DESIGN OF VGA (VARIABLE GAIN RESISTOR)

Graduation Project Year 2022/2023

2022 - 2023

FACULTY OF ENGINEERING

Suhag university

Book contents

1. Introduction to VGA circuit

- Brief description of VGA circuit and its components

- Introduction to layout

- Importance of analog layout in VGA circuit design

2. Analog Layout Fundamentals

- Fabrication process

- Matching

- Latch up

3. VGA Circuit Layout Considerations

- Component Placement for VGA Circuit Layout

- Routing Signals in VGA Circuit Layout

4. Simulation and Verification of VGA Circuit Layout

- Design Rule Checker (DRC)

- Layout Versus schematic (LVS)

- Parasitic Extraction (PEX)

5. Conclusion and Future directions for VGA circuit design

6. References

Chapter 1

Introduction to VGA circuit

Chapter 1

Brief description of VGA circuit and its components: -

Importance of layout

I.1   Motivation: -

Microelectronic products continue to connect and change to control our world in the 21st century thanks to unprecedented technological advances and rapid adoption in economic and industrial society. In addition to the successful parade of smartphones contemporary achievements such as activity trackers delivery drones and electric vehicles as well as bio-implants for driverless cars and ongoing innovations related to the Internet of Things will not only address many logistical and medical and environmental problems of Today. but they have profoundly revolutionized modern life in many respects.

Driven by a market demand for low-cost, multi-purpose, and densely integrated circuits (ICs), the semiconductor industry can be seen to follow a trend towards system-on-chip (SOC) solutions with a growing amount of mixed-signal content, i.e., both digital and analog IC sections. Due to the growing need for more computing power digital circuits have long been of great interest but now analog circuits are also of great interest., primarily incited by the desire for functional diversification and system integration. Amongst others, the soaring importance of sensor functionality and the sophistication of advanced human-machine interfaces make Analog circuit parts are more important.

Increasing chip complexity and shortening product lifecycles require electronic design automation in both the digital and analog domains to complete the task of creating IC designs. But even digital IC design follows a highly automated algorithm synthesis flow from the start. Analog circuits are typically produced by experienced designers with a low level of automation especially since the schematic design phase where the circuit diagram is converted to the actual implementation circuit is a major bottleneck in the overall design flow Figure (1.1).
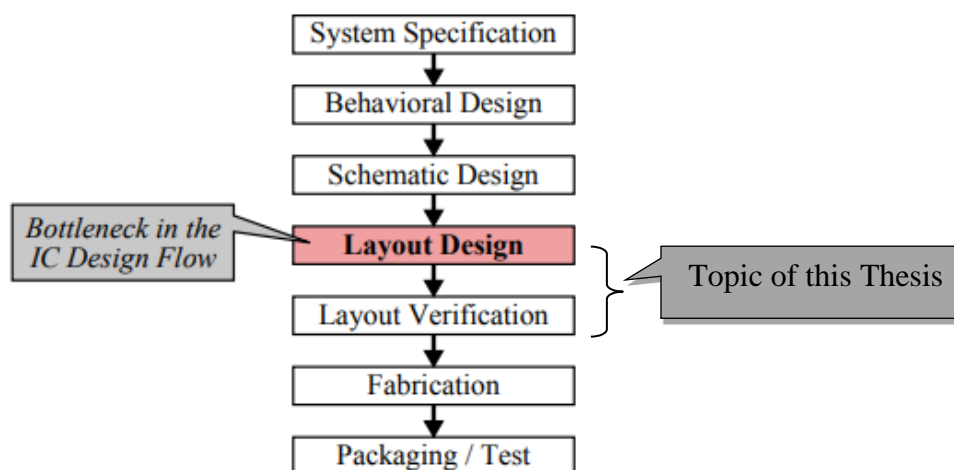
Figure 1. 1: Simplified illustration of the integrated circuit design flow.

The task of verifying the system specification as an electronic circuit is a very creative act. This is especially true for the analog layout design steps that this paper focuses on. Analog layout design relies heavily on the intuitive experience and creativity of the designers involved and is therefore considered an art by many experts in the field. This is also reflected in Alan Hastings title The Art of Analog Layout []. It is considered standard work in the analog display community. Unfortunately, these human characteristics cannot be easily automated on a scale that meets industrial needs.

In the past, layout design was done manually using drafting tools such as pencils, rulers, and compasses. The layout engineer would use these tools to draw the circuit on a sheet of paper or a Mylar film [BOPET] (Biaxially-oriented polyethylene terephthalate).

This process was a time-consuming and labor-intensive process that required a high level of skill and expertise. The layout engineer would need to visualize the circuit in three dimensions and create a layout that met the performance, reliability, and manufacturability requirements. As computer-aided design (CAD) technology advanced, the process of layout design began to be done using software. The first EDA layout tools were simple software programs that allowed designers to draw the schematic and layout of their circuits manually. However, these tools were limited in their functionality and were not very efficient.

Today, EDA layout tools have advanced significantly, and they are the primary means of layout design in the electronics industry. These tools use a combination of

algorithms, heuristics, and user input to optimize the placement and routing of components on a circuit board, while taking into account factors such as signal integrity, thermal management, and manufacturing constraints. EDA layout tools have many advantages over manual layout. They are much faster and more efficient, and they can handle the increasing complexity of modern electronics design. They also provide a high level of accuracy and consistency, which is critical for ensuring the performance and reliability of the circuit. Overall, the evolution of layout design from manual layout in the past to the EDA tools that are used today has been driven by advances in computer hardware and software, as well as the increasing complexity of integrated circuits.

The Problem of Analog Layout Design

This section describes some aspects of analog layout design that are important for understanding the work presented here and introduces technical terms that will be covered in the rest of this thesis.

As mentioned in Section 1.1, the problem of analog layout design is to take a given electronic circuit and turn it into a physical representation, which is itself also called a layout design. The purpose of that physical representation is to describe the detailed chip geometries on the photolithographically masks which need to be created for the various layers of the semiconductor manufacturing process.

These geometries include the layout of the circuit components and their electrical interconnections with interconnect holes between isolating layers, as well as the so-called bond pads for the chip's connections to its periphery, converting an electronic circuit into a practical layout requires consideration of many design restrictions and design objectives that arise from understanding the circuits' function. Thus, the readability of a schematic diagram is not only for circuit designers but also for layout designers.

I.1.1   Design Restrictions and Design Objectives: -

From a mathematical point of view, layout design is an optimization problem and can be regarded as a search for an optimal solution inside a huge solution space. Thereat, design restrictions define a valid region in the solution space, while the design objectives specify an optimum inside that valid region, as depicted in. Design restrictions are commonly divided into three categories:

- **Technological restrictions:** are meant to ensure the manufacturability of the (IC). They are derived from the chosen semiconductor technology and formulated as geometrical design rules. Design Rule Check (DRC) can be very complex, but most of them belong to one of the following groups: minimum width, minimum distance, minimum overlap, or minimum enclosure.

- **Functional restrictions** (electrical restrictions) are supposed to guarantee the circuit's proper electrical functioning. They can be separated into circuit-specific requirements (e.g., to prevent unwanted coupling effects) and process-specific requirements (such as the limitation of current density in electrical wires to avoid electromigration).

- **Design-methodical restrictions** are deliberately introduced to reduce the complexity of the layout design problem, thereby making the design task amenable to computer-aided automation approaches. An example is given by layer-dependent wire directions for the purpose of automated routing (e.g., metal1: horizontal, metal2: vertical).

Process Design Kit (PDK) and Design rules [1]

A **PDK** is a collection of files, libraries, and design rules that provide the information and tools necessary for designing ICs using a specific fabrication process. The PDK includes a range of design rule checks (DRCs) and layout versus schematic (LVS) checks to ensure that the layout of the IC meets the requirements of the fabrication process. The PDK also includes models for device simulation and characterization, allowing designers to simulate the behavior of the IC under different operating conditions.

In our work we use Saed_pdk_32_28 is a Process Design Kit (PDK) developed by the Semiconductor Advanced Electronics Design **(SAED)** group at Ain Shams

University in Egypt. it appears that the design rules for Saed_pdk_32_28 are based on micron (µm) measurements, rather than lambda-based rules. Lambda-based rules are typically used in deep submicron processes, where the feature sizes are smaller than the wavelength of light used in photolithography. The 0.35 µm technology node is not considered a deep submicron process,

I.2   Levels of Design Hierarchy

Analog IC layouts are usually built in a hierarchical fashion, which is achieved by putting design components inside other design components. Basic design units such as transistors are usually provided as primitive devices i.e., basic cells with no sub-hierarchies. At a higher level a group of design elements that together form a functional unit can be contained in a single (but in a hierarchical case) cell. Functional units become modular library components that can be instantiated in layouts like base units. To avoid semantic confusion when talking about units it is best to classify them according to these characteristics based on their position in the design hierarchy, for that purpose, this thesis proposes and adheres to the following terminology

	Hierarchy Level	Examples	Degree of (Re-)Utilization
Blocks	Block Level	Variable Gain Amplifiers (VGA)	Low
Advanced Modules	Module Level	Operational Transconductance Amplifiers (OTA), Differential Amplifiers,	Medium
Simple Module	Module Level	Differential Pairs, Current Mirrors,	High
Primitive Device	Devices Level	Transistors, Resistors, Capacitors, Guard rings	Very high

table: Classification of hierarchical cells in analog/mixed-signal design

I.3 Main Design Tasks:
Device Generation, Floorplanning, Placement, Routing

	Floorplanning	Placement	Routing
Considered Components	Circuit Blocks (treated as black boxes)	Primitive Devices and Modules	Wire Segments + Vias (to cross metal layers)
Quantities to be set by the Design Task	- Block Locations - Aspect Ratios - Pin Positions	- Locations - Orientations - Layout Variants	Wire Paths, Segment Layers and Widths + Via Positions and Sizes
Typical Restrictions	- Rect. Chip Outline - Block Distances - Chip Regions	- Block Outline - Space for Routing - Parasitics	- No Wires Above Devices - Available Metal Layers - Parasitics and Currents
Primary Objectives	- Minimize Area and Wirelength - Optimize Power Supply and Current Flow	Device Matching	Minimize Number of Vias and Number of Metal Layers

Table The main tasks in analog layout design

I.3.1 Device Generation

in layout refers to n is the task of creating the layouts for the individual components of the given input circuit. Here for, every component needs to be individually layouted according to its respective sizing (e.g., the channel width and channel length of a MOS transistor). In the past, this task has been an integral duty on the shoulders of an IC design team.

Today it is common practice that the primitive devices of a semiconductor technology are readily delivered by the vendor as part of a so-called process design kit (PDK), which is involves creating the metal interconnects, transistors, and other components that make up the circuit, and optimizing their placement, spacing, and routing to meet the requirements of the technology node. It is an iterative process that involves simulation, testing, and optimization.

Designers use simulation tools to model the performance of the circuit and identify potential layout issues, such as parasitics or timing violations. They then make adjustments to the layout, such as changing the placement of components or adjusting the width of metal interconnects, and test the circuit again to verify that the changes have improved its performance.

Device generation in layout is a critical step in the design of integrated circuits, as it determines the physical structure of the circuit and its performance characteristics.

Even primitive devices have an immense layout variability, and one major source of this variability is device folding. For example, a MOS transistor can be folded by changing its so-called number of fingers. As shown in Figure 2.1, the transistor variants thus have different aspect ratios while preserving the total channel width and channel length. Device generation is already important during floorplanning for estimating the total size of a layout block.

$$total\ width\ =\ width\ per\ finger\ *\ num\ of\ finger\ *\ num\ of\ multiplier$$
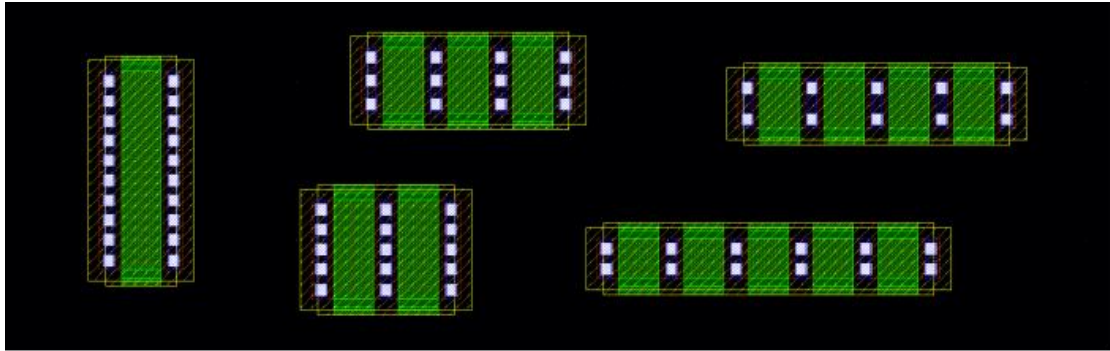
Figure 2.1: Different layout variants of a MOS transistor with the same total channel width and length.

In VLSI fingers and multipliers are two terms used to describe the layout transistors. finger means how many gates a transistor has while multiplier is how many times a transistor is replicated. Splitting transistors can either be done by using multiple transistors with a single gate" multiplier" or with transistors that have multiple gate "fingers". When laying out a MOSFET with a particular width and length, in an EDA tool, one has two options with regards to the **shape of the gate**:

1) **Single stripe** (classical case) (one finger);
2) **Several stripes** (several fingers).

 advantages and disadvantages of a transistor layout with **multiple fingers** (MF) vs **single finger**?

1) MF provide more **flexibility** in layout planning for transistor with high W/L or L/W. In other words, allows making a layout more square-like.
2) MF allow better **matching** of transistors, when needed. For example, if using common-centroid techniques.
3) MF layout **reduces gate resistance**.
4) MF **reduce current density** in the gate if there are technology limitations on this

Transistors with multiple fingers have the disadvantages also:

(1) the current direction is different for two neighboring fingers. E.g., if for the first finger the source is to the left then the source for the next finger will be to the right.

(2) The properties of transistor can change depending on the current direction. Therefore, extra care has to be taken when trying to achieve good matching.

(3) Using multiple fingers to obtained scaled current sources for example in a current mirror is also considered inferior to having multiple single gate transistors because of slightly different properties.
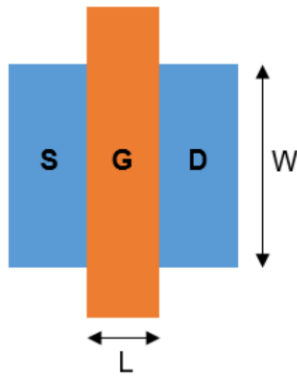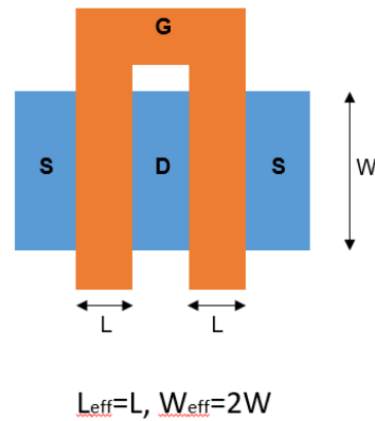

Figure 1: One finger.

