Basic Fabrication Process Steps and CMOS Layout

In this lecture, you will learn about

- Basic process steps used in fabrication.
- Brief usage of topography steps like lithography, etching and deposition.
- Brief usage of doping/thermal steps like oxidation, diffusion and ion implantation.
- Evolution of mask layout and design rules.
- Circuit implementation, stick diagram and full layout of various circuits.

Lithography

- MOS device structure shown earlier is implemented in silicon with help of layout masks. The layout has mask names that usually indicate either a function or a material layer.

- These geometrical patterns for each layer on the layout have to be transferred to Silicon. This is achieved by the lithography step in processing.

- The basic principle is to use a photosensitive material (X ray sensitive for X-ray lithography, e-beam sensitive for e-beam lithography) for exposing the layout geometries.

Lithography

- The photosensitive material called resist is deposited on Silicon Wafer which possibly has one or more materials grown or deposited on it as shown.

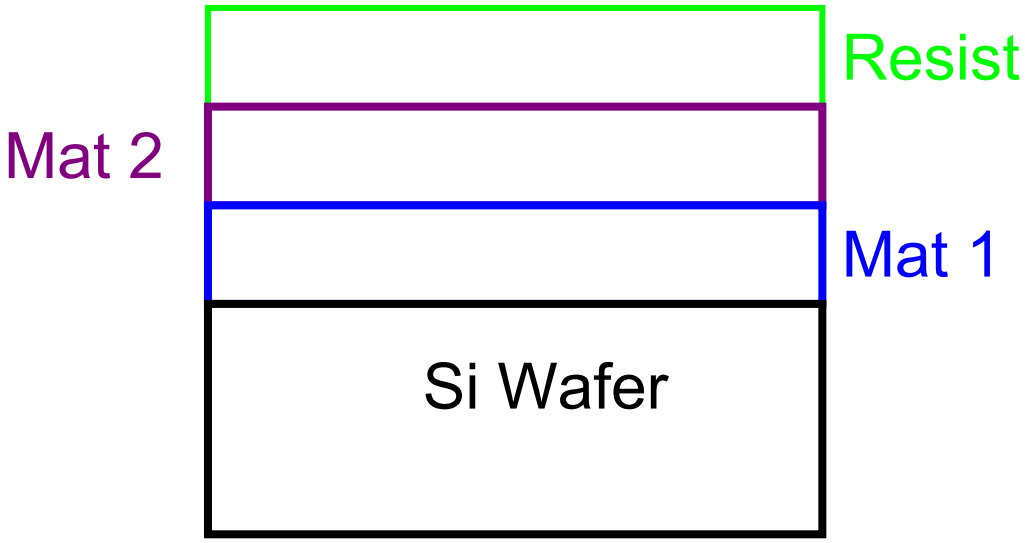


Figure: Silicon Cross Section after resist deposition

Lithography

- The resist is deposited on Silicon wafer by spinning the wafer at high speed with droplets of dissolved resist. The solvent is then evaporated by baking the wafer at about 50° - 60° C (Soft bake). The resist is typically 0.1-1μm thick. Then ultraviolet light is shone through the mask plate as shown.

