structure such as rectangular waveguides, the question arises about how we can further model more complex structures. The propagation along a waveguide made of conductor material has been modeled so far but, what if structures that have discontinuities?, or cavities are being filled with dielectric materials?. It is required to use alternative approaches to simplify the analysis for such intricate cases.

Let us recall from equation 2.36 the well known *wave impedance*, which tell us about how the medium impedes or prevents the electromagnetic wave to propagate. The impedance is a complex number in function of the frequency which is not a minor feature as it tells us how the material behaves at higher frequencies with respect to the conductivity σ. The impedance can be reduce in an ideal medium to $\mathcal{Z} = \sqrt{\mu/\epsilon}$. This presents us a major attribute of the complex wave impedance: its dependency on medium's constants.

Why is this relevant for our understanding of the transmission line concept?. The power transmission inside a guide relates to the characteristics of the materials thus to the impedance, more specifically, a field is established in the presence of certain conducting properties of the medium. Additionally, fields depend on the shape of the guide.

The parallel-plates case provides an example of a guiding structure and how E and H fields are related. As we study how waves evolve in parallel plates we expect to have different field values than while working with rectangular waveguides (see appendix A.1) and we obtain expression relating fields like,

$$\left|\frac{E_y}{H_x}\right| = \frac{\omega\mu}{\beta} = \frac{\omega\mu}{\sqrt{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2}} \quad . \tag{3.1}$$

This sort of relation represent the impedance in the $\hat{z}$ direction, being the direction of propagation for TE waves.

On the other hand, as we solve for the TEM-waves case, the resulting impedance has the form,

$$\left|\frac{E_x}{H_y}\right| = \frac{\beta}{\omega\epsilon} = \sqrt{\frac{\mu}{\epsilon}} \quad , \tag{3.2}$$

which was introduced earlier as the impedance of the medium $\mathcal{Z}$ and, in this case, also known as intrinsic impedance $\boldsymbol{\eta}$. This parameter can be traced back to a preliminary discussion about solving the *Helmholtz equation* and its representation in equation A.15.

The significance of these findings relies on the fact that models depicting fields and transmitted power in terms of voltages and currents stem from the behavior of impedance along the waveguide, in other words, knowing the impedance is the base on which we build a circuit representation for waveguides and, in a wider extent, for transmission lines.

We have explored so far a procedure to calculate impedance for waves propagating with parallel plates as guiding structure. Now, it is valuable to inspect how a rectangular cross-section waveguide transport waves and in particular the behavior of impedance.

As for rectangular waveguides, let us recall the result 2.44 and 2.45 from which we can obtain solutions for E_z and H_z. In addition to that, from the set equations 2.37 we can get expressions for fields in the guide in terms of E_z and H_z,

$$H_x = -\frac{\gamma}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial x} + j\frac{\omega\epsilon}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial E_z}{\partial y} \tag{3.3}$$

$$H_y = -\frac{\gamma}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial y} - j\frac{\omega\epsilon}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial E_z}{\partial x} \tag{3.4}$$

$$E_x = -\frac{\gamma}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial E_z}{\partial x} - j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial y} \tag{3.5}$$

$$E_y = -\frac{\gamma}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial E_z}{\partial y} + j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial x} \tag{3.6}$$

These expressions agree with the general case in which whether $E_z = 0$ or $H_z = 0$ are not considered. Just recall that the case studied in section 2.1.3.1 corresponds to $E_z = 0$. At this point is evident that either E_z or H_z is required in a rectangular waveguide for transmission to occur.
In section A.2 you can find a concise description of the procedure to solve equation 2.45 for TE waves in a rectangular guide. The final solution has the form,

$$H_z = C\cos\Big(\frac{m\pi}{a}x\Big)\cos\Big(\frac{n\pi}{b}y\Big) \quad . \tag{3.7}$$

This result jointly with equations (3.3 - 3.6) can be used to get all fields in a rectangular guide to TE waves,

$$H_x = \frac{j\beta}{\gamma^2+\omega^2\mu\epsilon}C\frac{m\pi}{a}\sin\Big(\frac{m\pi}{a}x\Big)\cos\Big(\frac{n\pi}{b}y\Big) \tag{3.8}$$

$$H_y = \frac{j\beta}{\gamma^2+\omega^2\mu\epsilon}C\frac{n\pi}{b}\cos\Big(\frac{m\pi}{a}x\Big)\sin\Big(\frac{n\pi}{b}y\Big) \tag{3.9}$$

$$E_x = \frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon}C\frac{n\pi}{b}\cos\Big(\frac{m\pi}{a}x\Big)\sin\Big(\frac{n\pi}{b}y\Big) \tag{3.10}$$

$$E_y = -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon}C\frac{m\pi}{a}\sin\Big(\frac{m\pi}{a}x\Big)\cos \quad , \tag{3.11}$$

where the presence of j simply means a phase difference between the components and H_z in the direction of propagation.

Moving one step forward, an impedance from (3.8 - 3.11) can be obtained,

$$\left|\frac{E_x}{H_y}\right| = \frac{\omega\mu}{\beta} = \frac{\omega\mu}{\sqrt{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2}} \quad . \tag{3.12}$$

As a final step, the set of equations can be further simplified by considering a particular mode namely $m = 1$ and $n = 0$, leaving us with $E_x = H_y = 0$ what finally yields the impedance for TE_{10} mode in a rectangular waveguide,

$$\left|\frac{E_y}{H_x}\right| = \frac{\omega\mu}{\beta} = \frac{\omega\mu}{\sqrt{\omega^2\mu\epsilon - \left(\frac{\pi}{a}\right)^2}} \quad . \tag{3.13}$$

This final result can also be obtained from equation 2.43, as you can easily check. For the reader we leave to check the case of TM waves, which can be solved in the same way as done before to find,

$$\left|\frac{E_x}{H_y}\right| = \frac{\beta}{\omega\epsilon} = \frac{\sqrt{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2}}{\omega\epsilon} \quad . \tag{3.14}$$

As final remarks, here are some important insights to account for:

- From 3.12 and 3.13 we can see how the *order* in which fields are operated indicates if we are looking impedance forward or backwards in the waveguide, but ultimately the value is the same. An additional observation is that, we can obtain a *total* impedance by considering all transverse components from (3.8 - 3.11),

$$\mathcal{Z} = \frac{E_{transverse}}{H_{transverse}} = \frac{\sqrt{E_x^2 + E_y^2}}{\sqrt{H_x^2 + H_y^2}} \tag{3.15}$$

- We can rewrite β by recalling equation A.36,

$$\beta = \omega\sqrt{\mu\epsilon}\sqrt{1 - (\omega_c^2/\omega^2)} \tag{3.16}$$

  which can be replaced in our expressions for TE mode impedance to get,

$$\mathcal{Z}_{TE} = \frac{\omega\mu}{\beta} = \frac{\eta}{\sqrt{1 - (\omega_c^2/\omega^2)}} \tag{3.17}$$

- By applying the procedures herein discussed we have formulas for TEM (parallel plates only), TE and TM modes in a rectangular waveguide, which also happens to apply for cylindrical guides, moreover, the expression just depend on the intrinsic impedance of the dielectric material and the ratio between ω and ω_c,

$$\mathcal{Z}_{TM} = \eta\sqrt{1 - (\omega_c^2/\omega^2)} \tag{3.18}$$

To grasp the meaning behind a **transmission line**, we return to the idea of *guiding structure* we have studied. There are three possible propagation modes namely TEM, TE and TM, and by combining them we can fully represent the propagation of a wave in a guiding structure. In addition to that, the geometry influences which modes are supported. For instance, in a rectangular cavity no TEM modes are allowed, in contrast to the parallel plate structure, which favors TEM propagation mode. In practice, a transmission line is defined as a system of two conductors that support TEM modes, and a waveguide is define as a system involving one enclosed conductor that supports the other extra propagation modes (Jordan and Balmain, 1968).

We can now move forward to define a representation for *transmission line* in the case of non-TEM waves and later we will include the circuit model to support our analysis. There is an issue to be addressed in the par-

ticular case of rectangular waveguides: how to relate $\vec{E}$ and $\vec{H}$ fields to $\vec{V}$ (voltage) and $\vec{I}$ (current)?, as they are the appropriate parameters to work in circuit representations. When referring to TEM waves it could seem something straightforward thinking about potentials between parallel plates or parallel wires, but as for TE and TM waves in rectangular waveguides, relations should be different.

A fundamental relation must be established between electric parameters and fields. It is easily remembered that there exists such relation in integral form,

$$\vec{V_{ab}} = \int_a^b \vec{E} \cdot d\vec{r} \quad . \tag{3.19}$$

Moreover, if the propagation is in $\hat{z}$ thus the voltage across a rectangular section is,

$$V_{ab} = \int_0^b E_y dy \tag{3.20}$$

while in order to find the current, Ampere's law allows to find I as,

$$I = \oint_c \vec{H}\, d\vec{l} \tag{3.21}$$

where the integral is done over the conductor's closed contour C.

From these relations one can obtain voltages and currents from the fields along a waveguide despite its geometry; it is valid whether you are working with TEM waves transported by parallel plates or coaxial lines, or while working with hollow waveguides with different geometries, what really matters is that there is a solution for Maxwell's equations which can be found by the procedures described in previous sections.

From these equations and by recurring to the customary circuit theory, we can obtain the impedance,

$$Z_0 = \frac{\vec{V}}{\vec{I}} \tag{3.22}$$

which is known as the characteristic impedance of the transmission line.

Let us study the TE_{10} case to find voltage and current. Consider the set of equation (3.8 - 3.11) representing all transversal components of both fields namely those laying in the plane perpendicular to the propagation direction. To find the voltage we can integrate E_y from 3.20,

$$\begin{aligned} V &= \int_0^b -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon} C \frac{m\pi}{a} \sin\Big(\frac{m\pi}{a}x\Big) \cos\Big(\frac{n\pi}{b}y\Big)\, dy \\ &= -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon} C \frac{m\pi}{a} \sin\Big(\frac{m\pi}{a}x\Big) \int_0^b \cos\Big(\frac{n\pi}{b}y\Big)\, dy \end{aligned} \tag{3.23}$$

and by replacing $m = 1$ and $n = 0$,

$$\begin{aligned} V &= -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon} C \frac{m\pi}{a} \sin\Big(\frac{\pi}{a}x\Big) \int_0^b dy \\ &= -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon} C \frac{m\pi b}{a} \sin\Big(\frac{\pi}{a}x\Big) \end{aligned} \tag{3.24}$$

The maximum voltage occurs at $x = a/2$ thus,

$$V = -\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon} C \frac{m\pi b}{a} \tag{3.25}$$

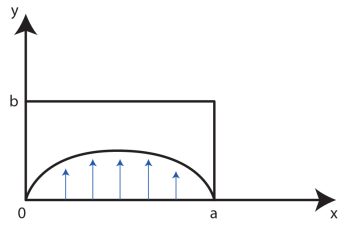


Figure 3.1: Electric field TE_{10} mode in a rectangular waveguide.

On the other hand, the current can be calculated from magnetic field. Continuing with our $\hat{z}$ propagation vector, thus we have E_y and $\vec{H} = -H_x$ furthermore, there is a surface current in $\hat{z}$, J_z, which is related to the magnetic field by,

$$J = \hat{n} \times H = -H_x \tag{3.26}$$

with $\hat{n}$ a unit vector normal to the conducting surface.

The longitudinal current for the lower conducting surface namely $b = 0$, is obtained from

$$I = \int_0^a J_z \, dx \tag{3.27}$$

which can be solved applying the same arguments as for the voltage using equations (3.8 - 3.11) and making $m = 1$ and $n = 0$,

$$I = -2\frac{j\beta}{\gamma^2 + \omega^2\mu\epsilon}C \tag{3.28}$$

If we use this result as well as 3.22 and 3.25, an expression can be obtained for the impedance,

$$Z_0 = \frac{\omega\mu}{\beta}\frac{\pi b}{2a} \tag{3.29}$$

which is known as *integrated impedance*, in this case for TE_{10} waves. This result can be further extended by realizing that it is related to the wave impedance 3.17 hence,

$$\begin{aligned} Z_0 &= Z_{TE}\frac{\pi b}{2a} \\ &= \frac{\pi b}{2a}\frac{\eta}{\sqrt{1 - (\omega_c^2/\omega^2)}} \end{aligned} \tag{3.30}$$

recalling that η is the intrinsic impedance of the material filling the waveguide.

The former results namely the **integrated impedance for TE_{10} waves in rectangular guides**, are of considerable importance as we count for a procedure connecting fields with circuit parameters (voltages and currents) inside a waveguide, moreover, this findings can be extended to higher order modes or geometries with arbitrary cross-sections. For the reader interested in further developments of this concepts, may find of help Jordan and Balmain (1968), Pozar (1998) chapter 3 as well as Marcuvitz (1986).

The waveguide can be studied by means of an equivalent model which contains a characteristic impedance, a model in which voltages and currents can be used to solve waveguide problems viz. the transmission line paradigm.

3.1.1 Circuit representation

The interest behind the developments presented so far, is to collect a set of instruments that could be used for the design of more devices derived from waveguides, for instance couplers, attenuators and isolators, among others, all of which are commonly used in instruments such as the receivers in radio telescopes.

In order to properly describe complex geometries and how electromagnetic waves behaves while transported along them, exist two alternatives namely the rigorous solution of Maxwell's equations and thinking of intricate devices as a mixture of more simpler ones. The latter has become the common practice as we can have a collection of *canonical* geometries carefully studied, and from them, we can compound more intricate devices. Montgomery et al. (1987) suggest that using a more abbreviated model of transmission lines described as inductance, capacitance and resistance, can vastly simplify the analysis instead of solving the electromagnetic field equations. The difficulties arise when connecting different waveguide sections or when elements are added inside the waveguides resulting in the loss of uniformity, hence creating a whole new problem to solve. Consider the case of connecting two rectangular waveguides by means of a third guide usually named **branch**; the electromagnetic analysis of that structure may be extremely difficult if trying to solve the Maxwell's equations, but on the other hand, the whole complexity is reduced when considering that we already have equations to explain how each *sub-waveguides* behaves.

After deriving impedance expressions for different guiding structures and wave propagation modes, they can be used to represent a guiding structure as circuit components (resistance, capacitance, inductance), which are *distributed* continuously along the structure.

The question arises on which voltage, current or impedance to use for an specific waveguide as we could have as many as modes are available, and just to make things worse, voltage is different along the integration line. For instance, recall that we assumed a value for x to obtain 3.25. Three criteria are stated in order to settle on the circuit parameters we use to describe a rectangular waveguide (Pozar, 1998):

- Voltage and current are selected for and specific mode, one in which voltage is proportional to the transverse component of the electric field.
- The product of equivalent voltages and currents must yield the power flow of the selected mode.
- The ratio between voltage and current namely the impedance, must be equal to the characteristic impedance of the line.

The perceptive reader may be aware that we stick to these guidelines during this section while finding expressions for voltage and current in rectangular waveguides. Though there is enough information to help us in definition of circuit description of a waveguide, something is missing namely how to properly represent impedance as resistance and reactance.

Let us recall the set of equations (2.37 - 2.42) which provides the description of TE waves in rectangular coordinates. Our interest is placed on the equations that relate the transverse components we have dealing with namely E_y and H_x,

$$\begin{aligned} \frac{\partial E_y}{\partial z} &= j\omega\mu H_x \\ \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} &= j\omega\epsilon E_x \end{aligned} \tag{3.31}$$

As for TE waves, there is no component of the electric field in $\hat{z}$ what necessarily means that $\boldsymbol{\nabla} \times H{=}0$ in the perpendicular $x-y$ plane. This condition allows us to conclude that in the $x-y$ plane a *magnetic scalar*

potential ψ can be defined. This potential can be used to express the magnetic field as the gradient,

$$\vec{H} = -\vec{\nabla}\psi \tag{3.32}$$

$$H_x = -\frac{\partial\psi}{\partial x} \quad , \quad H_y = -\frac{\partial\psi}{\partial y} \tag{3.33}$$

We already have wave relations for the transverse components in terms of H_z and E_z; in particular we can use the equation 3.6 remembering that $E_z = 0$, and to replace it in 3.31,

$$\frac{\partial}{\partial z}\left(j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial x}\right) = -j\omega\mu\frac{\partial\psi}{\partial x} \tag{3.34}$$

$$\frac{\partial H_z}{\partial y} - \frac{\partial}{\partial z}\left(-\frac{\partial\psi}{\partial y}\right) = j\omega\epsilon\left(-j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial y}\right) \tag{3.35}$$

which can be further simplified by operating the partial derivatives to get expressions for relating H_z and ψ. The equation 3.34 becomes,

$$\begin{aligned}\frac{\partial^2}{\partial z\partial x}\left(j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}H_z\right) &= -j\omega\mu\frac{\partial\psi}{\partial x}\\ \frac{\partial}{\partial z}\left(j\frac{\omega\mu}{\gamma^2+\omega^2\mu\epsilon}H_z\right) &= -j\omega\mu\psi\end{aligned} \tag{3.36}$$

and equation 3.35 yields,

$$\begin{aligned}\frac{\partial}{\partial y}\left(H_z - \frac{\partial}{\partial z}(-\psi)\right) &= \left(\frac{\omega^2\mu\epsilon}{\gamma^2+\omega^2\mu\epsilon}\frac{\partial H_z}{\partial y}\right)\\ H_z + \frac{\partial}{\partial z}(\psi) &= \left(\frac{\omega^2\mu\epsilon}{\gamma^2+\omega^2\mu\epsilon}H_z\right)\\ \frac{\partial}{\partial z}(\psi) &= -\left(\frac{\gamma^2+\omega^2\mu\epsilon}{j\omega\mu}+j\omega\epsilon\right)\left(\frac{j\omega\mu}{\gamma^2+\omega^2\mu\epsilon}H_z\right)\end{aligned} \tag{3.37}$$

The first observation from 3.36 and 3.37 is the tight relation between the fields and the direction of propagation. It is expected that the field changes along the waveguide giving a first taste of the importance of the *length*, which is coherent with the previously exposed concept of *distributed* characteristic impedance along the guiding structure.

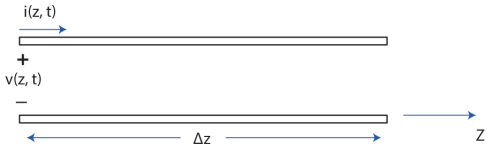
This result helps us to link circuit transmission line theory and electromagnetic waves in rectangular waveguides. But let us keep this result aside for a moment to enter into the transmission line modeling; we will comeback to it later.

3.1.2 Two-wire transmission line

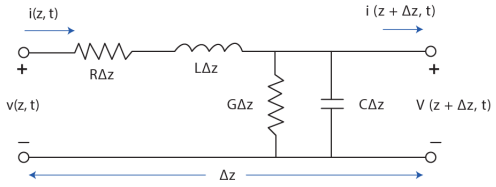
The key parameter in a transmission line is the length of the structure, furthermore, all other electric parameters are defined in terms of the length, as depicted in the Fig.3.2. The easiest approach to the transmission line model is the two-wire, moreover, it is a straight forward schema for guiding structures consisting of two conductors namely two parallel wires, two parallel plates or coaxial lines.

The physical variables involved in the wave transport can be represented as circuit elements. The conductors have a resistance R associated to their finite conductivity causing a voltage drop along the conductor, as well as a complex response in the frequency domain similar to a series inductance L. Think about what we have seen so far regarding the complex characteristic of the impedance. A second set of parameters are associated

to the presence of a dielectric medium which are represented as shunt elements in the circuit. The proximity between the conductors needs to be accounted in the form of a capacitance C which changes at the same time as the dielectric or the geometry of the conductors are changed. Finally, we must account for the losses in the material between the to conductors as a resistance, but in order to simplify the calculations, the reciprocal is normally used which is represented as a conductance G. Let us remember that conductance is the reciprocal of the resistance, but with circuit analysis in mind, it is easier to work with conductance with elements in the frequency domain. If the cross-section of the guide changes along the propagation axis, all these quantities will be functions of the position instead of constant values.



(a) Two-wire transmission line model.



(b) Equivalent circuit for a length-dependent transmission line.

Figure 3.2: Two-wire transmission line in terms of incremental length.

The most general impedance could be represented as,

$$Z(j\omega) = -\frac{j}{\omega C} + R + j\omega L \tag{3.38}$$

which includes a *resistive* and a *reactive* components. The latter are usually called *Reactance* (X).

Another important element is the *admittance* Y, which is used to represent the shunt elements,

$$Y(j\omega) = G + jB \tag{3.39}$$

with B regarded as the element that represents all "imaginary" elements.

Another basic consideration is the fact that when splitting the admittance Y in different elements, they should be combined in parallel. For instance, if we have components in parallel, the admittance is obtained as,

$$Y(j\omega) = \frac{1}{R} + j\omega C - \frac{j}{\omega L} \tag{3.40}$$

A third concept to remember is that *Kirchhoff's* voltages and current laws are still valid for circuit analysis. The voltage law states that the algebraic sum of the voltages around a circuit must be zero, while the current law states that such algebraic sum must be zero when performed on the currents flowing into each branch. The appropriate technique commonly used to set up the equation from the circuit using *Kirchhoff's* laws is the *mesh*

currents and voltages, which is something you can easily find in most circuit theory textbooks. The relevance of these concepts in our work is that by using them you get into a system of equations that describe the circuit network in terms of voltages, currents and impedances,

$$\begin{bmatrix} v_1 \\ v_2 \\ ... \\ v_n \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} & ... & Z_{1n} \\ Z_{21} & Z_{22} & ... & Z_{2n} \\ ... & ... & ... & ... \\ Z_{n1} & Z_{n2} & ... & Z_{nn} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ ... \\ i_n \end{bmatrix} \tag{3.41}$$

There is an additional parameter to consider namely ΔZ, which describes the length of the guiding structure, see Fig. 3.2. The distributed impedance depends on the resistive and reactive parameters calculated from the length of the waveguide hence the voltage $V(z,t)$ and currents $I(z,t)$, are also length-dependent. By taking a differential length in $\hat{z}$, we can derive a set of equations by means of *Kirchhoff's* laws,

$$v(z,t) - R\Delta z i(z,t) - L\Delta z \frac{\partial i(z,t)}{\partial t} - v(z+\Delta z,t) = 0 \tag{3.42}$$

$$i(z,t) - G\Delta z v(z+\Delta z,t) - C\Delta z \frac{\partial v(z+\Delta z,t)}{\partial t} - i(z+\Delta z,t) = 0 \tag{3.43}$$

which can be expressed in differential form by dividing by Δz and taking the limit $\Delta z \to 0$ (Pozar, 1998),

$$\frac{\partial v(z,t)}{dt} = -Ri(z,t) - L\frac{\partial i(z,t)}{\partial t} \tag{3.44}$$

$$\frac{\partial i(z,t)}{dt} = -Gv(z,t) - C\frac{\partial v(z,t)}{\partial t} \tag{3.45}$$

These equations can be expressed in phasor form, commonly known as *telegraphists' equations*,

$$-\frac{dV}{dz} = (R + j\omega L)I \tag{3.46}$$

$$-\frac{dI}{dz} = (G + j\omega C)V \tag{3.47}$$

Both equations can be used to find a solution for V or I,

$$\frac{d^2V}{dz^2} + (R + j\omega L)(G + j\omega C)V = 0 \tag{3.48}$$

$$\frac{d^2I}{dz^2} + (R + j\omega z)(G + j\omega C)I = 0 \tag{3.49}$$

which has solutions of the type,

$$V(z) = V_o^+ e^{-\gamma z} + V_o^- e^{\gamma z} \tag{3.50}$$

$$I(z) = I_o^+ e^{-\gamma z} + I_o^- e^{\gamma z} \tag{3.51}$$

and the propagation constant,

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \alpha + j\beta \tag{3.52}$$

Both functions 3.50 and 3.51 describe traveling waves propagating in the positive or negative direction along $\hat{z}$. This final result can be further assessed in lossless and lossy cases namely R and G are either zero or greater than zero.

There is a major conclusion that stems from the equation 2.32 and our last result in equation 3.52. As you may recall, while working with fields, γ has a complex value with real and imaginary components, in terms of parameters describing the physics of the propagation process. Herein we have the same constant in terms of passive circuit parameters. The reader may also recall our finding for γ in equation A.35 for the rectangular

waveguide. As mentioned before, if this constant is real valued the propagation is not allowed, moreover, though losses may exists, the reactive components of the impedance are key for the frequency response to take place and thus the propagation of the waves along the structures.

As for the current, we can get an expression in terms of voltage by using 3.48 and 3.50,

$$I(z) = \frac{\gamma V_o^+ e^{-\gamma z}}{R + \omega L} + \frac{\gamma V_o^- e^{\gamma z}}{R + \omega L} \tag{3.53}$$

where the expression $\frac{\gamma}{R+\omega L}$ may be regarded as the characteristic impedance Z_0.

In the particular case of a lossless line, in other words $\alpha = 0$, the propagation constant becomes imaginary thus it takes the value of β,

$$\gamma = \alpha + j\beta = j\omega\sqrt{LC} \tag{3.54}$$

After considering the fundamental two-wire transmission line model, let us go back to the preliminary results in equations 3.36 and 3.37 for TE waves in rectangular guides; let us contrast these equations with the foregoing results, in particular equation 3.46. Both results are similar regarding the functional form hence we can make,

$$V = j\frac{\omega\mu H_z}{\gamma^2 + \omega^2\mu\epsilon} \quad , \tag{3.55}$$

$$\psi = I \quad . \tag{3.56}$$

Thus the field equations are rewritten as,

$$\frac{\partial V}{\partial z} = -j\omega\mu I \tag{3.57}$$

$$\frac{\partial I}{\partial z} = -\left(\frac{\gamma^2 + \omega^2\mu\epsilon}{j\omega\mu} + j\omega\epsilon\right) V \tag{3.58}$$

which allows us to state that,

$$Z = j\omega\mu \tag{3.59}$$

and,

$$Y = \left(\frac{\gamma^2 + \omega^2\mu\epsilon}{j\omega\mu} + j\omega\epsilon\right) \tag{3.60}$$

Both 3.60 and 3.59 are the required expressions that encode physical information of the phenomena as well as information that enables a circuit modeling of the waveguide as showed in 3.3.

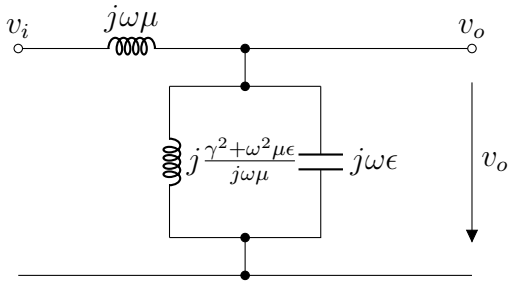


Figure 3.3: Equivalent circuit for TE waves in a rectangular waveguide.

Similar expressions can be derived for TM waves (Jordan and Balmain, 1968), allowing us to represent the case

as a circuit, as showed in 3.4.

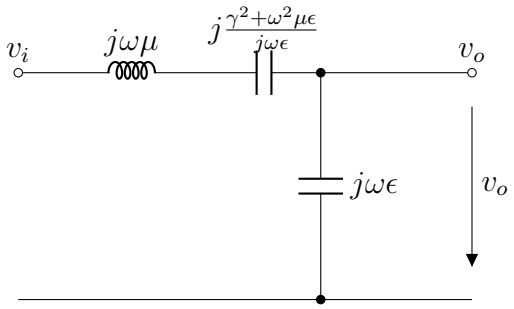


Figure 3.4: Equivalent circuit for TM waves in a rectangular waveguide.

An additional remark worth of mention. The specific values for L, C, R and G, must be obtained for every distinct transmission line problem, for instance, whether you are working with parallel wires or coaxial guides, these components have different values. Fortunately, they are tabulated in the literature, as provided by Franceschetti (1997) Table 7.2 , but in general,

$$L = \mu\frac{Z_0}{\eta} \quad , \quad C = \epsilon\frac{\eta}{Z_0} \tag{3.61}$$

Intuitively, we can approach the situations of open- and short-circuit ends in the line. In the former, the current is zero, while the voltage v_o reaches its maximum value, while in the latter, the current reaches its maximum and the voltage becomes zero. These conditions make sense in the light electromagnetic field analysis. For instance, in the short-circuit case, the zero voltage situation at the end of the line can be depicted as a wall of the same conducting material at the end of the line, which needs to comply with the boundary condition namely the field must be zero.

More important, there is an additional case in which the end of the line is neither open nor closed but a *load* is *attached* to it. This is called the *terminated* line, implying the existence of an impedance whose value can be associated with an adjoined waveguide with different characteristic impedance.

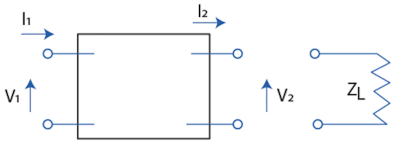


Figure 3.5: Two-terminal network.

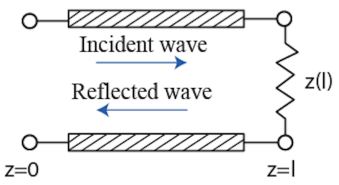


Figure 3.6: Terminated transmission line.

Consider a geometry as showed in Fig. 3.4. This structure can be seen as the junction between two different rectangular waveguides and each one with its own impedance hence there is an difference between them. A

perfectly *matched* impedance occurs when both impedances coincide, otherwise an imbalance takes place which needs to be accounted. The question arises as how to model the intersection between both guides; a preliminary approach is to think as if the opening waveguide terminates with an impedance load equivalent to the impedance difference.

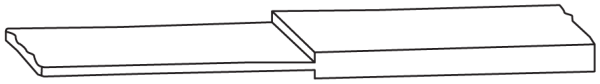


Figure 3.7: A discontinued waveguide.

The treatment of terminated lines is a matter of the utmost importance as it provides a widely implemented tool to analyze electromagnetic waves in guides namely *incident* and *reflected waves* .

We can recall the equations 3.46, 3.50 and 3.51 to express current in terms of the voltage,

$$I(z) = \frac{1}{Z_0} V_o^+ e^{-\gamma z} - V_o^- e^{\gamma z} \tag{3.62}$$

The presence of a terminating impedance has an electromagnetic interpretation which is the need to comply with a boundary condition, in particular, related to the fields hence voltages and current at the end of the line.

$$Z(z) = \frac{V(l)}{I(l)} = Z_0 \frac{V_o^+ e^{-\gamma z} + V_o^- e^{\gamma z}}{V_o^+ e^{-\gamma z} - V_o^- e^{\gamma z}} \tag{3.63}$$

If we take l as the length of the waveguide along the z direction, thus

$$Z_{line} = Z_0 \frac{V_o^+ e^{-\gamma l} + V_o^- e^{\gamma l}}{V_o^+ e^{-\gamma l} - V_o^- e^{\gamma l}} \tag{3.64}$$

Solving for the $V_o^- e^{\gamma l}$ term gives,

$$V_o^- e^{\gamma l} = \frac{Z_{line} - Z_0}{Z_{line} + Z_0} V_o^+ e^{-\gamma l} \tag{3.65}$$

It is worth of mention that the voltage wave is the superposition of two components, moreover, V_o^- represents a reflected wave while V_o^+ is the incident wave. Thus, the amount of *reflected* wave in relation to the incident wave gives what is know as the *reflection coefficient* Γ,

$$\Gamma = \frac{Z_{line} - Z_0}{Z_{line} + Z_0} \tag{3.66}$$

This coefficient appropriately represents the previously mentioned situations. In order to grant that the incident wave is fully transmitted hence there is no reflected component, Z_{load} must coincide with the impedance of the line Z_0. As long as there is an impedance difference between the load and the line, a reflected wave must be expected. The equations 3.51 can be written in terms of Γ,

$$V(z) = V_o^+ \left[e^{-\gamma z} + \Gamma e^{\gamma z} \right] \tag{3.67}$$

which evidently shows the presence of an incident wave and a reflected component. Additionally, the open- and short-circuit cases, can be revisited by considering them in the light of the equation 3.67. If $Z_{load} = 0$ thus $\Gamma = -1$,

$$V(z) = V_o^+ \left[e^{-\gamma z} - e^{\gamma z} \right] = -2jV_o^+ \sin \gamma z \tag{3.68}$$

while for the open-circuit case, $Z_{load} \to \infty$ thus $\Gamma = 1$,

$$V(z) = V_o^+ \left[e^{-\gamma z} + e^{\gamma z}\right] = 2V_o^+ \cos \gamma z \tag{3.69}$$

The reader can find a more extensive treatment of this topics in text like Montgomery et al. (1987) and Owyang (1989). In the interest of a straightforward development of following sections, attention must be paid to the case depicted in Fig. 3.6. As stated before, we need to solve for the case of two waveguides with different impedances thereon Z_0 and Z_1. If we grant no reflected wave within the *load line*, for example, by having an infinite length, then the input guide *sees* a load $Z_{load} = Z_1$ thus,

$$\Gamma = \frac{Z_1 - Z_0}{Z_1 + Z_0} \tag{3.70}$$

The power of the incident wave splits into two components; there is a reflected fraction, obtained from Γ, while the other part is transmitted onto the second guide. The transmitted fraction relates to the *transmission coefficient* T.

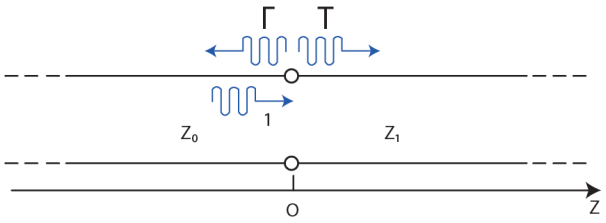


Figure 3.8: Transmission and reflection coefficients between a couple of two-wire transmission lines, each of them with different characteristic impedances.

The transmitted wave corresponds to a fraction **T** of the positive component of the propagating wave in 3.67, $V_o^+ e^{-j\gamma z}$,

$$V(z) = V_o^+ T e^{-\gamma z} \tag{3.71}$$

furthermore, voltage must be same at the intersection point hence as 3.71 is equated with 3.67, we can find an expression for coefficient T (Pozar, 1998),

$$T = 1 + \Gamma = \frac{2Z_1}{Z_1 + Z_0} \tag{3.72}$$

3.1.3 Power transfer

By carefully looking these results, we find the underlying consequence of power transfer. Whether we talk about voltage and current, or about fields transferred along or onto a waveguide, we are in fact treating with power propagation. The equation 2.24 gives the power in terms of the complex Poynting vector, but the interest lies on the real part of the complex quantity, $Re(\frac{1}{2}E \times H^*)$. As you may recall, there is a relation between electric field and voltage given by 3.20 and the same for currents in equation 3.21, and both are useful to find an expression for power in terms of voltages and currents. A comprehensive derivation of this expression can be consulted in Collin (1991) chapter 5 but, for our aim, it is enough to mention that the electric and magnetic energies are obtained, i.e.,

$$W_e = \frac{1}{4} V^2 C \tag{3.73}$$

$$W_m = \frac{1}{4} L I^2 \tag{3.74}$$

to move forward in finding the expression,

$$P = \frac{1}{2} Z_0 I^2 = \frac{1}{2} V I^* \tag{3.75}$$

An alternative perspective is to calculate the total power carried by the fields in the z direction, which is performed through the cross-section of the guide, but as only the transverse components propagate power. Once there is an expression available for the fields, we can compute the energy by integrating the Poynting's vector over the cross section of the guide (Montgomery et al., 1987),

$$P = \frac{1}{2} \frac{1}{Re(\mathcal{Z})} \int |E_{transverse}|^2 \, dS = \frac{1}{2} Re(\mathcal{Z}) \int |H_{transverse}|^2 \, dS \tag{3.76}$$

where the impedance $\mathcal{Z}$ is the wave impedance corresponding to TEM, TE or TM cases. P can be written as (Orfanidis, 2016),

$$P = Re(\frac{1}{2} E_{transverse} \times H^*_{transverse}) \tag{3.77}$$

By using the relations,

$$H = \frac{1}{\eta} \hat{z} \times E \tag{3.78}$$

and

$$|E_{transverse}| = \mathcal{Z} \, |H_{transverse}| \tag{3.79}$$

the equation 3.77 becomes,

$$P = \frac{1}{2\eta} |E|^2 = \frac{1}{2} \eta \, |H|^2 \tag{3.80}$$

As for the TE_{10} mode in rectangular waveguides, the transverse components E_y and H_x are known, thus an expression is obtained for power flow or transmited power recalling the impedance for TE waves in the equation 3.17,

$$P = \frac{1}{2} \mathcal{Z}_{TE} \frac{\beta^2}{\gamma^2} |\nabla H_z|^2 \tag{3.81}$$

$$P = \frac{1}{2} \mathcal{Z}_{TE} \int |H_x|^2 \, dS = \frac{a^3 b}{\lambda_g^2} \mathcal{Z}_{TE} \tag{3.82}$$

where λ_g has been derived in A.41. TM modes can be treated similarly remembering that TM_{11} is the lowest possible mode for TM waves as developed by Montgomery et al. (sec.2-18) and Orfanidis (chap.9).

$$P = \frac{1}{2\mathcal{Z}_{TM}} |E|^2 = \frac{1}{2\mathcal{Z}_{TM}} \frac{\beta^2}{\gamma^2} |\nabla E_z|^2 \tag{3.83}$$

The foregoing expressions for transmitted power provide a framework for the next sections as the behavior of the line is more often explicitly referred in terms of power than voltages and currents. The appropriate unit to use is the *decibel* or **dB** as it tells about relative measures namely increases (gains) or decreases (losses). For instance, 0 dB implies neither a decrease nor an increase in power; 3 dB corresponds to two fold increase in power.

The conversion from linear power ratios to dB is done as,

$$10 \log \left(\frac{\text{Power Level 1}}{\text{Power Level 2}} \right) \tag{3.84}$$

but, if you are comparing voltages or currents,

$$20\log\left(\frac{\text{Voltage Level 1}}{\text{Voltage Level 2}}\right) \tag{3.85}$$

Now consider again the transmission coefficient T. It gives the fraction of power fed onto a guide with Z_1 from a waveguide with Z_0. There is a power relation involved between the incident power and the power that is actually transferred, thus the transmission coefficient can be expressed in dB,

$$IL = -10\log|T|\,dB \tag{3.86}$$

what is commonly known as *insertion loss*. On the other hand, the *return loss*, is the ratio between the incident and the reflected power,

$$RL = -20\log|\Gamma|\,dB \tag{3.87}$$

3.1.4 Wave formalism and Microwave networks

The actual applications of the guiding structures are rather limited if the geometries are restricted to the few herein mentioned. The most exciting applications are driven by the use of complex devices that originate from a composition of canonical geometries such as the rectangular waveguide. Think about devices with many branches forming circuits; each branch has *its own* behavior, dissipates power and can conveniently be represented by a characteristic impedance.

Let us begin this section defining what *port* means. Easily said, a waveguide has terminal points or endings that connect to other guides or structures. Its relevance relies on the concept of a whole structure being characterized by its impedance at every termination or port. Keep in mind that the transmission of power has been modeled in terms of Γ and T, which implies that we must know the impedance at every port if our goal is to develop new devices from interconnected structures.