$L_{eff}=L, W_{eff}=2W$

*Figure 2: Two fingers. Two transistors in parallel (widths are summed).*

I.3.2 Floorplanning

is the task of specifying locations, aspect ratios, and pin positions for the layout blocks of a chip. Therein, each block is treated as a black box whose area is roughly estimated by the floorplan designer from generating the block's devices. For economical and electrical reasons, the primary objectives in floorplanning are to minimize the total layout area and the total wirelength, as well as to optimize the power supply and the current flow. A hard restriction concerning the layout area can be that the blocks must fit into a fixed outline, depending on the semiconductor package chosen for the physical sealing of the chip during the final stage of the fabrication. In general, the top-level chip boundary is demanded to be a rectangle whose aspect ratio should not depart too far from a square. In the context of wirelength minimization, some blocks are required to be positioned close to the chip boundary because they will later be connected with the periphery. On the other hand, it may also be necessary to keep a certain minimal distance between dedicated blocks such that sensitive signals are not disturbed by unwanted thermal and electrical influences. A large block can contain subordinate blocks that also need to be floor planned.

Pin-Driven Planning:

The pin-out is the first step of your floorplan in which you may participate as a mask designer. Some refer to it as the pad-out. This is the process in which the input and output pins that will surround the chip in its packaging are defined.

The pin-out directly affects your work. The quality of your pin-out directly affects how good your chip floorplan will be and how easy the chip will be to layout.

You could use the same package and the same signals, but use the various signals to be in very different places. Very different pins placement will produce a very different layout as same as in the next figures.
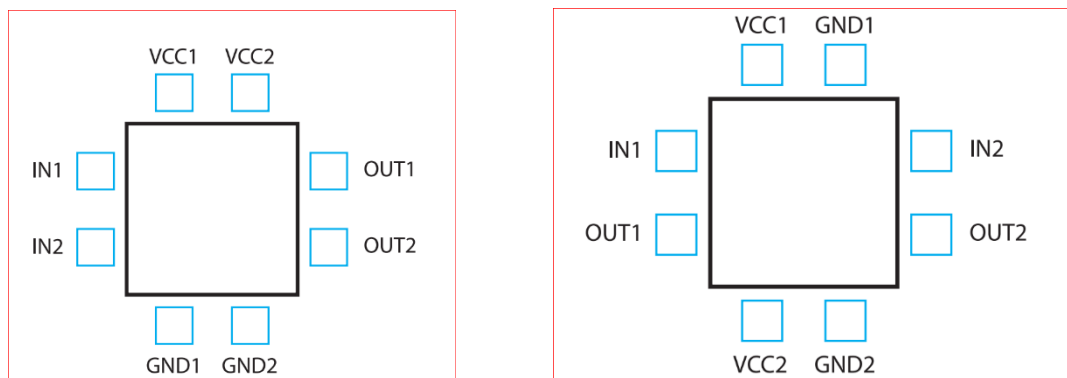
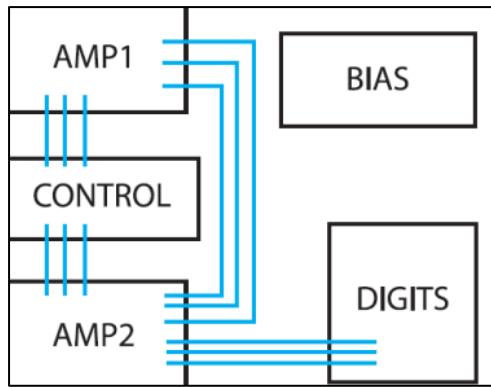
To Obtain A good floorplanning [5]:

- Start thinking about the layout when doing the schematics. Think about how this schematic translates to the layout.
- Industrial quality layout.
- Compact rectangular layout.
- Input and output on the two sides.
- VDD and VSS on the top and bottom.
- N-well all around
- User-defined connectivity requirements for every layer, in addition to net-based routing constraints, handle the results of analogue signal crosstalk, minimum capacitance, and resistance.
- Understands the various divisions or blocks of a design.
- Understands the important features of each block: size, aspect ratio, and pins.

- Dynamically displays the connectivity between blocks and connections to the pads.
- Allocates space for routing based on the number of routing layers.
- Places each block and optimizes the pin locations for each block based on the overall connectivity requirements and the feasibility of routing the signals between blocks.
- Places top-level ports based on constraints. floorplan made the electron's life travels from one side to the other easier.
- Use dummy elements to improve symmetry and avoid Gate Etching Effect.
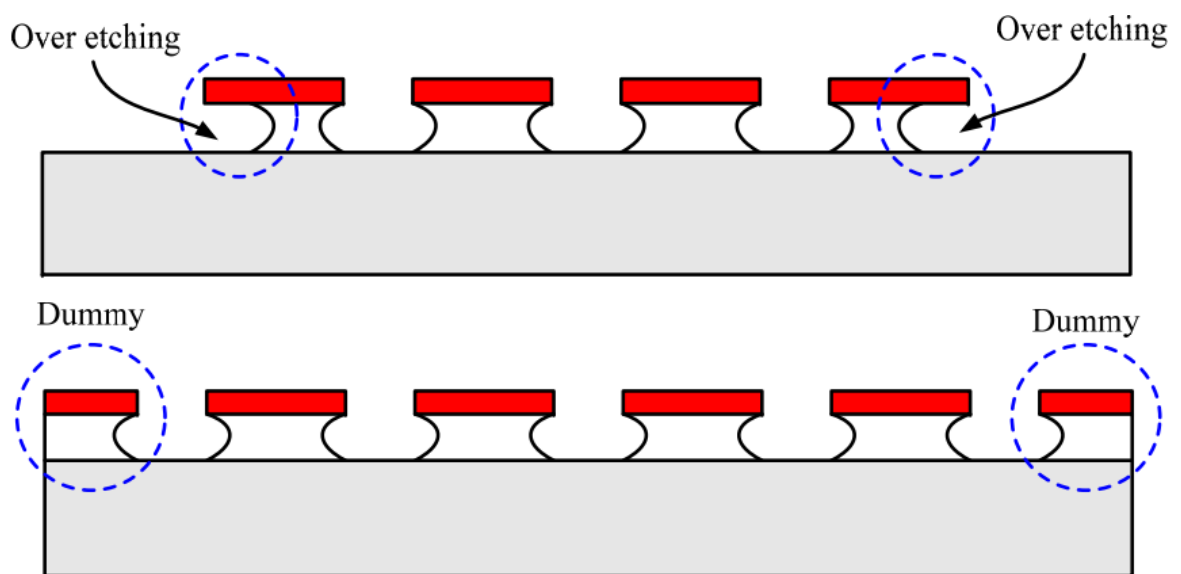
- Matched transistors are used extensively in both analog and digital CMOS circuits.
- Keep sufficient spacing between power blocks and sensitive blocks.
- Use Guard rings for isolation and guard Ring is a useful tool to protect our circuits from parasitic effects and so that chances of latch-up are reduced.
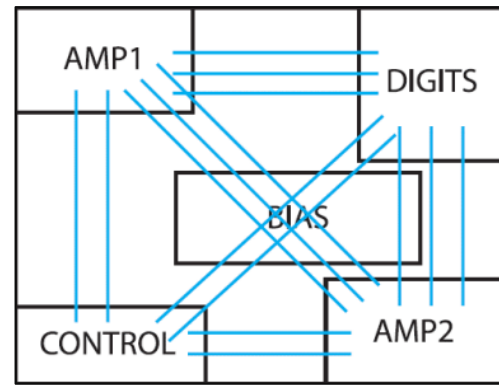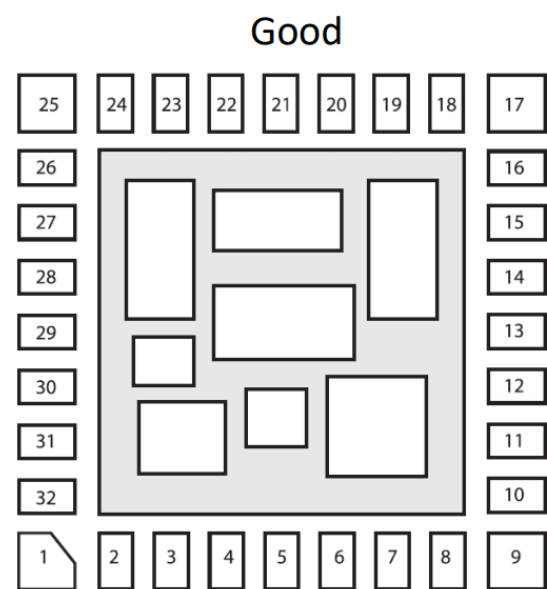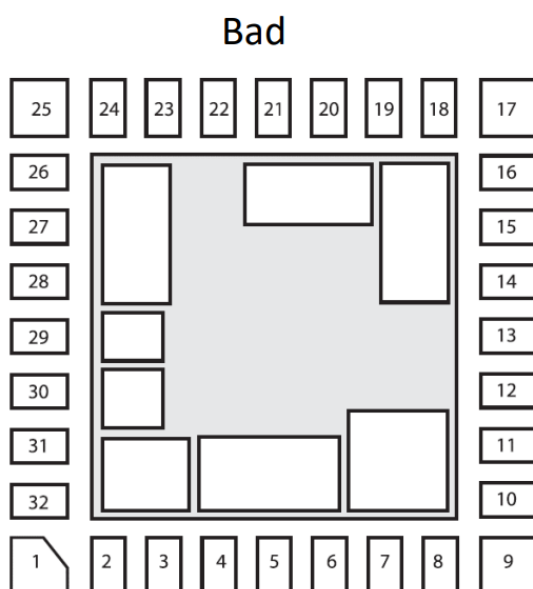
Good   Bad

Neat fly lines indicate good floor planning [5]:

o   Minimize The area as possible is Perfect but be care to leave a good space for routing and decoupling caps.

o   Inputs and outputs should be placed near the correct cell blocks.

I.3.3 Placement: it is not only means to move the layout components into appropriate locations, but may also require to rotate them and to vary their layout without affecting their electrical function (for example by changing the number of fingers discussed above), such that all components fit into the block outline defined during the floorplanning phase. As in floorplanning, it is desired to obtain a device placement wherein the total area and wirelength are minimized, but usually these objectives first and foremost stem from the need

to obtain a good device matching. The same is true for higher-level modules whose placement aims at achieving an overall symmetry of the layout block where the modules are placed in. Opposing the need to place the components close to each other, a common demand is that a sufficient amount of space must be reserved between the components to accommodate their routing.

I.3.4 Routing: refers to the process of creating the physical connections between the various components of an electronic circuit. This involves designing the metal interconnects that link the transistors, resistors, capacitors, and other components of the circuit, and optimizing their placement and routing to meet the specific requirements of the design. Routing in layout is a critical step in the design of electronic circuits, as it determines the physical structure of the circuit and its performance characteristics. The routing process involves creating the metal tracks that connect the different components of the circuit, and optimizing their placement and routing to minimize parasitic effects such as resistance, capacitance, and inductance.

The demand to allow for routing space between layout components is rooted in the fact that sensitive circuitry often forbids electrical wires to be drawn above these components. this has to be taken into account during the routing task. Another restriction is to confine which metal layers are available for the routing. On the lower design levels, a common agreement is that only the first two ($M2$, $M3$) of the available metal layers may be used in order to retain the remaining metal layers for the later top-level routing. Leading a wire across different metal layers requires to connect the respective wire segments with a VIA (vertical interconnect access). The size of a via and the width of a wire segment must be set with respect to the expected current load.

Primary objectives in routing are:

- o to minimize the number of vias (i.e., to avoid crossing between metal layers if possible),
- o to minimize the number of metal layers (and thus the number of necessary photolithographically masks),

Some important Rules which should follow in Routing:

- Choose routing layers based on process parameters and circuit requirements. For each process a standardized list of routing layers should be determined based on layer resistance and capacitance. Layers such as N-well, active, and high-resistance poly gate are not used for routing.
- Priorities between routing layers can also be standardized using the same criteria [5].
- Use of more numbers of columns and rows vias.
- The use of metals with an odd number for vertical wires and an even number for horizontal wires to avoid interference between metals and each other.
- Minimize total run time for carrying out routing process [5].
- Minimize total wire length.
- Avoid too much parallel routing of metals (to decrease parasitic capacitance [5].
- Increase the width of the wire to overcome Electro migration [6].

Manhattan routing rule:

also known as the Manhattan style or Manhattan distance, is a common technique used in analog layout design to route interconnects between components. It is named after the street layout of Manhattan, which is characterized by a grid of perpendicular streets. The Manhattan routing rule involves routing interconnects in a series of right-angled, orthogonal segments, either horizontally or vertically. This ensures that the interconnects are parallel to the edges of the layout grid and minimizes parasitics such as capacitance and inductance. Using it also to reduce the common area between metal layers. For example, if Metal 1 is routed in horizontal direction. Is required to route M2 in vertical direction as shown in Figure (1.3)
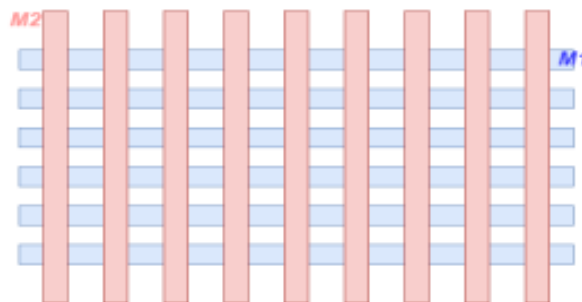

Figure 13. Manhattan routing rule

While design restrictions are strict confinements which must definitely be satisfied, design objectives represent gradual optimization goals that are pursued as good as possible. They can be roughly classified into economic optimization goals and functional optimization goals. Economic optimization goals include the reduction of product costs (e.g., by minimizing the total chip area and the number of required metallization layers) as well as reducing the development costs (e.g., by minimizing the design effort via design automation). Examples for functional optimization goals are the minimization of the total wirelength as well as optimizing the chip's heat dissipation to prevent critical hot spots. One of the key advantages of the Manhattan routing rule is that it is easily automated using computer-aided design (CAD) tools. This allows designers to quickly and efficiently route interconnects in complex circuits, saving time and reducing errors. However, the Manhattan routing rule also has some limitations. For example, it may not be the most efficient routing strategy for certain types of circuits, such as those with high-frequency signals or those with non-orthogonal components. []

As the name implies, functional restrictions and functional optimization goals pertain to the functionality of an integrated circuit. Herein, three basic issues can be embraced by the term functionality:

- Does the circuit work accurately enough to perform the desired function?
- How well is the circuit set up against long-term failure owing to effects of degradation?
- What measures are taken to prevent an instantaneous malfunctioning due to fabrication   problems?
  The answer of those issues will be postponed to chapter 3

I.4 Physical Verification

After the physical design is complete the layout must be fully verified to ensure correct electrical and logical functionality (correctness of a physical layout design with respect to the design rules and specifications), It is an essential step in the semiconductor manufacturing process, as it ensures that the final product will function correctly and meet the desired performance criteria. The physical verification process involves checking various parameters such as the width and spacing of metal and poly lines, the

alignment of cells, and the presence of shorts and opens. The verification process is performed using various tools and techniques such as:

- **Design rule checking (DRC):** verifies that the layout meets all technology-imposed constraints. These rules include the following categories: -
  - Typical geometrical design rules that require minimum or maximum values for widths, spacings, extensions, intrusions, and enclosures for layout polygons. These rules make sure that the layout structures can be correctly generated on the silicon due to process accuracy
- **Layout versus schematic (LVS):** is checking that the design is connected correctly. The schematic is the reference circuit and the layout is checked against it. In principle, the following is verified:
  - Electrical connectivity of all signals, including input, output, and power signals to their corresponding devices
  - Device sizes: transistor width and length, resistor sizes, capacitor sizes.
  - Identification of extra components and signals that have not been included in the schematic; floating nodes would be an example of this. The last item overlaps into the items checked in the electrical rules check, which is described previous

- **Parasitic extraction (PEX):** is extracting the parasitic effects of interconnects and devices in a chip. Parasitic effects are unwanted electrical characteristics that arise due to the physical properties of the components used in the chip design. These effects can cause delays, noise, and other performance issues that can impact the functionality of the chip.