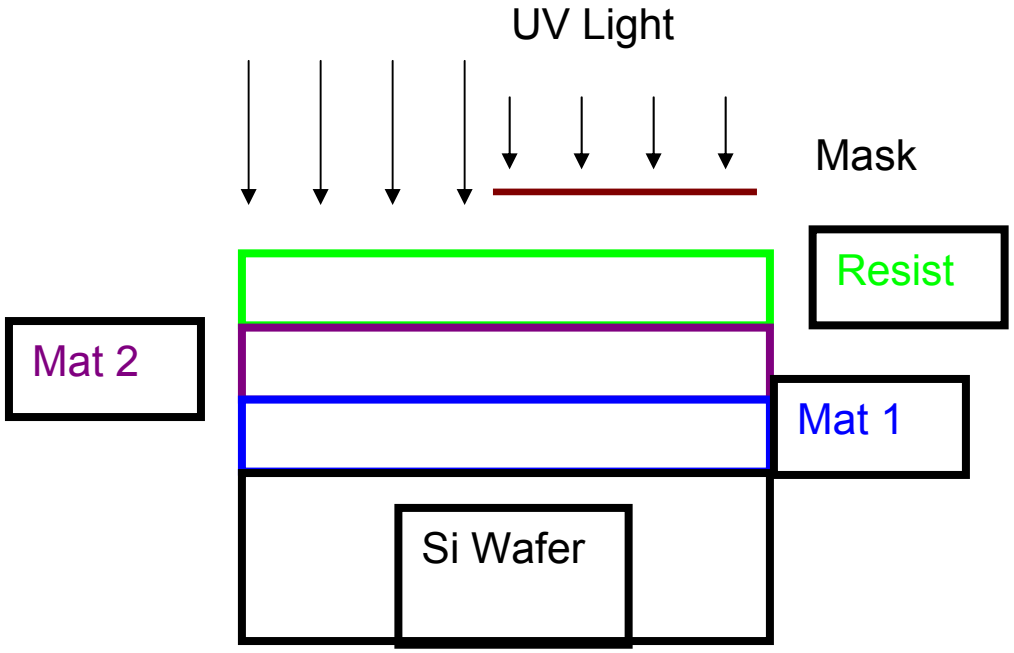


Figure : The process of exposure

Lithography

- This UV light comes from a very complicated optical system to ensure uniform intensity over the large extent of wafer and proper focusing on resist. The positive resist exposed to UV light softens and can be dissolved in a developer to expose mat 2. Now, mat 2 and mat 1 in the exposed area can be selectively etched to do processing in the masked Silicon.

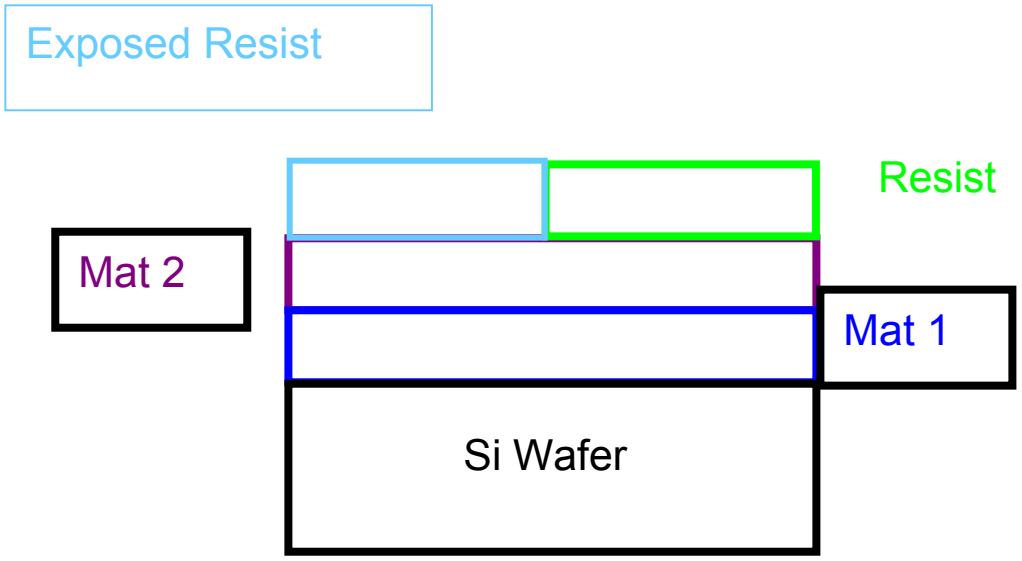


figure : The process of development

Lithography/Etching

- However, due to diffraction of light around mask edges, nonplanar surfaces and slight over or under exposures, it is difficult to get exact replica of open feature transferred to the resist. Hence actual mask sizes on layout are not transferred to silicon exactly and there is always variability. This variability has become a serious concern below 90nm technology where impact on circuit performance due to this is significant.

- The normal step, which follows lithography, is etching. Etching is a selective removal of a material from the wafer. The basic principle is that the material to be etched will only be removed from exposed areas and will not be removed from the areas protected by unexposed resist.

Etching

- Ideally, any material other than the material to be etched should not be affected in this process. The silicon wafer is in the state shown after mat 2 has been etched following the development previously shown. Note that sometimes the edge etched is not vertical and there is significant undercut in mat 2 under resist.

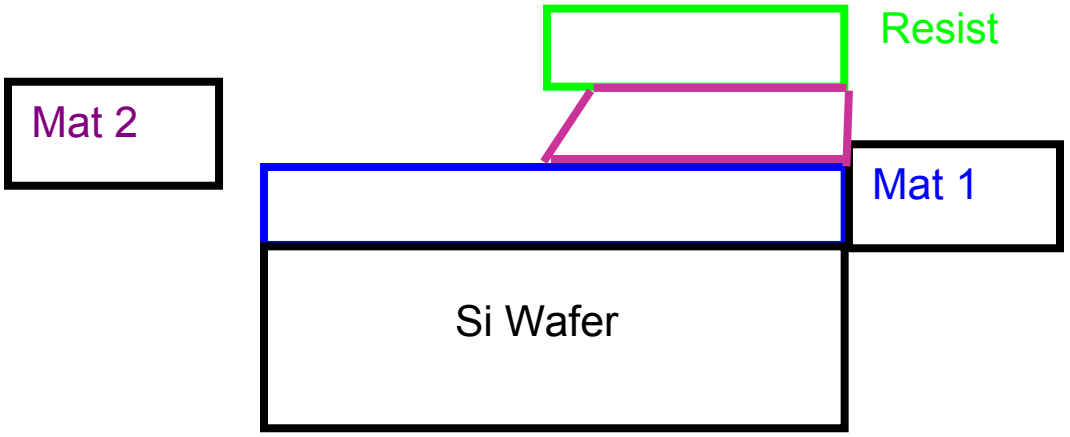


figure : The process of development

- This undercut in etching is a problem which increases some features and decreases others and hence some adjustments to layout are needed as a consequence.

