

Thesis Plan  
DOCTORATE IN ELECTRICAL AND COMPUTER ENGINEERING  
SPECIALIZATION IN MICRO-NANOELECTRONICS

NOVA University Lisbon

*Draft: June 25, 2025*

Thesis Plan  
DOCTORATE IN ELECTRICAL AND COMPUTER ENGINEERING  
SPECIALIZATION IN MICRO-NANOELECTRONICS  
NOVA University Lisbon  
*Draft: June 25, 2025*

## ABSTRACT

*Vagus nerve stimulation* (VNS) is an effective therapeutic medical procedure for the suppression of epileptic seizures, depression, mood and appetite control and chronic pulmonary obstructive disease (COPD). Low-intensity focused ultrasound (LIFUS) has been shown to be a promising non-invasive neuromodulation technique to perform VNS. The safety and effectiveness of the procedure is ensured by trained physicians and technicians operating the device while relying on real-time imaging to guide the procedure. Failing to acquire the target nerve results in poorly applied US sonication procedure, reducing the effectiveness of the therapeutic session. Mechanical or thermally-induced cellular tissue damage can also become evermore likely when aiming for finer and more precise stimulation procedures.

The need for an automated closed-loop control system is evident to ensure the focal spot is correctly positioned, ensuring optimal application of the stimulation procedure. Correct application of US sonication protocols not only has the potential to reduce therapeutic sessions duration, but also enabling improved exploratory medical procedures to study the regulatory function of the peripheral nervous system by reducing the number of control variables during scientific experimental procedures.

Integrating the acquisition of the target nerve's position closer to the sensor ultimately leads to a lower power and highly portable medical device that can evolve to a point-of-care device. This dissertation proposes a CMOS application-specific integrated circuit (ASIC) for real-time US imaging, integrating a capacitive gated recurrent unit (GRU) neural network within its readout signal chain capable of closing the loop of a control system aiming for automatic, secure and precise correction of the beamsteering procedure.

**Keywords:** One keyword, Another keyword, Yet another keyword, One keyword more, The last keyword

## RESUMO

Independentemente da língua em que a dissertação está escrita, geralmente esta contém pelo menos dois resumos: um resumo na mesma língua do texto principal e outro resumo numa outra língua.

A ordem dos resumos varia de acordo com a escola. Se a sua escola tiver regulamentos específicos sobre a ordem dos resumos, o template (L<sup>A</sup>T<sub>E</sub>X) NOVAthesis L<sup>A</sup>T<sub>E</sub>X (*novathesis*) irá respeitá-los. Caso contrário, a regra padrão no template *novathesis* é ter em primeiro lugar o resumo *no mesmo idioma do texto principal* e depois o resumo *no outro idioma*. Por exemplo, se a dissertação for escrita em português, a ordem dos resumos será primeiro o português e depois o inglês, seguido do texto principal em português. Se a dissertação for escrita em inglês, a ordem dos resumos será primeiro em inglês e depois em português, seguida do texto principal em inglês. No entanto, esse pedido pode ser personalizado adicionando um dos seguintes ao arquivo `5_packages.tex`.

```
\abstractorder(<MAIN_LANG>):={<LANG_1>,...,<LANG_N>}
```

Por exemplo, para um documento escrito em Alemão com resumos em Alemão, Inglês e Italiano (por esta ordem), pode usar-se:

```
\ntsetup{abstractorder={de={de,en,it}}}
```

Relativamente ao seu conteúdo, os resumos não devem ultrapassar uma página e frequentemente tentam responder às seguintes questões (é imprescindível a adaptação às práticas habituais da sua área científica):

1. Qual é o problema?
2. Porque é que é um problema interessante/desafiante?
3. Qual é a proposta de abordagem/solução?
4. Quais são as consequências/resultados da solução proposta?

**Palavras-chave:** Primeira palavra-chave, Outra palavra-chave, Mais uma palavra-chave, A última palavra-chave

# CONTENTS

## LIST OF FIGURES

# ACRONYMS

novathesis NOVAthesis L<sup>A</sup>T<sub>E</sub>X *ii*)

# INTRODUCTION

## 1.1 Motivation

### 1.1.1 World Health: Current context

The leading cause of illness and disability worldwide are neurological disorders, affecting approximately a third of the world's population [WorldHealthOrg]. Such disabilities can have a profound impact on memory, cognition, personality, movement, and essential physiological mechanisms such as breathing [empty citation]. The global trend on the occurrence of neurological disorders such as depression, epileptic seizures and Parkinson's disease has been toward a rapid increase, prevailing, in relevance, alongside other current health epidemics such as cancer and cardiovascular disorders [empty citation]. In

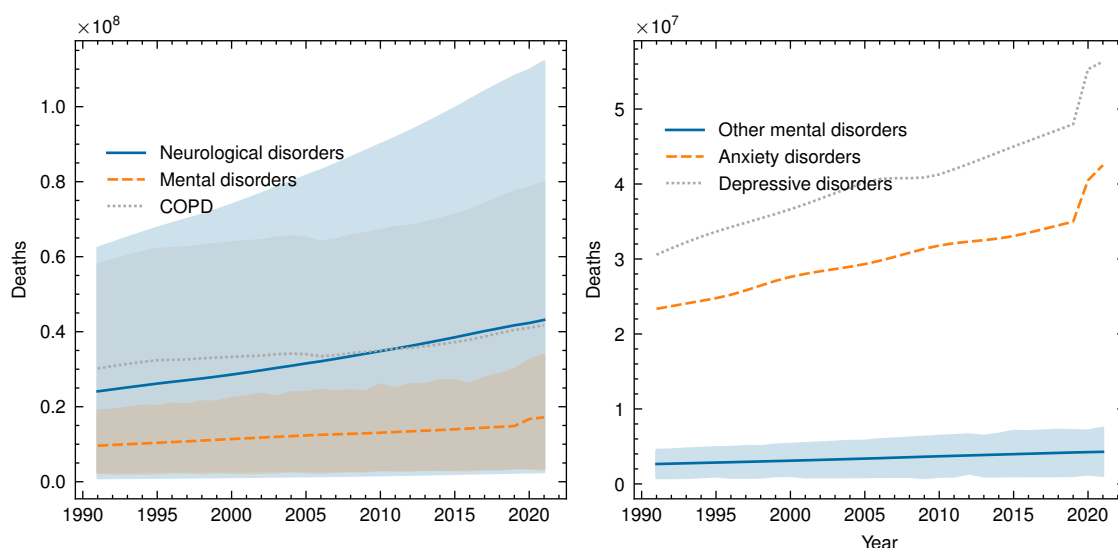


Figure 1.1: a) Yearly number of deaths associated to neurological, mental and chronic obstructive pulmonary diseases (COPD), according to the WHO, b) Discrimination of number of deaths associated to depressive and anxiety disorders relative to other mental disorders. Data retrieved from [<https://ghdx.healthdata.org/record/ihme-data/gbd-2021-nervous-system-disorders-1990-2021>]



light of the current world health context, the development of new clinical therapeutic and diagnostic interventions is of utmost importance, allowing safer, more comfortable, and highly effective alternatives to existing medical procedures.

Pharmaceuticals remain the traditional approach to the treatment of diseases. However, it is well established that the use of commercially available synthesized drugs affects the whole body, often inducing side effects that can lead to a decrease in patient quality of life [Goggins2022, Ji2020]. Moreover, recent pharmacoeconomic studies suggest that the research and development of new Pharmaceuticals costs range from \$161 million and \$4.54 billion, with additional costs of introducing new drugs into the market achieving \$880 million [https://www.nature.com/articles/d41573-024-00130-3, 10.1007/s40273-021-01065-y]. The complexity of the nervous system renders the aforementioned costs to be significantly optimistic [empty citation]. The development of drugs capable of tackling neurological diseases efficiently is limited by the available drug delivery methods, especially when delivering pharmaceuticals directly to the brain [empty citation]. The semipermeable blood-brain barrier regulates the transfer of chemical compounds between the circulatory and nervous system, rendering existing pharmaceutical drugs quite ineffective. Developing targeted drug delivery systems requires a parallel investment in research and development that can achieve similar investment magnitudes to the ones redirected towards drug development [empty citation]. This increase in spending has not produced positive results. While the global spending in new pharmaceutical drugs' research and development (R&D) has featured an increasing trend for the past 30 years, the same rising trend can be observed in the casualties associated to neurological, mental and chronic obstructive pulmonary disorders (COPD) (Fig. ??) - suffering a drastic increase during and post-COVID-19 years.

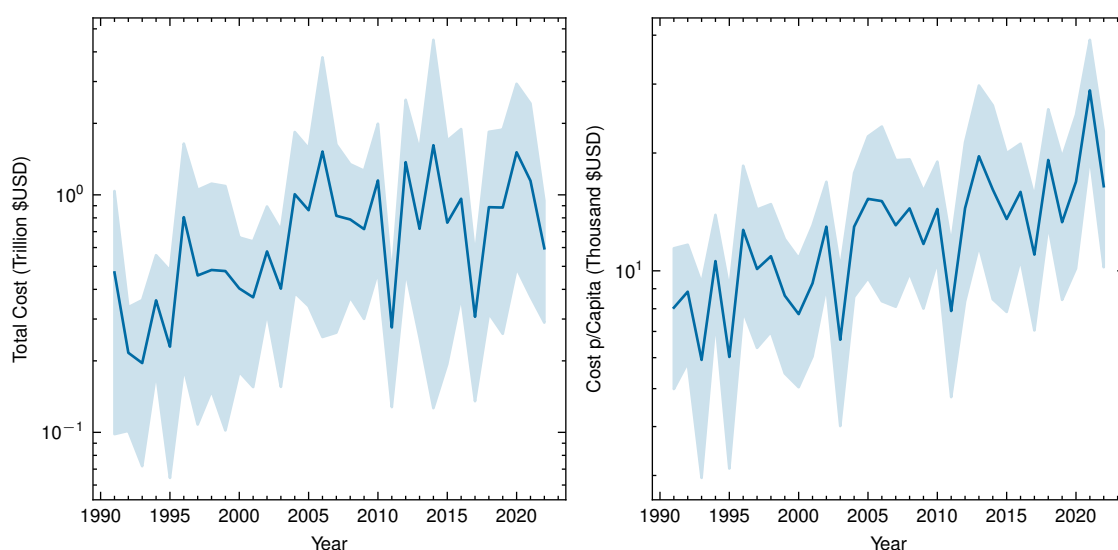


Figure 1.2: Yearly global pharmaceutical drug R&D spending in US Dollars (\$USD) a) in total b) per capita [https://github.com/datasets/pharmaceutical-drug-spending].

The total health-care spending is only projected to grow, leading scientists to believe that pharmaceutical drug R&D has been, so far, quite ineffective in the task of tackling diseases directly linked to the nervous system [**<empty citation>**]. While prevention will always be the most affordable and effective method to tackle most diseases, the genetic nature most neurological medical conditions feature remains a challenge that has been most effectively tackled by non-pharmaceutical treatments directly targeting the nervous system of each clinical subject [**<empty citation>**].

### 1.1.2 Vagus Nerve: A window of opportunity

The vagus nerve (VN) is the longest cranial nerve in the human body, extending from the brainstem to the abdomen through multiple organs such as the heart, lungs, stomach and intestines. The VN is responsible for the regulation of the parasympathetic nervous system (PNS). Due to its central role on the function of the PNS, multiple studies have shown that the VN can be stimulated to offer therapeutic procedures to conditions such as blood-pressure (BP) regulation, epileptic seizure suppression, mood and appetite control and chronic pulmonary obstructive disease (COPD) [**Goggins2022, Ji2020, Lescrauwaet2022, Browning2017, Undem2005**]. Regarding mood control, vagus nerve stimulation (VNS) has also been shown as a potent intervention to tackle depression, once it offers targeted therapeutic methods modulating the PNS [**DepressionVNS**]. Epileptic seizure suppression [**<empty citation>**] and depression and mood control [**<empty citation>**], are examples of clinically approved VNS therapeutic procedures that have already underwent clinical trials, and are currently making an impact on the lives of the respective patients. Currently undergoing pre-clinical and clinical trials are the therapeutic applications of blood pressure regulation, appetite control for tackling the obesity epidemic [**<empty citation>**], chronic obstructive pulmonary disease [**<empty citation>**] and Chron's disease [**<empty citation>**].

Neuromodulation-based therapeutic treatments have emerged as a promising method for tackling neurological disorders. Stimulation and suppression of neurological circuits by, respectively, activating or inhibiting neuronal activity enable the modulation of physiological and neurological functions [**Fomenko2018, Dalecki2004**]. Several modalities of clinical procedures exist to perform VNS.

\*Implantable VNS (iVNS)\* requires electrodes directly implanted into the VN, enabling its precise stimulation up to individual nerve fibers. However, this invasive method requires a surgical procedure to subcutaneously implant the electrical stimulation system and an electrode cuff wrapped around the VN near carotid artery (CA) and jugular vein (JV) in the neck region of the patient [**<empty citation>**]. The required surgical procedures introduces several unnecessary risks that also affect the time required to integrate implantable VNS applications in the market due to the required FDA approval [**<empty citation>**].

\*Transcutaneous VNS (tVNS)\* involves the stimulation of the VN using electrical currents induced by surface electrodes placed in direct contact with the skin of the patient [**<empty citation>**]. This non-invasive method of VNS usually targets the auricular and cervical branches of the VN, and requires an auricular device containing an electrode to induce the electrical currents in a skin surface that is close enough to the target nerve [**<empty citation>**]. While being less invasive, this method is incapable of achieving high-precision stimulation through the stimulation of specific regions of the VN. This limited spatial precision renders the VNS procedures to be very limited regarding the possible therapeutic outcomes and applications of this method [**<empty citation>**].