Chapter 2

Analog Layout Fundamentals

Fabrication process: -

Integrated Circuits are a collection of electrical circuits that are put on a tiny electronic chip. With the help of the fabrication process, a large number of devices are produced, including transistors, MOSFETs (Metal Oxide Semiconductor Field Effect Transistors), microcontrollers, computer processors, etc. Because of their stable structures, silicon and germanium are the most often used materials in the manufacturing of semiconductors.

Other materials favored in the fabrication sector include wood, thermoplastics, resins, silver, aluminum, magnesium, copper, and carbon steel. Let's first talk about fabrication and the needs it has in the semiconductor industry.

Construction of an industrial product is known as fabrication. It can also be described as a group of procedures used in the production of electronic goods. Silicon semiconductor chips, as an illustration, etc. Fabrication is a procedure used to turn raw materials into finished goods in the case of metals.

By using the smaller components created during the fabrication process, larger electrical components can be created.

The following are the fundamental techniques used in fabricating integrated circuits:

- Crystal growth
- Wafer processing
- Masking
- Oxidation
- Etching
- Deposition
- Ion implantation , diffusion and doping
- CVP
- Metallization
- Packaging

Manufacture

Semiconductor devices have long been used in electronics. Semiconductor device fabrication is the process used to manufacture semiconductor devices, typically integrated circuits (ICs) such as computer processors, microcontrollers, and memory chips (such as NAND flash and DRAM) that are present in everyday electrical and electronic devices. Electronic circuits are gradually built on a wafer, normally formed of pure single-crystal semiconducting material, using a multi-step photolithographic and physio-chemical process (including phases like thermal oxidation, thin-film deposition, ion-implantation, and etching). Silicon is almost always used, but various compound semiconductors are used for specialized applications.

Silicon Manufacture

In this chapter, we will consider only silicon-based (Si) technologies. Although other compound materials in groups III through V, such as gallium arsenide (GaAs) and aluminum gallium nitride (AlGaN), are also used to produce VLSI chips, silicon is still the most popular material, with excellent cost–performance trade-off. The mineral quartz consists entirely of silicon dioxide, also known as silica. Ordinary sand is chiefly composed of tiny grains of quartz and is there for also silica.

Despite the abundance of its components, elemental silicon does not exist in nature. The element can be synthesized artificially by heating silica and carbon in an electric boiler. The carbon reacts with the oxygen in the silica, resulting in nearly pure molten

silicon. As this cool, a slew of minute crystals forms and coalesces to form a fine-grained grey solid. Because it contains a large number of crystals, this type of silicon is called poly-crystalline. This metallurgical-grade polysilicon is unsuitable for semiconductor manufacturing due to impurities and a disordered crystal structure.

Metallurgical-grade silicon can be further refined to produce an extremely pure semiconductor-grade material. Purification begins with the conversion of the crude silicon into a violent compound, usually trichlorosilane. After repeated distillation, the extremely pure trichlorosilane is reduced to elemental silicon using hydrogen gas. The final product is exceptionally pure, but still poly-crystalline. Practical integrated circuits can only be fabricated from single-crystal material, so that the next step consists of growing a suitable crystal.

Crystal Growth

The principles of crystal growth are straightforward and well-known. Assume a few sugar crystals are added to a saturated solution that then evaporates. Sugar crystals act as seeds for the deposition of more sugar molecules. The crystals eventually grow to be quite large. Even in the absence of a seed, crystal growth would occur, but the result would be a swarm of small intergrown crystals. By suppressing unwanted nucleation sites, the use of a seed allows the growth of larger, more perfect crystals.

In principle, silicon crystals can be grown in much the same manner as sugar crystals. In practice, no suitable solvent exists for silicon, and the crystals must be grown from the molten element at temperatures in excess of 1400°C. the resulting crystals are at least a meter in length and ten centimeters in diameter, and they must have a nearly perfect crystal structure to be useful to the semiconductor industry. These requirements make the process technically challenging.

The usual method for growing semiconductor-grade silicon crystal is called *Czochralski process*. This process, illustrated in figure 2.1, use a silica crucible charged with pieces of semi-grade polycrystalline silicon. An elective furnace raises the temperature of the crucible until all the silicon melts. The temperature is then reduced slightly and a small seed crystal is lowered into the crucible. Controlled cooling of the melt causes layers of silicon atoms to deposit upon the seed crystal. The rod holding the seed slowly rises so that only the lower portion of the growing crystal remains in

contact with the molten silicon. In this manner, a large silicon crystal can be pulled centimeter-by-centimeter from the melt. The shaft holding the crystal rotates slowly to ensure uniform growth. The high surface tension of molten silicon distorts the crystal into a cylindrical rod rather than expected faceted prism.
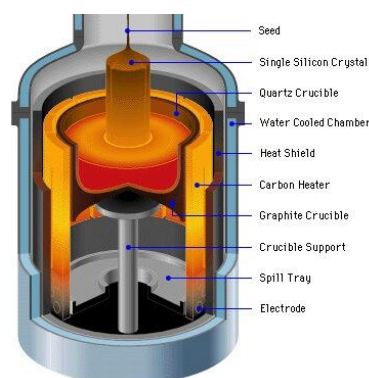

Figure 2.1 czochralski process for growing silicon crystal.

The czochralski process requires careful control to provide crystals of the desired purity and dimensions. Automated systems regulate the temperature of the melt and the rate of crystal growth. A small amount of doped polysilicon added to the melt sets the doping concentration in the crystals. In addition to the deliberately introduced impurities, oxygen from the silica crucible and carbon from the heating elements dissolve in the molten silicon and become incorporated into the growing crystal. These impurities subtly influence the electrical properties of the resulting silicon. Once the crystal has reached its final dimensions, it is lifted from the melt and is allowed to slowly cool to room temperature. The resulting cylinder of monocrystalline silicon is called an *ingot*.

Since integrated circuits are formed upon the surface of a silicon crystal and penetrate this surface to no great depth, the ingot is customarily sliced into numerous thin circular sections called wafers. That are 400μm to 600μm thick. Each wafer yields hundreds or even thousands of integrated circuits. The larger the wafer, the more integrated circuits it holds and the grater the resulting economies of scale.

Wafer Manufacturing

A series of mechanical processes are used to create wafers. The ingot's two tapered ends are sliced off and discarded. The remaining material is ground into a cylinder, the

diameter of which determines the size of the resulting wafers. After grinding, no visible indication of crystal orientation remains. The crystal orientation is determined experimentally, and a flat stripe is ground down one side of the ingot. Each wafer cut from it will have a facet, or flat, that clearly identifies the crystal orientation.

After grinding the flat, the manufacturer uses a diamond-tipped saw to cut the ingot into individual wafers. Approximately one-third of the precious silicon crystal is reduced to worthless dust in the process. The sawing process leaves scratches and pockmarks on the surfaces of the resulting wafers. Because the tiny dimensions of integrated circuits necessitate extremely smooth surfaces, each wafer must have one side polished. Mechanical abrasives are used first, followed by chemical milling. The resulting mirror-bright surface has a dark grey color and the characteristic silicon near-metallic luster.

The finished wafers' thickness and diameter must correspond to the mechanical strength and other physical qualities of the material the wafers will be used to make. The finished product must be sturdy enough to support its weight without shattering when handling the products of the wafers' diverse applications. As more material is added in the fabrication of the slice, the diameter of the wafer increases, as does the weight of the wafer. When enough weight has been added, the diameter cannot be increased because it compromises the slice's strength.

If the weight is not properly placed, a small amount of pressure will be enough to break down the wafer. Before they can be used in manufacturing, sliced wafers must be processed. Machines and abrasive chemicals are used to smooth out the rough surface of the wafer. The flawless surface facilitates printing circuit layouts on the wafer surface.

All silicon wafers are extremely useful components, just that there are different variations used for different purposes. It is therefore imperative to know about the different types of Silicon Wafers. There are mainly two varieties of silicon wafers commonly used today. These are undoped silicon wafer and Doped silicon wafer. Undoped Silicon Wafers, also known as Intrinsic or Float Zone (FZ) do not have any dopants in them. They are made of strictly pure crystalline silicon. This type of silicon wafer is recognized as the ideal semiconductor. Doped silicon wafers are formed by

introducing dopants (certain impurities) into the silicon crystal during the formation process. When boron is added into the mixture, a P-type doped silicon wafer is produced. P-type silicon wafers have numerous positively-charged holes. To produce an N-type doped silicon wafer, elements like phosphorus, arsenic, or antimony will be added. N-type silicon wafers have a negatively-charged electron in them. The quantity of dopant discovered in the wafer will ascertain if it is degenerate or extrinsic. To be degenerate means there's a higher concentration of dopants in it, while to be extrinsic means it has little or moderate dopants.

Photolithography

photolithography is used in all structuring process steps. Its purpose is to transfer a two-dimensional image of the required structures onto the wafer surface, so that subsequent processing (e.g. implantation, etching) can be applied to a restricted area.

The wafer is first coated with a thin radiation sensitive film, called a *photoresist* or *resist*. A photomask is then exposed to light to project a black and white image of the desired structure on the photoresist. The solubility of the photoresist changes with reference to. a specific fluid, called a *developer*, at the regions exposed to light. In the *developing process,* the soluble areas of the photoresist are removed from the wafer with this fluid. The insoluble parts of the photoresist remain.

Photoresist

Photoresist is deposited on the wafer by means of spin coating. Liquid coating material is applied to the center of the wafer, which is rapidly rotated. The centrifugal force causes the fluid to be distributed on the wafer. The solvent for setting the viscosity in the fluid is vaporized to produce a layer of constant thickness. Spraying the resist onto the wafer is another method that is sometimes used. Photoresists are radiation-reactive polymers. Positive and negative photoresists are available. A positive photoresist is a type of photoresist in which the portion of the photoresist that is exposed to light becomes soluble to the photoresist developer. The unexposed portion of the photoresist remains insoluble to the photoresist developer (see Fig. 2.2, top row). The effect is the opposite with negative photoresists where the portion of the photoresist that is exposed to light becomes insoluble to the photoresist developer. The unexposed

portion of the photoresist is dissolved by the photoresist developer (see Fig. 2.2, bottom row). In each case, the remaining photoresist serves as a mask for the next process step.
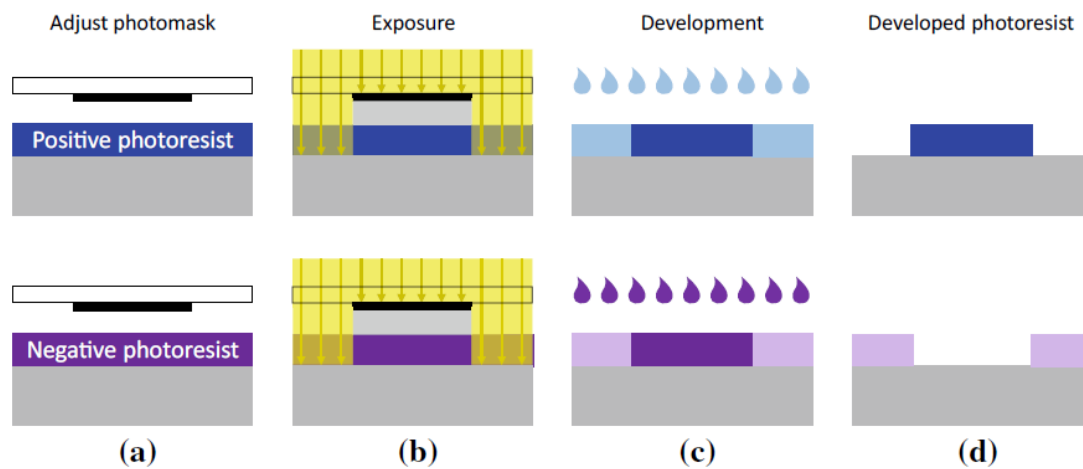

Figure 2.2 Schematic representation of the photolithography with positive resist (above), where the exposed regions are removed, and negative resist (bottom), where the exposed regions are kept.

Photomasks and Exposure

A *photomask* is used to expose a wafer. A photomask is a sheet of glass, on which a black-and-white image of the structures to be processed is applied to an opaque layer made of chromium. When this photomask is exposed, a shadow is cast on the exposed wafer to produce the desired image. There are two types of exposure: *direct exposure* and *projection exposure*.

Direct Exposure

With direct exposure, the photomask is placed above and in close proximity to the wafer, either in direct contact with the photoresist (*contact exposure*) or close to it (*Proximity exposure*). The result is in both cases a 1:1 (full scale) image produced by simple shadowing. Hence, the structure sizes on the photomask must correspond with the structures on the wafer (so-called "1X photomasks").

Contact exposure is unsuitable for volume manufacture as it can cause damage to and soiling on the resist and the photomask. Proximity exposure differs in that there is no contact between photomask and resist. However, the distance between mask and resist in proximity exposure, which cannot be reduced indefinitely but must lie between 10 and 40 μm, can cause resolution degrading.

If the dimensions of the structures being imaged are similar in size to the exposure wavelength, significant diffraction phenomena may occur. Direct exposure therefore approaches its limit for structure sizes to be imaged at this order of magnitude. The limit is even bigger (approximately 3 μ m) for proximity exposure. Hence, the structure sizes required in state-of-the-art semiconductor technologies that are smaller than this value cannot be imaged with direct exposure; this process is therefore not deployed any more today.

Projection Exposure

Projection exposure was developed to overcome the above-mentioned direct exposure issues. Here, the pattern on the photomask is projected through lenses onto a wafer coated with photoresist, as illustrated in Fig. 2.3. This system brings with it two fundamental benefits: (i) photomask and wafer are physically separated and (ii) the image can be optically reduced in size during the projection thus improving imaging accuracy. Among the most commonly used photomasks are 4X, 5X and 10X masks, in other words, photomasks with imaging ratios of 4, 5 and 10 to 1. A small part of a wafer can only be exposed in one step with this method, as the sizes of the photomasks and lenses are limited.