Deposition

- There are a variety of deposition methods available in processing. One would like to have uniform cover of deposited material on the surface of the wafer. It is quite easy to get uniformity if the surface on which deposition is done is planar. However, a typical deposition pattern with a CVD isotropic deposition system will be as shown (Notice?).

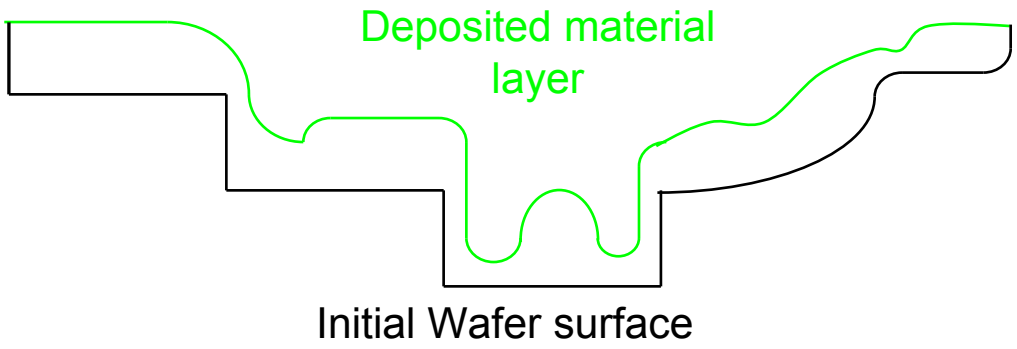


Figure : Profile after deposition on non-planar surface

- One can easily notice non-uniformity of the deposited layer. In fact, the layer is very thin on the vertical walls of initial surface. This may pose some problems for metal layer reliability due to higher current density. This sometimes dictates a higher line width for metal lines than minimum allowed width.

Oxidation

- This is one of the most important steps in Silicon Processing as oxide can be used as
- a block for impurity diffusion,
- as a dielectric for capacitance or interlayer insulation, or
- as a protective layer to guard Silicon against contamination.
- Oxide can also be selectively grown using nitride as a mask as nitride oxidation is much slower than Silicon oxidation. Thus the areas of Silicon covered by nitride will not oxidize.

- The oxide is normally grown in dry oxygen, steam, or high-pressure oxygen in 750°c to 1200°c range. The main reactions are

$$Si + O_2 \rightarrow SiO_2$$

$$Si + 2H_2O \rightarrow SiO_2 + 2H_2$$

The wet oxidation is much faster than dry oxidation.

Oxidation

There are certain general guidelines on which methods to use.

- (i) When very high quality, thin oxide is needed such as gate oxide in MOS processing, only dry oxidation is used. Nowadays Nitrated oxides are used.

- (ii) When good thick oxide is needed dry oxidation followed by wet oxidation is used. This reduces the time for oxidation by a factor of about 10. This is the only method for growth of good quality thick oxides of thickness around 1μm in practice. The high-pressure oxidation can also be used to enhance the growth rate.

- (iii) Deposited oxide layers are normally used for inter-metal isolation, since metals such as Aluminum cannot sustain high oxidation temperatures.

Oxidation

- There are many consequences of oxidation which affect designers or alter processing somewhat. These will now be discussed briefly.

- (1) While oxidizing, about 50% of Si layer is consumed. Hence the oxide layer will have higher elevation than original Si and will eventually lead to non-planar starting surfaces.

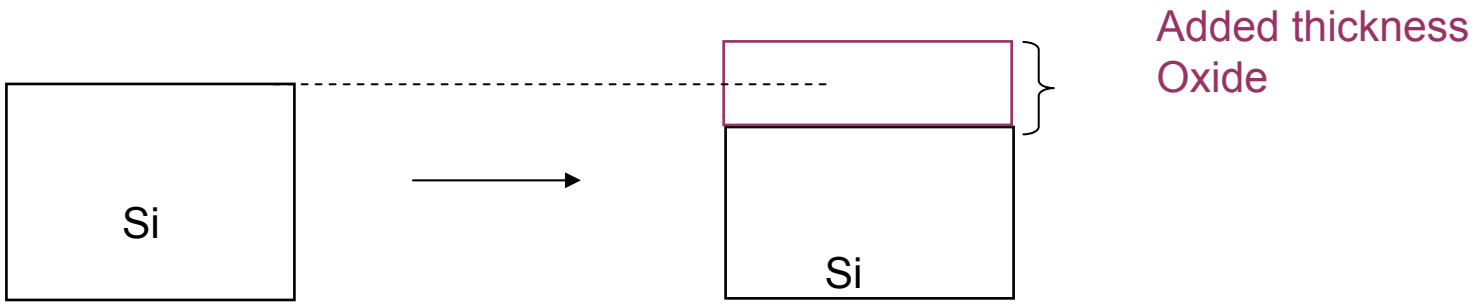


Figure : Increased Elevation due to oxidation

Oxidation

(2) Nitride is an effective mask for oxidation. If it were a perfect mask, the oxidation after patterning nitride will look as shown.