Due to its location is closer to the skin surface, targeting the VN through the human neck using low-intensity focused ultrasound (LIFUS) enables safer and effective non-invasive methods to perform VNS [Fomenko2018, Rivandi2023, Costa2021, Costa2019]. \*Ultrasound-based VN neuromodulation (fUS-VNS)\* has emerged as a promising method to enable spatially precise VNS procedures targetting a wide range of applications. Moderately increasing the fundamental, central frequency of the US focused beam enables a high lateral precision when delivering a significant amount of energy to the arbitrary VN region thorough the fUS focal spot [**<empty citation>**].

Moreover, existing studies on the impact of VNS on the nervous system of mammals are currently limited by the clinical apparatus required to conduct and measure the impact of VNS on the test subjects, especially when they're aimed at macaques or/and humans. The need for implantable VNS medical devices greatly increases the time it takes to obtain FDA approval to conduct such studies. The limited applicability of tVNS on the range of possible therapeutical applications makes fUS-VNS specially promising for next generation exploratory clinical devices [**<empty citation>**]. The ability to integrate both reliable imaging and stimulation systems in the same ASIC has become a much wanted goal and has sprouted an active research field on integrated US medical devices - not only for its implications on the future of point-of-care (POC) therapeutic devices for neurological disorders, but also for mitigating the procedural complexity of using fUS apparatus to target the VN upon conducting exploratory clinical procedures, as discussed in more detail in the next section [**<empty citation>**].

## 1.2 Clinical Ultrasound Practice

### 1.2.1 Ultrasound Neuromodulation Mechanisms

### 1.2.2 Ultrasound in the Clinic

## 1.3 Challenges in Ultrasound Image-Guided Neuromodulation Integrated Systems

### 1.3.1 Effective Vagus Nerve Stimulation

The efficiency of therapeutic procedures associated with VNS is guaranteed by ensuring that the sonication protocol is properly applied and US beam energy transfer to the target is optimized. The required lower US centre sonication frequency to achieve lower US signal attenuation when trying to perform non-invasive US imaging (and stimulation) of the vagus nerve through the human neck's muscle wall conditions the reconstructed US 2D US *brightness-mode* (B-mode) image to low spatial resolution [Szabo2014\_1\_8\_2\_Ultrasound]. Probe placement mismatch and the risk of US image mis-interpretation increases the influence of human-error when targetting the VN through the application of the control-loop presented in Fig. ?? e). Additionally, the vagus nerve can present six different positions according to reported anatomical studies, further increasing the difficulty of identifying the correct patient-specific location of the vagus nerve in real time, even for trained physicians [**<empty citation>**].

Finally, when targetting the VN while varying the US sonication protocol parameters to explore the influence of each parameter on the patient's response to the VNS procedure, the VN's location must be accurately recognized and targeted to ensure an higher confidence level on the results of the therapeutic procedure. Failing to target the VN leads to an ineffective clinical therapeutic or scientific exploration session [Ahmed2022].

### 1.3.2 Ultrasound Imaging ASIC Output Data-Rate and Power Dissipation Requirements

In general, two main methods of performing US imaging exist: 1) planar wave compounding (PWC) imaging and 2) scan-line (SL) imaging [Szabo2014]. SL imaging enables golden-standard US imaging quality, once ultrasound is two-way focused – an US beam-line is focused on a point in a numerically predefined imaging grid using a convex lens. Because the recovered echoes from that focused point in space are assumed to carry the majority of the power (assuming negligible secondary lobe energy), an inverse concave lens can be directly used to reconstruct the spatial information of the medium from the temporal information received. Contrarily, planar wave imaging methods such as PWC do not focus US in a single beam profile, but rather introduce controlled phase delays in the transmitted pulses to create a planar wavefront sonicating the whole imaging media at once. As such, during reception, the information of the medium arrives highly correlated

between each point of the imaging grid, and must be recovered simultaneously before starting the reconstruction of the US image. Both methods require high data-rates to transmit the high volume of samples generated by the US transducer array to the back-end data processing unit. Digital communication of the quantized information greatly increases the US imaging ASIC's output signal integrity. However, the high data-rate requirements lead to high power dissipation on the quantizer stages, the biasing circuitry, high-speed serial communication interfaces and ASIC output load drivers [Hopf2023, Chen2017, Tan2018, Jung2018].

$$P_{SER} = p_1 F_S C_{Line} V_{DD}^2 = 0.66 \cdot 50\text{MHz} \cdot 4\text{pF} \cdot (0.9\text{V})^2 = 27\text{mW} \quad (1.1)$$

As an example, let us consider an US imaging ASIC integrated within a PCB with a digital processing backend (Fig) performing an image acquisition. Depending on the metal trace width and length, the PCB connection between both chips can easily achieve a 1 - 4pF parasitic capacitance ( $C_{Line}$ ). The system features 32 channels, with an 8 bit quantizer per stage, operating at a nominal sampling frequency ( $F_S$ ) of 50 MHz. The total bit rate achieves  $32 \cdot 8b \cdot 50 \text{ MHz} = 1.6 \text{ Gbyte/s}$ . Considering that the US imaging ASIC's serial communication line, biased at a digital supply ( $V_{DD}$ ) of 0.9V, presents a 66% chance ( $p_1$ ) of transmitting a logic "1", then the total dynamic power dissipation of the serial transmission ( $P_{SER}$ ) can easily achieve 27mW, as observed from (??). If we consider that each AFE channel features a total power dissipation of 1.0mW, a value well within the state of the art, then the total power dissipation of the 32 channels is 32mW. Almost the same amount of power is dissipated in acquiring and processing the echo signals as well as exporting them out of the ASIC due to the very high data throughput required when performing US imaging. Moreover, this problem is aggravated when interfacing thousands of channels. A typical state-of-the-art US imaging system interfaces with a transducer matrix capable of 3D imaging by stacking multiple 2D B-mode image slices together, and to achieve enough resolution to enable the distinction of the complex anatomy of the human neck, at least 32-by-32 (1024) channels are required.

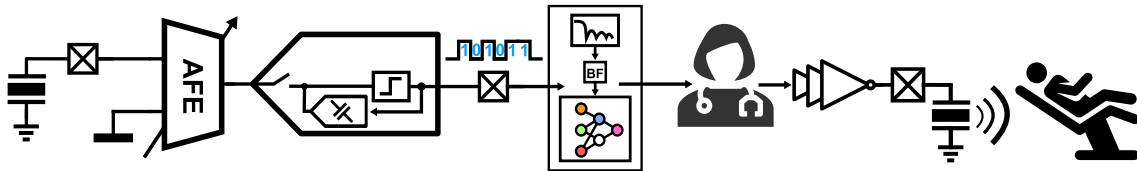


Figure 1.3: An envisaged US imaging ASIC in a typical fUS therapeutic clinical procedure information flux. An ultrasound image can be reconstructed from the recovered echoes, transported to the image reconstruction and data acquisition system in digital domain for higher signal integrity reliability.

## 1.4 Dissertation Goal and Thesis Plan Organization

Both preclinical and clinical studies can become great benefactors of miniaturized US transducer arrays that allow for targetting different regions of the VN, at different intensities and levels of spatial precision, once the variation of these parameters can lead to new insights and unlock new treatments. However, the targeting of the VN is a complex procedure that must be guaranteed to provide enough confidence level in the obtained results of the exploratory procedure, as well as guaranteeing the safety of future users of an envisaged wearable fUS VNS device. As discussed above, the guidance of the focused beam has been performed using bulky medical imaging devices, introducing a significant level of procedural and bureaucratic complexity that can lead to the cost increase of each study as well, as decreasing the celerity of the its conclusion. An integrated, minimally invasive, reconfigurable, wearable device capable of reliable US imaging as well as fUS stimulation has the potential remove these barriers.

The pursuit for the miniaturization of US imaging-enabled ASICs has led most research groups to tackle the problem of power dissipation reduction from an IC-design perspective solemnly. However, as further addressed, the decisions on how the integration between the sensor matrix and the interfacing integrated electronics are performed have the possibility of drastically impacting both the power dissipation of the data readout system as well as mechanically compressing the information on the phase of the received US echoes. Further compression can also be achieved by adopting a channel multiplexing strategy that effectively preserves the phase information of the received echoes, enabling an higher signal-to-noise (SNR) ratio in the reconstructed images. Moreover, these sensor integration decisions, combined with possible system-design choices, also allow vast amounts of free and additional space inside an IC that can be used to introduce a more complex, and refined control of power-efficient high-voltage (HV) pulsers to finally converge towards a wearable device capable of US imaging and stimulation that complies with FDA regulations regarding the safety of the device upon its direct interface with the patient. The focus of these work is the development of an highly miniaturized, and energy efficient US imaging device, capable of seamless integration within a low-form-factor wearable system with the ability of FUS beam-steering for VNS (Fig. ??).

A brief, but necessary introduction to US waves and their physical properties, the transducers used to perform a bidirectional translation of US into electrical signals, and US imaging algorithms and system specifications is discussed in Chapter ?. Insightful discussions on the current methods used to address the power dissipation limitations and the data throughput of an US imaging system are provided in Chapter ?. These discussions include a) which ultrasound transducer materials are best suited for each US application, with a special focus on the most suitable materials and transducer architectures for a FUS stimulation application and its implications regarding the image quality of these transducers, b) currently existing transducer matrix-interfacing electronics integration methods and their implications towards the quality and resolution of the reconstructed

US images and c) state-of-the-art ASICs for US echoes readout, with a special focus on the methods used to compress the number of output channels of the US RX system and the associated channel multiplexing schemes, promoting the power dissipation and output data-throughput reduction of the system. The implication of these channel multiplexing schemes on the compression of the sampled US echoes in the axis of time is also discussed, at it is of utmost importance to infer on the required number of discrete-time US echo samples required to reconstruct a reliable US image. Chapter ?? culminates in the research questions selected to be addressed in this dissertation. Finally, Chapter ?? provides a description of the methods used to select the scientific peer-reviewed material, and its sources, used to establish the ideas discussed throughout this dissertation plan, ending with a presentation of the devised work plan to be developed to achieve the goal of this dissertation.

# ULTRASOUND, TRANSDUCERS AND ULTRASOUND IMAGING

Fundamental concepts of \*what\* is an US wave, \*how\* it propagates and \*why\* does it present different behaviours upon its interaction with different materials is discussed in this introductory chapter. The spectral and temporal properties of a typical US echo in the two imaging modalities targeted within the scope of this dissertation - flow cytometry and 2D/3D B-mode - are also discussed. This chapter serves as a conceptual basis for justifying the system design premises presented in the rest of this dissertation.

## 2.1 Ultrasound Fundamentals

### 2.1.1 Wavefronts

Ultrasound is a mechanical wave that can propagate in a wide variety of shapes and \*wavefront\* morphologies throughout any material medium. It can propagate in solids, liquids and, with a little bit more difficulty, in gases. If it is a valid matter state, than sound can propagate in it.

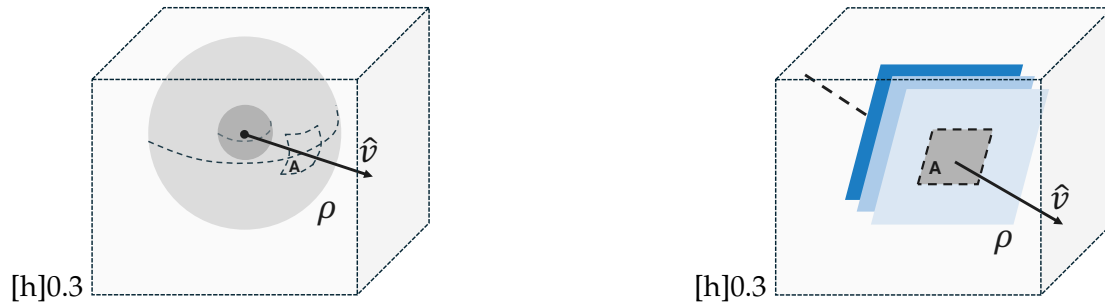


Figure 2.3: Three simple graphs

The two most important wave shapes for the aforementioned imaging and stimulation modalities are \*planar\* and \*spherical\*. Greater focus will be given to a planar wave propagating in the media in the following chapters, as it is the fundamental mechanism through which the modalities of USI discussed in this work are performed. A source



radiating a mechanical wave from a center point will lead to a spherical wavefront propagating with a velocity  $v$  in a medium with a density  $\rho$  ((Fig. ??)). Each wave can be decomposed into sequential infinitely extensible wavefronts that are collinear (or in other words, parallel) with each other - just like the ripples observed in the water after throwing a rock into a pond. As a single wavefront moves away from the source, the radius of the spherical surface defining this wavefront grows. An observer interacting with this wavefront at an infinite distance away from the source will observe no curvature in the wavefront - the velocity vector will be observed/sensed with the same phase in all the (square) finite elements in which the wavefront can be decomposed in. In this case the wavefront features a planar vectorial velocity field (Fig. ??). Upon interacting with an observing particle or \*scatterer\* in the medium, a \*force\* will be exerted on it. This force is directly proportional to the variation of momentum ( $d\hat{\gamma}/dt$ ) and inversely proportional to the area ( $A$ ) of the region in which this variational of momentum is observed. The ratio between the exerted force on the observer and the area of the observer's surface interacting with the wavefront defines the pressure  $p$  this same observer will be subjected to.