Furthermore, the composition of many guiding structures is called *microwave network*, where every port has an incident and reflected wave as well as a characteristic impedance. A microwave network can be treated by means of a system of equations similar to 3.41,

$$\begin{bmatrix} V_1 \\ V_2 \\ ... \\ V_n \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} & ... & Z_{1n} \\ Z_{21} & Z_{22} & ... & Z_{2n} \\ ... & ... & ... & ... \\ Z_{n1} & Z_{n2} & ... & Z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ ... \\ I_n \end{bmatrix} \tag{3.88}$$

or in terms of admittances,

$$\begin{bmatrix} I_1 \\ I_2 \\ ... \\ I_n \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} & ... & Y_{1n} \\ Y_{21} & Y_{22} & ... & Y_{2n} \\ ... & ... & ... & ... \\ Y_{n1} & Y_{n2} & ... & Y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ ... \\ V_n \end{bmatrix} \tag{3.89}$$

We can think of these equations as if every relation between a pair of ports is calculated while keeping all other ports open circuits.
Though it all seems familiar and somehow *easy* to grasp, the reality is that there is no direct form to measure quantities like voltages in non-TEM waveguides (Pozar, 1998). For this reason, the incident, reflected and transmitted wave model is of major significance by means of a more suited tool to model complex networks namely the *scattering matrix* .

We have seen, so far, how voltages relates to currents through impedances or admittances, but alternatively, scattering parameters describes the network in terms of incident and reflected waves at every port with the attribute that this quantity can be calculated or measured. When we refer to networks with multiple ports, we could use the model based on voltages and currents but, the characterization of a network is easier if we would be able to perform measurements of incident and reflected power in every port while all other ports remain open-circuited or loaded.

$$\begin{bmatrix} V_1^- \\ V_2^- \\ ... \\ V_n^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & ... & S_{1n} \\ S_{21} & S_{22} & ... & S_{2n} \\ ... & ... & ... & ... \\ S_{n1} & S_{n2} & ... & s_{nn} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \\ ... \\ V_n^+ \end{bmatrix} \tag{3.90}$$

An alternate view to describe a network is by expressing input quantities in terms of output quantities; in the case for a two-port network,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \tag{3.91}$$

The equation 3.91 is significant since, as you may find later in this text, many of the models and solutions presented in prior papers for quadrature hybrids and other microwave devices, rely on this matrix representation. For now, let us limit our scope to scattering parameters since they play a remarkable role in our pursue for optimization method for geometries of microwave device, more precisely, a quadrature hybrid.

Let us consider a two-port guide as a manageable model to further study scattering parameters. The equation 3.90 can be expressed as,

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{3.92}$$

where $[b]$ represents the reflected and $[a]$ the incident components of a generalized variable. Expanding 3.92 yields,

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{3.93}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{3.94}$$

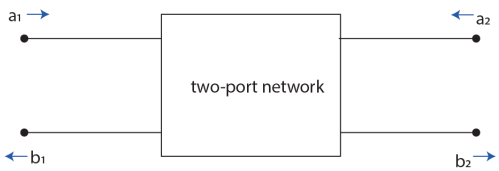


Figure 3.9: Two-port network with incident and reflected quantities.

A preliminary approach to define the S parameters, can be constructed from the prior expansion. The quantity b_1 is made of two components; a_1 is the incident wave in port 1 while a_2 corresponds to the incident wave through port 2 of our guiding structure. This means that S_{11} is equivalent to the reflection coefficient while S_{12} can be seen as the transmission coefficient from port 2 to port 1.

$$S_{11} = \left.\frac{b_1}{a_1}\right|_{a_2=0} \tag{3.95}$$

$$S_{12} = \left.\frac{b_1}{a_2}\right|_{a_1=0} \tag{3.96}$$

$$S_{21} = \left.\frac{b_2}{a_1}\right|_{a_2=0} \tag{3.97}$$

$$S_{22} = \left.\frac{b_2}{a_2}\right|_{a_1=0} \tag{3.98}$$

A *symmetrical* two-port network is the one in which the conditions $S_{11} = S_{22}$ is satisfied. On the other hand, in a *reciprocal network* the condition to satisfy is $S_{12} = S_{21}$. The former implies that the characteristics at either port are equal, while the latter means that, whether you use port 1 or port 2 as the entry point of signal, the behavior of the device is going to be the same in regard of the transferred power from one port onto the other.

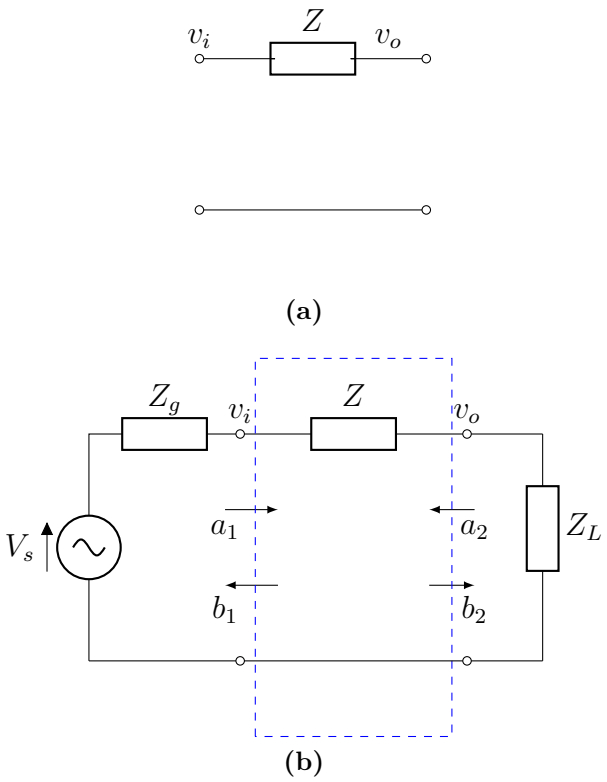


Figure 3.10: Two-port network modeled as a series impedance (a). The model can be generalized by adding a generator and matched load (b).

For instance, voltage v_i between terminals can be written as $v_i = a_1 + b_1$. If $b_1 = 0$ then $v_i = a_1$ which can only happen when $Z_g = Z$ namely the when generator and the line are perfectly matched. Can be obtained an expression for v_i in terms of the source voltage V_s,

$$v_i = V_s \frac{Z_i}{Z_g + Z_i} \tag{3.99}$$

where Z_g is the impedance of the sourcing device and Z_i corresponds to the impedance "seen" at the input

port. As $b_1 = v_i - a_1$ then,

$$b_1 = \frac{V_s}{2}\frac{Z_i - Z_g}{Z_g + Z_i} \tag{3.100}$$

with $Z_i = Z + Z_L$, we get,

$$b_1 = \frac{V_s}{2}\frac{Z + Z_L - Z_g}{Z_g + Z + Z_L} \tag{3.101}$$

Using the last formula together with 3.99 ,we can deduce an expression for S_{11},

$$S_{11} = \frac{Z}{Z + 2Z_g} \tag{3.102}$$

Using the same approach we can find an equation for S_{21} by finding v_o (Owyang, 1989),

$$S_{21} = \frac{2Z_g}{Z + 2Z_g} \tag{3.103}$$

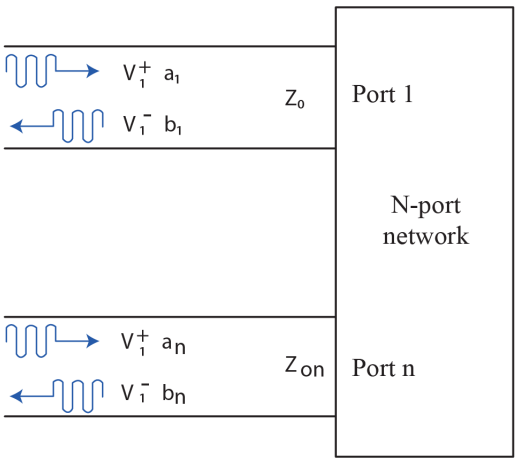


Figure 3.11: N-Port network.

Owyang (1989) presents a general procedure that seeks to obtain expressions for $[a]$ and $[b]$ quantities no matter the physical quantity in question, for instance, voltages, currents or power.

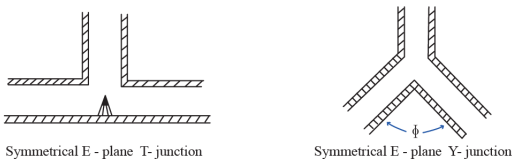


Figure 3.12: 3-Port network samples.

The same applies to an n-port device, where we can define junction planes through which the reflected and incident waves may cross, as depicted in Fig. 3.11. The Fig. 3.12 presents a three-port junction. The junction between the guides must be accounted in some way, and for this reason, the *shunt* or *series* impedance models are used (Owyang, 1989). Three guide lines are depicted clearly, while the difference relies on which type of

impedance is used to model the junction.

The suggested device can be represented in the form of S-parameters as well, moreover, if one port is chosen as the entry of signal, the other two ports may or may not have a perfectly matched load. The scattering matrix has the form,

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \tag{3.104}$$

If the load is matched, the component b_n is zero namely there is no return at ports,

$$S_{11} = S_{22} = S_{33} = 0 \tag{3.105}$$

The scattering matrix has the property of being symmetrical for reciprocal networks, which finally translates in the following unitarity relation (Franceschetti, 1997),

$$\sum_k S^*_{kn} S_{km} = \delta_{nm} \tag{3.106}$$

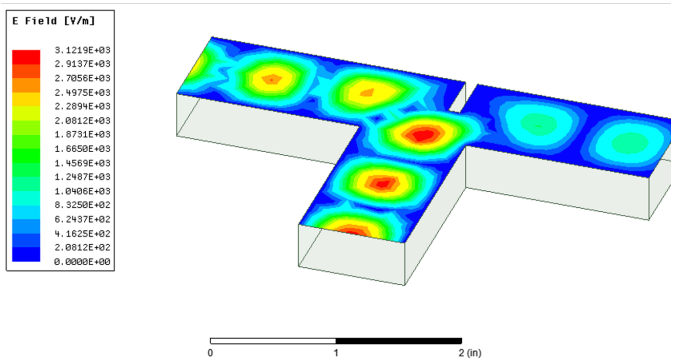


Figure 3.13: Simulated 3-Port T-junction asymmetrical power divider.

This equation represents that the dot product of any column with the complex conjugate of any other different column, is zero thus they are orthogonal. A matrix that satisfies these conditions is a unitary matrix. To be more specific, as the ports are perfectly matched, all components from the summation 3.106 that involves diagonal entries are zero, for instance,

$$\cancel{S^*11}^{\,0} S11 = \cancel{S^*11}^{\,0} S12 = \cancel{S^*11}^{\,0} S13 = 0 \tag{3.107}$$

and the same applies to every diagonal entry. Now, the sum sweeps through $k = 1, 2, 3$, while considering the cases when $n = m$ and $n \neq m$. If the latter is considered, the dirac delta condition gives,

$$S^*_{21} S_{23} + S^*_{31} S_{32} + S^*_{12} S_{13} + S^*_{32} S_{31} + S^*_{13} S_{12} + S^*_{23} S_{21} = 0 \tag{3.108}$$

If we enforce this condition,

$$S^*_{31} S_{32} = S^*_{21} S_{23} = S^*_{12} S_{13} = 0 \tag{3.109}$$

then the entries of the matrix with $n \neq m$ end up being zero (Pozar, 1998),

$$S_{32} = S_{21} = S_{13} = 0 \tag{3.110}$$

The same condition applies to the $n = m$ components thus

$$S_{12}^* S_{12} + S_{13}^* S_{13} = S_{12}^2 + S_{13}^2 = 1 \tag{3.111}$$

$$S_{21}^2 + S_{23}^2 = 1 \tag{3.112}$$

$$S_{31}^2 + S_{32}^2 = 1 \tag{3.113}$$

which finally yields,

$$|S_{12}|^2 = |S_{23}|^2 = |S_{31}|^2 = 1 \tag{3.114}$$

The ends result is a scattering matrix for a three-port device,

$$S = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{3.115}$$

Though many other subjects are of research interest for common and rare networks, we approach them to the extent of not losing the scope of this text. Other authors study topics like symmetry and commutability of the scattering matrix (Owyang, 1989), the relation between impedance, admittance and scattering parameters (Pozar, 1998; Owyang, 1989), and extensive details on how to obtain the entries of each of the matrices herein mentioned depending on the wave mode.

4

Power Dividers and Directional Couplers

4.1 Power Dividers

An immediate application from previous sections, is that of dividing or combining power in certain fraction. Think in terms of an input signal that splits into two or more signals with lower power than the original. A microwave network of that kind of capability is commonly known as *Power Divider*. A behavior like this can be achieved by means of three- or four-port networks. While three-port networks are mainly depicted as *T-type* junctions which ports behave either as inputs or outputs, the four-port devices usually have an input port and the remaining ports are signal outlets with particular properties linked to the specific geometric design.

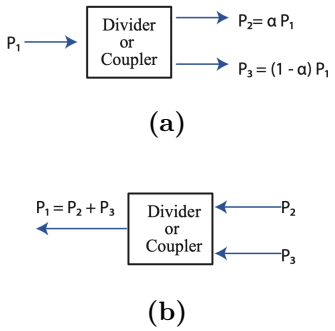
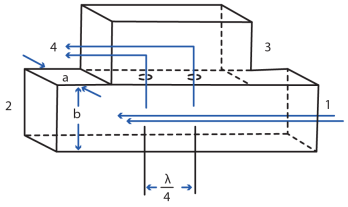


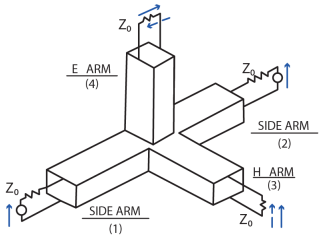
Figure 4.1: Directional Coupler as a 3-Port network application.

At this point, the reader may be aware that the three-port devices were studied in the last section by means of the S matrix for a reciprocal network. Lastly, for this type of devices, the power division takes place in fractions namely $P_2 = \alpha P_1$ and $P_3 = (1-\alpha)P_1$. This is achieved by using matched loads and reactances with an specific proportion, as derived by Collin (2001).

Attention must be paid to the four-port network, as it is requisite to go into the modeling of devices like the quadrature hybrid. The ways in which four-ports can be obtained are countless, furthermore, the coupling method between the guides definitely changes the circuit representation adding extra complexity to the problem. Consider the Fig. 4.2, in which a four-port device is created by the junction of two guides trough a hole. Though there is enough information regarding a two-port line namely a rectangular waveguide, this new device requires more work to model how power transfers between both lines and, consequently, between ports. Firestone (1952), defines *directional coupler* as a device which couples together transmission lines by means of a coupling mechanism, with the capacity to induce a proportion of the traveling wave from one line into the other.



(a) Bethe coupler.



(b) Magic T.

Figure 4.2: 4-Port devices examples.

To analytically solve a complex structure like this may be inconvenient thus the need of the scattering modeling. For instance, the design process can become extremely slow as we introduce modifications to the device such as a change in the angle in which the guides are intersected. For this reason, to speed up the design and optimization process and to obtain the scattering parameters for complex structures, the use of electromagnetic simulation software has become a widely spread practice.

By following a similar procedure to the one applied for the three-port network, a reciprocal perfectly-matched four-port device is represented by a S matrix of the form,

$$S = \begin{bmatrix} 0 & S_{12} & S_{13} & S_{14} \\ S_{21} & 0 & S_{23} & S_{24} \\ S_{31} & S_{32} & 0 & S_{34} \\ S_{41} & S_{42} & S_{43} & 0 \end{bmatrix} \tag{4.1}$$

A reciprocal four-port device is represented by a symmetric matrix with respect to its diagonal, while the zero-valued diagonal means that a perfectly matched load is supposed. A process can be applied to derive the entries of the matrix (see appendix A.4), which yields a matrix of the form,

$$S = \begin{bmatrix} 0 & S_{12} & S_{13} & 0 \\ S_{12} & 0 & 0 & S_{24} \\ S_{13} & 0 & 0 & S_{34} \\ 0 & S_{24} & S_{34} & 0 \end{bmatrix} \tag{4.2}$$

With the results gathered thus far, a few attributes can be inferred for a coupled four-port network:

a.) Matched terminals: As far the junctions and loads are matched, none reflected waves are expected.

b.) Amplitudes and phases: Power is split between the transmitted and coupled ports in different proportion. Phases experience the same at the output ports hence generating different devices to be used in particular cases depending of the amplitude and phase difference between the outlet ports.

c.) Symmetry: Devices are particularly symmetric what is of help to model the electromagnetic behavior.

d.) Coupling: An incident wave in port 1 couples power into port 2 and 3 but not into port 4 (Pozar, 1998) namely port 1 and 4 are uncoupled. On the other hand, port 2 and 3 are uncoupled as well.

e.) 3-dB coupling: As you may remember, a power relation of $\frac{1}{2}$ is equivalent to 3 dB, thus a directional coupler that divides incident power into two equal amounts if called a *3-dB directional coupler* or a *hybrid junction*.

The process gives information about the entries of the matrix, but the specific values have not been deduced. Let us consider that port 1 is the reference as entry point for an incident wave. Let us also recall equations 3.52, 3.67 and 3.71 and it is evident that if an incident wave from port 1 is fully transferred into port 2 namely there are no reflected waves, the wave at the end of the line must have a phase change due to the impedance along the waveguide.

By including the phase change due to the length of the waveguide, all entries can be fully deduced, moreover, we settle in a S-matrix with the form,

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & j & 0 \\ 1 & 0 & 0 & j \\ j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \end{bmatrix} \tag{4.3}$$

This final matrix represents a special case of directional couplers known as *hybrid couplers*. As the reader may notice, as the incident wave is fed at port 1, a $\pi/2$ phase difference takes place between ports 2 and 3. This particular device is commonly known as a *Quadrature Hybrid*. The antisymmetrical case, in which the phase difference is π, is known as *Magic-T hybrid*.

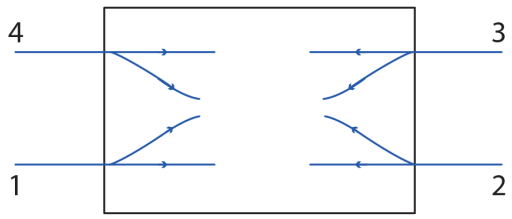


Figure 4.3: Directional coupler signal flow and schematic symbol.

Directional couplers are characterized by two parameters that measure the coupler's ability to separate waves and to couple ports. The *directivity* measures how well the coupler separates forward and reverse waves (Pozar, 1998), which must be infinite in an ideal situation. Numerically it is the relation between the input signal at the coupled port and the reflected signal at the same port thus a high directivity means that the reflections from the output port will cause less interference at the coupled port.

$$Directivity = D = 20 \log \frac{\text{amplitude in wanted port}}{\text{amplitude in isolated port}} = 20 \log \frac{b}{|S_{14}|} \quad dB. \tag{4.4}$$

As for the specific case of this research, the formula used to calculate the directivity is,

$$Directivity = 20 \log dB \left(\frac{S(3,1)}{S(4,1)} \right) \tag{4.5}$$

On the other hand, by design, the coupling between port 1-4 and ports 2-3 should be zero, while the coupling from port 1 into and 3 must have a value that agrees with the design constraints, for instance, a 3-dB coupling.

The relation between the emerging power from port 3 to the incident power in port 1 is called *coupling coefficient*,

$$Coupling = C = 10 \log \frac{P_1}{P_3} = -20 \log b \quad dB. \tag{4.6}$$

The final parameter is the *isolation* which measures the leakage of power from port 1 to port 4,

$$Isolation = I = 10 \log \frac{P_1}{P_4} = -20 \log |S_{14}| \quad dB. \tag{4.7}$$

An alternate definition for directivity is,

$$D = I - C \quad dB \tag{4.8}$$

Many devices can be conceived which are capable of delivering power by using the coupling effect. The reader may be interested in Pozar (1998) who extensively develops solutions for alternative configurations as the **Bethe Hole Couplers** and the **Multihole couplers**. Let us concentrate on modeling our subject: the multi-branch quadrature hybrid.

4.2 Quadrature Hybrid

Millimeter and sub-millimeter receivers, in specific those operating in ALMA's radiotelescopes, are permanently undergoing upgrades to accomplish their scientific objectives. One of the key improvements that the observatory is seeking until 2030, as part of its Wideband Sensitivity Upgrade Plan (Carpenter et al., 2019), is to boost the the sensitivity by improving the front-end capabilities. Sensitivity upgrades are enabled by wider bandwidths and the noise reduction operations, the latter attained using devices like mixers in Double-Side or Single-Side Band configurations.

The quadrature hybrid plays a crucial role in the down-converting operation of the incoming signal to be subsequently used in the image rejection for the noise reduction process. Thus, the slightest refinement in the operation of the quadrature hybrid has a major impact in global performance measures as the sensitivity or image fidelity.

In this section, we seek to set forth the background for quadrature hybrid design, moreover, we look to describe the different methods to solve complex structures like the multi-branch and cascading hybrid structures.

4.2.1 Modeling

As mentioned by Matthaei et al. (1980), these devices can be conceived in distinct forms. A TEM-mode directional coupler is inherently consisting of parallel plates or coupled transmission lines as micro-strips. On the other hand, there are TE- or TM-mode directional couplers in which the coupling takes place through mechanical means.

The coupling mechanism is the foremost modeling goal. There are wide variety of options to couple two main waveguides namely holes, coupled micro-strips and branches, but the choice actually depends on the type of application and constructive techniques. Moreover, while working in frequencies of the order of terahertz, it is presumed that the fabrication method would be that of micro-strip lines, on the contrary, in the lower frequencies regime the cavities are still useful and well understood.

The use of a mechanical coupling is commonly known as *Branch Line Coupler* which consists of parallel transmission lines coupled through branch lines. The branch line geometric dimensions modify its impedances hence allowing to adjust the electric performance of the complete device. Furthermore, the characteristic impedance

of the two parallel main lines may be changed creating different sections thus having an impact in the operation as well.

Regardless of the coupler's type under consideration, it is always necessary to account for the kind of transmission line hence getting different impedance expressions depending of the selected wave modes. In a more general case, the branches connecting the main lines must be represented as shunt or series junctions.

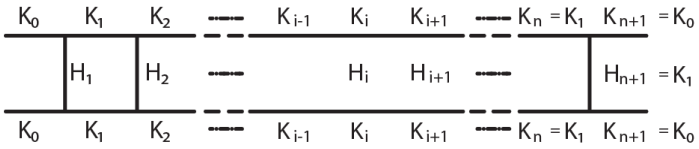


Figure 4.4: Branch-line coupler immittance schematic.

Though different modeling approaches are at hand, we must account for the modeling of discontinuities resulting from intersections or disruptions in the continuity of the transmission line, regardless of the modeling technique; the main waveguides are interrupted thus creating extra impedances altering how the coupler behaves in practical situations. A generalized two-terminal basic waveguide may be represented by a transmission line with some alteration in the middle thus creating an structure divided in three phases, as illustrated in Fig. 4.5 (Marcuvitz, 1986).

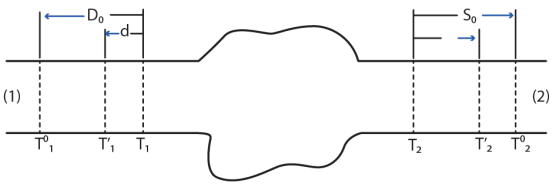


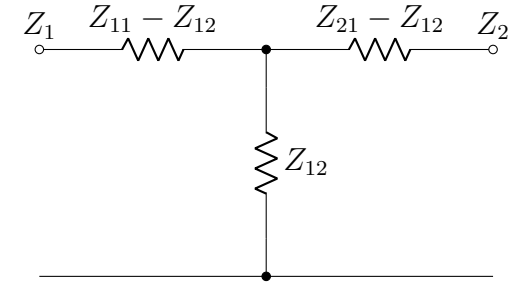
Figure 4.5: General transmission line with discontinuity.

The discontinuity is represented by two basic transmission lines with an additional circuit joining them together. The Fig. 4.6 shows the two types of intermediary connections, both of which are usually known as $\pi-$ and $T-$networks.

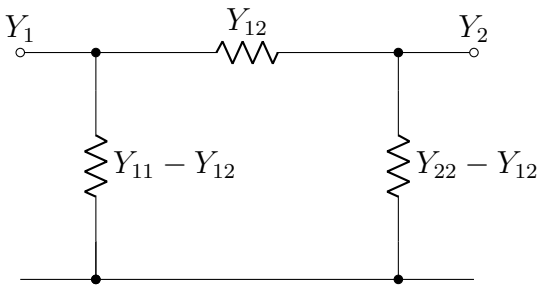
Now, if we try to model every intersection by means of the equivalent circuits presented in the Fig. 4.6, for a branch-line coupler as presented in Fig. 4.5, in which we find a large number of intersections between the main lines and the branches, we will soon discover an extremely intricate equivalent circuit and, in consequence, a hard-to-solve set of equations. To have a wider view of the circuit representation and calculation of different networks, please refer to Marcuvitz (1986) (Sec. 5), where you can find a treatment for a wide variety of two-port networks, for instance, a rectangular to circular change in cross section or the change in height of a rectangular guide.

4.2.2 Periodic structures

As a mean to easily study structures that are the product of joining repetitive sub-structures thus creating a sequential pattern, we require a method that provides a way to treat the device as a black box with input and



(a) T network.



(b) π network.

Figure 4.6: a) T and b) π intermediate connections to model discontinuities.

output signals.

Consider the fundamental junctions from Fig. 4.6, which can be connected in a repetitive manner (Matthaei et al., 1980). The transmission line is then treated as a whole by properly combining the parameters form each individual sub-network. The combining process is done through the ABCD-Matrix method, which was briefly mentioned in the equation 3.91.

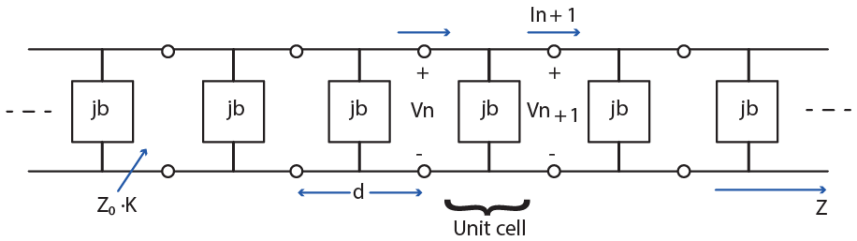
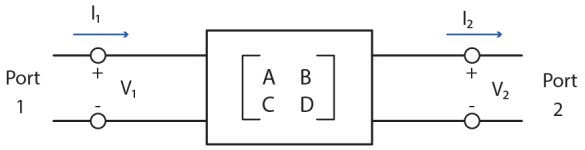


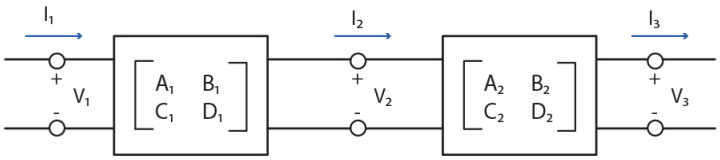
Figure 4.7: Periodic connection with intermediate admittances.

This type of network is usually symbolized in terms of susceptance B, which is the inverse of the reactive components from Fig. 3.2. With the aim of avoiding confusions with the B term in ABCD-Matrix, we will use the suceptance term as $\mathcal{B}$.

The ABCD parameters are defined in terms of voltages and currents presented in the Fig. 4.8, represented by the matrix form 3.91 and obeying the relations (Pozar, 1998),



(a)



(b)

Figure 4.8: ABCD-parameters representation to solve cascade networks.

$$V_1 = AV_2 + BI_2 \tag{4.9}$$
$$I_1 = CV_2 + DI_2 \tag{4.10}$$

The matrix may have a change in I_2 to account for current direction between two cascading connection of adjacent networks.

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} \tag{4.11}$$

where the index $'1'$ represents variables at port 1 and index $'2'$ represents the electric quantities at port 2.

This can be extended by thinking in two adjacent networks as depicted in Fig. 4.8, yielding a set of matrix equations,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \tag{4.12}$$
$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} V_3 \\ I_3 \end{bmatrix} \tag{4.13}$$

which can be rewritten by direct substitution,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} V_3 \\ I_3 \end{bmatrix} \tag{4.14}$$

Each ABCD-matrix is obtained for every individual two-port transmission line. Pozar (Tab. 4.1) presents a brief but useful set of two-port circuits and their ABCD parameters. Finding the specific parameters, as many of the procedures presented earlier in this text, requires to deal with impedances hence calculating voltages and currents, furthermore, all of this can be expressed in terms of $\vec{E}$ and $\vec{H}$ fields. The set of equations 4.15 provides the relations to obtain the parameters whether by strictly introducing voltages and currents in terms of fields

or by simply analyzing the circuits as open- or short-circuits.

$$A = \left.\frac{V_1}{V_2}\right|_{I_2=0} \tag{4.15}$$

$$B = \left.\frac{V_1}{I_2}\right|_{V_2=0} \tag{4.16}$$

$$C = \left.\frac{I_1}{V_2}\right|_{I_2=0} \tag{4.17}$$

$$D = \left.\frac{I_1}{I_2}\right|_{V_2=0} \tag{4.18}$$

Let us test the the power of this method. After working with the basics of transmission lines, some fundamental relations are found and usually tabulated (Ramo et al., 1994) (Tab. 5.11a). For example, the line impedance for a general open-ended line is,

$$Z_i = Z_0 \coth \gamma l \tag{4.19}$$

which is tightly connected with the concepts seen in 3.63 section 3.1.2. When finding the ABCD for a line like the one in Fig. 3.10.a, we get

$$A = \cosh \gamma l \tag{4.20}$$

$$B = Z_0 \sinh \gamma l \tag{4.21}$$

$$C = Y_0 \sinh \gamma l \tag{4.22}$$

$$D = \cosh \gamma l \tag{4.23}$$

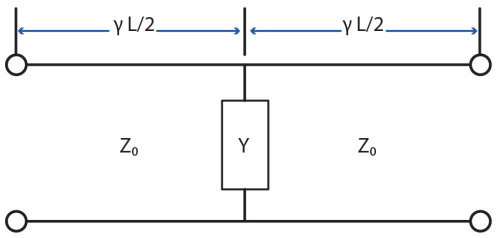


Figure 4.9: A basic two-port transmission line disruption that can be solved using ABCD-parameters.

Now, if we are asked to find expressions for a line with a disruption, it can be conceived as two $l/2$ lines joined by a shunt admittance in cascade. If the admittance is represented as,

$$A = 1 \tag{4.24}$$

$$B = 0 \tag{4.25}$$

$$C = Y \tag{4.26}$$

$$D = 1 \tag{4.27}$$

then the set of matrices to be combined which represent the complete system becomes,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh\frac{\gamma l}{2} & Z_0 \sinh\frac{\gamma l}{2} \\ Y_0 \sinh\frac{\gamma l}{2} & \cosh\frac{\gamma l}{2} \end{bmatrix} \begin{bmatrix} 1 & B \\ Y & 1 \end{bmatrix} \begin{bmatrix} \cosh\frac{\gamma l}{2} & Z_0 \sinh\frac{\gamma l}{2} \\ Y_0 \sinh\frac{\gamma l}{2} & \cosh\frac{\gamma l}{2} \end{bmatrix} \tag{4.28}$$

This multiplication can be made to get every ABCD-parameter, just as presented in 4.29, moreover, there is still the alternative to obtain expressions for S- and Z-parameters in terms of A, B, C and D (Ramo et al., 1994), likewise, there exists a way to convert between two-port network parameters namely from S-parameters into ABCD-parameters and from Z-Matrix into ABCD-parameters as well. Pozar (Tab. 4.2) provides a table comprising all fundamental transformations.

$$A = D = \cosh\gamma l + \frac{Y}{2Y_0}\sinh\gamma l \tag{4.29}$$

which complies with the forward-to-backward symmetry.

4.2.3 Branch-line design

During the course of this text, one of the most important concepts used is the one of impedance, which relates to other parameters that describe how input and output quantities are behaving namely admittance and ABCD parameters. Owyang and Matthaei et al. among others, when approaching the solving of symmetric branch-line couplers, use the extended concept of *immittance*, which is nothing more than the admittances and impedances matrices.

By referring to Fig. 4.4 and 4.7, it can be noticed the presence of series impedances corresponding to sections of the main guide and, shunt admittances representing the physical disruption of the branches. The introduction of reactive elements to a guiding structure is obtained by adding physical discontinuities in a uniform waveguide (Owyang, 1989). This discontinuity perturbs the fields withing the structure depending on the type of reactance produced by the physical alteration.

As we use the name immittance we refer to a general form of impedance/admittance, allowing us to talk in terms of a normalized to describe relative dimensions throughout the quadrature hybrid. As for the notation, Matthaei et al. proposes, and many other authors followed, the use of K_i for series impedance and H_i to shunt admittance.

Something else is required involve during the study quadrature hybrids. As reciprocal devices, they require what is called end-to-end symmetry. So, $K_0 = K_{n+1}$ meaning that the impedance for the input port is the same at the output; $H_1 = H_{n+1}$ namely the opening and closing branches have the same admittance. Besides, this configuration creates n sub-sections and $n + 1$ branches.

4.2.4 Reed's method

The multi-branch waveguide set as directional coupler has been treated for several authors, most of them focusing on how to solve for different setups of the device. One of the most notorious is Reed and Wheeler (Reed and Wheeler, 1956), who introduced a design method that fits for any number of branches. In this proposed design, the opening and closing branches are of the same height, while all other branches have reduced height with respect to the opening and closing ones. The method is based on the even- and odd-mode analysis **(see appendix A.5)**, recalling that the wave amplitude at every port, calculated from the reflection and

transmitted coefficient are,

$$A_1 = \Gamma_e/2 + \Gamma_o/2 \tag{4.30}$$
$$A_2 = T_e/2 + T_o/2 \tag{4.31}$$
$$A_3 = T_e/2 - T_o/2 \tag{4.32}$$
$$A_4 = \Gamma_e/2 - \Gamma_o/2 \tag{4.33}$$

where o and e stand for even and odd components. Further development was achieved through the ABCD

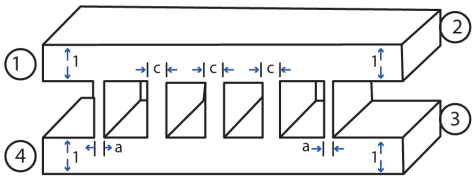


Fig. 1 -- Typical coupler with five branches

Figure 4.10: Normalized dimensions in a branch-line coupler. **Source:**Reed(1958)

parameters formalism. The values of Γ and T coefficients for each half is obtained from,

$$\Gamma/2 = \frac{A + B - C - D}{2(A + B + C + D)} \tag{4.34}$$
$$T/2 = \frac{1}{A + B + C + D+} \tag{4.35}$$

where T and Γ are the elements S_{12} and S_{11} of a basic two-port network.

The method is straightforward in the sense that thinking in sections enables the study of periodic structures. For instance, a three-branch device, with the special feature of having evenly-spaced sections of $\lambda/4$, yields the following matrix for the even mode,

$$Even = \begin{bmatrix} 1 & ja \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix} \begin{bmatrix} 1 & jc \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix} \begin{bmatrix} 1 & ja \\ 0 & 1 \end{bmatrix} \tag{4.36}$$

A three-branch coupler has opening and closing branches of height a and a single intermediate branch of height c. Consequently, it has two subsections consisting of flat two-port transmission lines of length $\lambda/4$ and additionally, as everything else is normalized against the main guide, its height is the unity. Hence the first, third and fifth matrices correspond to the T-junction disruptions for the even mode component of the analysis, while the second and fourth represent the main guide flat sections with characteristic impedance 1.

As the reader may guess, the analysis follow with the odd-mode components. The matrices are the same but a and c must be exchanged by $-a$ and $-c$.

$$Even = \begin{bmatrix} -a(-c) - a & -j(a^2(-c) + 2a) \\ j(-c) & -a(-c) - 1 \end{bmatrix} \tag{4.37}$$
$$Odd = \begin{bmatrix} ac - 1 & -j(a^2c - 2a) \\ jc & ac - 1 \end{bmatrix} \tag{4.38}$$

To proceed with solving for the quadrature hybrid, it is expected that, neither in even- or odd- modes, the device has reflections. This constraint is achieved by setting B and C equal in the total ABCD-matrix. To

obtain a in terms of c we solve,

$$-j(a^2(-c)+2a)=j(-c) \tag{4.39}$$

which finally yields,

$$a=\frac{1\pm\sqrt{1-c^2}}{c} \tag{4.40}$$

If this value is replaced in the matrices and, additionally, c is taken as the impedance depending of the electrical length $\theta=\gamma l$,

$$Even=\begin{bmatrix}\cos\theta & j\sin\theta\\ j\sin\theta & \cos\theta\end{bmatrix} \tag{4.41}$$

$$Odd=\begin{bmatrix}\cos\theta & -j\sin\theta\\ -j\sin\theta & \cos\theta\end{bmatrix} \tag{4.42}$$

By replacing these entries into 4.34 and by using 4.31 then the amplitudes at port 2 and 3 can be deduced, which are our aim.

$$A_2=\cos\theta \tag{4.43}$$

$$A_3=j\sin\theta \tag{4.44}$$

The process can be escalated to find solutions for devices with $n+2$ branches. Reed solves until five branches and arrives to a generalized approach by deducing general expression for the ABCD-Matrix from Tchebyshev polynomials (Owyang, 1989),

$$S_0(-c)=+1 \tag{4.45}$$

$$S_1(-c)=-c \tag{4.46}$$

$$S_2(-c)=c^2-1 \tag{4.47}$$

$$S_3(-c)=-c^3+2c \tag{4.48}$$

$$S_4(-c)=c^4-3c^2+1 \tag{4.49}$$

$$S_5(-c)=-c^5+4c^3-3c \tag{4.50}$$

$$S_6(-c)=c^6-5c^4+6c^2-1 \tag{4.51}$$

$$S_{n+1}(-c)=-c\quad S_n(-c)-S_{n-1}(-c) \tag{4.52}$$

If each branch has a length of $\lambda/4$ keeping the same spacing between branches, with a uniform main guide, a matrix can be found for the even mode,

$$Even_{n+2}=\begin{bmatrix}-aS_n(-c)-S_{n-1}(-c) & -j(a^2(S_n(-c)+2aS_{n-1}(-c)+S_{n-2}(-c))\\ jS_n(-c) & -aS_n(-c)-S_{n-1}(-c)\end{bmatrix} \tag{4.53}$$

For an even number of branches, the term C of the ABCD-matrix 4.53 is equal to A_2, which is used to compute the height c. on the other hand, if there is an odd number of branches, C is equal to A_3. By applying the same reasoning as before, a can be found,

$$a=\frac{\sqrt{1-S_n^2(-c)}-S_{n-1}(-c)}{S_n(-c)} \tag{4.54}$$

This method applies to an specific quadrature hybrid in which the inner branches are all the same and equally spaced by $\lambda/4$ length. It is clear that any further modification completely changes the behavior of the device

hence the power at each port, the reflections and directivity. Bhagat (1968) describes the methods proposed by other authors to deal with the situation in which the main guide is not uniform and the impedance is changing from section to section of the device. This task is not minor, it is time consuming and very inefficient as the goal is to find devices that properly work in different regimes and settings. For instance, to evaluate the impact on the overall bandwidth by a change in a section's impedance would be almost an impossible task.