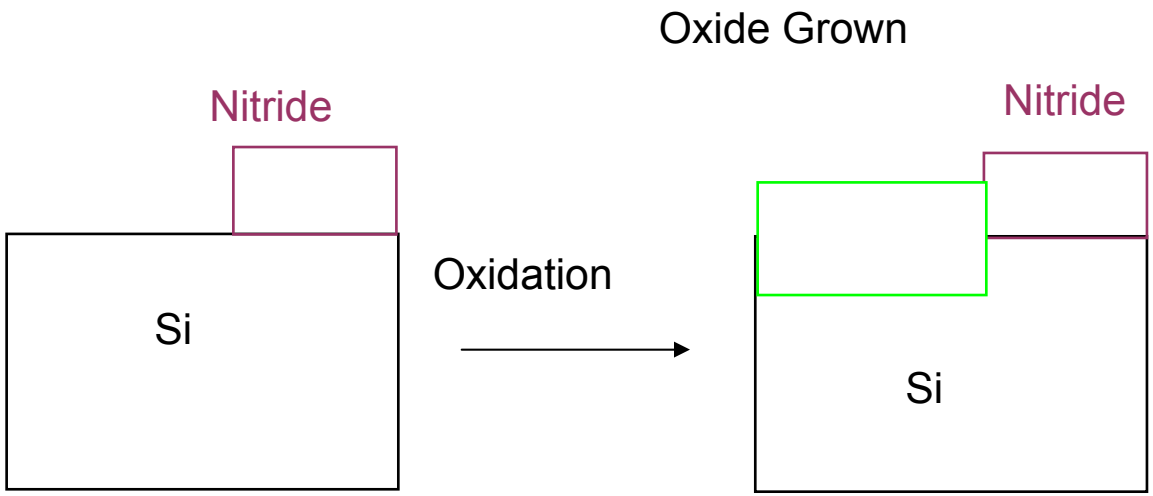


Figure : Selective Oxidation with Perfect Nitride

Oxidation

In reality, oxide encroaches under nitride and lifts it off to produce bird's beak structure.

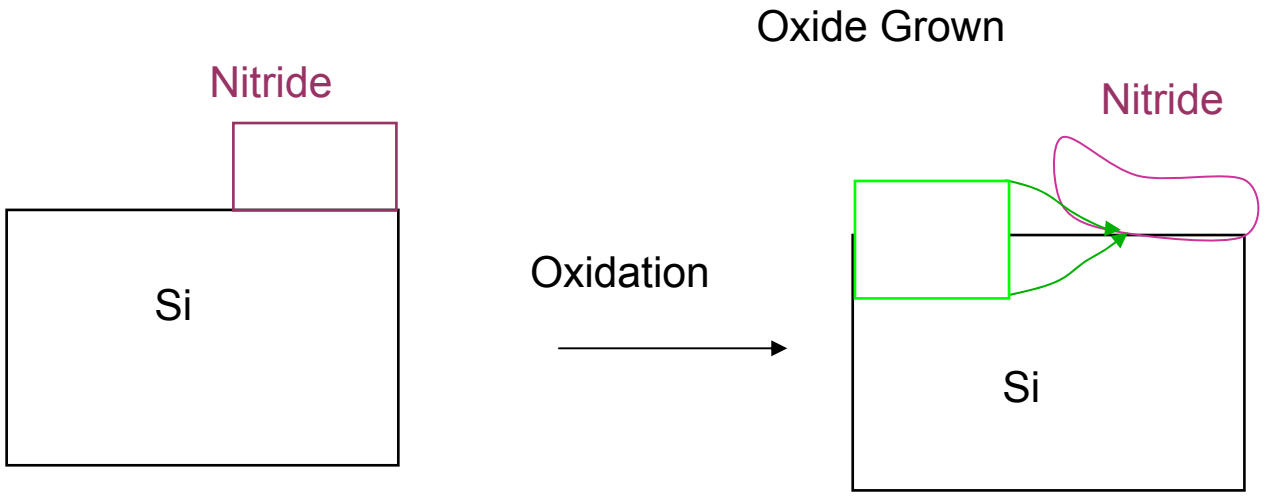


Figure: bird's beak structure.

Oxidation

This has very profound implications for width of MOS devices as shown.