Photomasks therefore have structures for one or a few chips. These types of photomasks are called *reticles*. A Reticle is a special type of photomask where the data for only part of the final exposed area is present. A Reticle is loaded into a Stepper or Scanner system where multiple exposures are made to cover the full patterned area. However, the term "reticle" has become a synonym for "photomask" as state-of-the-art photomasks are seldom designed for an entire wafer.

A wafer is exposed in many single steps in the so-called "step-and-repeat technique". The wafer is transported on a positioner table under the projection optics. This type of exposure equipment is called a *wafer stepper* or simply a *stepper*. Figure 2.3 shows the stepper principle of operation.
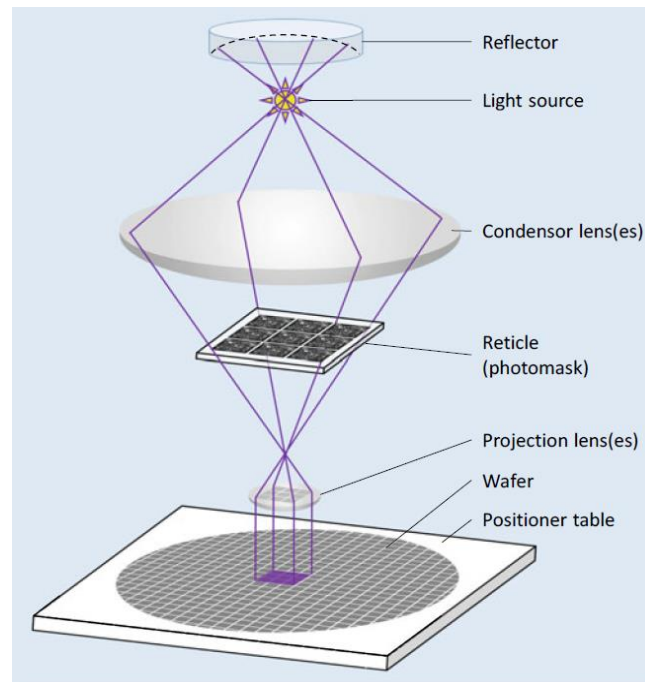
Figure 2.3 Exposing a wafer with step-and-repeat technology. Using one reticle, multiple exposures are made to cover the full patterned area

Technology is being continuously developed to improve imaging accuracy, and such Ultraviolet light is used for higher optical resolution. Exposure and photomask in addition to these developments, there are other approaches for downscaling structure sizes. Mirror systems are needed for projections using light of shorter wavelengths as the materials available for lenses are increasingly opaque.

Alignment and Alignment Marks

Progressive structural processes that interact with one another are used to construct integrated devices. This means that a reticle's (i.e., a photomask's) position for a specific layer must always be precisely aligned with the structures that are already present on the wafer. For instance, when the devices have been installed, the contact holes need to be precisely positioned where the connecting areas are.

Before each exposure in the step-and-repeat procedure, the wafer's position and attitude with regard to the exposure reticle must be modified. To do this, alignment marks on the reticle are automatically identified by optical methods (Fig. 2.4). These marks, which include geometric shapes like crosses, build structures on the wafer in accordance with the actual process step, just like all other geometric elements on the

reticle. If alignment marks' effects on these structures can be seen visually, the wafer position can be aligned with the help of alignment marks on subsequent photomask(s).

An opening in the photomask in the shape of a cross serves as the alignment mark in Fig. 2.4 to the right. The wafer underneath is modified so that the cross-shaped structure, which is also present, is situated in the center of the wafer.

However, some structural processes, like ion implantation doping techniques, for instance, leave no reliable optical traces. But this is not a problem. As mechanical tolerances result in variations in every adjustment, adjustments are made in as many process stages as possible using the same alignment mark on the wafer to prevent the deviations from building up. When a structure can no longer be detected, an alignment mark from a later process stage is used. The "newer" alignment mark is then used in the subsequent phases.

During the alignment procedure, the correct angular orientation of the wafer must be checked in addition to the correct location of the alignment marks. As a result, the alignment marks are placed at two distant points on the reticle to ensure accuracy. the wafer attitude is correct.

The alignment marks are placed outside the chips because they are not functional chip structures. The chips have a 50-100μ m clearance from one another to facilitate (subsequent) chip dicing. This clearance is known as a "sawing trench," "saw street," or "saw street clearance." "scribe line". In this clearance space, the alignment marks are placed (Fig. 2.4, left).
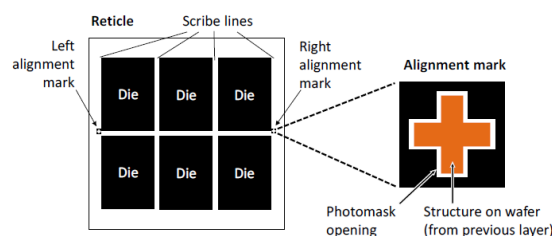
Figure 2.4 Alignment marks on a reticle (photomask) with six chip structures.

Reference to Physical Design

The extremely high effort required to create the photomasks derived from the layout design to perform exposure is the price for the degree of miniaturization currently achieved with structure sizes of the order of nanometers. Modern photomasks Semiconductor processes can be quite costly.

It's important to remember that a single design flaw can render photomasks and wafers totally worthless. In the event of a fault, this financial hit is accompanied by a significant delay in development time. Troubleshooting and a new production run can take up to six months. When one considers the delays in bringing the product to market, the financial losses become even more severe.

Imaging Errors

The structures on the photomasks are determined by layout data. So far, we've discussed how photolithographic images and subsequent targeted process steps are used to convert these mask structures to wafer structures. When the structures created "on" and "in" the wafer are compared to the original layout structures, various types of deviations become apparent. We then go over these unavoidable deviations and how we can deal with them in the layout design.

There are three categories of imaging errors: (1) *overlay errors,* (2) *diffraction effects* and (3) *edge shifts*. Overlay errors and diffraction effects occur during exposure, while edge shifts occur in subsequent structuring process steps.

Overlay Errors

The wafer and photomask (reticle) cannot be positioned with absolute accuracy relative to the exposure device due to mechanical tolerances and measurement inaccuracies that can occur during alignment. As a result, structures on the photomask are not precisely mapped to the wafer during exposure, as specified by the layout template.

Exaggerated representations of possible exposure faults are shown in Fig. 2.5, where we show (a) displacements and (b) rotations that can occur relative to the required positions. To adjust the depth of focus, the wafer and photomask are moved along the optical axis. The layers will be scaled relative to one another if the distances between

the masks, lenses, and wafer are changed during this focusing step (c). The perspective will be distorted if the photomask is tilted with respect to the optical axis (d).

Overlay errors are also caused by the fact that wafers and photomasks, like all materials, expand when heated. If wafer exposures take place at different temperatures, spacings between structures on the wafer will change. As a result, displacements (see Fig. 2.5a) occur, the magnitude of which is determined by where they occur on the exposed wafer. To reduce this effect, keep the exposure temperatures as constant as possible throughout the manufacturing run.
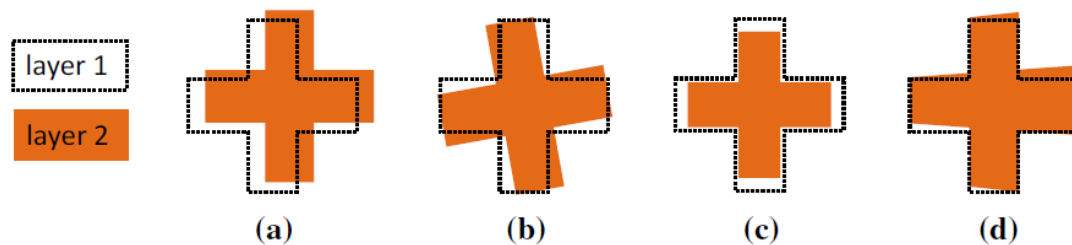
Figure 2.5 Possible overlay errors between two layers: a displacement, b rotation, c scaling,

This cause of overlay errors cannot be completely eliminated because it is not always possible to keep these temperatures constant. In high temperature steps of the process, the wafer can become irreversibly deformed. The severity of these deformations varies from wafer to wafer. This effect may result in displacements and scaling (see Fig. 2.5a, c).

The nature and extent of these overlay defects are unpredictable. Their effects are cumulative and can only be limited to specific areas by efforts that concentrate on specific devices and the manufacturing process. Obviously, the maximum allowable overlay error in a semiconductor process should not be greater than the minimum feature size. It should usually be much lower than this value.

Let us take a look at a typical overlay fault to see how it can affect a layout. To ensure proper electrical connection, contacts in chip fabrication must always be completely covered with metal. Because contacts and metallic interconnect layouts use different photomasks, this "design rule" for full coverage must also be met when overlay errors occur.

How overlay errors are handled in layout design is shown in Fig. 2.6. For creating the layout, an enclosure design rule is specified that requires all structures in the "Contact" layer to be covered by structures in the "Metal1" layer and to overlap on all sides by a minimum value. This is the same as the maximum allowable overlay error, which is the deviation that can occur in the worst-case scenario when all overlay errors are superimposed. This design rule is met by the layout structure in the centre of Fig. 2.6. The right-hand side of the figure depicts a possible fabrication situation. For clarity, the assumed overlay error is exaggerated here.
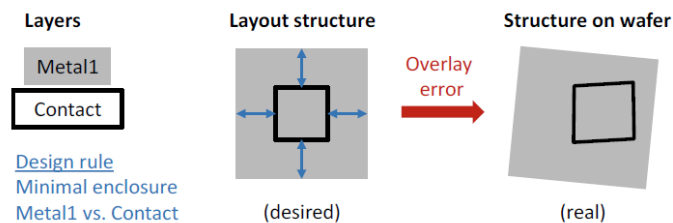
Figure 2.6 Handling of unavoidable overlay errors in physical design. Under all circumstances, contact holes must be completely covered by metal; this necessitates a design rule for a sufficiently large "minimal enclosure" between metal and contact layer that accounts for possible shifting, rotation, scaling, and perspective distortion.

Edge Shifts

Some technology layers experience effects that cause the graphics elements on the processed wafer to be enlarged or shrunk in comparison to the associated graphics elements in the layout. (Please keep in mind that we are not discussing scaling, in which the dimensions of the elements are changed by a certain factor, as with a "Zoom.") The changes caused by these enlarging/shrinking effects are additive: the structure's boundary lines are shifted outwards by a specific value (positive shift) or inwards by a specific value (negative shift) (negative shift). We call this category of imaging errors edge shifts because the individual structures in the layout are typically modelled as polygons (i.e., as geometrical elements bounded by edge strips).

Figure 2.7 depicts a simple example of a process step with a negative edge shift. The structure on the wafer contracts in relation to the element on the photomask (shown with red arrows in the figure on the right). These edge shifts have layer-specific sizes that are defined for each semiconductor process. As a result, the effect can be mitigated

by pre-sizing the photomask geometry. In particular, the prepared layout data are automatically modified in a layout to mask preparation process that is part of the layout post process. If a value k edge shift occurs during the process, the edges of the layout geometries are shifted by a value k before the new data is transferred to the photomask. This operation is shown in the bottom row (b) in Fig. 2.7.
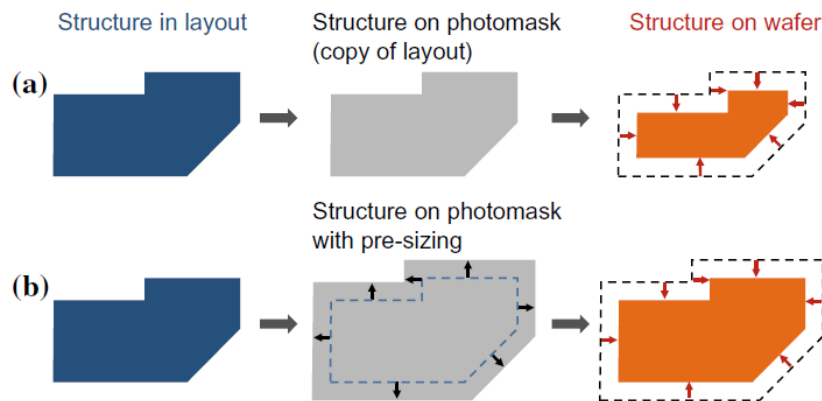
Figure 2.7 Edge shift inwards in the wafer process; a without pre-sizing, b with pre-sizing

Diffraction Effects

Diffraction phenomena occur at the structural edges of the chromium layer on the photomasks due to the wave properties of light, limiting the optical resolution of photolithography. The smaller the feature size (for a constant exposure wavelength), the greater the impact of diffraction effects on imaging accuracy. In Fig. 2.8, we show these effects with an L-shaped layout structure.

The layout element is transferred to the photomask (in grey) unchanged in the top row. As the feature size decreases, the shape of the area exposed in the photoresist (blue) deviates more from the shape of the photomask opening and thus from the desired layout structure (moving from left-to-right in the figure).

The small fillets at the corners are insignificant as long as the wavelength of light is greater than the feature size (see Fig. 2.8, top left). However, when the feature size to wavelength ratio falls below one (so-called sub-wavelength lithography), significant line-end shortening occurs. Corner rounding has also increased significantly (see Fig. 2.8, top center and right).

These diffraction effects can be corrected by slightly enlarging the photomask opening where it is underexposed and slightly shrinking it where it is overexposed. This procedure, known as optical proximity correction, is another example of a preventive measure (OPC).

As long as the diffraction effects are of the same order of magnitude as those shown in the middle of the figure, these corrective measures can be defined using simple rules based on structural shapes. If the line width is roughly the same as the feature size, "hammer heads" are attached to the ends of "thin" lines. Square elements, so-called "serifs", are added (this means more light) to the outer corners, or "punched" (means less light, hence also called "jogs") from the inner corners. These measures, known as *rule-based OPC*, are illustrated in Fig. 2.8, bottom center. In addition, the change in line widths, caused by interference from many parallel lines, can be corrected by rule-based OPC (not shown in Fig. 2.8).

The extent of imaging errors increases as feature sizes shrink in relation to light wavelength (see Fig. 2.8, top right). In this case, rule-based OPC is ineffective because the exposure result is increasingly influenced by surrounding features. Corrections to the photomasks must then be calculated individually for each structure. The algorithms used in this case are based on models that describe wave-optical effects. A model-based OPC result is shown in Fig. 2.8, bottom right.
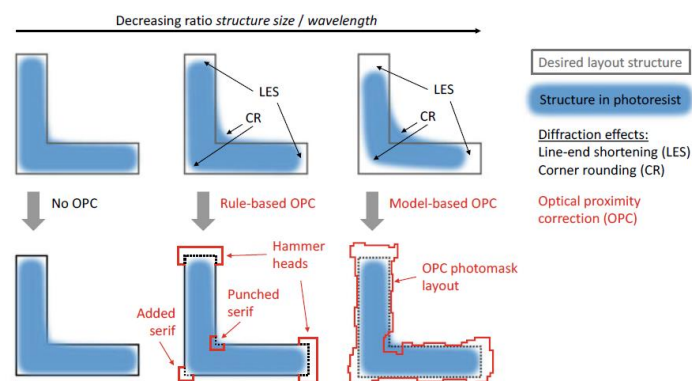
Figure 2.8 Diffraction effects in photolithography (top row) and possible corrective measures using optical proximity correction (OPC, bottom row). The imaging errors increase as the ratio of feature size to optical wavelength decreases (from left to right).