Reed gives a set of immittances for different settings of branches and couplings. Table 4.1 takes some of these values for reference and further discussion. Recall that a and c correspond to the outer and inner branches, in which all inner branches are of the same size c. Thus, if the input port uses a WR10 waveguide, with a rectangular cross-section of $2.54mm \times 1.27mm$, all other measure are taken relatively to $b = 1.27mm$, hence for a 3-branches coupler, $a = 0.5259mm$ and $c = 0.8980mm$.

Table 4.1: Selected immittances for 3-dB multi-branch couplers. **Source:** Reed (1958)

Branches	Coupling	a	c
3	3-dB	0.4141	0.7071
8	3-dB	0.1064	0.2257

Table 4.2: 3-dB multi-branch couplers dimensions from WR10 main waveguides. All sizes are given in mm.

Branches	Coupling	a	c
3	3-dB	0.5259	0.898
8	3-dB	0.1351	0.2866

Table 4.3: Selected immittances for 3-dB multi-branch couplers. **Source:** Levy and Lind (1968)

Branches	Coupling	a_1	a_2	a_3	a_4	a_5	b_1	b_2	b_3	b_4
3	3-dB	0.4149	1.3432				1.3775			
8	3-dB	0.0145	0.0913	0.3548	0.9210		1.0311	1.2245	1.7453	2.1476

Table 4.4: 3-dB multi-branch couplers to operate in the 75-110 GHz band.

Branches	Coupling	a_1	a_2	a_3	a_4	a_5	b_1	b_2	b_3	b_4
3	3-dB	0.5269	1.70				1.7494			
8	3-dB	0.018415	0.1159	0.4505	0.9210		1.1696	1.5551	2.2165	2.7274

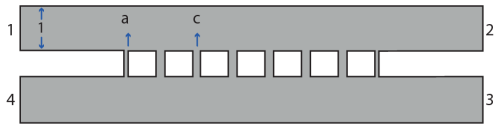


Figure 4.11: 8-branch simulated model from Reed's table. **Source:** Reed(1958)

Levy and Lind (1968), offer a method for a device with $\lambda/4$ sections at mid-band frequency and uniform impedance for the main guides. This method is close to the one of (Reed and Wheeler), nevertheless the use

of Butterworth polynomials and enables the use of branches of different heights along the quadrature hybrid, always complying to the end-to-end symmetry. In particular, they present immittance values for multiple couplings and branches, of which results for 3 and 8 branches are showed in the Table 4.3, where a_i are the branch's heights and b_i correspond to the main-guide's immittance. In practical terms, as the input and output ports are equal by the reciprocity and end-to-end symmetry, the immittance values are equivalent to geometrical sizes to be calculated as fractions of the main-guide size.

By selecting as our objective the 75-110 GHz band, the input port must be a WR10 waveguide, with a rectangular cross-section of $2.54mm \times 1.27mm$. In consequence, the reference dimension is 1.27mm, giving the sizes for branches and sections showed in the Table 4.4. A result worth to be mentioned is the significant reduction in branch sizes as the number of branches is increased.

5

Quadrature Hybrid Simulation

Though the modeling process introduced in the last chapter seems straight forward, it might be clear to this point how the complexity grows as the geometry changes. Alterations like adding branches or using sections with different electrical length hence varying the mainline impedance, just play a role in making hard to analytically derive the quadrature hybrid model. However, while looking for quadrature hybrids that fully comply with a set of requirements for specific applications, an optimization method must be applied.

Herein we present the most salient features of the quadrature hybrid by using Ansys HFSS as simulation tool. We use the design obtained by Reed and Wheeler in Table 4.2 as the simulation object. Concisely, the device has flat main guides, $\lambda/4 = 0.74mm$ separation between branches and equally-sized inner branches. As showed in Fig. 5.1, the incoming power to port 1 is divided into port 2 and port 3. Also, the progressive effect of the branches is clear in the sense of the phase shifting effect between the both outlets. As expected, the port 4 is isolated namely there is no power flowing through it.

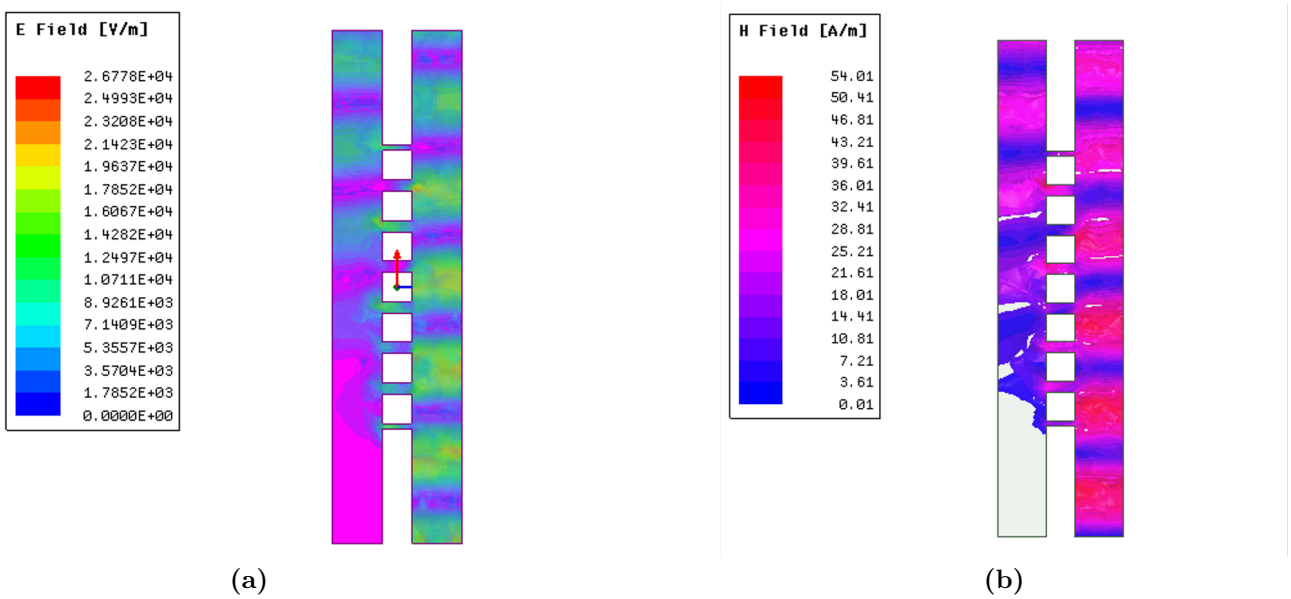


Figure 5.1: Simulated (a) E-field and (b) H-field for a 3-dB 8-branch quadrature hybrid.

A deeper understanding of this type of coupler stems from the simulation performed on the Reed's eight-branch hybrid (Fig. 5.2). There are four merit figures to describe the behavior, moreover, these parameters are of major significance during the optimization process to be applied during forthcoming sections.

Amplitude Imbalance: This is regarded as one of the most relevant features to observe as it becomes essential for the image rejection performance in ALMA's receivers (Fig. 5.2.a). As for the 3-dB multi-branch from

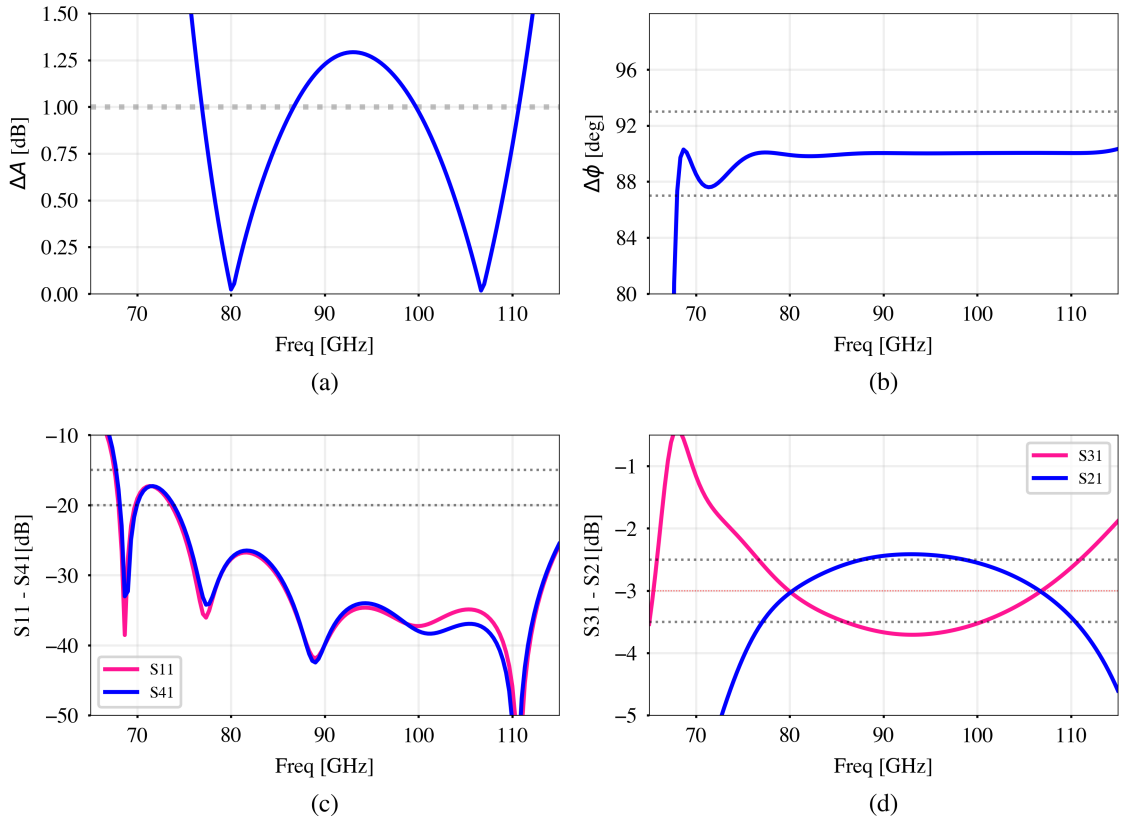


Figure 5.2: Amplitude imbalance (a), phase imbalance (b) and S-parameters (c,d) for 8-branch quadrature hybrid set for W-band (75-110 GHz).

Reed's research, it is evident that the mid-band amplitude imbalance is out of limits, therefore failing to fulfill the $\Delta A < 1$ requirement.

Phase Imbalance: Next, we find the $\Delta\phi$ or *Phase Imbalance*, which measures how phase deviate with respect to the 90° difference between the wave at ports 2 and 3. Herein we can see that this parameter behaves as expected throughout the operational band, in spite of failing in the range between 70 GHz and 78 GHZ.

Coupling: A significant amount of work is placed on finding the hybrid design that complies with the mentioned constraints, moreover, they are the result of the S-parameters measures in the operational band. The S_{31} correspond to the coupling effect supposed to happen in a directional coupler, hence a power transmission must occur between port 1 and port 3. It is expected that this parameters stays the same or as close as possible to -3 dB. Let us recall that this value tells us about the 1/2 power split between port 2 and 3. The ΔA parameters is directly dependent on the S_{31} and S_{21}; the farther they are apart from the -3 dB mark, the higher the amplitude imbalance. It is sufficient to notice the gap between S_{31} and S_{21}, which is specially large at mid-band. An additional feature must be noticed. The different researches, whether through simulations or by laboratory tests, show that as we try to have a wider useful bandwidth by adding more branches, the mid-band coupling parameter suffers a detriment namely gets higher. This is of concern as it reveals the presence of two conflicting parameters to account for during the optimization thus setting a limit to the number of branches or the bandwidth coverage.

The bandwidth: An increase in the number of branches allows for a wider bandwidth (Ishii, 1995). The physical consequence is that the height of each branch is reduced as the branch number is creased hence setting a practical design and fabrication limit for multi-branch couplers. This is confirmed by the branch sizes found

in Table 4.3. Trials show that two-branch 3-dB hybrids have a useful bandwidth of 10% while a Tchebyschev design allows for roughly a 50%.

5.1 The design and optimization problem

The design challenge consists of choosing the branch line locations and impedances namely their heights to achieve a desired coupling and directivity for an specific operational band. Any method used to model the hybrid must account for the T-junction effects, adding extra complexity to the design. As our goal is to find quadrature hybrid designs that can comply with all constrains along the whole bandwidth, attention is paid to the the operational bands ALMA Band 2 (67-90 GHz) and ALMA Band 3 (the 84-116 GHz).

The implications for radio astronomy are significant. Quadrature hybrids are proposed to be used in Single-Side Band heterodyne receivers for the ALMA bands 3 to 10 (Andoh and Minamidani, 2003) by undergoing a the same design process as herein introduced, while using assisted by HFSS simulator.

More recently, as part of the *ALMA Development Roadmap*, the observatory is pursuing a wider bandwidth, consequently, new receivers are being designed for ALMA Band 2 to extend the original 67-90 GHz to reach a 67–116 GHz atmospheric window coverage (Mroczkowski et al., 2019). Yagoubov et al. reported a proposal for a receiver intended to operate in the extended 67–116 GHz window which contains 2SB channels with 3-dB 90° couplers (Yagoubov et al., 2020). Therefore, there is an ongoing search for increased sensitivity in ALMA Receivers in a wider band, all this complying with the tight constraints established for radio astronomical detection.

Authors like Gonzalez et al. and Ding et al. have approached the problem of dealing with amplitude and phase imbalances while trying to achieve a wider bandwidth. The first approach is to extend the number of branches and to have non-uniform main waveguides hence modulating the behavior by changing the intermediate section height and length. Moreover, Gonzalez et al. managed to design and fabricate a quadrature hybrid to operate in W-band (75-110GHz) with an amplitude imbalance of $0.8dB$.

Rashid et al. (2016) presented a design based on the Reed's model for multi-branch couplers. Their work introduced the addition of ripple in the operational band by modifying the heights of the input and output branches and by adding discontinuities in the form of extruded cavities in the main-guide. By using these features they achieved a low amplitude imbalance as the ripple in the mid-band sections allows for closer values of the S_{31} and S_{21} to the -3 dB mark, all that while greater branches are permitted thus easing the fabrication process.

Accomplishing a full-band coverage while complying with the quadrature hybrid theoretical behavior namely zero return losses, 3-dB coupling and 90° phase shift.

$$\Delta A < 1dB \tag{5.1}$$
$$89° < \phi < 91° \tag{5.2}$$
$$-3.5dB < S_{31} < -2.5dB \tag{5.3}$$
$$-3.5dB < S_{21} < -2.5dB \tag{5.4}$$
$$S_{11} < -15dB \tag{5.5}$$
$$S_{41} < -15dB \tag{5.6}$$

where $\Delta A = \text{mag}\left(\frac{S21}{S31}\right)$ and $\Delta\phi = \arg\left(\frac{S21}{S31}\right)$. As already mentioned, the quadrature hybrid has a bandwidth coverage under 50% hence it is considered particularly narrow-band device, which is far from the 100% expected coverage.

6

Particle Swarm Optimization Algorithm (PSO)

- Evolutionary algorithms, such as the Particle Swarm algorithm, have a major feature namely random exploration of the solution space. This feature allows an exploration of the parameter space using sampling methods.
- The algorithm has been widely validated in numerous cases, particularly in geometry-oriented problems, but not exclusively.
- This research introduces a novel application of the PSO algorithm to quadrature hybrid design, particularly for ALMA receivers.
- We also introduce a novel application of the *absorbing wall* technique for radio astronomy instrumentation optimization.

6.1 Optimization

In this section, we provide a brief review of general optimization concepts to subsequently advance in applying and optimization algorithm for our quadrature hybrid design problem.

Though many authors can provide comprehensive definitions of the *optimization* concept, it can be simply put as the *detection of optimal solutions for a problem* (Parsopoulos and Vrahatis, 2010). Greater attention must be paid to the type of problem and the constraints posed by it. Additionally, even though a prospect solution fulfills all the requirements set by the problem, still needs to undergo a test until proved to be the optimal one, which implies that there must be a method to scientifically demonstrate the quality of the proposed solutions.

In the search of optimization methods, many phases have been exhausted due the fact that the need of optimization is primordial in science and engineering. Just consider the evolution of linear, quadratic programming or dynamic programming, all being major optimization sub-fields intended to satisfy specific needs. During the 80's a boom of different methods took place, mainly following the trend of gradient descent algorithm, but since the 90's there is a growing trend to employ Machine Learning, Genetics and Evolutionary algorithms (Rajesh Kumar Arora, 2015).

Despite the selected algorithm, the modeling phase always precedes the optimization allowing for a mathematically description of the actual problem. During the modeling phase, variables, parameters and constraints must be defined in a way that properly represents our system. Let us consider the modeling of a physical spin system using ISING (Hartmann and Rieger, 2001). By using this model, the idea of traditional optimization methods to find a state that satisfies the imposed constraints i.e., the temperature, is unfeasible. The system and the problem impose the use of stochastic methods, like Monte Carlo techniques, to infer probability distributions for the system's states. This situation depicts the importance of modeling and appropriately selecting the *right*

optimization method for the problem.

A rephrasing of the optimization problem would be that our goal is to maximize or minimize a function, which is something conceptually clear from calculus courses. This definition explains why so many disciplines find of great value the implementation of optimization methods. Under this criteria, an optimization problem can be expressed as,

$$\min f(x) \tag{6.1}$$

subject to,

$$g_i(x) \leq a \quad i = 1, 2, ..., m \tag{6.2}$$
$$h_j(x) = c \quad j = 1, 2, ..., n \tag{6.3}$$
$$\tag{6.4}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{bmatrix}$$

where g_i are inequality constraints while the h_j are equality constraints. The x_i corresponds to the design variables which are usually bounded (constrained optimization) by maximal and minimum values, in other words $x_{min} \leq x_i \leq x_{max}$. Therefore, the constraints are functions of the design variables and, if the set x of design variables satisfy all the imposed constraints is thus considered a *feasible solution*. The point x_s from the set of feasible solutions $\mathcal{S}$ that allows for the minimum of $f(x)$ to be reached is called an optimum point.

$$\min_{x_i \in \mathcal{S}} f(x) \tag{6.5}$$

Thus, the idea of minimizing a function implies a lack of desirability of solutions for a particular model, in other words, $f(x)$ represents the cost of using x_i as a solution. In a global view, cases could exist in which there are two "issues" to address but cannot be combined into a single measure of cost hence creating a situation in which more than one objective functions must be optimized. This is usually known as *Multi-objective optimization*. An objective function can be a mapping from n dimensional into a one-dimensional space,

$$\forall x \in \mathbb{R}^n, f : \mathbb{R}^n \to \mathbb{R} \tag{6.6}$$

A linearly constrained problem is characterized by functions whereby the constrains satisfy $g : \mathbb{R}^n \to \mathbb{R}$. On the other hand non-linear constraints are expressed by function with more than one feasible point to satisfy the equality constraint, for example a circle or elliptical functions.

6.1.1 Global and local minima

As for the objective function, in practical terms, the optimization problem frequently involves functions that have more than one minima point. This situation manifests itself as false global solutions emerging from an algorithm incapable of further exploring the solution space.

In the Fig. 6.1, two minima points can be found though one global minima can be identified. Functions like this pose a challenge for algorithms and become a major concern during the selection and tuning of an algorithm. As we will see later, algorithms like the Particle Swarm Optimization (PSO), can easily fall in local minima points hence unable to get out of the well to finally reach the global minima. It strongly depends on the type of algorithm and the hyper-parameters selection to ensure it finally reaches optimal solutions.

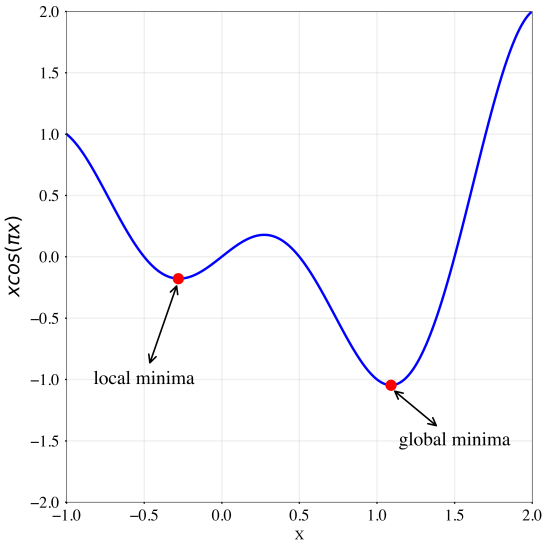


Figure 6.1: Function with two local minima points.

Techniques have been developed to help to establish a between the model and the characteristics of the selected algorithm. This is commonly known as *transformations* or *performance enhancing techniques*, which implicitly or explicitly introduces changes to the variables or constraints to improve the effectiveness of the algorithm. This is not a minor subject as there is a explicit need to prevent finding local minima.

There are two paradigms to apply some transformation to our objective function or the algorithm in order to reduce the probability of finding a local minima instead a global one. *Trajectory methods* are based on changing the direction as the algorithm descends into a local minima thus switching to ascend when reaching that minima point. As the algorithm explores the solution space in the opposite direction, a maximal point is expected to be found, assuming a smooth differentiable curve. If the maximal point is found, the algorithm is switched back to descent trajectory hence there is a chance for the process to end in a new minima.

The second transformation is known as *penalty methods*. This method directly modifies the objective function, preventing the local minima by excluding them through penalizing functions applied to each local minima. This type of transformation is tightly related to heuristic methods as the process is improved by the introduction of heuristics to avoid false local minima (Törn and Zilinskas, 1989).

Transformations go beyond the realm of local minima to become widely used techniques to face multi-objective optimization problems as well as to transform constrained problems into unconstrained ones. For instance, let us consider a situation in which our problem has n objective functions $f_n(x)$. We can define a new function,

$$\forall x \in \mathbb{R}^n, F : \mathbb{R}^n \to \mathbb{R} \tag{6.7}$$

with,

$$F(x) = w_1 f_1 + w_2 f_2 ... = \sum w_n f_n \tag{6.8}$$

in which w_n are weights defined mainly on the basis of experience or via heuristic methods. Hence a new function $F(x)$ is formed which maps to a one-dimensional real space and becomes our new minimization/maximization objective.

In a penalty method, the algorithm penalizes the objective function as constrains are violated. Let us consider a new objective function that includes constraints expressed as a Lagrange function,

$$F(x) = f(x) + \sum_j \lambda_j h_j(x) + \sum_i \mu_i g_i(x) \tag{6.9}$$

A penalty is applied by two means, by transforming constraints into functions and by adding a penalty parameter. constraints can be expressed as functions that add a value to the total function when the constraints are not satisfied, and become zero in the opposite case. If our constraint is $g_i \leq 0$, then,

$$\max[0, g_i(x)] \tag{6.10}$$

would be a valid function to apply. Thus equation 6.9 becomes,

$$F(x) = f(x) \pm penalty \tag{6.11}$$

$$F(x) = f(x) + r \sum_j h_j(x) + r \sum_i \max[0, g_i(x)] \tag{6.12}$$

where r is the penalty coefficient, which becomes a tunnable parameter during the optimization process.

6.1.2 Review on Optimization Algorithms

After broadly defining an optimization problem, we have to pay attention to the implementation of methods to solve them. The rise of new computational methods and the growth in the processing capabilities enabled the development of maximization/minimization methods, many of which were around for decades before being effectively applied to science and engineering problems. These methods are described in the form of operations to solve a problem which, in many cases, is traced back to ordered computational sentences (Baldick and Uni-, 2006). In this sense, an *algorithm* is the structure of operations that effectively solves a problem. Moreover, there are many adequate alternative structures to find an answer for an specific problem only distinguished by their speed to come up with a satisfying solution.

Algorithms are classified into direct and iterative. The former means that there is finite list of steps to arrive to a solution while the latter implies the need of recurrent cycles of the operation list to finally obtain a solution. Though this topic is of great interest, the scope must be narrowed down to iterative algorithms, to focus on new concepts as random exploration.

The idea that supports the types of method to solve multi-objective optimization problems is *search space*. This can be seen as the set of decision variables that satisfy the constraints or feasible set $\mathcal{S}$, furthermore, in an problem with n decision variables, each possible solution is an n-dimensional vector. Optimization algorithms are intended to search for vectors in the search space that optimizes the objective function by implementing *search strategies*. The Fig. 6.2 shows some common optimization methods.

6.1.2.1 Probabilistic strategies

This strategy applies sampling techniques to select points in the search space, to evaluate them and to choose the best (Bozorg-Haddad et al., 2017). As the reader may notice, the larger the decision space the greater the computational overhead; this situation can grow until becoming unfeasible to implement due to time an computational required resources. This type of algorithms are frequently performed through methods like *random search* in which the space is explored randomly choosing prospect solutions and then evaluating their objective functions.

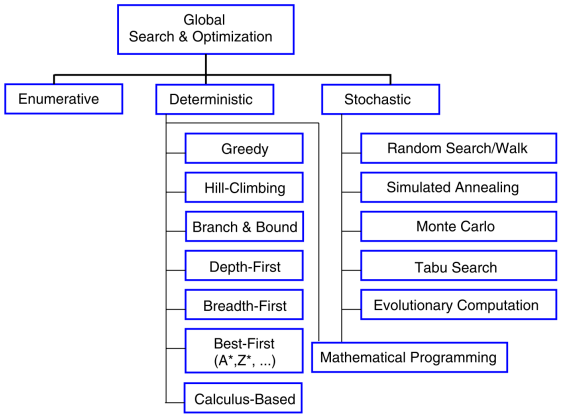


Figure 6.2: Optimization approaches taxonomy. Source: Coello et al., 2007.

As random search methods soon become inefficient and unpractical, solutions have been provided to solve the issue of having large search spaces hence algorithms like the *targeted sampling* have become widely used to the point of being the foundation for meta-heuristic methods. These methods have more systematic search strategy.

6.1.2.2 Heuristic and meta-heuristic strategies

Heuristics are widely seen as problem-solving methods to reach sufficiently good approximations. For instance, *the rule of thumb* is an heuristic method in which a guess is made based on prior information or some predetermined assessment parameter.

These methods take into account the experience gained previously to select feasible solutions. Each problem requires a particular heuristics in which the goal consists in finding potentially good solutions, but not solving the problem itself. This characteristic necessarily implies the use of gained experience through iterative processes to refine the selection of the proposed solutions. A side explanation from the search-space perspective, is that heuristics remove portions of the search space to reduce the number of solution to be tested thus targeted sampling gradually centers the attention in regions of the search space where is very likely to find the global optimum (Bozorg-Haddad et al., 2017).

An empirical salient feature is that these algorithms own what is called a *problem-independent design*, which means that these methods are well suited for all kind of complex problems. This is an advantage to tackle extremely complex problems that are hard to solve by traditional means. Moreover, an advanced version of heuristic methods are the Meta-heuristic and evolutionary algorithms which comply with the problem-independent optimization principle.

At a fundamental level, meta-heuristic algorithms simulate elements and procedures that translate into intelligent behaviors in nature, for instance, evolutionary computation and swarm intelligence belong to this category of algorithms. Meta-heuristic algorithms like genetic algorithms, particle swarm optimization, simulated annealing and variable neighborhood search, are widely used in different field such as genetic engineering, aerodynamics, and electromagnetic design, as a consequence of the adaptability to almost any sort of optimization problem.

6.1.2.3 Population based strategies

What distinguishes these type of algorithms from others is their population-based approach in which a population of feasible solutions is created and, through an iterative process, a refinement of that population is performed by selection and replacement procedures therefore giving birth to a new population that carries the

best inherited characteristics of their predecessors (Rao Kurada et al., 2013;Bozorg-Haddad et al., 2017).

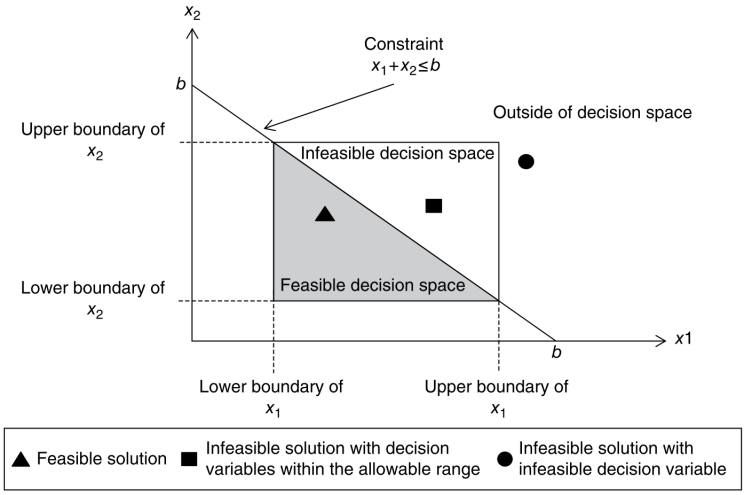


Figure 6.3: A feasible solution space spanned by two variables. Source:Bozorg-Haddad et al., 2017.

A remarkable enhancement, in comparison with other algorithms, is the parallel-search process. Instead of using a single proposed solution to evaluate its fitness, populations are formed by many individuals all having its own fitness value. The refinement also occurs in parallel manner because the heuristic rules are equally applied to all members. Therefore, whether a couple of individuals or a very large populations are used, the refinement still occurs hence reaching nothing less than a local minima solution.

Meta-heuristic algorithms and population based methods are rarely deterministic in nature. On the contrary, they are more stochastic as the population is randomly generated from the search space, mainly during the first steps of the algorithm. In this case, there is no meaningful difference in the nature of the decision variables, whether they are discrete or continuous is not a relevant issue. As for the continuous variables, a sample must be randomly taken between their lower and upper boundaries by means of a distribution function which, for all practical purposes, is usually chosen as a uniform distribution for random generation of the variables.

By including the random generation of the individuals and the sampling method governing the *mutation process* a region of feasible solution from the whole search space can be defined. This is showed in the Fig. 6.3 for a two-variable case in which a two-dimensional solution space is spanned. There is a feasible region established by the joint knowledge of the max and min values for each variable.

Our intention is to get more fully acquainted with the suitable methods to optimize the quadrature hybrid thus finding a design that complies with the constraints introduced in former chapters.

6.2 Particle Swarm Optimization (PSO) Algorithm

6.2.1 Theoretical background

Instead of getting right into the PSO algorithm and its structure, it is of greater value to briefly introduce the concepts involved in genetic and evolutionary algorithms. This will be of help to understand how the PSO works and how the structure must be developed to fulfill the guiding principles of the **population-based meta-heuristic algorithms**.

There is a very diverse field of study around these algorithms resulting from a significant number of real-world applications in which the concepts have been tested with exceptional results. As the field grows, many new alternative algorithms are proposed, for instance, the multi-objective Evolutionary Algorithm (MOEA) which is a generalization of Evolutionary Algorithms (EA), or the mimetic algorithm (MA) which is a hybrid between EA and local search algorithm. These algorithms differ from each other but all keep the same foundational and core process (Corne and Lones, 2018). Computationally speaking, PSO is an extension of *cellular automata* (CA), in which the cells in CA can be seen as the swarm. Both models follow the principles of a) individual particles updated in parallel, b) new values depend only on the previous and its neighbors and c) all particles update under the same rules (del Valle et al., 2008).

6.2.1.1 Genetic and evolutionary algorithms

Nature-inspired algorithms adopt the ways that nature uses to adapt to challenging environments or systems. On the one hand, Genetic Algorithms (GA) are based on mechanisms of natural selection namely survival of the fittest. Intuitively, we can realize that the population undergoes many transformations (*by cross-over*), furthermore, those changes manifests through *generations*. As the algorithm mimics nature, the individuals are selected from their fitness to environmental conditions leaving a final individual as the best at the end of the process. On the other hand, Evolutionary Algorithms (EA) combine adaptation and learning in order to mimic “Intelligent Behaviors", in which adaptation occurs by evolution (*by mutation*) exploiting information from previous steps (Parsopoulos and Vrahatis, 2010).

Algorithm 1 Genetic and Evolutionary core algorithm

```
Initialize:
    Population
Evaluate:  Population
while condition do
    Apply:  Selection
    Apply:  Crossover/combination
    Apply:  Mutation
    Evaluate:  New individuals
    Update:  Population
end while
```

Algorithm 1 shows the core operations performed when developing GA and EA. The differences sprout as the specific operators are adjusted for the problem we are dealing with, for instance, the selection can be done differently depending on whether binary or real-numbered representations are used.

Population initialization: A first generation generated from the search space. The most common procedure is a random generation of this initial population using a uniform distribution spanning the whole the search space. As the designer gathers information about the characteristics of the population or regions of a higher potential to find a solution, different distributions functions can be used to select a higher number of individuals from those special regions.

Selection: A variety of approaches can be used to select individuals by either deterministic or stochastic criteria. A deterministic criteria consists in selecting the best individuals directly namely those with the lowest fitness value. There are other stochastic methods like *tournament selection* and *roulette-wheel* selection (Parsopoulos and Vrahatis, 2010).

Crossover or combination: This is a procedure in which two individuals are crossed to produce a new individual which has information carried by its parents. The crossing can randomly happen between parents selected in the previous step by applying an *exchange* operation of sub-parts of them.

Mutation: This process is fundamental in the heuristic optimization process as it means a change in the properties of the individual to fit the environment. The mutation has a goal of keeping the population as diverse as possible thus avoiding undesirable behaviors, which in our case means getting stucked into local minima. If we consider a case in which the individuals are binary represented, a mutation operation could consist of random bit flips. Thus, mutations can be realized as randomly introduced perturbations in each individual's structure.

6.2.1.2 Particle Swarm

The ground for the PSO algorithm is the meta-heuristic of *swarm intelligence*. This behavior is seen in flocks of birds as they achieve the goal of reaching a destination despite the high number of different individuals, which finally inspired Eberhart and Kennedy (1995) to develop the algorithm. Swarm intelligence deals with many-individual systems who coordinates themselves without centralized control. There are many theories on how they achieve control in a decentralized manner, but the most accepted states that this behavior results from the local interaction between individuals which finally translates into a collective behavior.

The PSO algorithm is based on this kind of social interaction seen in bird flocks. An individual seeking for food in a region could have a hard time finding it, but instead of keep looking in the same spot, it is more "intelligent" to follow the bird that is known to be closer to food. This same principle is followed by the PSO in the search for optimal solutions.

The bird population is expressed as a set of points in the solution space called *particles*. These particles are evaluated against a fitness function which measures the *distance* to food. Additionally, each particle has a direction and speed guiding how it moves around the search space. As each particle has an associated fitness value hence the best particle becomes the leader of the flock, which is followed by the other particles around the search space.

A key ingredient of the heuristic process relies on how each particle determines its next position in the search space. This is achieved by using a) its best individual position and b) the best position achieve globally by the group. This particle updating is related to the mutation element of the more universal algorithm, in which the new position is set by a function that uses best individual position and best global position as well.

6.2.2 Structure

Two of the main structural elements of the PSO have already been mentioned namely *particle* and *velocity*. A third element that characterizes a particle is *position*. As a population-based algorithm, initializing the particles population is one of the first tasks, which is done by randomly initializing positions and velocities.
To be more specific, a particle is represented by a vector which entries are the decision variables that determine the search space. In an n-dimensional optimization problem, particles are defined by vectors as,

$$particle = X = [x_1, x_2, ..., x_n] \tag{6.13}$$

where X represents a possible solution to the optimization problem. The PSO starts by randomly generating many particles thus the population can be seen as a matrix of size $m \times n$ where m denotes the size of the population of feasible solutions.

$$population = \begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ X_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & ... & x_{1n} \\ x_{21} & x_{22} & ... & x_{2n} \\ . & & & \\ . & & & \\ x_{m1} & x_{m2} & ... & x_{mn} \end{bmatrix}$$

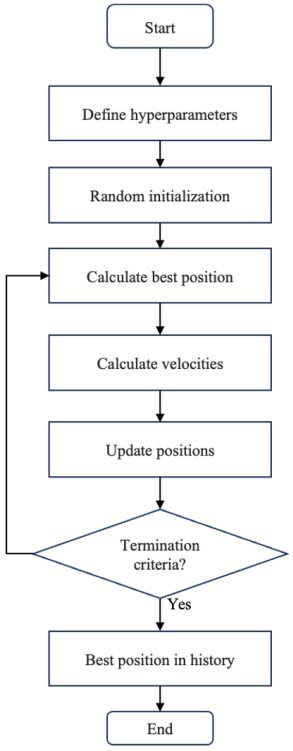


Figure 6.4: Flow chart for the PSO Algorithm.

After creating the population, every particle must undergo an evaluation to classify which one holds the best position in the search space. Two values are used to track every particle's performance namely the *pbest* and the *gbest*. The former the best position array achieved by every particle, while the latter is the best position among all particles.

$$pbest = [best\ x_1, best\ x_2, ..., best\ x_n] \tag{6.14}$$

$$gbest = [global_{best}\ x_1, global_{best}\ x_2, ..., global_{best}\ x_n] \tag{6.15}$$

Keep in mid that the x_i entries are explicit values given to every variable of the problem, is written in vector form to represent a position in a hyper-space. As we are working with an iterative algorithm, in every cycle of the optimization process the *pbest* and *gbest* are updated which means that the holder the best global position title could change in every iteration. This process must stop when the *convergence* of all particles around a consistent global best position takes place.

6.2.2.1 Velocity and particle update

A key feature of the algorithm is the use of *velocity* as the method for particles to change direction and follow the global best particle. This can be seen as a particle whose flying around the search space is determined by the velocity vector. This parameter is used to update every particle's next position by a calculation based on the *pbest* and *gbest* which represent available information regarding own experience and knowledge of the performance of other individuals. This calculation accounts for the relative distance between a particle and the global best, furthermore, if there is a significant relative difference a high speed in the direction of the global best is allowed to quickly follow the leader. That is not the case when the particle is very close to the global best particle, in which case the speed must decrease in order to remain close to the best individual.
As in the case of the position the velocity is an array for every particle as well, in which every entry is the

velocity corresponding to each variable of the n-dimensional space. Thus, for a particle j its velocity looks like,

$$V_j = [v_1, v_2, ..., v_n] \tag{6.16}$$

The first approximation to the velocity equation, must include local and global knowledge but, as these two factors change its relative importance for every variable involved in our problem, makes more sense to consider random weights for every factor, thus the velocity can be calculated by,

$$\vec{v_i}(t) = \vec{v_i}(t-1) + \phi_1 rand_1 \cdot \left(\vec{pbest_i} - \vec{X_i}(t-1)\right) + \phi_2 rand_2 \cdot \left(\vec{gbest} - \vec{X_i}(t-1)\right) \tag{6.17}$$

where ϕ_1 and ϕ_2 are hyper-parameters to assign specific weight to the individual or to the social prior knowledge hence the designer is able to put more or less stress on each one these factors. Lastly, The random numbers are sampled from an uniform distribution between 0 and 1.

The first component of the sum in equation 6.17 is known as *inertia* or *momentum*. It models the tendency to continue in the same direction for every particle. The second component describes an attraction towards the best individual position which is called *memory* or *self-knowledge*. Finally, the third component is an attraction towards the best global position which is known as *social knowledge*.

As for the particle's update, it is a process dependent of the velocity calculation which allows for a new position to be obtained as,

$$\vec{X_i}(t) = \vec{X_i}(t-1) + \vec{v_i}(t) \tag{6.18}$$

Algorithmic wise these functions are time dependent which is equivalent to iteration steps, therefore every current state depends on the prior positions and velocities.

