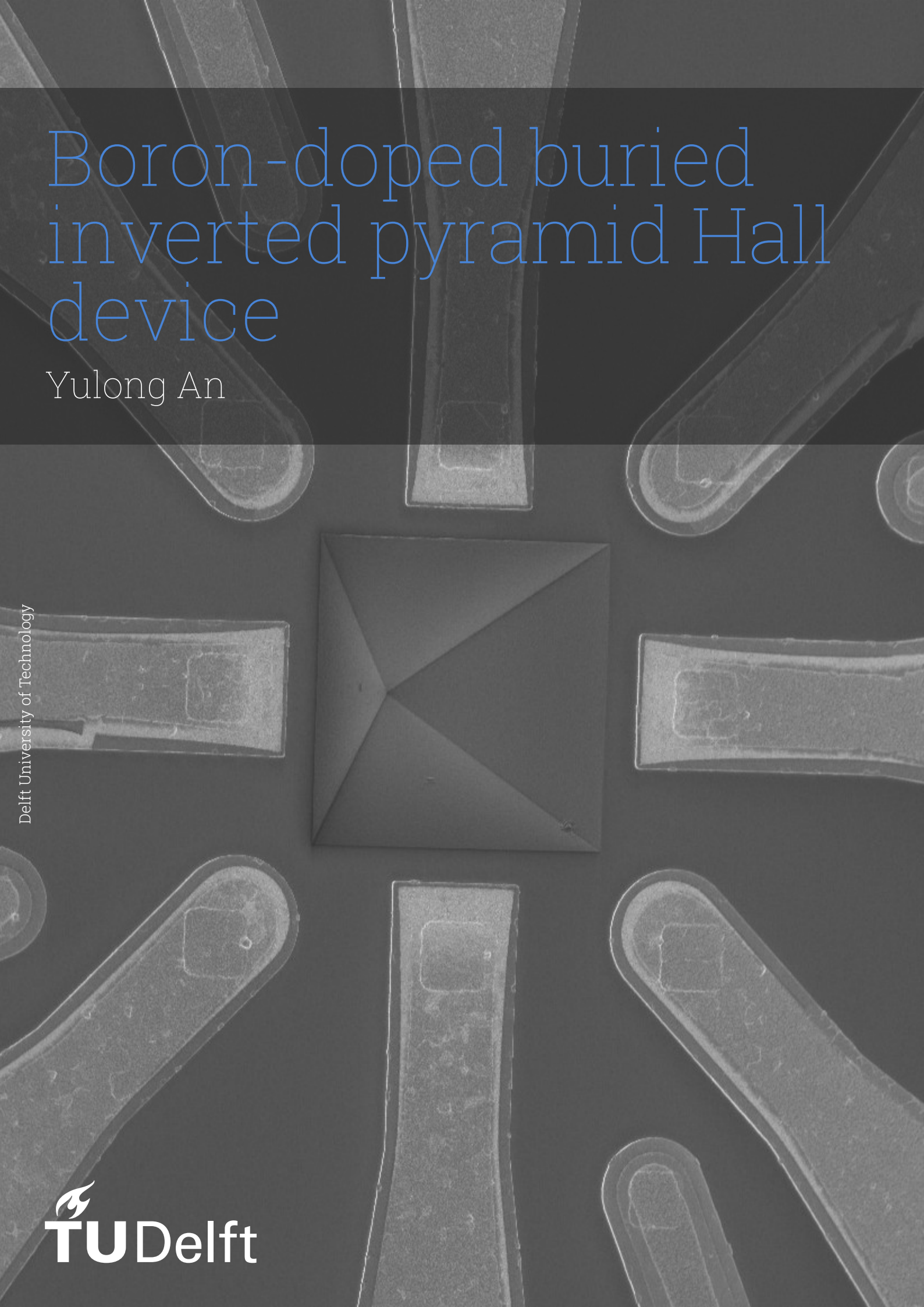
Boron-doped buried inverted pyramid Hall device

Yulong An



Delft University of Technology

TU Delft

Boron-doped buried inverted pyramid Hall device

by

Yulong An

student number: 6020933

Supervisor:	Karen Dowling
Daily supervisor:	Jacopo Ruggeri
Project Duration:	9, 2024 - 12, 2025
Faculty:	Faculty of EEMCS, Delft

TU Delft

Acknowledgements

First of all, I would like to express my sincere gratitude and respect to my supervisor, Karen. This project has provided me with countless learning opportunities, and it is thanks to her guidance that it was able to take shape. She has created an environment that was open, easygoing, and enjoyable, while always showing trust and respect. For all of this, I am truly grateful.

Secondly, I would also like to express my deep gratitude to my daily supervisor, Jacopo. His knowledge in his field impressed me from the very beginning of this project, and I am truly thankful for his continuous support. I still remember one time when the furnace I had reserved was suddenly taken by someone else—he immediately stepped in and helped me abort the running recipe. It was the only time I had ever seen someone abort a tool in the EKL, so it made him seem pretty cool to me.

Moreover, I want to thank my family, whose love and support have shaped who I am today. I am also grateful to my friends—here's a big toast to all the Friday dinners we enjoyed over the past two years.

I would also like to thank my badminton partners. The great games we played always kept me happy and energized.

Finally, I would like to cite here a few lines from one of my favorites, "My Way," the classic by Frank Sinatra.

For what is a man, what has he got?

If not himself, then he has naught.

To say the things he truly feels.

And not the words of one who kneels.

The record shows I took the blows.

And did it my way.

Although these lyrics are nothing like my life experience, it's cool to use them to conclude this part, isn't it?

Yulong An
Delft, December 2025

Summary

This thesis presents a comprehensive study on the fabrication and characterization of a buried inverted-pyramid Hall sensor realized using pure-boron LPCVD. Conventional inverted-pyramid Hall sensors typically suffer from high residual offset and elevated flicker noise, effects that are often attributed to charge trapping at the passivation surface. Introducing a buried sensing layer is an effective strategy to mitigate these surface-related disturbances.

Motivated by this concept, a p+ layer was integrated into the pyramid Hall sensor architecture. Pure-boron LPCVD was employed to form the required ultra-shallow junction, taking advantage of the absence of channeling and transient-enhanced diffusion effects commonly observed in ion-implanted boron. To minimize interference from the junction field effect, the boron deposition process was modeled and simulated using Synopsys Sentaurus. Based on these results, a one-minute LPCVD boron process was implemented to form the p+–n junction above the active region.

The fabricated sensors exhibit an equivalent magnetic residual offset ranging from 7.5 µT to 710 µT, with the buried sensor showing a lower residual offset in most modes. The sensor demonstrates a thermal noise floor of approximately 1 µT/√Hz at a 1.26 V supply. Flicker-noise analysis using the Hooge model reveals that the buried inverted-pyramid Hall sensors exhibit a lower α_H/N and a reduced corner frequency compared to conventional devices, indicating a possible improvement in low-frequency noise performance.

Overall, the buried-layer architecture offers an alternative pathway for reducing both residual offset and flicker noise in inverted-pyramid Hall sensors, providing a potential solution for high-precision magnetic field sensing in advanced applications.

Contents

1

Introduction

1.1. Motivation and Objectives

Magnetic sensing has become essential in many modern electronic systems, particularly as electrification, automation, and smart devices continue to expand[5]. These applications increasingly require reliable, contactless measurement of current, position, and motion—tasks for which magnetic sensors offer clear advantages over optical, capacitive, or resistive methods, especially in harsh or high-power environments.

Progress in magnetic sensing has followed two main directions: discovering new physical transduction mechanisms and improving semiconductor technologies used to implement them. Magnetoresistive devices, for example, achieve excellent sensitivity but rely on complex material stacks and precise control of layers. In contrast, Hall-effect sensors exploit a simple physical principle yet greatly benefit from advances in microfabrication, as they can be monolithically integrated in a CMOS process[18].

Recent research has shifted from maximising sensitivity to refining accuracy, offset performance, and multi-axis capability—key requirements that determine practical performance across diverse magnetic-sensor applications. This has led to the development of vertical and three-dimensional Hall structures that engineer current paths to extract richer magnetic-field information.

Different structures of 3D Hall sensors have been developed to enable the simultaneous detection of multiple magnetic-field components. Conventional implementations combine a Hall plate with two vertical Hall devices[41], but their performance is fundamentally limited by the vertical devices, which require deep n-wells for adequate sensitivity. Modern CMOS technologies provide only shallow active regions, making high-performance vertical Hall sensing increasingly difficult. An alternative approach proposes a 3D inverted-pyramid Hall sensor capable of measuring both in-plane and out-of-plane fields without the use of vertical Hall devices.

The 3D inverted-pyramid Hall sensor exhibits several significant advantages, including high in-plane and out-of-plane sensitivities, low crosstalk, and compatibility with current spinning techniques. Its fabrication process is relatively simple and almost CMOS-compatible, making the device a cost-effective alternative to state-of-the-art solutions. However, the device still suffers from a millitesla-level residual offset—higher than that of standard Hall plates—and exhibits high flicker noise, which may indicate surface instability of the device [36].

For Hall sensors, the top passivation layer can contain trap charges that interact with the active region, thereby introducing both flicker noise and offset. An efficient method to mitigate such surface-related instabilities is the implementation of a buried structure. This structure employs a reverse-biased pn-junction to isolate the Hall plate's active region from its surroundings, improving stability and reducing noise and offset [29].

Motivated by this effective approach, this thesis investigates the use of buried structures in a 3D inverted-pyramid Hall sensor, focusing on how tailored device architectures and process innovations

can overcome the limitations of conventional Hall sensors and enable more accurate and robust magnetic-field measurements. This project also investigates the influence of substrate grounding on the electrical behavior of the pyramid structures.

A 3D inverted-pyramid Hall sensor consists of a three-dimensional cavity etched into silicon, where eight contacts are strategically placed at the corners and mid-edges of the pyramid to serve as biasing and sensing terminals, respectively. The geometry is defined by anisotropic TMAH etching of (100) silicon, producing sidewalls at the characteristic 54.74° angle [35].

A central method for implementing the buried structure is the introduction of pure-boron LPCVD deposition to create ultra-shallow junctions within the device. These ultra-shallow junctions enable improved control of trap-induced charge distribution in the active region without introducing a large junction field effect or causing pinch-off of the conduction channel.

1.2. Thesis overview

This thesis is organised as follows.

First, the fundamental physical concepts relevant to Hall-effect devices are introduced, and the role of buried structures in mitigating flicker noise is discussed. A review of the state of the art in Hall magnetic sensors, together with the principles of pure-boron CVD, is also provided (Chapter 2). The subsequent chapter presents the modeling and simulation work carried out in Synopsys Sentaurus, including the boron deposition process, the corresponding fabrication steps, and the influence of p+ doping and biasing conditions (Chapter 3).

Chapter 4 details the fabrication of the Hall sensors in the Else Kooi Laboratory (EKL) at TU Delft, with emphasis on the critical microfabrication steps of the process flow. Electrical and structural characterization results of the pyramid devices are then reported (Chapter 5).

Finally, the main conclusions are summarized, and potential directions for future research are outlined (Chapter 6).

2

Background and state of the art

This chapter aims to provide a technical reference for the remainder of this thesis's discussion, beginning with the theory of the Hall effect, Hall device characteristics, and the parasitic effects associated with the Hall device.

The chapter begins with an overview of the Hall effect, its operating principle, and its fundamental characteristics. It then introduces the 1D Hall-plate sensor and outlines the extraction of its key performance parameters. The role of buried layers in Hall plates and the techniques used to realize such structures are subsequently discussed. This is followed by an introduction to the pure-boron CVD process and its relevance to the formation of ultra-shallow junctions. The chapter concludes with a review of the current state of the art in integrated Hall sensors.

2.1. The Hall effect device

The Hall effect is one of the most well-known and widely used phenomena in solid-state physics. Edwin Hall first discovered it in 1879, where he described an experiment that proved 'a permanent effect of a magnet on the distribution of currents in a system of wires'[11]. Although this observation was made years before the electron was identified, Hall's hypothesis was correct. This phenomenon found was, that current is more crowded on one side of a conductor under the effect of a magnetic field perpendicular to the current, and an electric field appears to counteract the magnetic Lorentz force.

2.1.1. Fundamental theory of the Hall effect

Lorentz force can be described in:

$$F = q_0 E + q_0[v \times B] \tag{2.1}$$

Where F is the Lorentz force acting on charge q_0 with velocity of v through an electric field E and magnetic field B. In Equation 2.1, the first term on the right-hand side is called the electrostatic force, and the second term is referred to as the magnetic Lorentz force. This second term indicates that a charged particle will experience a transverse force when moving perpendicular to a magnetic field. This force is transverse both to the velocity of this particle and the direction of the magnetic field.

2.1.2. Basic characteristics of Hall device

A straightforward example of a plate-like Hall device is shown in figure 2.1. This is a thin plate of conducting material with four contacts around it. Its length, width, and thickness are denoted by l, w, and t. C1 and C2 are biasing contacts, through which a biasing current I is fed by the applied biasing voltage V. External electric field E_e is induced by the biasing. The homogeneously applied magnetic field B is perpendicular to the plate plane and only in the z direction. This device is considered a thin and long device, as l/w » 1, which means that no current flows vertically in the plate.

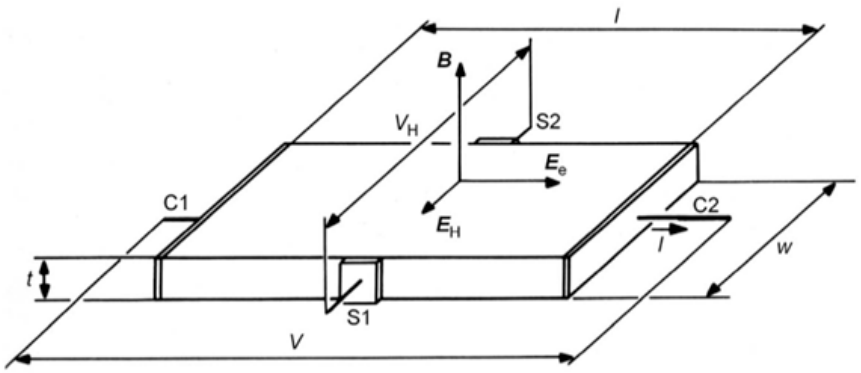


Figure 2.1: Hall device in the form of a rectangular plate[29].

Therefore, the first thing to note is that the Lorentz force acts on the free carriers. The carriers are pulled to deviate from their original path and accumulate on one side of the plate. As the carriers accumulate, an electric field is established, denoted by E_H. At equilibrium, this electric field produces an electric force that compensates for the effect of the Lorentz force. Now, the current flows only in the x direction. A set of equations can be drawn:

$$J_x = \sigma_B[E_x + \mu_H(E_y B_z - E_z B_y)] \tag{2.2}$$

$$J_x = \sigma_B[E_y + \mu_H(E_z B_x - E_x B_z)] = 0 \tag{2.3}$$

As $E_z B_x = 0$,

$$E_y = \mu_H E_x B_z \tag{2.4}$$

$$J_x = (1 + (\mu_H B_z)^2)\sigma_B E_x = \sigma E_x \tag{2.5}$$

Hence, the Hall voltage can be calculated by integrating the electric field from sensing contact S1 to S2:

$$V_H = \int_{S1}^{S2} E_y dy = \frac{\mu_H B_z V w}{l} \tag{2.6}$$

And if the Hall plate is biased with current rather than voltage, it starts from Ohm's law relationship:

$$V = RI = \frac{Il}{q\mu nwt} \tag{2.7}$$

Equation 2.6 can be expressed as:

$$V_H = \frac{\mu_H B_z I}{q\mu nwt} = \gamma_H \frac{B_z I}{qnt} \tag{2.8}$$

As the assumption of a long plate is not always the case. The geometrical correction factor G_H was introduced to characterize the geometry of a real plate. For long rectangular plates, G_H<1. For a square Hall plate with point contacts at the corners, G_H is close to 1. Bringing the geometrical correction factor into consideration, Equations 2.6 and 2.8 can be written as: Equation 2.6 can be expressed as:

$$V_H = \frac{B_z \mu_H V w}{l} G_H = S_v V B_z \tag{2.9}$$

$$V_H = \gamma_H \frac{B_z I}{qnt} G_H = S_I I B_z \tag{2.10}$$

Where $S_v = \frac{\mu_H w}{l} G_H$ denotes the voltage-related sensitivity and $S_I = \frac{\gamma_H}{qnt} G_H$ is the current-related sensitivity. The equation shows that the voltage-related sensitivity depends on both the geometry of the plates and the Hall mobility. The current-related sensitivity is inversely proportional to the carrier concentration. The product qnt is the charge per unit area of the plate [C/cm2]. In Hall magnetic sensors, low-doped (typically 10^{14}-10^{16} cm^2), small-doping-gradient layers are usually used, as it is easier to deal with more uniform quantities. Although a lower charge density means higher current-related sensitivity for Hall sensors, too low a value of charge density can cause harm due to the strong influence of the junction field effect and surface charge[1].

While sensitivity is one of the most important figures of merit of a Hall sensor, several additional performance metrics are essential for a complete evaluation of device behavior. The key figures of merit include the equivalent input magnetic noise spectral density, which determines the minimum detectable magnetic field and thus the sensor's resolution; the offset, a quasi-static output error that limits accuracy; and the non-linearity, which quantifies deviations from the ideal proportionality between output, magnetic field, and bias current. Alongside these, the cross-sensitivity serves as an important performance metric, characterizing the device's susceptibility to non-magnetic influences such as temperature, mechanical stress, or supply fluctuations[29].

These quantities are typically extracted from characterization curves such as noise spectra, offset-versus-biasing voltage plots, and input–output transfer characteristics. By distinguishing the primary figures of merit from supporting performance metrics, a clearer and more structured assessment of Hall sensor performance can be achieved[29].

2.1.3. Parasitic effects and trapped charges in Hall device and its influence on sensor properties

In applications, realizing a Hall device with high performance is challenging in the presence of parasitic effects. These effects are usually associated with device material, interface, surroundings, and fabrication imperfections. Some consequences of these imperfections lead to effects such as high offset, noise, non-linearity, and instability.

A Hall device can be modeled as a Wheatstone bridge because it is a four-contact device, as shown in Figure 2.2. Two diagonal contacts are biasing contacts, and the other are sensing contacts. The resistors represent the resistance between adjacent contacts. The offset voltage V_{offset} appears due to the mismatch of the arm resistance. As the resistors are quasi-static, the offset voltage is also a quasi-static output value that appears together with the Hall voltage at the output. The offset is usually in the order of a few millivolts. Assuming a sensitivity of 100V/A·T, that is equivalent to tens of milli-Tesla, which is much larger than the magnitude of the signal[14]. This is no doubt problematic in precision applications.

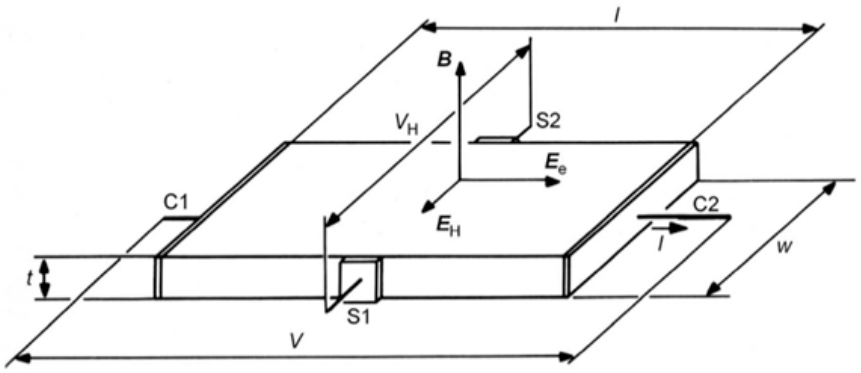


Figure 2.2: Hall device modeled as a Wheatstone bridge[14].

Offset in Hall sensors is primarily caused by fabrication imperfections and piezoresistive effects, both of which introduce asymmetry into the device. A well-established solution for reducing offset in symmetric

Hall plates is the current-spinning method. This technique exploits the non-reciprocity of symmetric Hall geometries—specifically, the interchangeability of bias and sense contacts. In a simple Hall plate, the Hall voltage rotates with the current as the plate spins, whereas the offset remains fixed. By averaging the output over the different spinning phases, the offset is effectively cancelled. Geometric imperfections and thermally induced voltages from temperature gradients can also be suppressed when the biasing direction is switched orthogonally [24].

Hall plates typically achieve offset reduction factors of 500–1000 using current spinning [2]. However, real devices often exhibit resistive nonlinearity, for instance due to the junction field effect (JFE), which limits the completeness of offset cancellation and leaves a finite residual offset [38]. The JFE refers to the modulation of a semiconductor's conductive channel by the electric field of a nearby reverse-biased p–n junction: as the depletion region expands, it reduces the local carrier concentration and effectively narrows the conduction path. In Hall sensors, this occurs when a p-type layer is located too close to the n-type active region, causing distortion of the current flow or, in extreme cases, pinch-off of the channel.

Additionally, spatial fluctuations in interface charges can introduce further asymmetry, contributing to residual offset.

There are always surface states at the surface of a solid-state semiconductor due to the discontinuity at the surface. Similarly, at an oxidized semiconductor surface, the transition from the ordered semiconductor crystal lattice to the oxide layer creates discontinuity, resulting in localized electron states called traps. As the surface is covered with an oxide layer, the oxide itself also contains more traps due to the defects in this layer[43]. According to most studies, a Si/SiO2 system has two families of defects: Pb centers and E' centers. Pb centers are silicon dangling bond centers dominating interface traps at the Si/SiO2 boundary, while E' centers are traps in bulk silicon dioxide caused by a positively charged oxygen vacancy[9][17]. As shown by Surya et al.[42], the interaction time of traps increases with their distance from the Si/SiO2 interface. In other words, traps located deeper in the oxide exchange charge with the underlying silicon more slowly, whereas traps closer to the interface respond more rapidly. Therefore, most traps in bulk silicon oxide are beyond the scope of this thesis due to their slow exchange time. Instead, we consider the characteristics of E' centers that are very close to the Si/SiO2 border and Pb centers. In the case of a thermally oxidized silicon surface, which is the most stable semiconductor surface now known, the surface densities of these charge carriers may be in the range $10^9 - 10^{13}$ cm^{-2}[29].

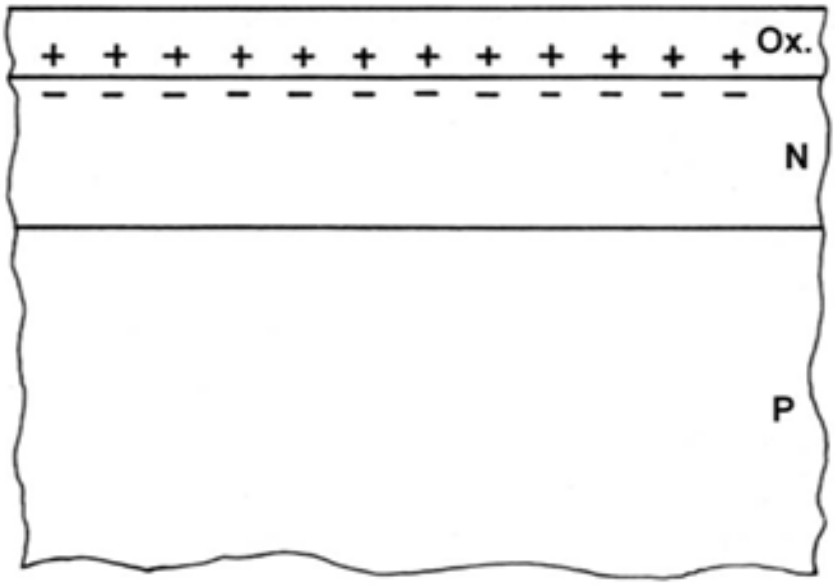


Figure 2.3: The influence of surface charges on a Hall device. N is the active layer, P is the substrate, and Ox is the oxide layer. [29].

Traps can exchange charges with the underlying region through tunneling or thermal activation. This change is dynamic as traps can capture and emit carriers[9]. Therefore, a charge at the interface induces a charge of opposite polarity in the underlying layer, see 'Ox.' And 'N' in Figure 2.3. The induced charge adds to the existing charge in the active layer, altering the device's performance[29]. To be more specific, a variation in the interface charge density δQ_{s} will induce a corresponding variation in the surface charge density, $-\delta Q_{\text{s}}$, in the active layer. The equivalent offset field of the Hall devices is given by:

$$B_{\text{offset}} \approx \frac{\Delta R}{R}\left(\frac{1}{\mu_H}\right) \tag{2.11}$$

Where R is the nominal resistance of resistors in Wheatstone bridge model in Figure 2.2. ΔR is the variance in resistance. Since:

$$R \sim \frac{1}{qnt} = \frac{1}{Q_H} \tag{2.12}$$

where $qnt = Q_H$ is the surface charge density in the Hall plate, the mismatch of resistance due to surface charge fluctuation is:

$$\frac{\Delta R}{R} = -\frac{\delta Q_s}{Q_H} \tag{2.13}$$

From the current-related sensitivity equation:

$$S_I = \frac{r_H}{qnt} G_H \tag{2.14}$$

The relationship between the sensitivity of the Hall device and the surface charge variation can be drawn:

$$\frac{\delta S_I}{S_I} = -\frac{Q_s}{Q_H} \approx S_I Q_s \tag{2.15}$$

If we assume $r_H G_H \approx 1$. Using this same assumption, Equation 2.11 and Equation 2.13, an equivalent offset field can be obtained:

$$|B_{\text{offset}}| \approx \frac{S_I}{\mu_H}\,\delta Q_s \tag{2.16}$$

From this relationship between equivalent offset and surface charge variation, it is evident that the higher the sensitivity, the more effectively trapped charges can affect the offset. From equation 2.15, one can also see that the trap charges affect the sensitivity stability in a Hall device . Similarly, the higher the sensitivity, the more unstable the sensitivity of a Hall device can be. Therefore, the sensitivity of a Hall sensor should have an upper limit; otherwise, the sensor would be spoiled.

As for the noise, it is a random, spontaneous perturbation of a deterministic signal inherent to the device's physics. It is especially emphasized because its magnitude determines the Hall device's precision limit. It is often expressed in terms of power spectral density. Generally, it is observed that the noise PSD is dependent on frequency at low frequencies and becomes more uniform at high frequencies[10].

There are several sources of noise inside the Hall device, but the most important ones for this thesis's scope are the thermal and flicker noise. Thermal noise, also known as Johnson-Nyquist noise, is the random fluctuation of voltage or current caused by the thermal motion of charge carriers in a conductor at equilibrium. It is a uniform contribution that cannot be removed[35]:

$$PSD(f) = 4k_bTR \tag{2.17}$$

Flicker noise, also called 1/f noise, is the common name or fluctuation with a PSD proportional to $\frac{1}{f^\gamma}$ with γ close to 1, usually in the range 0.7-1.3.

The Hooge model has been successful in explaining the 1/f noise in metals and bulk semiconductors[13]. For surface conduction devices like MOSFETs, the current path is confined to the narrow channel below the gate oxide surface, making them more sensitive to surface charge fluctuations. However, the Hooge noise model is empirical; it does not provide insight into the physics of the flicker noise. It is written as:

$$S_{V_\alpha} = \frac{V^2 \alpha_H}{N} \cdot \frac{1}{f^\gamma} \tag{2.18}$$

Where N is the number of carriers and α_H is the Hooge parameter. The exact value of it depends on the material. Interestingly, the lowest Hooge parameter value is found in silicon junction field-effect transistors[29]. It would be anticipated that a similar structure show a better low-frequency noise performance.

2.1.4. Hall plate and buried Hall devices

The Hall plate is a planar Hall sensor configuration that measures the magnetic field component perpendicular to the active layer of the Hall device. This type of Hall device is also referred to as a Horizontal Hall Device (HHD). The geometry of the Hall plate enables full CMOS integration, making cointegration with signal conditioning electronics easy. The Hall plate integrated in CMOS technology is shown in figure2.4.