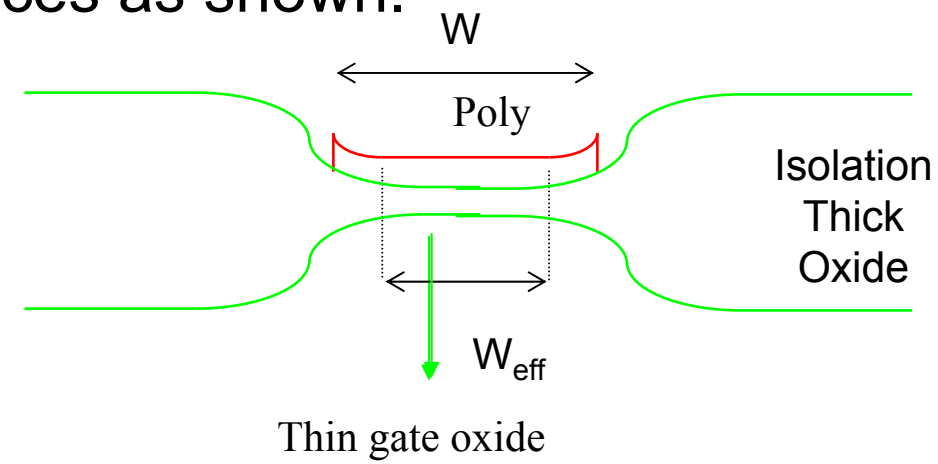


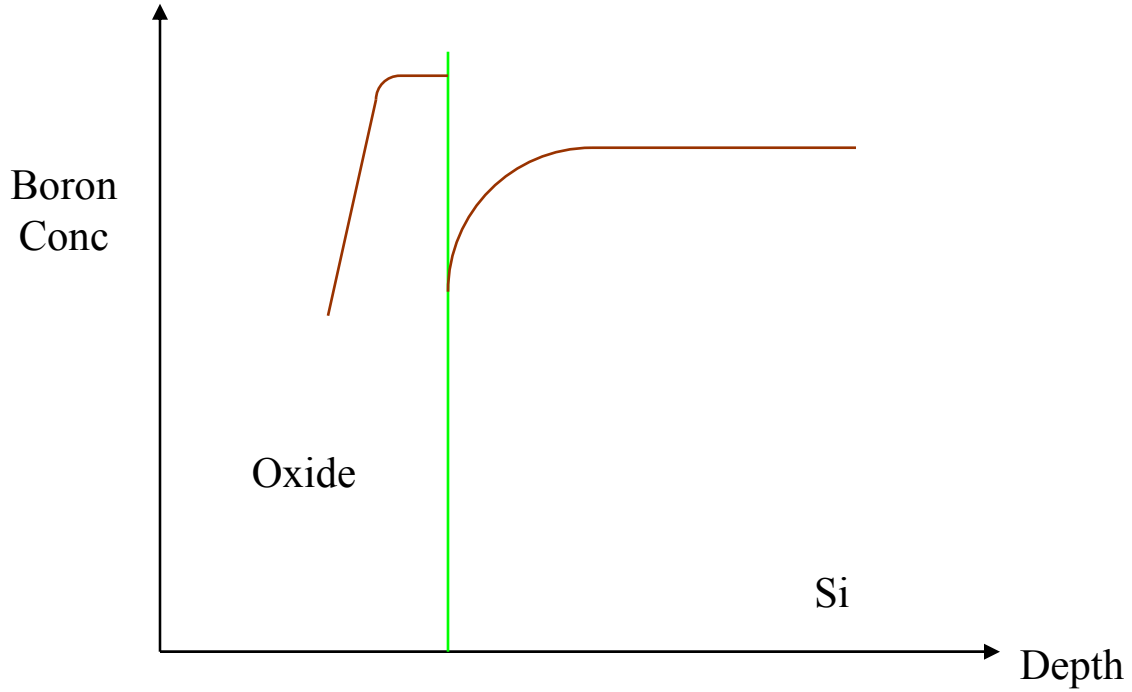
Figure : Width Reduction due to Bird's Beak

The thickness of oxide away from W_{eff} is larger than gate oxide thickness. Since the threshold voltage increases with increasing oxide thickness, there will be width reduction. For technology below 0.18μm, shallow trench isolation is used instead to avoid this problem.

Oxidation

(3) Impurity Segregation: Due to high temperature during oxidation, the impurities tend to diffuse deeper. The redistribution, however, is affected by the presence of oxide – Si interface.