### 2.1.2 Mechanical Wave Equation in Fluids

As the target application of this dissertation is a medical imaging modality, it is important to take an in depth look at the underlying physical phenomena explaining the interactions briefly described in the previous subsection for fluids. Most of the *in-vivo* setups in which the target applications' scope fits into relies on the propagation of US waves in materials best-modelled by incompressible fluids - water, blood, soft tissue, brain tissue, muscle and fat.

In an ideal incompressible fluid, a particle is **displaced** ( $\mathbf{u}$ ) with a velocity  $\mathbf{v}$  when a wavefront propagates through a fluid with a resting-density  $\rho$  (??).

$$\mathbf{v} = \frac{d\mathbf{u}}{dt} \quad (2.1)$$

A velocity *potential* (vectorial field) between any point within the medium can be defined for convenience (??).

$$\text{equation)} \mathbf{v} = \nabla \phi$$

The pressure  $p$  to which any particle is subjected to is then defined by (??).

$$\text{equation)} p = -\rho \frac{\partial \phi}{\partial t}$$

The generalized wave equation describing the propagation of sound in an incompressible fluid can then be defined through (??).

$$\nabla \cdot \mathbf{v} - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0 \equiv \nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = 0 \quad (2.2)$$

When a wave is propagating in the medium, the wave equation enables us find to find regions of local pressure disturbance, with higher pressure *compressional* regions and - since the fluid is incompressible - lower pressure *rarefactional* regions (Fig. ??). These

regions arise from the displacement of particles within the fluid, and can occur in the collinear (Fig. ?? - longitudinal waves) or orthogonal (Fig. ?? - transversal) axis relative to the source of the US waves.

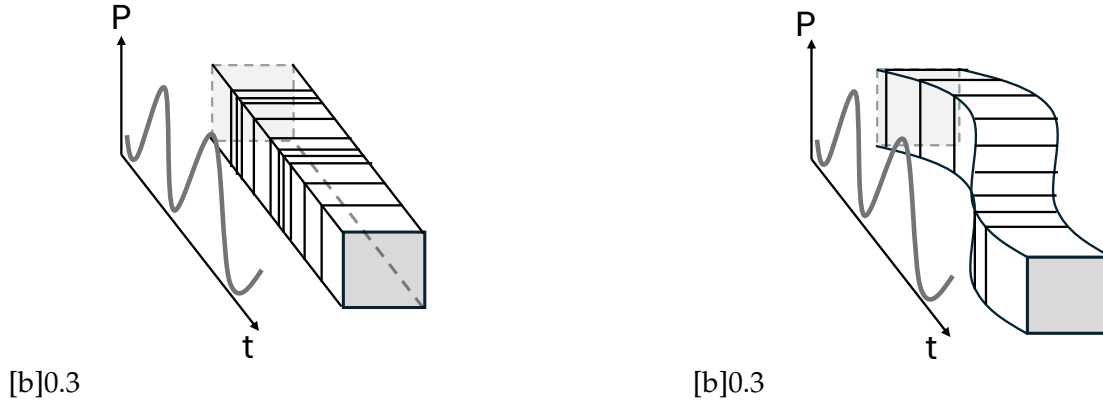


Figure 2.6: Three simple graphs

### 2.1.3 Characteristic Acoustic Impedance and Elastic Wave Propagation

In a contained volumetric medium, the propagation of an elastic (sound) wave in an incompressible liquid medium bounded by elastic boundary surface is well-modeled through a longitudinal wave propagating through the liquid. In such an environment, the higher the (particle/matter) density of the medium ( $\rho$ ), the faster a sound wave can propagate through it with a speed  $c_L$ .

Any particle travelling within the medium features some *impedance* to its displacement. The ratio of pressure ( $p$ ) of the mechanical wave displacing it to the particle's velocity ( $v_L$ ) defines the *acoustic characteristic impedance* ( $Z$ ) of the medium (??). A more useful definition of medium's  $Z$  - in the context of sound propagation - can be given through the product of the medium's density and the speed of sound in it.

$$Z = p/v_L = \rho c_L [\text{Rayl} \equiv \text{kg/m}^2\text{s}] \quad (2.3)$$

The propagation of a sound wave between two media with different acoustic characteristic impedances  $Z_1, Z_2$  and sound velocities  $c_1, c_2$  is conditioned by two phenomena - *reflection* and *diffraction* (Fig. ??). Diffraction is described using an equivalent of Snell's Law for elastic waves (??).

$$\frac{\sin \theta_{11}}{\sin \theta_{21}} = \frac{c_1}{c_2} \quad (2.4)$$

The mismatch (dissimilarity) between acoustic characteristic impedances determines the magnitude of the reflected wave, where the reflected wave features the same angle as the incident wave (Fig. ??). As in electromagnetic wave propagation, the *reflection factor* ( $RF$ ) can be defined - through the analysis of the propagation between media as a two-port

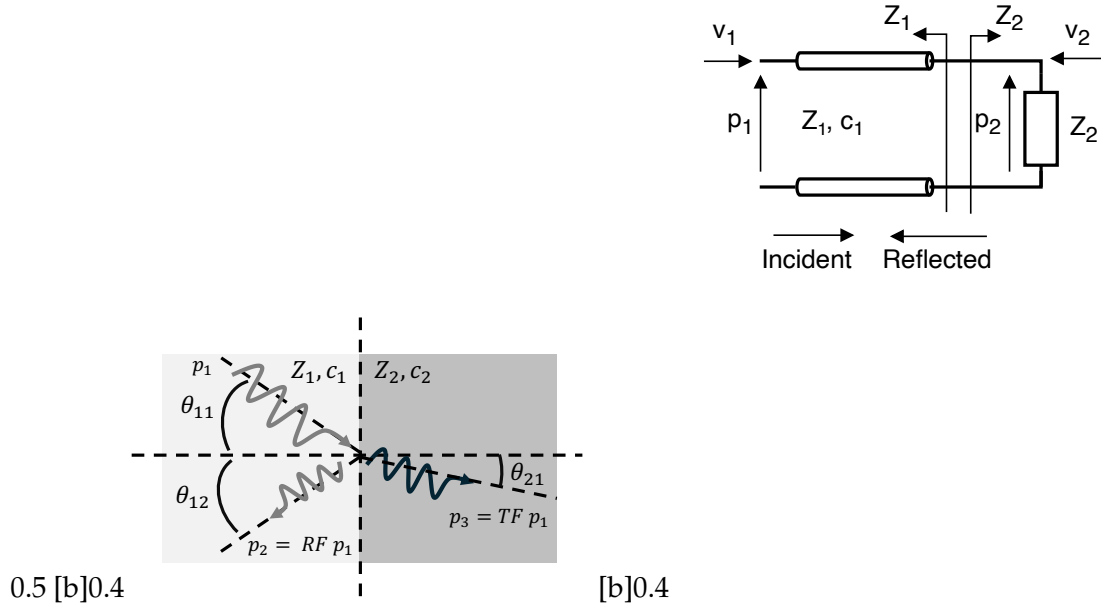


Figure 2.7: Three simple graphs

propagation problem - to determine *how much* of the incident wave's energy is *reflected back into medium 1* (??). Similarly, the *transmission factor* ( $TF$ ) is defined as the energy ratio of the signal that is transmitted between the two media (??), linearly proportional to  $RF$  - if  $RF$  grows in magnitude,  $TF$  will inevitably reduce in magnitude as well by the same proportion.

$$RF = \frac{Z_2 - Z_1}{Z_2 + Z_1} \quad (2.5)$$

$$TF = \frac{2 Z_2}{Z_2 + Z_1} \quad (2.6)$$

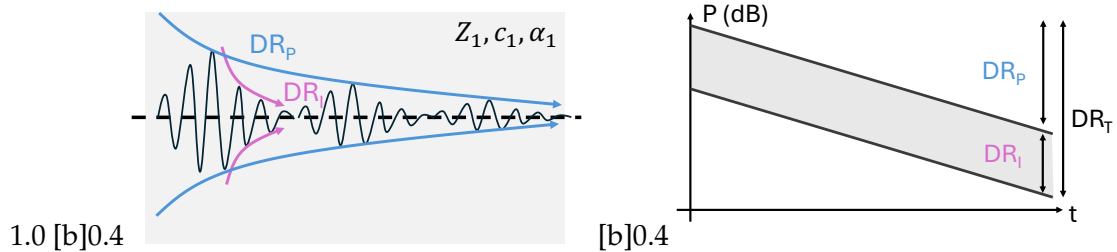


Figure 2.8: Three simple graphs

*Attenuation* is the governing phenomena leading to the progressive loss of energy of an elastic wave propagating in any media - mainly due to *energy scattering* and *energy absorption* phenomena intrinsic to the incompressible liquid(s) constituting the media (Fig. ??). While *absorption* is the main factor in the propagation of an elastic wave in an

*homogeneous medium*, scattering should also be taken into account for *heterogeneous media* - *i.e.* the human body. In diagnostic US applications it is often the case that the sonicated volumetric regions are relatively small, enabling soft tissue regions of the body to be well approximated by an homogeneous medium. Attenuation through absorption becomes then the main issue to address when developping readout systems for diagnostic US. Attenuation of an acoustic wave propagating within a medium ( $DR_P$  in dB) is linear-in-dB, and is represented by an exponent power factor that is directly proportional to the center frequency of the carrier of the wave ( $f_0$  in MHz), the *roundtrip* distance the wave (and in diagnostic US, its corresponding echo) travels ( $D_L$  in cm) and the attenuation coefficient of the material it propagates in ( $\alpha$  in  $\text{dB MHz}^{-1} \text{cm}^{-1}$ ) (??). Throughout this thesis, this propagation-induced attenuation ( $DR_P$ ) is also referred to as *propagation dynamic range*. US waves can be decomposed into a superposition of vector-basis functions that are gaussian-modulated cosine temporal impulse responses [**<empty citation>**]. Each gaussian distribution modulating the center tone has a 6-dB pass-band bandwidth that most often creates a 40 dB *instantaneous dynamic range* ( $DR_I$ ) on the impulse response of any low Q-factor, wide-bandwidth transducer preferable in most diagnostic US applications [**<empty citation>**] (Fig. ??). Superposed to  $DR_P$ , the *total dynamic range* ( $DR_T$ ) of a typical US echo verified in most diagnostic US applications can range from 60 dB to 120 dB. *instantaneous dynamic range* and other transducer-related concepts are further discussed in Sec. ?? with a greater depth.

$$DR_P = f_0 D_L \alpha \text{ [dB MHz}^{-1} \text{cm}^{-1}] \quad (2.7)$$

$$DR_T = DR_P + DR_I \text{ [dB]} \quad (2.8)$$

Common bio-materials found within the body of mammals feature attenuation coefficients typically leading to 30 - 60 dB of *propagation dynamic range*. The attenuation coefficients of such bio-materials can be found in table ??.

## 2.2 Ultrasound Transducers

In this section, the fundamental concepts enabling the representation of the US trasnducer in the electrical domain are described. The electrical modelling of any trasnducer geometry and material is fundamental to optimize the design of *unipolar* and *multipolar* drivers in *therapeutic* US. An accurate electrical impedance modelling of the trasnducer also enables the optimization of *DC-decoupling* and *input impedance* matching networks in the *low-noise amplifier* (LNA) located at the interface between the transducer and the AFE.

The *Butterworth-van-Dyke* (BVD) model [**VanDyke1928poti**] (Fig. ??) is a simplification of the Krimholtz, Leedom, and Matthaei (KLM) model [**Leedom1971itosau**] considering the latter is operated as a resonating network. While the KLM model is a complete two-port model allowing for the bidirectional direct translation between the acoustic

Table 2.1: Attenuation coefficients for commonly found biological mediums in the body of mammals [Culjat2010uim\&b]

Medium	Attenuation ( $\alpha$ ) [ $dB\ MHz^{-1}\ cm^{-1}$ ]
Air	1.64
Water	0.0022
Blood	0.2
Bone (Cortical)	6.9
Bone (Trabecular)	9.94
Brain	0.6
Breast	0.75
Cardiac	0.52
Connective Tissue	1.57
Fat	0.48
Marrow	0.5
Muscle	1.09
Tendon	4.7
Soft Tissue (Average)	0.54

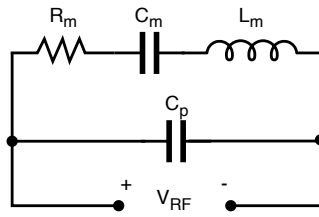
pressure interfacing with the piezoelectric transducer's surface and the electric potential within the top and bottom plate, the BVD model outputs an electrical output from the electrical input equivalent to a generated *displacement current* ( $I_d$ ) or *displacement voltage* ( $V_d$ ). Such a transformation in analysis-domain greatly simplifies the electrical modelling of the piezoelectric model from a pre-characterization process of its *charge-displacement* system matrix representation [Szabo2014duiioTransducers]. The BVD model includes a resonant branch ( $R_m$ ,  $L_m$ ,  $C_m$ ) modelling the mechanical resonance of the transducer, and a parasitic capacitor ( $C_p$ ) representing the electrical losses dampening the resonator branch over time Fig. ??).

$$Q = f_0/BW \quad (2.9)$$

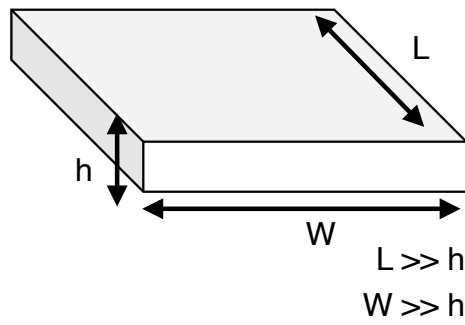
## 2.3 Therapeutic & Diagnostic Ultrasound

The pressure generated at the surface of the elements depends on the amplitude of the applied pulse, and the transmit efficiency (expressed in Pa/V) of the transducer. In the case of focused transmission, the pulses are timed by the transmit beamformer such that the generated acoustic waves converge at the desired focal point. The pressure at this focal point can then be found by taking into account the TX beamforming gain and the propagation attenuation in the medium.