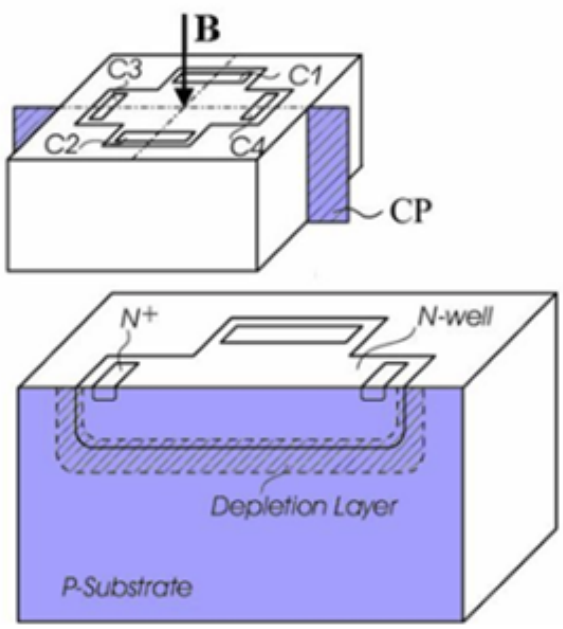


Figure 2.4: Hall plate in CMOS technology. This is an n-type device made on a p-type substrate. N-well is the active layer of the device. N+ regions are the sensing and biasing contacts[29].

The device's contacts are positioned at the four boundaries of the device, equally far away from each other. Two parallel contacts provide biasing to the device, while the other two are for sensing the Hall voltage under a magnetic field. The current spinning method, which is an offset cancellation method that constantly switches the sensing and biasing contacts, can reduce offset voltage by three orders of magnitude. Silicon Hall plates based on the current spinning method feature high precision , e.g., the residue offset achieves $3.4 \pm 2\,\mu$T in [7].

As the influence of surface instability discussed in section 2.1.3 is so negative in various aspects, there's a simple method to address is: to get away from it. Instead of contacting the active region directly with

the noisy oxide surface, one can use a reversed-bias p-n junction to separate the active and passivation layers.

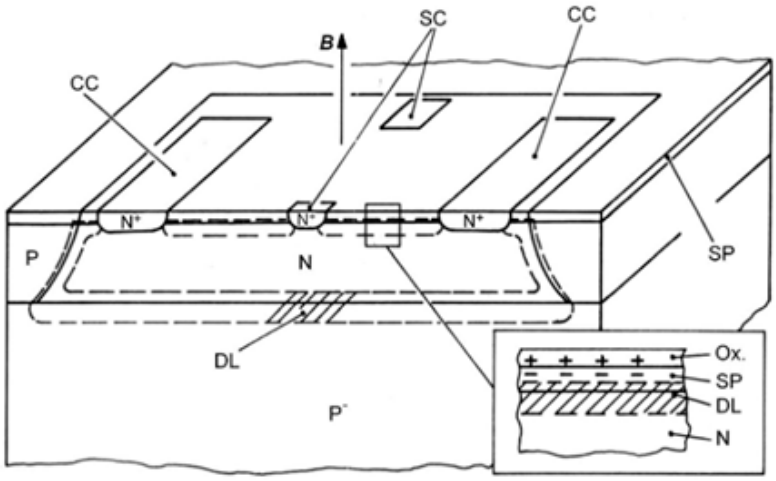


Figure 2.5: Buried Hall plate is realized using silicon bipolar integrated circuit technology. N is n-type active region, p^- the substrate, P the deep-diffused isolation walls, SP the shallow p layer, DL the depletion layer, CC the current contact, SC the sensing contact[29].

Figure 2.5 illustrates such a structure, known as a buried Hall device. Compared with the conventional structure of the Hall plate, an extra layer of a shallow p-type layer is implemented. This shallow p-type layer is used to push the active region away from the oxide surface, creating a depletion layer when encountering the active layer. This depletion layer plays the role of an electrostatic shield between the silicon/oxide interface and the active region, blocking all the fluctuations from the active layer. The shielding greatly improves the stability of the Hall device and reduces its l/f noise. Buried structures are also used to realize stable Hall devices of other semiconductor materials, such as GaAs[30].

This buried structure not only reduces surface noise and improves stability, but it also helps modulate the current-related sensitivity. As the current-related sensitivity is inversely proportional to the effective thickness of the Hall plate, the extra p-n junction depletes the upper part of the n-type active layer. It improves current-related sensitivity by reducing effective thickness. However, this extra p-n junction also has a drawback: The junction field effect, or JFE. It causes non-linearities and offsets because, when bias is applied, the region closer to high potential tends to deplete more, resulting in non-uniformity in active layer thickness[32]. The worst part is that, as current spinning can be used to reduce the offset of a device, JFE can't be mitigated by the current spinning method, leading to a worse residual offset.

2.1.5. Vertical Hall devices

The devices described in the earlier sections are Hall plates sensitive to a magnetic field perpendicular to the chip plane. For the magnetic field parallel to the device surface, the so-called vertical Hall device was devised by Popovic in 1984[7]. The first vertical Hall devices were fabricated using N-substrate P-well technology.

The most common and simplest structure of a vertical Hall device is the five-contact vertical Hall device, as shown in Figure2.6.

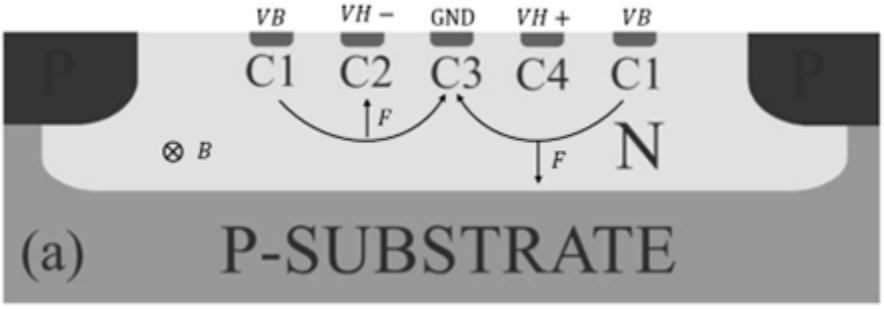


Figure 2.6: Cross-section of five contact vertical Hall device based on deep N well. If contacts C1 and C3 inject a current, the Hall voltage appears at C2 and C4, and vice versa[29].

C1 and C3 serve as the biasing contacts, while C2 and C4 function as the sensing terminals. When current flows between C1 and C3 in opposite directions, an applied magnetic field orthogonal to the current induces a Lorentz force that deflects the charge carriers in opposite directions. This lateral deflection creates a charge imbalance, which establishes an electric field and produces a potential difference between C2 and C4 that varies linearly with the applied magnetic field. A five-contact vertical Hall device (VHD) is inherently sensitive only to the magnetic-field component perpendicular to its cross-section. Therefore, by combining two such VHDs oriented orthogonally, it becomes possible to sense both in-plane magnetic-field components simultaneously. Figure 2.7 illustrates the merging of two orthogonally arranged vertical Hall devices for full in-plane magnetic-field detection.

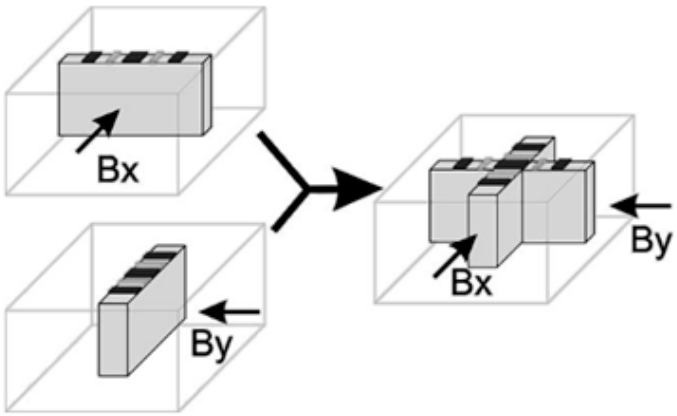


Figure 2.7: Merging of two orthogonally placed vertical Hall devices. The two devices share the central current contact[29].

Vertical Hall devices have been realized in n-well CMOS technologies by sitting the device on its junction-isolated n-type islands. However, the vertical Hall device integrated into modern CMOS technology has shortcomings due to the inherent features of CMOS technology, including shallow junctions and strongly doped layers. Based on ion implantation and diffusion, CMOS doping steps are predominantly shallow and optimized for lateral gradient control. This makes it difficult for low-offset and high-sensitivity VHD operations. Additionally, the junction field effect plays a crucial role in offset generation. The p-n junction between the n-well region and the p-type substrate generates a depletion layer around the device's active area. On one hand, it isolates the active region from the substrate. On the other hand, it modulates the thickness of the active layer unevenly under uneven potential distribution.

2.2. Various methods for creation of buried layers

In section 2.1.4, it is described that a buried structure brings enormous advantage to a Hall plate. The extra-shallow p-type layer can push the active region deeper away from the surface interference, reducing noise and instability in the Hall device. Regarding the implementation of the p-type layer, three common processes are used: ion implantation, diffusion, and in-situ epitaxy.

2.2.1. Ion implantation

Ion implantation is the introduction of energetic, charged particles into a substrate such as silicon. Dopant ions are accelerated to high energy and physically embedded into the substrate. The implantation energy can range from 1keV to 1MeV, leading to ion distribution depth from 10nm to 10μm[22].

At the beginning of the implantation process, a dopant source gas, such as BF_3 for B^+ and AsH_3 for As^+, is broken up into charged ions by a heated filament. After that, an analyzer utilizes a magnetic field to select ions with the desired mass-to-charge ratio. Then, these selected ions are accelerated to the implantation energy as they move from high voltage to ground, and the ion beam is scanned all over the wafer surface using electrostatic deflection plates. The energetic ions lose energy when colliding with electrons and nuclei in the substrate and stop at some depth within the lattice. The average depth is therefore controlled by the implantation energy. This process creates a non-equilibrium distribution of dopants in the substrate.

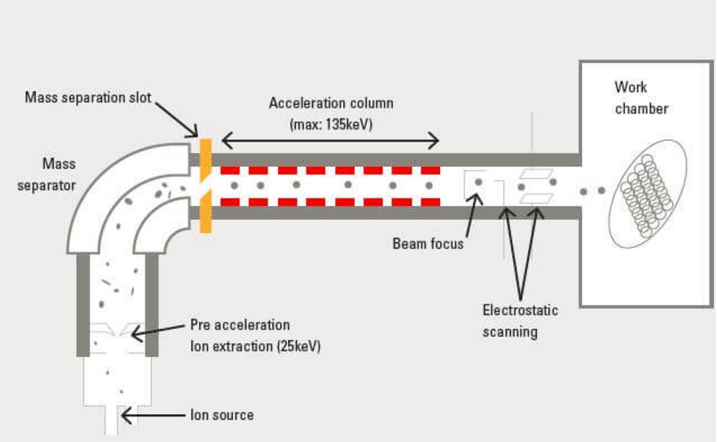


Figure 2.8: Workflow of ion implantation[3]

The main advantages of ion implantation are its precise control of doping distribution, reproducibility of doping, low thermal budget, and integration in CMOS in most cases. However, there are also a few drawbacks. The principal side effect is the damage to the semiconductor lattice due to ion collisions. The energetic ions displace silicon atoms, creating point defects, dislocations, and even amorphous layers. These structural disruptions degrade carrier mobility and can introduce traps, which are particularly detrimental in mobility- or noise-sensitive devices. Also, most of the ions as implanted are not located in substitutional sites. Therefore, a subsequent annealing treatment is needed to remove this damage and to activate the implanted ions. Conventional annealing uses an open-tube, batch furnace system to anneal. This process requires a fast and high temperature to fully remove the implantation damage, known as rapid thermal annealing (RTA).

The other big issue with ion implantation is ion channeling. In ideal situations, the projected range in amorphous polycrystalline substrates should be a Gaussian distribution. However, even for a purposely misoriented implantation on the <111> axis, the impurity profile shows an exponential ‘tail’ at deeper depth, meaning a higher doping concentration in the deeper substrate than expected. This tail is related to the ion channeling effect, which occurs when energetic ions enter a single crystal along certain crystallographic directions, and they are guided between rows of atoms in a crystal. Figure2.9 illustrates a crystal lattice viewed from the <110> direction. Trajectories of ions that are implanted in such a direction would not bring them close enough to atoms that can lose energy. So, the only resistance

force is the electrical force from atoms, resulting in these ions stopping deeper in the substrate. This effect can be simply minimized by having a thin layer of silicon dioxide to serve as a blocking layer. Or, tilting the substrate around 7 degrees so it's not aligned with any crystallographic orientations. However, ion channeling is still particularly critical for shallower implantation, which requires low energies[4].

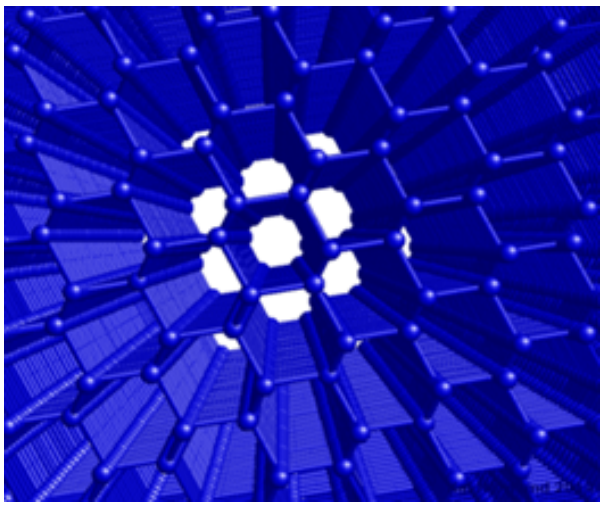


Figure 2.9: A 12nm thick silicon crystal viewed down the <110> direction[46].

For p-type doping, the drawback of transient-enhanced diffusion (TED) proves to be more evident because low-energy B^+ ion implantation is conventionally used for p-type doping. During ion implantation, the energized ions would damage the crystal, creating a lot of point defects. These defects become 'highways' for dopant atoms, where they diffuse way faster and deeper than they would in normal silicon. This phenomenon only lasts during the early stage of annealing until the excess defects recombine or diffuse away. However, for small atoms like Boron, the transient-enhanced diffusion is particularly severe, making it difficult to create an ultra-shallow junction.

2.2.2. Diffusion

Diffusion is another key method and alternative to impurity doping. Diffusion of impurities is typically done by replacing wafers in a carefully controlled furnace and passing a specific gas mixture through the furnace tube at high temperature. This method relies on Fick's law of diffusion: dopant atoms move from a high concentration region into the silicon lattice due to a concentration gradient. In contrast to ion implantation, this is a high-temperature process , which usually ranges from 600°C to 1200°C. The usage of high temperature exploits the fact that silicon's crystal lattice becomes permeable enough for dopant atoms to migrate through substitutional sites at high temperatures.

Diffusion is generally divided into two stages: predeposition and drive-in diffusion. During the predeposition stage, the wafer is exposed to a dopant-rich environment, which is usually a gas or liquid dopant source. The dopant atoms diffuse into the silicon surface until the solubility limit is reached. This creates a shallow dopant layer with a high surface concentration. During drive-in diffusion, the wafer is heated up in an inert atmosphere without dopant. This redistributes and pushes dopants deeper into the wafer while reducing the surface concentration[25].

The advantages of diffusion are: simple and low-cost compared to ion implantation, high throughput, no TED effect and no channeling effect. However, the nature of diffusion causes poor lateral control compared to ion implantation. Another huge problem is its thermal budget, as high temperature would cause unwanted dopant redistribution, which is not good for creating an extremely shallow junction.

2.2.3. In-situ doping during epitaxy

Similarly to diffusion, in-situ doping during Si or SiGe chemical vapor deposition (CVD) has been increasingly used. To be specific, diborane (B_2H_6) is widely used as p-type dopant gas. However, a few studies have demonstrated p-like doping behavior of n-type Si surfaces after exposure solely to

B_2H_6 in an oxygen-free atmosphere without any extra addition of silane-based sources[48][26]. As in the traditional diffusion method, the maximum doping concentration is limited by the solid solubility of boron in Si at the process temperature.

2.2.4. Pure Boron Deposition

In the pure B_2H_6 case of in-situ doping, boron atoms can accumulate at the Si surface beyond the usual solubility limit by varying source gas parameters and the exposure time. When all the silicon surface sites are occupied by boron atoms, the boron coverage can exceed one monolayer and lead to a distinct boron layer – known as pure boron deposition[37].

A model of the surface reaction doping for boron atoms on silicon during exposure to diborane in a CVD reactor was proposed by [16], as schematically shown in Figure 2.10. Assuming that BH_3 is the dopant species generated by thermal dissociation of B_2H_6 gas, Si—H are silicon atoms at the surface whose dangling bonds are terminated by hydrogen, and Si are silicon atoms with dangling bonds, the sequence of the mechanisms is:

1. Physisorption: $Si - H + BH_3 \rightarrow Si - HBH_3$
2. Reflection: $Si - HBH_3 \rightarrow Si - H + BH_3$
3. Migration: $Si - H + Si - HBH_3 \rightarrow Si - H + Si - HBH_3$
4. Recombination: $Si - HBH_3 + Si - HBH_3 \rightarrow 2Si - H + B_2H_6$
5. Chemisorption: $Si + BH_3 \rightarrow SiB + 3H$
6. Diffusion: $SiB \rightarrow Si + B$(diffused)

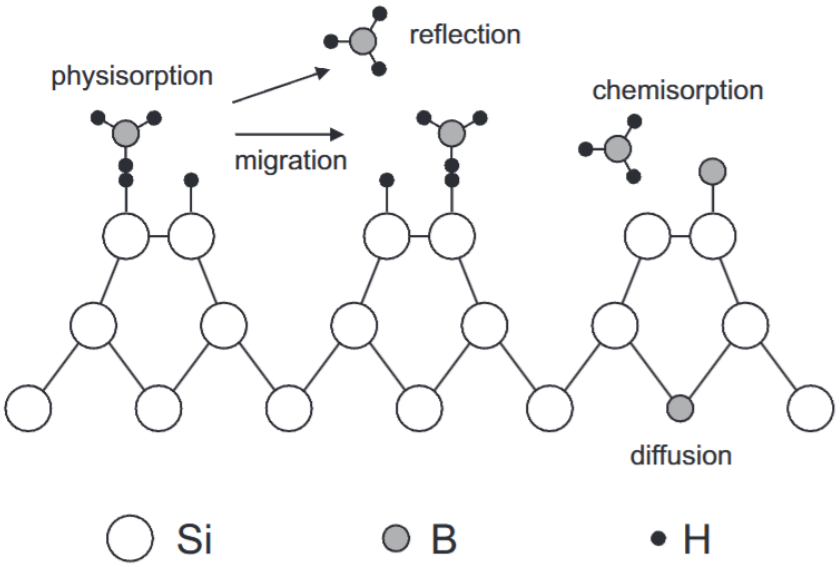


Figure 2.10: Doping reaction model for Si surface exposure to B2H6 dopant gas[16].

The boron atoms at the silicon surface will chemically react with silicon atoms, and act as a source for boron diffusion into the silicon substrate during the CVD process itself. In this process, the temperature and Si-B bond formation rate determine the boron doping of the substrate and the composition of the pure boron film. As the temperature increases, boron-silicon compounds start to be the dominant formation product; at lower temperatures, amorphous boron is formed on the silicon surface. For the case of pure boron deposition, the deposition temperature is generally between 400°C to 700°C[23].

2.2.5. Considerations for ultra-shallow junctions

The buried structure of the Hall effect device requires a shallow p+ layer. It is important for this p+-type layer to be ultra shallow and well-controlled, even though a thinner active layer results in higher

current-related sensitivity. This is because the buried structure is very similar to a junction field effect transistor in geometry, as can be seen in Figure 2.11. When the underlying n-type active layer is biased, the depletion region between the shallow-p+ layer and the n-type active layer may pinch off the active layer if the shallow-p+ layer is too thick. Even if the channel is not pinched off, the thick p+ layer would modulate the thickness of the channel due to the junction field effect, greatly affecting the linearity of sensitivity with biasing. This need for ultra-shallow junctions makes ion implantation unsuitable due to the TED effect and channeling effect it brings. The conventional diffusion method is also not sufficient for ultra-shallow junction formation. However, Boron doping via Boron CVD is a reasonable solution.

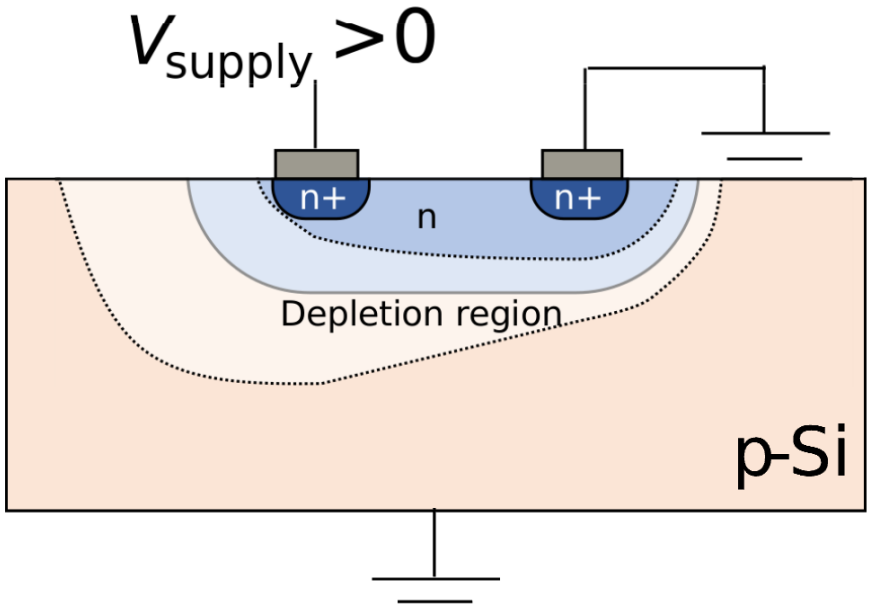


Figure 2.11: Conceptual drawing of junction field effect in Si Hall-effect plates, with constricted conduction path due to reverse bias with substrate.[6].

Along with the surface reactions that induce formation of a B-layer, diffusion of boron atoms into the Si substrate is also accomplished during the B_2H_6 exposure. The drive-in of boron atoms during the CVD process itself is very limited, given the fact that it is a relatively low-temperature process compared with the conventional diffusion method. Compared with ion implantation, there is no significant TED effect or channeling effect of boron atoms. Either the deposited or diffused boron atoms can be quantitatively controlled by varying the exposure time, while temperatures below 700°C can ensure an extremely limited junction depth even after prolonged deposition times[40].

2.3. Overview of modern Hall sensor types and performance

This section will briefly introduce recent advancements in Hall-effect devices, both 1D and 3D. Various metrics will be considered, such as residue offset, and compared. The section will conclude with a description of the new pyramid 3-axis Hall device, and identify opportunities for it to be improved compared to the literature.

2.3.1. Overview of 1D devices

1D Hall devices represent a significant evolution in magnetic sensing technology, offering high resolution and linearity in a specific sensing direction. Among the various 1D Hall sensor architectures, the Hall plates stand out due to their ability to sense magnetic fields perpendicular to the wafer surface while maintaining compatibility with planar CMOS processes.

Although the square Hall plate geometry is the most common one, the device geometries, such as cross-shaped and multiple strips designs, allow for higher bias currents and significantly boost absolute sensitivity compared to conventional layouts. Optimized horizontal Hall sensors in modern BCD tech-

nology have demonstrated sensitivity as high as 964V/A/T [49]. Generally speaking, current-related sensitivity of state of the art ranges from 200-500V/A/T , and voltage-related sensitivity of state of the art ranges from 40-60mV/V/T[7][47][19]. Techniques like spinning current and time-division multiplexing circuits reduce 1/f noise and offset. Four-phase and even eight-phase spinning schemes are used for higher-order offset cancellation, with residual offsets as low as tens of microtesla[33].

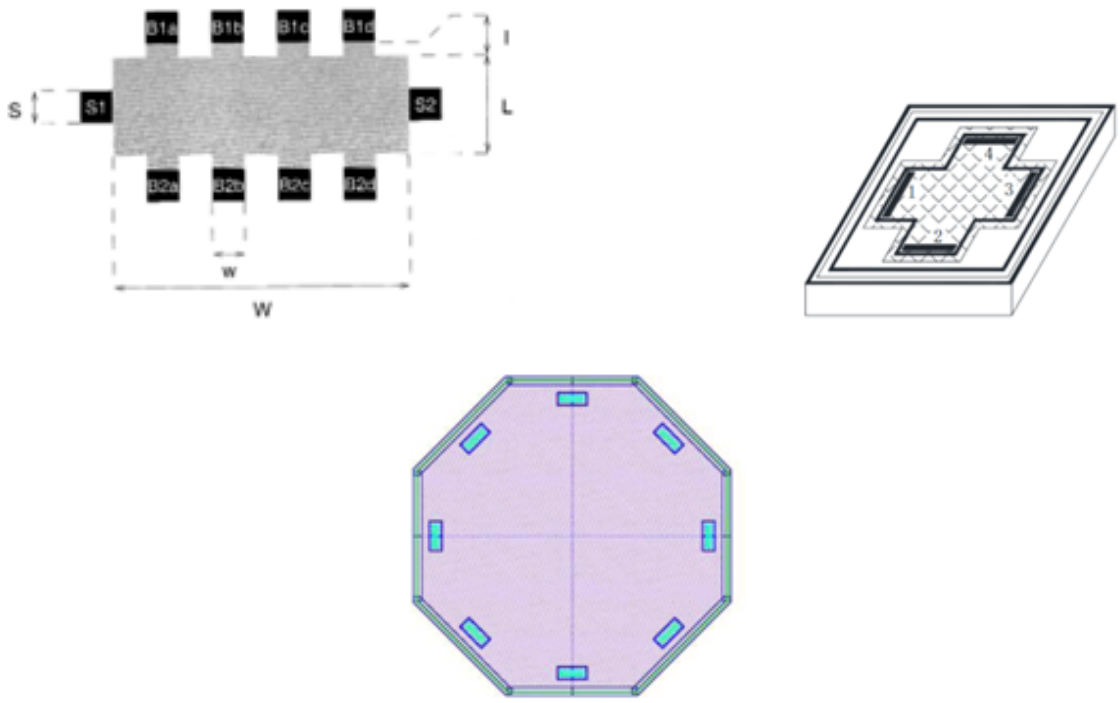


Figure 2.12: Different Hall plate geometries. Multi strips, cross-shaped and octagonal-shaped[49][15][28].

The first vertical Hall device developed by Popovich exhibited high current-related sensitivity of around 400V/A/T[31] and overall good performance. Nevertheless, the devices are not compatible with the standard CMOS technology. The next alternative was trench-Hall technology, partially compatible with standard CMOS technology. The trench-Hall technology yields highly sensitive vertical Hall devices, around 320 V/A/T, at the expense of costly pre-processing[45].

Vertical Hall devices compatible with standard CMOS technology exhibit low sensitivity due to the shallow junction depth. For example, a vertical Hall device fabricated in 0.18 µm CMOS technology and mounted on a flexible substrate has a sensitivity of 59 V/AT[12]. Although these vertical Hall devices have a good signal-to-noise ratio due to the alternative biasing and sensing of the Hall voltage, they feature a current-related sensitivity that is one to two orders of magnitude lower than horizontal Hall devices[20].

Sensor	S_I(V/A/T)	S_V(mV/V/T)	$B_{off,res}$ (uT)
AlGaN/GaN 2DEG Hall plate[7]	89	57	3.4
180nm CMOS Hall plate [47]	310	-	2000
Hall plate [24]	250	-	2000
VHD [20]	59	-	41.66
The first VHD [31]	400	-	-

Table 2.1: Comparison of 1D Hall devices

Optimized geometry and fabrication, combined with techniques such as current spinning, further enhance performance, while CMOS compatibility ensures practical integration.