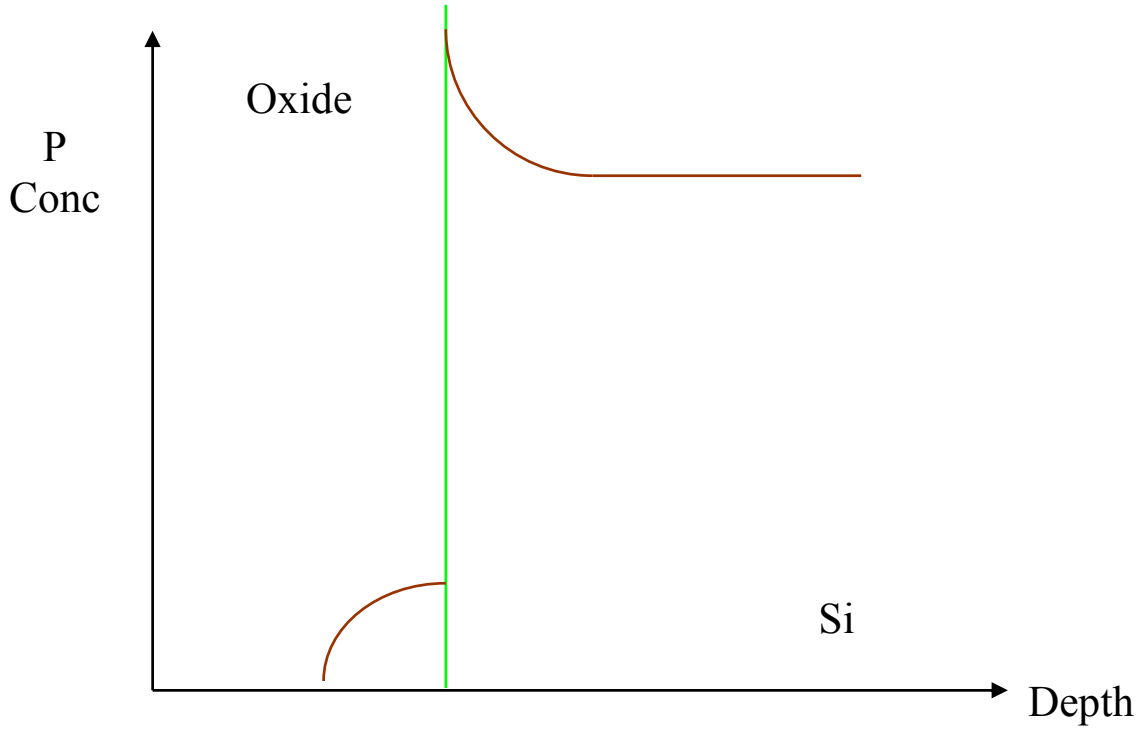
The concentration of Boron near the interface is reduced in Si compared to normal concentration and affects the threshold voltage.



Boron redistribution due to Segregation

Oxidation

The concentration of P and As is increased as shown.



As and P redistribution due to Segregation

Ion Implantation

- This is a process of introducing impurity or other ions into Si by giving them enough energy to penetrate into Si.

- Higher energy ions will penetrate deeper into Silicon. Hence the ion energy and number of ions to be introduced measured by dose determine the impurity profile.

- Normally this step is used to selectively dope silicon where the parts covered by resist will not allow implant to pass through.

Ion Implantation

- If the scattering of impurity projectiles is random, it is easy to find the average depth of penetration (range R_p) (This is also normally the location of the largest concentration), standard deviation in the depth (ΔR_P), and standard deviation perpendicular to beam direction (ΔR_L).
- The typical values of range are in 0.01μm - 2μm interval.
- For P, As and Sb, the distribution of ions is a Gaussian. For Boron, it is found that the distribution has a long tail which is modelled by a Pearson IV or V distribution function.

Ion Implantation

- High-energy Oxygen implants were popular for forming oxide deep inside Silicon for Silicon on Insulator (SOI) devices.

- Relatively precise control of doping profile and integrated doping concentration (this is actually the implant dose typically specified in cm^{-2}) is possible by ion implants.

- However, it is not totally free of problems, which will be discussed in brief.