Applying and Structuring Oxide Layers

One of silicon's major advantages over other semiconductors is that it forms a very stable intrinsic oxide: silicon dioxide (SiO_2). Silicon dioxide, which we will refer to as "oxide" in the following for simplicity, has many beneficial properties.

Oxide is a good electrical insulator and a dielectric in capacitive applications. It is mechanically stable and thus appropriate for a strong layer structure. It is simple to create from a process standpoint, and it serves as a masking layer for many process steps, as we will see. It is also visible. Because alignment features can be detected beneath the oxide, this is a useful factor in fabrication. This also enables a wide range of applications, including LEDs (silicon can emit light), solar cells, and photodiodes (light can penetrate silicon from outside). Different processes are available for producing and structuring oxide layers. We will examine these now.

Thermal Oxidation

Thermal oxidation is used to create oxide on the wafer surface. Once formed, an oxide layer can only grow if oxygen atoms diffuse through it until they reach the silicon beneath. As a result, as the thickness of the oxide layer increases, the rate of oxide growth slows. There are two different thermal oxidation processes: *dry oxidation* and *wet oxidation*.

Dry oxidation

Wafers are heated in an oxidation furnace and exposed to pure oxygen (O_2) at temperatures ranging from 1000 to 1200 °C (approximately 2000 °F). The oxide grows slowly and produces good quality oxide with few vacancy defects.

Wet oxidation

In this process, the oxygen first flows through boiling water. The wafer is thus exposed to steam, as well. It is much faster than dry oxidation. It is more difficult to control however and produces a lower quality oxide. This is the first thick oxide layer that is created directly on the silicon.
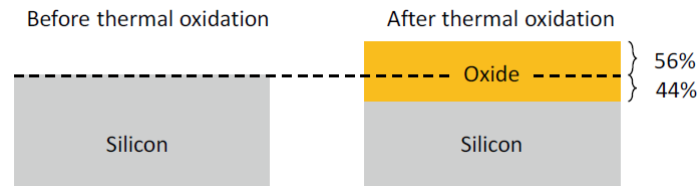
Figure 2.9 Thermal oxidation of silicon (Si) to silicon dioxide (SiO2)

Oxide Structuring by Etching

Etching is a chemical removal process that is frequently repeated for different materials in chip fabrication. It is critical that etchants that remove material be used selectively with respect to the substance, i.e., not etching away other substances (or at least only very slightly).

Figure 2.10 shows how etching can be used to structure an existing oxide layer. A photomask is used to expose and develop an etch-resistant photoresist. There are two types of etching that can be used: wet etching and dry etching.
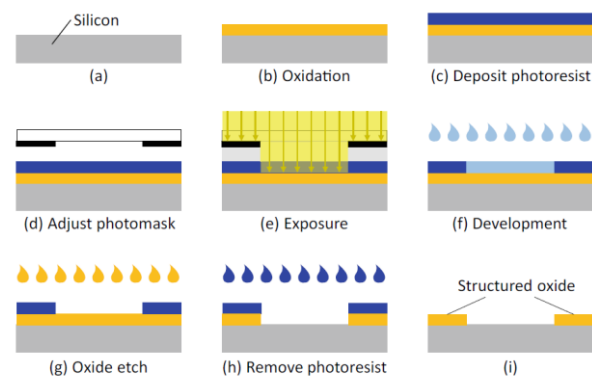
Figure 2.10 Structuring an oxide layer

Wet Etching

A fluid chemical etching agent dissolves and removes oxide during wet etching. This is a straightforward and widely used method, and the fast etching rate can be adjusted. The disadvantage of wet etching is that it is isotropic, meaning that it acts in all directions. This causes undesirable lateral etching beneath the photoresist. Because of these so-called undercuts, the oxide openings are always larger than the photoresist openings. This results in an edge shift (Fig. 2.11, left). The undercut etching rate is slightly lower than the vertical etching rate because the etching agent cannot circulate as easily under the photoresist and is thus more highly saturated. The lateral undercut is typically 80% of the etching depth.

Due to this undercutting effect, wet etching is no longer suitable for imaging typical feature sizes in advanced processes. As a result, in these processes, wet etching is only used to dissolve and remove entire layers.

Dry Etching

Reactive ion etching (RIE) is a popular dry-etching method. In general, the etching agent is ionized and used as a gas plasma. An electrically alternating field causes the ions to oscillate. The field is parallel to the wafer surface. Chemically active ions oscillate in this direction and etch material only vertically. The main advantage of RIE is that there is no edge shift in this process (Fig. 2.11, right).

In this process, the etching effect is a combination of a physical effect (where the material to be etched is bombarded with particles in a specific direction) and a chemical effect, i.e., etching. The RIE process can produce extremely fine structures. Furthermore, the trenches formed can be much deeper than they are wide.
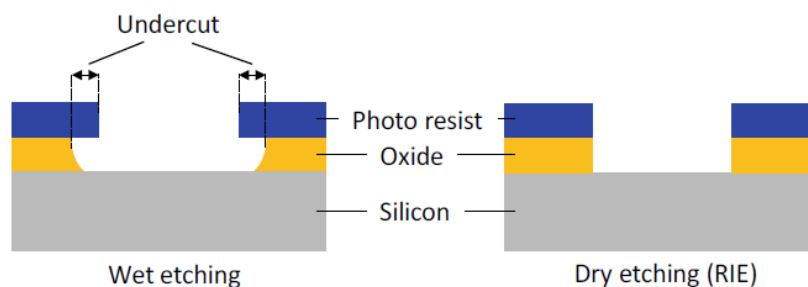
Figure 2.11 A comparison of wet and dry etching (RIE, reactive ion etching).

Example:

CMOS fabrication process

1- **Substrate:** First we choose a substrate as a base for fabrication. For N- well, a P-type silicon substrate is selected.
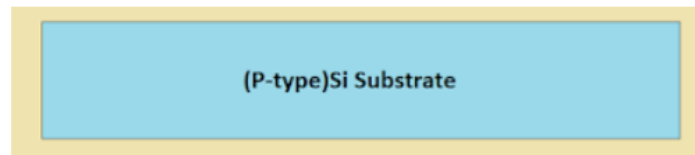

Si substrate p-type

2- **Oxidation:** Creating a protective SiO_2 layer on the wafer surface. It protects the junction from moisture and other atmospheric contaminants. It is used to isolate one device from another.
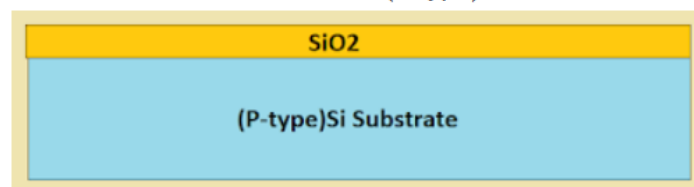

3- Photoresist coating: A photoresist is a light-sensitive material used in several processes, such as photolithography, to form a specific patterned coating on a surface
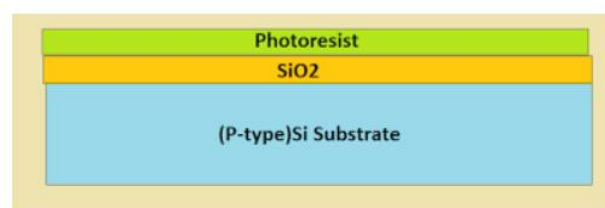
4- **Masking:** The process of photolithography includes masking with a photographic mask and photo etching. a mask with the desired pattern is used as a medium to expose UV (Ultra-Violet) lights. The UV light through the mask reaches the photoresist material. The exposed resist remains on the surface and the unexposed part is removed from the surface.
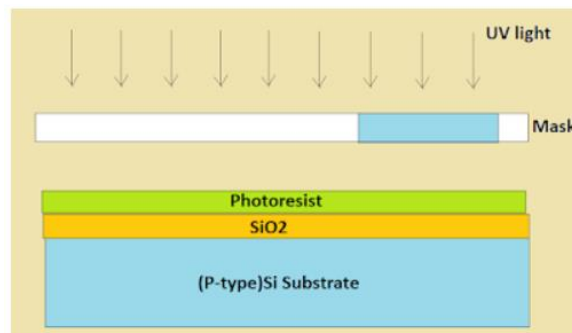
5- **Removal of Unexposed Photoresist:** The mask is removed and the unexposed region of photoresist is dissolved by developing wafer using a chemical such as Trichloroethylene.
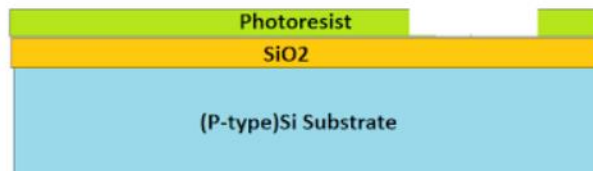

6- **Etching:** The unexposed window is removed from the surface and the regions are etched together to form a clean wafer surface. It removes the oxide from the areas through which dopants are to be diffused.
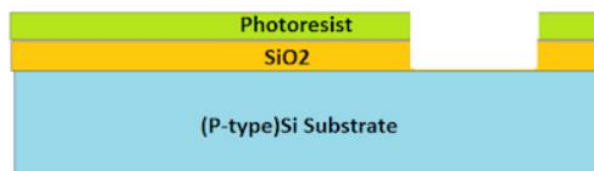

7- **Removal of Whole Photoresist Layer:** During the **etching process**, those portions of SiO2 which are protected by the photoresist layer are not affected. The photoresist mask is now stripped off with a chemical solvent.
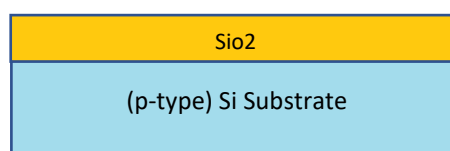
8- **N-WELL formation and SiO_2 removing:** The n-type impurities are diffused into the p-type substrate through the exposed region thus forming an N- well.
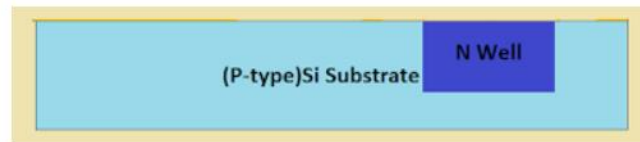
9- **Deposition of Polysilicon:** Chemical Vapor Deposition (CVD) process is used to deposit a very thin layer of gate oxide.
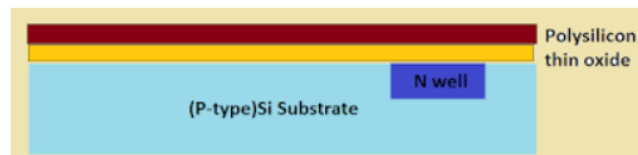

10- **Ion implementation:** it is the process of adding impurities to a silicon wafer.

11- **Metallization:** This step is used for the formation of metal terminals which can provide interconnections. Aluminum is spread on the whole wafer.

12- **Removal of Excess Metal and Formation of Terminals:** In the gaps formed after removal of excess metal terminals are formed for the interconnections.
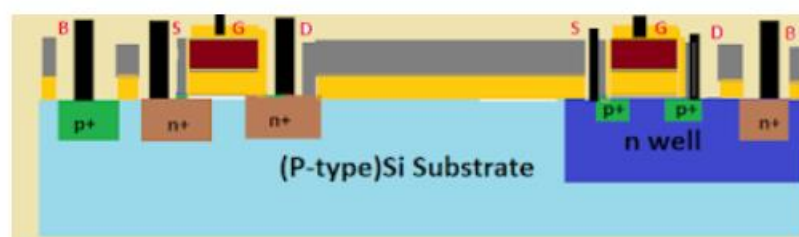
Following photolithography, one of two procedures is usually carried out. One method involves etching thin films, which is used to transfer photoresist patterns to the thin film(s) beneath. The second involves employing ion implantation and photoresist patterns to prevent certain dopants from entering certain areas.

a wafer's surface in certain places. We go through wet chemical etching and dry etching methods for thin-film etching in this part. We also cover chemical mechanical polishing (CMP), a technique used to remove unpattern thin coatings.

Thin Film Etching

Wet etching and dry etching are two commonly used methods for transferring patterns into underlying layers after a photoresist pattern has been created. The most important variables for both wet and dry etching are etching rate (amount of thickness reduced per unit time), selectivity, and degree of anisotropy. Temperature and solution concentration are frequently significant determinants of etch rates. The etch rate ratio of one material to another, as determined by the selectivity equation, is the definition of selectivity (S).

$$S = \frac{R_2}{R_1}$$

where R2 is the desired material's etch rate and R (7.11) is the intended material's planned removal-resisting etch rate for the material beneath, around, or next to it. The rate at which an etchant dissolves material in various directions is expressed mathematically as the degree of anisotropy, A_f.

$$A_f = 1 - \frac{R_l}{R_v}$$

where R_l, is the rate of lateral etch, and R_v, is the rate of vertical etch. Keep in mind that the etchant is entirely anisotropic if Af = 1. The etchant is entirely isotropic, though, if $A_f = 0$. The level of anisotropy in conjunction with photolithography plays a significant role in the resolution that can be achieved. The effects of etch bias (i.e., $d_{film} - d_{mask}$) on the final feature size are shown in Figure 7.13. Dry etch methods are referred to wet etch methods for the submicron features needed in CMOS. This is

because dry etching processes often have a higher level of anisotropy. Metal, semiconductor, and insulator removal are accomplished using wet and dry etching.
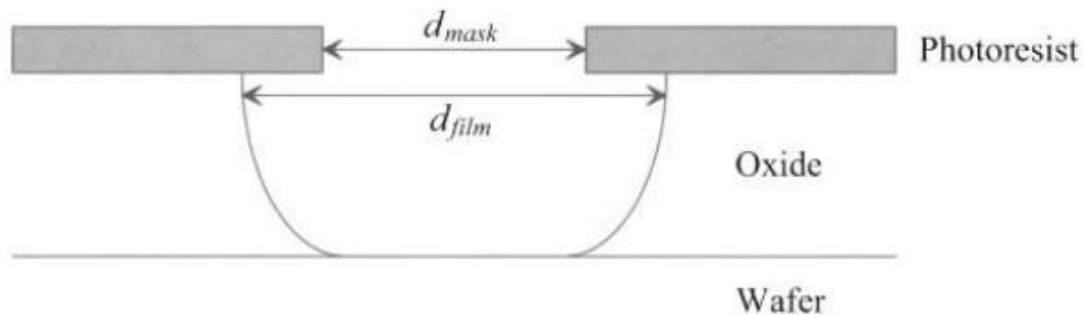

Figure: Picture of a post-etch profile. Observe that the mask opening does not match the manufactured opening in the underlying oxide film due to isotropy in the etch process. Etch-bias refers to the variation between these dimensions.

Deposition

a method of making semiconductors in which a thin coating is used to add electrical properties to a wafer surface. A thin film of a certain chemical is placed to a wafer's surface during the deposition process to give it electrical properties. The procedure entails depositing several molecule or atomic-thick layers of a substance onto a wafer. The degree of homogeneity and thinness of the film define the quality of the semiconductor, making it an essential step in semiconductor production. Because the covering is so thin, sophisticated technology is required.