2.3.2. Overview of 3D devices

Three-dimensional (3D) Hall sensors are designed to simultaneously measure all three orthogonal components of the magnetic field vector within a single sensing structure. The simplest structure of a 3D Hall device can be made by associating two orthogonal vertical Hall devices with a horizontal Hall device[27]. However, the performance of this configuration is limited by the VHDs because they require deep wells to reach a high sensitivity, which is not possible in standard CMOS technology. Furthermore, the number of contacts remains an issue.

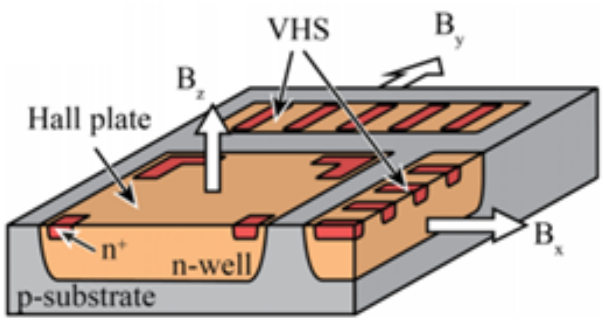


Figure 2.13: Schematic of an integrated three-axis magnetic field sensor with a Hall plate and 2 VHSs[39].

New structures of 3D silicon Hall devices have been invented to compress this schematic into a shared active region. Integrated devices utilizing a single active region can reduce the required contacts and potentially enhance spatial resolution. For example, [34] proposed a novel structure that functionally integrates two parallel-field Hall devices for the in-plane components of the magnetic field and one orthogonal Hall magnetic field version for the perpendicular to the chip. The advantage of this structure is low crosstalk and equal voltage surface conditions. The sensitivities of three channels are 130 mA/T for Bx, 220 mA/T for By, and 150 mA/T for Bz, respectively.

However, the absence of fourfold symmetry and inherent disparities in channel sensitivity and inter-axis cross-coupling can significantly degrade overall performance[39]. There are several alternatives to gain equal sensitivities and similar offsets from the three orthogonal components of the magnetic field.

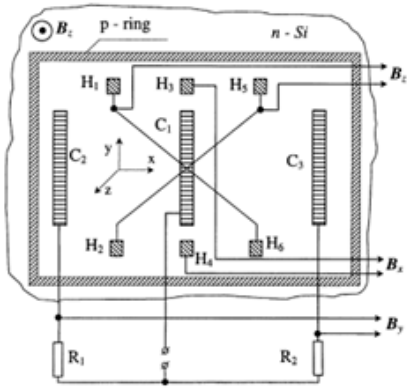


Figure 2.14: Top view of 3D Hall sensor developed by Roumenin[34].

In [39], Sander et al. designed a silicon-based 3D hall device with a hexagonal prism with symmetric sets of three contacts on its top and bottom surfaces. Currents are sent obliquely across the device, allowing it to function as three mutually crossing, identical, and effectively orthogonal Hall sensors. This configuration enables the device to measure all three spatial components of the magnetic field with nearly equal sensitivity. Its voltage-related sensitivity is S_{Vx} = 33.0 mV/V/T, S_{Vy} = 33.9 mV/V/T, S_{Vz} =

33.3 mV/V/T, meaning a nearly isotropic sensitivity. However, the fabrication is quite complex, involving non-standard CMOS processing such as double-sided patterning and deep reactive ion etching.

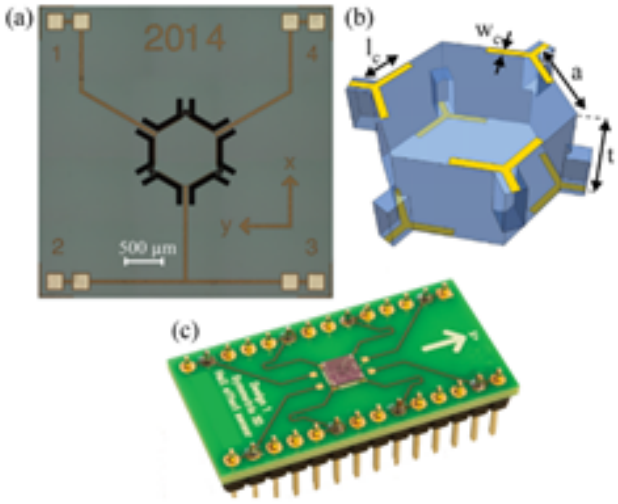


Figure 2.15: Novel sensor designed by Sanders[39].

In [36], Ruggeri et al. designed a novel 3-axis Hall-effect sensor element based on an inverted pyramid structure, realized by leveraging MEMS micromachining and CMOS processing. The device can selectively detect in-plane (X, Y) and out-of-plane (Z) magnetic field components within a single structure by applying different bias and sense configurations. The current-related sensitivities in the X and Y directions are around 70 V/AT, while the current-related sensitivity in the Z direction is around 100 200 V/AT. This design also reaches 0.2-4 mT offset by employing the current-spinning method.

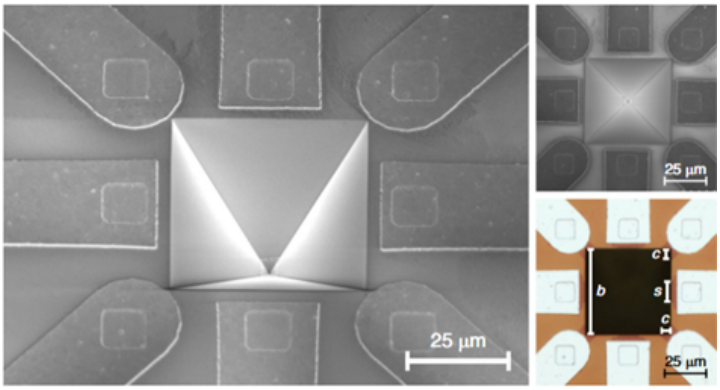


Figure 2.16: Novel sensor designed by Ruggeri[36].

Sensor	XY S_I(V/A/T)	Z S_I(V/A/T)	XY S_V(mV/V/T)	Z S_V(mV/V/T)	$B_{off,res}$ (uT)
PHD + VHDs [39]	6.27	90.1	1.9	27.3	-
Hexagonal [34]	8.6-8.8	8.7	33-33.9	33.3	40
PYRAMID [36]	64.1-82.2	14.8-17.1	19.6-21.4	200-4000	

Table 2.2: Comparison of 3D Hall devices

2.3.3. Conclusion

To overcome the limitations observed in the current 3D inverted-pyramid Hall sensor—namely its residual offset, flicker noise, and sensitivity to surface-related charge fluctuations—this thesis proposes the integration of a buried-layer structure to isolate the active region from the unstable device surface. By introducing an ultra-shallow, well-controlled reverse-biased pn-junction beneath the PYRAMID cavity, the buried layer can shield the Hall region from interface traps. Doping by pure boron CVD is chosen for creating an ultra-shallow p+ layer to avoid channeling and TED effects and to minimize the instability caused by JFE. The following chapters, therefore, focus on the design, simulation, and fabrication of buried-layer configurations tailored to the pyramid geometry, demonstrating how refined device architectures and optimized diffusion processes can yield a more accurate and robust 3D Hall sensor.

3

Simulation of the PYRAMID buried structure

This chapter presents simulations of the boron deposition process used to form the ultra-shallow junction required for the buried structure in the pyramidal Hall sensors. To minimize the junction field effect, the junction depth must be kept as shallow as possible. The chapter begins with simulations of boron ion implantation, followed by diffusion simulations that model the subsequent thermal drive-in. A comparison is then made to demonstrate the advantages of pure boron deposition in achieving ultra-shallow junctions relative to conventional implantation. The influence of deposition time on the resulting boron profile is also investigated. Finally, a device-level structure simulation is performed to assess the extent of the junction field effect in the modeled geometry.

3.1. Sentaurus Sprocess simulation

Two doping techniques—ion implantation and diffusion—were discussed in the previous chapter. Pure-boron diffusion during the deposition phase is preferred over implantation because it avoids the transient enhanced diffusion (TED) and channeling effects typically observed in implanted boron. To illustrate these differences more clearly, this section presents simulation results for both doping methods using Sentaurus Sprocess. Furthermore, the influence of deposition temperature and deposition duration on the resulting boron profiles is examined. Additional simulations performed with the Sentaurus SDevice tool are also included in this section. The simulations were performed for two types of phosphorus implantation, with the corresponding implantation parameters summarized in Table 3.1. Since this project made use of wafers that had already been etched and implanted, the implantation parameters were predetermined and therefore represented a fixed constraint for all subsequent process steps.

implantaion type	Low-doped n-well	High-doped n-well
$Dose\,(cm^{-2})$	$2 \cdot 10^{12}$	$4.5 \cdot 10^{12}$
$Energy\,(keV)$	150	100
$Expected\,junction\,depth\,(\mu m)$	0.8	0.5
$Expected\,peak\,doping\,(1/cm^3)$	$2 \cdot 10^{16}$	$6 \cdot 10^{16}$

Table 3.1: Parameters of the phosphorus implantation used for simulation

3.1.1. Boron ion implantation simulation

The setup of the boron ion-implantation simulation is shown in Figure 3.1. The sample consists of a tilted p-type silicon substrate with a (111) surface orientation, representing one sloped face of the PYRAMID structure. The tilt angle is set to 54.736°, consistent with the actual geometry. A 22 nm-thick silicon oxide layer is deposited on the surface to act as a dirt barrier layer. The substrate undergoes phosphorus implantation first, followed by an annealing step. The energy is 100keV and the implantation

dose is $4.5 \cdot 10^{12} cm^{-2}$, which are consistent with the intended implantation parameters. Subsequently, boron ions are implanted perpendicular to the tilted surface. A post-implantation anneal at 1050°C for 500 minutes is then performed to activate the dopants and repair implantation-induced damage.

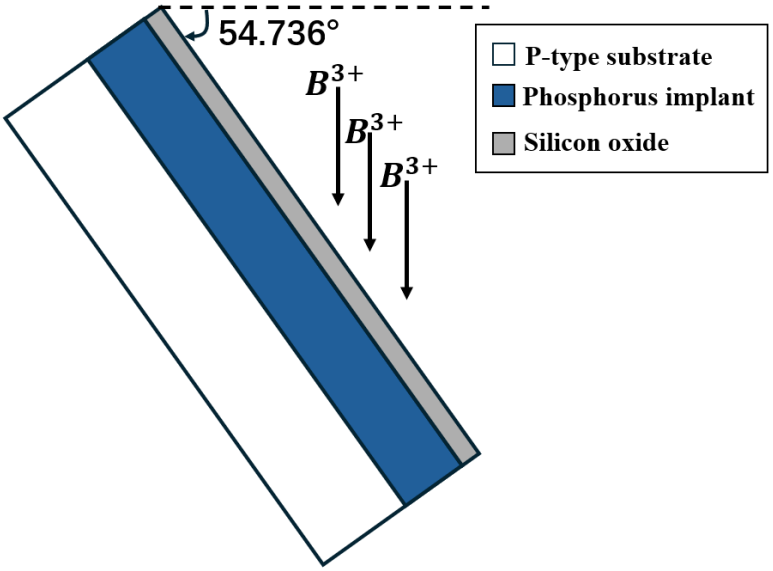


Figure 3.1: Setup of the boron ion implantation simulation. The black vertical arrows represent the direction of boron ions.

The objective of this implantation step is to form an ultra-shallow junction; therefore, a thinner P+ layer is preferred. To investigate the minimum junction thickness level with ion implantation, an implantation dose of $1 \cdot 10^{15} cm^{-2}$ and an energy of 15 keV were selected. These values were selected because they correspond to the parameter set used in this project that yielded the lowest simulated junction depth. The energy is very low to create a shallower junction, while the implantation dose is relatively high to ensure the p+ layer is not depleted. The resulting boron concentration profile is shown in Figure 3.2. This result was used as a reference to compare with the junction depth obtained from pure boron deposition. In the event that the boron CVD process did not perform as expected, this implantation-based junction profile would serve as a viable backup option.

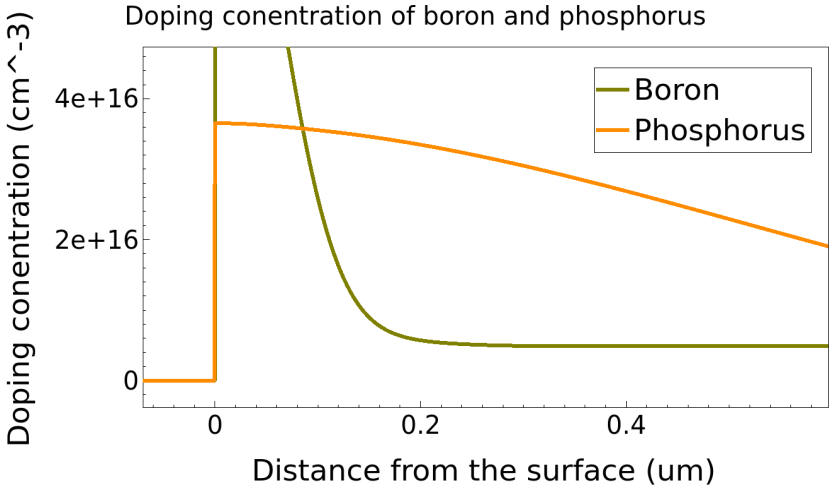


Figure 3.2: Doping profile of boron and phosphorus from implantation simulation.

3.1.2. Boron diffusion simulation

The in-situ diffusion of boron during boron CVD also needs to be simulated. During the first few seconds of boron deposition, the B_2H_6 molecules decompose on the Si surface, generating highly reactive boron atoms that readily participate in exchange reactions, where they replace surface Si atoms and occupy substitutional lattice sites. After the surface silicon solubility is reached, the boron atoms begin to accumulate and form an amorphous layer. From this moment on, diffusion becomes the dominant mechanism governing the doping of the substrate.

The setup of the boron diffusion simulation is shown in Fig. 3.3. In this setup, a few simplifications are made. First, as there is no pure boron material available in SProcess, a thin layer of silicon is created and uniformly doped with boron at a concentration equal to the atomic density of silicon. In this way, all lattice sites are effectively occupied by boron atoms. However, this approximation may lead to a faster diffusion than in reality, because the kick-in (exchange) process of boron is neglected and only diffusion is considered. A second simplification is that the initial few seconds required for the build-up of the surface solubility are ignored. As a result, the simulated profiles may overestimate the actual boron penetration depth.

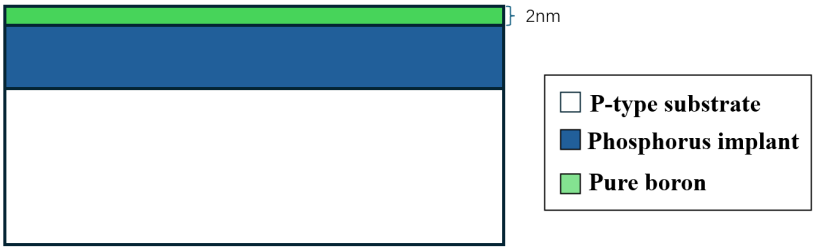
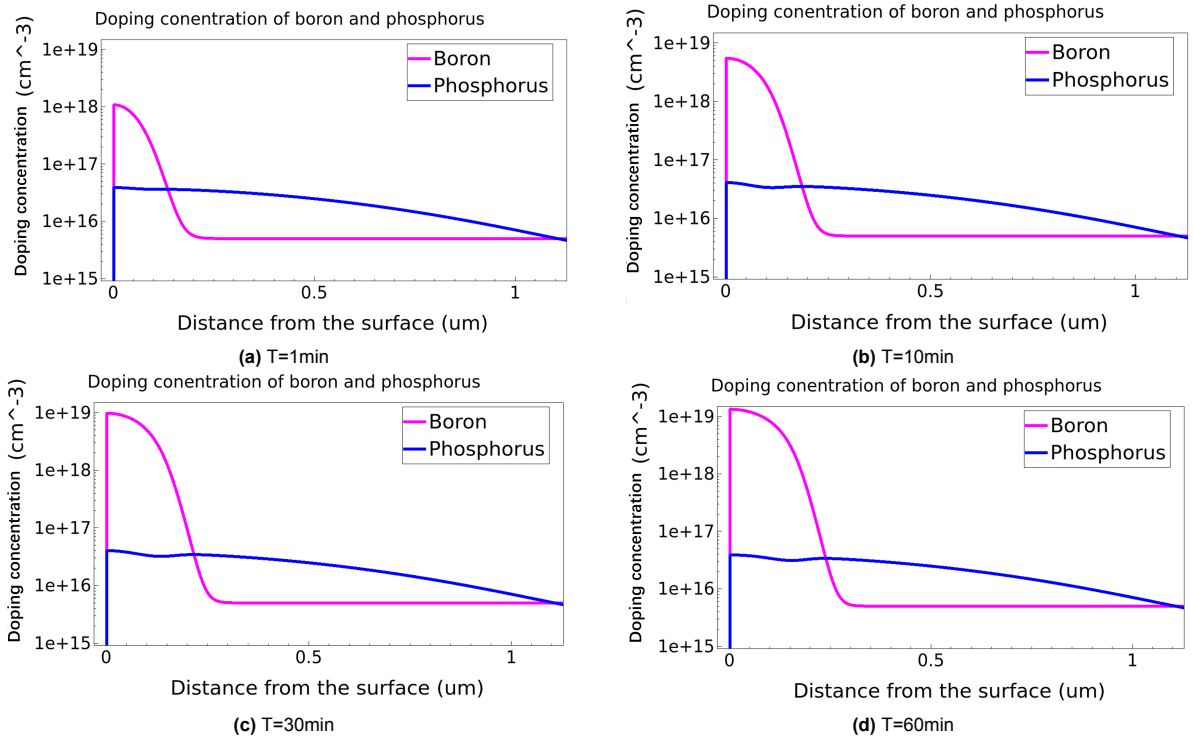


Figure 3.3: Setup of the boron diffusion simulation.

In the diffusion simulation, a 2nm-thick boron-rich layer is introduced on top of the silicon surface to emulate the pure-boron deposition obtained during LPCVD. Subsequent thermal diffusion is performed at 600°C for various durations—1 min, 10 min, 30 min, and 1h. The resulting boron concentration profiles for these annealing times are presented in Fig. 3.4. For comparison, the phosphorus doping profile corresponding to the implanted n+ region is also included. All profiles shown correspond to the high-doping wafers used in this work.



(a) T=1min

(b) T=10min

(c) T=30min

(d) T=60min

Figure 3.4: Doping profile of boron and phosphorus from diffusion simulation for duration time of 1min, 10min, 30min and 1h..

From the doping concentration curves for the different diffusion durations, the junction depth of the boron-diffused p+ layer can be estimated by identifying the point where the boron profile intersects the phosphorus profile. At this crossing point, the net active concentration approaches zero. Based on this criterion, the estimated junction depths are approximately 130 nm for 1 minute of diffusion, 180 nm for 10 minutes, 210 nm for 30 minutes, and 240 nm for 60 minutes.

Given these junction depths, the remaining active region maintains a thickness of more than 300 nm. Therefore, the p+ layer is not expected to pinch off the active layer at zero bias.

Compared with ion implantation, the 1-minute boron diffusion yields a junction depth that is broadly comparable to the implanted case. However, the junction depths predicted by the diffusion simulation are likely overestimated. This discrepancy arises because the simplified diffusion model does not accurately capture the boron deposition physics, particularly the exchange mechanism occurring at the pure-boron/silicon interface, which slows down the effective diffusion in reality. By neglecting this exchange process, the model exaggerates the junction broadening. Nevertheless, the simulation still provides a useful worst-case estimate and offers valuable insight for designing a robust fabrication flow.

3.2. Sentaurus structure simulation

The main concern lies in the junction field effect (JFE) introduced by the buried-layer structure. Because the top layer is p+-type, the middle layer is n-type (the active channel), and the bottom substrate is p-type, the stack behaves similarly to a JFET. In this configuration, the central n-type region forms the conduction channel, while the two surrounding p-type regions function as “gates.” When a reverse bias exists across either p–n junction, the depletion regions on both sides expand into the n-type layer. This reduces the effective channel width, and in the extreme case, the channel can be fully depleted. Thus, the conduction width of the active n-layer is modulated by the electrostatic biasing applied to the surrounding p-regions.

To verify that the fabricated devices will not experience unintended pinch-off, a simplified 2D cross-section of one side of the PYRAMID was constructed using Sentaurus SDE, and the electron density distribution was simulated in Sentaurus SDevice. The simulated structure is shown in Figure 3.5. The lateral dimension of the model is 25μm, matching one of the widths of the PYRAMID sidewall.

The substrate was first implanted with phosphorus to define the n-type active layer, after which the p+ region was introduced. The doping profiles used in the device simulation were directly exported from the Sentaurus Process results, which are simulated from the parameters in Table3.1, ensuring that the electrical simulation reflects the actual implant, diffusion, and activation conditions of the fabricated structure. Two N-well contacts were placed symmetrically at the top corners of the structure and biased at 1.3V. The p+ layer was contacted and biased at 0V, and the p-type substrate was also held at 0V. Under these boundary conditions, the simulation computes the electron density distribution and allows evaluation of whether the active n-layer remains sufficiently open and does not approach the pinch-off regime.

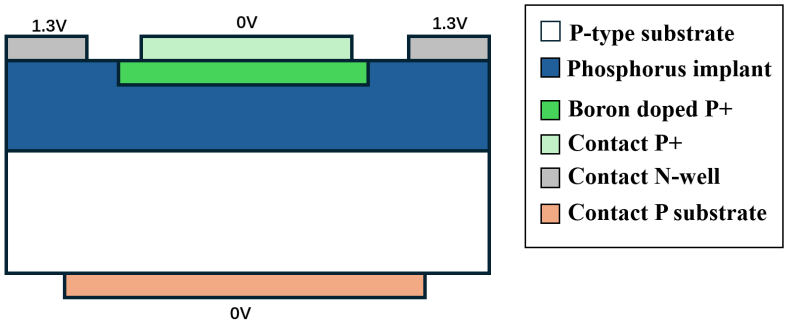


Figure 3.5: Structure of Sentaurus Device simulation.

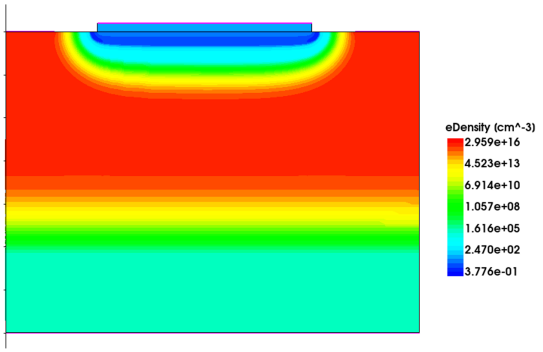
Boundary conditions were defined in Sentaurus SDevice, and the electron density distribution was obtained by solving the coupled semiconductor equations under these applied biases. First, the Poisson equation is solved using the specified doping profiles, material parameters, and electrode voltages, yielding a self-consistent electrostatic potential throughout the device. This potential determines the band bending and establishes the driving forces for subsequent carrier transport.

Next, the electron and hole continuity equations are solved in the drift–diffusion framework together with Poisson's equation, using the selected physical models for mobility degradation, bandgap narrowing, and recombination. These models ensure that transport behavior reflects realistic semiconductor physics, such as reduced mobility in heavily doped regions and carrier loss through recombination. Sentaurus employs an iterative Newton-based algorithm to simultaneously balance drift, diffusion, recombination, and generation until a steady-state solution is achieved.

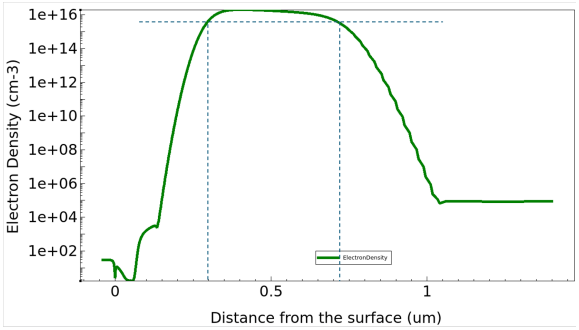
The simulated electron density distributions are presented in Figure 3.6, Figure 3.7, and Figure 3.8. These correspond respectively to:

1. a high-doped n-well with a 1-minute p+ diffusion,
2. a high-doped n-well with a 60-minute p+ diffusion, and
3. a low-doped n-well with a 1-minute p+ diffusion.

Each simulation allows direct visualization of how the p+ junction depth and the underlying n-well doping concentration influence the electron distribution and the effective width of the active region.

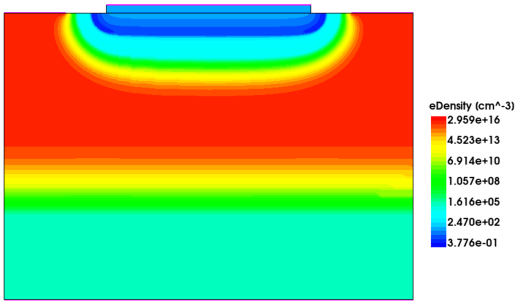


(a) Electron density distribution in the structure

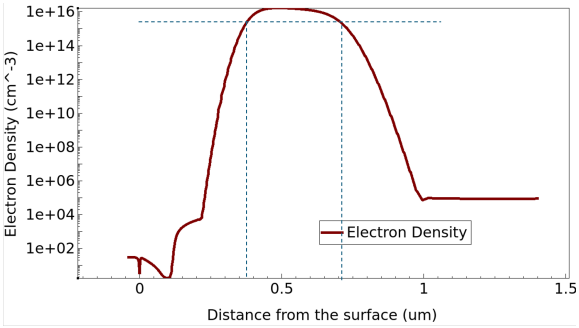


(b) Electron density distribution in the middle cross section

Figure 3.6: Electron density distribution of high-doped phosphorous profile and 1min diffusion boron profile. N-type region is biased with 1.3V.

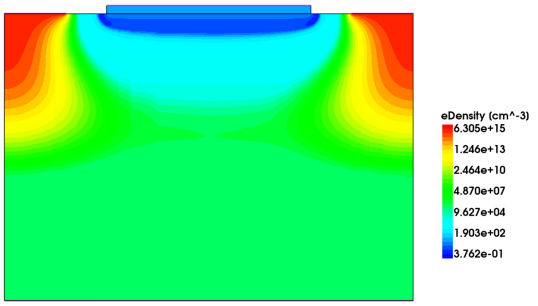


(a) Electron density distribution in the structure

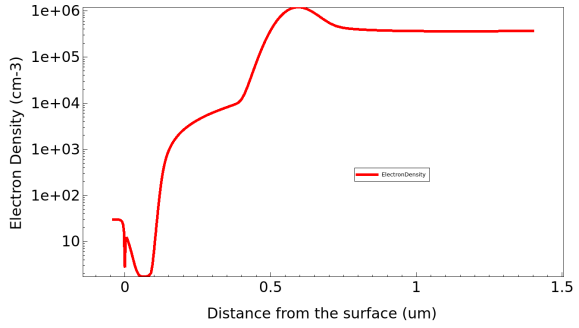


(b) Electron density distribution in the middle cross section

Figure 3.7: Electron density distribution of high-doped phosphorous profile and 60min diffusion boron profile. N-type region is biased with 1.3V.



(a) Electron density distribution in the structure



(b) Electron density distribution in the middle cross section

Figure 3.8: Electron density distribution of low-doped phosphorous profile and 1min diffusion boron profile. N-type region is biased with 1.3V.

The electron density plot directly reflects where the n-type channel remains conductive under the applied biases, making it a reliable basis for determining the active region width. In a drift–diffusion simulation, the local electron concentration is the primary quantity governing current flow: regions with sufficiently high electron density constitute the conductive path, whereas regions in which the electron concentration decreases by several orders of magnitude become effectively depleted and contribute negligibly to conduction.