The acoustic wave will reflect from interfaces between regions with different acoustic impedance in the medium, leading to echoes that return to the transducer. The amplitude of an echo depends on the reflection coefficient associated with the interface, and the

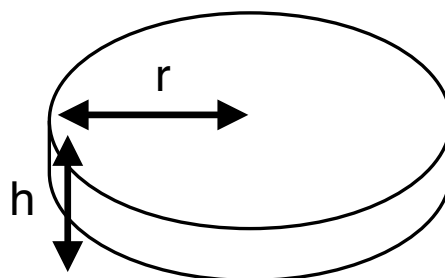


□□[t]0.3



□□[t]0.3

□□[b]0.3



geometrical spreading and attenuation that the echo experiences as it travels back to the transducer. The resulting surface pressure at one of the elements then translates to a voltage through the transducer's receive sensitivity (expressed in V/Pa). This signal is amplified by the LNA and TGC and digitized, and finally combined with the signals from other elements by the receive beamformer. Upon performing delay-and-sum (DAS), the correlated signals from the different channels add up constructively, while uncorrelated noise does not, giving an RX beamforming gain and leading to the creation of a pressure profile through time, leading to the reconstruction of an US B-mode image upon quadrature-filtering and log-compression of the beamformed RX US signals.

## STATE-OF-THE-ART

### 3.1 Ultrasound Transducers

#### 3.1.1 Ultrasound Transducer Materials

It is impossible to establish a review of existing ultrasound imaging (USI) and ultrasound stimulation (USS) systems without first highlighting the current materials used to implement the ultrasound transducers (USTs). The choice of material impacts frequency range, transducing power efficiency, biocompatibility, the architecture of the interfacing electronics and thus the cost of high-volume manufacturing of the transducers, the interfacing electronic system and the integration of the former with the latter. Due to their miniaturisation, moderate-to-high centre frequency range and moderate-to-high electromechanical transducing efficiency (resulting in lower driving power dissipation) capabilities, this sub-section focuses on the review of the benefits and limitations of UST implemented using bulk piezoelectric materials, piezoelectric micromachined ultrasound transducers (PMUTs) and capacitive micromachined ultrasound transducers (CMUTs) - transducing materials dominant in medical applications. Notice that bulk piezoelectric and PMUTs both use piezoelectric materials for implementing the UST, but differ in manufacturing process, geometry and transducer electromechanical coefficients - affecting also their equivalent electric model used upon the development of the interfacing electronics.

Modern bulk piezoelectric transducer materials have the advantage of exhibiting moderate-to-high electromechanical coupling coefficients while presenting high quality factors. Lead zirconate titanate (PZT), a ceramic-based piezoelectric material, and lead magnesium niobate-lead titanate (PMN-PT), a relaxor ferroelectric-based piezoelectric material, have been widely adopted in recent years due to their electromechanical properties enabling the miniaturisation and compact integration of the transducers with the interfacing electronics while still providing for a wide range of existing and emerging medical applications [<https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>].

Bulk PZT transducer materials present thickness-mode electromechanical coupling coefficients ( $k_{33}$ ) typically ranging from 0.5 to 0.75, while presenting mechanical quality factors ( $Q_m$ ) ranging between 30 and 80 at a high acoustic impedance typically found within 33



to 35 MRayl [<https://link.springer.com/book/10.1007/978-0-387-76540-2> Fig 10.12] [<https://iopscience.iop.org>]. The presented coupling coefficient enable a lower driving power when aiming for high-intensity focal spot profiles. Due to its high acoustic impedance, PZT transducers require matching layers to minimize acoustic energy losses associated to the impedance mismatch between the human body and the transducer to improve the efficiency and efficacy of the US therapeutic system [<https://pubmed.ncbi.nlm.nih.gov/32708159/>] [<https://www.nature.com/articles/srep42>]. While therapeutic applications profit from high  $Q_m$ , resulting in narrow-band, long response time transducers to maximize energy transmitted versus provided driving stimuli, diagnostic applications require broad-band, short impulse response transducers featuring low  $Q_m$  to maximize the lateral resolution of the reconstructed US image. As such, for diagnostic applications PZT transducers require a mechanical or electrical mechanism to lower their  $Q_m$  [[empty citation](#)][[empty citation](#)]. Damping backing layers have been explored in the context of US applications capable of diagnostic and therapeutic operation modes [[empty citation](#)]. However, this mechanical approach renders the backed transducers ineffective for the transmission of high-intensity focused US beams, and transducer arrays must feature a configuration of diagnostic and therapeutic-optimised transducers that usually compromises on the diagnostic capabilities of the system [[Hassan Random distribution](#)][[Chao Chen Cross](#)]. Electronic feedback systems have been recently started to be explored in the context of electrically lowering the  $Q_m$  factor of the transducer when operating in diagnostic (imaging) mode [<https://ieeexplore.ieee.org/document/8589657>]. By coupling an electrical resonator with a low quality factor to the transducer with a high quality factor, the latter can be lowered by controlling the strength of the coupling between both resonators. Increasing the amplitude of the electrical resonator's output will lead to a lower quality factor for the transducer, enabling it to maintain optimal impulse response in both therapeutic and diagnostic applications. This is a very recent and active research track, and it is being actively pursued at the research.

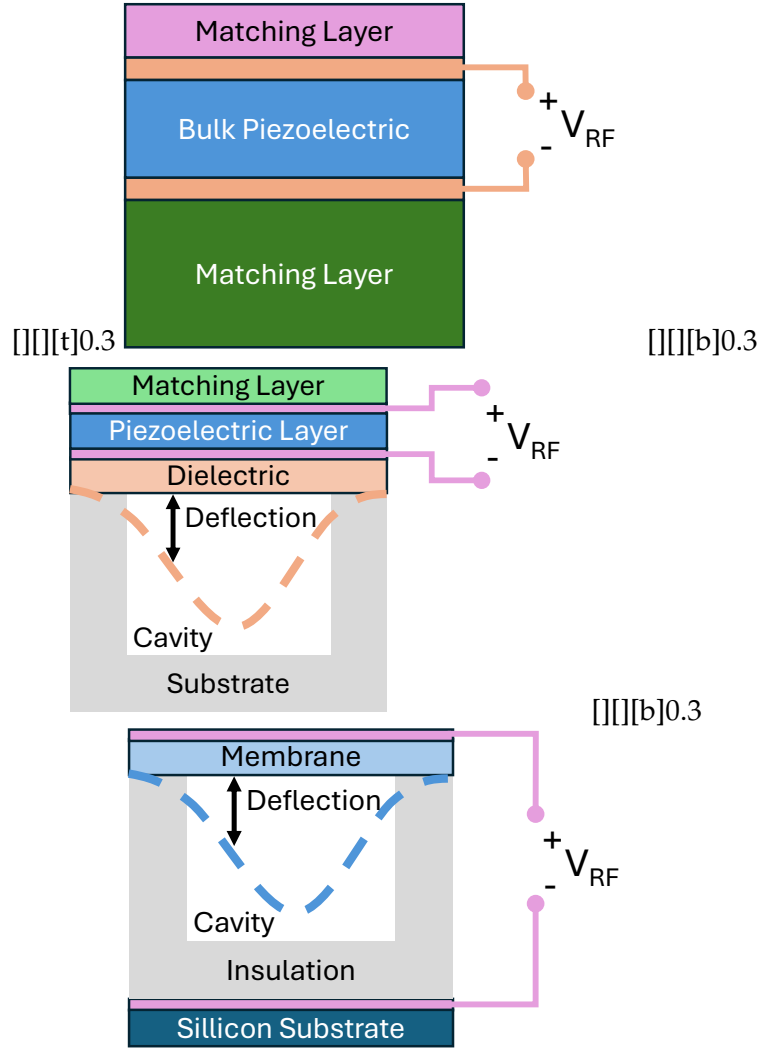
Similar to PZT, PMN-PT also experiences significant acoustic impedance mismatch and acoustic energy losses when being interfaced with human tissue, due to its high acoustic impedance ranging from 35 to 40 MRayl depending on the percentage of lead titanate (PT) present in the composite [<https://pubmed.ncbi.nlm.nih.gov/32708159/>]. PMN-PT transducers also feature a high  $Q_m$  in the same order of magnitude as that of PZT, ranging from 150 to 180, needing to be compensated for when aiming for their usage on US diagnostic applications as well. However, the electromechanical coupling coefficient of PMN-PT ranges from 0.85 up to 0.95, providing for a lower required driving power to achieve similar focal spot intensities when compared to PZT transducers. Additionally, the higher coupling coefficient of PMN-PT transducers also enables for an higher sensitivity, meaning that a similar strain force stimuli can generate a greater charge displacement when compared to PZT, and thus an higher voltage signal can be read by the interfacing electronics in diagnostic applications, relaxing their input referred noise requirements and reducing their area and power dissipation requirements [<https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>].

The Curie point of a piezoelectric materials is the temperature at which the piezoelectricity properties are permanently altered from its original conditions, and in general, piezoelectric materials lose the ability of charge-displacement past this point [<https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>]. The development of suitable manufacturing and integration procedures of piezoelectric transducer arrays that do not change their underlying physical properties is, as such, an active research field and crucial for the pursuit of evermore miniaturised and energy efficient US-based systems and their applications. PZT presents an higher Curie temperature (350 °C) than that of PMN-PT (130 °C–170 °C), introducing significant challenges in the fabrication of PMN-PT devices, as well as transferring the material onto backing substrates without damage. This introduces harsh limitations on the range of applications of PMN-PT. The electromechanical transducing capabilities of PMN-PT come at the expense of an increased complexity regarding the manufacturing and integration process of the transducer arrays when compared to PZT [<https://doi.org/10.1533/9781845699758.1.318>, <https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>]. Another great advantage of PMN-PT-based transducers is tied to its DC-bias sensitiveness. Some US transducers provide a acoustic to electric transduction sensitivity that depends on the DC-bias voltage level between both electrode plates of the transducer. Example of this phenomena are CMUTs and PMN-PT or PIN-PMN-PT based transducers. More specifically, bulk PMN-PT transducers only exhibit electromechanical conversion capability while a bias voltage is applied, and it has been verified that an inversion of the DC-bias level introduces a 90 ° phase shift in the received or transmitted US waves relative to a forward-biased transducer [[10.1109/ULTSYM.2016.7728854](https://doi.org/10.1109/ULTSYM.2016.7728854), [10.1109/ULTSYM.2018.8579821](https://doi.org/10.1109/ULTSYM.2018.8579821)]. This effect has been successfully explored to enable the implementation of an electronically-reconfigurable acoustic Fresnel Lens paving new unexplored research paths towards the miniaturisation and reduction of the power dissipation of imaging/HIFU-capable US row-column transducer array applications [[10.1109/ULTSYM.2018.8579821](https://doi.org/10.1109/ULTSYM.2018.8579821), [10.1109/ULTSYM.2018.8579821](https://doi.org/10.1109/ULTSYM.2018.8579821), [10.1109/ULTSYM.2019.8926257](https://doi.org/10.1109/ULTSYM.2019.8926257)].