Physical vapor deposition (PVD) and chemical vapor deposition are the two forms of deposition (CVD). Physical vapor deposition (PVD), which does not involve a chemical reaction, is typically used to deposit metal layers.

Chemical vapor deposition (CVD), on the other hand, entails supplying outside energy to the vapor of particles produced by a chemical reaction of gas. On semiconductors, insulators, and conductors alike, it can be utilized to deposit thin films.

Types of deposition: -

- Thin film deposition
- Chemical-vapor deposition (CVD)
- Low-pressure chemical-vapor deposition
- Plasma-assisted chemical-vapor deposition
- Sputter deposition
- Materials deposited
  - Silicon nitride (Si3N4)
  - Silicon dioxide (SiO2)
  - Aluminum
  - Polysilicon

Chemical Vapor Deposition

Chemical vapor deposition (CVD) is a process by which gases or vapors are chemically reacted, leading to the formation of solids on a substrate. CVD can be used to deposit various materials on a silicon substrate including SiO2, Si3N4, polysilicon, and so on. For instance, if silane gas and oxygen are allowed to react above a silicon substrate, the end product, silicon dioxide, will be deposited as a solid film on the silicon wafer surface. The properties of the CVD oxide layer are not as good as those of a thermally grown oxide, but they are sufficient to allow the layer to act as an electrical insulator.

The advantage of a CVD layer is that the oxide deposits at a faster rate and a lower temperature (below 500°C).

If silane gas alone is used, then a silicon layer will be deposited on the wafer. If the reaction temperature is high enough (above 1000°C), the layer deposited will be a crystalline layer (assuming that there is an exposed crystalline silicon substrate). Such a layer is called an epitaxial layer, and the deposition process is referred to as epitaxy instead of CVD. At lower temperatures, or if the substrate surface is not single-crystal silicon, the atoms will not be able to aligned along the same crystalline direction. Such a layer is called polycrystalline

Sputter Techniques

Sputter deposition is well known for being a method for creating thin films of alloys and complicated materials for use in industry. Its foundation is a discharge of free ions and electrons in a gaseous environment (see Figure 11). Arc discharge, glow discharge, and dark discharge are the three types of discharges that can be recognized. In addition to their luminescence, they can be identified by their breakdown voltage, current density, and current-voltage characteristic (Figure 11(b)). These key features are influenced by the electrode material, process gas, and geometry of the electrodes and deposition chamber.
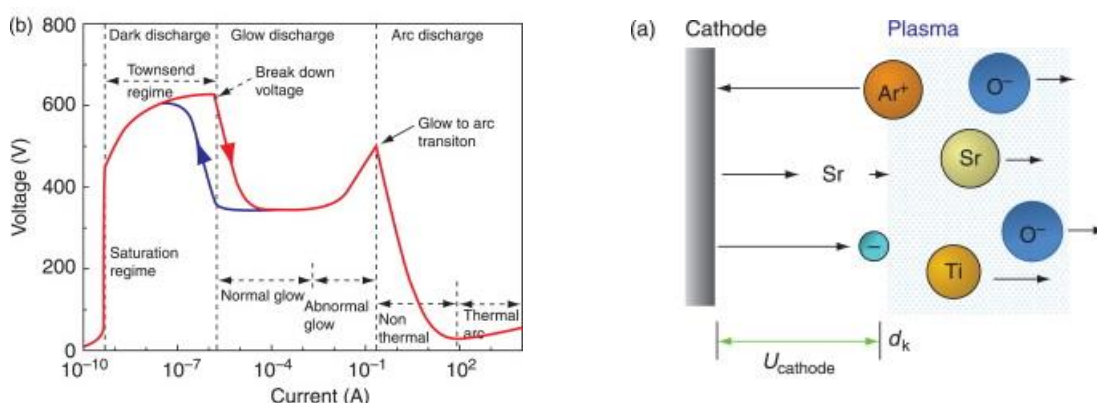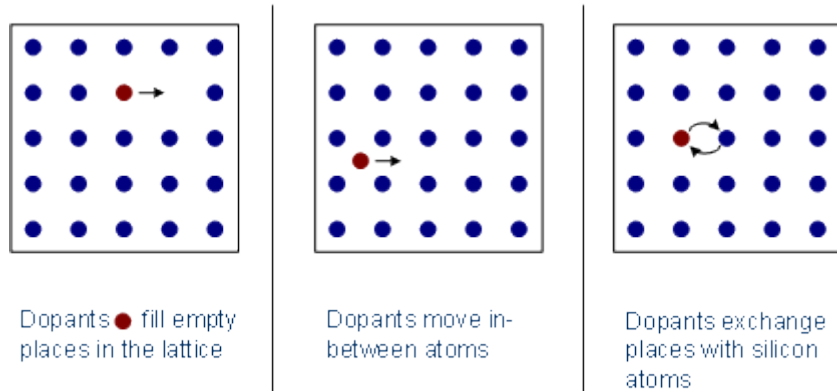

Doping

Doping is the intentional addition of impurities to a semiconductor crystal in order to change the conductivity of the crystal due to a lack or surplus of electrons. This article explains the partial doping of silicon, which differs from wafer manufacture doping, which involves doping the complete wafer. Diffusion and ion implantation are two methods that can be used to introduce foreign materials (or alloy).

Diffusion

Molecular diffusion, also known as just diffusion, is the net movement of molecules by random molecular motion from one area of higher concentration to another. Diffusion causes materials to gradually mix together. To give an example, after a specific amount of time, an ink drop in a glass of water spreads out equally. A solid lattice of atoms is present in a silicon crystal, through which the dopant must pass. There are several ways to accomplish this:

- **empty space diffusion:** Even in flawless single crystals, there are always gaps in the crystal lattice that impurity atoms can fill.
- **inter lattice diffusion:** The impurity atoms in the crystal lattice migrate between the silicon atoms.
- **changing of places**: The silicon atoms trade places with the impurity atoms in the crystal lattice.

Dopants ● fill empty places in the lattice

Dopants move in-between atoms

Dopants exchange places with silicon atoms

Metallization

The process of depositing a metal layer on a metallic or non-metallic surface is known as "metallization." The coating may be made of aluminum, zinc, or silver. Aluminum is the metal used in CMOS production to protect the surface from external environmental elements like dust, air, water, etc.

deposition

Metal film
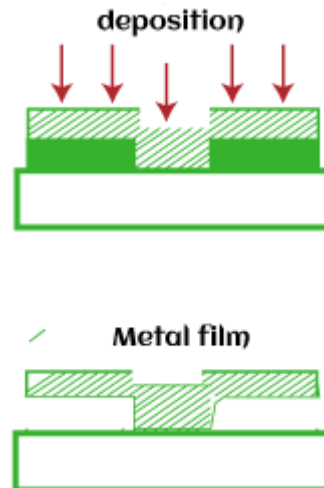
Metallization is also used to interconnect many components that form an Integrated circuit (IC). The many components that make up an integrated circuit are also connected using metallization. Components include things like relays, transistors, capacitors, and resistors. According to the previous discussion, the metal layer is initially placed on the silicon wafer's surface. After that, as illustrated below, the necessary pattern or region for connecting is carved.

etch

Etched
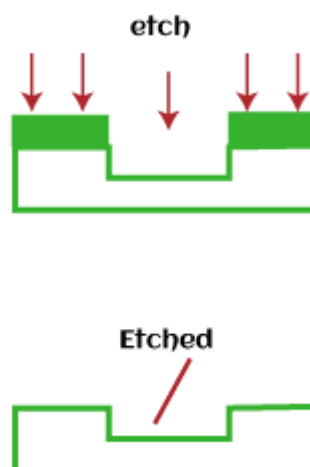
The metallization process uses chamber to apply metal layer. The wafer is placed inside the chamber, which coats the entire surface inside it. The thickness of the metal layer can vary depending on the requirements.

Conductive films enable electrical connection between devices and the environment. And here are Desired properties of the metallization for integrated circuits: -

1. Low resistivity
2. Easy to form
3. Easy to etch for pattern generation
4. Should be stable in oxidizing ambientes and oxidizable
5. Mechanical stability, good adherence, and low stress
6. Surface smoothness
7. Stability throughout processing, including high temperature sinter, dry or wet oxidation, gettering, phosphorus glass (or any other materials) passivation, and metallization
8. No reaction with final metals
9. Should not contaminate devices, wafers, or working apparatus
10. Good device characteristics and lifetimes
11. For window contacts - low contact resistance, minimal junction penetration, and low electromigration

In addition to offering contacts, gates, and connections, metallization serves two crucial functions. The resistance and capacitance of the connecting runners serve to regulate the circuit speed. The flatband voltage V_{FB} is also determined by it:

$$V_{FB} = \phi_m - \phi_s \equiv \phi_{ms}$$

where ϕ_m represents the work function of the gate metallization and ϕ_s represents the work function of the semiconductor. The voltage known as V_{FB} is necessary to maintain a flatband state in the semiconductor by balancing the work function difference between the metal and semiconductor. Thus, V_{FB} adds to the threshold voltage V_T , which is the voltage needed at the gate metal to achieve conduction between the source and drain regions.
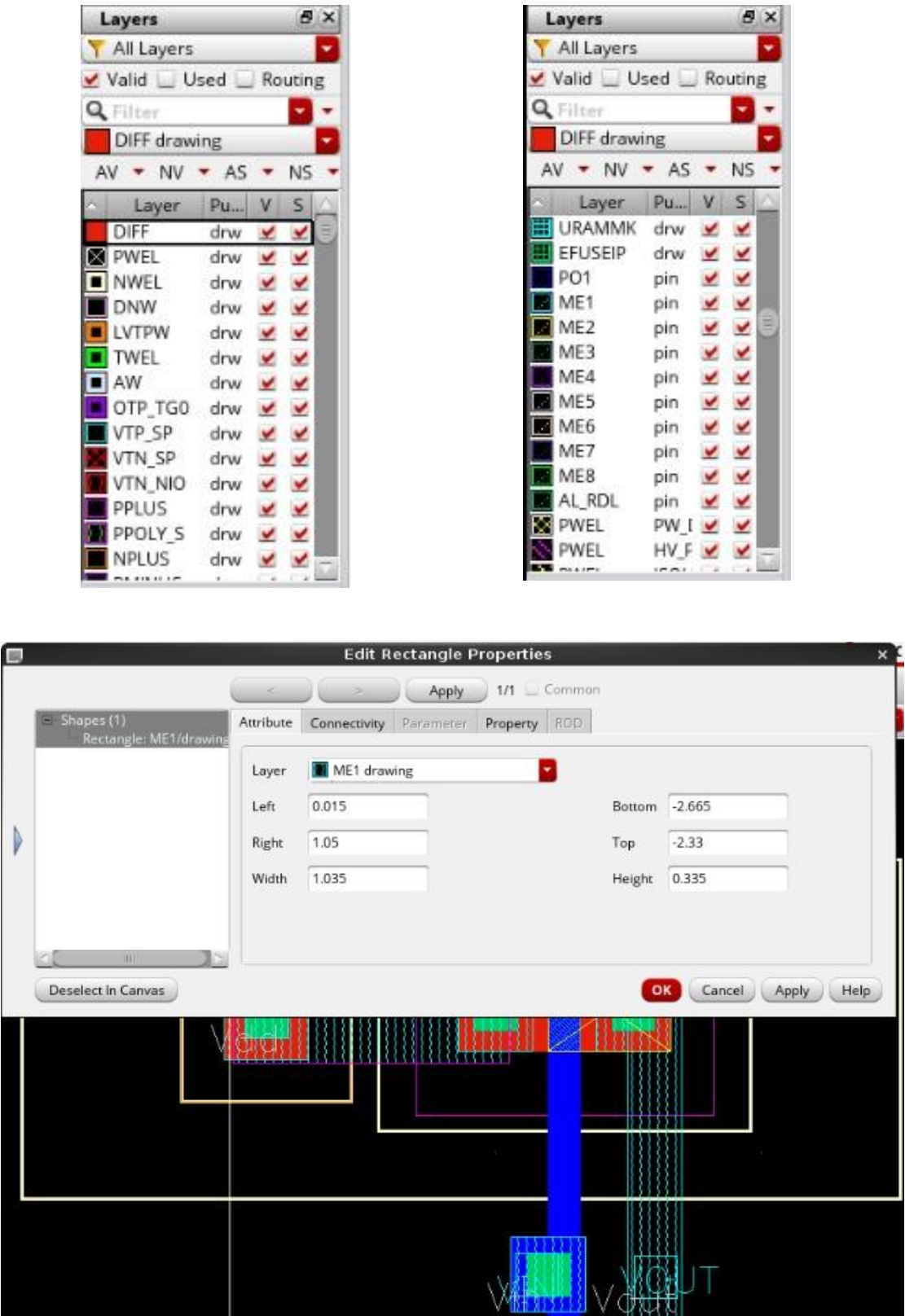
Signal Routing Matching:
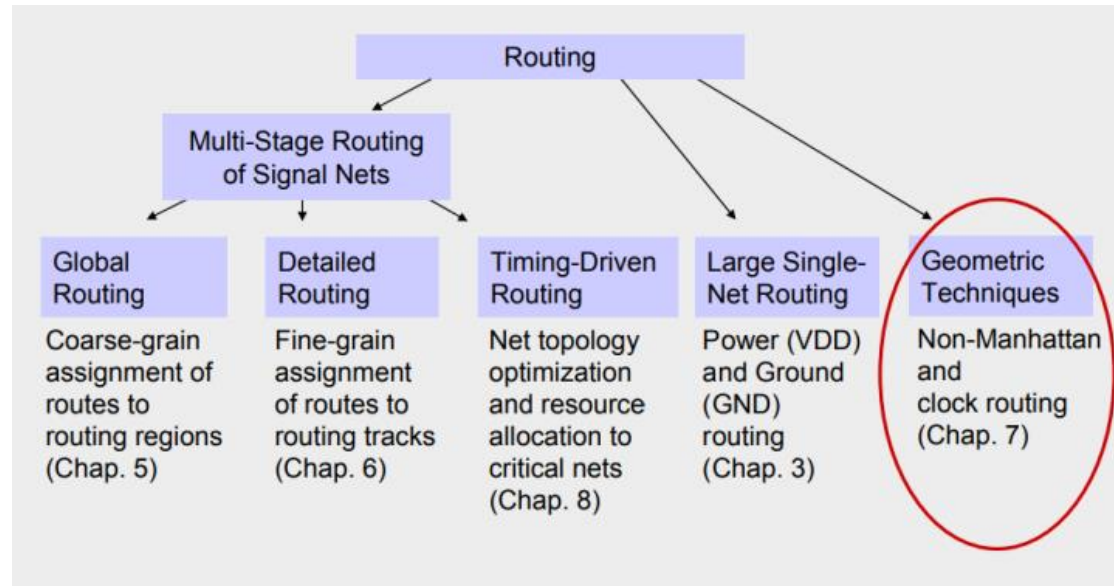
Routing examples
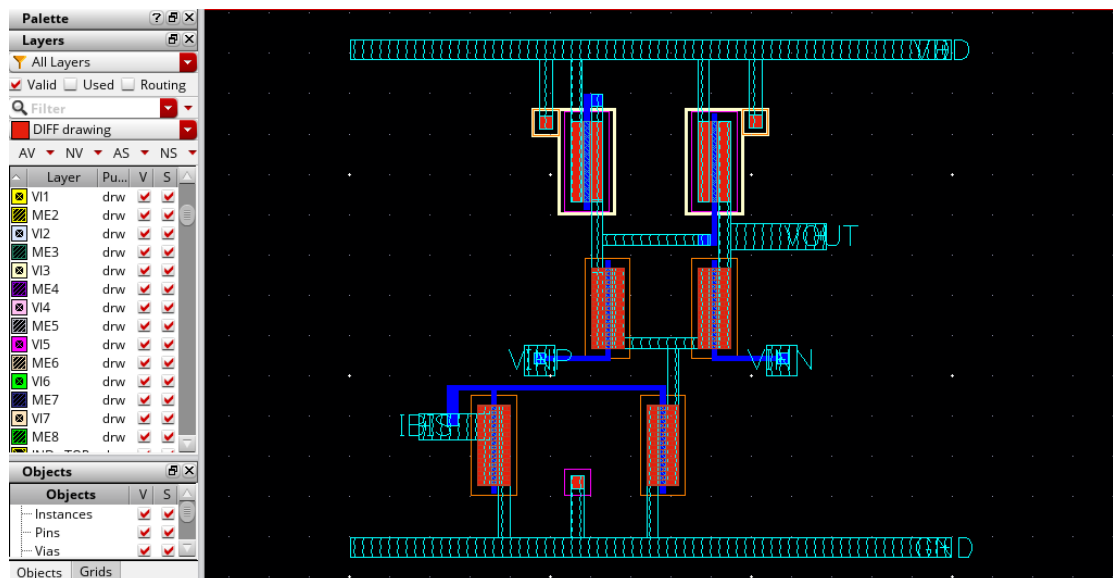
Metals used in the example:

Physical Design Routing Process:
Metal routing:

a) It involves generating metal wires to connect the pins of same signal while obeying manufacturing design rules.

b) Before routing is performed on the design, cell placement has to be carried out wherein the cells used in the design are placed.

c) The connections between the pins of the cells pertaining to same signal need to be made. At the time of placement, there are only logical connections between these pins.

d) The physical connections are made by routing. More generally speaking, routing is to locate a set of wires in routing space so as to connect all the nets in the netlist taking into consideration routing channels' capacities, wire widths and crossings etc.

e) The objective of routing is to minimize total wire length and number of vias and that each net meets its timing budget. The tools that perform routing are termed as routers.

f) You typically provide them with a placed netlist along with list of timing critical nets. These tools, in turn, provide you with the geometry of all the nets in the design.

Routing stages:

1. Global Routing

2. Track Assignment

3. Detail Routing

4. Search and repair

Challenges and Solutions in Analog Layout Design

Analog layout design is a critical step in the development of analog circuits, as it involves creating a physical representation of the circuit that can be fabricated and tested. However, this process can be complex and challenging, as a range of factors can impact the performance of the circuit. In this chapter, we will explore some of the common challenges that can arise in analog layout design, including matching and symmetry, parasitics , crosstalk and electromigration. We will also discuss strategies and techniques for addressing these challenges, including layout by understanding these challenges and approaches to addressing them, designers can create analog layouts that meet the required performance specifications and are manufacturable.

One of the most prominent challenges in analog layout Mismatch problem

Mismatching in (IC) integrated circuits is generally of two types:

- **Random mismatches:** the error resulting from that one can't be identified or controlled while implementing the layout, it will happen during the fabrication and the reasons behind that are the non-uniform etch rate, the doping and finally the wafer itself. errors by variations in the thickness of oxide layers, variations in the doping level of transistors, and variations in the width and length of metal interconnects. These variations can lead to differences in the performance of individual components, even when the components are designed to be identical

- .**Systematic mismatches :** which is due to non-uniform thermal distribution during fabrication process, it can be solved by proper layout techniques, best device matching so as to be as close as possible to the needed performance. Systematic mismatches are caused by: Process biases, Mechanical stress, Temperature gradients, and Polysilicon etch rates, etc.

What is matching?

Two integrated device or more have same desired value are exposed to the same conditions whether pressure, temperature, current etc.

Why needs matching?

What if your headphones started playing songs with more volume in left than the right? Sounds bad right. So, the voice sensitivity needs to be matched for both left & right side by same.

Actually, due to process variation in fabricating the transistors (like non uniform doping, oxide thickness, etching etc.), no 2 transistors have exactly the same electrical properties. By applying techniques such as Interdigitzer & Common centroid, it is possible to match transistors reasonably

In a sensitive amplifier (differential pair), matching would greatly enhance the performance by improving the ability to reject noise in its differential input.

How to do matching?

We can achieve matching when avoid mismatch sources and use relative accuracy Mismatch is a deviation of values of integrated component after fabrication desired value used for simulation Let's assume we want design low pass filter in 1MHZ

$\omega = \frac{1}{RC}$ in simulation design $R = 1к\Omega$, $C = 1nf$

after fabrication process $R = 1.3к\Omega$, $C = 1.3nf$, $\omega = 0.6MHZ$

We will identify the most prominent sources of mismatch

- Process Variations
- Placement and Gradient

Mismatch (Process variations)

Process variation during fabrication process limits accuracy and desired performance of analog circuit:

- **mask production and alignment**
- **latera diffusion**: Diffusion widens implanted region can affect doping of neighboring devices. <u>Solution:</u> Increase distance and use dummy structures that affect all transistors the same.
- **over etching**: Poly silicon does not always etch uniformly.
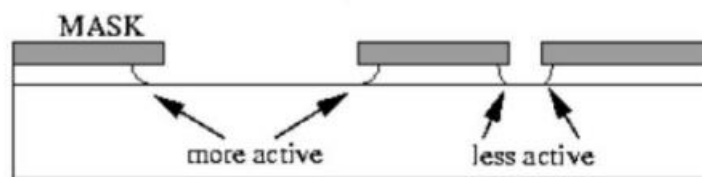

- Large openings etch faster than small openings in mask. <u>Solution:</u> is to use dummy structures.

And all of this effects on Dimensions of device $\frac{W}{L}$ Ratio and consequently the threshold voltage value will change with it.
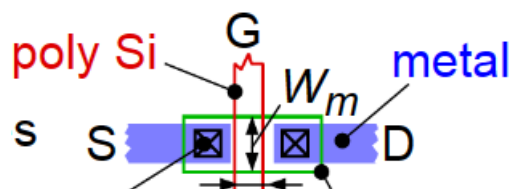
- Mismatch (Placement and Gradient)
- Orientation
- Stress gradient: occurs mainly from wafer dicing where stress is highest at edge of the chip, Packaging can cause stress in chip. **Solution:** Keep critical matched devices in center of chip, avoid using corners for matched devices
- Temperature gradient We use matching so that the devices are exposed to the same thermal effect
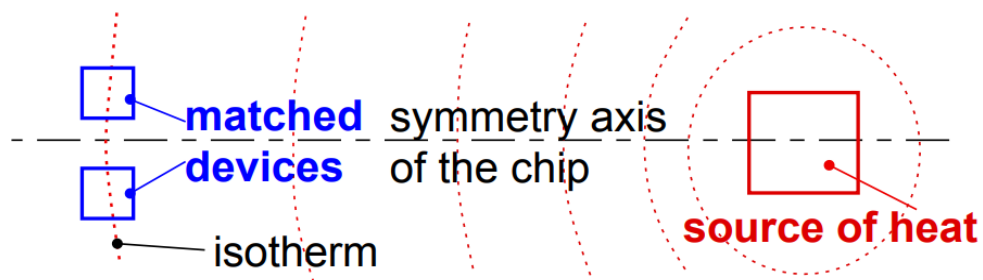
The devices have the same orientation, are close to each other, and are interleaved as neighbors to each other. We have satisfied the three main rules, When these components are etched, the ones in the middle of the block see very different conditions during processing than the ones on the ends. The resistors on the ends might etch more, making them slightly narrower than the ones in the middle

Figure 4 1Edges of blocks etch differently than middles of blocks.

To reduce the impact of process variations and parasitics (will be illustrated) and improve performance we use dummy components.
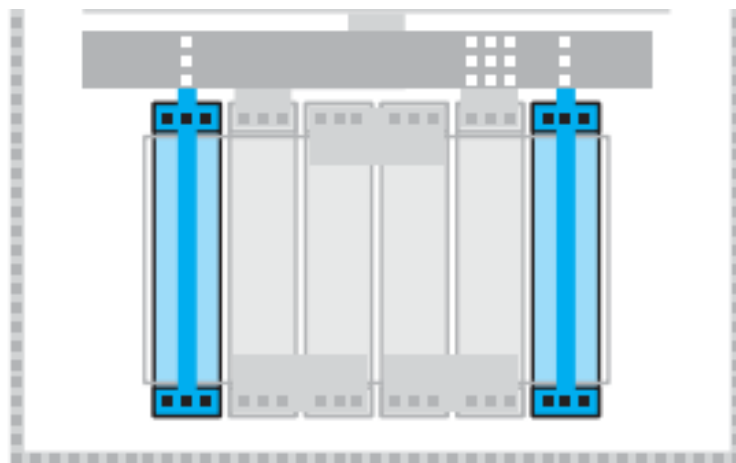

Figure 4: Two dummies to help the four circuit resistors etch

Dummy components, also known as "fillers," are non-functional components that are added to the layout of an integrated circuit to improve performance and reduce the impact of process variations and parasitics. They are typically made of the same materials as the functional components and are designed to match the physical characteristics of the functional components. There are several types of dummy components that can be used in layout design, including dummy resistors, dummy capacitors, dummy transistors, These components are strategically placed in the layout to achieve specific goals, such as matching the resistance . Dummy resistors are often used in analog layout design to match the resistance of functional resistors and reduce the impact of process variations. They are typically placed in parallel with the functional resistors and have the same dimensions and materials as the functional resistors. Dummy transistors are non-functional transistor structures that are added to the layout to match the parasitic capacitance and resistance of functional transistors and reduce the impact of process variations. They are typically placed in parallel with the functional transistors and have the same dimensions and materials as the functional transistors. designers can identify the optimal placement and number of dummy components to achieve optimal performance. Notice we shorted the dummy resistors together and tied them to ground or (VDD). They are not part of the circuit. They are only there to give the outer two circuit resistors similar conditions for etching as the inner two circuit resistors.

dummies are used to balance the effects on the lateral transistors, they may be used as well to shield all around the devices in smaller and more sensitive technologies.

Before talking about matching techniques let's see the Rules for MOS transistor matching:

Rules for optimum matching:

1- **Minimum distance**
2- **Same surroundings (same neighborhood)**

When we see in this figure of two transistor, we see that the transistors have been placed directly next to each other. Good in terms of distance however, if you continue thinking about these two transistors, you notice another problem.

For a CMOS transistor, the parameters that most affect the characteristics of the transistor are the gate length (L)and the gate width(W). Some etches used in processing etch preferentially in one direction. That is the problem. One device is placed sideways. What etching errors occur in one transistor's width, will occur in the other transistor's length.
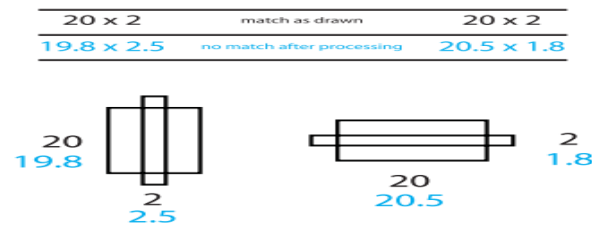

After etching process, we found one device width equal 20.5 and other 19.8

3- Same orientation on the chip

If we follow these three basic rules, throughout all of our layout, we are guaranteed a certain amount of good matching, plus the advantage of better device performance. There will be times when trying to keep all our transistors, resistors, and capacitors in the same orientation makes your layout very difficult in this case can split up and reshape our devices, there will be times when we cannot split our devices sensibly, or we are not allowed to. In these cases, how much we know about what the circuit is doing can help our layout. Find out what components are the least important in the circuit and maybe, just maybe they can be rotated to make your layout smaller.

For example, you might have a problem device that just will not fit well in our layout. If you think the device is non-critical, it may be a good to rotate. Go to our circuit designer and ask, "Is it ok to rotate this transistor?" communicate with our circuit designer help us to find suitable solutions.

Orientation

Current flow direction
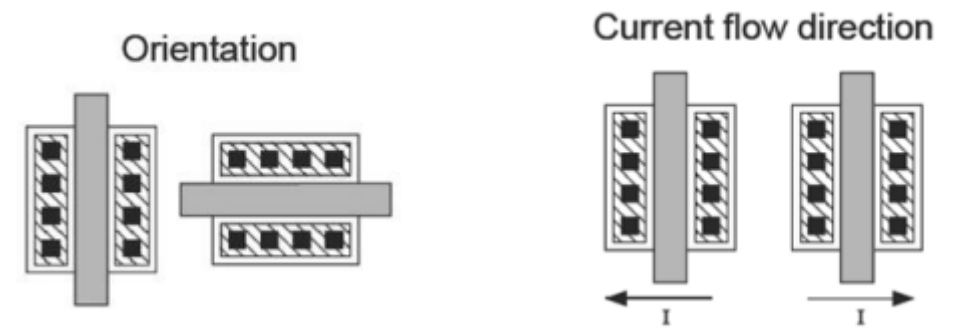
4- choose appropriate unit to be used in matching

Sometimes we have more than two devices that must match each other. There might be five or six devices, all needing to match.

First step we can use is placing the resistors as close together as we can, second step is to keep them in the same orientation, third step is called a root component I mean one resistor from which you will make all the others. Using a root component, if the resistors are all the same size, all the same shape, and all the same orientation, and they are all close to each other, you get very good matching. If the resistors over-etch, they all over-etch the same way and still match each other.

For example
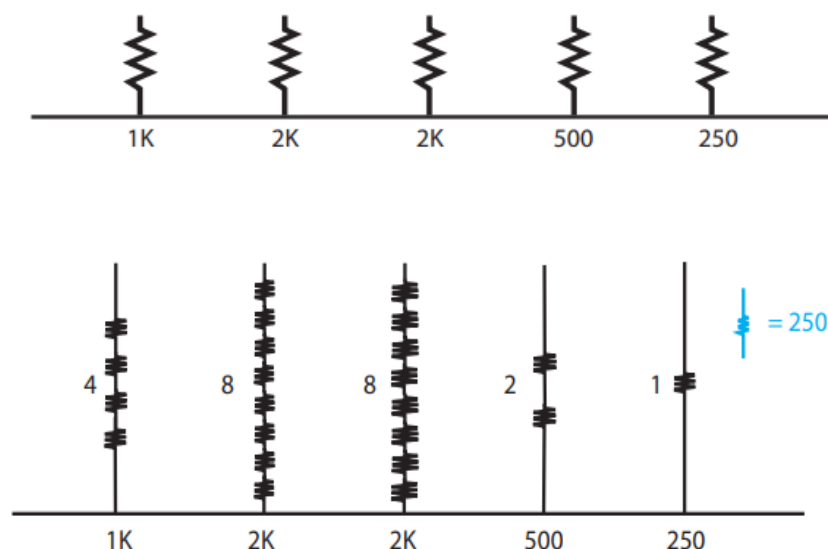


if use 250Ω It is good but in this example, we have some large resistors, like the $2K\Omega$ resistor, for instance. Consider this, contacts on a resistor are typically quite variable and using a 250Ω resistor as the root component could have a

significant portion of its resistance made up from contact resistance. The contact resistance could then create a significant amount of the total resistance on your larger resistors, which have eight times the number of contacts.