The threshold for defining the effective active region width is set to $5 \cdot 10^{15} cm^{-3}$, which corresponds to the background doping level of the p-type substrate. Using this value as the cutoff ensures that the

extracted active width includes only the portion of the n-type layer where the electron concentration remains significantly above the substrate doping and is therefore capable of supporting lateral conduction. Regions in which the electron density falls below this level are dominated by the underlying p-type background and can be considered effectively depleted, contributing negligibly to current flow.

For the high-doped n-well with a 1-minute p+ diffusion, the estimated effective active region width is approximately 420 nm, with a reduction of a channel width of about 400 nm, which is sufficient for proper Hall sensor operation. In the case of the high-doped n-well with a 60-minute p+ diffusion, the effective width is reduced to about 340 nm, but still remains above the minimum required thickness. In contrast, for the low-doped n-well with a 1-minute p+ diffusion, the simulated electron density indicates that the active channel becomes fully pinched off. This result should not be interpreted as a definitive failure of the low-doped devices, as the simulation is known to slightly overestimate the p+ junction depth for a given diffusion time. However, it does highlight a non-negligible risk in fabricating the low-doping PYRAMID structures, particularly regarding channel depletion and potential loss of conductivity.

3.3. Conclusion

To conclude this chapter, the following key results were obtained from the simulations:

- A boron implantation simulation was performed to serve as a viable backup option.
- Four different diffusion durations were evaluated at a fixed temperature, and the corresponding p+ junction depths were estimated. Although the simulated junction depths may be somewhat overestimated, the results still indicate junction profiles that are broadly comparable to those obtained through ion implantation.
- A device-level structural simulation was performed to evaluate the influence of the junction field effect on the width of the active channel. The results indicate that the high-doping samples operate safely without channel pinch-off under a bias voltage of 1.3 V, whereas the low-doping samples would require a reduced bias voltage to avoid significant JFE-induced channel narrowing.

Based on these findings, an optimal device design should employ the high-dose phosphorus implantation parameters together with the shortest feasible boron diffusion (deposition) time, thereby minimizing the JFE-induced narrowing of the active region.

4

Fabrication of the Hall sensor

In this chapter, the fabrication of the inverted-pyramid Hall sensor is reported. The process flow is first summarized and then broken down into detailed sections for a better understanding. To begin, an overview of the fabrication process is presented.

4.1. Fabrication overview

Figure 4.1 shows the complete process flow of this pyramid device. This device is based on a p-type silicon substrate with a <100> orientation. The substrate is first grown with LPCVD silicon nitride, and then the silicon nitride layer is dry etched to create the pyramid square base. This silicon nitride layer served as a hard mask for TMAH etching of the bulk silicon.

After creating the pyramid using TMAH wet etching, a thin layer of silicon oxide is thermally grown on the slanted walls of the pyramid. This layer of silicon oxide serves as a dirt barrier before the pyramid is implanted with phosphorus. After ion implantation, the sample is annealed to activate the dopants and achieve a deeper junction.

The silicon nitride layer and the dirt barrier are then stripped by wet etching. A layer of PECVD TEOS oxide is deposited, and features are opened using wet etching. This oxide serves as a hard mask for boron LPCVD, as a pure boron layer does not deposit on top of silicon oxide. The atoms from the pure boron layer also diffuse into the silicon during LPCVD, resulting in a shallow p+ layer on top of the phosphorus implantation. The pure boron layer is then stripped in an aluminum etchant and HNO 69% at 110 °C. The oxide hard mask is also stripped using wet etching.

A second dirt barrier is thermally grown not only for subsequent N+ implantation but also to activate the diffused p+ layer. Then, spray coating is used to pattern the area to be implanted. The open areas are then implanted with low-energy high-dose arsenic ions. The photoresist layer and the second dirt barrier are then stripped with plasma etching and BHF, respectively.

In the next stage, a short and high-temperature thermal oxidation is applied. The goal is to activate the N+ dopant and create a thin oxide layer with a good quality near the interface. A silicon oxide passivation layer is deposited using PECVD TEOS. To mitigate the relatively inferior film quality typically associated with PECVD oxides, an additional short-duration high-temperature annealing step is introduced to increase the oxide density and enhance its structural integrity. Finally, wet etching of the silicon oxide layer is carried out to define the contact openings.

Thereafter, an HF dip etching process is performed to remove the native oxide from the contact openings, thereby ensuring reliable ohmic contact between the metal and the silicon surface. Subsequently, an aluminum–silicon layer is deposited onto the wafer surface by sputtering. The wafer is then spray-coated with photoresist, patterned, and subjected to dry etching to define the metal features. Due to the formation of thin residual sidewalls (often referred to as “fences”) along the pattern edges resulting from the dry etching process, a dedicated fence-removal step in the aluminum etcher is carried out

to eliminate these residues. Finally, a low-temperature alloying anneal in the furnace completes the fabrication sequence.

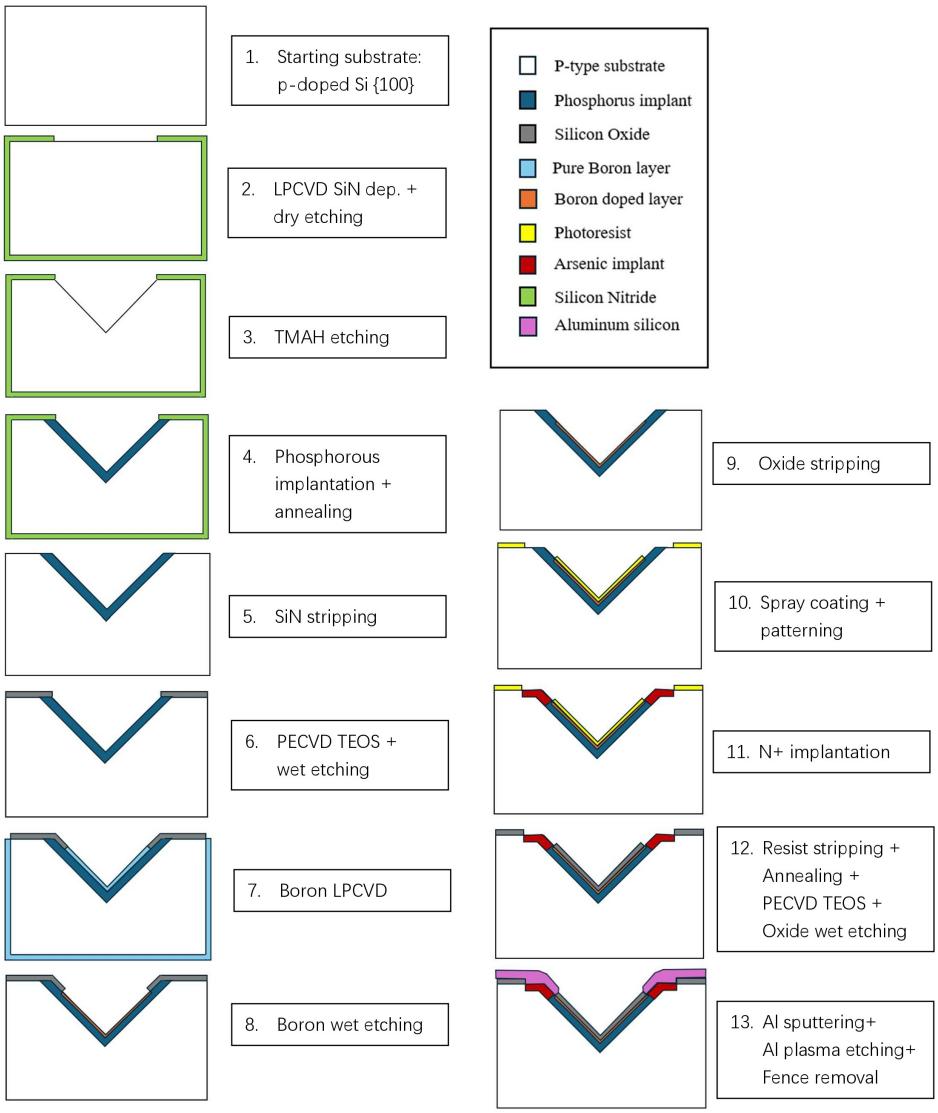


Figure 4.1: The complete process flow of the boron-capped pyramids.

4.2. Fabrication details

In this section, the detailed fabrication procedures of the proposed device are presented and discussed, with particular emphasis on the sequence of process steps, their underlying purposes, and the rationale behind key technological choices.

4.2.1. Lithography considerations and mask design

Lithography, being one of the most fundamental and critical steps in microfabrication, must be a primary consideration when developing a process plan. The fabrication of this design was conducted at the Else Kooi Laboratory (EKL), where two exposure systems were available: the SUSS MicroTec MA/BA8 mask aligner and the ASML PAS 5500/80 wafer stepper.

The SUSS MA/BA8 mask aligner employs ultraviolet light to transfer a pattern from a photomask onto a photoresist-coated substrate using either contact or proximity exposure modes. Its high-resolution optical system enables the definition of micron level features in positive-tone resists. Because the

mask is in direct contact with the wafer during exposure, the resulting pattern has a 1:1 magnification, meaning the feature dimensions on the wafer are identical to those on the mask.

In contrast, the ASML PAS 5500/80 wafer stepper operates in projection mode, where the photomask is positioned above a high-resolution reduction lens. The chromium side of the mask faces downward toward the lens, and the pattern is projected onto the wafer surface through this optical system. The projection lens provides a 5× reduction, so the features on the photomask are five times larger than the corresponding features printed on the wafer. This tool gives a resolution of around 500 nm, with an overlay error around 100 nm.

Due to the presence of the cavity, spray coating is employed to obtain a more uniform photoresist coverage. Spray coating typically produces a thicker resist layer, which is generally more compatible with the exposure characteristics of contact and proximity lithography. Cost considerations also play a significant role. Photomasks for contact aligners are substantially less expensive than the reduction reticles used in steppers, making contact lithography a more economical choice for prototype fabrication or small-volume runs. However, because this design requires eight images—equivalent to eight individual masks for a contact aligner but only two masks for a stepper—the additional cost can be effectively recovered by employing four-quadrant masks in the stepper process.

The ASML PAS 5500/80 wafer stepper, on the other hand, provides superior resolution and overlay accuracy compared to the mask aligner. The critical feature size in this design is on the order of 1µm, making resolution the foremost consideration when evaluating the advantages and limitations of each lithography tool. Moreover, as this fabrication run involves multiple wafers processed in separate batches, the automatic operation of the stepper offers a significant advantage. Once the optimal exposure dose and process recipe are established, the stepper can consistently reproduce the same patterning quality across all wafers, thereby minimizing the need for manual fine-tuning and improving process efficiency and repeatability.

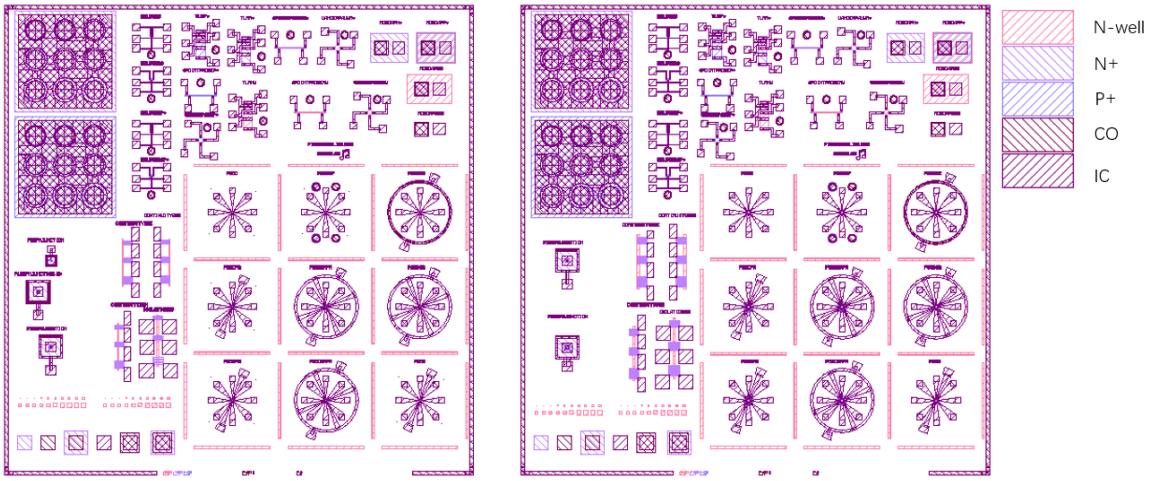


Figure 4.2: The overview of complete mask design.

The stepper masks are realized using KLayout. The overview of mask design can be seen in figure 4.2. The design incorporates two different pyramid dimensions, and each dimension requires five process layers. The layers are N-well layer, P+ layer, N+ layer, CO layer and IC layer. Table 4.1 illustrates the definition for every layer.

Given the relatively large number of layers, implementing a multi-image exposure layout was considered the most cost-effective approach. In this design, a die size of 10x10 mm^2 is adopted. For PAS5500/80 waferstepper, which has a usable image field of 21x21 mm^2 measured on wafer level, this configuration allows four complete lithographic images to be placed within a single mask field. This multi-image arrangement significantly reduces mask fabrication costs while maintaining efficient wafer utilization. The multi-image layout for two masks used in this project is shown in Figure 4.3.

Layer name	Definition
N-well layer	This layer defines the pyramid dimension and the phosphorus-implanted active region.
P+ layer	This layer defines the boron deposition region.
N+ layer	This layer defines the high-dose Arsenic-implanted regions.
CO layer	This layer defines where contacts or vias will be etched through the dielectric to expose the underlying silicon.
IC layer	This layer defines the metal routing that connects different device regions or pads.

Table 4.1: Definitions of different layers

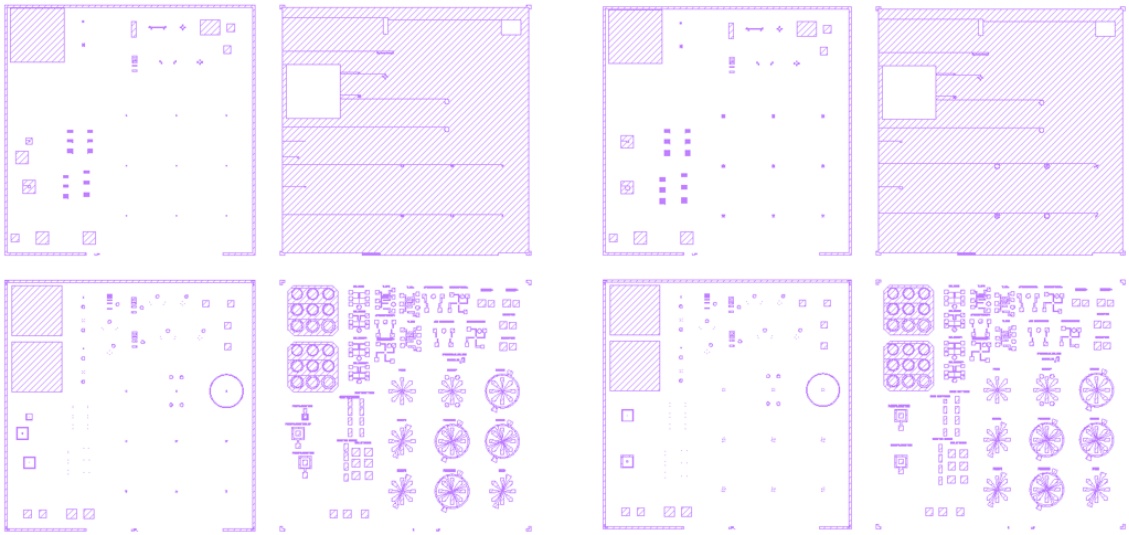


Figure 4.3: The realistic mask layout for $25\,\mu m$ and $50\,\mu m$ pyramid respectively.

There are two masks used for fabricating 25 µm and 50 µm pyramid devices, respectively. Each mask consists of four quadrants, with each quadrant corresponding to the mask image for one individual layer. For the N-well layer, existing masks from previous runs were reused, and all subsequent layers were designed and aligned based on this reference.

Each die contains nine pyramid structures, arranged in a 3 × 3 matrix. Although each pyramid device features a different layout design, the contact sizes and metal line widths are kept identical across all devices to ensure consistent electrical conditions. The variations among the layouts are intentionally introduced to facilitate a consistent comparison of different effects.

In this project, the focus is on evaluating the effect of the buried structure in the 3D pyramid Hall sensor. Therefore, a direct comparison between a pyramid with and without the buried structure is essential. Since fabrication variations can occur even across the same wafer, placing both types of pyramid structures within the same die is crucial to minimize the influence of process non-uniformities and enable a more reliable comparison.

A floating P+ layer may induce time-dependent drift in the measured signal, thereby complicating the characterization process. To prevent this, it is preferable to create a P+ contact that allows the layer to be grounded. However, when forming such a contact, the P+ region unavoidably connects to the underlying p-type substrate, which means that grounding the P+ layer also grounds the substrate. Consequently, this project also investigates the influence of substrate grounding on the electrical behavior of the pyramid structures.

These design goals yield the design concepts in this project: P ring, Ground ring, and Boron sheet. Specific layouts of each design concept are shown in Figure 4.4.

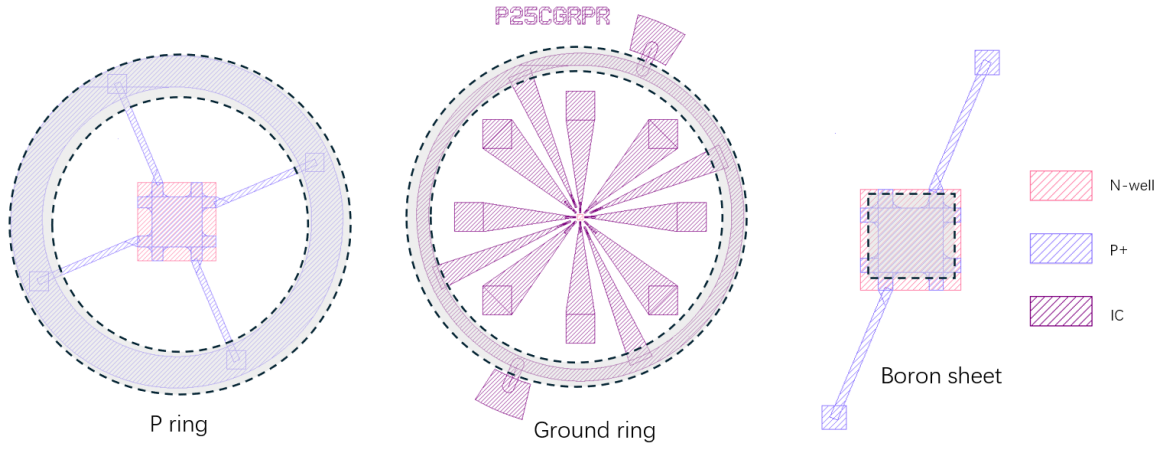


Figure 4.4: Design concept layout: P ring, Ground ring and Boron sheet.

The P-ring is a ring-shaped P+ region that connects to the P+ area located at the center of the pyramid. This ring not only provides an electrical contact to the central P+ region but also ensures a uniform grounding of the substrate. The ground ring, in turn, is a larger metal ring that can either be connected to the P-ring and routed to the bonding pads or implemented without the boron structure and placed directly in contact with the substrate. The boron sheet refers to the P+ layer embedded inside the pyramid, along with two P+ arms extending outward toward the contact region, designed to maintain minimal physical contact with the substrate while ensuring reliable electrical connectivity.

With these design concepts, various combinations are created for the 3x3 matrix, allowing the effect of every design aspect to be investigated. The 3x3 matrix is illustrated in figure 4.5

Control	**Control With Ground pad**	**Control With Ground ring**
Control With P(Boron) ring	**Boron sheet With P ring And Ground ring**	**Boron sheet With Ground ring**
Boron sheet With P ring	**Control With P ring And Ground Ring**	**Boron sheet**

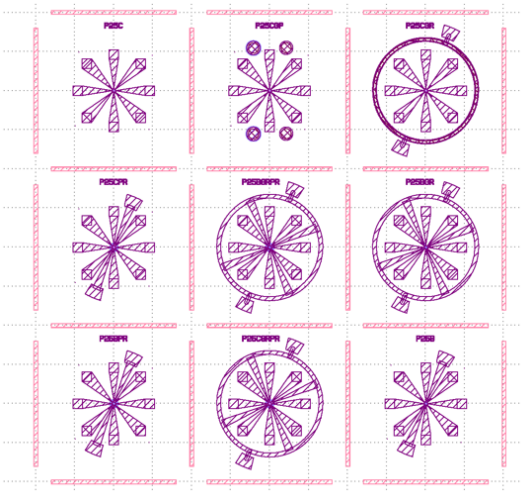


Figure 4.5: 3x3 matrix layout of different configurations of pyramid.

The zoomed-in view of one of the pyramid devices is shown in Figure 4.6. The device presented in this figure corresponds to the design incorporating both the P ring and the Ground ring. The image on the left provides a wider view of the Ground ring region, whereas the image on the right provides a narrower view of the P ring region. It can be clearly observed that the metal interconnects adopt a fan-out layout,

which helps to minimize parasitic resistance along the signal paths. The metal lines establish electrical connections to the underlying P+ and N+ regions through designated contact openings.

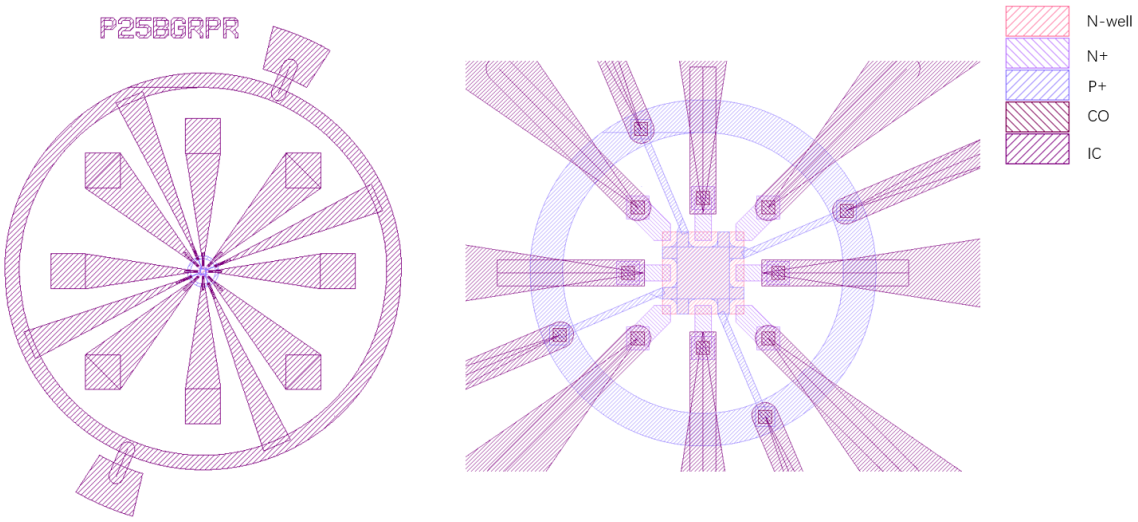


Figure 4.6: Zoom-ins of pyramid device with P ring and Ground ring.

The remaining area of the die is populated with various test structures, including TLMs, four-point probes, MOSCAPs, and Van der Pauw structures for the N, N+, and P+ layers. These structures are specifically designed to characterize the sheet resistance and contact resistance of the respective layers. In addition, two diode structures are included to evaluate the electrical performance and junction characteristics of the doped regions.

4.2.2. Control experiments description

Two parallel process runs are carried out in this project. One focuses on fabricating the actual inverted-pyramid devices, while the other omits the pyramid etching step and produces flat devices. The rationale behind including both flat and inverted-pyramid samples is that, in the event of issues with the inverted-pyramid devices, the flat counterparts can be used for debugging and process verification. Both process runs follow nearly identical fabrication steps, except for a few specific differences that will be highlighted in later sections.

These two processes are based on the same substrate, which is 4-inch single-sided polished wafers. These wafers are (100) silicon wafers, meaning the surface of the wafer is parallel to the (100) crystal plane of silicon. These wafers are uniformly doped with boron, exhibiting p-type conductivity. The resistivity of the wafers is within the range of $1 - 5\,\Omega \cdot cm$, which corresponds to a doping range of 2.68e15 - 1.47e16cm^{-3}. This indicates that each processed wafer inherently exhibits a variation in net doping, depending on the initial doping concentration of the substrate. Nevertheless, for the sake of simplicity, the substrate doping will be assumed to be $1 \cdot 10^{16}\, cm^{-3}$in all subsequent analyses.

The devices were realized in two pyramid side dimensions: 25 µm and 50 µm. These two dimensions were processed on different wafers. An individual wafer for each dimension ensures that the best process parameters can be implemented accordingly. For the flat devices, only the side dimension of 25µm was fabricated to simplify the fabrication process. Since most processing steps employ automated equipment capable of handling multiple wafers in a continuous batch, maintaining a uniform wafer dimension helps avoid the need for fine-tuning or recalibration between different batches.

It is noteworthy that the flat devices are fabricated on new wafers, whereas the inverted-pyramid devices are fabricated on previously processed wafers. The top and side views of the processed structure are illustrated in Figure 4.7.

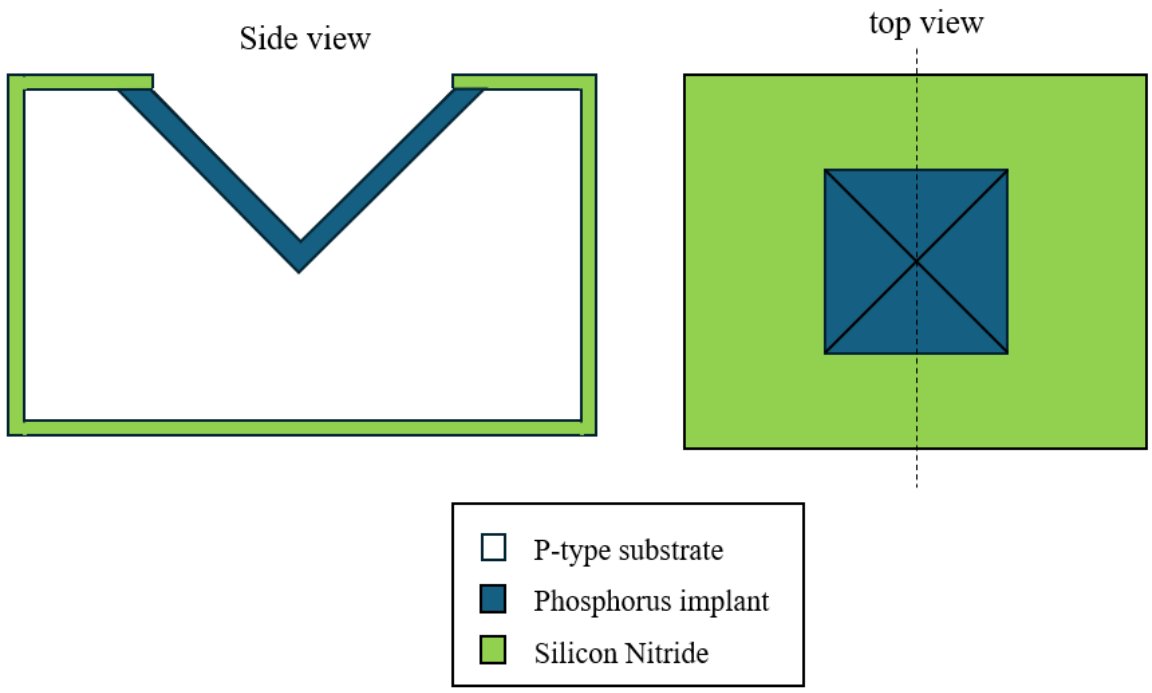


Figure 4.7: Illustration of the starting material of the inverted-pyramid devices.

The starting wafers for the inverted-pyramid devices had already undergone the following process steps: zero-layer marks etching, LPCVD SiN deposition and nitride dry etching, TMAH etching of the bulk silicon, and phosphorus implantation.