Though this simpler version of the velocity formula is a kick-starter to introduce the algorithm and its components, efforts have been made by different authors to improve convergence and the effectiveness of the algorithm. One of the first major modifications introduced by Shi and Eberhart was the so-called *inertia weight* ω.

$$\vec{v_i}(t) = \omega \times \vec{v_i}(t-1) + \phi_1 rand_1 \cdot \left(\vec{pbest_i} - \vec{X_i}(t-1)\right) + \phi_2 rand_2 \cdot \left(\vec{gbest} - \vec{X_i}(t-1)\right) \tag{6.19}$$

which is called *canonical PSO*. A modern version of the algorithm proposes an alternative definition for velocity, which achieves a higher convergence rate (Clerc and Kennedy, 2002),

$$\vec{v_i}(t) = \chi \left(\vec{v_i}(t-1) + \phi_1 rand_1 \cdot \left(\vec{pbest_i} - \vec{X_i}(t-1)\right) + \phi_2 rand_2 \cdot \left(\vec{gbest} - \vec{X_i}(t-1)\right)\right) \tag{6.20}$$

These modifications stem from a number of researches to face some undesirable dynamical behaviors of the traditional version of the algorithm. One of the most noticeable conducts is the *explosion* of the swarm if velocities are not limited to control trajectories. Thus, the use of χ coefficient is recommended by Clerc and Kennedy (2002) for particles to take smaller steps. Also, the velocity clamping by setting strict bounds is used to prevent the extreme large steps hence avoiding the swarm divergence (Parsopoulos and Vrahatis, 2010).

6.2.2.2 Hyper-parameters selection

To properly implement the PSO, several factors must be considered to favor convergence and avoiding explosion of the swarm. Those considerations are expressed in a set of configuration parameters that need carefully selected.

Regarding the constriction and inertia factors, solutions have been proposed by Van Den Bergh and Engelbrecht (2006) and Clerc and Kennedy (2002) in which, depending on the kind of selected velocity equation, the values for ω or χ are dynamically assigned, whether by using random generated numbers to introduce a chaos

Table 6.1: PSO hyper-parameters.

Parameter	Symbol	Description	References
Inertia or Constriction factor	χ	Control convergence time	Shi and Eberhart (1998), del Valle et al. (2008)
Velocity max. limit	V_{max}	Setting the velocity limit to avoid explosion of the swarm	Parsopoulos and Vrahatis (2010)
Social coefficient	ϕ_2	Relative importance of social behavior to calculate next position	Wang et al. (2018), Carlisle and Dozier (2001)
Self-knowledge coefficient	ϕ_1	Relative importance of individual best performance.	Wang et al. (2018), Carlisle and Dozier (2001)

mechanism into the system or by having a calculation rule. Despite of how carefully the number is theoretically obtained, the designer must fine-adjust these values to avoid premature convergence hence reaching local minima. Shi and Eberhart (1998) suggested to set the parameter to [0.9, 1.2] and a linearly time-decreasing inertia weight to significantly enhance the PSO performance. Further researches suggested the use of quadratic functions to adjust the parameter along the iterative process (Tang et al., 2011). In general, the adoption of an adaptive mechanism that reduces the inertia as the iterative process advances, has a significant impact in the positive performance of the algorithm.

As for the learning factors or social and self-knowledge coefficients, let us recall that their function it to pull particles toward the particle best and the global best. These values have been obtained mainly through empirical evidence. They are usually selected to be 2.0 but experiments have proved that setting values to $\phi_1 = 2.8$ and $\phi_1 = 1.3$ can be beneficial for certain types of optimization problems (Carlisle and Dozier, 2001). An interesting work is presented by Mashayekhi et al. (2019) in which a parametric study is performed for different scenarios of a simulated optimization problem.

Finally the velocity limit must be set. The selection of this parameter is not as straight forward as one may think due to the lack of knowledge regarding of how good a choice is. Moreover, if we use a dynamic inertia factor, the velocity limit should vary through the iterative process as well. Furthermore, a velocity limit must be set for each variable involved in out optimization problem, therefore the velocity limit is an n-dimensional array, not just a single value to govern the dynamics of each particle. A common choice is to use the available range of each variable to set velocity within those limits, for instance, if a variable x_1 has a $x_1max = 10$ and $x_1min = 3$, its dynamic range is 7 thus the velocity is maintained in that range,

$$V_{max} = \frac{X_{max} - X_{min}}{2} \tag{6.21}$$

Other alternatives are open such as having a dynamically reduced velocity according to the success of search or linearly decreased with time, however, there is situation that need to be looked after that is avoiding particles flying out of the physical solution space.

6.2.2.3 Absorbing walls

In this research we introduce a novel implementation of a method call **absorbing wall** or **bouncing technique**. Though this method is implemented in the optimization of electromagnetic systems (Robinson and Rahmat-Samii, 2004), but it is the first time is has been implemented to optimize a quadrature hybrids for radio astronomy applications.

Let us consider an initial point in a two-dimensional search space, $X = [x_1, x_2]$. As we know, the variables have a range defined by its max and min values. It is possible that a new position obtained by the particle

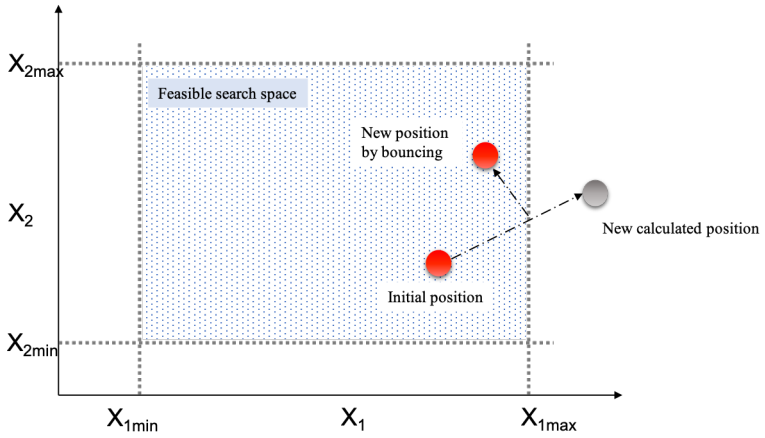


Figure 6.5: Reflecting wall or bouncing method.

position update equation 6.18, throws the particle outside of the feasible region, which is called *invisible wall*. An immediate solution is to restrain the particle to leave the search space by setting the new position to the boundary values x_{1max}, x_{2max}, x_{1min} or x_{2min}, if the calculation overcomes the limits. This solution harms the free movement of the particle around the search space hence affecting the convergence of the algorithm.

There are three approaches to allow the point (ball) to return into the search space thus keeping the exploration inside the feasible region. The first is called *absorbing walls* which sets the velocity to zero for the dimension which hits its boundary namely upper or lower limit. The boundary absorbs the energy of the particle and eventually during the iterative process, the velocity will change in the opposite direction.

A second approach is the *reflecting wall*. When the particle hits the boundary, the sign of the velocity is immediately changed in that dimension so the particle is reflected back into the solution space.

This research implements the third approach known as *damping boundaries*. In this case the boundaries behave like dampers which absorb energy but reflects the ball in the opposite direction with a smaller velocity magnitude. As it seems, this is a hybrid solution between the absorbing and reflecting methods.

7

PSO Architecture for Quadrature Hybrid Optimization

The on-going effort to develop a quadrature hybrid that fully covers the ALMA Band 2+3 is a driver for this work to take place. Many parties got involved but, in particular, the work done by Restrepo and Lucero (2020) must be acknowledged. They provided an initial implementation of the PSO for a blade and bow-tie antenna intended to operate in the *Mapper of the IGM Spin Temperature* (MIST) experiment. From there, improvements and new techniques where applied to fulfill the quadrature hybrid requirements and to improve the algorithm's performance.

- The code was developed to fulfill the coding best practices, including modularity, attending the DRY principle and avoiding of memory leaks.
- Python 3 was the language of choice not only to have a faster development process but for the available tools to apply posterior data analysis as well; this is a key activity to implement the *surrogate-based optimization*.
- A fitness function is proposed in which multiple optimization objectives and constraints are condensed. The function was constructed using a *weights* transformation and a *Tikhonov*-like penalty term.
- A fully functional version of the code can be found at this github repository. The user must have a fully operational version of the ANSYS HFSS commercial software to perform the simulations during the optimization process.

7.1 Surrogate-based meta-heuristics and particle encoding

Before getting into details of the classes and coding structures used along the solution, we want to introduce the global optimization framework which contains within itself the PSO algorithm.

Surrogate-base optimization (SBO) technique is based on the use of surrogate models to speed up the finding of optima points. At some point, other algorithms are considered as sub-optimization as the surrogate model allows for a faster but raw selection of the feasible solution space and, through and iterative process, the selection in refined by adding information from previous optimization batches. Although the method can be coded and automated a "manual" version can be applied by reviewing the optimization results and analyzing them to fine adjust parameters.

The central concept is the one of *surrogate model* which consists of a raw approximation of the model to optimize. In our research the surrogate is a raw model of the quadrature hybrid represented by three vectors namely the nominal model, maximum and minimum ranges.

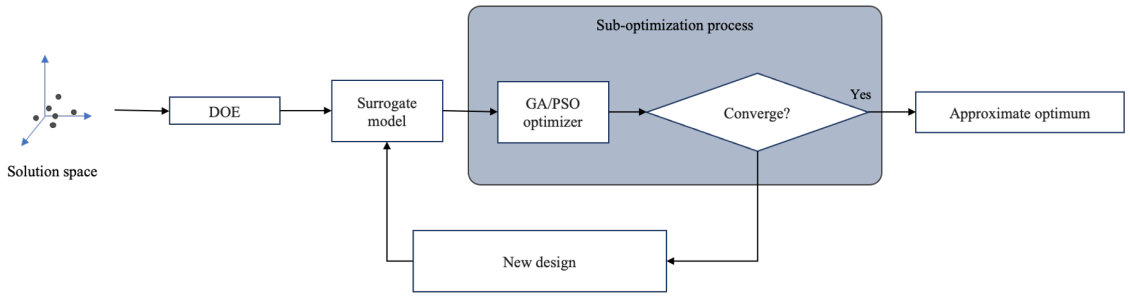


Figure 7.1: Surrogate-based meta-heuristics.

Every quadrature hybrid design is described a nominal vector whose entries are the physical dimensions to reproduce it at every point during of the optimization process. Let us consider the quadrature hybrid depicted in Fig. 7.2. This model is not necessarily supposed to be a hybrid that fully complies with all the requirements for our problem but it is a auxiliary model to search for an optimized version. The model is describe by a vector which, in fact is the particle containing the n-dimensional information of our optimization problem.

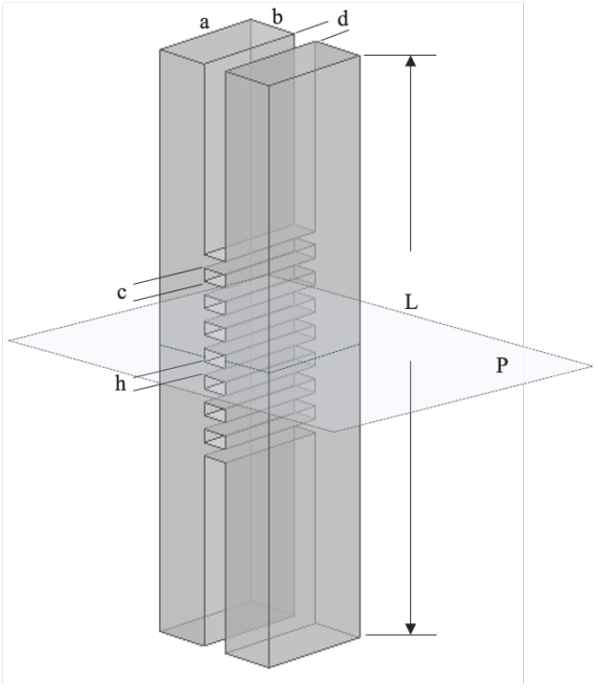


Figure 7.2: Meta-model for a quadrature hybrid.

There are a number of ways to write a vector that could represent our geometry, depending on the attributes of our surrogate model. For instance, let us conceive a quadrature hybrid whose branches and separations are all the same, therefore a vector to describe the geometry could take the following forms,

$$particle = [a, b, c, h, d, L] \tag{7.1}$$

$$particle = [a, b, d, c, c, c, c, c, c, c, h, h, h, h, h, h, h, h, L] \tag{7.2}$$

In a highly symmetric hybrid in which the periodic structure is recurrent all along the device, these vectors should be enough to reconstruct it. If want to try more flexible geometries in which all branches have a different

height, then the particle vector must incorporate that new information in some way, for example,

$$particle = [a, b, d, c, c, c, c, c, c, c, h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, L] \tag{7.3}$$

We still have to account for three factors to fully define a quadrature hybrid in terms of an optimization object:

- **The analytical model**: In Chap. 4 the quadrature hybrid was fully defined thus we must take into account all prior knowledge at this step of the optimization process. The end-to-end symmetry is critical to obtain a feasible quadrature hybrid, hence the branch height for the opening and closing branches must be the same. In fact, the whole device must be symmetric with respect to the plane P showed in Fig. 7.2.
- **The simulation tool**: In this research we have used the ANSYS HFSS software as the tool to simulate every proposed design encoded in the particle vector. The software must have an interface through which our algorithm can easily pass the vector and to receive the simulated results back. In our case, a base model must be created beforehand to work as a template on which the dimensions are put in to draw the proposed geometry.
- **The range of the variables**: The maximum and minimum ranges must be encoded in vector form as well. Furthermore, in our surrogate-based process, these vectors become the key object to put all the knowledge gained in every optimization batch. For example, a max vector for our proposed particle 7.1 takes the form,

$$x_{max} = [a_{max}, b_{max}, c_{max}, h_{max}, d_{max}, L_{max}] \tag{7.4}$$
$$x_{min} = [a_{min}, b_{min}, c_{min}, h_{min}, d_{min}, L_{min}] \tag{7.5}$$

  As the reader may notice, these vectors are the key resource to establish the hyper-dimensional boundary to apply the absorbing wall technique, moreover, they are the place where search space regions are defined. Recall that, as we refine the selection of portions of our search space, we are getting closer to global minima while simultaneously reducing the computational cost of optimization process

After defining our particle vectors, the surrogate-based model operates as showed in Fig. 7.1. An initial model is submitted to be sub-optimized by the PSO algorithm, which throws a resulting quadrature hybrid design after a full optimization batch. Therefore, the SBO assists the core PSO optimizer in narrowing the search space but the heavy lifting is done by the PSO algorithm as showed in Fig. 7.3. This graph resumes the optimization process stemming from the more general PSO flow introduced in Fig. 6.4 but customized to our needs.

After a full optimization batch, the final design must be reviewed by an expert to assess the quality of the resulting hybrid as the goal is to provide a refined surrogate model for the next full optimization batch of the PSO algorithm. To achieve this goal, the expert must adjust the nominal model but, higher priority must be assigned to modify the max/min vectors to produce a narrower search region for the PSO. This may seems difficult at first, but as the number of optimization runs increases, so does the experience of the expert regarding the particular optimization problem.

7.1.1 Core PSO process

To implement the PSO every optimization problem must be assessed to precisely understand how every step applies. Optimizing an electromagnetic system such as the power dividers and antennas, requires a mean to obtain electromagnetic performance information of every model proposed by the algorithm. As the reader may notice, these performance measure will change from one problem to another, furthermore, the fitness function will be certainly different. As for electromagnetic and microwave devices the designer has two alternatives either a) to develop an analytical model of the device and hard-code its own simulation tool or b) to use a commercial

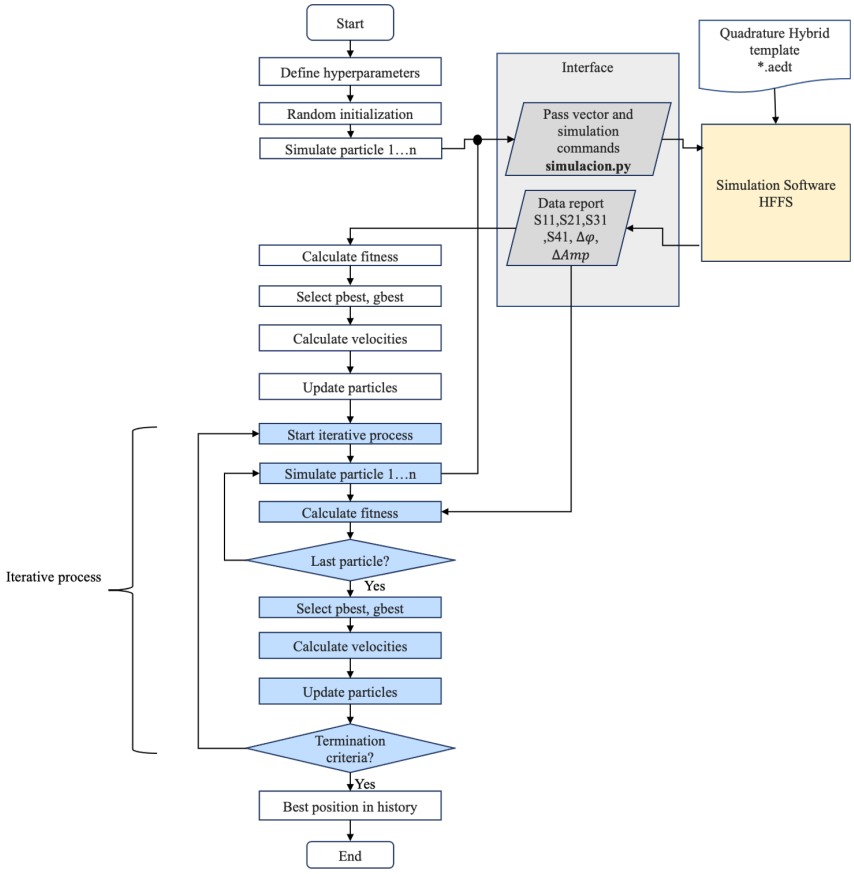


Figure 7.3: Quadrature hybrid optimization architecture using PSO algorithm and HFSS simulation tool.

simulation tool.

In this research we adopted the second approach by using ANSYS HFSS simulation software. The software provides an interface that enables users to run python scripts to share information back and forth, moreover, it allows an external execution control which in our case is very convenient to manage the launch of every simulation and to receive back specific reports like S_{11}, S_{21}, S_{31}, S_{41}, ΔAmp and $\Delta\phi$.

The optimization process is divided in two stages. The first stage is a single iteration through all randomly generated particles to submit every one of them to the simulation process and, consequently, to receive back the electromagnetic performance data. The information is shared back to the algorithm in CSV files which are used to calculate the fitness measure for every particle. Recall that the fitness will provide a measure to sort particles thus identify the best global and individual position. Finally, the velocities are calculated allowing for the new positions to be obtained.

The second stage is an iterative process of recurrent executions of simulation, evaluation, selection and particle updates. Is during this stage that the optimization takes place by a continuous cycle of mutation and application of the principles of the PSO, to be ended either by exhausting the number of pre-set iterations or by finding a convergence point.

7.2 PSO classes

The use of object-oriented programming (OOP) became one of the most salient features of the main-stream languages today. Though commonly mentioned, it is rarely consciously used being applied more as a default feature in many frameworks. Let us recall that the idea is to encapsulate data structures and procedures into separate entities which can be treated as objects. These objects can be further instantiated into *child* entities thus allowing a replications of the parent object while maintaining its features.

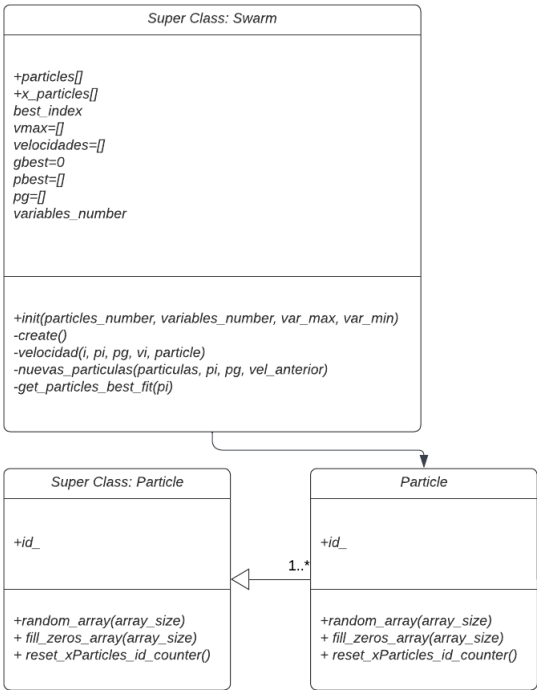


Figure 7.4: UML diagram for Swarm and Particle classes.

The PSO algorithm has a particular component that matches perfectly with the OOP paradigm namely the population-based characteristic. In this research we developed two main objects to implement a robust version of the PSO: Swarm and Particle Classes. These basic abstractions allowed a solution that easily generates and replicates particles.

There are many benefits on using this type of programming paradigm, for instance, if there is distributed-processing environment (cluster) where multiple instances of the PSO can concurrently run, we could easily have many optimization threads simultaneously executing.

There are multiple advantages in this proposed structure, most of them related to the traceability and management of the optimization process. There is self-describing relation between the swarm and the particle population as the population is treated as an array of particles. We can also track the global behavior of the PSO algorithm encapsulated in the *gbest*, *pbest* and *pg* variables. As for the particles, we can have reproducible entities with homogeneous methods in which every particle has a unique randomly-generated identifier to track it throughout the full optimization batch.

Table 7.1: Swarm and Particle Classes detailed.

Class	Method	Description
Swarm	init	Initilizes the swarm instance by defining the number of particles, the number of variables of the optimization problem. Also initializes the max/min boundaries and velocities array as well.
Swarm	create	This method is used to create the population of particles as an array of particles. It invokes the Particle class to randomly create as many particle objects are required.
Swarm	nuevas_particulas	This is the core method in which the particles are updated by using the velocity equation hence this is the place where ϕ_1, ϕ_2, χ and *damping* coefficient are used.
Swarm	get_particle_best_fit	Reports the index of the best particle from the whole population. This index is important to use it as reference the global best particle through the full optimization process.
Particle	init	A particle is an array named *values_array* which has the same dimensionality as the problem. As the particle is created the array is filled with random values taken from a uniform distribution.
Particle	fill_zeros_array	This method provides a mean to easily reset the values in the array.

7.3 Parameters

Section 6.2.2.2 gives an introduction to the PSO's hyper-parameters and its selection based on prior experimental findings. Herein is given a briefly description of the actual set of hyper-parameters and how the tuning process was applied to our specific optimization problem.

Since the hyper-parameters are mainly related to the velocity equation, we ended up working with the equation 6.19. Though tests were performed using the equation 6.20, the most promising results were obtained with the former. An computationally feasible optimization problem in which the simulation process is relatively fast namely in the range of 0.1∼10 seconds, the population size can be of the order of hundreds of particles while the number of iterations can be in the range of 50∼100. In a situation like that, there is enough room for convergence to take place but, as it happens in our case, a single particle simulation can easily take ∼250 seconds thus limiting the computational resource and therefore imposing limits to the number of particles and iterations.

In our initial simulations, every run could take around 30 iterations with a population of 10∼15 particles, but under those conditions, we had to count with a convergence control. We found this to be reachable by using the fore-mentioned equation and the parameters ω. This parameter provided a tool to have control over the convergence time.

As already mentioned, the literature suggest that this parameter could be dynamically changed through iter-

Table 7.2: PSO hyper-parameters.

Parameter	Symbol	Value
Inertia or Constriction factor	ω	$0.85 \cdot \omega_0^{iteration-\lambda} \quad w_0 = 0.9 \quad \lambda = 0.15$
Velocity max. limit	V_{max}	$\frac{3(\vec{x}_{max}-\vec{x}_{min})}{5}$
Social coefficient	ϕ_2	2.1
Self-knowledge coefficient	ϕ_1	2.0
Damping	damping	0.7

ations to favor the convergence by starting the process with a relatively high inertia a allowing the system to get looser as the iterations pass. We finally came up with an empirical equation that controls the dynamics of the convergence time,

$$\omega = 0.85 \cdot \omega_0^{iteration-\lambda} \tag{7.6}$$

where $w_0 = 0.9$ is the initial inertia coefficient to be modified as the current iteration number increases. The value $\lambda = 0.15$ can be modified to accelerate the convergence at convenience.

By using this setting, the PSO can easily reach a quadrature hybrid optimized model between the 8 and 12 iterations. This is a great accomplishment similar to the convergence obtained by Han and Zhang (2012) who used the surrogate model to optimize an airplane wing design.

As for the velocity limit, the calculation is straightforward with the help of equation,

$$\vec{V}_{max} = \frac{3(\vec{x}_{max} - \vec{x}_{min})}{5} \tag{7.7}$$

where $\vec{x}_{max}$ and $\vec{x}_{min}$ are vector containing the maximum and minimum values a variable can take in our search space. Regardless of the fraction, this equation is similar to equation 6.21, both having the same swarm-explosion issues as well.

Finally the values for ϕ_1 and ϕ_2 are showed in table 7.2, where both are the result of a trial-and-error evaluation. Although many researches support a value of 2 for both parameters, we evaluated the impact of using a different relation between them. A first trial consisted of assigning a higher weight to the social factor, furthermore, the used values were $\phi_1 = 16$ and $\phi_2 = 2.0$. The outcome was that the first proposed global solution did not allowed for other particles to explore the search space, which is an unwanted behavior. In an ideal scenario the best global particle allows for other individual to improve their own behavior and to compete between them to take the best global position.

Moreover, as we switched the emphasis in favor of the best individual performance, the particles were inclined to follow their best historical value therefore producing a very long convergence time, if ever happened. Consequently, the best performance coincided with the one suggested by the literature, with a small difference in the social parameters, which provided an extra push for particles to follow the best global particle.

7.3.1 Absorbing wall implementation

As mentioned in Sec. 6.2.2.2, this work introduces a novel implementation the **reflecting wall** or **bouncing technique** as there is no reference in the literature that suggests this method has been previously applied in radio astronomy optimization.

Algorithm 2 Absorbing wall implemented algorithm

```
1: for part ← 1 to m do
2:     for var ← 1 to n do
3:         if velocity[part][var] > V_max[var] then
4:             velocity[part][var] ← V_max * sgn(velocity[part][var])
5:         end if
6:
7:          particle_new.array[var] ← particle_old.array[variable] + velocity[part][var]
8:
9:         if particle_new[var] > x_max[var] then
10:            velocity[part][var] = damping × velocity[part][var]
11:             particle_new.array[var] ← x_max[variable] − |velocity[part][var]|
12:
13:        else if particle_new[var] < x_min[var] then
14:            velocity[part][var] = damping × velocity[part][var]
15:             particle_new.array[var] ← x_min[variable] + |velocity[part][var]|
16:
17:        else
18:            particle_new.array[var] ← particle_old.array[variable] + velocity[part][var]
19:        end if
20:    end for
21: end for
22: return particle_new,velocity
```

Our goal is to properly deal with exploding-swarm condition, which is usually contained by using a very restrictive V_{max} value not without negatively degrading the convergence of the algorithm. Thus the absorbing wall method allows for a more convenient way to deal with the particles crossing the boundaries of the search space.

In the algorithm 2 we present the detailed procedure to implement the bouncing wall. First thing to notice are the nested loops running through all particles and simultaneously running through each variable within each particle. There is V_{max} vector which contains very maximum velocity for each dimension of the search space.

The bouncing wall and the particle update operation are applied simultaneously thus every new particle is calculated based on a *damped* velocity. In the line 7 we find a first approximation to the new particle without prior knowledge of the particle leaving the search space. From line 9 through 19, each dimension of the recently updated particle is evaluated to conclude if the upper or lower boundaries are infringed. If those limits are crossed, the point is forced in the opposite direction with a reduced velocity by a *damping* factor. In case the maximum boundary of an specific variable is crossed, the particle is updated according to,

$$Particle_{new}[dimension] = X_{max}[dimension] - |velocity[dimension] \cdot damping| \tag{7.8}$$

Alternatively, if the minimum is violated, the operation to update the particle is,

$$Particle_{new}[dimension] = X_{min}[dimension] + |velocity[dimension] \cdot damping| \tag{7.9}$$

7.4 Constraints

To appropriately define the fitness function, a deeper understanding of our optimization problem is required in terms of the constraints. With explanatory purposes, some of the early optimized quadrature hybrids will be used, to depict with more detail some of the problems the algorithm will encounter frequently to evaluate the fitness of the proposed design.

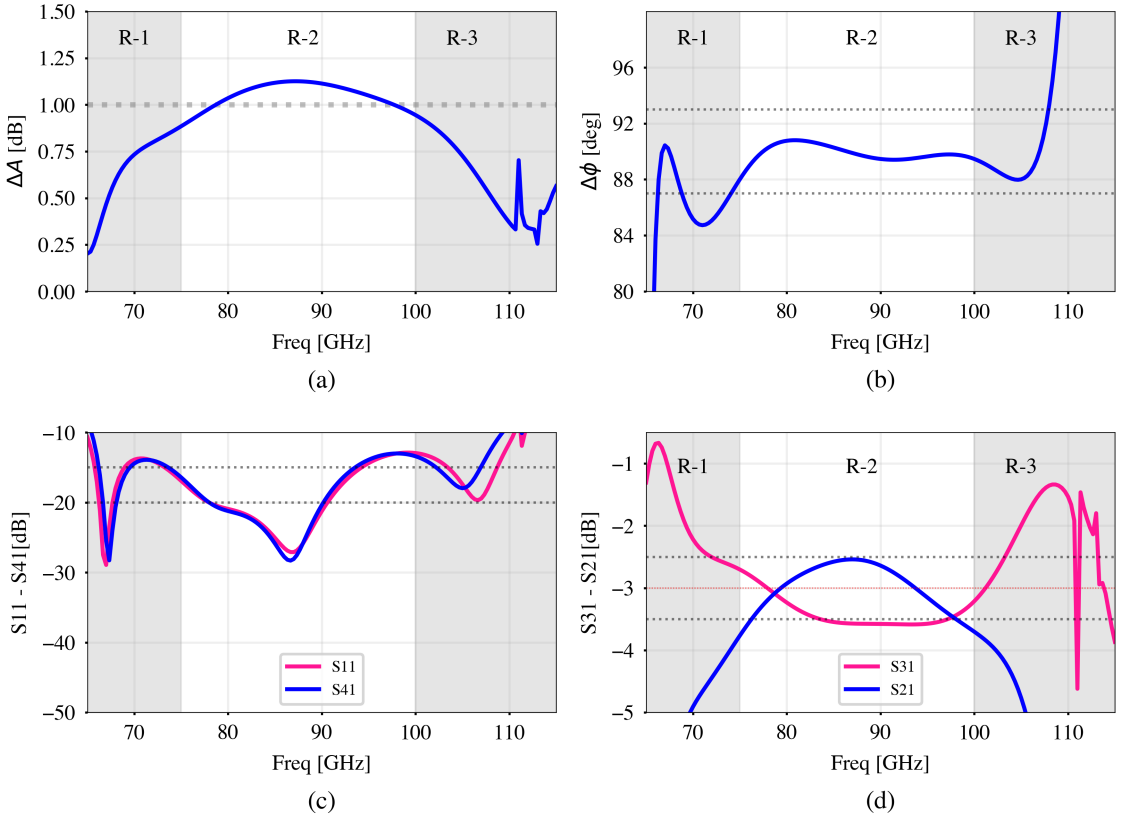


Figure 7.5: Initial Amplitude imbalance (a), phase imbalance (b) and S-parameters (c,d) for an 8-branch quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

In the Fig. 7.5, preliminary results are displayed for a quadrature hybrid in the 67 – 116 GHz Band. This plot presents the specific behaviors to be found with almost every quadrature hybrid. Let us concentrate on the plot d, where S_{31} and S_{21} are showed.

The band can be partially divided in three regions namely, lower-band (Region 1), mid-band (Region 2) and upper-band (Region 3). Every region present its own characteristic behaviors, particularly, the R-1 and R-3 regions tend to deviate significantly from the -3dB mark established in 5. Remember that our expectation is to stay in the range [-3.5dB , -2.5 dB] for the device to produce a $\Delta A < 1$. Another characteristic of these regions is the oscillatory behavior, evident in the form of undesired spikes as they degrade the performance of a microwave device that is supposed to operate with very faint signals coming from interstellar objects.

The mid-band region behaves differently as it is usually continuous and smooth. As we studied in preceding chapters, this region is seriously affected as we add branches to our hybrid with the goal of extending the band coverage. The amplitude imbalance in this region (Fig. 7.5.a) does not comply with the constraint of staying

below the 1 dB mark, thus a bigger effort must be made to keep the S_{31} and S_{21} within the [-3.5dB , -2.5 dB] gap, furthermore, as they are pushed closer to the -3 dB mark the ΔA gets lower.

Therefore, to properly represent the oscillatory and well behaved functions in three separate regions, we developed python methods to split the data for each variable in three sub-bands, namely 65-70GHz, 71-110GHZ and 111-115GHz. Though this limits were established on the basis of experience, they can easily modified to test different regions.

Moreover, in every region we evaluate five constraints. Here we provide the constraint established in mid-band region, but these values are different for the upper and lower sub-bands.

Table 7.3: Mid-band smooth and oscillatory constraints.

Constraint
-3.5dB$<S_{31}<$ -2.5 dB.
-3.5dB$<S_{21}<$ -2.5 dB.
$0<\Delta A<$ 1 dB.
-3.5 dB$<Overshoot_{S_{31}}<$ -2.5 dB.
-3.5 dB$<Overshoot_{S_{21}}<$ -2.5 dB.

An important observation is that we use neither S_{11}, S_{41} and $\Delta\phi$ as these parameters will respond the the improvement he obtain in the ΔA, S_{31} and S_{21} parameters thus all optimized candidate designs must be manually inspected to ensure the proposed quadrature hybrid does comply operate under ALMA's observatory constraints.

7.5 Fitness function

Since our optimization problem shows all the stated behaviors namely oscillation in the lower an upper regions and a smooth activity in mid-band region, the fitness function has to include them in its final form.

First and foremost we evaluate the S-parameters and ΔA vis-à-vis the pre-set constraints. For instance, we take the S_{31} measure which spans from 65 through 115 GHz, and evaluate every point of the data against the expected constraints resulting on binary results informing us if the S_{31} complies with each constraint. The same parameter is evaluated in the three regions as every one has its own set of limit values to comply with a) being within the [-3.5dB , -2.5 dB] gap and b) its value staying below the overshoot limit.

Afterwards, the penalties functions are applied as we need to punish the parameter under review depending on how much it complies with the constraints. Let us consider the same S_{31} example. If the parameters complies with the [-3.5dB , -2.5 dB] constraint along the full range of frequencies (65-115 GHz), hence the evaluated constraint would be a stream of binary '1's telling us that the figure had a perfect "score". On the contrary, if the figure has some point out of limits, the evaluation will throw a stream of '1's and '0's.

As every merit figure is evaluated using the same method, the final result is a binary representation showing how every figure do or do not comply with the established constraints. This is valuable as it provides a mean to give a quality measure of the whole figure encoded in a real value between 0 and 1, as follows,

$$penalty_{S_{31}} = 1 - \overline{E(x)} \quad ; \quad x \in Binary : E = evaluated\_constraint_{S_{31}} \tag{7.10}$$

$$penalty_{S_{21}} = 1 - \overline{F(x)} \quad ; \quad x \in Binary : E = evaluated\_constraint_{S_{31}} \tag{7.11}$$

$$penalty_{S_{31}Overshoot} = \overline{G(x)} \quad ; \quad x \in Binary : G = evaluated\_constraint_{S_{31}Overshoot} \tag{7.12}$$

$$penalty_{S_{21}Overshoot} = \overline{H(x)} \quad ; \quad x \in Binary : H = evaluated\_constraint_{S_{21}Overshoot} \tag{7.13}$$

$$penalty_{S_{AmpImb}} = \overline{I(x)} \quad ; \quad x \in Binary : I = evaluated\_constraint_{AmpImbalance} \tag{7.14}$$

Since the weights are defined to emphasize certain behaviors in specific regions, the next step is, by penalizing them, to affect how each behavior adds into the full fitness function. Let us recall that the weights and the penalties are matrices thus,

$$W = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \\ w_{51} & w_{52} & w_{53} \end{bmatrix} \tag{7.15}$$

$$\tag{7.16}$$

$$Penalties = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \\ p_{41} & p_{42} & p_{43} \\ p_{51} & p_{52} & p_{53} \end{bmatrix} \tag{7.17}$$

where the columns represent each of the three defined regions for our problem and the rows represent each of the five constraints to evaluate. A penalized version of the weights is obtained by,

$$W_{penalized} = 100 \times \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \\ w_{51} & w_{52} & w_{53} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \\ p_{41} & p_{42} & p_{43} \\ p_{51} & p_{52} & p_{53} \end{bmatrix} \tag{7.18}$$

To be consistent, this result must be normalized to keep all weights in the range $0 \sim 1$,

$$w_{max} = \max\left(W_{penalized}\right) \tag{7.19}$$

$$w_{min} = \min\left(W_{penalized}\right) \tag{7.20}$$

$$W_{penalized} = \frac{W_{penalized} - w_{min}}{w_{max} - w_{min}} \tag{7.21}$$

In despite of having obtained weights and penalties applied, there is one missing task that is to calculate the actual amount of oscillation or deviations from the target marks as these numbers are our actual measure to calculate the fitness value. Two functions are used to obtained those numbers,

$$deviation_{S_{11},S_{21}...} = MSE\left(Simulated_{S_{11},S_{21}...}, Ideal_{S_{11},S_{21}...}\right) \tag{7.22}$$

and

$$Y = \nabla^2 Simulated_{S_{11},S_{21}...} \tag{7.23}$$

$$Oscillation\_intensity = \left| \int_{f_{min}}^{f_{max}} Y dx \right| \tag{7.24}$$

The equation 7.22 calculates the Mean Squared Error for each simulated parameter against the ideal or expect value, thus a parameter like the S_{31} is expressed as single MSE real value. The same applies for parameter hence we obtain a matrix with the measure of the adjustment of each figure to the expected behavior. As for the oscillatory behavior, by using the equation 7.23 we evaluate the intensity of the spikes or the overshoot from one point to the next by derivation. The result is a data set with very acute angles where a significant change in amplitude happens. If the data in a particular region is smooth and well behaved, the final outcome must be a flat zero-centered plot, but as the data presents intense oscillations, the outcome is a number higher than 0, which represents the presence of oscillations.

This numbers are written in matrix form as,

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \\ f_{41} & f_{42} & f_{43} \\ f_{51} & f_{52} & f_{53} \end{bmatrix} \tag{7.25}$$

in which every row represents a constraint and the columns represent each region of the operational band, thus the entries are the calculated measures of fit to the established constraints.
The next step is to apply the penalized weights to the equation 7.25 therefore a new matrix is obtained in which every entry is a measure of the fitness of each feature, either amplified or damped, as follows,

$$product = W_{penalized} \cdot F \tag{7.26}$$

where the dot operation is done using the numpy's "dot" function. The importance of these values reside in the diagonal of the matrix thus a python diagonal operation is applied. As these numbers could have wild variation hence producing extremely big or small fitness values, a logarithmic function is applied.

$$log_{scaled} = \log 100 \times product \tag{7.27}$$

$$diagonal = Diag(log_{scaled}) \tag{7.28}$$

where *diagonal* is a 5-entries vector each of them representing the cost for each one for the constraints throughout the full operational bandwidth.