Both previously discussed transducer materials contain lead, which can lead to significant environmental damage when the devices are improperly disposed off [<https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>, <https://doi.org/10.1533/9781845699758.1.318>, <https://doi.org/10.1016/j.ceramint.2021.03.054>]. Additionally, although not well understood, existing concerns have lead recent studies regarding the biocompatibility of using lead-based materials in medical implantable devices that are placed in direct contact with the patient's tissue. Lead is a toxic heavy metal that can lead to significant health complications of patients when their body tissue comes into contact with the lead-based material for long periods of time, resulting in a growing research effort to implement lead-free ultrasound transducer materials [<https://doi.org/10.1016/j.ceramint.2021.03.054>, <https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>]. Polymer-based, lead-free transducer materials have emerged as a promising alternative to PZT and PMN-PT based piezoelectric materials, offering greater biocompatibility and less environmental damage risk [<https://iopscience.iop.org/article/10.1088/1361-6463/ac8687>]. Polymer-based piezoelectric materials such as polyvinylidene fluoride (PVDF), barium

titanate ( $\text{BaTiO}_3$ ) and potassium sodium niobate (KNN) have emerged as potential replacement for lead-based ceramic composite piezoelectric materials. Compared to lead-based bulk piezoelectric materials (presenting a brittle texture due to their composite ceramic matrix lattice), polymer piezoelectric materials are flexible and lightweight, while enabling a wider range of manufacturing processes in different substrates, presenting lesser complexity manufacturing and electronic integration of transducer arrays [<https://doi.org/10.1016/j.ceramint.2021.03.054>]. As an example, PVDF provides an acoustic impedance ranging from 4-10 MRayl, being closer to that of human-tissue [<https://ieeexplore.ieee.org/abstract/document/9354791>]. In the scope of biomedical applications, the increased flexibility lowers the acoustic impedance of the transducer, offering a greater conformability and lower acoustic impedance mismatch towards human tissue - increasing the efficiency of acoustic energy transfer between the human body and the transducer, making them specially suitable for US diagnostic applications [<https://doi.org/10.1016/j.ceramint.2021.03.054>]. While still requiring matching layers to maximise the mechanical energy transfer between the transducer and human body, the lower transducer array manufacturing and integration complexity also increases the yield of transducer elements within the large-area transducer arrays, which is currently an important challenge in transducer arrays employing bulk lead-based piezoelectric materials containing >1000 elements [<https://www.nature.com/articles/s41467-024-47074-1>]. However, the increased flexibility, decreased manufacturing and integration complexity and increased biocompatibility of polymer-based piezoelectric transducer materials comes at the expense of lower electromechanical coupling and charge displacement coefficients [<https://doi.org/10.1016/j.ceramint.2021.03.054>]. Despite presenting the highest piezoelectricity performance among piezoelectric polymers, PVDF presents an electromechanical coupling coefficient ranging from 0.2 to 0.5 in stacked PVDF-based composite materials [<https://doi.org/10.1016/j.jmbbm.2021.104669>, <https://doi.org/10.1002/adfm.201908724>, <https://doi.org/10.1002/adfm.201908724>, <https://doi.org/10.1002/adfm.201908724>]. Compared to PZT or PMN-PT-based materials, the reduced electromechanical coupling coefficient of polymer-based piezoelectric materials can lead to a doubling of the power dissipation of the driving electronics in therapeutic US applications. The additional power dissipation reduces the portability of the therapeutic device and, in some cases, can lead to a violation of FDA guidelines regarding the maximum power of 0.5 W to prevent the uncontrolled temperature increase of human tissue, rendering polymer-based piezoelectric materials unsuitable for the design of US stimulation devices [**Fomenko2018**, **Dalecki2004**].

The last two classes of transducer discussed in this section, PMUTs and CMUTs, consist of material membranes micromachined to a few micron-level thickness, enabling the softened materials to mechanically interact with US pressure waves with significantly lower acoustic impedance, working on the principle of the flexural vibration of a thin membrane film. The differences between PMUTs and CMUTs lie on the physical phenomena governing the transduction between electrical signals and the mechanical displacement of the membrane [<https://doi.org/10.1109/ACCESS.2024.3359906>]. In

Figure 3.1: Creating subfigures in  $\LaTeX$ .

general, the enhanced acoustic impedance matching removes the need for matching layers when interfacing the transducer with soft tissue, consequently decreasing the complexity of transducer integration processes [<https://doi.org/10.1038/s41378-023-00555-7>, <https://doi.org/10.1109/TUFFC.2021.3112917>]. PMUTs are micromembrane ultrasound transducers exhibiting thin-film piezoelectricity, backed by an acoustic cavity and encapsulated between two flexible metal electrode layers. PZT (lead-based), PVDF and more recently aluminium nitride (AlN) (lead-free alternatives to PZT) are amongst the most widely adopted PMUT materials, and the lower acoustic impedance reduces the quality factor of the transducers, increasing their bandwidth and consequently axial resolution of obtained US images - making PMUTs especially suitable for US diagnostic applications [<https://doi.org/10.1109/ACCESS.2024.3359906>, [10.1109/TRANSDUCERS.2019.8808774](https://doi.org/10.1109/TRANSDUCERS.2019.8808774), [10.1109/ULTSYM.2019.8926015](https://doi.org/10.1109/ULTSYM.2019.8926015)]. PMUT transducers are still operated in its poling direction (3-3 electromechanical coupling mode), but the significantly lower thickness geometry significantly changes the electromechanical properties of the piezoelectric material when

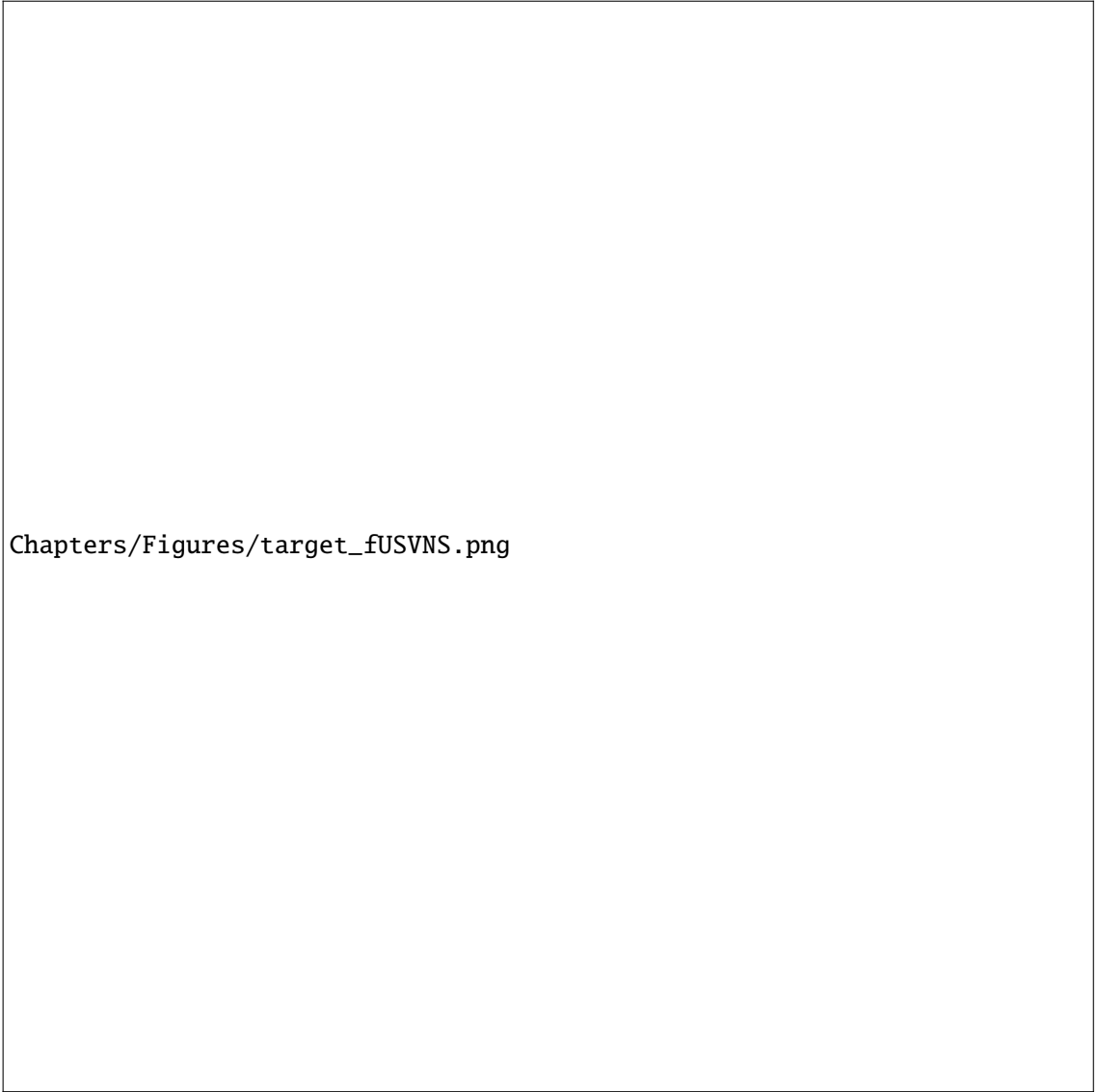
compared to its bulk transducer material counterpart, and together with the air-backing cavity provide for low quality factor, wide-bandwidth transducer suitable for diagnostic applications. The electromechanical coupling coefficient of PMUTs is significantly lower when compared to bulk PMN-PT transducers, making them less suitable for power-efficient, portable, miniaturised US therapeutic applications [<https://www.nature.com/articles/s41378-024-00783-5>, <https://doi.org/10.1038/s41378-023-00555-7>].

The basic operational principle of CMUTs relies on the change of capacitance while the transducer mechanically interacts with US waves. CMUTs are implemented in a silicon substrate, using a flexural membrane consisting of a metal electrode and a flexural micromachined dielectric substrate supported by an insulator [<https://doi.org/10.1109/TUFFC.2021.3112917>]. This membrane is suspended over a cavity acting as an air-backing matching layer providing, similarly to PMUTs, significant transducer bandwidth and axial resolution improvements - making CMUTs especially suitable for US diagnostic applications [<https://doi.org/10.1016/j.jmu.2012.10.1109/ACCESS.2024.3359906>]. The substrate is then integrated on top of the bottom electrode pad. The US pressure wave-induced mechanical displacement of the insulated top electrode provides for capacitance changes relative to bottom electrode. These capacitance changes are then transduced into electrical signals with an energy directly proportional to the energy of the sensed mechanical waves. Applying an AC electrical signal between the top and bottom electrodes charges/discharges the equivalent capacitor of the CMUT transducer, leading also to the mechanical displacement of the suspended membrane relative to the substrate, enabling also the transmission of US waves. The silicon substrate integration of CMUTs enables the utilization of fabrication methods are easily integrated into a CMOS manufacturing process, making the low transducer array manufacturing and integration complexity the main advantage of CMUTs [[10.1109/TUFFC.2021.3112917](https://doi.org/10.1109/TUFFC.2021.3112917), [10.1109/ACCESS.2024.3359906](https://doi.org/10.1109/ACCESS.2024.3359906), <https://www.nature.com/articles/s41378-024-00783-5>]. However, while the low quality factor of CMUTs provides suitability towards US medical imaging devices, the low electromechanical coupling coefficient renders CMUTs not suitable for US wave transmission due to the high driving power required by CMUTs to achieve moderate-to-high intensity focal spot pressures required by therapeutic applications [[10.1109/TUFFC.2021.3112917](https://doi.org/10.1109/TUFFC.2021.3112917), [10.1109/ULTSYM.2012.0020](https://doi.org/10.1109/ULTSYM.2012.0020)].

### 3.1.2 Ultrasound Transducer Array Integration Methods

Achieving a miniaturized ultrasound 3D imaging system requires the compact integration of the two fundamental components of the system - an ultrasound transducer matrix and the interfacing integrated electronics. Three main methods can be used to achieve this integration (Fig. ??).

The highest integration possible is achieved through the use of *monolithic integration* (Fig. ??). Regarding the integration of bulk piezoelectric transducers, an initial commercially available slab of transducer material is diced to the total width and length composing the



Chapters/Figures/target\_fUSVNS.png

total aperture of the transducer array. A first sub-dicing procedure cutting from 20% to 60% of the total thickness of the diced slab is then performed in both longitudinal and transversal directions to delimit each column and row, respectively, of the transducer matrix. The sub-diced slab is then flipped and aligned with the respective pads reaching from the top-most layer of the CMOS technology metal stack. The bonding is most commonly performed through the use of a chemical bonding process, where a conductive adhesive epoxy is cured at a temperature low enough to not surpass the Curie point of the transducer while still allowing for a strong bond. An alternative to the use of conductive epoxy is the use of an anisotropic conductive film (ACF) applied directly to the CMOS chip. After curing, a final sub-dicing through the rows and columns is applied to finally separate each transducer pillar, finally filling the kerf sections with a non-conductive epoxy to increase the durability of the micron-scale apparatus. To prevent damage of the CMOS chip in this final dicing procedure, a sacrificial epoxy layer is added, finalizing with the

introduction of a conductive sheet establishing the virtual ground (common potential) in the top plate of every transducer pillar (Fig. ??) [TiagoCosta2018]. Monolithic integration has several important advantages compared to other alternatives. The direct bonding between the transducer and the top-level metal layer of the CMOS technology stack allows for the optimal minimization of the electric parasitic passive components that are known to alter the impulse response of the transducer [<empty citation>]. More specifically, from the point-of-view of a therapy-oriented device design, the minimization of the parasitic capacitance in the ASIC-to-transducer optimizes the power transfer between the HV driving electronics and the transducer matrix, consequently enabling for maximum power delivered at the beamformed focal spot [TiagoCosta2018, HassanTiago2022]. Conversely, without proper control and empirical optimization of the curing time for the bonding conductive adhesive layer, and an incorrect choice of the cutting depth in each sub-dicing procedure, the yield of each successfully integrated transducer pillar with each respective pad can become exceedingly low with pillars being broken out off of their bond and being projected towards the remaining pillars, also destroying them, due to the high rotational speed of the dicing saw (Fig. ??). A yield of 40% for the totality of the transducer elements of each transducer matrix is settled as the acceptable minimum to consider the transducer matrix capable of operating as an ultrasonic transduction device, and the incorrect calibration of the underlying manufacturing parameters of this integration process can lead to this threshold not being met far too often [<empty citation>]. Moreover, the width of the dicing saw sets the minimum kerf size establishing the physical separation between each transducer element, consequently limiting the minimum pitch achievable with a monolithic integration method. This renders this integration method too restrictive regarding the development of high frequency ultrasound stimulation and imaging systems [<empty citation>].