As discussed earlier, the flat devices are fabricated from scratch on new wafers. The first process step is identical to that of the pyramid devices, namely, the etching of the zero-layer alignment marks. However, from the subsequent step onward, the process flow for the flat devices diverges.

4.2.3. Coating consideration

As spin-coating is straightforward for flat devices, it becomes impractical for pyramid structures. Spin coating relies on centrifugal forces to spread the photoresist uniformly across the wafer surface: the resist is dispensed onto the sample, which is then rotated at low speed to distribute the polymer, followed by a high-speed spin to achieve the target thickness and uniformity. This method provides excellent uniformity, repeatability, and throughput on planar substrates. However, it is unsuitable for wafers containing deep cavities or high-aspect-ratio topographies. Due to resist reflow and gravitational effects, the photoresist accumulates at the bottom of the cavities, leaving the upper sidewalls and edges insufficiently coated.

Therefore, spray coating was a better option for pyramid devices. Spray coating applies photoresist by atomizing the liquid into fine droplets and directing them onto the wafer surface using a controlled spray nozzle. Unlike spin coating, the resist is deposited from above while the wafer remains stationary or slowly rotates, allowing droplets to reach deep cavities and high-aspect-ratio structures. As the droplets accumulate, they form a continuous film that can uniformly cover complex 3D topographies. However, the surface roughness of the spray-coated resist is significantly higher, mainly because there is almost no resist reflow to smooth the deposited droplets. Moreover, spray coating involves a large number of deposition parameters that must be re-optimized whenever the wafer topography changes, making the optimization difficult and time-consuming.

4.2.4. N-well implantation of flat devices

The deposited and patterned Si_3N_4 serves a dual purpose for pyramid devices: as an etching mask and an implantation mask. It is indeed a convenient choice for pyramid devices, and yet etching is not

necessary for flat devices. Therefore, a simpler alternative would be to use patterned photoresist to serve as the implantation mask. This is simple for the reason that it saves more process steps after lithography and there's no concern of enlarged patterns due to the undercut. As an implantation mask, this layer of photoresist should be thick enough to be able to guarantee no dopant penetration in the covered areas.

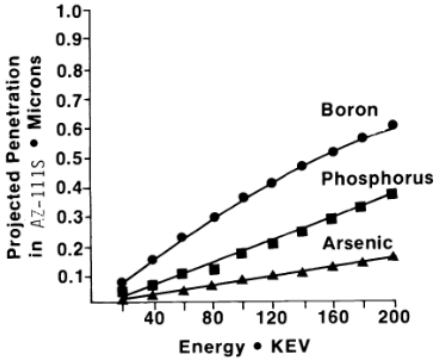


Figure 4.8: Ion implantation in AZ-111S positive photoresist.[8]

Figure 4.8 shows the projected range (implantation penetration) of three commonly used dopant ions in AZ-111S positive photoresist. The curves indicate that a sub-micron resist film can block ions with energies up to $\sim$200 keV, subject to the ion species. Implant stopping in organic resists depends primarily on the film's elemental composition and mass density. Because SPR 3012—the positive photoresist used in this work—has composition and density comparable to the AZ-111 series, Figure 4.8 provides a reasonable first-order guideline for the required mask thickness in this project.

A spin-coated Shipley SPR3012 positive photoresist with a thickness of approximately 1.4 µm was selected as the masking layer for the n-well implantation. The additional $\sim$ 1 µm thickness provides a sufficient safety margin beyond the theoretical stopping requirement. Standard photolithography procedures, including exposure and development in MF-322 developer, were carried out to define the desired implantation patterns. Prior to photoresist coating, the previously mentioned dirt barrier layer was grown on the wafer surface to prevent contamination and suppress the channeling effect during implantation. After the ion implantation, the photoresist mask was removed by means of oxygen plasma etching.

4.2.5. Annealing after N-well implantation

For flat devices, a duplicate of the pyramid devices' doping profile is desired. As mentioned earlier, the desired net peak doping concentrations are $1 \cdot 10^{16} cm^{-3}$ and $5 \cdot 10^{16} cm^{-3}$. As the flat device batch contains 15 wafers, only one of the doping concentrations is adopted for simplicity. The aimed net doping concentration is $1 \cdot 10^{16} cm^{-3}$ for the flat devices, and the desired junction depth is 1 µm. With the design objectives defined, the implantation and annealing parameters were simulated using Synopsys Sentaurus SProcess. One-dimensional simulation was carried out on a (100)-oriented silicon substrate with a background boron doping concentration of $1 \cdot 10^{16} cm^{-3}$. A thermal oxidation step at 950 °C was simulated to form the dirt barrier layer, followed by phosphorus ion implantation and subsequent annealing. As the furnace used for annealing in the EKL supports a maximum temperature of 1050 °C, this temperature was selected for the annealing process in this project. One-dimensional simulation on a (111)-oriented silicon substrate was also simulated, resulting in a similar result. Simulated doping profile is shown on the right of Figure 4.9.

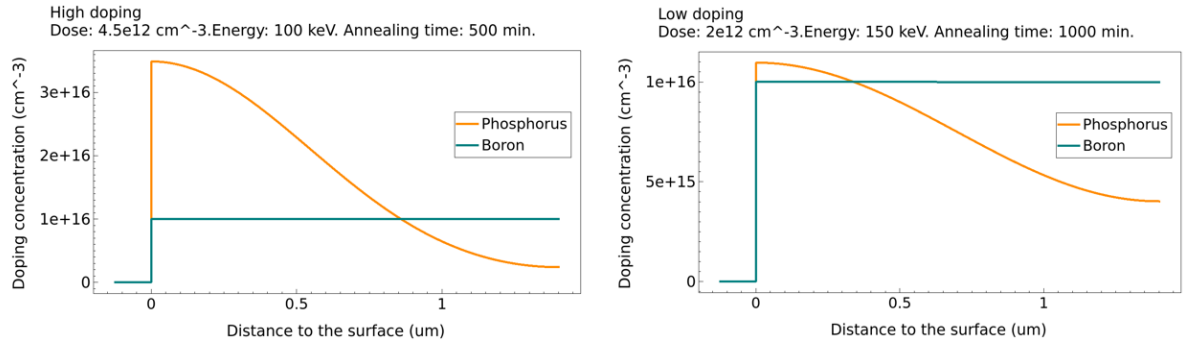


Figure 4.9: High and low N well doping profile after annealing.

For the low-doping samples, an annealing step of 1000 minutes under 1050 °C in an inert atmosphere is needed to achieve the expected junction depth. As discussed, pyramid samples are divided into low-doped and high-doped batches. For the low-doped pyramid samples, the same annealing remains. And for the high-doped wafers, whose doping profile can be seen from the left of Figure 4.9, an annealing of 500 minutes under 1050 °C in an inert atmosphere is required to create a deep enough junction. The parameters' values that fulfill the requested specifications are reported in Table 4.2.

The silicon nitride layer was ultimately removed through wet etching in orthophosphoric acid at 157 °C, which effectively stripped away both the nitride hard mask and the underlying silicon-oxide dirt barrier in a single step.

$N_D\,(cm^{-3})$	$1.2 \cdot 10^{16}$	$3 \cdot 10^{16}$
$Dose\,(cm^{-2})$	$2 \cdot 10^{12}$	$4.5 \cdot 10^{12}$
$Energy\,(keV)$	150	100
$Annealing\,Temperature\,(C)$	1050	1050
$Annealing\,Time\,(min)$	1000	500

Table 4.2: Parameters of the n-well implantation and annealing

4.2.6. Pure boron deposition

The experimental implementation of pure boron deposition on the Hall sensor constitutes a key element of this project. A thin, pure-boron film is deposited onto the active region and subsequently driven into the desired P+ diffusion layer. Following diffusion, the boron layer must be removed for two main reasons. First, pure boron is intrinsically non-conductive and would therefore obstruct the formation of metal–silicon contacts. Second, if left in place, it would continue to supply boron during later high-temperature steps, resulting in uncontrolled over-doping of the substrate. Overall, the formation of the P+ layer comprises three sequential stages: patterning, deposition, and post-diffusion removal.

Broadly speaking, a layer can be patterned in one of two ways: either the layer is first formed and subsequently etched, or a patterned mask is defined in advance, and the layer is selectively formed only in the exposed regions. In this design, the pure boron layer is deposited by Low-Pressure Chemical Vapor Deposition (LPCVD), during which in-situ boron diffusion occurs simultaneously with the film growth. Consequently, if the boron layer was fully deposited first and patterned afterwards, the underlying silicon in the undesired regions would already have been doped during deposition. To prevent unintentional doping of these areas, the pre-patterning approach—i.e., defining the openings before LPCVD— is therefore preferred.

Device structuring can be implemented relying on the selectivity of the boron deposition. The pure boron would only deposit on pure silicon and boron atoms will not be absorbed on the SiO_2[44]. Therefore, a patterned SiO_2 layer was used as a hard mask for selective boron deposition. The oxide was deposited by Plasma-Enhanced Chemical Vapor Deposition (PECVD) using TEOS (tetraethyl orthosilicate) as the

precursor. The PECVD process was carried out at 350 °C for 8.2 s, yielding a target film thickness of 300 nm.

AZ nLOF 2020 negative photoresist was used for this step. The key motivation is that the planar wafer surface can be accurately exposed using the stepper, and for a negative resist the exposed regions cross-link and harden. Since the large opening for boron deposition lies inside the cavity, the planar surface—rather than the cavity interior—is defined as the exposed and hardened area, making negative resist the more suitable choice. If a positive resist was used, the features inside the cavity would lie outside the depth-of-focus of the stepper, resulting in blurred features.

Negative resist also offers additional advantages for this geometry. As shown in the mask layout in Figure 4.4, part of the P+ outline lies within the cavity, where resist coating is inherently non-uniform.

During the optimization of the process, a trade-off was identified between surface uniformity and resist thickness. It was observed that, likely due to the high viscosity and optical characteristics of the AZ nLOF 2020 negative photoresist, the lithography pattern exhibits a measurable offset after development when the resist thickness exceeds 4 µm. This creates a very narrow thickness window: the resist must remain below this threshold to avoid pattern shift, yet reducing the thickness too much increases the surface roughness, which can exceed the tolerance required for accurate stepper alignment. At higher roughness levels, the stepper can no longer reliably detect or align to the zero-layer alignment markers.

Figure 4.10 was captured under the optical microscope. The Vernier clearly indicates an upward shift of the P+ pattern relative to the zero layer, as evidenced by the misalignment of the long Vernier bars. In addition, the grey halo surrounding the features suggests that the wafer experienced over-development, which is typically manifested as resist undercut.

Figure 4.11 presents an example of a developed die exhibiting severe surface roughness. The curly black lines and scattered dark spots across the grey photoresist regions indicate nonuniform resist dissolution and poor film integrity after development.

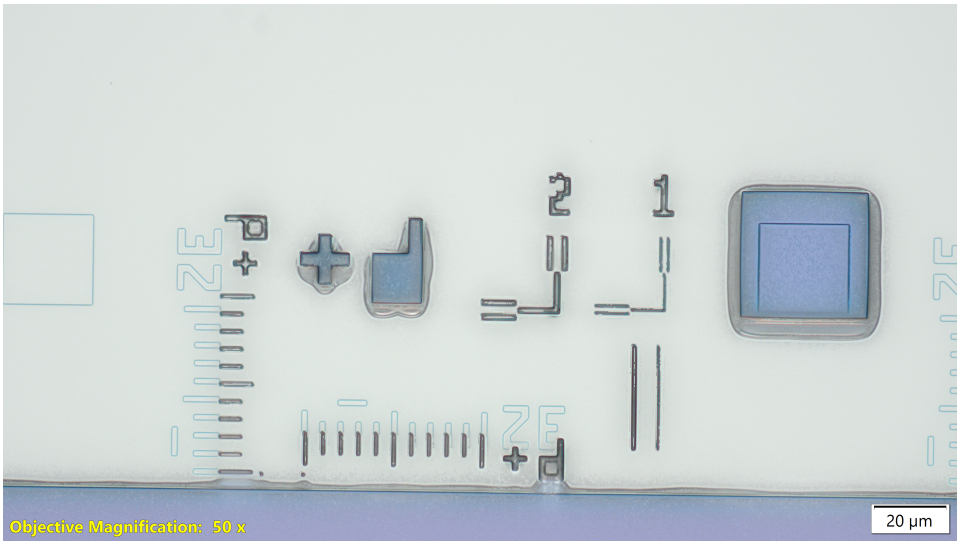


Figure 4.10: Example of shifted pattern and over-developed photoresist.

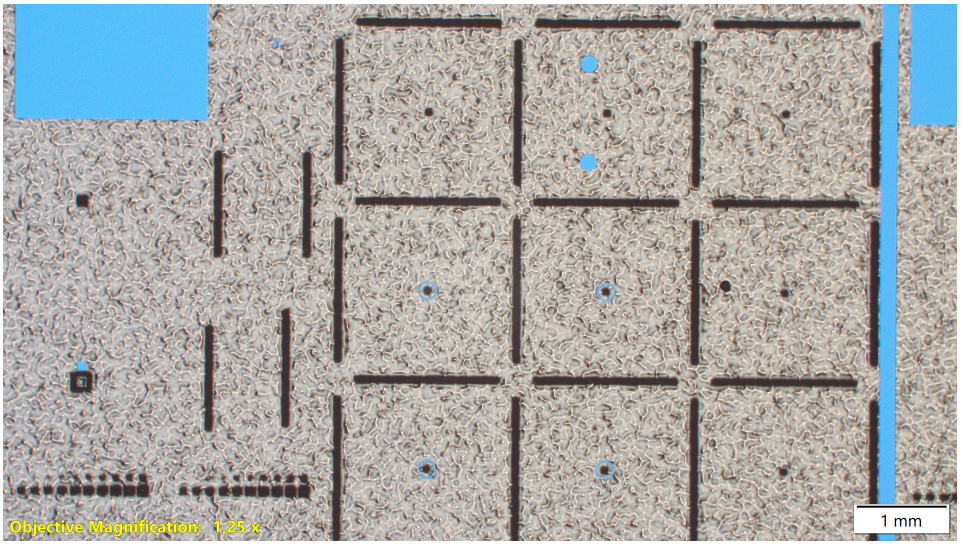


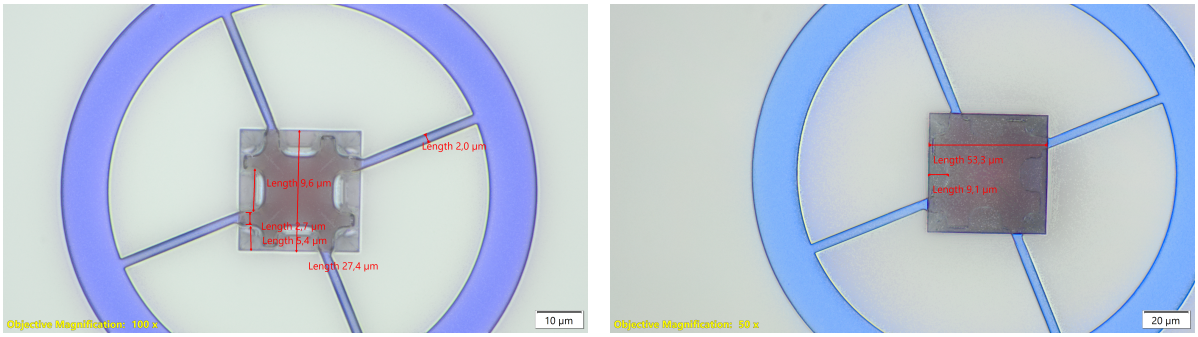
Figure 4.11: Example of rough surface of the spray-coated photoresist.

After extensive experimentation and optimization, a reliable recipe was developed. 10 minutes of manual HMDS was first conducted to improve adherence. The coating recipe is 1 set of 4 layers. The resist dispensed was 2 mL per set, and the nozzle pressure was 1000 mbar, followed by a soft bake of 4 minutes at 100°C.

The deposited resist thickness was also measured. It exhibited significant non-uniformity and uneven distribution of thickness range, with values of approximately 3-4 µm at the center of the wafers and 2.8-3.5 µm near the edges. This thickness variation originates from the operating principle of the spray coater: the nozzle scans from one side of the wafer to the other with a varying velocity, moving more slowly at the edges and more rapidly toward the center to ensure that the rotating wafer receives a comparable overall resist dose. Despite this compensation strategy, the non-uniform distribution of sprayed droplets remains clearly observable. Such variations complicate both exposure and development, as certain dies inevitably become underexposed while others are overexposed, and similarly, some dies end up underdeveloped while others are overdeveloped.

The correlation between exposure dose and development outcome also arises from another issue. Spray-coated resist tends to accumulate at the apex of the cavities, and these residual pockets require substantially longer development times compared with the resist on the planar surface. This extended development introduces a risk of lifting or partially removing the exposed, cross-linked regions. Therefore, a moderately higher exposure dose—resulting in more strongly cross-linked resist—was targeted to improve the mechanical stability of the patterned features during development.

The coated wafers were subsequently exposed using the stepper with an exposure dose of 120 mJ for the 25 µm PYRAMIDs and 160 mJ for the 50 µm PYRAMIDs. A post-exposure crosslink bake at 110 °C for 90 s was required, as the negative resist used is a chemically amplified resist that relies on thermal activation to initiate the cross-linking reaction. The wafers were subsequently transferred to the developer track for development in MF322. A total development time of 208 s was used for the 50 µm PYRAMIDs, whereas the 25 µm PYRAMIDs required 120 s. The extended development time is selected to ensure complete removal of resist residues at the bottom of the etched cavities, which is critical for achieving uniform feature profiles. A hard bake at 100 °C for 90 seconds was performed at the final stage of the development process to stabilize the remaining resist and enhance its mechanical robustness for subsequent steps. The results on a 25 µm pyramid can be observed in Figure 4.12a.



(a) 25um pyramid. **(b)** 50um pyramid.

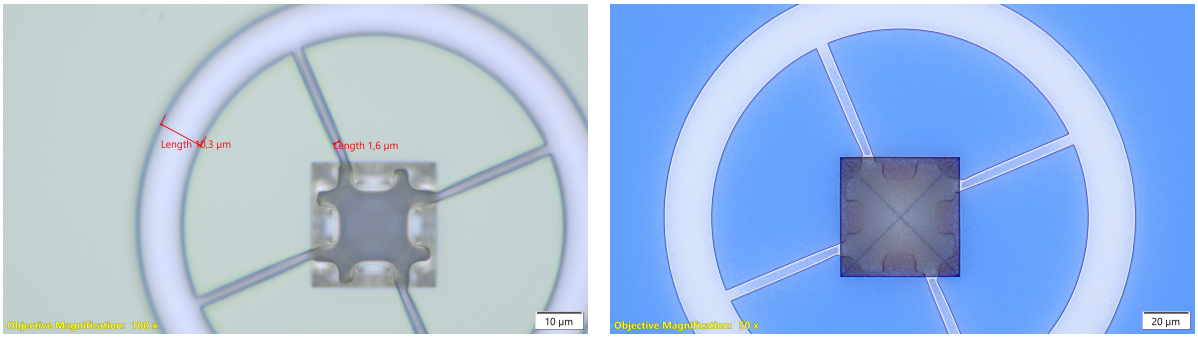
Figure 4.12: The microscopic result of the lithography step of the P+ layer on PYRAMIDs.

The structures inside the etched cavities were examined using the microscope's focus-sweep function, which enables all points within the cavity to be brought into focus simultaneously. Dimensional measurements were performed during inspection. The feature sizes on the planar surface closely match the design values. However, the exposed areas inside the cavities are approximately 0.5 micrometers larger than intended. This deviation is attributed to the increased photoresist thickness within the cavities, which leads to reduced exposure efficiency and consequently enlarged developed regions.

Furthermore, the patterns display rounded edges. This originates partly from the intentional design choice to avoid sharp corners, but it is also a consequence of the non-uniform resist thickness inside the cavities. The uneven resist distribution results in spatially non-uniform development rates, further contributing to edge rounding. The out-of-focus exposure also contributed to this issue, resulting in a blurry feature boundary.

The results for the 50 µm pyramid are shown in Figure 4.12b. The outcome exhibits the same type of issues observed in the 25 µm pyramid structures.

After the resist patterning, silicon oxide etching is needed to open the hard mask window for boron deposition. The patterned silicon oxide was later wet etched with BHF diluted with seven parts of water. Wet etching was selected in this case over dry etching to avoid physical bombardment of the spray-coated resist. The non-selective sputtering characteristic of dry etching would have removed the resist in the less-protected upper corners of the cavities, thereby exposing and partially attacking the underlying silicon oxide. The wet etchant, on the other hand, exhibited excellent selectivity for silicon oxide with respect to the spray-coated photoresist, ensuring the clean removal of the oxide while preserving the integrity of the resist. However, the inherently isotropic nature of wet etching led to a lateral undercut, resulting in a further enlargement of the patterned features. The enlargement of the structure depends on controlling the etching time. But an enlargement of around 0.5-1 µm was expected.



(a) 25um pyramid. **(b)** 50um pyramid.

Figure 4.13: The microscopic result of the etched silicon oxide hard mask for the P+ layer on PYRAMIDs.

Figure 4.13 presents the microscopic results of the two pyramid sizes after BHF etching of the silicon-dioxide hard mask. The exposed regions, shown in white, appear noticeably larger than the designed openings due to lateral under-etching inherent to the isotropic BHF process. In addition, the rounded-corner effect is more pronounced in the 50 µm pyramid structures, which is attributed to the increased cavity size. The photoresist layer was then removed by means of oxygen plasma etching.

For nanometer-thin layers, the integrity of the Pure Boron deposition is highly sensitive to any native oxide, particle contamination, or residual films on the silicon surface. Therefore, prior to loading the wafers into the CVD reactor, a pre-clean was performed by immersing the wafers in a diluted HF (0.55%) solution for 4 minutes to remove the native oxide and to hydrogen-passivate the surface, thereby suppressing immediate re-oxidation. The wafers were subsequently dried using Marangoni drying, which serves as an effective alternative to spin-rinse drying by preventing the formation of drying marks and minimising watermark defects.

The boron deposition is performed by means of Low-Pressure Chemical Vapor Deposition (LPCVD). In LPCVD, the precursor gases react on the heated wafer surface under reduced pressure, enabling uniform and conformal film growth. For electrically active pure boron layers, deposition temperatures around 400°C and 700°C are of particular interest. In this temperature regime, the process typically forms a bilayer stack consisting of an amorphous boron film on top of a thin boron–silicon compound layer at the interface.

In general, a higher deposition temperature increases the deposition rate of PureB, promotes interdiffusion between boron and silicon during deposition, and results in a smoother, lower-roughness film. Conversely, lower deposition temperatures yield rougher films but significantly suppress boron diffusion into the silicon substrate. This reduced diffusion is highly desirable from a device-integration perspective: excessive in-situ diffusion can narrow the conductive region and induce a parasitic JFET effect, in which the pn junction depletion region partially pinches off the active layer. Minimizing diffusion during deposition also preserves thermal budget flexibility for subsequent high-temperature fabrication steps, where the diffusion profile can be more precisely controlled during the dedicated annealing process.

As the LPCVD furnace in the EKL is still undergoing testing and characterization, the process parameters were required to remain fixed except for the deposition duration time. Consequently, the deposition temperature was preset to 600°C and kept unchanged throughout all experiments. The specific workflow of the boron LPCVD process is outlined.

The furnace is first ramped to 600°C at a rate of 10°C/min. When the center of the furnace reaches 300°C, the tube is pumped down. After pump-down, the process gases are introduced with flow rates of 1.1 slm Hydrogen and 900 sccm Diborane.

The units sccm (standard cubic centimeters per minute) and slm (standard liters per minute) are commonly used to specify gas flow rates in semiconductor processing equipment. The unit sccm indicates the volumetric flow of gas in cubic centimeters per minute under standardized temperature and pressure conditions, while slm represents the corresponding flow in liters per minute. By definition, 1slm=1000sccm.

As the furnace requires approximately 30 minutes to reach the target temperature, deposition already takes place during this ramp-up period at relatively low temperatures. Once the desired process temperature is reached, the effective deposition time starts to be counted. After completion of the deposition step, the furnace is cooled down at the same rate of 10 °C/min, while the gas flow is switched to 2 slm Nitrogen. The wafers are unloaded only after the system has fully cooled down.

To systematically investigate the influence of deposition time on film properties and device performance, this project employed flat samples and flat wafers to characterize four different boron LPCVD deposition durations: 1 minute, 10 minutes, 30 minutes, and 1 hour. For the pyramid devices, only the 1-minute deposition was used in order to minimize the risk of pinching off the active region.

The thickness of the PureB layer was measured immediately after deposition using spectroscopic ellipsometry. The extracted thickness values for the different deposition durations are summarized in Figure 4.14.

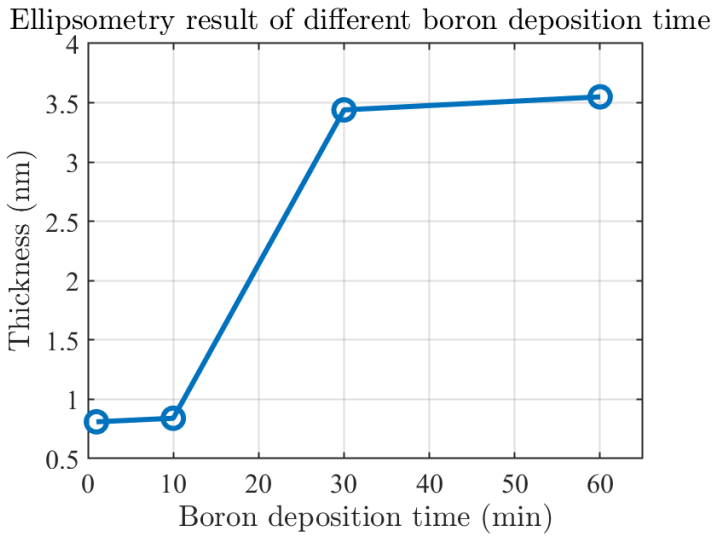


Figure 4.14: Ellipsometry result of different boron deposition time.

Aluminum metal etchant was found to be effective for removing the PureB layers. In this work, a standard Al etchant consisting of 80 % phosphoric acid, 16 % acetic acid, and 4 % nitric acid was used at 40°C. The etch rate of PureB in this solution was measured to be approximately 0.5 nm/min. The boron stripping step was therefore performed by immersing the samples in the Al etchant for 20 minutes, ensuring the complete removal of the thin boron layer. However, for samples with thicker boron deposits, this method proved inadequate due to the intrinsically slow etch rate of PureB in the Al etchant.

Because boron is known to react readily with HNO_3-based acidic solutions, facilitating the dissolution of the amorphous α-B layer during standard cleaning processes [37], an additional etching route was employed. A boiling 65 % nitric acid solution at 110°C was used in the standard cleaning wet bench to accelerate the removal of thicker PureB films. Following the boron-removal steps, spectroscopic ellipsometry was performed again to confirm the complete disappearance of the boron layer.

Both etching approaches rely on the same fundamental mechanism: the boron layer is first oxidized by nitric acid to form boron oxide or boric acid, which is subsequently dissolved through hydration and solvation. The effectiveness of these oxidation–dissolution cycles supports the interpretation of the deposited PureB film as a metal-like, highly reactive boron layer rather than a chemically inert crystalline phase.

The silicon dioxide mask was subsequently removed using buffered hydrofluoric acid (BHF). The annealing step for boron diffusion was carried out together with the annealing of the N+ implantation.