Since we want to segregate the effect of the S-parameters from the oscillatory behavior, we use a functional used in regularization problems called *Tikhonov* method (Queipo et al., 2005). The equation is divided in two parts, the first contains the loss or cost functions to quantify our empirical error while the second has a penalty term. In Tikhonov regularization the second term is frequently used with a derivative term that penalizes high local curves. Therefore, the first term enforces closeness to the data while the second provides information of the smoothness of the curves.

$$fitness = \sum_{i=1}^{j} Cost_i + \lambda \int \|D^m f\| dx \tag{7.29}$$

where λ is the regularization term and D^m is a m-order derivative of the function f. As we already applied a

method to evaluate the oscillatory behavior, our function reduces to,

$$fitness = \sum_{i=1}^{m} Cost_i + \lambda \sum_{j=1}^{n} Osc_j \tag{7.30}$$

In the equation 7.30 i runs through all smooth functions and j does the same for the functions that evaluate oscillatory behavior, hence our *diagonal* vector must be divided into two namely the entries that represent the cost for S_{31}, S_{21}, ΔA, and the entries that evaluate the oscillatory behavior $Overshoot_{S_{31}}$ and $Overshoot_{S_{21}}$,

$$fitness = \sum_{i=1}^{3} diagonal[i] + \lambda \sum_{j=1}^{2} diagonal[j] \tag{7.31}$$

which yields a final fitness real value. First thing to notice is that, the fitness equation allows for new costs to be included as far as they are part of the *diagonal* cost vector. In addition, after a series of trials to fine-tune the algorithm, the ultimate $\lambda = 0.008$ value is the one allowing a better mix between the smooth and the oscillatory parts of the fitness function.

7.6 Initial optimization results

In order to fine tune the algorithm an 8-branch model was selected to be our meta-model due to the available information and designs the could be used as a benchmark. This type of hybrid was studied in Sec. 4.2 and simulated in 5.2, the latter being our reference to assert that the optimization process actually works.

In the table 7.4 the optimization settings are showed. Attention must be paid to the vectors, which finally encode the geometry of the meta model and the allowed ranges to each variable that define the search space. These vectors, as mentioned in Sec. 6.2.2.2, are subjected to modification by the expert as the knowledge of the system increases thus refining the search space to speed up the optimization.

Table 7.4: PSO optimization settings.

Setting	Value
Id	ac5717cd
Iterations	20
Nominal Vector	$[0, 0.287, 0.251, 0.242, 0.129, 0.3, 1.3, 2.3, 3.2, 0.9, 2.8, 1.4]$ mm
X_{min}	$[0, 0.15, 0.15, 0.15, 0.15, 0.2, 1.0, 2.0, 3.0, 0.75, 2.6, 1.35]$ mm
X_{max}	$[0, 0.8, 0.8, 0.8, 0.8, 1.1, 2.2, 3.2, 5.0, 1.0, 3.1, 1.54]$ mm

Table 7.5: PSO optimization results.

Iteration	Particle	Array
0	0	$[0.0, 0.5269, 0.6301, 0.257, 0.278, 0.29766, 1.4875, 2.458, 3.215, 0.855, 2.7178, 1.535]$ mm
1	1	$[0.0, 0.527, 0.4270, 0.257, 0.279, 0.298, 1.488, 2.458, 3.215, 0.856, 2.718, 1.536]$ mm
9	1	$[0.0, 0.325, 0.325, 0.193, 0.193, 0.26, 1.0, 2.0, 3.0, 0.75, 2.6, 1.35]$ mm
18	1	$[0.0, 0.325, 0.325, 0.325, 0.325, 0.45, 1.0, 2.0, 3.0, 0.75, 2.6, 1.35]$ mm

After executing the full optimization batch, the obtained results are resumed in 7.5. In this table we find how the best global particle changes along the process, furthermore, the best particle in iteration 0 namely the random-generated particle, belong to the member 0, but as the iterations go by, the new best particle title belong to the member 1. Something to notice is that the best particle is changing its entries, even though the

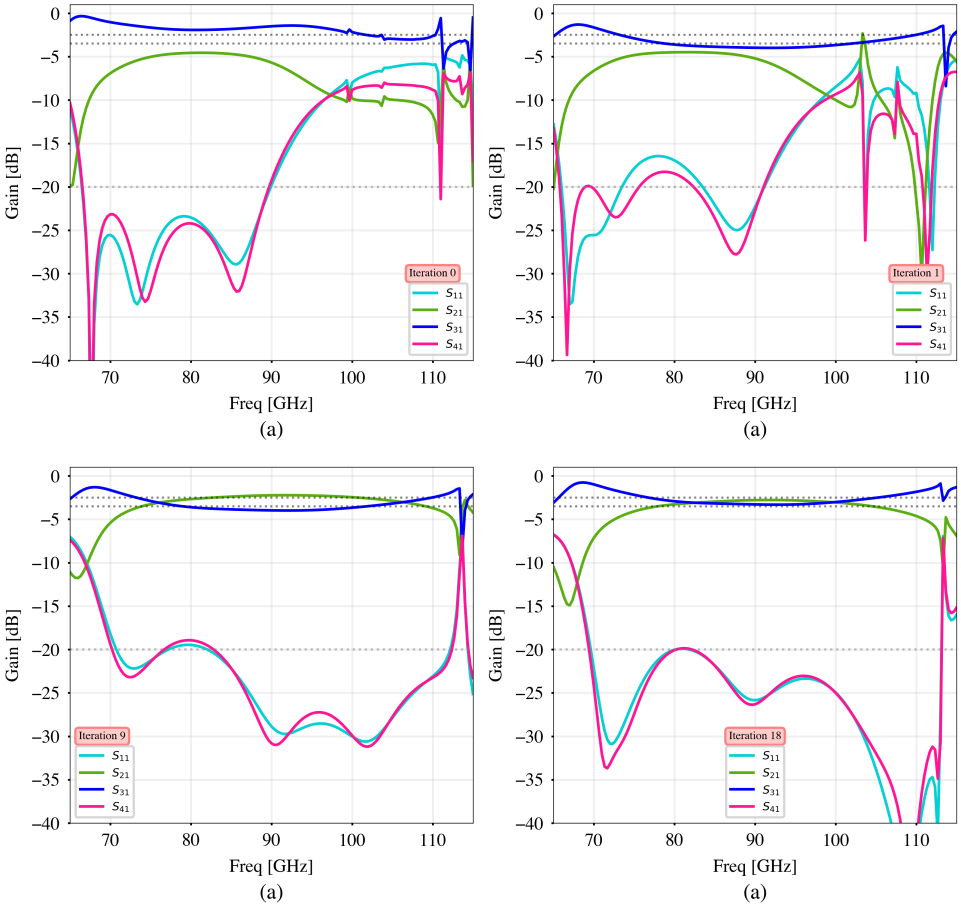


Figure 7.6: Optimization progress through iterations 0 (a), 1 (b), 9 (c) and 18 (d)for an 8-branch quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

results at iteration 9 and 18 are very close, both particles still have differences.

All these results can exposed and understood by plots 7.6 and the geometric view of the quadrature hybrids in 7.7. Fig. 7.6.a shows the S-parameters for the best global particle in the initial iteration. As the reader may notice, these results are far from the ideal behavior depicted in the Fig. 5.2. Let us recall that S_{31} and S_{21} must be as close as possible to the -3 dB mark, which is not happening for the winner particle.

Something promising happens when the particles are updated by using the process mentioned in previous sections. Both figures S_{31} and S_{21} start to move towards the established mark, though many oscillations still occur in the region 3 (upper band). The fitness function must be able to penalize this type of behavior to achieve smoother figures, which certainly happens for iterations 9 and 18, where the global behavior approaches the expectation presented in 5.2. In iteration 9, the S_{31} violates the [-3.5dB, -2.5 dB] range, something that gets fixed as the iteration 18 is reached, where a more desirable operation of the device. Moreover, the The S_{11} and S_{41} both descend below the -20 dB mark, while their shapes are getting closer, which is a desirable behavior as well.

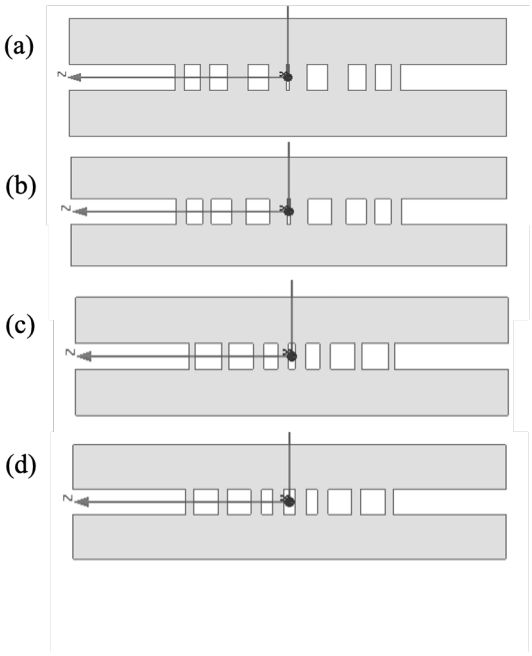


Figure 7.7: Quadrature hybrid geometry changes through the iterative process during the optimization.

Finally, the Fig. 7.7 shows the geometries represented by the particles. Fig. 7.7.a is the random generated quadrature hybrid over the 8-branch template. As the process advances, it is evident how the geometry is progressively changing.

An additional but significant finding is the fact that the obtained quadrature hybrid design has an irregular separation between branches, contrary to the reeds hybrid. The constant distance between branches is a simplification of a more general case, which our optimization process revealed without imposing an specific constraint in that regard.

7.7 PSO Validation

To validate the results obtained from the algorithm, our approach consisted in using a quadrature hybrid geometry previously studied for ALMA Band 5. The selected geometry consists of and 8-branch design with flat main waveguides.

As we will show, the resulting optimized design greatly improved amplitude imbalance when compared to the results presented by Billade et al. (2011). Hence, the operation of the algorithm and the fitness function as well, were validated. The trials demonstrated that the algorithm can effectively enhance the figures from other quadrature hybrid designs.

7.7.1 Band-5 hybrid benchmark

Billade et al. (2011) presented a first production cartridge for ALMA Band 5, covering frequencies from 163 to 211 GHz. The device is a side-band separating heterodyne receiver which involves the use of a 8-branch RF hybrid with the goal of a broadband performance. The hybrid used for this receiver had equally-spaced branches ($380\mu m$) and equally-sized branches ($175\mu m$).

Since our intention is to prove the effectiveness of the proposed method to generate optimized designs under certain constraints, including the operational band. Though our objective band is different, the algorithm should still provide feasible solution under a different regime.

The structure of the particle vector is defined on the basis of the 8-branch geometry presented in the Fig. 7.8, in which every entry represent a free geometric parameter. Thus a vector we provided to HFSS through the interface has the following form,

$$particle = [0, b_1, b_2, b_3, b_4, d_1, d_2, d_3, d_4, h, a, b]\,. \tag{7.32}$$

The b_x entries correspond to the size of every branch and, as you may notice, despite working with an 8-branch hybrid, four dimensions are enough due to the end-to-end symmetry. The same situation happens with the location d_x for each branch, which is measured from the middle plane. The length L is constant throughout the optimization process and is equal to $6.482mm$.

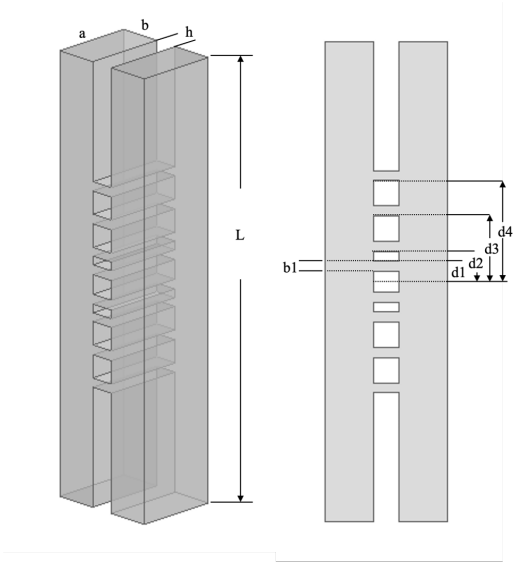


Figure 7.8: 8-branch quadrature hybrid particle structure.

7.7.2 Optimization results

After applying the optimization process introduced in former sections, we obtained a geometry that improves the results presented by Billade et al. (see Fig. 7.9). There are some remarkable outcomes worth of mention.
The amplitude imbalance ΔA is greatly improved as the optimized version in under 0.5 dB while the initial design has an imbalance over 0.75, which is equivalent to a 33%.The phase imbalance $\Delta\phi$ is also well behaved as it remains within the desired gap of $\pm 1°$. Let us recall that there is compromise between these two variables thus while one improves its counterpart degrades.

It is evident that the S_{31} and S_{21} are closer to the -3 dB mark than its benchmark counterpart. The S_{11} and S_{41} behave as expected as both figures are very similar. This happens in our optimized hybrid, not without a consequence resulting from the improved imbalance; ideally, these figures should stay below the -20 dB mark, but in our case, they are moved upwards when compared to the benchmark figures. This is the result of a modified geometry that improves the amplitude imbalance, but not without a cost. In the figures 7.8 and 7.10 is presented the resulting design, which is expressed in vector form as,

$$[0, 0.142, 0.128, 0.143, 0.134, 0.214, 0.47619, 0.95238, 1.42857, 0.35714, 1.2380, 0.6428] \quad \text{millimeter} \tag{7.33}$$

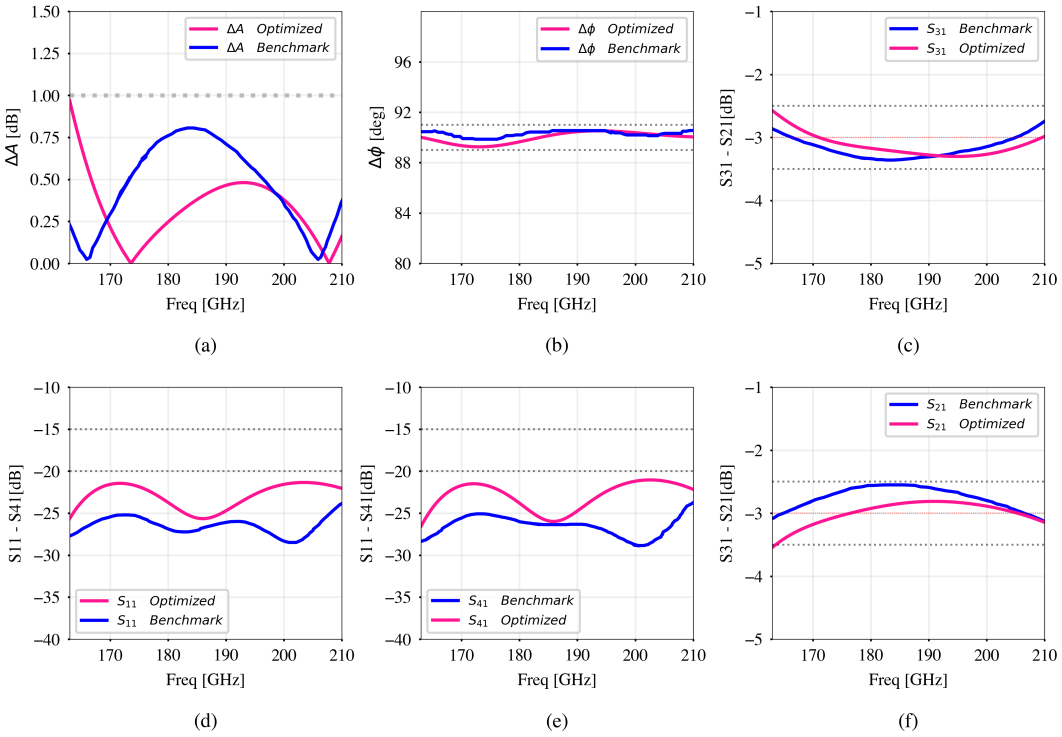


Figure 7.9: 8-branch quadrature hybrid optimized for ALMA Band 5.

Finally, the resulting design is very appealing for two reasons. First, the separation of the branches are all different hence breaking a very symmetric traditional design. The same happens with the size of the branches as, in our design, all are different. Our design uses completely flat main waveguides which is an advantage when thinking about fabrication costs.

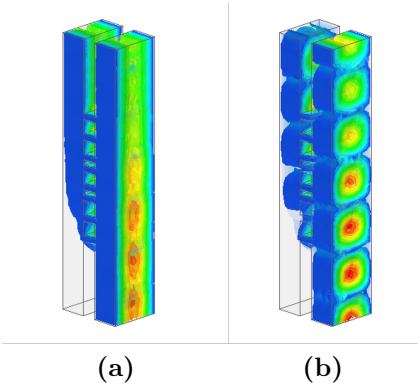


Figure 7.10: Simulated (a) E-field and (b) H-field for an optimized 3-dB 8-branch quadrature hybrid for ALMA Band 5.

8

Quadrature Hybrid Optimization

At this point we have studied the physics behind the directional couplers, modeled and simulated a quadrature hybrid and provided evidence of an effective optimization process for quadrature hybrids compliant with ALMA's performance requirements.

Multi-branch couplers are widely used in heterodyne receivers, but special attention has been paid to standard designs as those studied so far in this document. Thus far, Srikanth and Kerr (2001) designed and fabricated eight-branch model which proved to comply with the requirements for ALMA Band 3. Still, the problem remains on how to extend the operational band while complying with all constraints, hence it has been proposed by Gonzalez et al. (2017) to increase the number of branches or to use overmoded waveguides to extend the operational band. Further significant modifications have been introduced to the hybrid geometry by drilling extra cavities to the main guides to modify the impedance along the main guides and, consequently, to add ripples to the signal and consequently reducing the amplitude imbalance at the center of the band (Ding et al., 2019; Rashid et al., 2016; Hamid and Yunik, 1967).

In this section we present the optimization results of applying our method to 8-, 12- and 14-branch geometries, while also including extrusions in the main waveguide to assess their impact on the amplitude imbalance. Additionally, a fabrication constraint has been considered as the branches get tighter as a consequence of adding branches to extend the operational band. The size of the branch sets a fabrication limit as the use of standard drilling tools is highly desirable for cost reduction and ease of fabrication for other small-size application. For specific radio astronomy receivers which require cost-effective devices, the access to standard fabrication facilities and tools is of great help to speed up the design and commissioning process.

This fabrication restriction to allow the use of standard drilling tools, is defined as the ratio between the height of the main waveguide and the width of the branches (≤ 7) according to

$$Relation = \frac{a/2}{branch_{size}} \tag{8.1}$$

where the variable a corresponds to the main waveguide width. The 1/2 factor is present due to the fabrication technique called "split block" in which the hybrid is split in two halves while the drilling process takes place.

Thus, as we obtain different quadrature hybrid prospects, the relation must be kept. To facilitate the optimization process, the drilling limit is included as an added tolerance quantity to the x_{max} and x_{min} vectors used in the optimization process. Therefore, the allowed upper and lower limits for the branches are higher. This restriction imposes an actual physical limit to the process of adding branches to widening the operational band hence it is possible that the standard quadrature hybrid geometries herein studied are not sufficient to achieve a 100% band coverage for the ALMA Band 2+3 (67-116 GHz).

Let us begin by recalling that as we extend the operational band by adding branches to our design, the amplitude imbalance degrades at mid-band thus the addition of ripples to the S_{31} and S_{21} to produce a intertwined waves thus lowering the amplitude imbalance is proposed and studied by Ding et al., achieving a $\Delta A = 0.8$dB in W

Band (75 to 110 GHz). A similar approach is proposed by Rashid et al. with an ultra-low amplitude imbalance for the 159 - 216 GHz which easily covers the ALMA Band 5 (163 – 211 GHz). Their design uses waveguides discontinuities placed symmetrically along the main guides to achieve the ripple effect.
Therefore, the vector structure of each particle must vary according to the geometry under review. In Fig. 8.1 we show a proposed vector structure involving the cavities drilled in along the main guides, moreover, the dimensions e_x correspond to the depth of the extrusion in the outwards direction.

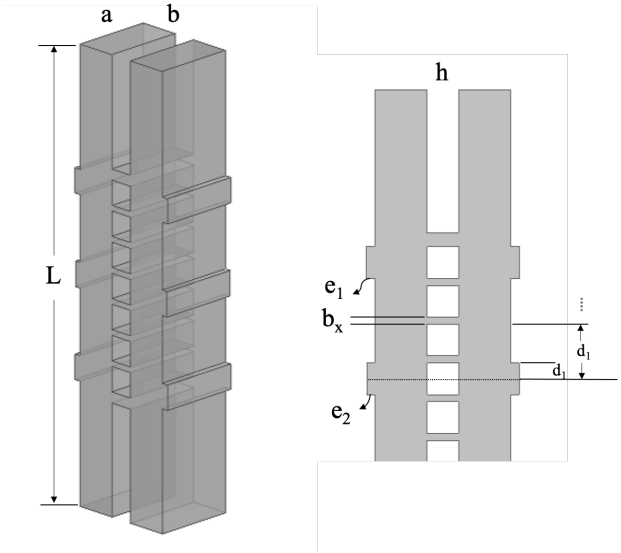


Figure 8.1: 8-branch quadrature hybrid particle structure for an extruded main guide design.

As for the particle's vector form, the b_x dimension correspond to each branch size, thus the particle looks like,

$$particle = [0, b_1, b_2, b_3, b_4, d_1, d_2, d_3, d_4, h, a, b, e_2, e_1]\,. \tag{8.2}$$

Alternatively, we described the geometry by using a second vector structure (see Fig. 8.2). Conversely, in the previous cases the position for each branch is measured from the middle plain, with this representation we provide directly the separation between adjacent branches. Though the vector looks exactly the same, the meaning behind each variable and how the HFSS interprets the vector to reconstruct the geometry is very different.
In despite of the great volume of tests and optimization results gathered following the structure presented in Fig. 8.3, we provide the most representative and significant quadrature hybrids. As the reader may notice, different number of branches and operational bands with the goal in mind of finding the best quadrature hybrid for each band of interest.

Multiple optimization runs were performed with two objectives namely to find optimized geometries for quadrature hybrids intended to operate in ALMA Band 2+3 or ALMA Band 3. Two main categories were used to discriminate the hybrids under test: *Flat main guides* and *Extruded main guides*; thus optimization runs were set to appropriately use vector describing those geometries.

The use of 8 branches as lower limit and 14 as the higher is due to prior information of the former being the most tested and implemented geometry throughout the literature, while the latter constitutes, as we will show, a constructive limit as the branches get so narrow that the fabrication relation restriction is not met.

Finally, the reader will find tables with brief information of the geometry and the optimization results, including a) the optimized particles which dimensions are given in millimeters, b) the fabrication ratio (see Sec. 8) and c) the directivity for the operational band. Supplementary information is provided in the appendix which contains

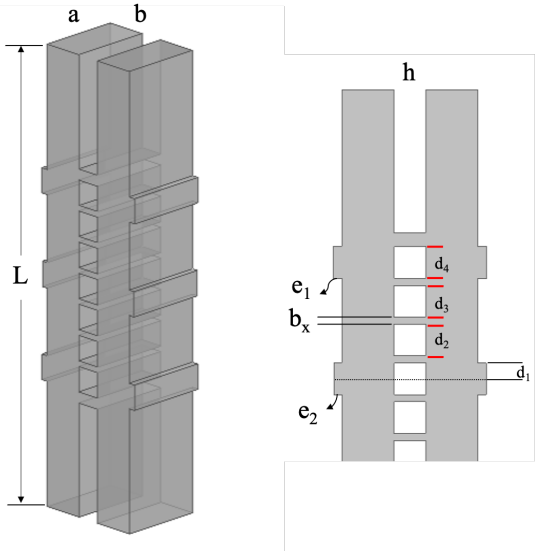


Figure 8.2: 8-branch quadrature hybrid alternative particle structure.

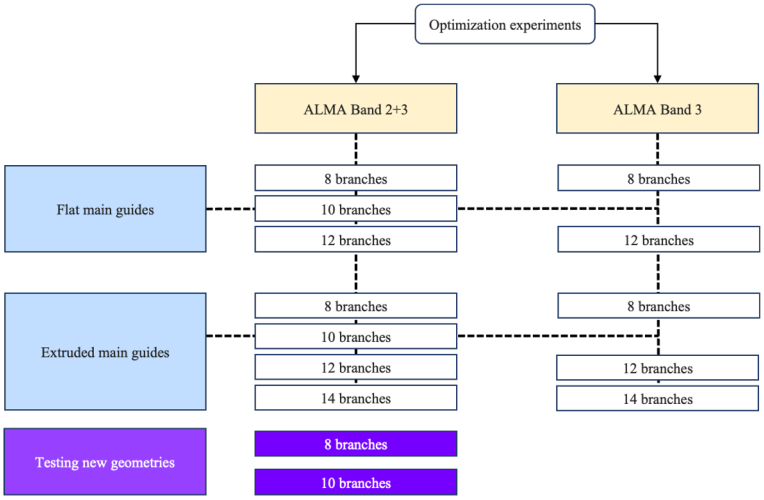


Figure 8.3: Taxonomy of the optimization experiments.

the graphical display of the simulated geometries and its fields

8.1 ALMA Band 2+3

8.1.1 8-branch hybrid

The optimized geometry and the vector structure coincides with the one presented in the Fig. 7.8 with the difference of being set to work on the ALMA 2+3 Band.

Table 8.1: 8-branch hybrid uniform main guides.

Setting	Value
Id	b8dc6220
Band	ALMA 2+3 (67-116 GHz)
Operational Band	75.66 - 107 GHz
Optimized vector	$[0, 0.264, 0.260, 0.264, 0.274, 0.245, 1.229, 2.260, 3.491, 0.741, 2.611, 1.305]$ mm
Fabrication Ratio	$a = 2.611, b_1 = 4.95, b_2 = 5.02, b_3 = 4.95, b_4 = 4.76$
Directivity	> 22.5 dB

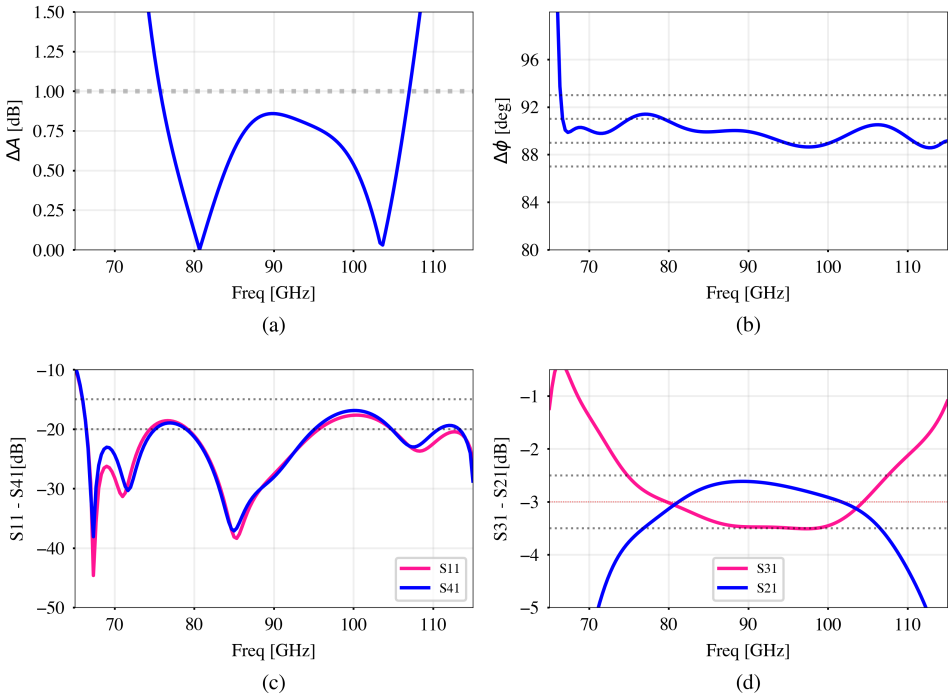


Figure 8.4: 8-branch quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

The final design showed in Fig. B.1, presents the optimized hybrid with flat main guides, unequal branch sizes and non-uniform distance between branches. As for the physical dimensions, the fabrication ratio is well below the limit of 7 which means that this proposed geometry can be easily fabricated with standard drilling tools.

The Fig. 8.4 shows an interesting result as the ΔA remains below the -1 dB mark in the 75-106 GHz range. The S_{11} and S_{41} show a good performance in the sense of they staying below the -15dB while their waves behave similarly. Finally, though the great part of $\Delta\phi$ stays within the desired range, a crossing occurs at some specific points.

This quadrature hybrid is still a narrow-band device, not fully compliant with ALMA Band 2+3 constraints, however is valuable hybrid as it can be re-scaled resulting in device useful to operate in other bands.

To reach an improved amplitude imbalance, as learned in prior sections, a design is introduced by adding extra cavities to the main guides with the purpose of creating ripples to make mid-band S_{31} and S_{21} waves tighter to the -3 dB constraint hence reducing the ΔA. This design still has 8 branches thus the expected optimized hybrid must resemble the previous with a lower ΔA.

Table 8.2: 8-branch hybrid extruded main guides.

Setting	Value
Id	3ef9a9f4
Band	ALMA 2+3 (67-116 GHz)
Operational Band	77.66 - 110 GHz
Optimized vector	$[0, 0.254, 0.278, 0.2614, 0.307, 0.786, 0.7678, 0.7875, 0.778, 0.766, 2.515, 1.25, 0.19, 0.1859]$mm
Fabrication Ratio	$a = 2.515, b_1 = 4.95, b_2 = 4.52, b_3 = 4.81, b_4 = 4.1$
Directivity	$>$ 20 dB in the range of 66.6 - 108 GHz

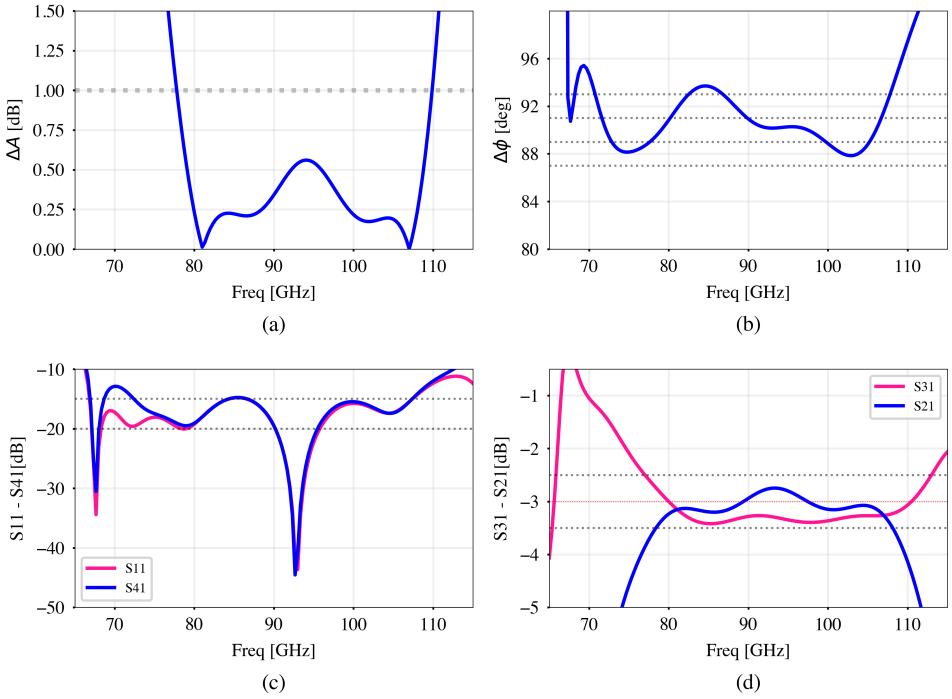


Figure 8.5: 8-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

As depicted in Fig. 8.5, evidently the expected ripple is added therefore the amplitude imbalance gets lower. Although this behavior es desirable, there is a trade-off between ΔA and other figures namely $\Delta\phi$ and $S_{11} - S_{41}$.

Adding the ripple has a negative impact in $\Delta\phi$ producing a parameter that is way out of the gap. As for the $S_{11} - S_{41}$, the wave are very close to each other at mid-band but the -20 dB mark is compromised. This result clearly shows the difficulties of improving one figure while degrading the others.

A valuable insight is that an 8-branch flat hybrid would be a more feasible device as it has simpler geometry thus reducing the fabrication cost.

8.1.2 10-branch hybrid

An important part of the work is to further evaluate the impact of adding branches to the hybrid design. In this part we introduced 2 additional branches and proceeded to optimize the design.

Table 8.3: 10-branch hybrid extruded main guides.

Setting	Value
Id	2d60535f
Band	ALMA 2+3 (67-116 GHz)
Operational Band	75.66 - 104 GHz
Optimized vector	$[0, 0.207, 0.3454, 0.7823, 0.721, 2.766, 1.355, 0.1046, 0.1216]$mm
Fabrication Ratio	$a = 2.766, b_{inner} = 6.68, b_{outer} = 4.0$
Directivity	> 21.1 dB

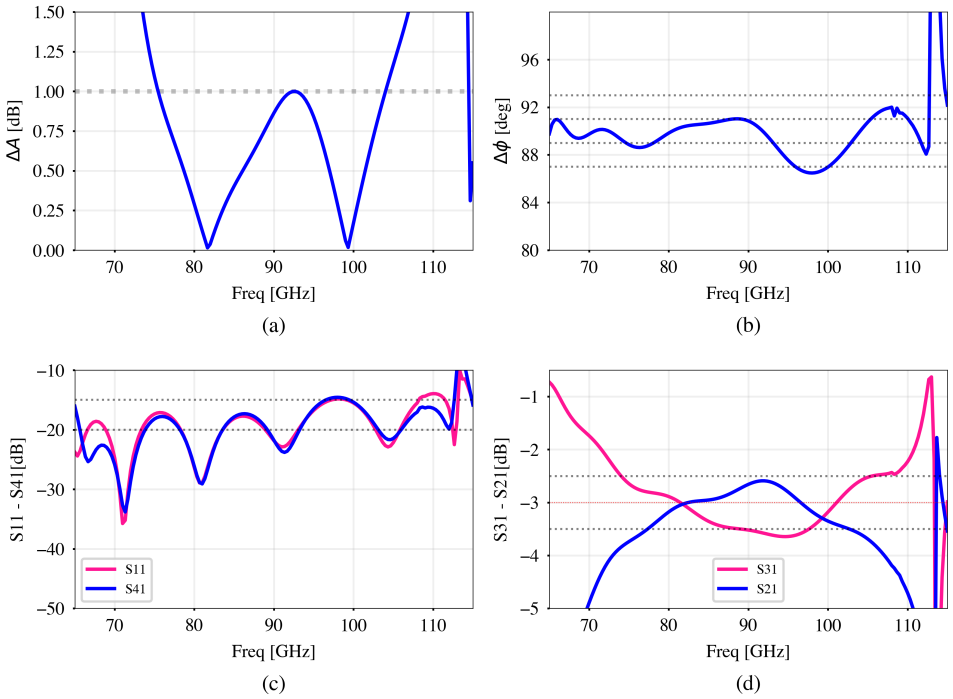


Figure 8.6: 10-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

A couple of relevant things can be noticed. First, the ripple is effectively introduced but the effect is not as considerable as in the previous 8-branch case. Second, the device still has a narrow band, moreover, the addition of just two branches do not provide a significant increase in the band coverage.

With this in mind, we proceed to optimized quadrature hybrids with a greater number of branches in order to find out if there is a feasible hybrid to cover ALMA 2+3 Band. Recall that, as we add branches, the design requires to have smaller branch sizes, thus a fabrication limit is set.

8.1.3 12-branch hybrid

Since the preceding hybrids remain in the narrow-band regime, further modification of the geometry must be introduced which in this case means the addition of two more branches.

At this point, it is clear that, while adding branches to extend the band, other parameters are negatively impacted namely the ΔA. This has been partially solved by adding ripples to waves through main waveguides modifications. Optimizing devices with flat main guides with a great number of branches do not provide meaningful results, therefore optimizing the extruded designs is a better research objective.

In this section we make use of two different ways to describe the hybrid's geometry; the first is, as we call it, the Rashid-Reed vector due to the way these two authors treat the free parameters of the hybrid, and the second, is a "flexible vector" design. The former keeps the separation between branches all the same along the hybrid's length, while the former allows for different separation size, hence this is equivalent to the vectors used before.

The tables 8.4 and 8.5 contain the optimized vectors, in which the Rashid-Reed vector structure can be appreciated. The first entry corresponds to the size of inner branches while the second corresponds to the size of the outer branches. The third entry is the separation between branches. The remaining entries are all the equivalent to the used in equation 8.2.

Table 8.4: 12-branch hybrid extruded main guides.

Setting	Value
Id	8df72fd2
Band	ALMA 2+3 (67-116 GHz)
Operational Band	80.66 - 113 GHz
Optimized vector	$[0, 0.187, 0.323, 0.724, 0.743, 2.58, 1.29, 0.185, 0.1730]$mm
Fabrication Ratio	$a = 2.580, b_{inner} = 6.9, b_{outer} = 3.99$
Directivity	> 21.5 dB

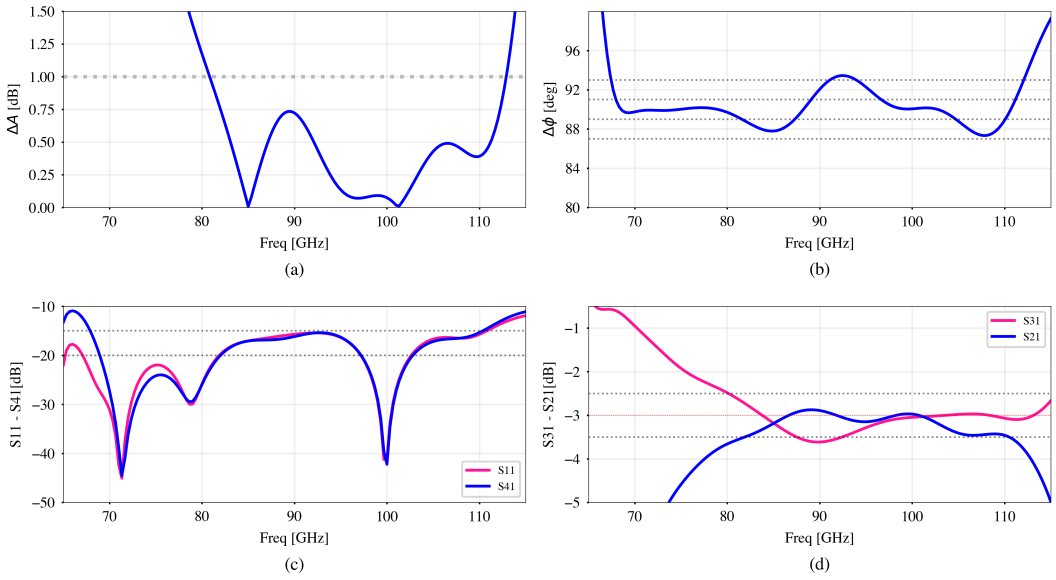


Figure 8.7: 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

The first thing to notice in this set of results (see Fig. 8.7 and 8.8), using the most symmetric version of the extruded quadrature hybrid, is that the band is not effectively extended thought there is significant improvement in the ΔA, even reaching values under 0.2 dB. As the $S_{11} - S_{41}$ figure reveals, there is a sacrifice in isolation and return losses thus the curve barely stay under -15 dB. This can be seen in the simulations B.4 and B.5 in which a residual power is traveling through Port 4. This result coincide with those of Rashid et al. (2016) for

ALMA Band 5.