The second integration method addressed in this discussion is *flip-chip bonding* (Fig. ??). This technique is most often used in the industry to enable direct integration between CMUTs and the respective CMOS technology [<empty citation>]. Two main types of flip-chip can be discriminated in the literature. An intermediate interposer printed circuit board (PCB) using vias that traverse through the PCB's layers establishing a connection between the transducer elements on the opposite side of the board and the dedicated pads in the top-level metal layer of the CMOS technology stack used for the design of the ASIC (Fig. ??) [HassanThesisRef182, 194]. The second approach replaces the intermediate interposer PCB with a silicon substrate, using a through-(silicon) substrate via for direct bonding with the transducer elements, in which CMUTs and PMUTs can be directly manufactured in (Fig. ??) [HassanThesisRef182, 193]. Although the use of an intermediate PCB to establish the connection between the transducer elements and each respective pad of the ASIC is a relatively inexpensive method for lower frequency applications, higher frequency applications, where the transducers' pitch is significantly smaller and the size of the required micro-vias for establishing the bonds, leads to a substantial cost increase of the PCB manufacturing process. Moreover, a commonly used fiber-glass

based material in PCB manufacturing is fiber-glass flame-retardant epoxy resin (FR-4), presenting a typical dielectric constant of 3.4 [<https://www.atlasfibre.com/material/fr-4/>, <https://link.springer.com/article/10.1007/BF02657420>]. The parasitic elements associated with the use of this method lead to a lower power distribution efficiency between the ASIC and each transducer element. On the other hand, a pure silicon (Si) wafer presents a typical dielectric constant of 11.7, but can vary depending on the oxidation level of its surface [<https://iopscience.iop.org/article/10.1088/2053-1591/abf684>]. The higher dielectric constant means the use of a through-silicon metal vias can lead to higher integration scale while lowering the effect of parasitics and increasing the power efficiency of a FUS stimulation application. The use of through-wafer vias is a technique currently facilitating the development of high-yield automated manufacturing and integration of CMUT transducer arrays directly with CMOS, which consequently ties this method to a transducer array with lower transmission efficiency compared to bulk piezoelectric transducers, as discussed in the previous section [[empty citation](#)]. Both monolithic and flip-chip integration have been proven as crucial methods to drive the miniaturization of US phased-array-based applications, either for diagnostic or therapeutic applications. The direct accessibility to the bottom plates of each individual transducer element, combined with the use of top-level metal pads within the ASIC's internal active area, is essential to allow cost-effectiveness and efficiency of the miniaturisation of the device.

Finally, the last method discussed in this section is a *PCB-based* transducer array integration, connecting the ASIC's I/O pads with the transducer array elements using standard PCB design methods and its metal traces to establish a connection to the bottom plate and wire-bonding for establishing connection to the top plate of the transducer elements (Fig. ??). The use of a PCB integration method allows for the cheapest option for the interface between the transducer array and the ASIC in a miniaturized US imaging/stimulation solution. However, it comes with several limitations that has led researchers to pursue both of the aforementioned methods. The first limitation, in similarity with a through-substrate via flip-chip bonding solution, is that in higher frequency applications, due to the reduced pitch of US transducer elements, the cost of PCB manufacturing can significantly increase to the customized reduced widths for the interfacing metal tracks that require laser erosion techniques [[empty citation](#)]. The second limitation is once again tied to the significantly increased parasitics of the PCB board interconnects when compared to the previous methods. In high frequency FUS stimulation or US imaging applications, the incorrect control of the phase-delay the PCB interconnect introduces in the signals due to the transmission line effects of each trace can greatly degrade the quality of the transmitted / received US signal, degrading the efficacy of the application as well. PCB integration is especially useful when interfacing ASICs with 1D transducer arrays. However, throughout the reviewed literature, the main limitation preventing the use of a PCB integration methodology in a matured, highly miniaturised device prototype is the high complexity of interfacing each transducer in a 2D transducer matrix with the I/O pads in the periphery of an ASIC (Fig. ??). The pitch requirements for each micro-via carrying the



US signals for the different layers of the PCB layer stack, and the space limitation capping the manufacturing costs of the ASIC, limiting the available I/O pads, quickly drive the costs of the prototype device using this method up while providing for a lower application efficacy when compared to the previous methods.

What if, when interfacing with a 2D transducer element matrix, to enable a FUS stimulation application through the use of a phased-array, while also allowing for the 3D US imaging of the sonicated medium, only a 1D set of metal interconnects was necessary between the ASIC and each column of the transducer matrix? Would a PCB integration methodology be suitable in that case, when aiming to drive the device prototyping costs down, while avoiding the loss of efficacy and functionality? This was the question asked and successfully answered at the School of Biomedical Engineering, Dalhousie University, Canada, in 2016 [<https://ieeexplore.ieee.org/document/7728854>, <https://ieeexplore.ieee.org/document/8926257>, <https://ieeexplore.ieee.org/document/9957541>]. The use of PMN-PT for the implementation of a crossed row-column transducer enables a DC-bias sensitivity towards the impulse response of each transducer array. By using either a positive or negative HV biasing in the top-plate of each row, a phase shift of  $90^\circ$  or  $-90^\circ$ , respectively, is introduced in the transmitted US waves (Fig. ??). This Hadamard-encoded electronically programmed phase shift is used to create a Fresnel lens discretized to two binary levels in anti-phase, creating zones of constructive and destructive interference in the sonication medium, ultimately leading to focusing of the transmitted waves in the elevational direction (Fig. ??).

The shorted bottom plates of each column of the crossed-electrode array are then used to recover the 3D spatio-temporal information of each 2D slice recovered. Each Hadamard code used to provide set the biasing and the electronically steered Fresnel lens through each row of the crossed-electrode array is ultimately associated with each elevational slice of the sonicated medium, allowing for 3D US images to be recovered while only interfacing with the columns of the array (1D set of metal interconnects). By combining the row-wise bias-encoded Fresnel lens for elevational focusing with a 1D phased array using programmed phased delays in the column-wise US TX pulses for azimuthal focusing, this method effectively enables the use of this crossed electrode configuration for US therapeutic applications (Fig. ??) [<https://ieeexplore.ieee.org/document/7728854>].

Regarding the suitability of the crossed-electrode configuration for US diagnostic applications, both main US imaging algorithms, SL and PWC, can be implemented in this crossed-electrode configuration, with the main difference between both implementations being the overall complexity of the PCB integration with the ASIC to allow for reliable US images to be reconstructed. When performing SL imaging, while the shorted bottom electrodes are used to establish the crossed-electrode's column wise 1D phased-array for azimuthal focusing, the shorted row-wise top electrodes are used to establish the bias-Hadamard-encoded Fresnel lens enabling the elevational focusing, creating a FUS scan-line that can be swept across the 3D medium being sonicated. During RX events, the two-way focusing required by SL US imaging can then be achieved by using the reverse beamformed

lens used in the column-wise 1D phased array during TX events. However, to mitigate the effect of secondary lobes in the decrease of the reconstructed image's spatial resolution, the RX channels of the US imaging readout system must be connected in the top shorted row-wise crossed electrodes used to establish the bias, leading to a significant increase in PCB integration complexity and a reduction of the miniaturisation form-factor of the device prototype in general. Additionally, the use of wire-bonding techniques to connect the RX channels to the row-wise shorted top electrodes can significantly decrease the signal integrity of the transduced echoes during RX in high-frequency US diagnostic applications, due to the additional parasitic inductance and resistive voltage dividers inherent to the wire-bond. Finally, the use of a HV biasing network in the same interconnect as the low-voltage (LV) RX channel introduces the need to well matched DC biasing networks that can either invalidate the use of a chiplet solution for the US imaging and stimulation-capable device, or introduce the risk of permanent damage to the RX channel integrated electronics leading to the loss of imaging functionality (Fig. ??). Alternatively, to mitigate ASIC damaging risks and drive the integration complexity downwards, a PWC imaging method can be used with this crossed-electrode configuration. When performing PWC US imaging, both TX and RX events can be performed using the shorted column-wise bottom electrodes of the array, with the row-wise electrodes being used solemnly for biasing of the bias-sensitive transducer elements, reducing PCB integration complexity without significantly (negatively) affecting the functionality and the reliability of the reconstructed US images. Due to the lack of TX focusing upon the transmission of steered planar/divergent waves, the reception of the echoes can be performed in the same set of 1D electrodes (and focusing direction) as the steered plane-wave TX pulses are transmitted (Fig. ??). Secondary lobes can be further mitigated in the digital backend of the data acquisition (DAQ) system by reconstructing the US image using Hadamard-encoded elevational information of each transmitted steered plane wave [<https://ieeexplore.ieee.org/document/8926257>, <https://ieeexplore.ieee.org/document/9957541>].

The previously described technique, designated as *Simultaneous Azimuth and Elevational Compounding* (SAEC), combines planar wave imaging techniques with a PCB integration method dedicated to the integration of electrostrictive transducer elements (PMN-PT) with the interface electronics of the US imaging system [<https://ieeexplore.ieee.org/document/9957541>]. The use of SAEC reduces the number of interconnections required to interface with a  $N \times M$  transducer matrix to  $N$  bottom electrode interconnects, facilitating the integration of US diagnostic and therapeutic systems with high-aperture 2D transducer arrays containing  $> 1000$  transducer elements. This does not necessarily impact the power required to drive all the transducer elements of the same column during TX events, once the parasitic capacitances encoding the mechanical losses of each transducer in resonance are added in parallel (Fig. ??). However, during RX events, upon the readout of the received US echoes  $\times M$  lesser receiver channels are required to enable a reliable reconstruction of the target US image, driving the dissipated power of the system's receiver down  $\times M$ , respectively.

Moreover, the use of SAEC places the viability of a PCB integration method on par

with monolithic and flip-chip bonding interfaces, with the main advantage of having the possibility for drastically reducing ASIC-array integration costs without the loss of device efficiency and functionality, both in the US diagnostic and therapeutic domains. Finally, while CMOS technologies supporting the integration of HV Bipolar-CMOS-DMOS (BCD) devices for driving the transducer elements are essential for miniaturized systems on chip (SoC) capable of transmitting US waves, the readout of the received echoes can be performed with standard CMOS technologies at a more advanced node that do not necessarily support such BCD devices. Consequently, the use of a PCB integration enables the use of a chiplet solution that can further reduce the total power dissipation of the US image-guided neuromodulation system's receiver, without compromising on the readout signals' integrity.

### 3.2 Transmitting Integrated Circuit Ultrasound Interfaces

Conventional US systems generate the required HV signals outside of the probe [**<empty citation>**]. The HV signals would then be applied to each transducer directly, using an interconnection per transducer [**<empty citation>**]. Driving the miniaturization of US medical devices, while guaranteeing suitable functionality and precise control for increasing the potential of US as a multi-purpose therapeutic and diagnostic technology for true POC solutions, requires the TX functionality to be integrated closer to the US transducer array. HV switching and multiplexing techniques significantly reduce the amount of interconnections required to drive the TX system. Despite its simplicity and compact integration, multiplexing ultimately limits the frame-rate of an US imaging system, while also diminishing the peak pressure that is possible to be achieved in a FUS system. The time-division-multiplexing (TDM) access of each transducer to its HV driving signal has to be performed at low-to-moderate frequencies (0.5 MHz - 15 MHz) due to the speed limitations of HV BCD technology devices. This inevitably introduces phase errors in the distributed waves, reducing the peak pressure observed in the regions where maximum constructive interference between TX US waves should be observed [**<empty citation>**]. Effectively, as it will be discussed in the following subsections, the element-level integration of a TX HV driver is essential for US image-guided neuromodulation applications, allowing for precise FUS beam-steering while providing for higher imaging resolution US diagnostic systems [**beam\_steering\_tiago\_hassan**][**milian\_tan\_pertjis**].

Although it is not the focus of this work, the following subsections discuss recent works achieving compact, wearable and implantable devices capable driving integrated US transducers for both diagnostic and stimulation applications. The discussion will mainly focus on observing the evolution towards higher power efficiency for the transmitting circuits while achieving a lower supply voltage without compromising on the achieved peak pressure sent to the medium. Understanding the limitations and advantages of the transmitting circuits described will enable for a more productive discussion of the power and area specifications that must be imposed on a receiving analog front-end integrated

within the same die and/or PCB as the driving circuits.

### 3.2.1 Supporting Semiconductor Technologies

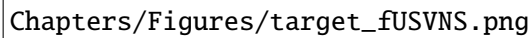
In Sec. ?? it was discussed that in order to achieve sufficient pressure amplitude in the generated US waves in both therapeutic and diagnostic applications, the transducer device requires voltages in the range of 30 V-160 V, depending on the type of transducer material used. Although bulk piezoelectric transducers require the minimum voltage potential observed in the state-of-the-art while still generating enough pressure to achieve intensities close to the required parameters in LIFUS stimulation procedures [TiagoCosta2018], the required voltage potentials are still far too high for standard CMOS technologies. Previous designs have proposed to generate HV pulses using stacked LV transistors (Fig. ??) [https://ieeexplore.ieee.org/document/6823174][https://ieeexplore.ieee.org/document/9340356]. However, the increased circuit complexity, potential reliability concerns and the lower switching speed-power efficiency tradeoff this solution offers has led research groups to seek more reliable alternatives. The amount of charged carriers being distributed to the substrate of standard threshold voltage (standard-VT) (LV) devices can easily trigger a phenomenon known as *latch-up* (Fig. ??). This phenomenon happens easily if the electric fields have an high enough magnitude that the finite resistance between the doped source contacts and the substrate leads to a positive feedback that eventually *latches* both parasitic bipolar junction transistors (BJT), completely turning them on, and leading to a massive amount of current being drawn from the integrated circuit's supply into the substrate (Fig. ??) [RahzaviDesignofAnalogCMOSCircuits2ndEd].