Ion Implantation

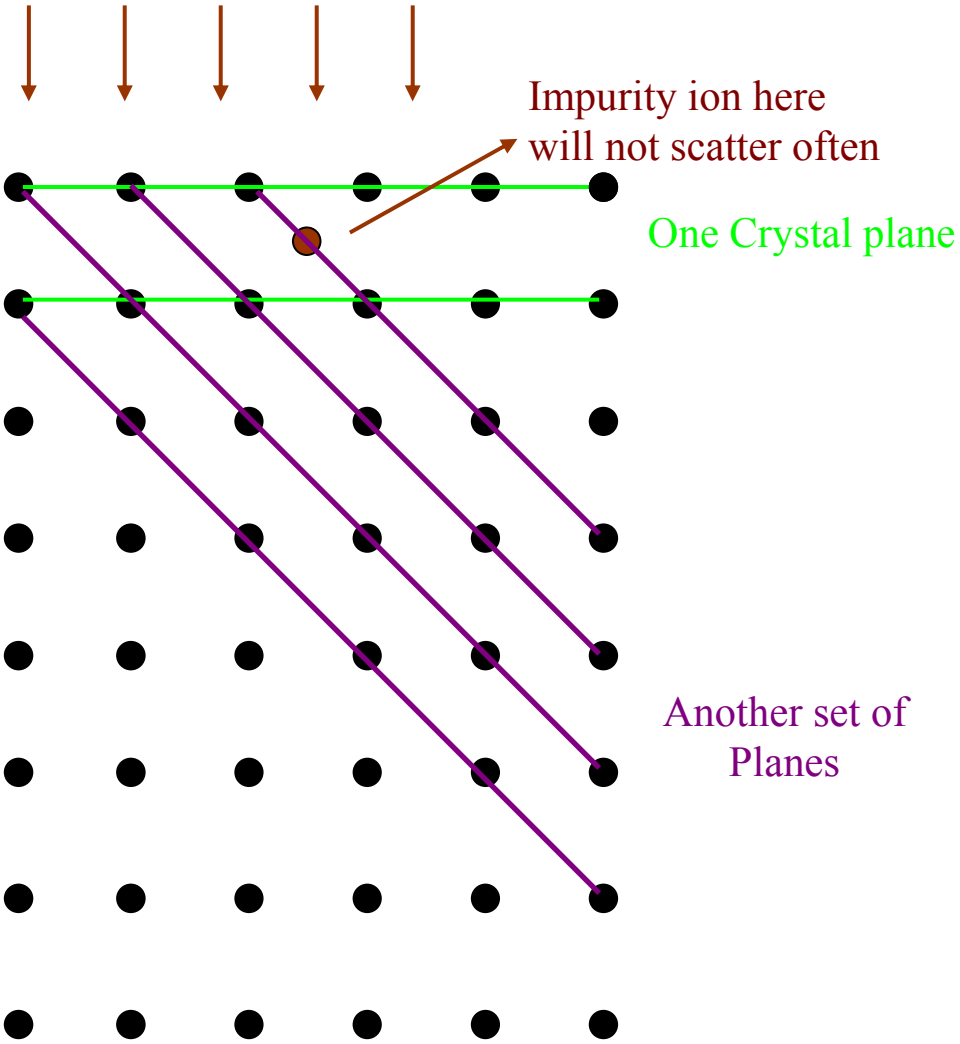


Figure : Channeling in Single Crystal Si

(1)Channeling: The incident ion beam could sometimes be perpendicular to one of the crystallographic planes in Si. In this case, due to less scattering along the path, the ion may penetrate 2 times deeper than the normal depth predicted by random scattering.

To avoid this, the beam may be sent at 3° - 7° to the normal.

Ion Implantation

- (2) Damage – Crystalline structure of Si is damaged by implants as Silicon atoms are scattered randomly and displaced from their normal positions in the periodic crystal lattice. Also, impurity (dopant) atoms will automatically be in random positions in Si.

- This damage is to the extent that Silicon becomes amorphous after a part of the dose has been implanted.

- Hence a post implant anneal, which gives thermal energy for Si to restore to the lowest energy single crystal state, becomes essential.

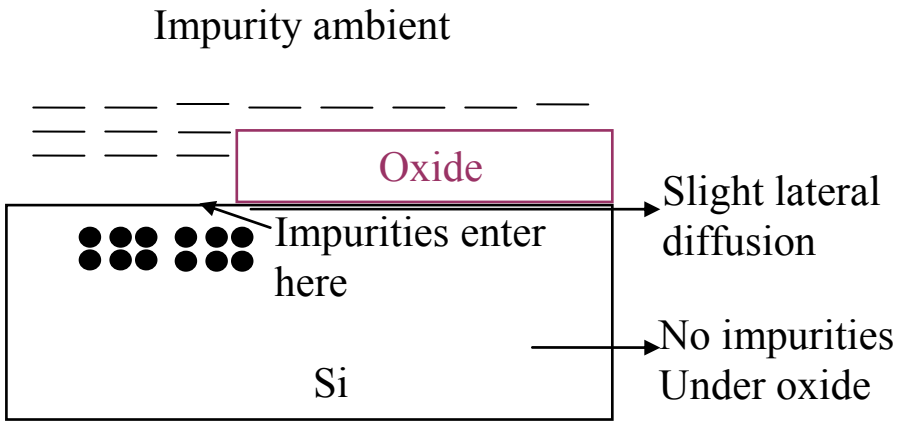
Ion Implantation

- Normally, an anneal of about 30 minutes at 850°c is adequate to remove the crystalline damage and activate impurities by moving them into substitutional sites.
- Nowadays, rapid thermal annealing is often used as thermal budget is quite limited for short channel MOSFETs. RTA for 30 seconds at around 1020°c is typically sufficient. SPIKE or FLASH or LASER anneal is also used in case of tight thermal budget.
- This, however, will redistribute the impurities by diffusion and some of the implant advantages are adversely affected.