Table 8.5: 12-branch hybrid extruded main guides.

Setting	Value
Id	a0ed
Band	ALMA 2+3 (67-116 GHz)
Operational Band	77.33 - 108 GHz
Optimized vector	$[0, 0.175, 0.313, 0.828, 0.819, 2.530, 1.265, 0.143, 0.1683]$mm
Fabrication Ratio	$a = 2.530, b_{inner} = 7.23, b_{outer} = 4.04$
Directivity	> 20 dB

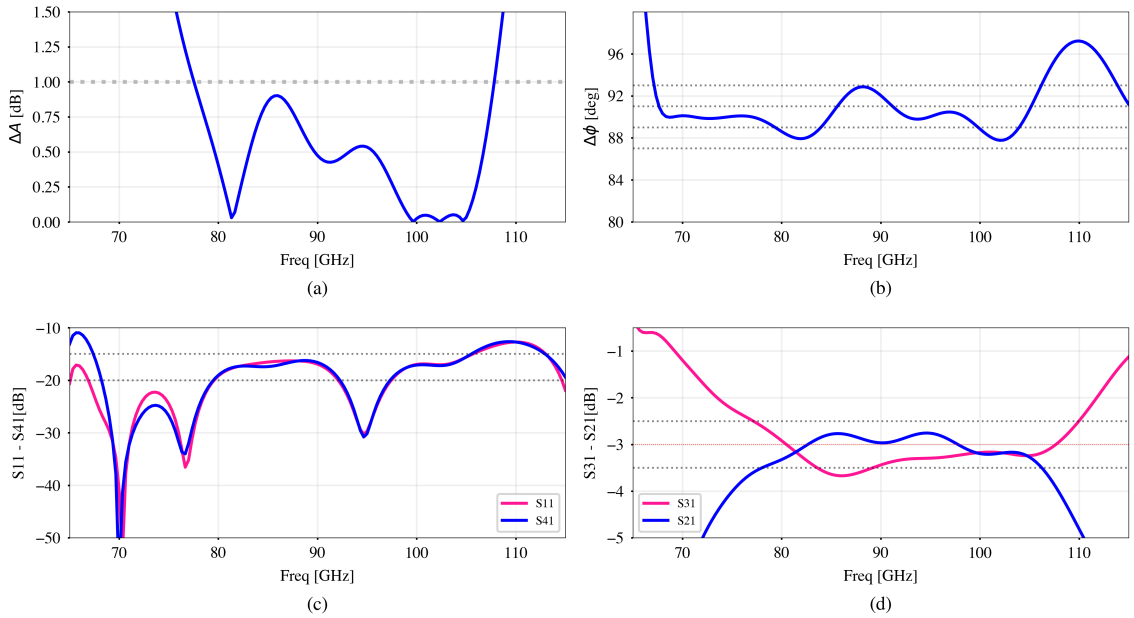


Figure 8.8: 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

Whereas the ΔA is improved, the $\Delta\phi$ is degraded presenting phase changes greater than $\pm 2°$, which coincides with the findings of Rashid et al. (2016) confirming that the introduction of the extra cavities create this sort of response.

The second approach, in which the parameters describing the inter-branch distance are loosen, is used. In this case, in the table 8.6 can be seen the optimized vector which allows for distinct branch sizes and separations as well.

The Fig. 8.9 shows a behavior which is better than the preceding designs in the regard that the ΔA stays under 1 dB mark while the $S_{11} - S_{41}$ stay well below -15 dB, which manifests itself in a smaller flow of power in Port 4 (see Fig. B.6). As for the $\Delta\phi$, its values are more controlled staying within the tolerance gap.

Table 8.6: 12-branch hybrid extruded main guides.

Setting	Value
Id	8ba81ea3
Band	ALMA 2+3 (67-116 GHz)
Operational Band	76 - 108 GHz
Optimized vector	$[0, 0.178, 0.17, 0.17, 0.179, 0.17, 0.216, 0.435, 0.55, 0.477, 0.604,$ $0.75, 0.42, 0.79, 2.62, 1.31, 0.18, 0.182]$mm
Fabrication Ratio	$b_1 = 7.36, b_2 = 7.71, b_3 = 7.71, b_4 = 7.32, b_5 = 7.71, b_6 = 6.06$
Directivity	> 22.3 dB

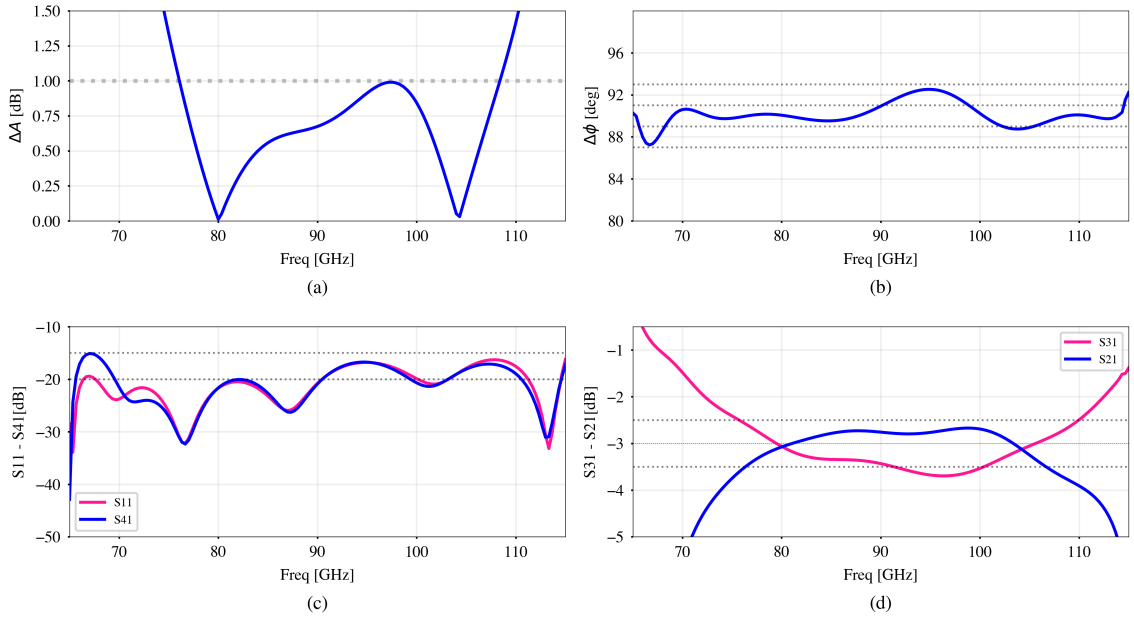


Figure 8.9: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

8.2 ALMA Band 3

ALMA Band 3 spans the range between 84 and 116 GHz, which is of particular scientific interest as it can be used to observe small-scale structure in cold gas clouds and to explore into galaxies to study how stars form. For this reason, there is a need to continuously improve the instruments to reduce atmospheric noise for the sake of higher quality observations.

With that purpose in mind, the use of our method can support a sustained improvement in the microwave devices used in this particular band. Based on the previous optimization results, it is clear that our optimized geometries could potentially improve those of other researches for quadrature hybrids set for specific bands as ALMA Band 3, as similarly proven for ALMA Band 5 during the validation stage (see Sec. 7.7).

The optimization focuses on 12-branch hybrids due to our prior knowledge of good performance while widening the operational band of the device. An 8-branch extruded design optimized to evaluate how such a hybrid behaves for the ALMA Band 3 regime, while a 14-branch optimization is performed to establish the trade-off of adding extra branches beyond 12.

8.2.1 8-Branch hybrid

As our first optimization goal is the 8-branch design, it makes sense to include the feature that could improve the amplitude imbalance namely adding the extruded discontinuities to the main guides.

The result is very compelling as it actually improves the ΔA performance while maintaining the $\Delta\phi$ within acceptable limits. This outcome is supported by the $S_{11} - S_{41}$ staying below the -20 dB line for most of the bandwidth and under -15 dB mark for the 100% of the band.

Table 8.7: ALMA Band 3, 8-branch hybrid with extruded main guides.

Setting	Value
Id	0ec1a
Band	ALMA 3 (84-116 GHz)
Operational Band	86.13 - 117.8 GHz
Optimized vector	$[0, 0.29, 0.28, 0.39, 0.29, 0.58, 0.55, 0.69, 0.5, 0.62, 2.54, 1.27, 0.18, 0.2]$mm
Fabrication Ratio	$a = 2.54, b_1 = 4.8, b_2 = 4.53, b_3 = 3.27, b_3 = 4.38$
Directivity	> 23.4 dB

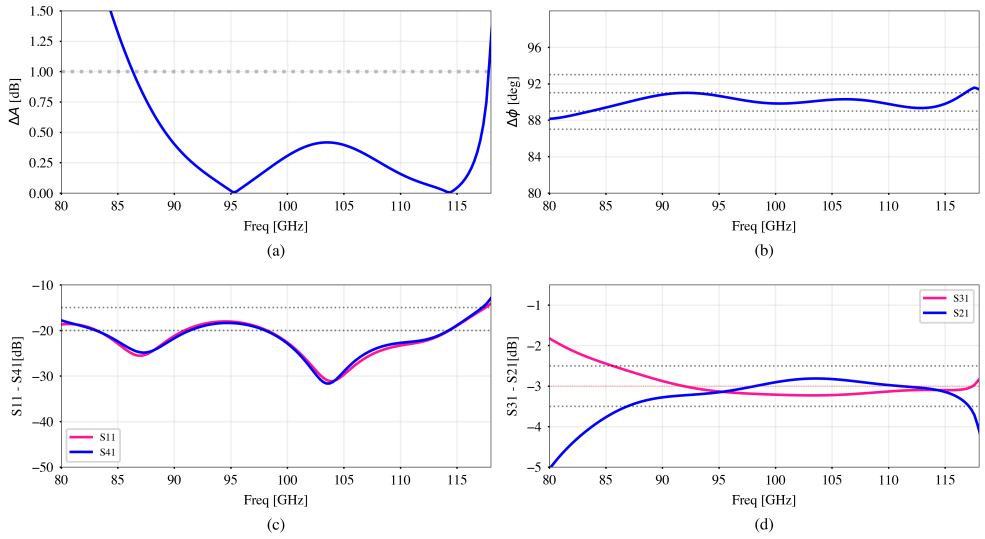


Figure 8.10: Simulation of an optimized 8-branch extruded quadrature hybrid set for ALMA Band 3 (86 – 116 GHz).

8.2.2 12-Branch hybrid

This optimization is performed using a full-length flexible vector to enable the use of different branch and separation sizes.

Table 8.8: ALMA Band 3, 12-branch hybrid with extruded main guides .

Setting	Value
Id	b11083
Band	ALMA 3 (84-116 GHz)
Operational Band	83.06 - 114 GHz
Optimized vector	$[0, 0.187, 0.188, 0.178, 0.173, 0.173, 0.232,$ $0.408, 0.35, 0.49, 0.478, 0.33, 0.247, 0.726, 2.506, 1.253, 0.208, 0.327]$mm
Fabrication Ratio	$a = 2.506$, $b_1 = 6.70$, $b_2 = 6.66$, $b_3 = 7.03$, $b_4 = 7.24$, $b_5 = 7.24$, $b_6 = 5.4$
Directivity	$>$ 24 dB in the 80-106 GHz range; >21 dB in the 106-116 GHz range

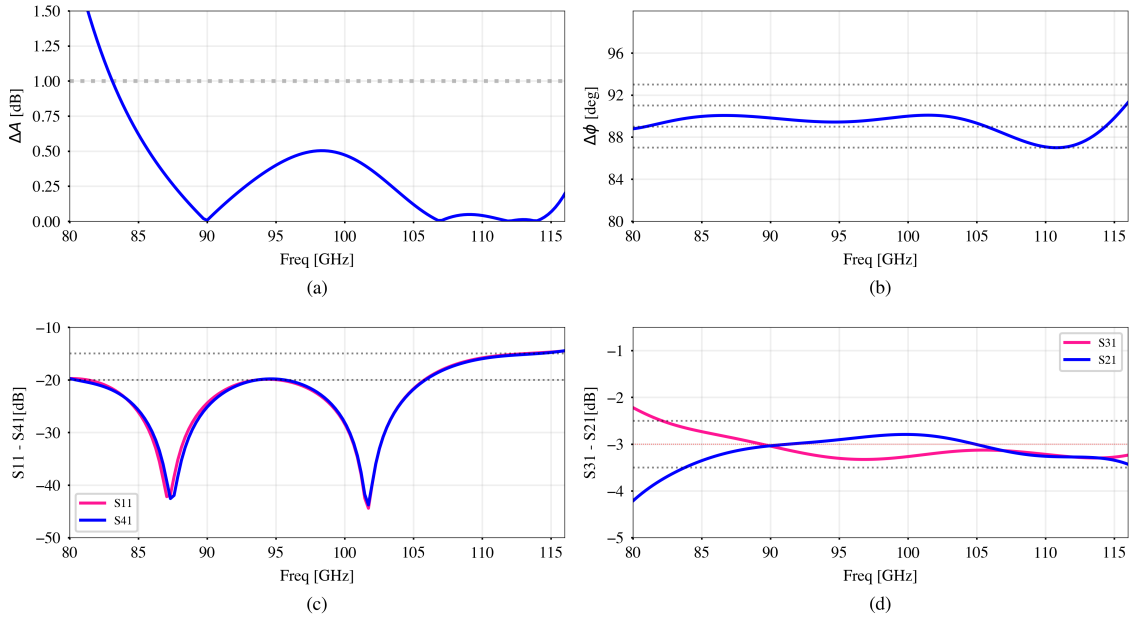


Figure 8.11: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 3 (86 – 116 GHz).

In the Fig. 8.11, the results are presented. There is a good performance with respect three markers. The first is the amplitude imbalance ΔA staying below the 0.5 dB mark, moreover, there is a range between 106 and 115 GHz in which a ultra-low imbalance takes place; secondly, the $S_{11} - S_{41}$ stays under the -15 dB line, even better, the figure presents stays below -20 dB level for a great part of the band. Finally, the phase imbalance $\Delta\phi$ behaves smoothly within the $\pm 1°$ range. Consequently, the isolation is as good as in the classical 8-branch devices (see Fig. B.9) while fully covering the 84-116 GHz band.

A remarkable result is the fabrication relation of the branches which value stays under 7. This ratio has a tolerance due to a $10\mu m$ fabrication margin of error hence the resulting values in the table 8.8 can be accepted or adjusted by tuning the branch sizes, moreover, the actual sizes adequate for fabrication purposes require not more than two decimal places due to the tooling error margin.

Two additional results of our interest are presented in table 8.9 and 8.10. In the Fig. 8.12 a better amplitude imbalance is presented in the operational band, lower than 0.3 dB; on the other hand, in the Fig. 8.13, we find a ΔA under 0.5 dB with a good performance for the return loss, isolation, namely $S_{11} - S_{41}$, and the $\Delta\phi$.

Table 8.9: ALMA Band 3, 12-branch hybrid with extruded main guides .

Setting	Value
Id	0bd63
Band	ALMA 3 (84-116 GHz)
Operational Band	83.6 - 115.86 GHz
Optimized vector	$[0, 0.17, 0.2, 0.18, 0.19, 0.2, 0.2, 0.32, 0.4,$ $0.46, 0.38, 0.67, 0.37, 0.68, 2.57, 1.285, 0.17, 0.18]$mm
Fabrication Ratio	$a = 2.57$, $b_1 = 7.55$, $b_2 = 6.425$, $b_3 = 7.14$, $b_4 = 6.76$, $b_5 = 6.426$, $b_6 = 6.425$
Directivity	> 22.5

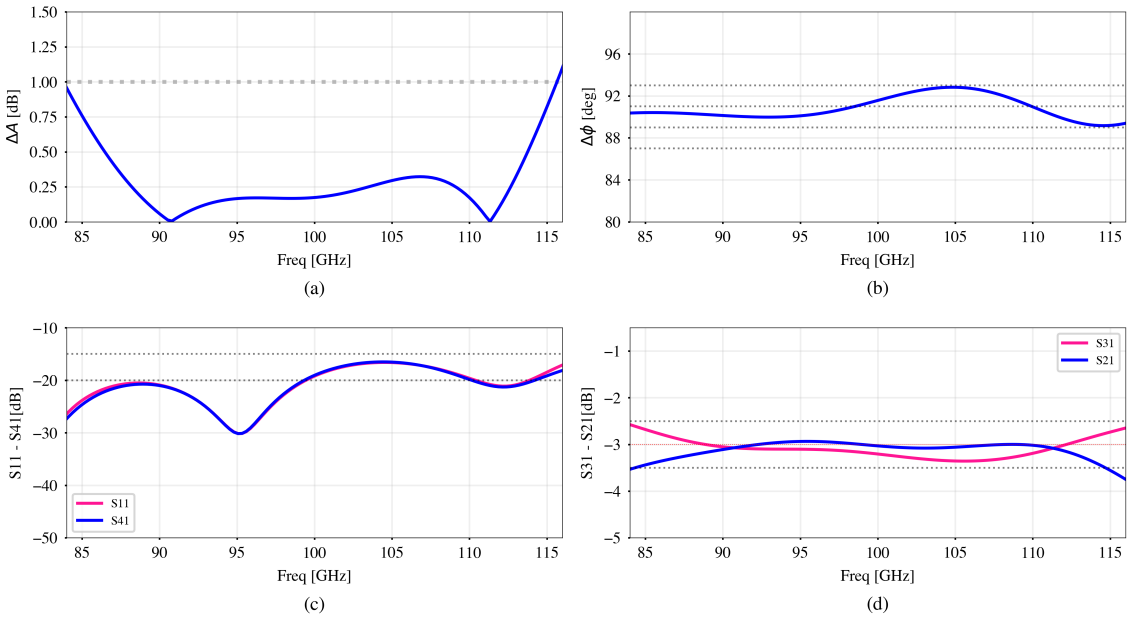


Figure 8.12: S-parameters and imbalances of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

Table 8.10: ALMA Band 3, 12-branch hybrid with extruded main guides .

Setting	Value
Id	70fc2
Band	ALMA 3 (84-116 GHz)
Operational Band	83.86 - 116.4 GHz
Optimized vector	$[0, 0.17, 0.2, 0.18, 0.19, 0.2, 0.2, 0.32, 0.4,$ $0.46, 0.38, 0.67, 0.37, 0.68, 2.57, 1.285, 0.17, 0.18]$mm
Fabrication Ratio	$a = 2.57$, $b_1 = 7.55$, $b_2 = 6.425$, $b_3 = 7.14$, $b_4 = 6.76$, $b_5 = 6.426$, $b_6 = 56.425$
Directivity	> 22.5

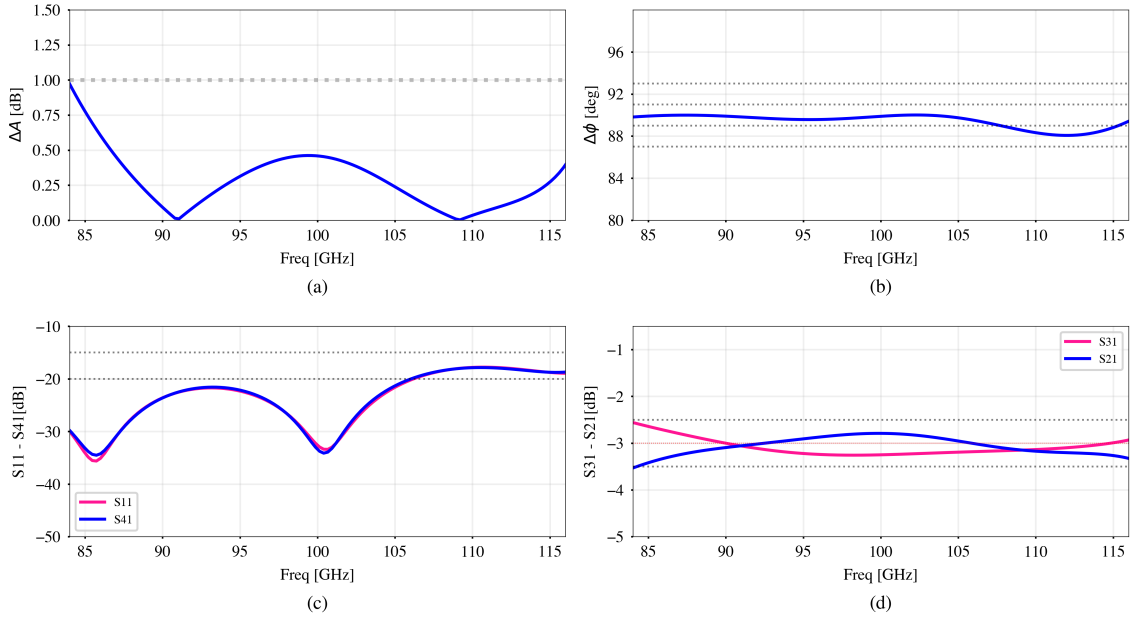


Figure 8.13: S-parameters and imbalances of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

8.2.3 14-branch hybrid

In order to test the limits of the classical branch geometry regarding the fabrication restriction imposed by the standard drilling relation, we add two extra branches to the design. In this case a Rashid-Reed vector is used to ease the optimization by reducing the number of free parameters.

Table 8.11: 14-branch hybrid extruded main guides.

Setting	Value
Id	3dfd6337
Band	ALMA 3 (84-116 GHz)
Operational Band	80.5 - 116.5 GHz
Optimized vector	$[0, 0.157, 0.154, 0.364, 0.705, 2.587, 1.2935, 0.186, 0.234]$mm
Fabrication Ratio	$a = 2.587$ $b_1 = 8.2388$ $b_2 = 8.399$
Directivity	> 22 dB

The results are valuable as they present a quadrature hybrids that complies with the established constraints. The ΔA stays under 1 dB for an 80-116 GHz range, hence spanning a 36 GHz bandwidth, which is wider than any other so far. The S-parameters fit the expected behavior while the $\Delta\phi$ stays within the tolerance ranges for most of the band.

Of greater importance is the fact that the inner branches (b_1) and the outer branches (b_2) as well, both have sizes that yield a fabrication relation higher than the factor of 7 which guarantees that standard drilling tools can be used (see Sec. 8). To achieve the fabrication of a geometry like this, a customized drilling tool is required thus elevating the fabrication cost.

As comparing this result with the ones with a smaller number of branches, we conclude that, while the number of branches grows the band coverage is widened, and therefore requiring narrower branches. An mentioned in previous sections, this is highly undesirable as the drilling of the hybrid gets more difficult hence more expensive.

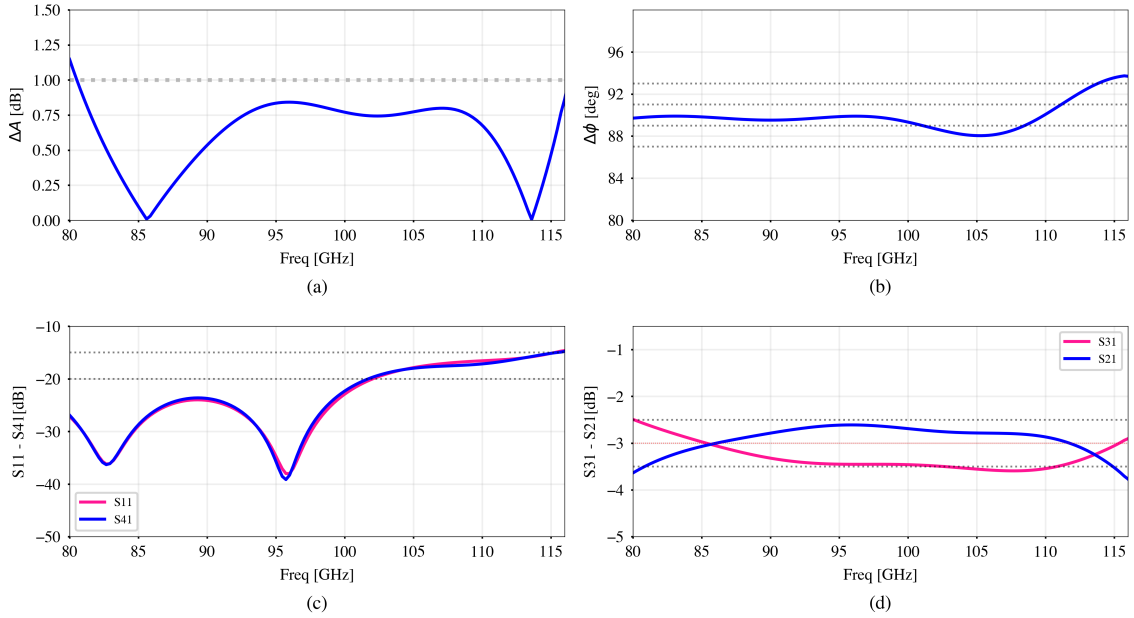


Figure 8.14: Simulation of an optimized 14-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

8.3 Quadrature Hybrid Performance Benchmarks

The quadrature hybrids and, in particular, directional couplers are extensively studied microwave devices hence there are plenty of bibliographical resource in which we could find simulated and experimental measurements to understand how those devices behave. In contrast, as the field of application is narrowed down, the number of sources and experimental results are dramatically reduced.

Moreover, as we focus on our research objective namely the ALMA 67-116GHz band, there is handful of results worth of mention in order to rule out the quality of the designs obtained through our method. Hereafter a concise comparison is made between some of those designs and our optimization results.

The reader must be aware of two things. In the first place, the designs are intended to operate in the 75-110GHz range which happens to be the intersection between ALMA Band 2 (67 - 90 GHz) and ALMA Band 3 (84 – 116 GHz). Second thing is that there is no reference in the literature that shows a quadrature hybrid design that fully covers the ALMA Band 2+3 (67 - 116 GHz), which for practical purposes classify as a ultra-wide band.

8.3.1 Quadrature hybrid for W-Band (75-110GHz) by Ding et al.

This paper reports an 8-branch hybrid in which the low amplitude imbalance is reached by using the main waveguides extrusions and discontinuities. The reported amplitude imbalance is under 0.8 dB while the phase imbalance $90°(+2.38°/-4.25°)$.

One of the most interesting features of this design is the changes in the main guides, which proved to be beneficial to improve the $\Delta\phi$. The hybrid was manufactured and measured proving a) $S_{11} - S_{41}$ under -16 dB, b) a $S_{21} - S_{31}$ figure not fully complying with the -3 dB mark which has the negative impact on the return loses and c) a ΔA fully covering the band under 0.8 dB.

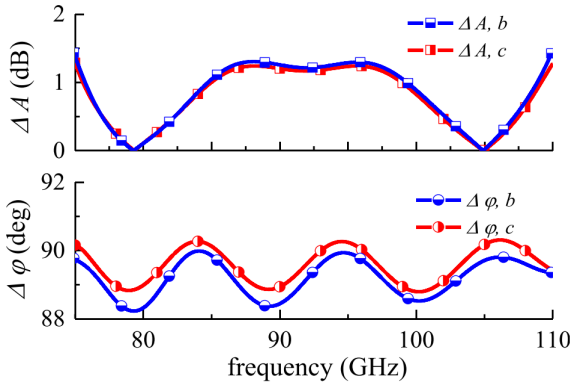


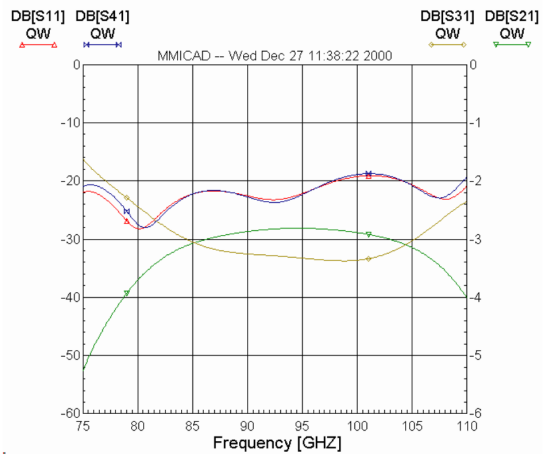
Figure 8.15: 3-dB 8-branch hybrid designed for W Band. **Source:** Ding et al., 2019

As we compare the initial results provided by Ding et al. to our optimization of a flat quadrature hybrid (see Fig. 8.4), we find that an optimized version can easily improve the performance in the band of interest. Now, as we introduce the inner extrusions, the optimization gives an improved amplitude and phase imbalance, as showed in the Fig. 8.5.

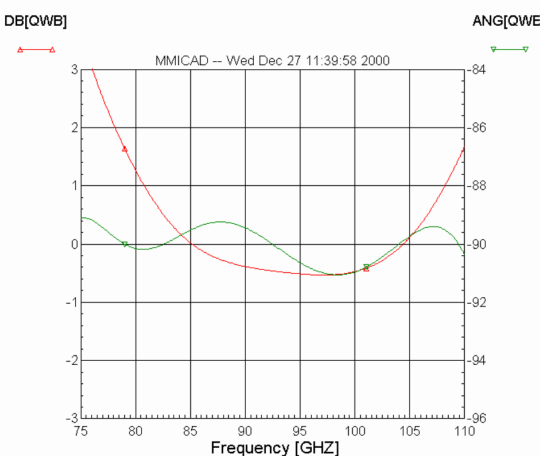
8.3.2 Quadrature hybrid (75-110GHz) by Srikanth and Kerr

This ALMA memo presents a number of results using the Quick Wave EM simulator. The authors had the same objective of finding those designs with $\Delta A < 1dB$.

In the design proposed by Srikanth and Kerr, they use different sizes for the outer most branches and different branch separations as well. Though other results were obtained, the Fig. 8.16 and Fig. 8.17 are the more interesting findings as they throw the lowest ΔA.



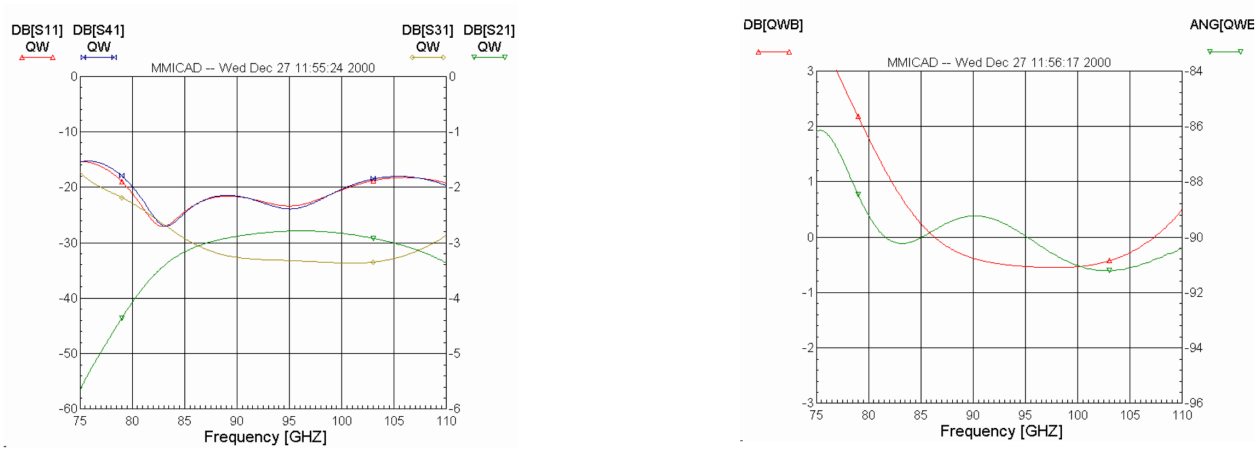
(a)



(b)

Figure 8.16: Simulated 6-branch hybrid with Quickwave (a) S-parameters (b) amplitude and phase imbalance.**Source:** Srikanth and Kerr, 2001

An interesting fact is that $S_{11} - S_{41}$ figure stays under -20 dB during a great part of the band. In contrast to the previous cases, the $S_{31} - S_{21}$ looks more centered at -3 dB which is important to have a better $\Delta\phi$. Though the amplitude imbalance is 0.5 dB at mid-band, it does not fully cover the band, furthermore, it effectively covers the region between 82.6 and 106.5 GHz.



(a) **(b)**

Figure 8.17: Simulated 6-branch hybrid with Quickwave (a) S-parameters (b) amplitude and phase imbalance.**Source:** Srikanth and Kerr, 2001

These results can be easily outperformed by using our 8-branch optimized design in Fig. 8.4, without compromising the simplicity of the flat hybrid and ensuring the use of standard drilling tools.

8.3.3 Quadrature hybrid W-Band (75-110GHz) by Monasterio et al.

This report do not strictly focuses on the quadrature hybrid itself, but on a complete design of a side-band separating down converter. The central device in a heterodyne receiver is the *mixer* which is based on the 90° hybrid, hence the same constraints must be fulfilled.

The multi-branch hybrid consists of WR-10 main guides based on the reed design to cover the full ALMA 2+3 Band. As showed in the report, the simulation (see Fig. 8.18) do not present the expected behavior due to the fabrication restrictions. However, the design was set to cover the upper-band hence showing a good performance in range between 85 and 118 GHz approximately.

This quadrature hybrid is particularly well behaved regarding the S_{11} as the figure stays below the -20 dB line. Unfortunately they do not provide information for the ΔA thus limiting the objective evaluation against our optimized designs. Nevertheless, the design of 12 branches corresponding to the figures are showed in Fig. 8.11, can be considered to improve the performance, mainly due to the low ΔA and smooth $\Delta\phi$.

8.3.4 Quadrature hybrid ALMA Band 3 by Andoh and Minamidani

Andoh and Minamidani (2003), presented a set of quadrature hybrids designed for ALMA Band 3 to 10. As in our case, they used HFSS as simulation tool to optimize quadrature hybrids to comply with the ΔA and $\Delta\phi$ requirements.

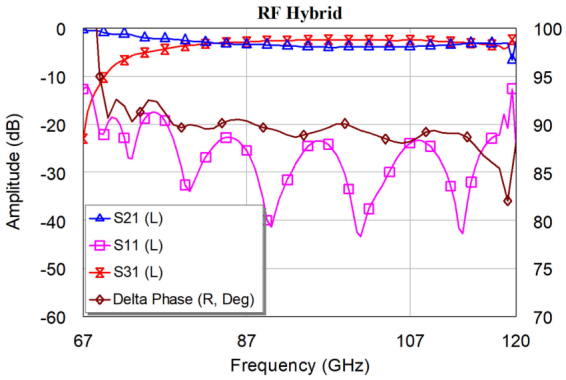


Figure 8.18: Scattering parameters of a 3-dB 90° hybrid designed for ALMA Band 3 **Source:** Monasterio et al., 2019

Of particular interest is the proposed design for ALMA Band 3. They used a flat 5-branch coupler, which behavior is presented in the Fig. 8.19. As in any flat design, the $S_{11} - S_{41}$ is very well behaved, staying under -20 dB for the whole band, even lower. So is the case of the phase imbalance which remains within the limits.

As we compare this design with our optimized hybrids on the basis of the fore-mentioned parameters, there is no immediate alternative to improve this performance. Now, as we take a look to the amplitude imbalance, Andoh and Minamidani's design is in the limit of $\Delta A = 1dB$ at mid-band.

If priority is given to the amplitude imbalance and the trade-off of degraded $S_{11} - S_{41}$ is accepted, then an extruded design can do the trick. In such a case, our Fig. 8.11 provides a feasible solution to improve the performance in ALMA Band 3.

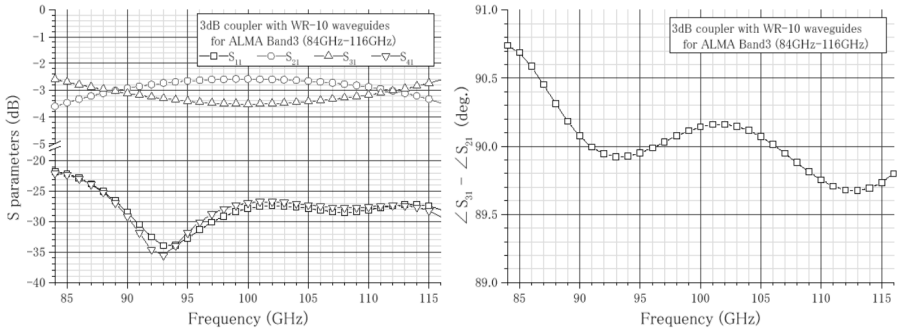


Figure 8.19: 3-dB 5-branch quadrature hybrid designed for ALMA Band 3 **Source:** Andoh and Minamidani, 2003

8.4 Assessment of Quadrature Hybrid Fabrication

The culmination of the development of an optimization process like ours, is to have a set of feasible geometries to be finally fabricated and deployed in an radio astronomy receiver. Though the scope of our work do not involve the fabrication and characterization of a quadrature hybrid, we want to set the ground for those subsequent stages of the project.

The key feature to consider is that a standard tolerance for a CNC machining service is typically $\pm 0.127mm$, the tightest machining tolerances possible are in the range of $\pm 0.0127mm$, though some milling machines like the KERN CNC could reach the $0.002mm$ accuracy (Rashid et al., 2016). Though higher precision can be obtained, custom tools should be made thus making the fabrication process very expensive. Our interest in having a fabrication ratio to constraint the branch sizes, relies in avoiding the use of customized tools while keeping the best performance possible for the optimized geometry.

In despite the good results showed by other geometries, the optimized geometries with the best performance are the ones in the tables 8.9, 8.9 and 8.10 for ALMA Band 3, due to their band coverage, fabrication ratios and amplitude imbalance. Additionally, a remarkable result is the $\Delta\phi = \pm 3^\circ$ for very low amplitude imbalances, as the one obtained in the upper band of the Fig. 8.11; this result improves the $\Delta\phi = \pm 4^\circ$ for ΔA near to 0.15dB, as reported by Rashid et al. (2016).

Let us take the case of the geometry represented by the vector in table 8.8. In this particular case, the branches 3, 4 and 5 are the those with the highest fabrication ration namely with the tighter branch size. We want to assess the behavior of the hybrid as the branches are bigger hence yielding a lower fabrication ratio. In the Fig. 8.20 we present a sweep analysis of the hybrid's behavior under changes in the branch sizes constrained to a ± 0.01 milling accuracy. As for the $S_{31} - S_{21}$ we find a good behavior in the regard of both figures keeping its shape and staying within the desired range. On the other hand, the ΔA slightly changes in the lower part of the band, while in the upper-band there is a more significant change reaching levels closer to 0.25dB. Consequently, the $\Delta\phi$ is well behaved during the most part of the band only reaching levels greater than 3° around the 110GHz, which still is viable if compared with the results reported by Rashid et al.

Another feasible design is the one in 8.8 which parametric analysis is presented in the Fig. 8.21. These results show a well behaved design even under the application of fabrication tolerance measure on the branch sizes. In particular, the $\Delta\phi$ stays under 3° in every one of the tested settings.

In the table 8.12 the nominal fabrication vectors to model two quadrature hybrid is provided complying with the fabrication ratio that ensures that a standard tool can be used for the machining process.

Table 8.12: Nominal vectors recommended for fabrication. Dimensions are given in millimeters.