4.2.7. Definition of the N+ regions

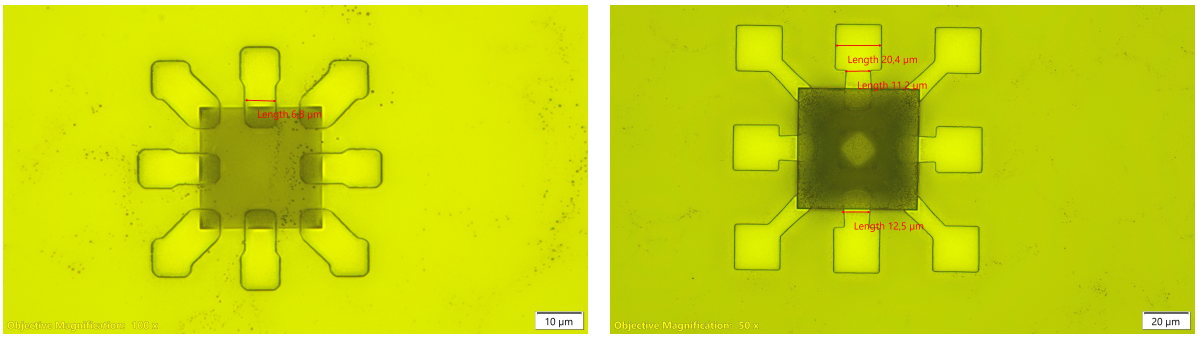
The samples were placed back into the furnace at 950 °C for 35 min to form the dirt-barrier layer. For boron dopants, this high-temperature process does not introduce too much diffusion, according to the simulation. This step was followed by coating the wafers with photoresist and patterning the N+ contact regions. As with the P+ region, spray coating was again chosen to define the N+ contact regions of the pyramid samples. The spray-coated resist acted as the implantation mask, and a positive resist was used because most of the patterns are located on or near the planar wafer surface.

Since the purpose of this resist layer is to serve as the implantation mask, its thickness must exceed the minimum stopping requirement. For the 40 keV arsenic implantation used to form the N+ region, the minimum thickness requirement is 90 nm[35]. This requirement is extremely low compared with the spray-coating capability, as the minimum achievable spray-coated thickness is around $2\mu m$. Therefore, the spray-coated resist layer provided more than sufficient protection against dopant penetration.

As in the P+ lithography, optimizing the spray-coating recipe for the pyramid samples was challenging. After multiple tests, the final recipe adopted was two sets of 8 layers recipe. A soft bake at 115 °C

for 2 min was performed between the two coating sets, followed by an additional 5 min soft bake after completing both sets. For each set, 2 mL of resist was dispensed, and the nozzle pressure was maintained at 1000 mbar. The resulting resist thickness was measured to be 5-6 µm from the wafer edge toward the center. The wafers were then aligned and exposed in the stepper.

Development was performed manually in diluted AZ-400K (1:2 with water) for approximately 2 min. The lithographic result is shown in Figure 4.15. A slight enlargement of the features was observed, which can be attributed to die-to-die resist thickness variations and the limited controllability of the manual development process. Despite this, the structures remain fully functional, with the contacts maintaining sufficient separation and appearing geometrically symmetric.

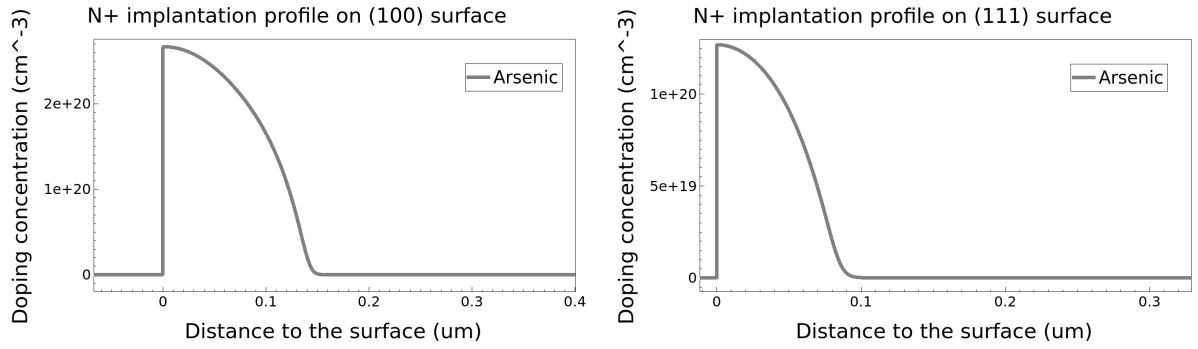


(a) 25um pyramid.

(b) 50um pyramid.

Figure 4.15: The microscopic result of the developed spray-coated resist as the hard mask for N+ implantation.

The wafers were then implanted with arsenic to form the highly doped N+ regions, which are designed to provide low-resistance ohmic contacts. The adopted implantation parameters were a dose of $5 \cdot 10^{15} cm^{-2}$ and an energy of 40 keV. A tilting angle of 0° was used to avoid the risk of short-circuiting adjacent contacts that could occur with an inclined beam. The drawback of a 0° tilt is the potential for a significant channeling effect, which may result in a deeper-than-intended junction. However, the dirt barrier formed earlier in the process helps to mitigate channeling by randomizing the ion trajectories. The simulated doping profile is shown in Figure 4.16. The photoresist mask was then removed after implantation.



(a) Implantation on (100) plane

(b) Implantation on (111) plane

Figure 4.16: Simulated doping profile of the N+ implantation after annealing.

4.2.8. Passivation layer formation

The next step for samples was the creation of a passivation layer, which protects the surface from contamination and electrically stabilizes the device. The passivation layer was designed to be a dual-layer, consisting of a thermally grown silicon oxide and a thicker layer of PECVD TEOS oxide. The samples were first undergone a pre-clean where the wafers were immersed in a diluted HF (0.55%) solution for 4 minutes to remove the unwanted native oxide.

The samples were then placed back into the furnace at 950 °C for 30 minutes to grow a thin thermal oxide layer of approximately $20nm$. This oxidation step serves multiple purposes:

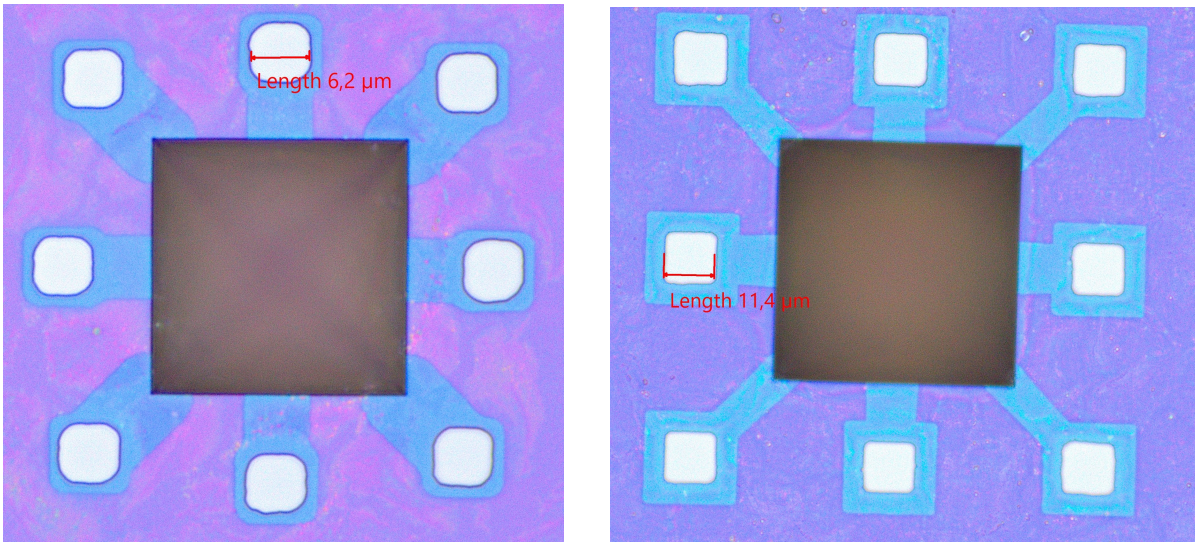
- **Formation of a high-quality Si–SiO_2 interface:** As discussed in earlier chapters, device offset and noise can be strongly influenced by interface traps located at or near the Si/SiO_2 boundary. Thermally grown oxide provides the highest interface quality, exhibiting the lowest interface trap density, which is critical for stabilizing the electrical behavior of the boundary region.
- **Activation of the implanted dopants:** Arsenic atoms implanted into silicon occupy interstitial positions and are electrically inactive. During thermal annealing, these ions diffuse into substitutional lattice sites where they become electrically active donors. This process increases carrier concentration in the N+ contact regions.
- **Recovery of implantation-induced damage:** Ion implantation introduces point defects, dislocation loops, and amorphized regions in the silicon lattice. High-temperature annealing promotes solid-phase epitaxial regrowth and defect recombination, restoring the crystalline structure and reducing leakage paths associated with implantation damage.

A second oxide layer was then deposited using PECVD at 300 °C for 8.2 s, employing tetraethoxysilane (TEOS) as the precursor. This process resulted in an additional silicon dioxide film approximately 300 nm in thickness. Following the deposition, the samples underwent an additional annealing step at 900 °C for 30 minutes in an inert atmosphere. This high-temperature treatment primarily serves to densify the PECVD TEOS oxide.

PECVD TEOS oxide, as deposited, typically exhibits lower density and a higher concentration of structural defects compared with thermally grown oxide or LPCVD oxide. These defects lead to reduced chemical resistance and poorer electrical properties. The post-deposition anneal promotes film densification, reduces defect density, and improves the film's etch resistance, dielectric quality, and overall stability.

The lithography step is the same as the one in the previous subsection. After development, the exposed silicon oxide was removed by wet etching in BHF diluted with 7 parts DI-water. Wet etching was preferred over dry etching to avoid physical bombardment of the photoresist. However, the isotropic nature of BHF etching caused a slight enlargement of the patterned features. After oxide removal was completed, the remaining resist was stripped in a 40 °C acetone bath.

The optical results are shown in Figure 4.17. The N+ regions can be identified by their distinct colour contrast relative to the surrounding areas. This contrast arises from the different thermal oxide growth rates on highly doped N+ silicon compared with the low-doped p-type regions. N+ regions contain a high concentration of implantation-induced point defects, which enhance oxidant diffusion through the growing oxide and increase the interfacial reaction rate. As a result, the thermal oxide on N+ areas grows faster. In this design, the oxide on the N+ regions is approximately 20 nm thicker than on the adjacent regions, resulting in a pronounced and easily observable color difference. Additionally, part of the dopants diffuse from silicon to the oxide, altering the refractive index. The contact openings are well within the N+ region.



(a) 25um pyramid.

(b) 50um pyramid.

Figure 4.17: The optical result of the etched contact openings on the passivation layer.

4.2.9. Metal interconnections

After defining the contact openings, the native oxide was removed by a 4-minute dip in 0.55% HF to ensure a clean silicon surface and low-resistance metal–silicon contact. Subsequently, a 675 nm-thick aluminium layer containing 1% silicon (Al–1%Si) was sputtered at 350 °C. Pure aluminium is known to induce spiking during subsequent high-temperature steps, such as alloying, and even during sputtering. Silicon from the wafer dissolves into the aluminum film, leaving behind voids that potentially penetrate into junctions and cause device failure. By using Al–1 %Si, the deposited film is already saturated with silicon, preventing further dissolution of substrate silicon and thereby eliminating the spiking risk.

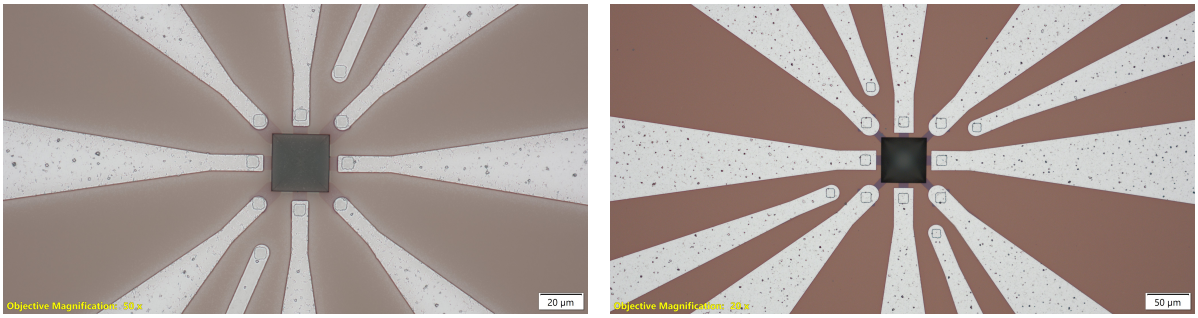
The target metal thickness for the interconnect layer is 600 nm. However, an additional 75 nm must later be sacrificed during the aluminium fence removal step. Therefore, 675 nm was deposited to compensate for this planned loss.

The metal layer was then coated with negative photoresist using a 4-layer spray-coating recipe, followed by a 4-minute soft bake at 100 °C. Negative resist was chosen because all metal-interconnect structures lie on the planar wafer surface, minimizing UV reflection issues that occur when exposing features near sloped topography. After coating, the wafers were exposed, cross-link baked at 110 °C for 90 s, and developed in MF-322.

Pattern transfer into the Al–1%Si layer was performed by plasma etching. During aluminium plasma etching, ion bombardment can sputter aluminium atoms into the resist sidewalls. These embedded Al particles cannot be removed during resist ashing in O_2 plasma, leaving metallic residues along feature edges—commonly referred to as aluminium fences. These residues must be removed prior to downstream processing to avoid electrical shorts and defects.

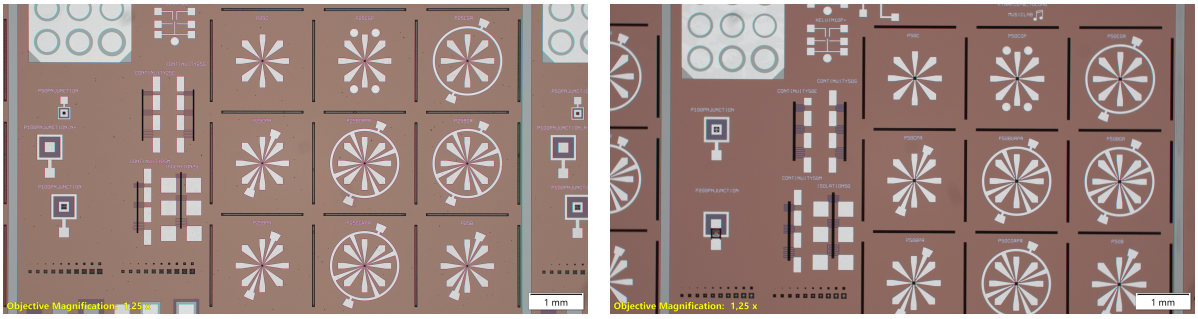
To eliminate the fences, a dedicated aluminium fence-removal step was carried out using buffered aluminium etchant. A 30-second wet etch effectively removed the sidewall residues but also consumed approximately 75 nm of the metal film, leaving the intended 600 nm-thick interconnect layer.

Finally, an alloying anneal was performed to reduce metal–silicon contact resistance. This forming-gas anneal passivates interface traps at the metal–silicon junction and improves both contact integrity and long-term device stability.



(a) 25um pyramid. **(b)** 50um pyramid.

Figure 4.18: The optical result of the pyramid devices



(a) 25um PYRAMIDs. **(b)** 50um PYRAMIDs.

Figure 4.19: The optical result of the pyramid matrices with different pyramid configurations

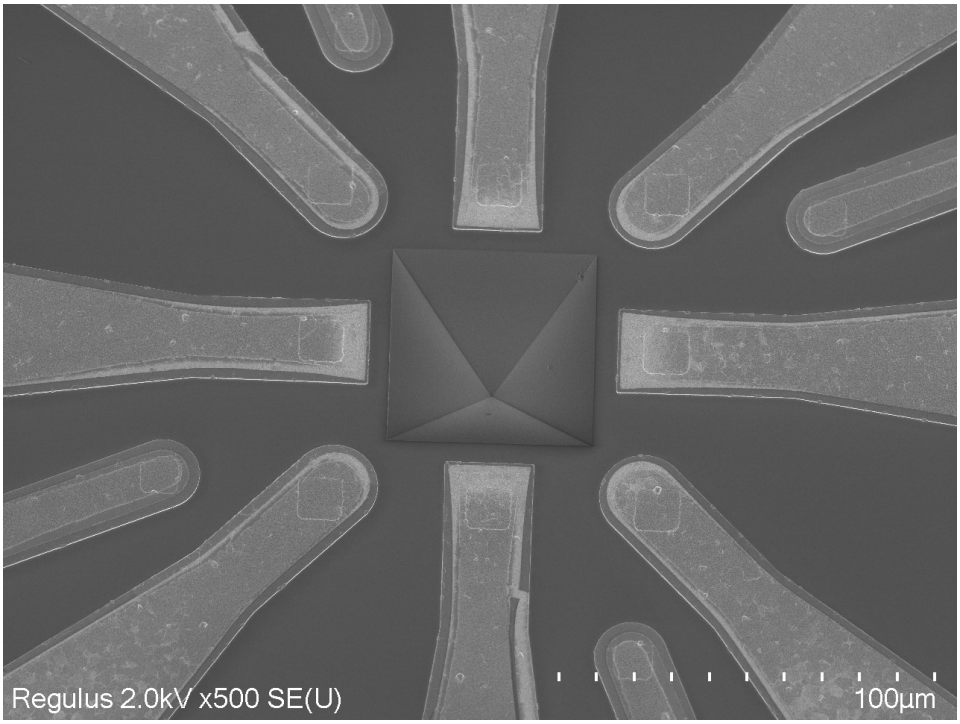


Figure 4.20: SEM result of the 50 micron pyramid device.

4.3. Conclusion

To summarize, the pyramid samples of both dimensions, as well as the flat reference samples, were successfully fabricated. All dies on the pyramid wafers were correctly patterned, although slight fabrication variations were observed across the wafer. In contrast, the flat samples exhibited a more uniform and reliable fabrication outcome. For subsequent electrical characterization, the central dies were preferred, as they exhibit fewer lithographic imperfections compared with those located near the wafer edges.

5

Characterization results of the boron-capped pyramidal Hall-effect sensors

This final chapter presents the characterization of the boron-capped pyramidal Hall-effect sensors. The chapter begins with a description of the die packaging procedure, followed by an overview of the experimental setup used. Subsequently, the simple electrical measurement results of the various semiconductor layers fabricated in this work are reported. Finally, the Hall-measurement results are presented, and key performance metrics of the devices are extracted and discussed.

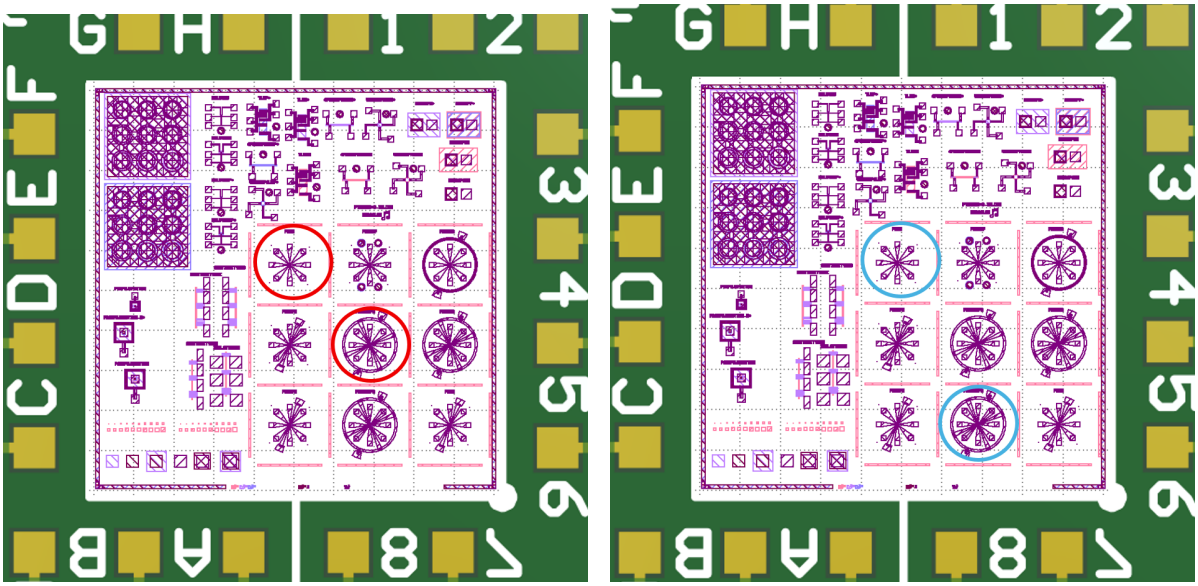
5.1. Sample packaging and working modes

5.1.1. Sample packaging

The high-doped 25 µm wafers, high-doped 50 µm wafers, and one of the flat device wafers (the one receiving a double anneal with a 1-minute boron deposition) were coated with a 3 µm-thick photoresist layer and subsequently diced. Dies located neither at the wafer edge nor at the wafer center were selected, as these regions are prone to non-uniform spray coating during fabrication. The remaining flat wafers and the low-doped PYRAMID wafers were reserved for on-wafer characterization using a probe station.

The diced dies were cleaned sequentially with acetone, isopropanol, and deionized water. Two dies of each type were mounted onto PCBs and wire-bonded for further testing.

Each PCB used for wire bonding contains 16 pads, allowing a maximum of two PYRAMID devices to be wired and characterized per die. For each wafer type, two wire-bonding configurations were selected, resulting in a total of four PYRAMID devices being wire-bonded per wafer type, as illustrated in Figure 5.1.



(a) Two devices selected for the first wire-bonding configuration (red circles).

(b) Two devices selected for the second wire-bonding configuration (blue circles).

Figure 5.1: Schematic diagram of the chosen PYRAMID devices for wire-bonding.

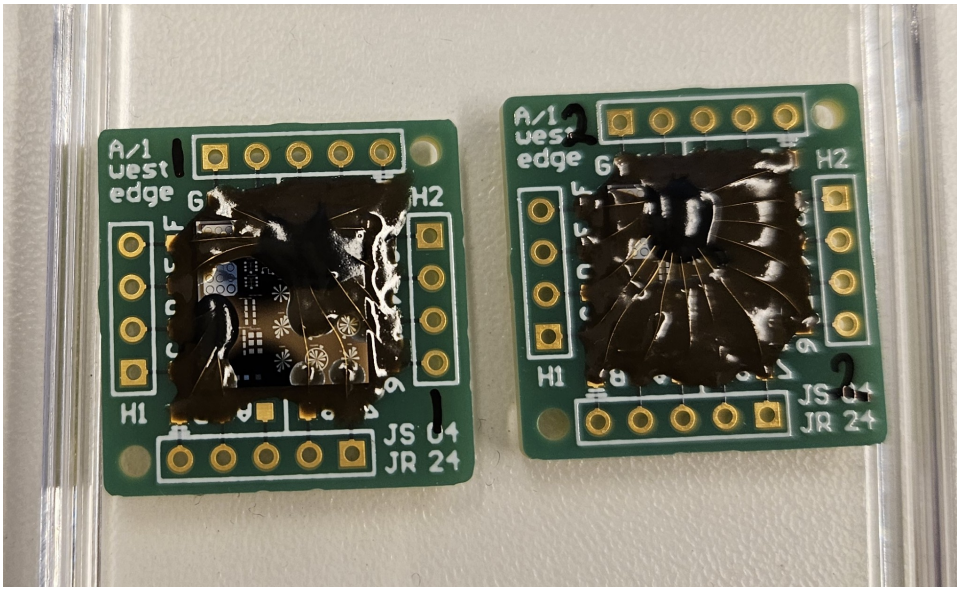


Figure 5.2: Photo of two wire-bonded dies.

For the first wire-bonding configuration, the plain PYRAMID structure was selected as the control group, along with the PYRAMID variant that included the boron pad, P-ring, and ground ring. This pairing allows a direct comparison of device performance across different structural designs while minimizing fabrication-related variations. The objective of this configuration is to isolate and study the influence of the buried structure.

For the second wire-bonding configuration, the plain PYRAMID and the PYRAMID including the P-ring and ground ring were wire-bonded. This setup aims to investigate the impact of grounding the substrate. Accordingly, two devices were wire-bonded and characterized for each die.

5.1.2. Working modes and samples

In the Hall-effect measurement, three operation modes—referred to as X mode, Y mode, and Z mode—are employed to enable full 3D magnetic field detection. The X and Y modes are identical, if rotated 90 degrees, and operate according to the same principle. In the X/Y mode, an alternating biasing pattern is applied, generating four anti-parallel current paths across the sloped faces of the pyramid. Two middle-side contacts aligned along the x-axis sense the Hall-voltage variation induced by an x-directed magnetic field, while the corresponding pair aligned along the y-axis senses the response to a y-directed field.

The two anti-parallel current components experience equal-magnitude but oppositely directed Lorentz forces for in-plane magnetic fields. Electrons flowing along one sloped face are deflected toward the sensing contact, whereas electrons on the opposite face are pushed toward the pyramid apex. For a z-directed magnetic field, however, the Lorentz forces on both faces act in the same direction, and the resulting contributions cancel by symmetry, thereby suppressing crosstalk in X/Y mode.

The Z mode operates differently. Here, a single pair of supply contacts is positioned at opposite corners of the structure, and the Hall voltage is sensed at the remaining two corners. When the supply is applied, the current divides into two parallel branches, producing a geometry well-suited for detecting the out-of-plane magnetic-field component[36].

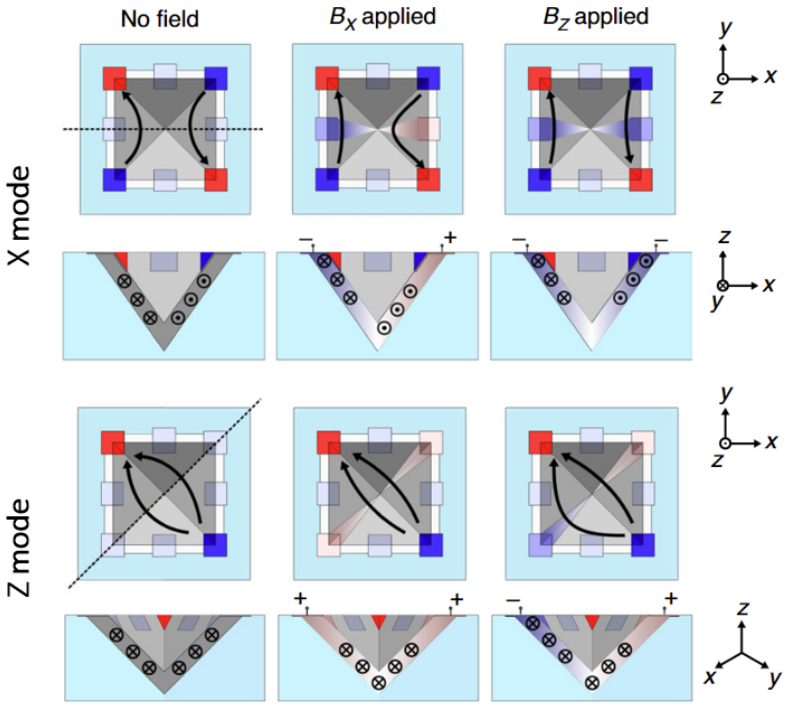


Figure 5.3: The sensing modes in X direction and Z direction. The black arrows represent the carrier velocities and their deflection due to an applied magnetic field.[36]

Four PYRAMID sensors were measured: the plain PYRAMIDs (referred to as the control sensors) and the PYRAMID variants that include the boron pad, P-ring, and ground ring (referred to as the boron sensors in the following discussion). Both sensor types were taken from the high-doped $25\mu m$ and $50\mu m$ wafers. The control sensors serve as the baseline for comparison.

The boron sensors can be configured in two ways: (i) grounding the P+ region, which also effectively grounds the substrate, or (ii) leaving the P+ region floating. With two configurations for each of the two boron sensors, a total of six measurements were conducted. In Table 5.1, the different sensor variants are assigned unique abbreviations to facilitate a clearer and more concise discussion.