At the expense of more silicon area to effectively spread out large electrical fields (preventing them from reaching critical values) and consequently increased distances between device terminals, diffusion regions and additional spacing for isolation barriers between the devices, the introduction of a bipolar-CMOS-DMOS<sup>1</sup> (BCD) HV integrated circuit technology (Fig. ??) solves this issue by enabling the co-existence between standard, (low) core-voltage CMOS and HV devices in the same integrated circuit die. Several notable works driving the miniaturization of US HV drivers that are required in wearable US diagnostic and therapy-oriented applications have exploited this technology [works\_on\_stimulation\_US\_devices].

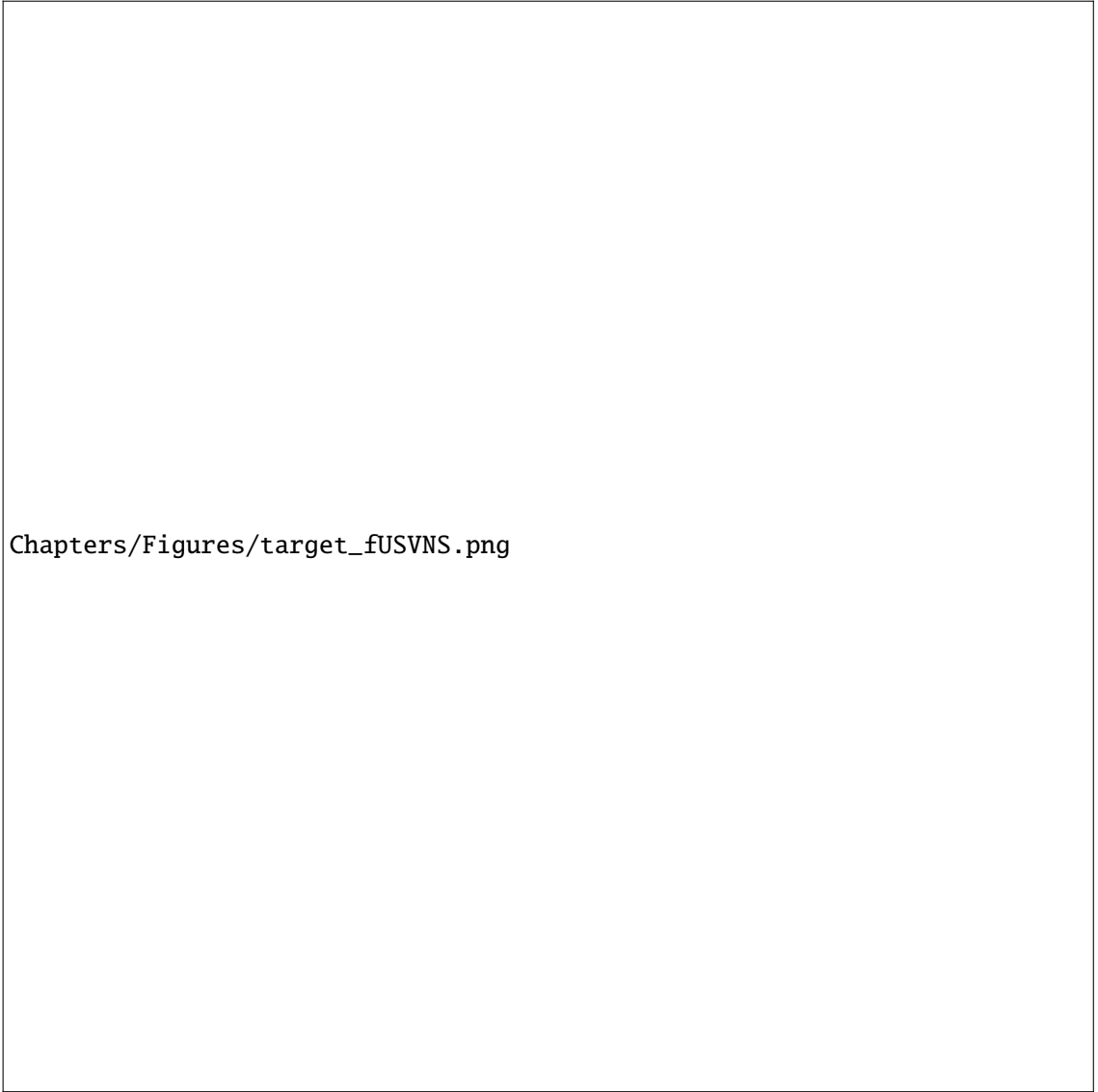
In DMOS devices, the source and bulk are connected and aligned with the gate, forming a channel with a length determined by the diffusion between the p-body and the n+ source, rather than the gate's dimensions. The width of the device is typically scaled, not the length. The lightly doped n-well region between the drain and channel acts as a drift region, determining the device's on-resistance ( $R_{on}$ ) and allowing it to handle higher voltages at the drain. DMOS devices are asymmetrical and have limited overdrive, constrained by the gate oxide breakdown voltage (typically around 5–5.5V in TSMC's 180nm BCD technology, being the most commonly observed technology for miniaturized

---

<sup>1</sup>DMOS: double-diffused metal-oxide-semiconductor technology

Chapters/Figures/target\_fUSVNS.png

US devices). The gate extends over both the channel and drift region to reduce the electric field along the drift region, though this increases gate capacitance. The HV deep n-well isolates the device's active parts from the substrate. The lateral structure of the DMOS shows a parasitic BJT formed by the doping profile of the p-body and n-well. To prevent this parasitic transistor from activating, the source and bulk are connected. The n-guard is often connected to a positive voltage to enhance electron collection. Isolation around DMOS devices is achieved with guard rings, whose width and spacing depend on the peak voltage the device needs to handle. However, wider isolation barriers can increase parasitic capacitance and significantly raise the area cost for small transistors. The deep n-well creates a parasitic PNP transistor, which can become forward-biased if the source voltage exceeds the drain, causing hole injection into the substrate. These holes can lead to substrate potential changes, causing dissipation and issues in nearby devices. To mitigate this, a p-guard ring is used to stabilize the substrate potential, and n-guard rings are also



Chapters/Figures/target\_fUSVNS.png

employed to reduce electron injection and latch-up risks.

### 3.2.2 Unipolar Pulsers

Unipolar pulsers are often referred to as *square-wave pulsers*, and their counter part are the class of *linear-amplifiers*. These are the two main classes allowing for compact integration between TX driver and US transducer. Linear amplifiers can drive the US transducer with a continuous amplitude signal. However, it is enough for most medical US applications to drive the US transducer with discrete-amplitude pulses. Consequently, the higher energy efficiency, reduced die area footprint, and increased bandwidth of pulsers will be the focus of the architectures discussed in this and the following subsection.

Unipolar pulsers implement the functionality of alternate settling of the top plate of the US transducer element towards HV level and ground. Thus, this circuit-level architecture

provides the lowest complexity possible for integrated HV US transducer drivers, using a low number of HV devices. The two main topologies observed in the SoTA inverter-based and source-follower (SF)-based HV driver circuit architectures. Both architectures exploit the aforementioned DMOS devices. A common-source (CS) inverter-based, single N-type HV device, is presented in [**<empty citation>**] to drive the US transducer between HV and ground. The use of a pull-up resistor leads to a significant static power dissipation when the driven top plate of the transducer is pulled to ground. This issue can be mitigated through the use of a complementary P-type HV device to eliminate static power dissipation of the driver. However, driving the gate of the P-type device requires a level-shifting of the gate-driving signal from 0 V to the  $V_{gs}$  required to close the N-type HV device to a pulsing signal leveled between the HV supply and  $(HV - V_{sg})$  to drive the P-type HV device without introducing reliability issues to the circuit. This has been achieved in [**<empty citation>**] through the use of a crossed-coupled LV PMOS close to the HV supply to pull the source of the P-type featuring a DC gate-biasing using a supply greater than the supply gate-biasing the N-type HV-devices (Fig. ??). This solution introduces possible reliability issues due to the use of LV devices close to the HV supply, although the reference is also shifted towards the mid-level supply biasing the gates of the P-type HV devices. Moreover, the use of three different HV supply voltages leads to an increased complexity in the layout of the architecture, while observing an increased die area footprint due to the required number of dedicated pads for the supply of the driver. Alternatively, to remove the reliability concerns regarding the use of LV devices in the HV driver, several works have adopted an inverter-based HV driver (half H-bridge topology) architecture using complementary HV devices. A single driving signal to drive both inverter devices, while using a dedicated pull-up resistor-based level shifter with protection (voltage-clamping) Zener diodes to shift the same driving signal to a suitable range enabling activation of the P-type device (Fig. ??) [**inv\_based\_1, inv\_based\_2, inv\_based\_3, ...**]. The use of a floating ground at the voltage of  $HVDD - V_{SG}$  enables the introduction of additional digital driving buffers (cascaded inverters) to increase the settling speed of the HV drivers[**<empty citation>**]. Latched HV devices have also been observed to increase the switching speed due to the positive feedback the topology introduces in the dynamic of the circuit (Fig. ??)[**HassanLatchedHVDriver**]. However, careful control of the settling time of each HV gate-driver must be performed to guarantee the HV latches can actually be latched between  $HVDD$  and ground. Bulk biasing feedback can be used to both perform settling time delay control and increase the energy efficiency of the driver [**<empty citation>**].

Despite the low complexity and power efficiency of the aforementioned solutions, the use of a single-ended unipolar square-wave HV architecture driving the US transducer's top plate between  $HVDD$  and ground leads to a great portion of the power dissipation to be related to even-harmonics due to the presence of a DC component in the driving signal (Fig. ??) [**<empty citation>**]. This spectral power leakage associated to distortion components of the square wave diminish the SNR of a US imaging being

reconstructed using such a TX driving pulse (Fig. ??) [empty citation]. It also renders this driving method not suitable for other imaging techniques such as second-order tissue harmonic imaging [empty citation], pulse inversion [empty citation] and chirp encoding [empty citation]. At the expense of  $4\times$  increase in power dissipation, the use of a bipolar, symmetric square-wave driving the top-plate of the US transducer between HVDD and HVSS, with  $HVSS = -HVDD$ , removes the signal's DC component (Fig. ??), therefore mitigating this disadvantage. The two most prominent bipolar pulser architectures found within the literature are the HV source-follower (SF) and the inverter-based driver. While an inverter-based topology offers superior energy efficiency at moderate-to-higher central frequencies [InverterBasedBipolarDriver-<https://ieeexplore.ieee.org/document/9365808>], the SF enables greater settling speeds at the expense of a greater architectural complexity, being suitable for high-frequency imaging and FUS applications [SourceFollowerBipolarDriver-<https://ieeexplore.ieee.org/document/9365808>].

### 3.2.3 Multi-Level Pulsers

Multi-level pulsers have become critical for next generation US HV pulsers. Apodization modulates the system response of the transducer element, enabling the concentration of the transmitted mechanical energy in the main lobe of the transmitted FUS wave [SzaboElementApodization]. As such, for SL imaging applications, the use of a multi-level pulsers allows for transducer element-level temporal and spatial apodization (Fig. ??) that leads to significant spatial resolution improvement by reducing the side-lobe magnitude (Fig. ??). The use of a multi-level pulsed wave to approximate a sinusoidal, continuous transmitting wave enables for a significant increase in the linearity of the received US echo due to the concentration of the energy in the main lobe, leading to an increased SNR for the reconstructed image [SzaboSpatialResolutionImage]. Regarding stimulation applications, this spatial and temporal apodization leading to the concentration of the mechanical energy in the main lobe of transmitted FUS beam leads to an increase efficacy for the therapy application while improving spatial resolution of the highest pressure region of the generated focal spot [empty citation]. Furthermore, although the generation of a multi-level pulsed wave inevitably requires an increase in hardware complexity, it has also been observed that significantly improves the energy efficiency of the HV drivers in two different ways. On one hand, upon the discharging of the transducer's parasitic capacitor (associated with its mechanical losses) between consecutive supply levels, the removed charge can be recycled back into the respective HV supply, leading to power dissipation reduction up to 40% [ChaoChenReviewPaperRef38]. On the other hand, it has been observed that the contribution of each individual element in a phased-array exciting each element of an US transducer matrix is different and dependant on the relative position of each element to the focal spot generated through the transmission of a FUS wave [MasoumehIUSPublication, AssociatiedRefs]. Thus, scaling the amplitude of the excitation pulse of each individual element to its contribution to the focal-spot's peak pressure leads to a significant reduction in the total power dissipation



of the ASIC, while leading to minimal losses in pressure at the intended focal-spot's volumetric region (Fig. ??).

$$Q_{UP} = C_p (HVDD - 0 V) = C_p HVDD [C] \quad (3.1)$$

$$Q_{MP} = C_p \left( N \cdot \frac{HVDD}{N-1} - (N-1) \cdot \frac{HVDD}{N-1} \right) = C_p \frac{HVDD}{N-1} [C] \quad (3.2)$$

Where  $C_p$  is the parasitic capacitance of the US transducer element.

The concept behind charge recycling in  $N$  – level pulsers is demonstrated in Fig. ?.  $N - 1$  HV switches are used to charge and discharge the US transducer element between consecutive HV levels by connecting the top-plate of the transducer each respective supply. Compared to a unipolar HV driver, where the transducer's parasitic capacitance is completely discharged to ground in each cycle (?), in a step-wise downward transition, a  $1/(N - 1)$  fraction of the total charge being discharged from the transducer is recycled back to the lower regulated supply (?) [ChaoChenReviewPaperRef38]. Consequently, the dynamic power dissipation (?) of multi-level pulsers with  $N$  levels is reduced by a factor of  $(N - 1)$  (?).

$$P_D = f_0 C_p V^2 [W] \quad (3.3)$$

$$f_0 C_p HVDD^2 \longrightarrow f_0 C_p HVDD \cdot \frac{HVDD}{N-1} \quad (3.4)$$

Where  $f_0$  is the cycle frequency driving each step of the pulse.