So, we Choose a middle value for our root component:
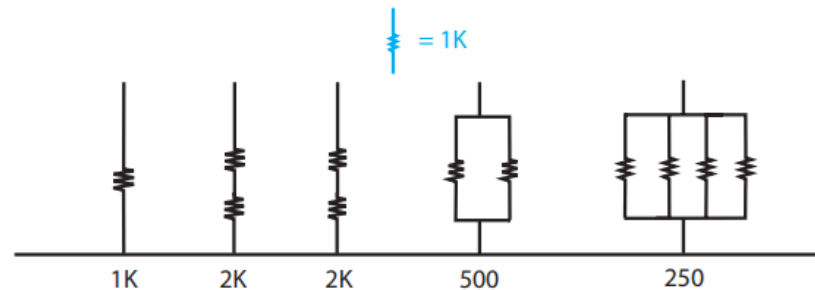
If use 1k better way to use the root component strategy is to pick a medium value. Let's pick one from our previous example. Let's choose 1K as our root resistor. The 2K's would each be two resistors in series. The 500-ohm is two resistors in parallel, and the 250 would be four resistors in parallel. We have made all our required values based on a 1K resistor. We used both series and parallel arrangements.

The root device method can be used with any type of device, not only resistors. The same issues.

Summery Rules for optimum matching:

Devices to be matched should have:

1. Same structure
2. Same temperature
3. Same shape and same size
4. Minimum distance,
5. use relative accuracy not absolute(true) accuracy.
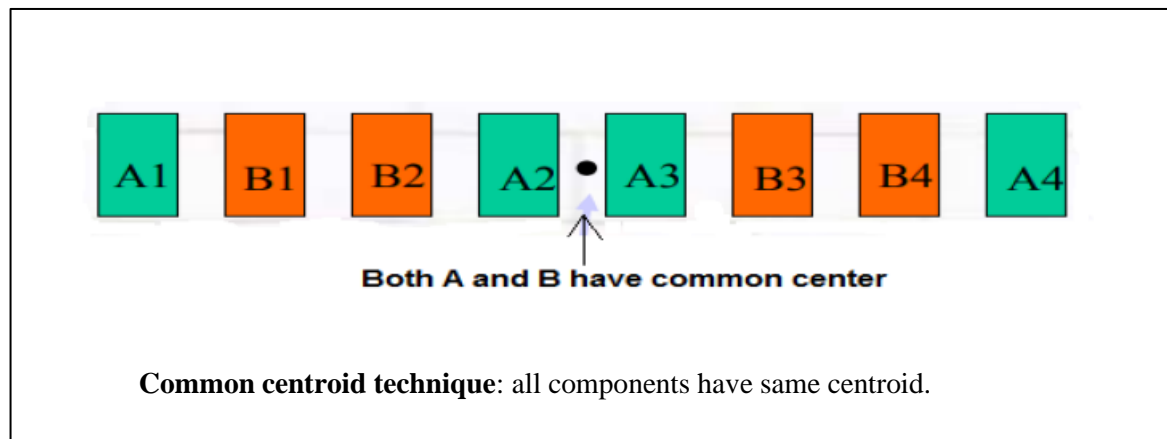6. choose appropriate unit size.

How the matching method is actually applied?

There are two common ways:
1. **inter-digitation**
2. **common centroid**

Let's say we have two devices A,B (A and B can be anything likes transistor, resistor, and capacitor) and split A and B into 4 small multiplier A1-A4 and B1-B4.

Center of **A2B2** and **A3B3**

Center of **A1B1** and **A2B2**

Center of **A3B3** and **A3B4**

Inter-digitization technique: Placing alternate components.

Both A and B have common center

Common centroid technique: all components have same centroid.
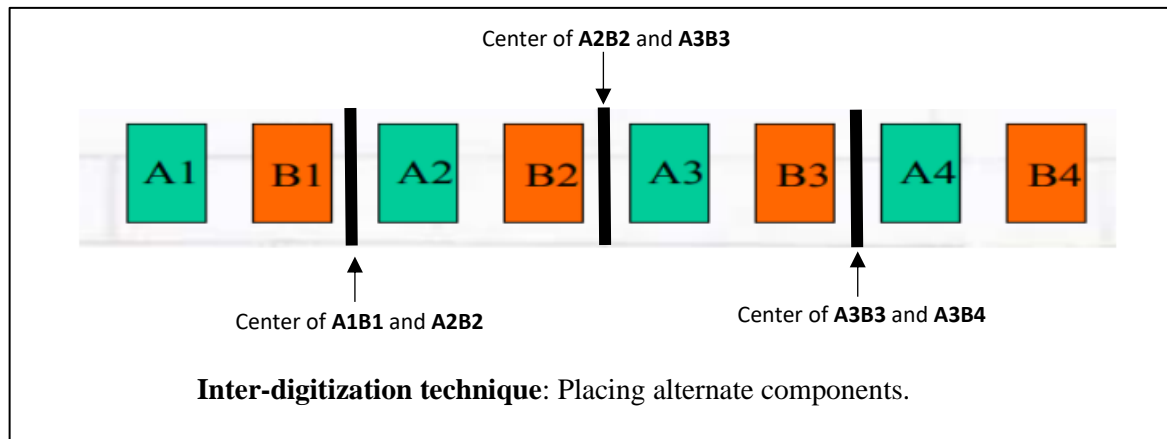
Example of the two way
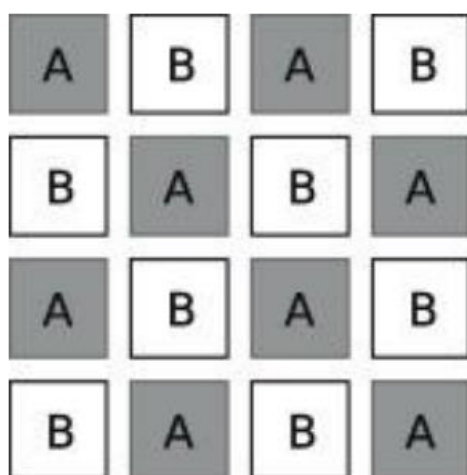
Inter-digitation
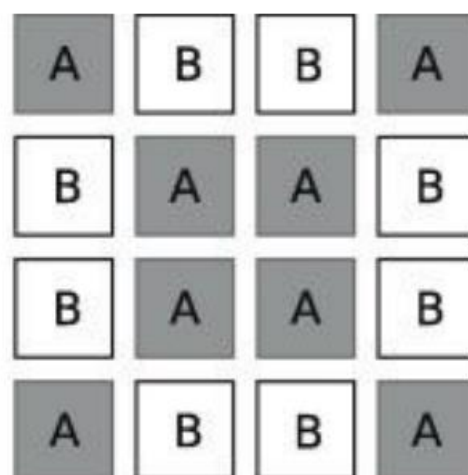
common centroid

Chapter 2

Chapter 3

VGA Circuit Layout Considerations

Chapter 3

Chapter 4

Simulation and Verification of VGA Circuit Layout

Physical Verification

After the physical design is complete the layout must be fully verified to ensure correct electrical and logical functionality (correctness of a physical layout design with respect to the design rules and specifications), It is an essential step in the semiconductor manufacturing process, as it ensures that the final product will function correctly and meet the desired performance criteria. The physical verification process involves checking various parameters such as the width and spacing of metal and poly lines, the alignment of cells, and the presence of shorts and opens. The verification process is performed using various tools and techniques such as:

- **Design rule checking (DRC):** verifies that the layout meets all technology-imposed constraints. These rules include the following categories: -
    - Typical geometrical design rules that require minimum or maximum values for widths, spacings, extensions, intrusions, and enclosures for layout polygons. These rules make sure that the layout structures can be correctly generated on the silicon due to process accuracy
- **Layout versus schematic (LVS)**: is checking that the design is connected correctly. The schematic is the reference circuit and the layout is checked against it. In principle, the following is verified:
    - Electrical connectivity of all signals, including input, output, and power signals to their corresponding devices
    - Device sizes: transistor width and length, resistor sizes, capacitor sizes.
    - Identification of extra components and signals that have not been included in the schematic; floating nodes would be an example of this. The last item overlaps into the items checked in the electrical rules check, which is described previous

- **Parasitic extraction (PEX):** is extracting the parasitic effects of interconnects and devices in a chip. Parasitic effects are unwanted electrical characteristics that arise due to the physical properties of the components used in the chip design. These effects can cause delays, noise, and other performance issues that can impact the functionality of the chip.

Chapter 4

Parasitics Extraction (PEX):

Introduction: -

Reliability of electronic circuits, which has always been an important issue, is becoming an increasing concern due to the ongoing downscaling of the structural dimensions and the continuous increase in performance requirements. (Lienig & Scheible, 2020).

All integrated circuits contain electrical elements not required for their operation. These include reverse-biased isolated junctions, and resistances and capacitances between various diffusions and depositions. The circuit does not benefit from the presence of these parasitic components, but they can negatively affect its operation sometimes.

Nothing in an integrated circuit operates perfectly. An IC is built of layers. You have metals running over other metals. You have transistors next to other transistors. You have transistors built in substrates. Whenever you introduce two different materials like this, you end up creating extra capacitances. It's like we deliberately placed lots of tiny capacitors all over our circuit. And worst of all, we cannot get rid of them. (Saint & Saint, 2002).

Sensitive circuitry may be subjected to noise, for instance, through capacitive coupling. This section will talk about parasitic devices that cause junctions, which are typically reverse-biased, to become forward biassed. When the forward bias of these junctions is applied, current starts to flow between circuit nodes that are typically kept apart. These leakages may only cause minor parametric shifts if these currents are small and the circuit is relatively insensitive to their presence. The circuit's functionality may be catastrophically damaged by larger currents. Even after the triggering event has been removed, the malfunctioning circuit may latch up and continue to act improperly. Because latchup generates excessive power dissipation and consequent overheating, an integrated circuit may actually be physically destroyed. Even if the circuit does not self-destruct, it can only be brought back to normal operation by cutting the power. (Hastings, 2001, p. 139)

Chapter 4

Chapter 5

Conclusion and Future Work

Conclusion and Future Work

5.1. Conclusion

Variable gain amplifiers (VGAs) are important building blocks in many analog circuits such as wireless communication systems, audio amplifiers, and instrumentation. The design of VGAs using analog layout techniques is a complex task that requires careful consideration of various factors, including signal quality, power consumption, and area. In this article, we will discuss some potential areas for future improvement in the design of VGAs using analog layout techniques.

The design of VGAs using analog layout techniques is a challenging task that requires a deep understanding of the underlying circuit topology and layout rules. Improving linearity and noise performance, reducing power consumption, increasing bandwidth, integrating with other circuits, optimizing the layout for manufacturability, and using advanced packaging techniques are potential areas for future improvement in the design of VGAs. These improvements will enable the development of more advanced and efficient analog circuits for a wide range of applications.

5.2 Recommendations for Future work

- **Improving linearity and noise performance**
  Linearity and noise performance are critical factors in the design of VGAs. Nonlinearities in the gain response of VGAs can lead to distortion in the amplified signal, while noise can degrade the signal-to-noise ratio (SNR) and reduce the dynamic range. To improve linearity and noise performance, advanced circuit topologies and layout techniques can be used. For example, folded cascode amplifiers can provide high linearity, while common-mode feedback circuits can reduce noise.

- **Reducing power consumption**

  Power consumption is a critical factor in many applications, and reducing power consumption can extend the battery life of portable devices. Low-voltage and low-power circuit design techniques, such as sub-threshold operation and dynamic voltage scaling, can be used to reduce power consumption.

- **Increasing bandwidth**

  The bandwidth of VGAs is limited by various factors, such as the gain-bandwidth product of the amplifier and parasitic capacitances. High-speed amplifiers and optimizing the layout to reduce parasitic capacitances can increase the bandwidth of VGAs.

- **Integrating VGAs with other circuits**

  Integrating VGAs with other circuits, such as mixers, filters, and frequency synthesizers, can create more advanced systems that provide additional functionality. For example, variable gain mixers are commonly used in RF applications, while variable gain filters are used in audio applications.

- **Optimizing the layout for manufacturability**

  The manufacturability of VGAs can be improved by optimizing the layout for design-for-manufacturability (DFM) and design-for-yield (DFY). Layout pattern matching and lithography simulation can be used to minimize process variations and improve yield.

- **Using advanced packaging techniques**

  Advanced packaging techniques, such as flip-chip and wafer-level packaging, can improve the performance and reliability of VGAs. Flip-chip packaging can reduce parasitic effects and improve heat dissipation, while wafer-level packaging can improve the integration density and reduce the form factor.

Chapter 5

References

1. Alan Hastings, "The Art of Analog Layout", Second Edition, Pearson Prentice Hall, New Jersey, 2006, ISBN: 978-0-13-146410-0.

2. Jürgen Scheible, "Constraint-Driven Design – Eine Wegskizze zum Designflow der nächsten Generation", Proc. of ANALOG 2008, pp. 153–158, VDE Verlag, Berlin, Offenbach, Apr. 2008, ISBN: 978-3800730834.

3. M. A. El-Moursy, M. I. Elmasry, and M. A. Awad, "SAED32: A standard cell library for 0.35 μm CMOS technology," Proceedings of the 17th International Conference on Microelectronics (ICM), 2005, pp. 1-4.

4. [1] M. A. Shayan, S. M. R. Hasan, and M. A. Matin, "A Comparative Study of Routing Strategies for Analog Layout Design," in 2016 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 2016, pp. 1-6.

5. 1 J. Saint and C. Saint, IC Mask Design, McGraw-Hill, 2002. (Shubham, 2023).

6. https://semiconductor.samsung.com/us/support/tools-resources/dictionary/semiconductor-glossary-deposition/

7. https://www.sciencedirect.com/topics/chemical-engineering/low-pressure-chemical-vapor-deposition

8. EE501Lecture1Process

9. https://www.sciencedirect.com/topics/materials-science/plasma-enhanced-chemical-vapor-deposition

10. https://www.halbleiter.org/en/waferfabrication/doping/

11. CMOS Circuit Design Layout and Simulation 3rd Edition

12. https://www.javatpoint.com/ic-fabrication-process Clein, Dan. CMOS IC layout: concepts, methodologies, and tools. Elsevier, 1999. [1]

13. Haseloff, Eilhard. "Latch-up, ESD, and other Phenomena." Texas Instrument application report (2000). (http://www.ti.com/lit/pdf/SLYA014) [2]

14.

Chapter 6