Id	Nominal Vector
b11803	$[0, 0.18, 0.18, 0.19, 0.18, 0.18, 0.23, 0.41, 0.35, 0.49,$ $0.48, 0.33, 0.25, 0.72, 2.5, 1.25, 0.21, 0.33]$
0bd63	$[0, 0.18, 0.19, 0.18, 0.18, 0.2, 0.2, 0.32, 0.4, 0.46,$ $0.38, 0.67, 0.37, 0.68, 2.57, 1.285, 0.17, 0.18]$

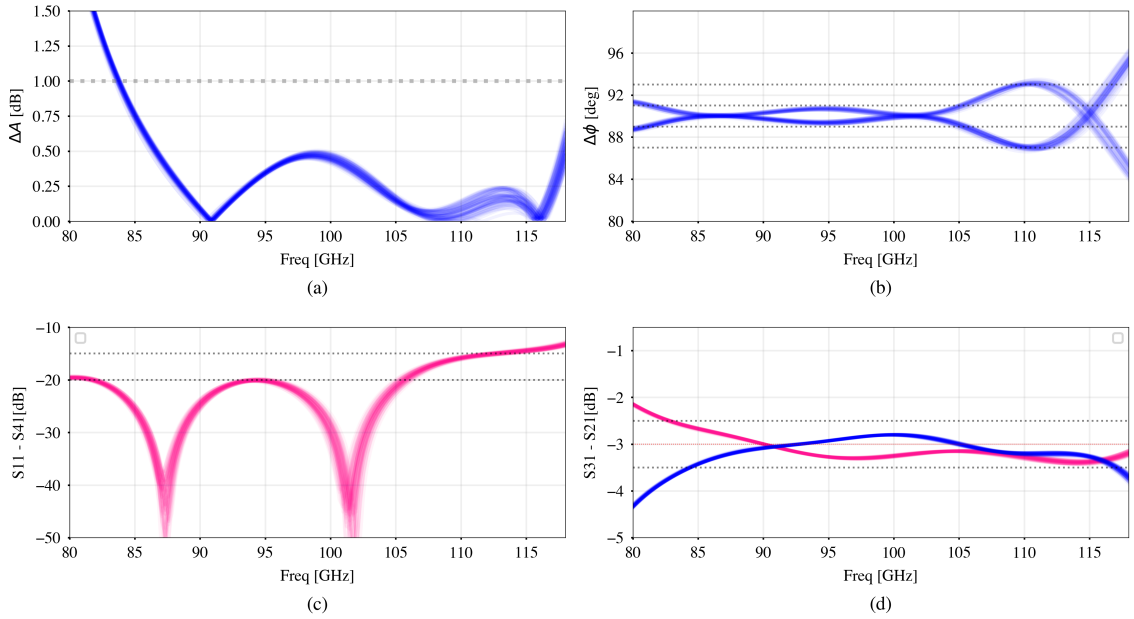


Figure 8.20: Parameter sweep to evaluate the performance under branch size change within the tolerance range.

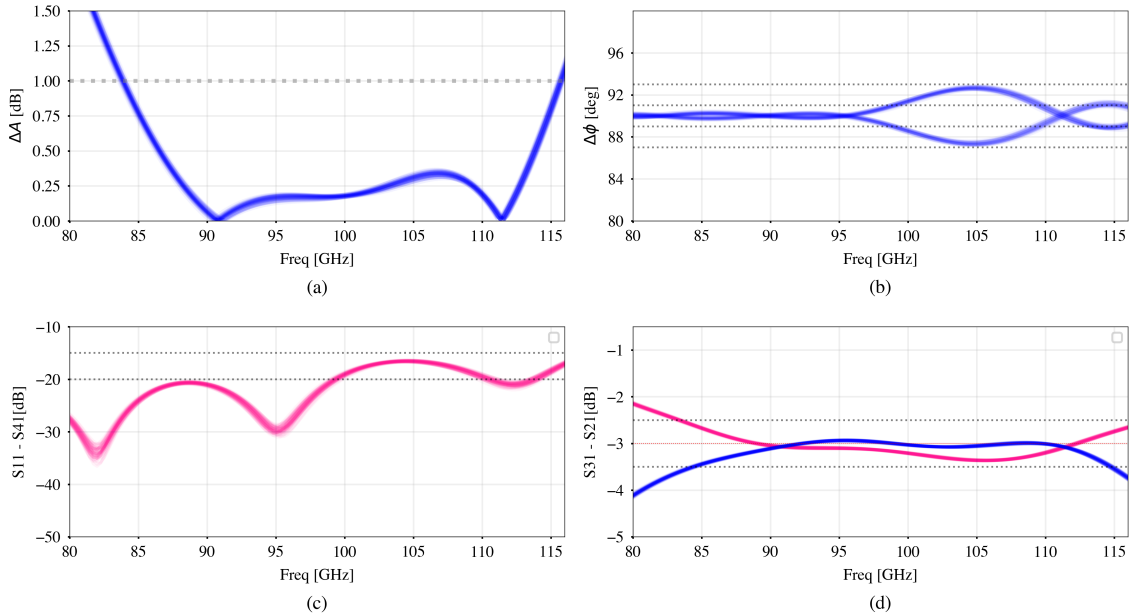


Figure 8.21: Parameter sweep to evaluate the performance under branch size change within the tolerance range.

9

Conclusion

After carefully studying the underlying concepts of microwave devices and guiding structures, an electromagnetic and transmission-line model was provided for the directional couplers and the multi-branch quadrature hybrid. This research approached the design challenges imposed by the highly demanding standards of radio astronomical observations. Moreover, measures as the amplitude (ΔA) and phase ($\Delta\phi$) imbalances, determine a set of constraints to satisfy in order for the hybrid to be eligible for fabrication and characterization stages in observatories as ALMA and others.

To facilitate the design process of microwave devices, in this thesis was introduced a novel optimization method based on the genetic algorithm, PSO. Dealing with an algorithm of stochastic nature gives advantages such as a random exploration of the search space, a property that is reinforced by the use of a population-based technique hence allowing for a robust exploration method of the solution space.

Our objective of getting the best ΔA possible while complying with the set of constraints imposed to the S-parameters was achieved after coding and fine tuning the algorithm. Likewise, a fabrication restriction is imposed to allow the use of standard drilling tools, which creates an additional side-issue of setting a limit to the size a branch can take. This unwanted effect was addressed in our method to reach our goal of producing a device that operates in wider-band regime than the classical quadrature hybrids while presenting a feasible design for fabrication purposes.

The use of the *Particle Swarm* algorithm together with the *Surrogate-Based* meta-heuristics, produced optimized multi-branch quadrature hybrids that comply with the set of constraints, improving other designs intended to operate in ALMA Band 3 and ALMA Band 5. The optimized quadrature hybrids can fully cover either the ALMA Band 2 or 3, spanning a 32 GHz bandwidth.

As the objective geometries correspond to the classical hybrids namely those with flat main guides and multi-branch designs, the improvement in ΔA is reached by using added cavities within the main guides while extension in the operational bandwidth is achieved by the addition of branches to our meta-models.

In despite of the successful deployment of the algorithm to provide optimized geometries, the results show that there exists physical restrictions manifested in constructive limits that impede the use of classical quadrature hybrids to achieve an extended operational bandwidth of 49 GHz corresponding to the ALMA Band 2+3 (67 - 116 GHz). Therefore, a subsequent step in this line of research is the pursue of novel designs expressed as new meta-models to extend the operational bandwidth in a more effective fashion.

Our resulting optimized designs provide feasible alternatives for ALMA W-Band and Band 3. Evidence of the performance improvement is provided as we use the benchmark established by different authors in the field.

In order to take further steps in the path that leads to a quadrature hybrid deployment in a production cartridge intended to operate in ALMA Band 3, this research provides two fully optimized and simulated designs presented in the table 8.12. Both geometries comply with the constraints set along this thesis as well as with a

suitable fabrication ratio to allow the use of standard milling tools. Furthermore, both have good ΔA (<0.5dB), and $\Delta\phi$ ($\pm 3.15°$). A parametric analysis was performed to study how these geometries behave under different branch sizes which could stem from the drilling accuracy. This analysis showed very reliable designs even under changes due to the fabrication errors.

Finally, this work can be of great help for other observatories, for instance, those looking for lower commissioning cost of their receivers without sacrificing the quality of their measurements. By allowing the use of standard drilling tools, we enable a better cost-benefit relation for radio astronomy receiver's design and fabrication.

Bibliography

Andoh, H., & Minamidani, T. (2003). ALMA MEMO 468 Designs of Wideband 3dB Branch-line Couplers for ALMA Bands 3 to 10, 1–10.

Baldick, R., & Uni-, T. (2006). *Applied optimization: formulation and algorithms for engineering systems* (Vol. 44). https://doi.org/10.5860/choice.44-2127

Bhagat, N. C. (1968). *Design Considerations Relating to Directional Couplers* (Doctoral dissertation). University of London.

Billade, B., Nystrom, O., Meledin, D., Sundin, E., Lapkin, I., Fredrixon, M., Desmaris, V., Rashid, H., Strandberg, M., Ferm, S. E., Wang, H., Xu, H., Obrocka, M., Ellison, B., Pavolotsky, A., & Belitsky, V. (2011). Performance of the first ALMA band 5 production cartridge. *2*(2), 56–56.

Bozorg-Haddad, O., Solgi, M., & Loáiciga, H. A. (2017). *Meta-Heuristic and Evolutionary Algorithms for Engineering Optimization*. https://doi.org/10.1002/9781119387053

Carlisle, A., & Dozier, G. (2001). An Off-The-Shelf PSO. *Proceedings of the workshop on particle swarm optimization (Indianapolis, IN)*.

Carpenter, J., Iono, D., Testi, L., Whyborn, N., Wootten, A., & Evans, N. (2019). The ALMA Development Roadmap. http://arxiv.org/abs/1902.02856

Claude, S., Jiang, F., Niranjanan, P., Dindo, P., Erickson, D., Yeung, K., Derdall, D., Duncan, D., Garcia, D., Leckie, B., Pfleger, M., Rodrigues, G., Szeto, K., Welle, P., Wood, I., Caputa, K., Lichtenberger, A., & Pan, S.-K. (2008). Performance of the pre-production band 3 (84-116 ghz) receivers for alma. *Millimeter and Submillimeter Detectors and Instrumentation for Astronomy IV*, *7020*, 70201B. https://doi.org/10.1117/12.788128

Clerc, M., & Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, *6*(1), 58–73. https://doi.org/10.1109/4235.985692

Coello, C. A. C., Lamont, G. B., & Veldhuizen, D. A. V. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. https://doi.org/10.1007/978-0-387-36797-2

Collin, R. E. (1991). *Field Theory of Guided Waves*. John Wiley & Sons.

Collin, R. E. (2001). *Foundations for Microwave Engineering* (second). Wiley Insterscience.

Corne, D. W., & Lones, M. A. (2018). Evolutionary algorithms. *arXiv*, 1–19. https://doi.org/10.1201/9781420034349.ch23

del Valle, Y., Venayagamoorthy, G. K., Mohagheghi, S., Hernandez, J. C., & Harley, R. G. (2008). Particle swarm optimization: Basic concepts, variants and applications in power systems. *IEEE Transactions on Evolutionary Computation*, *12*(2), 171–195. https://doi.org/10.1109/TEVC.2007.896686

Ding, J., Zhao, Y., Ge, J.-X., & Shi, S. (2019). A 90° Waveguide Hybrid with Low Amplitude Imbalance in Full W-Band. *Journal of Infrared, Millimeter, and Terahertz Waves*, *40*(4), 429–434. https://doi.org/10.1007/s10762-019-00577-1

Eberhart, R., & Kennedy, J. (1995). New optimizer using particle swarm theory. *Proceedings of the International Symposium on Micro Machine and Human Science*, 39–43. https://doi.org/10.1109/mhs.1995.494215

Ellingson, S. W. (2020). *Electromagnetics (volume 2)* (Vol. 2). Virginia Tech Publishing. https://doi.org/https://doi.org/10.21061/electromagnetics-vol-2

Finger, R. (2013). *Design and construction of a digital sideband separating spectrometer for the 1.2-meter southern radio telescope*. Universidad de Chile. https://repositorio.uchile.cl/handle/2250/113943

Firestone, W. L. (1952). *Transmission-line directional couplers* (Doctoral dissertation).

Franceschetti, G. (1997). *Electromagnetics: Theory, Techniques, and Engineering Paradigms*. Springer Science.

Fuller, G. A., Avison, A., Beltran, M., Casasola, V., Caselli, P., Cicone, C., Costagliola, F., Breuck, C. D., Hunt, L., Jimenez-Serra, I., Laing, R., Longmore, S., Massardi, M., Mroczkowski, T., Paladino, R., Ramstedt, S., Richards, A., Testi, L., Vergani, D., ... Wagg, J. (2018). The science case for alma band 2 and band 2+3. http://arxiv.org/abs/1602.02414

Fuller, G. A., Avison, A., Beltran, M., Casasola, V., Caselli, P., Cicone, C., Costagliola, F., Breuck, C. D., Hunt, L., Jimenez-Serra, I., Laing, R., Longmore, S., Massardi, M., Mroczkowski, T., Paladino, R., Ramstedt, S., Richards, A., Testi, L., Vergani, D., ... Wagg, J. (2020). The science case for alma band 2 and band 2+3. http://arxiv.org/abs/1602.02414

Gonzalez, A., Kojima, T., Kaneko, K., & Asayama, S. (2017). 275-500 GHz Waveguide Diplexer to Combine Local Oscillators for Different Frequency Bands. *IEEE Transactions on Terahertz Science and Technology*, *7*(6), 669–676. https://doi.org/10.1109/TTHZ.2017.2758789

Guilloteau, S. (2001). Spectral lines in high redshift galaxies with alma. *ASP conferences series*, *235*.

Hamid, M. A., & Yunik, M. M. (1967). On the Design of Stepped Transmission-Line Transformers. *IEEE Transactions on Microwave Theory and Techniques*, *MTT-15*(9), 528–529. https://doi.org/10.1109/TMTT.1967.1126525

Han, Z.-H., & Zhang, K.-S. (2012). Surrogate-Based Optimization. *Real-World Applications of Genetic Algorithms*, (May 2016). https://doi.org/10.5772/36125

Hartmann, A. H., & Rieger, H. (2001). *Optimization Algorithms in Physics*. Wiley. https://doi.org/10.1002/3527600876

Henderson, B. C., & Cook, J. A. (2001). The communications edge ™ tech-note image-reject and single-sideband mixers. *12*. www.wj.com

Hesper, R., Gerlofsma, G., Mena, P., Spaans, M., & Baryshev, A. (2009). A sideband-separating mixer upgrade for alma band 9. *20th International Symposium on Space Terahertz Technology*, 20–22.

Ishii, T. K. (1995). *Handbook of Microwave Technology*.

Jackson, J. D. (1975). *Classical Electrodynamics* (A. J. Wiley, Ed.; 2nd). https://doi.org/10.1007/978-1-4471-7284-0_43

Jackson, J. D. (1999). *Classical Electrodynamics* (J. Wiley, Ed.; 3rd). https://doi.org/10.1007/978-1-4471-7284-0_43

Jordan, E. C., & Balmain, K. (1968). *Electromagnetic Waves and Radiating Systems*. Prentice Hall.

Khudchenko, A., Hesper, R., Baryshev, A. M., Barkhof, J., & Mena, F. P. (2017). Modular 2sb sis receiver for 600-720 ghz: Performance and characterization methods. *IEEE Transactions on Terahertz Science and Technology*, *7*, 2–9. https://doi.org/10.1109/TTHZ.2016.2633528

Levy, R., & Lind, L. (1968). Synthesis of Symmetrical Branch-Guide Directional Couplers. *IEEE Transactions on Microwave Theory and Techniques*, *16*(2), 80–89.

Marcuvitz, N. (1986). *Waveguide Handbook* (second). Peter Peregrinus. https://doi.org/10.1016/s1063-5823(16)30025-4

Marr, J. M., Snell, R. L., & Kurtz, S. (2016). *Fundamentals of radio astronomy : Observational methods* (Vol. 1).

Mashayekhi, M., Harati, M., & Estekanchi, H. E. (2019). Development of an alternative PSO-based algorithm for simulation of endurance time excitation functions. *Engineering Reports*, *1*(3), 1–15. https://doi.org/10.1002/eng2.12048

Matthaei, G. L., Young, L., & Jones, E. (1980). *Microwave filters, impedance-matching networks, and coupling structures* (Vol. 1). https://doi.org/10.1109/proc.1965.4048

Mena, F. P., Kooi, J. W., Baryshev, A. M., Lodewijk, C. F., Zijlstra, T., Hesper, R., Gerlofsma, G., Klapwijk, T. M., & Wild, W. (2011). Design and performance of a 600-720-ghz sideband-separating receiver using alox and aln sis junctions. *IEEE Transactions on Microwave Theory and Techniques*, *59*, 166–177. https://doi.org/10.1109/TMTT.2010.2090417

Monasterio, D., Jarufe, C., Gallardo, D., Reyes, N., Mena, F. P., & Bronfman, L. (2019). A Compact Sideband Separating Downconverter with Excellent Return Loss and Good Conversion Gain for the W Band. *IEEE Transactions on Terahertz Science and Technology*, *9*(6), 572–580. https://doi.org/10.1109/TTHZ.2019.2937955

Montgomery, C. G., Dicke, R. H., & Purcell, E. M. (1987). *Principles of Microwave Circuits*. https://doi.org/10.1049/pbew025e

Mroczkowski, T., de Breuck, C., Kemper, C., Phillips, N., Fuller, G., Beltrán, M., Laing, R., Marconi, G., Testi, L., Yagoubov, P., George, D., & McGenn, W. (2019). Wide Bandwidth Considerations for ALMA Band 2. *arXiv*, 1–22.

Orfanidis, S. J. (2016). *Electromagnetic Waves and Antennas* (Vol. 2). https://doi.org/10.1016/B978-075064947-6/50011-3

Owyang, G. H. (1989). *Foundations for Microwave Circuits*. https://doi.org/10.1007/978-1-4613-8893-7

Parsopoulos, K. E., & Vrahatis, M. N. (2010). *Particle Swarm Optimization and Intelligence* (Vol. 270). http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-61520-666-7

Pozar, D. (1998). *Microwave Engineering* (John Wiley & Sons, Ed.; Vol. 1).

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Kevin Tucker, P. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, *41*(1), 1–28. https://doi.org/10.1016/j.paerosci.2005.02.001

Rajesh Kumar Arora. (2015). *OPTIMIZATION Algorithms and Applications*. CRC Press.

Ramo, S., Whinnery, J. R., & Duzer, T. V. (1994). *Fields and Waves in Communication Electronics* (3rd). John Wiley & Sons.

Rao Kurada, R., Pavan, K. K., & Rao, A. D. (2013). A Preliminary Survey on Optimized Multiobjective Metaheuristic Methods for Data Clustering Using Evolutionary Approaches. *International Journal of Computer Science and Information Technology*, *5*(5), 57–77. https://doi.org/10.5121/ijcsit.2013.5504

Rashid, H., Desmaris, V., Belitsky, V., Ruf, M., Bednorz, T., & Henkel, A. (2016). Design of Wideband Waveguide Hybrid with Ultra-Low Amplitude Imbalance. *IEEE Transactions on Terahertz Science and Technology*, *6*(1), 83–90. https://doi.org/10.1109/TTHZ.2015.2502070

Reed, J., & Wheeler, J. (1956). A Method of Analysis of Symmetrical Four-Port Networks. *IRE Transactions on Microwave Theory and Techniques*, *4*(4), 246–252. https://doi.org/10.1109/TMTT.1956.1125071

Reed, J. (1958). The Multiple Branch Waveguide Coupler. *IEEE Transactions on Microwave Theory and Techniques*, *6*(4), 398–403. https://doi.org/10.1109/TMTT.1958.1125213

Restrepo, O., & Lucero, F. (2020). *MIST Memo* (tech. rep.). University of the Chile and University ECCI.

Robinson, J., & Rahmat-Samii, Y. (2004). Particle swarm optimization in electromagnetics. *IEEE Transactions on Antennas and Propagation*, *52*(2), 397–407. https://doi.org/10.1109/TAP.2004.823969

Schieven, G. (2022). *Observing with alma – a primer*. ESO. www.almascience.org.

Shi, Y., & Eberhart, R. (1998). A Modified Particle Swarm Optimizer. *IEEE International Conference on Evolutionary Computation Proceedings.*, *1998*(5), 1144–1153.

Srikanth, S., & Kerr, A. (2001). ALMA Memo 343: Waveguide quadrature hybrids for ALMA receivers. *ALMA Memo Series*, *9*(January). http://legacy.nrao.edu/alma/memos/html-memos/alma343/memo343.pdf

Tang, Y., Wang, Z., & Fang, J. A. (2011). Feedback learning particle swarm optimization. *Applied Soft Computing Journal*, *11*(8), 4713–4725. https://doi.org/10.1016/j.asoc.2011.07.012

Thompson, A. R. (R., Moran, J. M., & Swenson, G. W. (W. (2001). *Interferometry and synthesis in radio astronomy*. Wiley.

Törn, A., & Zilinskas, A. (1989). *Lecture Notes in Computer Science: Global Optimization*. Springer-Verlag.

Van Den Bergh, F., & Engelbrecht, A. P. (2006). A study of particle swarm optimization particle trajectories. *Information Sciences*, *176*(8), 937–971. https://doi.org/10.1016/j.ins.2005.02.003

Vassilev, V., Belitsky, V., Risacher, C., Lapkin, I., Pavolotsky, A., & Sundin, E. (2017). Design and characterization of a sideband separating sis mixer for 85-115 ghz. *Softwaretechnik-trends*, 173. https://ui.adsabs.harvard.edu/abs/2004stt..conf..173V/abstract%0Ahttp://www.nrao.edu/meetings/isstt/papers/2004/2004173180.pdf%0Ahttps://research.chalmers.se/en/publication/9670%0Ahttps://lens.org/152-614-049-768-722

Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: an overview. *Soft Computing*, *22*(2), 387–408. https://doi.org/10.1007/s00500-016-2474-6

Yagoubov, P., Mroczkowski, T., Belitsky, V., Cuadrado-Calle, D., Cuttaia, F., Fuller, G. A., Gallego, J. D., Gonzalez, A., Kaneko, K., Mena, P., Molina, R., Nesti, R., Tapia, V., Villa, F., Beltrán, M., Cavaliere, F., Ceru, J., Chesmore, G. E., Coughlin, K., ... Wadefalk, N. (2020). Wideband 67-116 GHz Receiver Development for ALMA Band 2. *Astronomy and Astrophysics*, *634*(36777), 1–23. https://doi.org/10.1051/0004-6361/201936777

A

Appendix 1

A.1 Solving Maxwell's equations

The idea of energy transport in the electromagnetic fields can be further modeled in terms of propagation of waves. This assumption means that there is a dependence on the factor $e^{j\omega t}$, allowing a solution in the Fourier domain.

We can express fields in harmonic form with complex amplitude coefficients,

$$\vec{E} = E_0 e^{j\omega t} \tag{A.1}$$
$$\vec{H} = H_0 e^{j\omega t} \tag{A.2}$$
$$\tag{A.3}$$

Therefore, Maxwell's equations can be rewritten,

$$\boldsymbol{\nabla} \times \vec{E} = -j\omega\vec{B} \tag{A.4}$$
$$\boldsymbol{\nabla} \times \vec{H} = j\omega\vec{D} + \mathcal{J} \tag{A.5}$$
$$\boldsymbol{\nabla} \cdot \vec{D} = \rho \tag{A.6}$$
$$\boldsymbol{\nabla} \cdot \vec{B} = 0 \tag{A.7}$$

Equations A.4 and A.5 can be rearranged by replacing $\vec{D}$ and $\vec{B}$ by their fields counterparts; $\mathcal{J}$ is substituted using 2.8, to get the *phasor form* of Maxwell's equations. The phasor form is commonly used to work with oscillatory waves. Jordan and Balmain (1968) provide a clear and comprehensive description of the phasor representations,

$$\boldsymbol{\nabla} \times \vec{E} = -j\omega\mu\vec{H} \tag{A.8}$$
$$\boldsymbol{\nabla} \times \vec{H} = (\sigma + j\omega\epsilon)\vec{E} \tag{A.9}$$

From A.8 and A.9 we can derive the *Helmholtz wave equation* by solving for one of the two variables. Hence we obtain a general wave equation for the particular field $\vec{E}$ in a lossy medium,

$$\nabla^2\vec{E} + \omega^2\mu\epsilon\left(1 - j\frac{\sigma}{\omega\epsilon}\right)\vec{E} = 0 \tag{A.10}$$

Moreover, if we consider a lossless medium the *Helmholtz equation* becomes,

$$\nabla^2\vec{E} + \omega^2\mu\epsilon\vec{E} = 0 \tag{A.11}$$

or in terms of $\vec{H}$,

$$\nabla^2\vec{H} + \omega^2\mu\epsilon\vec{H} = 0 \tag{A.12}$$

We can then define the wave number k as,

$$k = \sqrt{\mu\epsilon}\omega \tag{A.13}$$

In order to simplify our analysis, we can think about a plane wave with the propagation direction in the z axis and a electric field in the x direction. The solution to the Helmholtz equation for $\vec{E}$ in the frequency domain is,

$$E_x = E^+ e^{-jkz} + E^- e^{jkz} \tag{A.14}$$

which is evidently composed by two traveling waves in opposite directions where E^+ and E^- are complex functions. The calculation of the $\vec{H}$ can be made by applying A.14 and A.8,

$$H_y = \frac{1}{\eta}[E^+ e^{-jkz} + E^- e^{jkz}] \tag{A.15}$$

where we define the wave impedance $\eta = \omega\mu/k = \sqrt{\mu/\epsilon}$, which can be seen as the *intrinsic impedance* of the medium. More important it the fact that $\vec{E}$ and $\vec{H}$ are orthogonal to the direction of propagation and to each other.

A.2 Rectangular cross section - TE mode

Solving the equation 2.44 can be done by using separation of variables, as presented by Jordan and Balmain (1968) and Montgomery et al. (1987), where $H_z = X(x)Y(y)$. If we substitute this expression for H_z in the wave equation, we get

$$Y\frac{\partial^2 X}{\partial x^2} + X\frac{\partial^2 Y}{\partial y} + (\gamma^2 + \omega^2\epsilon\mu)XY = 0 \tag{A.16}$$

Next step is to divide by XY what leaves the third term as a constant that must be equal to the first and second terms allowing the separation of A.16 into,

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} = -k_x^2 \tag{A.17}$$

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} = -k_y^2 \tag{A.18}$$

where $k_x^2 + k_y^2 = (\gamma^2 + \omega^2\epsilon\mu)$. By multipying each equation by X and Y respectively results in typical one-variable differential equations,

$$\frac{\partial^2 X}{\partial x^2} + k_x^2 X = 0 \tag{A.19}$$

$$\frac{\partial^2 Y}{\partial y^2} + k_y^2 Y = 0 \tag{A.20}$$

The solutions are also familiar,

$$X = A\cos(k_x x) + B\sin(k_x x) \tag{A.21}$$

$$Y = C\cos(k_y y) + D\sin(k_y y) \tag{A.22}$$

The coefficients must be determined from boundary conditions of the problem; in this case we look for solutions

that make zero every tangential component of the field $\vec{E}$,

$$E_y(x=0)=0 \tag{A.23}$$
$$E_y(x=a)=0 \tag{A.24}$$
$$E_x(y=0)=0 \tag{A.25}$$
$$E_x(y=b)=0 \tag{A.26}$$

If we use these boundary conditions together with A.21 to calculate required fields from equations 2.39, the results are,

$$-Asin(k_x\cdot 0)+Bcos(k_x\cdot 0)=0 \tag{A.27}$$
$$-Asin(k_x\cdot a)+Bcos(k_x\cdot a)=0 \tag{A.28}$$
$$-Csin(k_y\cdot 0)+Dcos(k_y\cdot 0)=0 \tag{A.29}$$
$$-Csin(k_y\cdot b)+Dcos(k_y\cdot b)=0 \tag{A.30}$$

hence B and D must be zero, finally reducing the solution to,

$$sin(k_x a)=0 \tag{A.31}$$
$$sin(k_y b)=0 \tag{A.32}$$

The parameters k_x and k_y are the wave numbers in $\hat{x}$ and $\hat{y}$. From A.31 and A.32 we can deduce that,

$$k_x=\frac{m\pi}{a} \quad , \quad m=0,1,2,3... \tag{A.33}$$
$$k_y=\frac{n\pi}{b} \quad , \quad n=0,1,2,3... \tag{A.34}$$

In consequence, in every direction we get discrete harmonics that stem from the parameters m and n. Remembering that γ is the phase propagation constant of the waves and that $k_x^2+k_y^2=(\gamma^2+\omega^2\epsilon\mu)$, ultimately we find,

$$\gamma=\sqrt{\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2-\omega^2\epsilon\mu} \tag{A.35}$$

which is only valid for some specific values of ω allowing for the propagation to occur in the z direction, found if γ is an imaginary value. In the situation of real-valued γ, the propagation is not allowed since attenuation is occurring. The critical value, usually known as **cut-off frequency**, is the one that satisfies,

$$\omega_c=\sqrt{\frac{1}{\epsilon\mu}\left[\left(\frac{m\pi}{a}\right)^2+\left(\frac{n\pi}{b}\right)^2\right]} \tag{A.36}$$

Here we can stop to make some considerations about the cut-off frequency. This critical value explicitly depends on the geometric size of the cross-section of the waveguide and, at the same time, depends on the values **m** and **n**, thereby different cut-off frequencies can be found for "higher" modes, namely higher values of m and n. This principle plays a role in novel designs which consider the use of *overmoded* waveguides to modify the behavior of the electromagnetic wave, for example, to expand the operational band.
Let us remember that γ is the propagation constant, which can be divided into real and imaginary parts. The latter is called *phase constant*, and is usually established as equal to the *wave number* $\vec{k}$, what is perfectly valid for TEM devices such as coaxial cables but not so for TE or TM waves. We can still derive information from A.35 by recalling that γ is related to the wavelength by,

$$\gamma=j\frac{2\pi}{\lambda_g} \tag{A.37}$$

where λ_g is the wavelength of the wave in the waveguide. Another important relation resulting from the free-space wave velocity (Pozar, 1998) is

$$\omega = \frac{\gamma}{\sqrt{\mu\epsilon}} \tag{A.38}$$

what also applies for the cut-off frequency, thus

$$\omega_c^2 = \frac{\frac{2\pi}{\lambda_c}}{\mu\epsilon}. \tag{A.39}$$

We can use A.37 and A.39 to replace into A.35, thus,

$$\left(\frac{2\pi}{\lambda_g}\right)^2 = \left(\frac{2\pi}{\lambda}\right)^2 - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2 \tag{A.40}$$

and finally we get the *wavelength of the guide*,

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2}} \tag{A.41}$$

The variable λ is the wavelength corresponding to the material filling the waveguide namely dielectric, air, etc. If we define λ_0 as the wavelength of a TEM wave in the free-space,

$$\lambda_0 = \lambda\sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}} \tag{A.42}$$

we can then rewrite A.41,

$$\lambda_g = \frac{\lambda_0}{\sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0} - \left(\frac{\lambda_0}{\lambda_c}\right)^2}} \tag{A.43}$$

One final observation must be that when assigning values to m and n, specific modes arise and, simultaneously, cut-off frequency changes. As for the particular case of TE waves, modes are notated by using subscripts TE_{mn}; for instance, TE_{10} yields,

$$\omega_c = \frac{\pi}{\epsilon\mu a} \tag{A.44}$$

$$\lambda_c = 2a. \tag{A.45}$$

A.3 Impedance for TEM waves in parallel plates

Let us briefly explore the parallel plates case. As in every other case, Maxwell's equations can be solved for the boundaries defined by our objective geometry, parallel plates as depicted in Fig. A.1. Boundary conditions considering perfect conductors are $E_{tang.} = 0$ and $H_{normal} = 0$. Wave equations can be derived (Jordan and Balmain, 1968) using similar arguments as in the previous sections,

$$\frac{\partial^2\vec{E}}{\partial x^2} + (\gamma^2 + \omega^2\mu\epsilon)\vec{E} = 0 \tag{A.46}$$

$$\frac{\partial^2\vec{H}}{\partial x^2} + (\gamma^2 + \omega^2\mu\epsilon)\vec{H} = 0 \tag{A.47}$$

Solutions can be found for the TE, TM or TEM waves. Let us take the TE waves again for explanatory purposes. The wave equation can be rewritten as,

$$\frac{\partial^2 E_y}{\partial x^2} + (\gamma^2 + \omega^2 \mu\epsilon) E_y = 0 \tag{A.48}$$

By using the same methods from earlier sections, this simple-harmonic equation is solved yielding,

$$E_y = C_1 \sin\Big(\frac{m\pi}{a}x\Big) e^{-\gamma z} \tag{A.49}$$

$$H_z = -\frac{m\pi}{j\omega\mu a} C_1 \cos\Big(\frac{m\pi}{a}x\Big) e^{-\gamma z} \tag{A.50}$$

$$H_x = -\frac{\gamma}{j\omega\mu} C_1 \sin\Big(\frac{m\pi}{a}x\Big) e^{-\gamma z} \tag{A.51}$$

where a is the separation between the plates and C_1 is a constant to be determined. This looks different when compared to the rectangular waveguide, as there are no boudaries in every direction; the parallel plates case presents a boundary set by the two plates meanwhile running to the "infinite" in $\hat{y}$ and $\hat{z}$.

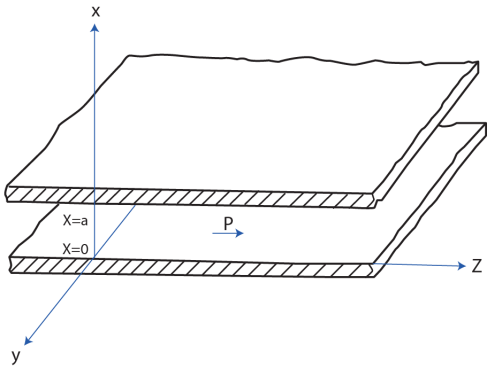


Figure A.1: Parallel plates guiding structure.

Once we have expressions for the fields, we can compute the corresponding impedance as the relation between E and H fields. Recall from equation 2.32, that the general γ is a complex number, but for perfectly conducting plates is an imaginary number β,

$$\left|\frac{E_y}{H_x}\right| = \frac{\omega\mu}{\beta} = \frac{\omega\mu}{\sqrt{\omega^2\mu\epsilon - \left(\frac{m\pi}{a}\right)^2}} \tag{A.52}$$

which happens to be the impedance in the z direction and, additionally, is constant over the cross section of the guide.

Let us consider an additional example. A guide made of two ideally conductive parallel plates in which we transport TEM waves and, if we set the propagation direction along z axis, then we have a magnetic field $\vec{H} = -H_x\hat{x}$ parallel to the surface. Thus, we have a current density in the surface flowing in the z direction $\vec{J}_s = \hat{n} \times \vec{H}$, which is simply $\vec{J}_{sz} = -\vec{H}_x$.

By applying the same procedure, we can compute fields for a TEM wave between parallel plates, which is of major importance, as it is the type of wave propagated along two-conductor transmission lines furthermore it is usually know as the *principal wave*. For the entirely transverse waves we can get the fields from the resulting equations for TM or TE waves,

$$H_y = C_1 e^{-j\gamma z} \tag{A.53}$$

$$E_x = \frac{\gamma}{\omega\epsilon} C_1 e^{-j\gamma z} \tag{A.54}$$

$$E_z = 0 \tag{A.55}$$

These results can be used to calculate the impedance for TEM waves in parallel plates,

$$\left|\frac{E_x}{H_y}\right| = \frac{\beta}{\omega\epsilon} = \sqrt{\frac{\mu}{\epsilon}} \quad . \tag{A.56}$$

In the more general case of not perfect conductivity, there is a surface impedance Z_s of the conductor that needs to be overcome by the surface current, creating the need for a tangential component of $\vec{E}$ (Jordan and Balmain, 1968). The importance relies on the fact that, as we include the surface impedance, we are working with a complex quantity which further involves a change in the *phase* of the electric wave. When an electromagnetic wave is guided by a structure, it is done along the surface of the conductor; currents are flowing in the conductor and charges manifest in its surface.

A.4 S-matrix for four-port devices

Condition 3.106 must be used to derive the entries of the matrix. Let us take $n = 1$, $m = 2$, while $k = 1...4$; recall that the sum must be equal to zero if $n \neq m$,

$$S_{11}^* S_{12} + S_{21}^* S_{22} + S_{31}^* S_{32} + S_{41}^* S_{42} = 0 \tag{A.57}$$

$$S_{31}^* S_{32} + S_{41}^* S_{42} = 0 \tag{A.58}$$

Similarly, by applying the same process to all cases in which $n \neq m$, a set of equations are obtained,

$$S_{31}^* S_{32} + S_{41}^* S_{42} = 0 \quad , \quad n = 1 \quad , \quad m = 2 \tag{A.59}$$

$$S_{21}^* S_{23} + S_{41}^* S_{43} = 0 \quad , \quad n = 1 \quad , \quad m = 3 \tag{A.60}$$

$$S_{21}^* S_{24} + S_{31}^* S_{34} = 0 \quad , \quad n = 1 \quad , \quad m = 4 \tag{A.61}$$

$$S_{12}^* S_{13} + S_{42}^* S_{43} = 0 \quad , \quad n = 2 \quad , \quad m = 3 \tag{A.62}$$

$$S_{12}^* S_{14} + S_{32}^* S_{34} = 0 \quad , \quad n = 2 \quad , \quad m = 4 \tag{A.63}$$

$$S_{13}^* S_{14} + S_{23}^* S_{24} = 0 \quad , \quad n = 3 \quad , \quad m = 4 \tag{A.64}$$

keep in mind that all other possible combinations are redundant because of the symmetry of the matrix. The next step is to iterate while taking $n = m$, for instance, if $n = m = 2$,

$$S_{12}^* S_{12} + S_{22}^* S_{22} + S_{32}^* S_{32} + S_{42}^* S_{42} = 1 \tag{A.65}$$

$$|S_{12}|^2 + |S_{32}|^2 + |S_{42}|^2 = 1 \tag{A.66}$$

As for the $n \neq m$ case, a set of equations may be obtained,

$$|S_{21}|^2 + |S_{31}|^2 + |S_{41}|^2 = 1 \quad , \quad n = m = 1 \tag{A.67}$$

$$|S_{12}|^2 + |S_{32}|^2 + |S_{42}|^2 = 1 \quad , \quad n = m = 2 \tag{A.68}$$

$$|S_{13}|^2 + |S_{23}|^2 + |S_{43}|^2 = 1 \quad , \quad n = m = 3 \tag{A.69}$$

$$|S_{14}|^2 + |S_{24}|^2 + |S_{34}|^2 = 1 \quad , \quad n = m = 4 \tag{A.70}$$

The behavior of an specific device can be modeled as specific entries of the matrix and, in consequence, determines the rest of the entries. For instance, if the expected feature is that of no coupling should occur between port 1 and 4, hence $S_{14} = S_{41} = 0$. Same happens if ports 2 and 3 are expected to be isolated, thus $S_{23} = S_{32} = 0$.