Sensor type	Abbreviation
$25\mu m$ control sensor	P1
$25\mu m$ boron sensor grounded	P2
$25\mu m$ boron sensor floating	P3
$50\mu m$ control sensor	P4
$50\mu m$ boron sensor grounded	P5
$50\mu m$ boron sensor floating	P6

Table 5.1: Abbreviation of different sensor types

5.2. Simple electrical measurement

During the electrical characterization of semiconductor layers, Transfer Length Method (TLM) structures, four-point probe structures, and van der Pauw structures of different layers were measured. In addition, the resistance of the low-doped PYRAMID Hall devices was extracted, and these devices were also used as van der Pauw structures to determine the sheet-resistance parameters of the active layer.

5.2.1. Four-point probe structure

In the 4-point-probe structures, a current IFORCE is driven through the I-1 FORCE and I-2 FORCE pads, and the resulting voltage VSENSE is measured across the V-1 SENSE and V-2 SENSE pads. The sheet resistivity can then be extracted using Equation (5.1), where W and L denote the width and length of the 4-point-probe structure, respectively. The use of a 4-point-probe configuration, rather than a conventional 2-point measurement, eliminates the influence of series contact resistances, thereby enabling an accurate determination of the intrinsic sheet resistance.

$$R_{sheet} = R_{meas}\frac{W}{L} = \frac{V_{sense}}{I_{force}}\frac{W}{L} \tag{5.1}$$

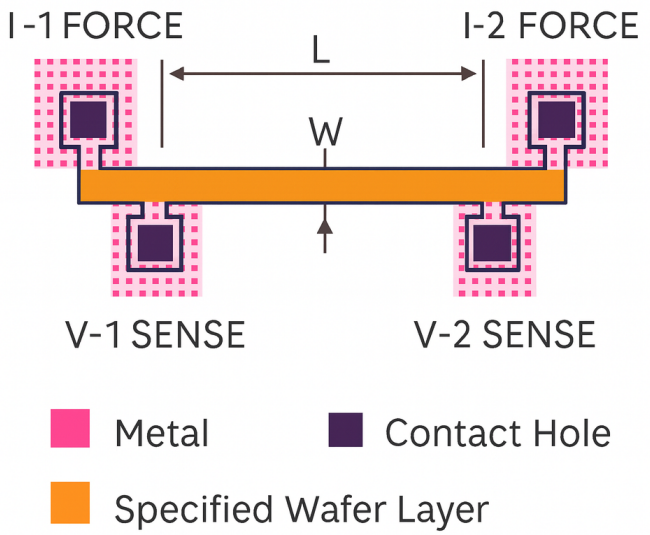


Figure 5.4: Example of a four-point probe structure.

The W/L parameter in this design is 1/16. The sheet resistance extracted from the four-point probe measurements is summarized in Table 5.2. As discussed in Chapter 4, two annealing steps follow the N^+ implantation. The first annealing step activates the dopants and forms a high-quality thermal oxide at the Si–SiO2 interface. The second annealing step densifies the PECVD TEOS oxide. In the table,

“0 Annealing” indicates that no annealing was applied to the sample; “1 Annealing” refers to samples that underwent only the first annealing step; and “2 Annealing” denotes samples that completed both annealing steps.

For the entries labeled “1min P+”, the time corresponds to the boron deposition duration. Thus, “1min P+” represents a P+ layer formed by one minute of boron deposition, and the same interpretation applies for the remaining deposition times.

5.2.2. Van der Pauw structure

The Van der Pauw structures consist of four contacts (VDP-1, VDP-2, VDP-3, and VDP-4, arranged in a counter-clockwise order) positioned symmetrically around an octagonal active area. The sheet resistance is extracted by forcing a current between two adjacent terminals and measuring the resulting voltage across the opposite pair of terminals. By repeating the measurement for the two orthogonal current–voltage configurations and applying Equation (5.2), the sheet resistance of the material can be accurately determined.

$$R_{sheet} = R_{meas}\frac{\pi}{ln2} = \frac{V_{12}}{I_{43}}\frac{\pi}{ln2} \tag{5.2}$$

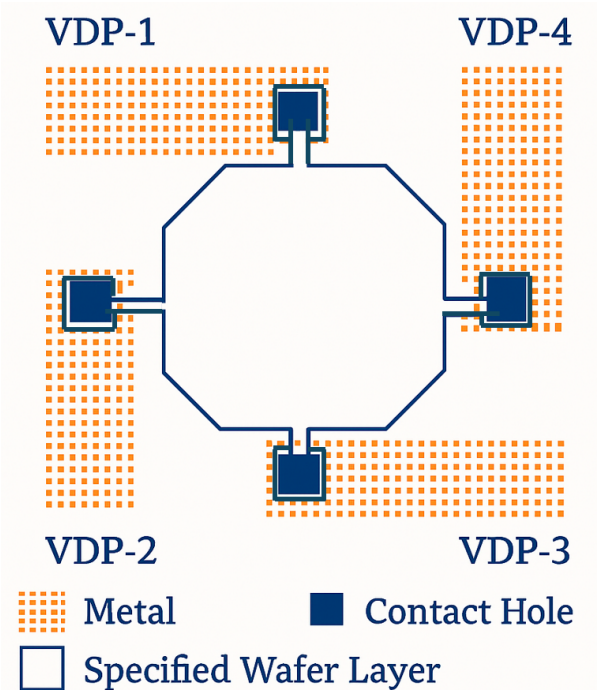


Figure 5.5: Example of a Van der Pauw structure.

The sheet resistance extracted from the Van der Pauw measurements is summarized in Table 5.2. The PYRAMID structures were also characterized using the Van der Pauw method, as they effectively function as plate-like geometries. Consequently, the sheet resistance of the low-doped PYRAMID devices was also evaluated and included in the analysis.

5.2.3. Transfer Length Method structure

The Transfer Length Method (TLM) structure consists of a bar of semiconductor layer and a series of equally wide semiconductor bars contacted by multiple metal pads with systematically varied inter-pad spacings. Each pair of adjacent metal pads forms a two-terminal resistor whose total measured resistance includes contributions from both the sheet resistance of the semiconductor layer and the metal–semiconductor contact resistance, as shown in Equation 5.4, where R_c denotes contact resistance, d is the distance between two adjacent contact pads and W is the width of the contact pad. By

measuring the resistance between multiple pad pairs with different spacings and fitting the resulting values to a linear model, the contact resistance and sheet resistance can be simultaneously extracted.

$$R_{\text{total}}(L) = 2R_c + R_{\text{sheet}}\frac{d}{W} \tag{5.3}$$

If the linear TLM model is plotted, the resulting graph should resemble the plot shown in Figure 5.6. The contact resistance is obtained from the y-intercept of the fitted line, where the intercept corresponds to $2R_c$; therefore, the contact resistance is given by the y-intercept divided by 2. The sheet resistance is extracted from the slope of the fitted line, which equals R_S/Z. Multiplying the slope by the width of the contact pad yields the sheet resistance of the semiconductor layer.

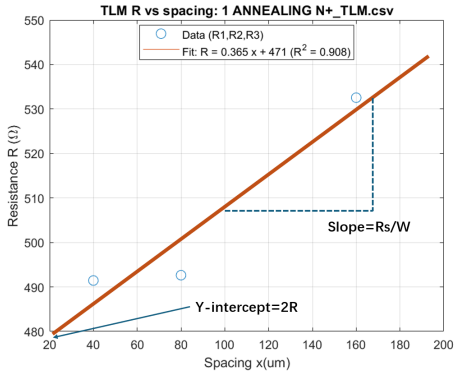


Figure 5.6: Example curve of the transfer length method.

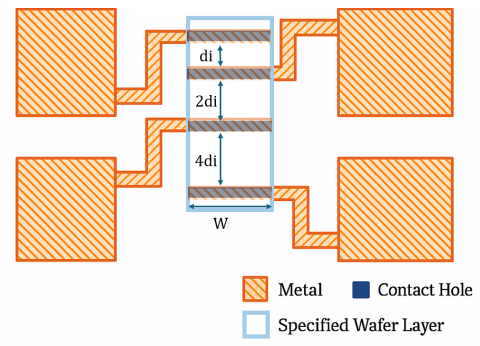


Figure 5.7: Example of a TLM structure.

The measurement results of the TLM structures are listed in Table 5.2. Only two reliable datasets are reported, as most measurements did not yield consistent results. In principle, the sheet resistance and contact resistance are obtained by fitting the measured resistance values to a linear model as a function of the pad spacing. However, in most cases of this thesis, the measured resistance did not exhibit a clear linear dependence on the spacing, preventing a meaningful extraction of these parameters. In Figure5.8, one example of non-linear data is presented. This deviation is likely caused by an operational mistake during the measurement procedure.

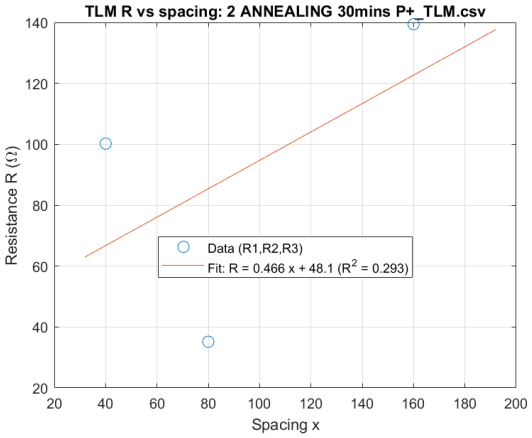


Figure 5.8: Example of a non-linear result from TLM measurement due to measurement mistake.

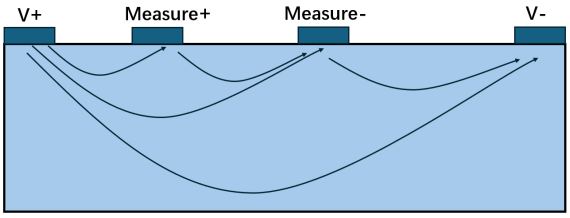


Figure 5.9: Graphical explanation for measurement error. The V+ and V- are the biasing, Measure+ and Measure- are the measuring contacts. The black arrow lines represent the current flow.

During the measurement, the current was inadvertently forced between the two outermost contacts of the TLM structure, while the voltage was measured using the intermediate pads. Instead of biasing and sensing across each adjacent contact pair, as required by the TLM method, this configuration spans multiple pad spacings simultaneously. Such an arrangement introduces a systematic error because the measured voltage drop no longer corresponds to the resistance of a single, well-defined pad pair. The resulting non-uniform current spreading, inclusion of unintended series resistances, and distortion of the local potential distribution violate the assumptions of the linear TLM model, as shown in Figure5.9. Consequently, both the sheet resistance and the contact resistance extracted from these measurements become inaccurate and cannot be reliably interpreted.

Layer type	4PP R_{sheet} ($\Omega/\square$)	VDP R_{sheet} ($\Omega/\square$)	TLM R_{sheet} ($\Omega/\square$)	TLM $R_{\text{contact}}(\Omega)$
0 Annealing N+	1349.8	-	7854	888
1 Annealing N+	92.56	115.6	-	-
2 Annealing N+	82.9	95.8	123.7	111.8
1min P+	6.2	128	-	-
10mins P+	6.2	130	-	-
30mins P+	6.1	127	-	-
1h P+	5.9	125.6	-	-
25um Boron Pyramid low-doped n-well	-	640	-	-
25um Pyramid low-doped n-well	-	372	-	-
50um Boron Pyramid low-doped n-well	-	1825	-	-
50um Pyramid low-doped n-well	-	1693	-	-
25um Boron Hall plate low-doped n-well	-	0.46	-	-
25um Hall plate low-doped n-well	-	0.57	-	-

Table 5.2: Van der Pauw structure, Four-point probe, TLM sheet resistance, and TLM contact resistance result. “0 Annealing” indicates that no annealing was applied to the sample; “1 Annealing” refers to samples that underwent only the first annealing step; and “2 Annealing” denotes samples that completed both annealing steps. Label “1min P+” corresponds to the boron deposition duration, and the same interpretation applies for the remaining deposition times.

5.2.4. Resistance measurement

The resistance of the PYRAMID’s active layer can be determined by applying a voltage between two contacts and measuring the resulting current. The resistivity of the material is given by

$$\rho = 1/\sigma = 1/e(\mu_n n + \mu_p p) \tag{5.4}$$

Where ρ is the resistivity, σ is the conductivity, $\mu_{n,p}$ is the mobility of electrons and holes, and n, p are the electron and hole densities. Increasing the doping concentration raises the carrier density, which tends to reduce the resistivity. At the same time, higher doping degrades the carrier mobility due to increased impurity scattering. At low to moderate doping levels, the increase in carrier concentration dominates over the reduction in mobility, resulting in a decrease in resistivity with increasing doping. At very high doping levels, however, the mobility degradation becomes strong enough to compensate for the higher carrier concentration, so the resistivity no longer decreases as rapidly and may eventually approach a nearly constant value.

This resistivity–concentration relationship enables the use of the measured resistance of a layer with a known geometry as an indicator of its doping level when comparing with the layer resistance of other samples. The resistance measured between diagonal terminals in [36] is used for benchmarking, as this project uses a very similar fabrication flow to [36]. Additionally, knowing the resistance provides useful estimates of the expected thermal noise and helps determine the appropriate biasing conditions for the device. The IV curve of resistance measurement on 50 µm boron sensor is shown in Figure 5.10. The linear curve means that the contact is on an n-region.

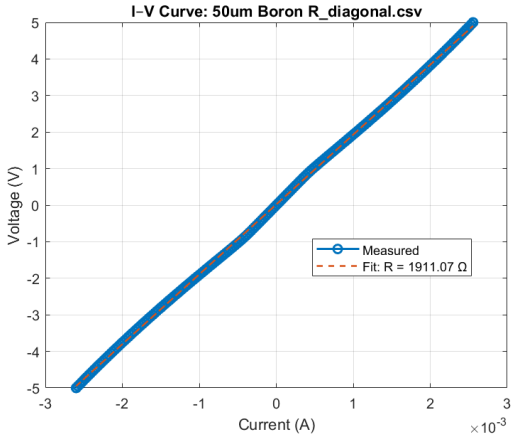


Figure 5.10: The IV curve of resistance measurement.

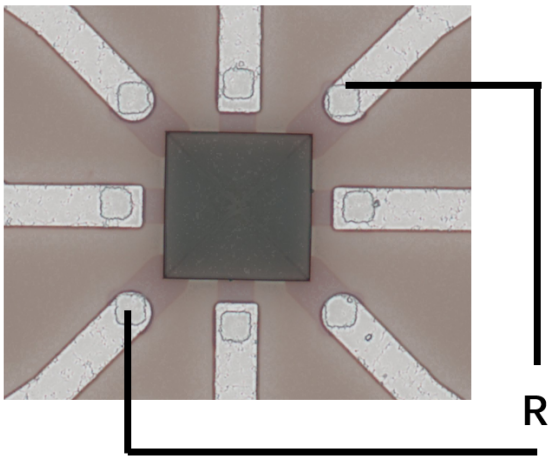


Figure 5.11: The diagonal terminals are used to measure resistance.

The resistance measurements were performed between the two diagonal terminals of the PYRAMID structures, corresponding to the configuration that yields the maximum two-terminal resistance, as shown in Figure 5.11. The extracted resistance values are summarized in Table 5.3. Measurements were conducted for both the high-doped and low-doped PYRAMID devices. For verification and debugging purposes, the flat devices—which effectively function as simple Hall plates—were also characterized.

Device type	Resistance ($k\Omega$)
Low-doped 25μm control	0.57
Low-doped 25μm boron	0.75
Low-doped 50μm control	1.91
Low-doped 50μm boron	1.74
High-doped 25μm control	1.47
High-doped 25μm boron	1.68
High-doped 50μm control	1.31
High-doped 50μm boron	1.42
Low-doped 25μm flat control	10.8
Low-doped 25μm flat boron	10.7
PYRAMID 25μm from [36]	9.3

Table 5.3: Resistance measured between the diagonal ports of the devices

5.2.5. Discussion

In the four-point probe measurement, the N+ layer results show that the 1-annealing N+ sheet resistance is significantly reduced compared to the 0-annealing N+ sample, indicating that the first annealing process activates most of the dopant. The P+ sheet-resistance values are relatively similar. Because the P+ test structures are fabricated directly on top of the p-type substrate, they are not electrically isolated. Therefore, during the measurement, the extracted resistance corresponds to the P+ layer in parallel with the resistance of the p-type substrate. As a result, the measured sheet resistance should be lower than the sheet resistance of P and P+ layer, assuming a uniform doping level.

A rough estimation can be made from the relationship between mobility and boron concentration [21]:

$$\mu_p = \mu_0 e^{-p_c/p} + \frac{\mu_{max}}{1 + (p/C_r)^\alpha} - \frac{\mu_1}{1 + (C_s/p)^\beta} \tag{5.5}$$

where $\mu_0 = 44.9\, cm^2/(V \cdot s)$, $\mu_{max} = 470.5$, $\mu_1 = 29$, $C_r = 2.23 \times 10^{17}$, $C_s = 6.1 \times 10^{20}$, $\alpha = 0.7$, $\beta = 2$, $p_c = 9.23 \times 10^{16}$ are empirical values. From simulation and from the literature [40], a peak boron concentration of $1 \times 10^{21} cm^{-3}$ is estimated. The corresponding mobility is $24.766(cm^2/(V \cdot s))$. Using Equation 5.4, a resistivity of $0.000252(\Omega \cdot cm)$ is obtained. With a simple conversion from resistivity to sheet resistance,

$$R_s = \frac{\rho}{t} \tag{5.6}$$

and assuming a junction depth of about 100 nm, a sheet resistance of $0.03\Omega/\square$ is roughly estimated. Performing the same calculation for the substrate doping yields 19.05–$95.24\Omega/\square$ for a bulk resistivity range of 1–$5(\Omega \cdot cm)$. Therefore, measured sheet resistance close to $0.03\Omega/\square$ is reasonable after considering the two layers in parallel. The measured four-point-probe sheet resistance is around $6\Omega/\square$ for different P+ layers. These values are not within the reasonable parallel-combination range.

The VDP measurement results show similar sheet resistance values for the N+ layers. However, the P+ values deviate significantly from the four-point-probe results. This discrepancy is likely caused by the fact that the P+ test structures are not electrically isolated; current spreads both vertically and laterally during the measurement, and part of the current flows outside the intended region, distorting the VDP measurement result.

The sheet resistance extracted from the PYRAMIDs can be used to roughly estimate the doping concentration using the previous equations. However, the calculated doping levels are unrealistically low (around $\approx 10^{13} cm^{-3}$), indicating that the measurement is incorrect. The likely reason is the large, invasive contacts on the PYRAMIDs, which distort the current distribution and violate the assumptions of the extraction formulas.

Most TLM measurements also failed because the measurement procedure was incorrect. The TLM structure was biased only at the first and last contacts, while voltages were measured at the intermediate pads. Under this configuration, the current distribution becomes nonuniform, and the voltage at each intermediate pad reflects the entire distributed network of sheet resistances and contact resistances. As a result, the extracted resistance no longer corresponds to a single spacing, the R–d plot becomes nonlinear, and the extracted sheet and contact resistances are incorrect.

The resistance measured on the PYRAMIDs (around $1.5k\Omega$) is generally much smaller than the benchmarking value ($9.3k\Omega$) , about three times lower than the estimation. The low-doped flat samples yield more reasonable values, suggesting that the issue is not related to doping or annealing but is instead linked to spray coating, which is the major fabrication difference between the PYRAMID wafers and the flat wafers. The boron-doped samples generally show higher resistance than the control samples, which is due to their lower carrier concentration and slightly narrower active channel.

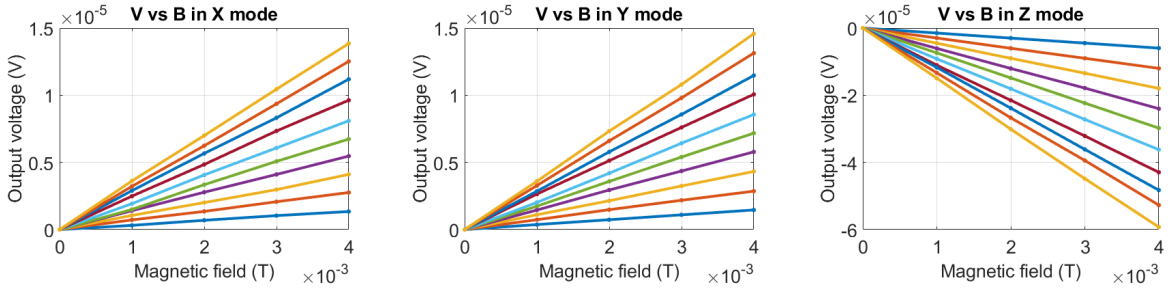
5.3. Hall-effect measurement

In this section, three key aspects of sensor performance are evaluated. Sensitivity determines the sensor's ability to convert an input magnetic field into an electrical output. Residual offset limits the achievable resolution. Noise affects the sensor's accuracy, particularly at low magnetic field levels.

5.3.1. Sensitivity

To measure the sensitivity of the sensors, they were placed inside a 3D coil capable of generating magnetic fields in all spatial directions, allowing for the automated characterization of each device. A constant current was sourced through the input contacts while the corresponding supply voltage was recorded. The Hall voltage was measured at the output contacts. The eight current-spinning phases were controlled using a switching relay matrix, and each spinning sequence was repeated ten times to obtain a more stable averaged output voltage. The magnetic field was swept from 0 to 4 mT in four steps for each supply current.

The magnetic-field sensitivity was extracted at three supply-current levels: $100\mu A$, $550\mu A$, and $1mA$. In addition, one dedicated measurement set employed ten equidistant supply currents in the range of 100–$1000\mu A$ on sensor P2. For this set, the Hall voltage as a function of the magnetic field is shown in Figure 5.12. The output voltages appear approximately equidistant at each magnetic field value, demonstrating that the device exhibits a linear response to the applied magnetic field.



(a) X mode. **(b)** Y mode. **(c)** Z mode.

Figure 5.12: Processed output voltage in function of the magnetic field, for ten equidistant supply currents in the range 100-1000μA, measured for the sensor P2.

In general, the sensitivity of Hall-effect devices is reported normalized to the supply current or supply voltage. Two metrics, known as current-related sensitivity (S_I) and voltage-related sensitivity (S_V), are defined as the sensitivity normalized by the supply current or voltage, respectively. These metrics should maintain a constant value throughout the biasing range, with little fluctuation caused by noise.

The sensitivity measurements were repeated for the six samples. The results are reported in the table 5.4.

Sensor	Mode	S_I (V/AT)	S_V (mV/VT)
P1 ($25\mu m$ control)	X	2.85	3.7
	Y	2.92	3.74
	Z	-11.9	-8.04
P2 ($25\mu m$ boron grounded)	X	3.417	3.86
	Y	3.646	4.117
	Z	-14.96	-8.91
P3 ($25\mu m$ boron floating)	X	3.429	3.87
	Y	3.7	4.2
	Z	-15.04	-8.95
P4 ($50\mu m$ control)	X	2.932	3.73
	Y	2.99	3.84
	Z	-2.55	-1.82
P5 ($50\mu m$ boron grounded)	X	2.826	3.62
	Y	3.085	4
	Z	-2.273	-1.62
P6 ($50\mu m$ boron floating)	X	3.054	3.92
	Y	3.19	4.15
	Z	-2.16	-1.54

Table 5.4: Sensitivity measurement results

The measured sensitivities are very low, approximately an order of magnitude below the design expectations, and they show little variation across the different samples. Overall, the boron sensors tend to exhibit slightly higher sensitivities than the control sensors. This is consistent with expectations, since the deposited P+ layer narrows the active region and reduces its peak doping concentration, which can increase carrier mobility and thereby enhance the voltage-related sensitivity. However, the negative sign observed in the Z-mode sensitivity is unexpected and suggests a possible issue in the fabrication process.

Although the exact cause of the observed sign flip is not yet fully determined, several plausible explanations can be considered. First, the sign inversion cannot be attributed to the presence of the p+ layer or to a reversal of the dominant charge carriers, since the phenomenon occurs in both the boron-diffused samples and the control samples. This rules out carrier-type inversion as the source of the effect. Therefore, the most reasonable hypothesis is that the sign flip originates from an unintended alteration of the current distribution within the device rather than from changes in the semiconductor polarity.

If an unintended N+ short is present along the periphery of the Hall plate, the injected current no longer flows through the intended N-well region when operating in Z mode. Because the N+ layer exhibits a much lower sheet resistance than the lightly doped N-well, a significant portion of the bias current is diverted into this low-resistance edge path. As the current re-enters the N-well only near the contact points, the resulting current distribution becomes highly non-uniform, and the local direction of current flow near the sensing contacts can deviate substantially from the ideal diagonal path. Since the Hall voltage is proportional to the product of the current density vector and the magnetic field vector, any distortion in the current direction may directly affect the sign of the measured signal.

5.3.2. Residual offset

A proper test setup for residual offset measurements requires external stray magnetic fields to be well below the electrical measurement limits. To eliminate such contributions, the sensors were placed

inside a zero-gauss chamber during testing. The sensor offset was reduced using the current-spinning technique. A constant current is sourced through the input contacts, and the corresponding supply voltage is recorded. The Hall voltage is measured at the output contacts. The eight current-spinning phases are implemented using a switching relay matrix.

For each mode, the output voltage is measured over the eight phases of the spinning sequence and then averaged; the remaining value after this averaging represents the residual offset. This measurement is performed for ten equidistant supply currents in the range of 100–$1000\mu A$ and repeated for all devices.

Current spinning effectively suppresses offset contributions originating from external thermal gradients and geometric asymmetries. However, it cannot cancel second-order terms, like junction field effect, velocity saturation, and self-heating, which ultimately determine the residual offset. For a fair comparison across devices and technologies, the residual offset is typically normalized by the sensitivity and expressed in magnetic-field units. The absolute value of equivalent magnetic residual offsets is shown in Figure 5.13. Their values vary stochastically depending on the detection mode and the supply current. However, it appears that the boron sensors with a floating P+ region exhibit the lowest residual offsets in both PYRAMID geometries. In contrast, the control sensors tend to show the highest residual offsets in most cases. This behavior may be attributed to the fact that the residual offset originates from second-order terms, which may be partially suppressed by the presence of the buried structure.

At a supply voltage of 0.5 V—the typical operating voltage for VHDs—the measured residual offset spans from approximately 20 μT and 1.5mT. Another noteworthy observation is that grounding the substrate or the P+ layer does not reduce the residual offset. In fact, it appears to slightly increase the residual offset in some cases.

This behaviour is reasonable in light of the junction field effect. When the p+ region is grounded, the junction experiences the full reverse bias of approximately 0.5 V. In contrast, when the p+ region is left floating, its potential can shift slightly through capacitive coupling and leakage, effectively reducing the reverse-bias voltage across the junction. A lower reverse bias weakens the modulation of the active n-region by the surrounding p+ layer. As a result, the JFE influence is reduced, leading to a smaller residual offset in the floating-p+ configuration.

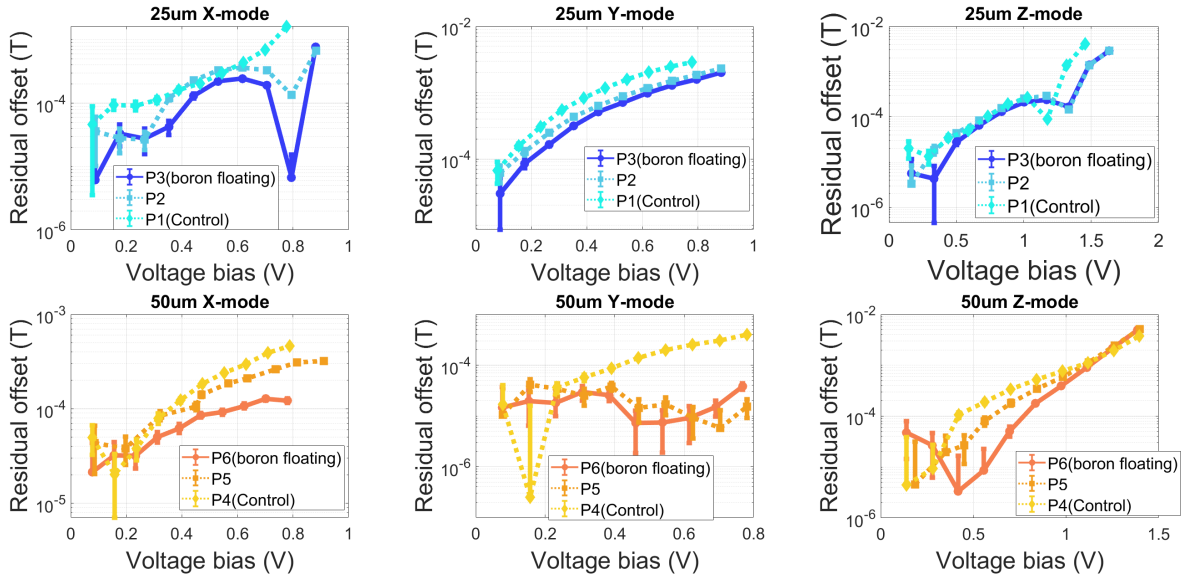


Figure 5.13: The absolute value of equivalent magnetic residual offsets for the three sensors of 25μm and 50μm PYRAMIDs reported in log scale. The dotted lines are the boron sensor grounded and the control sensor, respectively. The continuous line is the boron sensor floating.