$N$ -level pulsers require multiple HV supply voltages distributed off-chip, reducing the integration level of a miniaturised US stimulation ASIC due the increase number of supply pads and die area. Moreover, the additional power dissipation upon driving the HV switches to connect individual transducer elements to each available HV supply counterbalances the energy efficiency increase through the use of this driving architecture. To address these issues, multiple schemes tackling the increased hardware complexity of this method have been proposed [ref1, ref2, ref3, ...]. Instead of switching both transducer electrodes to ground, the use of an HV switch in parallel with a PMUT transducer has been used to short-circuit the transducer's plates creating a virtual ground in the driving pulse signal, while the bipolar pulser drives the top-plate of the transducer element between symmetric HVDD supplies [<empty citation>]. This method has the main disadvantage of being suitable for transducer materials exploiting the 31 oscillation mode for generating the US mechanical waves, namely CMUTs and PMUT, once these materials enable a low complexity integration and access between both of the transducer's plate electrodes and the ASIC (Fig. ??, ??). In bulk piezoelectric transducers, the bottom electrode of the transducer element is vertically opposed to the ASIC, and must rely on a high-complexity, dedicated wire-bond interconnection to allow for its

access (Fig. ??). Other works targeting the increased hardware complexity of a multi-level pulser capable of more than 3 pulsing levels have used bulky, and often off-chip, stacked-capacitors to generate a multi-level supply [[empty citation](#)]. HV level-crossing detection using a scaled-down supply of the HV supply and a zero-crossing detector (ZCD) together with a programmable (resistive) ladder to control the charging and discharging of the transducer's top-plate between 7 voltage levels have provided a very compact, pitch-matched solution to a multi-level pulser architecture [[empty citation](#)]. Finally, the introduction of an inductor per element has also been explored to allow for a 75% dynamic power dissipation reduction achieved through charge recycling [<https://ieeexplore.ieee.org/abstract/document/9365826>, <https://ieeexplore.ieee.org/abstract/document/10904513>]. At the expense of a significant increase in the form factor of the system, the individual inductors create an LC tank in parallel with the transducer elements being driven at the resonance frequency by the resonant pulser. The energy stored in the magnetic field of the inductor upon the discharging of the transducer can be used to charge it again in the next cycle (Fig. ??). The use of an individual inductor per element leads to a prohibitively high form-factor for miniaturised systems capable of driving US transducers under a HV regime. To tackle this issue, a single-inductor system has been recently introduced to reuse the energy stored in the passive inductor element with groups of transducer elements in the array at different phases, multiplexing the inductor's stored energy redistribution and achieving 88% dynamic power dissipation reduction [<https://ieeexplore.ieee.org/document/10904513>] (Fig. ??).

### 3.3 Receiving Integrated Circuit Ultrasound Interfaces

Due to propagation-media absorption, scattering and reflection phenomena, the received US echo signal is subjected to a linear-in-dB decay associated to a *propagation dynamic range* ( $DR_P$ ) (Fig. ??). Moreover, the band-pass characteristic of the transducer element's system response can be modulated through a Gaussian-pulse-modulated cosine in time, leading to an intrinsic *instantaneous dynamic range* ( $DR_I$ ) for the transduced US echoes (Fig. ??). The first interfacing circuits to the received US echoes are responsible for 1) linear amplification of the transduced US echo, 2) noise and inter-channel electrical correlation suppression to increase the signal's integrity and suppress noise contributions of the following circuits and systems, and 3) time-gain compensation (TGC) of the received US echo. The TGC functionality, usually implemented in the analog domain, enables the compression of the high *total dynamic range* ( $DR_T = DR_I + DR_P$  [dB]) of the US echoes to a limited dynamic range relaxing the required specifications of the following circuits and systems in the signal processing chain (Fig. ??). Alternatively, a channel quantization interfacing exploiting noise-shaping for high-input  $DR$  enables direct quantization of the received US echoes without performing TGC in the analog domain [[DirectQuantizationSensorToQuantizerWorks](#)]. The following subsections provide a discussion on the existing architectures establishing the SOTA regarding power, area

and noise-efficiency performance employed in miniaturized, integrated diagnostic US systems. The discussion will be mainly guided from a point of view that prefers aggressive miniaturization of the whole analog signal processing chain, reducing it to its core fundamental functionalities.

### 3.3.1 Low-Noise Amplification

The low-noise-amplifier (LNA) interfaces directly with the transducer element, providing for a linear amplification of the received US echo, increasing its signal-to-noise-distortion ratio (SNDR) for an enhanced signal integrity upon being processed by the following circuits of the signal processing chain. A secondary variable-gain-amplifier (VGA) cascaded with the LNA is commonly employed to further compress the total dynamic range of the received US echoes through TGC.

Selecting the correct LNA topology depends, mainly, on the type of US transducer the LNA is interfacing with, the input impedance of the circuits the LNA is driving, and the topology enabling an optimal area, noise and power tradeoff that is specifically tailored for each specific diagnostic application. The competitive urge to develop LNA with the best noise figure-of-merit (FoM) in relation to the SOTA commonly leads IC designers to over-design their circuits, leading to excessive power dissipation and die-area being used to suppress input-referred noise (IRN) far beyond its required specification. Depending on the center frequency of the TX US waves to achieve the necessary spatial resolution, the total dynamic range of the received US echoes can reach 90 dB in vascular catheter US probes, due to the elevated attenuation the heart's muscle tissue introduces in the US signal. As such, an higher power dissipation enabling higher IRN suppression capabilities is required [ChaoChen2018]. On the other hand, the total dynamic range of the US echoes in trans-cutaneous vagus nerve imaging applications is usually within 60-70 dB, relaxing the IRN specification of the employed LNA architecture. For main LNA architectures can be distinguished throughout the SOTA: 1) voltage amplifiers (VA), usually employed in capacitive feedback configuration, 2) trans-impedance amplifiers (TIA), usually employed in resistive feedback for DC bias-point establishment, 3) current amplifiers (CA), usually employed in capacitive feedback configuration and 4) trans-conductance amplifiers (TCA), usually employed in open-loop (Fig. ??). The output impedance of each transducer material holds a complex relationship between TX driving frequency, material properties, device geometry, the presence of backing and matching layers, and top and bottom electrodes' acoustic impedance. For custom integrated US diagnostic systems, a dedicated characterization of the available US transducer material should be carried out in media applying similar acoustic loads to those verified in the intended diagnostic application [ChaoChenReviewPaper, ChaoChen2018]. Bulk piezoelectric transducers, such as PZT and PMN-PT, and other materials typically employed in PMUTs, feature a relatively low output impedance in resonance ( $<2\text{ k}\Omega$ , leading to the output voltage to be the electrical

signal of interest to readout. Maximizing the signal power transfer between the transducer and the interfacing signal processing chain requires the sensing of the transducer's voltage in this case. For this case, VA [empty citation] and TCA [empty citation] architectures provide a high input impedance through the direct interface between the transducer's signal and input-stage devices' gates, facilitating power and area efficient voltage-sensing. On the other hand, CMUTs and lead-free, polymer-based transducer materials such as PVDF typically feature an high output impedance ( $> 10 \text{ k}\Omega$ ), and their output current is the electrical signal of interest to readout. In this case, the low input impedance CA [empty citation] and TIA [empty citation] LNA topologies feature facilitate lower power dissipation and die are footprints when performing current-sensing.

Several circuit-level techniques have been employed throughout the SOTA to further optimize power dissipation and the total area of the LNA architecture. Exploiting the duty-cycled TX/RX phase transitions, the use of capacitive negative feedback has been used in VA and CA topologies to improve linearity and remove the noise contribution of the feedback network, at the expense of a greater die area and limited amplifier bandwidth [empty citation]. The power-noise efficiency tradeoff has been commonly optimized through the use of a current-reuse (self-biased) inverter-based topology, for the maximization of the LNA's bandwidth under the same DC current biasing point and lower gain error against process-voltage supply-temperature (PVT) variations [empty citation]. In general, open-loop LNA topologies have been exploited to minimize the power dissipation and die area FoMs in wide-bandwidth US diagnostic applications, at the expense of an elevated noise floor and decreased linearity [empty citation]. For ultra-wide bandwidth (UWB) US diagnostic applications, the incorporation of analog-domain coherent down-conversion-based envelope-detection (ED) has been exploited to greatly increase the power efficiency of the A-mode imaging ASIC [empty citation]. Conversely, despite the increased noise floor and decreased linearity, open-loop amplification topologies are very interesting when aiming to develop a US diagnostic ASIC in advanced CMOS process nodes for further miniaturization of the imaging ASIC. The reduced intrinsic voltage gain of devices in sub 40 nanometer nodes (Fig. ??) often lead to design choices preferring the use of open-loop inverter-based VA architectures to 1) avoid the increased feedback gain error due to the reduced open-loop gain, 2) achieve sufficiently high output voltage swing and input voltage headroom (mainly avoiding cascode device topologies) while 3) achieving a sufficiently high bandwidth in an energy efficient manner [empty citation]. However, the use of inverter-based amplifier topologies can increase the difficulty of introducing common-mode-feedback (CMFB) circuitry and supply-rail regulation for increased power supply rejection-ratio (PSRR) into the LNA without significantly increasing the power and area per channel of the AFE in general, while decreasing the distortion-free dynamic range [empty citation].

### 3.3.2 Power Supply Rejection

Correlated signals, such as digital interference coupled to the analog supply, can be amplified by a factor of  $N$  (array size) during averaging, summing or beamforming operations in parallel LNA channels interfacing with closely integrated transducer arrays, potentially causing artifacts in the reconstructed ultrasound image. The correlated noise can lead to an increase in the input-referred noise floor, [Chen2015, Tan2018, Hopf2022], and even when being below the individual LNA channels' it can still affect image quality (Fig. ??). Additionally, electrical crosstalk between neighboring LNA channels can elevate side lobes and cause unwanted echoes, requiring careful layout to shield sensitive signals from interference (Fig. ??). A PSRR figure that scales with array size is essential. The single-ended nature of the signals read from US transducers predominantly used in diagnostic US ASICs are most often read out using single-ended LNA architectures to accommodate the received US echoes while achieving a low power dissipation and area-per-channel FoM [empty citation]. The use of current-reuse inverter-amplifiers has been commonly observed in the SOTA. Although their main objective is the increase of power-dissipation efficiency upon the suppression of the IRN figure, it also enables the decoupling of the LNA devices' source nodes from the supply rails [empty citation]. The single-ended nature of the current-reuse inverter LNA topologies often employed leads to a relatively inefficient method for increasing the PSRR of the amplifier, once perturbations on the supply rails are directly represented in the single-ended output of the topology. Dedicated supply low-dropout regulators (LDO) to each LNA sub-array can be used to mitigate the aforementioned issues,. However, the use of a power-intensive LDO for each LNA sub-array prevents significant advances in the power and area-per-channel figures of merit inverter-based LNA AFEs enable (Fig. ??) [Chen2016, Chen2017, Chen2018, Guo2022]. The use of fully differential LNA topologies using common-mode feedback (CMFB) networks decoupling the source-nodes of the amplifier's devices from the supply rails has been identified as a promising method to further increase the PSRR of the amplifier [empty citation]. Once again, at the expense of a doubling of the area and power dissipation for each LNA, fully-differential topologies represent any type of signal variations present in the source and gate of the core amplifier's devices, and amplified to the corresponding drain node, in the symmetric output drain node counterpart, leading to a significantly increased PSRR of the differential output signal of the LNA[empty citation]. Finally, the ability of floating inverter amplifiers (FIA) to operate decoupled from the supply rail in their corresponding tracking phase has been identified as a promising way of increasing the PSRR of the associated output (differential) signal in recent years [empty citation].

### 3.3.3 Time-Gain Compensation (TGC)

The performance of TGC in the analog domain of the AFE significantly relaxes the IRN specifications of cascaded channel quantization circuits, offering significant power

dissipation reduction and miniaturization capabilities for the whole AFE in general [**<empty citation>**]. Two main methodologies to perform TGC in the analog domain can be discriminated throughout the SOTA. Variable-gain amplifiers (VGAs) apply a continuous, linear-in-dB, amplification to the input signal through time through the means of analog tuning of the DC operating point of the amplifier's core devices [**<empty citation>**], and continuous interpolation of discrete gain steps performed using programmable feedback networks [**<empty citation>**].

On the other hand, programmable-gain amplifiers (PGAs) perform TGC through the application of linear-in-dB discrete gain-steps to the input US echo signal [**<empty citation>**]. The use of a digital word to discretely tune the electrical parameters of passive devices can enable a PGA functionality through discrete (resistive) source-degeneration and negative capacitive feedback loop-gain programmability. PGAs offer

#### **3.3.4 Sub-Array Beamforming ( $\mu$ BF) and Channel Multiplexing**

#### **3.3.5 Channel Quantization**

### **3.4 Research Questions**

## LITERATURE REVIEW METHODOLOGY & WORK-PLAN

- 4.1 Literature Review Methodology**
- 4.2 Technology Readiness Level (TLR) - Assessment and Objectives**
- 4.3 Work-Plan**

A

## APPENDIX 1



| B

## APPENDIX 2

Something something.

I

ANNEX 1