Take A.59 and A.63 while considering the property of the symmetry of the matrix. It allows to rewrite entries

in the form $S_{31} = S_{13}$ hence,

$$S_{13}^* S_{23} + S_{14}^* S_{24} = 0 \tag{A.71}$$
$$S_{13}^* S_{14} + S_{23}^* S_{24} = 0. \tag{A.72}$$

Now multiply the first by S_{23}^* and the second by S_{14}^*, and subtract to get,

$$S_{13}^* \left(|S_{23}|^2 - |S_{14}|^2 \right) = 0 \tag{A.73}$$

As our device is expected to have coupling between port 1 and 3 thus $S_{13} \neq 0$, and in consequence,

$$S_{23} = S_{14} = 0. \tag{A.74}$$

Further results can be obtained from (A.67 - A.70), for instance, by using the first two equations we can conclude that,

$$|S_{13}| = |S_{24}| \tag{A.75}$$

Operating the same way with other pair of equation, other relations are determined,

$$S_{21} = S_{12} = S_{34} = S_{43}. \tag{A.76}$$

Thus the S-matrix takes the form,

$$S = \begin{bmatrix} 0 & S_{12} & S_{13} & 0 \\ S_{12} & 0 & 0 & S_{24} \\ S_{13} & 0 & 0 & S_{34} \\ 0 & S_{24} & S_{34} & 0 \end{bmatrix} \tag{A.77}$$

There is something that still needs to be addressed. Though we already identified how entries distribute, the specific values have not been deduced, moreover, the phase at every port must be found with respect a reference phase. Let us consider that port 1 is the reference as entry point for an incident wave. Let us also recall equations 3.52, 3.67 and 3.71 and it is evident that if an incident wave from port 1 is fully transferred into port 2 namely there are no reflected waves, the wave at the end of the line must have a phase change due to the impedance along the waveguide. If we consider a lossless guide thus $\alpha = 0$ then,

$$S_{21} = \frac{v_2^-(\text{Z=l})}{v_1^+(\text{Z=0})} = \frac{V_0^+ e^{j\beta l}}{V_0^+} = e^{j\beta l} \tag{A.78}$$

The same happens if other coupling route is chosen, from port 1 to port 3, for instance. Let

$$S_{21} \equiv a e^{j\theta} \quad , \quad S_{31} \equiv b e^{j\phi} \tag{A.79}$$

With this two equivalences at hand, let us replace them into A.62 while recalling the symmetries found until this point,

$$a e^{-j\theta} b e^{j\phi} + b e^{-j\phi} a e^{j\theta} = 0 \tag{A.80}$$

which can be reduced to,

$$e^{-j(\theta-\phi)} + e^{+j(\theta-\phi)} = 0 \tag{A.81}$$

Since the last relation can be treated as a trigonometric identity, we obtain,

$$\theta - \phi = \pm \left(\frac{\pi}{2} + n\pi \right) \qquad n = 0, 1, 2... \tag{A.82}$$

Frequently, both θ and ϕ are chosen in a way that the device has a characteristic response, for example, by

choosing $\theta = 0$, the following is obtained,

$$S = \begin{bmatrix} 0 & a & be^{j\phi} & 0 \\ a & 0 & 0 & be^{j\phi} \\ be^{j\phi} & 0 & 0 & a \\ 0 & be^{j\phi} & a & 0 \end{bmatrix} \tag{A.83}$$

which is simplified by directly replacing a value for ϕ,

$$S = \begin{bmatrix} 0 & a & bj & 0 \\ a & 0 & 0 & bj \\ bj & 0 & 0 & a \\ 0 & bj & a & 0 \end{bmatrix} \tag{A.84}$$

Finally, we need to account for the values of a and b. These values can be obtained by complying to the relations (A.67-A.70) in the following form,

$$|S_{21}|^2 + |S_{31}|^2 + |S_{41}|^2 = |a|^2 + |bj|^2 = a^2 + b^2 = 1. \tag{A.85}$$

There are possible solutions to A.85, all of which can be assessed based on the expected behavior. One solution is $a = b = 0$, which is trivial as it implies $S_{12} = S_{13} = S_{24} = S_{34} = 0$; this case corresponds to a fully decoupled network. Another solution comes from the fact that the expectation is a coupling to occur between ports 1-2 and ports 1-3, hence a and b must also comply to a power relation. For instance, in directional couplers where the coupling factor is 3-dB, the coefficients are determined,

$$a = b = 1/\sqrt{2}. \tag{A.86}$$

The complete S matrix has the form,

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & j & 0 \\ 1 & 0 & 0 & j \\ j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \end{bmatrix} \tag{A.87}$$

A.5 Even- and Odd-Mode theory for four-port devices

This method is very useful as it enables a simpler analysis by dividing the four-port network into two-port lines, each one to be studied separately. This is specially useful for symmetrical directional couplers as they can be treated by means of the superposition principle in which the circuit is decomposed into even- and odd-mode excitation.

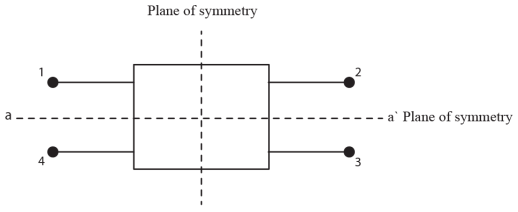


Figure A.2: Symmetry planes in a quadrature hybrid.

The even-mode corresponds to the use of two in-phase inputs (excitation) while the odd-mode corresponds to two out-of-phase inputs, with the conditions that as both solutions are superposed, yields the original expected

behavior. Let us think about the quadrature hybrid; the design demands that the input at port 1 is 1 while the resulting signal at port 4 must be zero, by the isolation principle. Thus, as soon as the circuit separation takes place in even and odd modes, the signal at port 1 is 1/2 while the signal at port 4 is 1/2 (even mode) and $-1/2$ (odd mode), where the in-phase/out-of-phase condition is accounted by the positive or negative value.

Consider a symmetrical four-port device as showed in Fig. 4.3, in which a plane of symmetry can be drawn horizontally. If a signal of equal magnitude and phase a_e is introduced into port 1 and 4, the plane of symmetry behaves as an open circuit or magnetic wall. Thus the scattering parameters become,

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix} \begin{bmatrix} a_e \\ 0 \\ 0 \\ a_e \end{bmatrix} \tag{A.88}$$

By solving the equations, it can be found that,

$$b_1 = a_e(S_{11} + S_{14}) \tag{A.89}$$
$$b_2 = a_e(S_{21} + S_{24}) \tag{A.90}$$
$$b_3 = a_e(S_{31} + S_{34}) \tag{A.91}$$
$$b_4 = a_e(S_{41} + S_{44}) \tag{A.92}$$

This can be further developed by recalling the symmetry relations A.77 and A.76, allowing to reshape the entries of the matrix, hence,

$$b_1 = b_4 = a_e(S_{11} + S_{14}) \tag{A.93}$$
$$b_2 = b3 = a_e(S_{12} + S_{13}) \tag{A.94}$$

After having found the incident and reflected quantities, the transmitted and reflected coefficients can be obtained,

$$\Gamma_1 = \frac{b_1}{a_e} = (S_{11} + S_{14}) \tag{A.95}$$
$$\Gamma_4 = \frac{b_4}{a_e} = (S_{11} + S_{14}) \tag{A.96}$$
$$T_1 = \frac{b_3}{a_e} = (S_{12} + S_{13}) \tag{A.97}$$
$$T_4 = \frac{b_2}{a_e} = (S_{12} + S_{13}) \tag{A.98}$$

The odd-mode relations are obtained by exciting the device with opposite-phase waves causing the plane of symmetry to work as an electric wall or short-circuit.

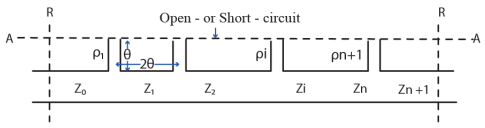


Figure A.3: Eve and odd sub-circuit.

As showed in Fig. A.3, two sub-circuits are then created thereby leaving a two-port transmission line with stubs along the guide. The stubs are open-ended, in the even-mode case, or short-circuit, in the other case.

B

Appendix 2

B.1 Optimized Quadrature Hybrids simulations

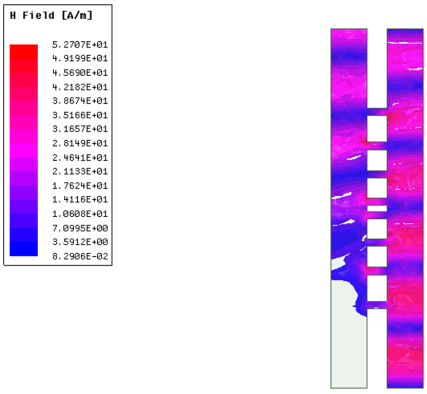


Figure B.1: Simulation of an optimized 8-branch quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

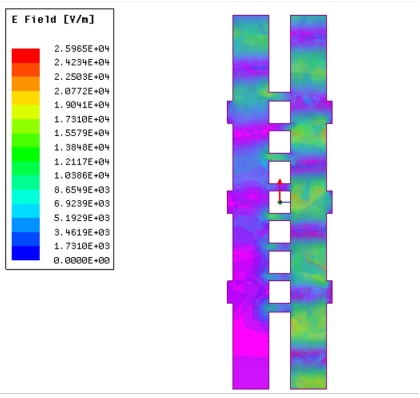


Figure B.2: Simulation of an optimized 8-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

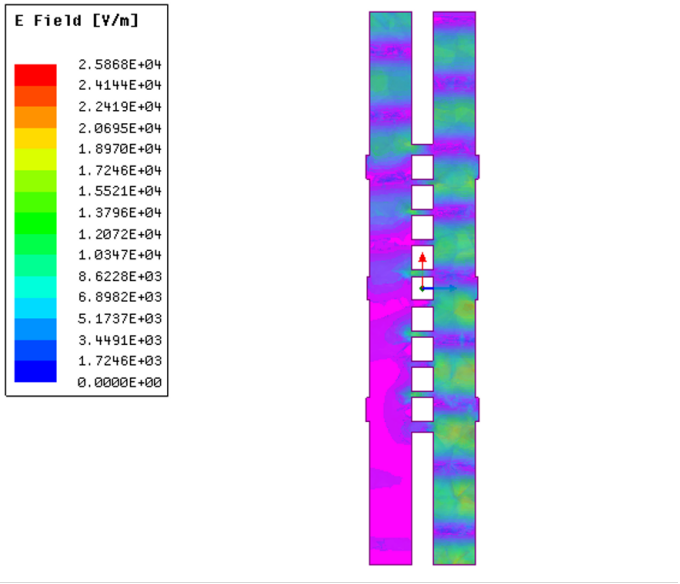


Figure B.3: Simulation of an optimized 10-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

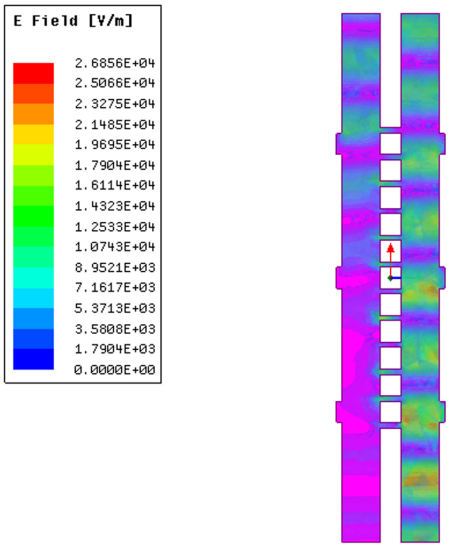


Figure B.4: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

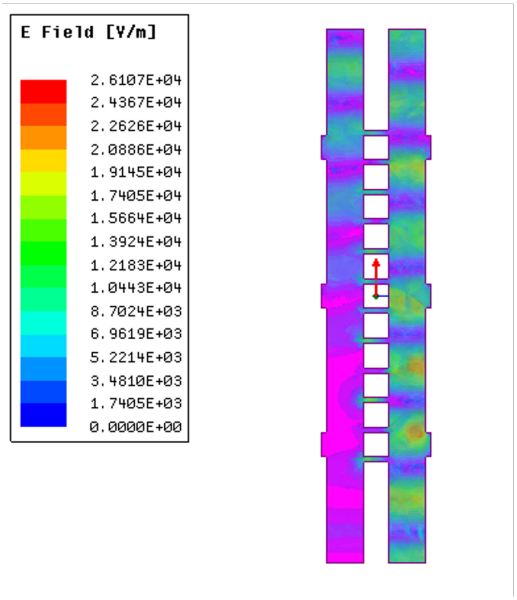


Figure B.5: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz).

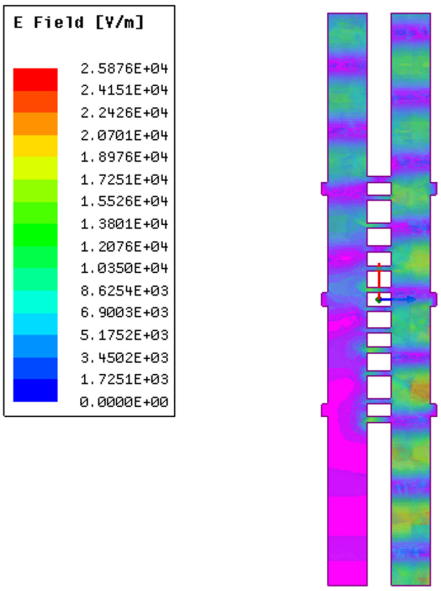


Figure B.6: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 2+3 (67 – 116 GHz) using a flexible branch distribution.

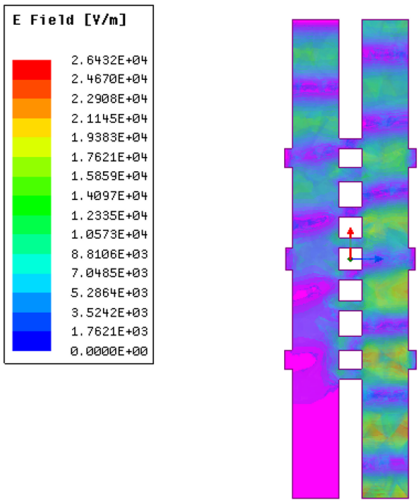


Figure B.7: Simulation of an optimized 8-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

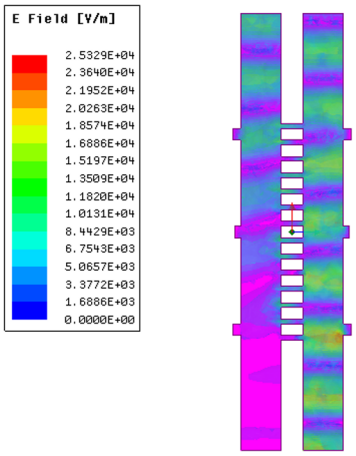


Figure B.8: Simulation of an optimized 14-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

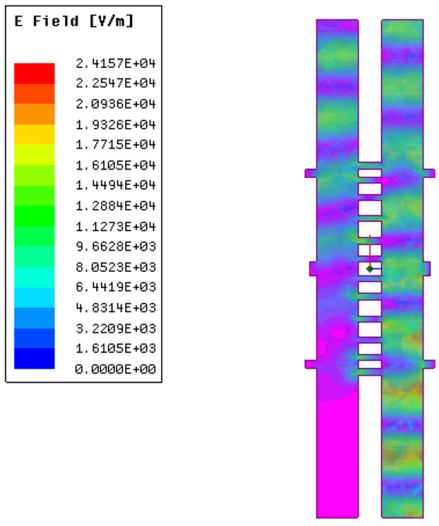


Figure B.9: Simulation of an optimized 12-branch extruded quadrature hybrid set for ALMA Band 3 (84 – 116 GHz).

C

Appendix 3

C.1 Abstract submitted for the *32nd IEEE International Symposium on Space Terahertz Technology (ISSTT 2022)*

Particle Swarm Algorithm Applied to Quadrature Hybrid Multi-Branch Directional Coupler Optimization for ALMA Band 3

Jorge Hernán Cárdenas [*1], Óscar Restrepo[2], Fausto Patricio Mena[3], David Monasterio[4] and Germán Chaparro[1]

***Abstract*—We introduce a novel optimization method based on the evolutionary algorithm Particle Swarm Optimization (PSO). Simulations were conducted for the design of quadrature hybrids intended to operate in the 85-115 GHz band using a fully tuned and validated version of the algorithm. We present quadrature hybrids designs which are optimized with respect to operational requirements for the scattering parameters and amplitude imbalance. Furthermore, the resulting designs take into account machining constraints related to cost and feasibility requirements. This method can be easily extended to optimize other microwave devices and waveguides for radio astronomy applications, with the benefit of speeding up the design process as well as reducing the computational costs.**

Keywords—quadrature hybrid, directional coupler, optimization algorithm, amplitude imbalance, millimeter and submillimeter device.

I. Introduction

Radio astronomy has highly demanding standards for microwave devices, often requiring ultra-wide band operability, high sensitivity, and noise rejection. For this reason, research into the design of microwave components for band widening within realistic manufacturing constraints is an ongoing effort [1][2]. In this paper we introduce a novel geometry optimization based on an evolutionary algorithm, that has been applied to quadrature hybrids operating in ALMA (Atacama Large Millimeter/submillimeter Array) Band 3 (85-115 GHz). Our reported results are part of an ongoing research project focusing on the design of a quadrature hybrid covering the ultra-wide ALMA Band 2+3 (67-115 GHz). Our approach to the problem of quadrature hybrid optimization is based on the exploration of the geometric parameter space of standard multi-branch couplers using the Particle Swarm algorithm.

II. Classical Quadrature Hybrids

Multi-branch couplers are widely used in heterodyne receivers. Special attention has been paid to standard designs consisting of two main waveguides connected by branches without further modifications to the guides [4]. Even though an eight-branch model has been proven to comply with the requirements for ALMA Band 3 [3], it has been proposed to increase the number of branches or to use overmoded waveguides to extend the operational band [2]. Further significant modifications have been introduced to the hybrid geometry by drilling extra cavities to the main guides, to add ripples to the signal and consequently reducing the amplitude imbalance at the center of the band [5]-[7]. We applied our method to 8-, 12- and 14-branch geometries, while also including extrusions in the main waveguide to assess their impact on the amplitude imbalance.

III. Quadrature Hybrid Optimization

Our optimization method is based on the Particle Swarm (PSO) evolutionary algorithm together with the Surrogate-Based Optimization (SBO) heuristic method. The algorithm seeks to find the waveguide and branches dimensions stored in a vector called a *particle,* that minimizes a fitness function (FF), while complying with any electromagnetic constraints. In this work we have used those set by ALMA, namely: a) scattering parameters S_{31} and S_{21} are expected to be close to -3 $\pm$0.5 dB, which also ensures keeping the amplitude imbalance below 1 dB [5], and b) the S_{11} and S_{41} parameters must stay below -15 dB. To allow the use of standard drilling tools, we set an additional constraint, defined as the ratio between the height of the main waveguide and the width of the branches, with a maximum value of 7.

As for hyperparameter and FF weights tuning, we settled on a scheme that we benchmarked against a well-defined and characterized quadrature hybrid designed for ALMA Band 5 (163–211 GHz) [9]. After we ran several optimization batches with this geometry as a starting point to fine-tune the algorithm and validate our resulting hybrids, we obtained an optimized geometry that improves the result obtained in [9] (see Fig. 1(a)).

[1]Institue of Physics, University of Antioquia, Medellín, Colombia (jorge.cardenas3@udea.edu.co) and (german.chaparro@udea.edu.co); [2]ECCI University and Universidad de Chile, Bogotá, Colombia (orestrepog@ecci.edu.co); [3]NRAO, Charlottesville, VA, USA (pmena@nrao.edu); [4]DAS Universidad de Chile, Santiago 8370451, Chile (David.monasterio@raig.uchile.cl).

NOTES:

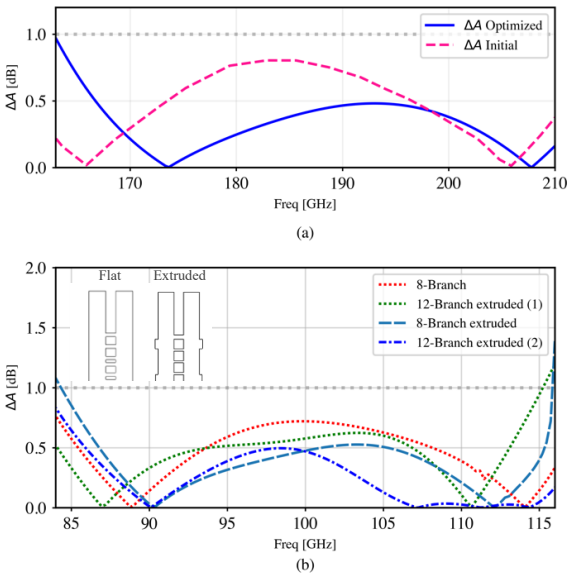


Fig. 1. Optimized quadrature hybrid designs (a) for ALMA Band 5 and (b) ALMA Band 3.

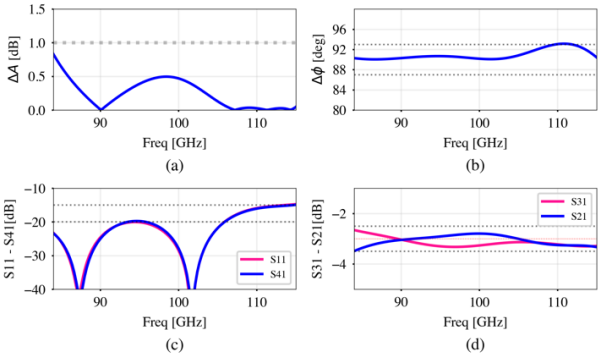


Fig. 2. Resulting (a) Amplitude imabalance, (b) phase imbalance and S-parameters (c,d) for a 12-branch quadrature hybrid incluiding extrusions optimized for ALMA Band 3.

IV. Results in ALMA Band 3

The resulting quadrature hybrid designs obtained for ALMA Band 3 using our optimization method show an excellent simulated performance vis-à-vis the S-parameter and amplitude imbalance requirements. Moreover, resulting geometries are found to be comparable or to improve upon prior results for ALMA Band 3 by fully covering the band with an amplitude imbalance below 1 dB. Fig. 1(b) shows the results for different standard geometries with amplitude imbalance under the expected value, with some hybrids staying below 0.5 dB. The 12-branch hybrid also shows a good S_{11}-S_{41} behavior (Fig. 2) thus an excellent amplitude imbalance contributing to an appropriate IRR [8], all that making it subject of interest for the further analysis and possible fabrication. We performed an additional optimization batch based on our 14-branch hybrid result aimed at widening the current band by an additional 10 GHz, yielding a band spanning from 80 GHz to 116 GHz with full compliance to the established constraints.

V. Conclusions

This novel optimization method provides validated and consistent results for quadrature hybrid designs that comply with electromagnetic constraints that would allow it to operate and improve upon current ALMA Band 3 designs. This method speeds up the design process and it can be extended to the development of other microwave devices for radio telescopes such as feedhorns, orthomode transducers (OMT) and even microstrips. We also found that standard geometries do not provide a viable solution to find a quadrature hybrid to fully cover the entire ALMA band 2+3, as we were only able to reach a solution compatible with (but not wider than) Band W under realistic manufacturing constraints. To achieve the whole coverage, attention must be paid to alternative and novel geometries that can be optimized using the tool herein proposed; this is part of our ongoing work.

References

[1] P. Yagoubov et al., "Wideband 67-116 GHz receiver development for ALMA Band 2," Astronomy and Astrophysics, vol. 634, no. 36777, pp. 1–23, 2020.

[2] A. Gonzalez, T. Kojima, K. Kaneko, and S. Asayama, "275-500 GHz Waveguide Diplexer to Combine Local Oscillators for Different Frequency Bands," IEEE Transactions on Terahertz Science and Technology, vol. 7, no. 6, pp. 669–676, 2017.

[3] S. Srikanth and A. Kerr, "ALMA Memo 343: Waveguide quadrature hybrids for ALMA receivers," ALMA Memo Series, vol. 9, January 2001.

[4] Reed, "The Multiple Branch Waveguide Coupler," IEEE Transactions on Microwave Theory and Techniques, vol. 6, no. 4, pp. 398–403, Oct. 1958.

[5] J. Ding, Y. Zhao, J.-X. Ge, and S. Shi, "A 90° Waveguide Hybrid with Low Amplitude Imbalance in Full W-Band," Journal of Infrared, Millimeter, and Terahertz Waves, vol. 40, no. 4, pp. 429–434, Apr. 2019.

[6] M. A. Hamid and M. M. Yunik, "On the Design of Stepped Transmission-Line Transformers," IEEE Transactions on Microwave Theory and Techniques, vol. MTT-15, no. 9, pp. 528–529, 1967, doi: 10.1109/TMTT.1967.1126525.

[7] H. Rashid, V. Desmaris, V. Belitsky, M. Ruf, T. Bednorz, and A. Henkel, "Design of Wideband Waveguide Hybrid with Ultra-Low Amplitude Imbalance," IEEE Transactions on Terahertz Science and Technology, vol. 6, no. 1, pp. 83–90, 2016.

[8] A. Khudchenko, R. Hesper, A. M. Baryshev, J. Barkhof, and F. P. Mena, "Modular 2SB SIS Receiver for 600-720 GHz: Performance and Characterization Methods," IEEE Transactions on Terahertz Science and Technology, vol. 7, no. 1, pp. 2–9, 2017.

[9] B. Billade et al., "Performance of the first ALMA band 5 production cartridge," 22nd International Symposium on Space Terahertz Technology 2011, ISSTT 2011, vol. 2, no. 2, pp. 56–56, 2011.

NOTES:

D

Appendix 4

D.1 Poster presented in the *32nd IEEE International Symposium on Space Terahertz Technology (IS-STT 2022)*

Particle Swarm Algorithm Applied to Quadrature Hybrid Multi-Branch Directional Coupler Optimization for ALMA Band 3

UNIVERSIDAD DE ANTIOQUIA

Jorge Hernán Cárdenas, Germán Chaparro (Institute of Physics, University of Antioquia),
Óscar Restrepo (ECCI University and Universidad de Chile), Patricio Mena (NRAO), David Monasterio (DAS Universidad de Chile)

Introduction

We introduce a novel optimization method based on the evolutionary algorithm Particle Swarm Optimization (PSO). Simulations were conducted for the design of quadrature hybrids intended to operate in the ALMA Band 3 namely 85-115 GHz band using a fully tuned and validated version of the algorithm.

We present quadrature hybrid designs which are optimized with respect to operational requirements for the scattering parameters and amplitude imbalance.

Furthermore, the resulting designs consider machining constraints related to cost and feasibility requirements.

This method can be easily extended to optimize other microwave devices and waveguides for radio astronomy applications, as used by Restrepo, O. et al. (in prep.), with the benefit of speeding up the design process as well as reducing the computational costs

PSO Algorithm and Optimization Process

Our method was applied on 8-, 12- and 14-branch geometries, while also including extrusions in the main waveguide to assess their impact on the amplitude imbalance.

The algorithm seeks to find the waveguide and branches dimensions stored in a vector called a particle, that minimizes a fitness function (FF), while complying with any electromagnetic constraints. We have used those set by ALMA, namely: a) scattering parameters S_{31} and S_{21} close to -3 ± 0.5 dB, which also ensures keeping the amplitude imbalance below 1 dB [2], and b) the S_{11} and S_{41} parameters must stay below -15 dB.

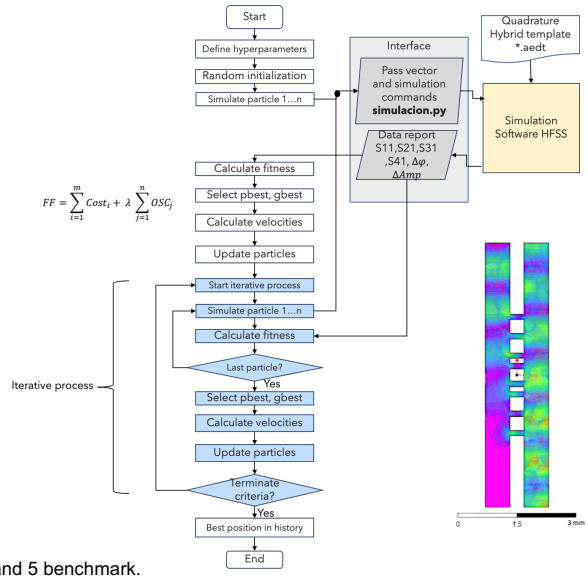
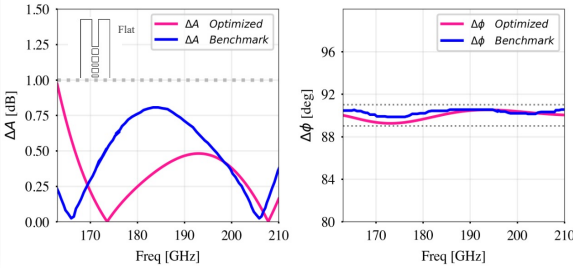


Figure 1. Validation results using ALMA Band 5 benchmark.



Our approach to validate the method consisted in using a quadrature hybrid previously studied for ALMA Band 5 [1]. The selected geometry consists of an 8-branch design with flat main waveguides.

The algorithm effectively improves the prior results by reducing the ΔA, while fulfilling other constraints as the $\Delta \phi = 90° \pm 1°$.

Optimized Quadrature Hybrids

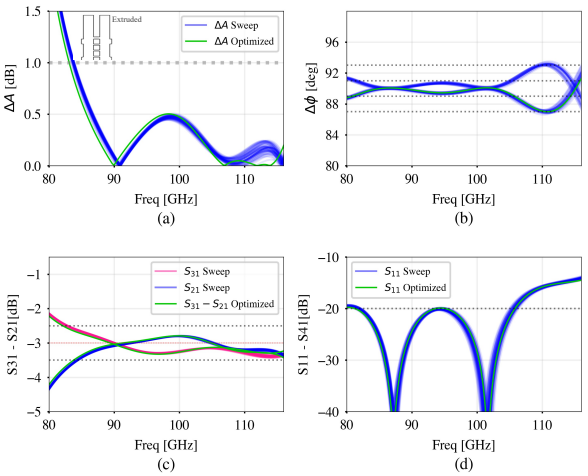


Figure 2. ALMA Band 3 Quadrature Hybrid parametric sweep.

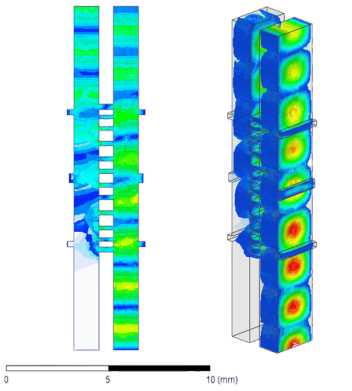


Figure 3. ALMA Band 3 optimized Quadrature Hybrid.

The resulting quadrature hybrid designs obtained for ALMA Band 3 show an excellent simulated performance vis-à-vis the S-parameter and amplitude imbalance requirements. Moreover, resulting geometries are found to be comparable or to improve upon prior results for ALMA Band 3 by fully covering the band with an amplitude imbalance below 1 dB.

The 12-branch hybrid also shows a good S_{11}-S_{41} behavior (Fig. 3) thus an excellent amplitude imbalance contributing to an appropriate IRR [3], all that making it subject of interest for the further analysis and fabrication.

Conclusions

1. The algorithm can effectively enhance the figures from other quadrature hybrid designs.
2. This method speeds up the design process and it can be extended to the development of other microwave devices for radio telescopes such as feedhorns, orthomode transducers (OMT) and even microstrips.
3. The 12-branch selected geometry fulfills the constraints even under parametric sweep analysis.
4. Attention must be paid to alternative and novel geometries that can be optimized using the tool herein proposed to achieve ultra-wide band performance namely ALMA Band 2+3.

References

[1] Billade, B., Nyström, O., Meledin, D., Sundin, E., Lapkin, I., Fredrixon, M., Desmaris, V., Rashid, H., Strandberg, M., Ferm, S. E., Wang, H., Xu, H., Obrocka, M., Ellison, B., Pavolotsky, A., & Belitsky, V. (2011). Performance of the first ALMA band 5 production cartridge. 22nd International Symposium on Space Terahertz Technology 2011.

[2] J. Ding, Y. Zhao, J.-X. Ge, and S. Shi, "A 90° Waveguide Hybrid with Low Amplitude Imbalance in Full W-Band," Journal of Infrared, Millimeter, and Terahertz Waves, vol. 40, no. 4, pp. 429–434, Apr. 2019.

[3] A. Khudchenko, R. Hesper, A. M. Baryshev, J. Barkhof, and F. P. Mena, "Modular 2SB SIS Receiver for 600-720 GHz: Performance and Characterization Methods," IEEE Transactions on Terahertz Science and Technology, vol. 7, no. 1, pp. 2–9, 2017.

Acknowledgements

Thanks to Germán and Oscar for their guidance and support; to Patricio and David for their valuable input; to the University of Antioquia for its financial support. To my wife Juliana and baby Emma.

CONTACT JORGE CARDENAS
University of Antioquia
Jorge.cardenas3@udea.edu.co

STUDENT POSTER

Baeza, Andalusia (Spain) | October 16 – 20, 2022

32ND INTERNATIONAL SYMPOSIUM ON SPACE TERAHERTZ TECHNOLOGY

E

Appendix 5

E.1 Poster presented in the VII CONGRESO COLOMBIANO DE ASTRONOMÍA Y ASTROFÍSICA *(COCOA 2022)*

UNIVERSIDAD DE ANTIOQUIA
Facultad de Ciencias Exactas y Naturales

Optimización de Híbrido en Cuadratura para la Banda 3 de ALMA Usando el Algoritmo Particle Swarm Optimization (PSO)

Jorge Hernán Cárdenas [1†] Oscar Restrepo [2‡] Germán Chaparro [1§]

[1]Instituto de Física, Universidad de Antioquia, Colombia. [2]Universidad ECCI & Universidad de Chile.

Objetivo

Optimizar el diseño geométrico y el comportamiento electromagnético para un híbrido en cuadratura (HC) a través del uso del algoritmo genético *Particle Swarm Optimization*, aplicado al desarrollo de un receptor en la banda 3 (84-116 GHz) para el observatorio ALMA (Atacama Large Millimeter/submillimeter Array).

Método

Se desarrolló un nuevo método de optimización basado en el algoritmo evolutivo *Particle Swarm Optimization* (PSO), aplicado al diseño de híbridos en cuadratura para receptores en la Banda 3 de ALMA. El algoritmo se desarrolló considerando como restricciones aquellas establecidas por ALMA: a) los parámetros S_{31} y S_{21} deben estar cerca de -3 ± 0.5 dB, asegurando un desbalance de amplitud inferior a 1 dB, y b) los parámetros S_{11} y S_{41} deben estar por debajo de -15 dB [1].

Diseño para fabricación

Cómo resultado llegamos a un diseño optimizado para operar en la **Banda 3 de ALMA**. El diseño cumple con las restricciones impuestas, lo que lo convierte en candidato para la etapa de fabricación.

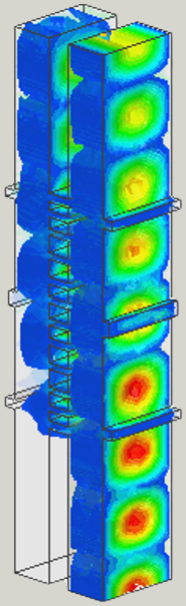


Figure 1. Híbrido en cuadratura optimizado para operar en ALMA Banda 3 (84 – 116 GHz).

Resultados

Las geometrías halladas muestran mejores resultados que diseños previos para la Banda 3 con un desbalance de amplitud inferior a 1 dB. La figura 2 muestra el resultado de optimización de dos geometrías estándar, con valores de ΔA que llegan a estar por debajo de 0.5 dB [2].

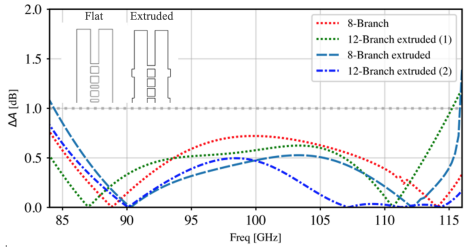
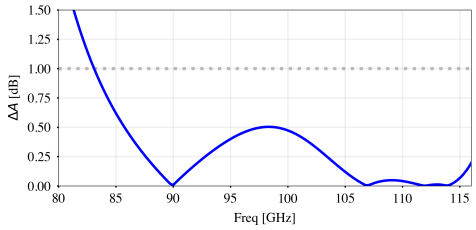
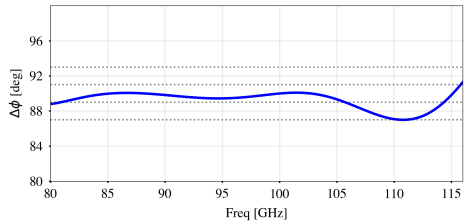


Figure 2. ΔA para diferentes geometrías de HC optimizados para la ALMA Banda 3.

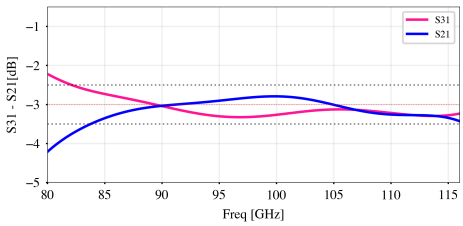
Con el proceso de optimización se encontraron HC que **(a)** cumplen con las restricciones de fabricación impuestas por ALMA y **(b)** se ajustan a una restricción impuesta para garantizar la fabricación usando herramientas de maquinado estándar (ver Fig. 3).



(a) Desbalance de amplitud.



(b) Desbalance de fase



(c) Parámetros $S_{31} - S_{21}$

Figure 3. Parámetros del híbrido en cuadratura seleccionado para fabricación.

Conclusiones

- El método provee resultados válidos y consistentes para el diseño de híbridos en cuadratura que satisfagan las restricciones electromagnéticas para operar y mejorar diseños actuales en la Banda 3 de ALMA.
- Este novedoso método se puede extender al desarrollo de otros dispositivos de microondas como *feedhorns* y *orthomode transducers (OMT)*, que son parte de radiotelescopios como ALMA.
- Este trabajo es de gran ayuda para otros observatorios que busquen poner en marcha receptores a un menor costo sin sacrificar la calidad de las mediciones.
- Se obtuvieron geometrías que cumplen con las restricciones impuestas por el observatorio ALMA para los híbridos en cuadratura que operen en sus receptores.

Perspectivas

El siguiente paso en esta investigación es la fabricación y caracterización de un híbrido en cuadratura resultante del proceso de optimización que aquí se reporta.

Referencias

[1] Itziar De Gregorio, Bill Dent, James Di Francesco, Jennifer Donovan Meyer, Darrel Emerson, and Daniel Espada. *ALMA Cycle 3 Technical Handbook*. 2015.

[2] Andrey Khudchenko, Ronald Hesper, Andrey M. Baryshev, Jan Barkhof, and Fausto Patricio Mena. Modular 2SB SIS Receiver for 600-720 GHz: Performance and Characterization Methods. *IEEE Transactions on Terahertz Science and Technology*, 7(1):2–9, 2017.

[3] Tony Mroczkowski, Carlos de Breuck, Ciska Kemper, Neil Phillips, Gary Fuller, Maite Beltrán, Robert Laing, Gianni Marconi, Leonardo Testi, Pavel Yagoubov, Danielle George, and William McGenn. Wide Bandwidth Considerations for ALMA Band 2. *arXiv*, pages 1–22, 2019.

[4] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

†jorge.cardenas3@udea.edu.co
‡orestrepog@ecci.edu.co
§german.chaparro@udea.edu.co

Faculty of Exact and Natural Sciences
Universidad de Antioquia
Medellín, Colombia
www.udea.edu.co