5.3.3. Noise

Noise performance is also evaluated for the Hall sensors, as it determines both the detection limit and the achievable resolution. The devices were biased using a simple voltage divider consisting of an alka-

line battery and foil resistors (50–100 Ω). Ceramic capacitors implemented in a low-pass configuration provided a stable DC bias. The sensor outputs were AC-coupled to a low-noise amplifier with a gain of 200, and the amplified signals were then fed into a spectrum analyzer for noise characterization.

The noise spectral density was characterized in modes X and Z, where modes X and Y share a semi-identical configuration. Measurements were performed at three supply voltages—1.24 V, 0.62 V, and 0.31 V. Additionally, the intrinsic thermal noise floor of the device was measured by applying no supply voltage.

Because the sensors are first connected to a low-noise amplifier (LNA) before the spectrum analyzer, the noise contribution of the LNA must be taken into account. To quantify this, a special calibration measurement was carried out by shorting the input terminals of the LNA. This setup captures the instrument noise, which can later be subtracted during sensor noise post-processing. All measurements were conducted over a frequency range from 10 Hz to 100 kHz.

For mode X, noise spectral density was measured for two orthogonal spinning phases—phase 1 and phase 3. For mode Z, only one phase (phase 1) was measured, because this mode uses only the corner contacts, resulting in symmetry between orthogonal spinning phases. Figure 5.15 displays the resulting noise characterization.

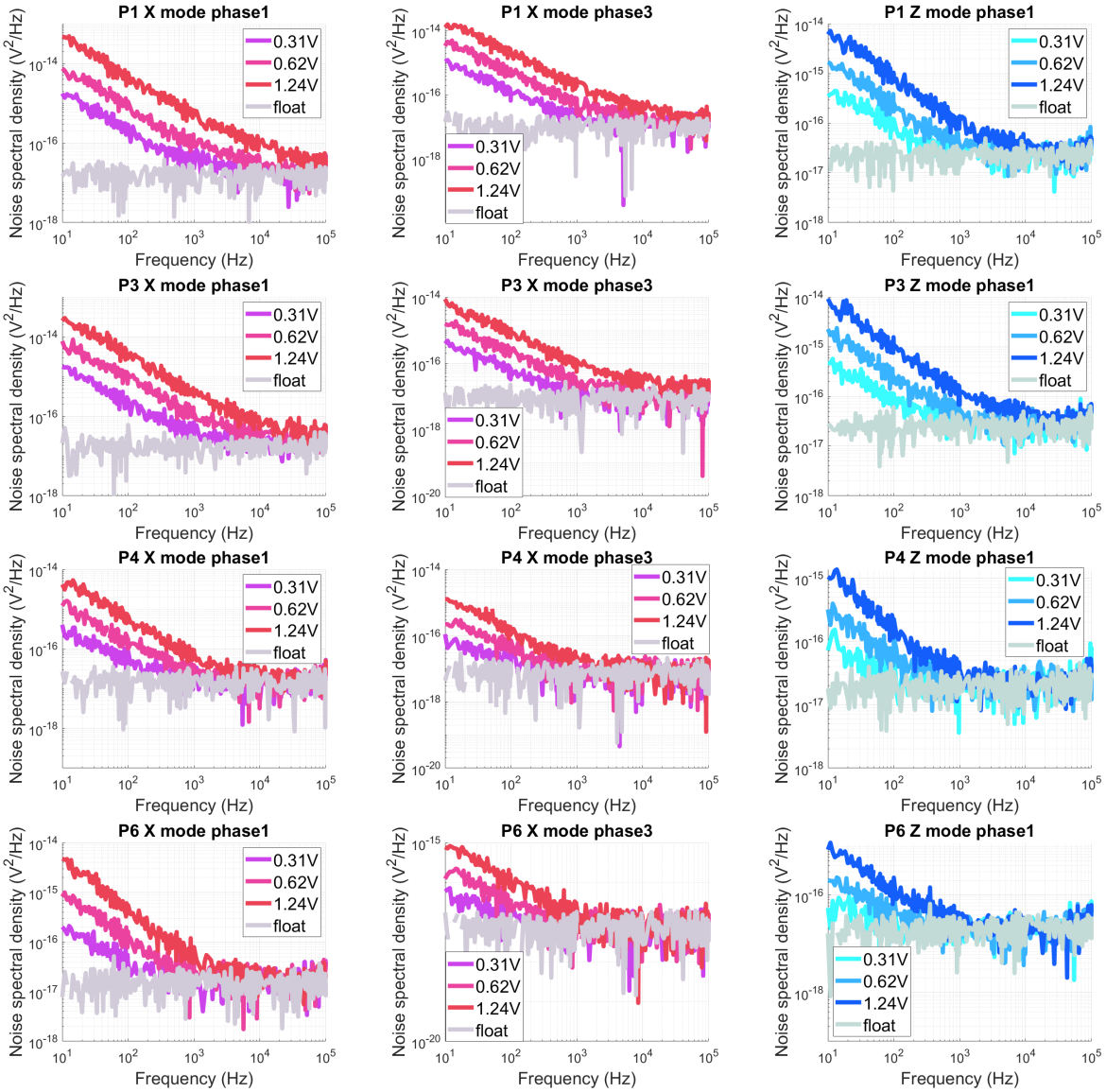


Figure 5.14: Noise power spectral densities (PSD) of sensor P1, P3, P4 and P6. The noise is measured for three supply voltages at room temperature, in two orthogonal phases, for modes X and Z. The dotted lines are the performance of control sensors and the continuous lines are the performance of boron sensors.

The noise performance of the control sensors and the boron sensors is also compared in Figure 5.15. The y-axis is plotted on a linear scale to facilitate a clearer visual comparison between the curves. Overall, it can be observed that sensors originating from the same die generally exhibit lower noise for boron sensors than for control sensors. In a few cases, the performance of the two sensor types is comparable.

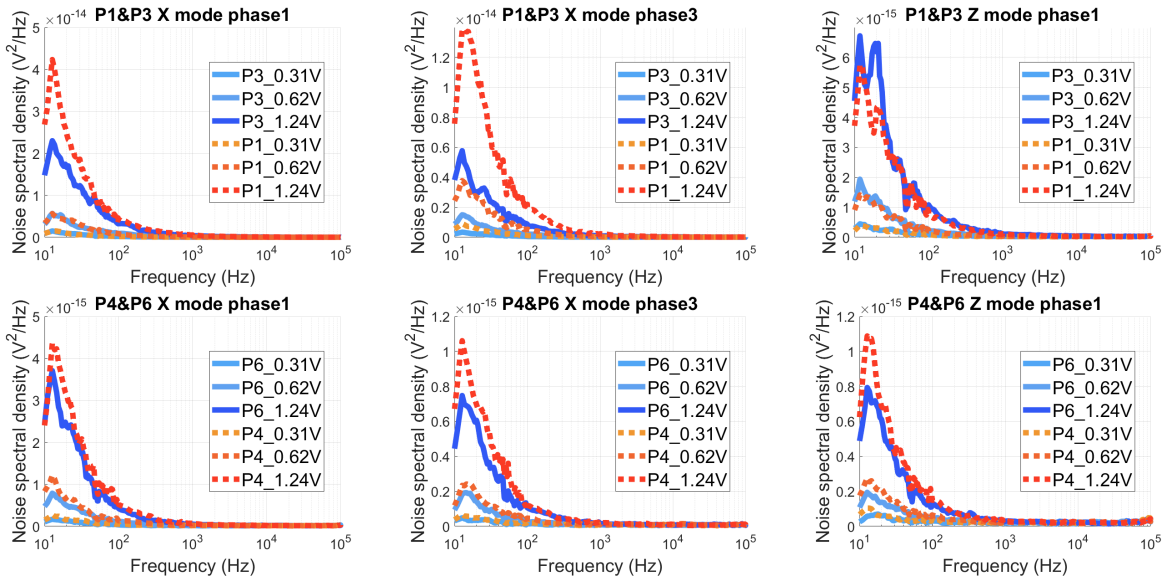


Figure 5.15: Filtered Noise PSD of sensor P1, P3, P4 and P6. The noise is measured for three supply voltages at room temperature, in two orthogonal phases, for modes X and Z. The results are compared between P1 vs. P3, P4 vs. P6.

The sensor noise is generally dominated by two primary components: flicker noise and thermal noise. In this work, flicker noise is of particular interest because it is strongly influenced by traps and defects at the semiconductor interface—precisely the mechanisms that the buried-layer structure in this design aims to mitigate.

γ and α_H/N of the Hooge model were extracted for a supply voltage of 1.26 V by fitting the $1/f$ portion of the spectrum with the Hooge model. The corresponding corner frequencies and thermal noise floors were then derived. These parameters are reported in Table 5.5.

Sensor	Mode & Phase	α_H/N	γ	Corner freq. (kHz)	Thermal noise floor (nV/$\sqrt{Hz}$)	Equivalent magnetic thermal noise floor (uT/$\sqrt{Hz}$)
P1	**X** mode phase 1	2.388x10^{-13}	0.951	21.86	5.312	1.139
	X mode phase 3	1.217x10^{-13}	0.992	12.23	4.119	0.884
	Z mode phase 1	5.968x10^{-14}	1.030	2.678	5.279	0.521
P3	**X** mode phase 1	2.109x10^{-13}	0.986	11.53	5.745	1.181
	X mode phase 3	4.116x10^{-14}	0.957	5.025	4.338	0.892
	Z mode phase 1	4.858x10^{-14}	0.961	3.076	5.846	0.518
P4	**X** mode phase 1	4.356x10^{-14}	1.050	2.601	4.239	0.902
	X mode phase 3	1.292x10^{-14}	1.098	1.189	2.939	0.625
	Z mode phase 1	1.137x10^{-14}	1.079	0.4723	4.853	2.116
P6	**X** mode phase 1	3.889x10^{-14}	1.082	1.936	4.147	0.840
	X mode phase 3	7.552x10^{-15}	1.047	1.009	2.926	0.592
	Z mode phase 1	8.208x10^{-15}	1.061	0.3996	4.754	2.45

Table 5.5: Power spectra fitted parameters at 1.31 V. The equivalent magnetic thermal noise floor is calculated at a biasing voltage of 1.26V.

5.3.4. Discussion

Based on the measurement results, the sensors exhibit two pronounced issues: low sensitivity and an inverted Z-axis magnetic-field response. According to the expression for the voltage-related sensitivity, the sensitivity scales with the Hall mobility. Since the Hall mobility decreases with increasing doping concentration, a reduced sensitivity is typically associated with an effectively higher doping level in the region carrying the current. In the present context, this reduction in sensitivity may originate from two possible mechanisms: (i) the annealing step may not have driven the dopants sufficiently deep into the substrate, leaving a highly doped active region near the surface, or (ii) unintended N+ regions may be shorting the Hall plate, diverting most of the lateral current from the N-well into a highly doped, low-mobility parasitic path.

To further investigate the origin of the problem, a quick Hall measurement was performed on the flat reference sensors. A Z-direction magnetic field was applied, and two repetitions of the eight-phase spinning sequence were used to extract the averaged output voltage of the $25\mu m$ low-doped Hall plate. The resulting current-related sensitivity was approximately 200V/A/T in Z-mode, which is more than twenty times higher than that of the PYRAMID sensors fabricated with identical nominal doping and device dimensions. Moreover, resistance measurements show that the flat devices exhibit substantially higher resistance compared to the PYRAMID devices. These observations strongly indicate that the intrinsic material quality and doping are adequate, and that the degraded performance originates from fabrication differences between the PYRAMID wafers and the flat samples. The most significant differences between the two process flows lie in the use of spray coating for the PYRAMID wafers, the additional processing steps required to etch the pyramid cavities, and the fact that the implanted n-type layer remained at room temperature for approximately 1.5 years before undergoing annealing.

A plausible hypothesis is that the reduced resist coverage near the PYRAMID corners during spray coating allowed arsenic ions to penetrate the resist during implantation. This would create a narrow N+ ring along the upper edge of the cavity. Such a parasitic N+ ring constitutes a low-resistance shunt around the Hall plate: the majority of the current flows through this highly doped region, which exhibits very low Hall mobility and therefore a weak magnetic response due to its small lateral width. Only a minority of the current flows through the intended N-well region and contributes to magnetic-field detection. Furthermore, the presence of the N+ ring alters the lateral current distribution within the Hall plate, which can explain the observed inversion of the Z-direction Hall polarity. The distorted current pathways effectively modify the potential distribution sampled at the sensing contacts, thereby reversing the measured Hall sign under Z-axis excitation.

In the sensitivity measurements, the sensors incorporating the buried structure do not exhibit a measurable improvement compared to the control devices. This outcome is likely attributed to the very shallow P+ implantation, which produces only a limited modification to the active-region doping profile and therefore has minimal influence on the carrier mobility within the Hall plate. An additional observation from the sensitivity data is that grounding the substrate (P+) slightly reduces the sensitivity. This reduction may originate from additional vertical current leakage paths or increased capacitive coupling through the grounded substrate.

In the residual offset measurements, the control sensors of both device dimensions generally exhibit a higher residual offset. However, the offset–voltage curves do not follow a consistent trend across devices, indicating that multiple error mechanisms contribute simultaneously. For example, in the $50\mu m$ Z-mode comparison shown in Fig. 5.13, the control and boron-doped sensors show comparable residual offsets, which is reasonable because the buried structure does not mitigate thermal gradients. In other cases, however, the boron-doped sensors demonstrate lower residual offsets, possibly due to the suppression of charge-fluctuation–induced offset components, arising from longer carrier recombination times in the boron-modified buried layer. Nonetheless, variations in residual offset performance may also arise from device-to-device fabrication differences, especially given the process complexities involved in the PYRAMID structures.

In the noise measurements, all sensors exhibit a similar thermal-noise floor, consistent with the expected values derived from their measured resistances. From the noise spectra and the extracted Hooge-model parameters, it is evident that the flicker noise level is lower in the boron-doped sensors across both device dimensions. Specifically, the boron-modified devices show a reduced α_H/N parameter and a lower corner frequency. These improvements suggest that the buried structure introduced

by boron doping may have the ability to partly block fluctuating charges—likely in the passivation or surface-adjacent layers—thereby suppressing 1/f noise and improving overall low-frequency performance.

To sum up, the key findings from the measurement are:

- The pyramid sensors exhibit low sensitivity and an inverted Z-axis response due to the shorting of contacts, redirecting current distribution.
- Grounding the substrate slightly reduces sensitivity and increases residual offset due to the increased JFE.
- Buried pyramid sensors show lower residual offset, potentially due to charge-fluctuation suppression from the buried layer.
- The buried pyramid sensors show lower flicker-noise levels, indicating that the boron-induced buried structure may help block or screen fluctuating charges, improving low-frequency noise performance.

6

Conclusion and future work

The results of this thesis can be summarized as follows:

- The limitations of conventional inverted-pyramid Hall sensors—namely high residual offset and pronounced flicker noise—were identified and analyzed. The underlying cause was explained. To address this, an existing approach based on a buried sensing structure was reviewed, motivating the integration of such a structure into the pyramid sensor. Pure-boron LPCVD was selected as the doping method due to its cost-effectiveness and its ability to form ultra-shallow junctions.
- Process simulations using Synopsys Sentaurus were conducted to model the boron deposition and diffusion mechanism during LPCVD. To evaluate and minimize the junction field effect, a simplified 2D device structure was simulated to study how the applied bias modulates the active conduction channel.
- Buried inverted-pyramid Hall sensors were successfully fabricated in two layout sizes and for two doping conditions. Flat buried Hall plates were also fabricated for layer characterization and process debugging. The lithographic masks were designed, and a complete fabrication process flow was developed. Furthermore, the boron LPCVD process was experimentally investigated for different deposition durations.
- The fabricated samples were characterized in all three sensing modes. Although certain issues, such as reduced sensitivities and altered detection directions, were observed, the buried pyramid Hall sensors functioned correctly and demonstrated significantly improved residual offset and flicker noise performance compared with their conventional counterparts.

However, further work is necessary to fully exploit the potential of the buried pyramid Hall sensor. To be specific:

- A more accurate simulation framework is required. The current boron diffusion model neglects the chemisorption phase of the LPCVD boron deposition mechanism, leading to overly aggressive diffusion and an overestimation of the junction depth. A more suitable process model—capable of capturing both chemisorption and subsequent diffusion—should be implemented to provide reliable doping profiles and guide fabrication process optimization.
- Additional p-type doping techniques should be explored, including ion implantation and epitaxial deposition, to evaluate alternative approaches for achieving controlled ultra-shallow junctions.
- The suspected cause of the reduced sensitivity is the unintended shorting of the n+ contacts, likely originating from insufficient photoresist coverage at the pyramid edges. The resist-coating procedure should be revisited, and detailed inspection using tools such as SEM is recommended to ensure proper edge coverage.
- Further simulations in COMSOL are needed to investigate the observed sign inversion in the Z-field response. One hypothesis attributes this behavior to altered current distribution caused by partial contact shorting; however, identifying the specific topology responsible for a sign change

only in the Z-mode is challenging. Sensitivity simulations in COMSOL would greatly aid in diagnosing and validating these effects.

- The relationship between junction depth, flicker noise, and residual offset reduction can be quantitatively modeled. The junction depth used in fabrication was likely overestimated; resistance measurements show minimal difference between the control and boron-doped pyramids, indicating a very shallow junction. The linear sensitivity characteristics further suggest a negligible junction field effect. This implies that deeper junctions may still offer improved flicker noise and offset performance without introducing significant JFET interference, providing a possible margin for future optimization.
- Finally, although the boron-doped sensors show a clear trend of reduced flicker noise and residual offset, the current dataset is too limited to draw statistically robust conclusions. Variations due to fabrication spread and device-to-device randomness remain possible explanations. Characterizing a larger number of samples is essential to substantiate the observed improvements.

References

[1] Mirjana BANJEvIC. “High bandwidth CMOS magnetic sensors based on the miniaturized circular vertical Hall device”. In: *École Polytechnique Fédérale de Lausanne, Lausanne* (2011).

[2] Sandra Bellekom. “CMOS versus bipolar Hall plates regarding offset correction”. In: *Sensors and Actuators A: Physical* 76.1-3 (1999), pp. 178–182.

[3] Bodycote. *What Is Ion Implantation?* 2025. url: `https://blog.bodycote.com/specialist-technologies/what-is-ion-implantation/` (visited on 12/07/2025).

[4] K Cho et al. “Channeling effect for low energy ion implantation in Si”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 7 (1985), pp. 265–272.

[5] Marina Díaz-Michelena. “Small magnetic sensors for space applications”. In: *Sensors* 9.4 (2009), pp. 2271–2288.

[6] Karen M Dowling et al. “Low offset and noise in high biased GaN 2DEG Hall-effect plates investigated with infrared microscopy”. In: *Journal of Microelectromechanical Systems* 29.5 (2020), pp. 669–676.

[7] Karen M Dowling et al. “Micro-tesla offset in thermally stable AlGaN/GaN 2DEG Hall plates using current spinning”. In: *IEEE Sensors Letters* 3.3 (2019), pp. 1–4.

[8] David J Elliott. “Positive photoresists as ion implantation masks”. In: *Developments in Semiconductor Microlithography IV*. Vol. 174. SPIE. 1979, pp. 153–172.

[9] DM Fleetwood et al. “Effects of oxide traps, interface traps, and ‘‘border traps’’on metal-oxide-semiconductor devices”. In: *Journal of applied physics* 73.10 (1993), pp. 5058–5074.

[10] Martin Haartman and Mikael Östling. *Low-frequency noise in advanced MOS devices*. Springer Science & Business Media, 2007.

[11] EH Hall. “On a new action of the magnet on electric currents”. In: *American Journal of Science* 3.111 (1880), pp. 200–205.

[12] Hadi Heidari et al. “CMOS vertical Hall magnetic sensors on flexible substrate”. In: *IEEE Sensors Journal* 16.24 (2016), pp. 8736–8743.

[13] FN Hooge, TGM Kleinpenning, and Lode KJ Vandamme. “Experimental studies on 1/f noise”. In: *Reports on progress in Physics* 44.5 (1981), p. 479.

[14] Junfeng Jiang. “CMOS wide-bandwidth magnetic sensors for contactless current measurements”. In: (2019).

[15] J-B Kammerer et al. “Horizontal Hall effect sensor with high maximum absolute sensitivity”. In: *IEEE Sensors Journal* 3.6 (2004), pp. 700–707.

[16] Yukihiro Kiyota et al. “Role of hydrogen during rapid vapor-phase doping analyzed by x-ray photoelectron spectroscopy and Fourier-transform infrared-attenuated total reflection”. In: *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* 16.1 (1998), pp. 1–5.

[17] Patrick M Lenahan and JF Conley Jr. “What can electron paramagnetic resonance tell us about the Si/SiO 2 system?” In: *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena* 16.4 (1998), pp. 2134–2153.

[18] James Lenz and S Edelstein. “Magnetic sensors and their applications”. In: *IEEE Sensors journal* 6.3 (2006), pp. 631–649.

[19] Bin Liu et al. “Low-power and high-sensitivity system-on-chip hall effect sensor”. In: *2017 IEEE SENSORS*. IEEE. 2017, pp. 1–3.

[20] Fei Lyu and Yu Wang. “Current-related sensitivity optimization of CMOS five-contact vertical Hall sensor”. In: *Journal of Magnetism and Magnetic Materials* 497 (2020), p. 166069.

[21] Guido Masetti, Maurizio Severi, and Sandro Solmi. “Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon”. In: *IEEE Transactions on electron devices* 30.7 (2005), pp. 764–769.

[22] Gary S May et al. *Fundamentals of semiconductor fabrication*. Wiley, 2004.

[23] V Mohammadi, WB De Boer, and LK Nanver. “Temperature dependence of chemical-vapor deposition of pure boron layers from diborane”. In: *Applied Physics Letters* 101.11 (2012).

[24] PJA Munter. “A low-offset spinning-current Hall plate”. In: *Sensors and Actuators A: Physical* 22.1-3 (1990), pp. 743–746.

[25] P Negrini, D Nobili, and S Solmi. “Kinetics of phosphorus predeposition in silicon using POCl3”. In: *Journal of the Electrochemical Society* 122.9 (1975), p. 1254.

[26] J Nishizawa, K Aoki, and T Akamine. “Ultrashallow, high doping of boron using molecular layer doping”. In: *Applied physics letters* 56.14 (1990), pp. 1334–1335.

[27] Joris Pascal et al. “3D Hall probe integrated in 0.35 μm CMOS technology for magnetic field pulses measurements”. In: *2008 Joint 6th International IEEE Northeast Workshop on Circuits and Systems and TAISA Conference*. IEEE. 2008, pp. 97–100.

[28] Arup Polley, Srinath M Ramaswamy, and Baher S Haroun. “Residual offset in silicon Hall-effect sensor: Analytical formula, stress effects, and implications for octagonal Hall plate geometry”. In: *IEEE Sensors Journal* 20.19 (2020), pp. 11283–11291.

[29] Radivoje S Popovic. *Hall effect devices*. CRC Press, 2003.

[30] RS Popovic. “Hall devices for magnetic sensor microsystems”. In: *Proceedings of International Solid State Sensors and Actuators Conference (Transducers’ 97)*. Vol. 1. IEEE. 1997, pp. 377–380.

[31] RS Popovic. “The vertical Hall-effect device”. In: *IEEE Electron Device Letters* 5.9 (1984), pp. 357–358.

[32] RS Popovic, JA Flanagan, and PA Besse. “The future of magnetic sensors”. In: *Sensors and actuators A: Physical* 56.1-2 (1996), pp. 39–55.

[33] Robbe Riem, Johan Raman, and Pieter Rombouts. “A 2 MS/s full bandwidth Hall system with low offset enabled by randomized spinning”. In: *Sensors* 22.16 (2022), p. 6069.

[34] Ch Roumenin, K Dimitrov, and A Ivanov. “Integrated vector sensor and magnetic compass using a novel 3D Hall structure”. In: *Sensors and Actuators A: Physical* 92.1-3 (2001), pp. 119–122.

[35] Jacopo Ruggeri. “Realization of a novel 3D magnetic sensor”. PhD thesis. Politecnico di Torino, 2023.

[36] Jacopo Ruggeri et al. “Inverted pyramid 3-axis silicon Hall-effect magnetic sensor with offset cancellation”. In: *Microsystems & Nanoengineering* 11.1 (2025), p. 26.

[37] Naoto Saitoh et al. “Composition and growth mechanisms of a boron layer formed using the molecular layer doping process”. In: *Japanese journal of applied physics* 32.10R (1993), p. 4404.

[38] Christian Sander et al. “From three-contact vertical Hall elements to symmetrized vertical Hall sensors with low offset”. In: *Sensors and Actuators A: Physical* 240 (2016), pp. 92–102.

[39] Christian Sander et al. “Isotropic 3D silicon Hall sensor”. In: *2015 28th IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*. IEEE. 2015, pp. 893–896.

[40] Francesco Sarubbi. “Characterization of pure boron depositions integrated in silicon diodes for nanometer deep junction applications”. In: (2010).

[41] Ch Schott, D Manic, and RS Popovic. “Microsystem for high-accuracy 3-D magnetic-field measurements”. In: *Sensors and Actuators A: Physical* 67.1-3 (1998), pp. 133–137.

[42] Charles Surya and Thomas Y Hsiang. “Theory and experiment on the 1/f γ noise in p-channel metal-oxide-semiconductor field-effect transistors at low drain bias”. In: *Physical Review B* 33.7 (1986), p. 4898.

[43] Lewis M Terman. “An investigation of surface states at a silicon/silicon oxide interface employing metal-oxide-silicon diodes”. In: *Solid-State Electronics* 5.5 (1962), pp. 285–299.

[44] H-C Tseng, Fu-Ming Pan, and CY Chang. “Characterization of boron silicide layer deposited by ultrahigh-vacuum chemical-vapor deposition”. In: *Journal of applied physics* 80.9 (1996), pp. 5377–5383.

[45] R Steiner Vanha et al. “Trench-hall devices”. In: *Journal of microelectromechanical systems* 9.1 (2000), pp. 82–87.

[46] Wikipedia contributors. *Channelling (physics)*. 2025. url: `https://en.wikipedia.org/wiki/Channelling_%28physics%29` (visited on 12/07/2025).

[47] Yue Xu et al. “A highly sensitive CMOS digital Hall sensor for low magnetic field applications”. In: *Sensors* 12.2 (2012), pp. 2162–2174.

[48] Ming L Yu, DJ Vitkavage, and BS Meyerson. “Doping reaction of PH3 and B2H6 with Si (100)”. In: *Journal of applied physics* 59.12 (1986), pp. 4032–4037.

[49] Zhenyan Zhang et al. “High Sensitivity Horizontal Hall Sensors in 0.35 um BCD Technology”. In: *2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*. IEEE. 2015, pp. 510–514.