Diffusion

- This is a natural phenomenon due to the tendency for impurities to move from a region of high concentration to a region of low concentration. Hence if some impurities are introduced near Silicon surface, at higher temperature they will diffuse deeper into Silicon.

- Diffusion was a very popular way of selectively introducing impurities in Silicon before implants came along.



- Due to slow impurity diffusion in SiO2, oxide was used as a mask for selectiv introduction of impurities.

Figure : Oxide as Diffusion Mask

Diffusion

- Even if this technique is not so much used in the present implant age, the diffusion is still a part of processing life. The post implant anneal makes the implanted impurities redistribute by diffusion.

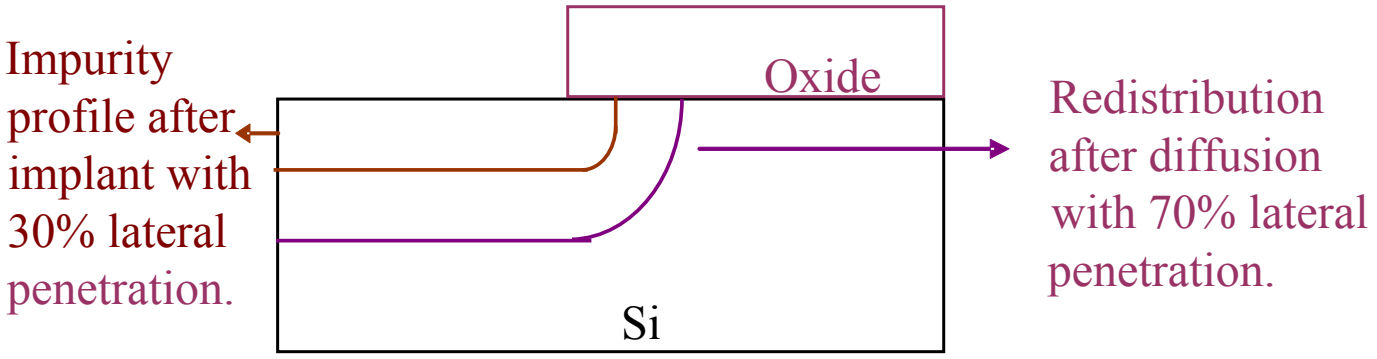


Figure : Redistribution of Implanted Impurities after anneal

Diffusion

- The diffusion takes place any time Silicon is heated. The presence of lateral diffusion at about 70% of depth decides some of the critical device features and inherently limits some of the device features and properties.
  - For very deep diffusions, the profile is a near Gaussian. For very shallow ones, it is a near error function.
  - The characteristic depths for the functions are determined by the diffusion coefficient D which is a primary exponential function of the temperature and the time t of diffusion. $\sqrt{Dt}$ is typically the characteristic depth for the distribution. The diffusion coefficient D is a very strong function of concentration of vacancies and interstitial within the crystal. It also depends on the presence of other impurities and materials.

Diffusion

- One such effect we saw in segregation before. The other one is shown in the figure below and has strong implications on Bipolar Junction Transistors (BJTs).

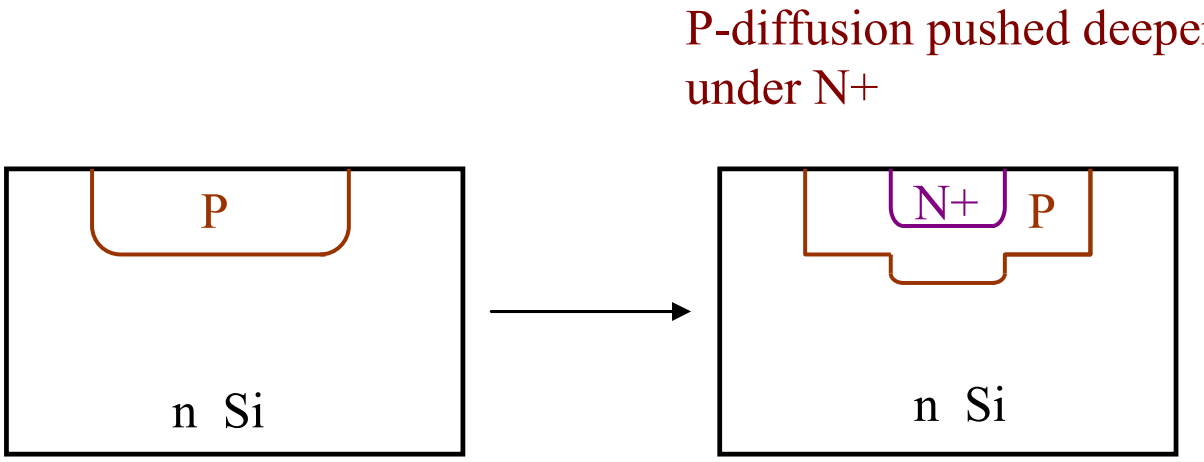


Figure : Effect of Impurity on Diffusion

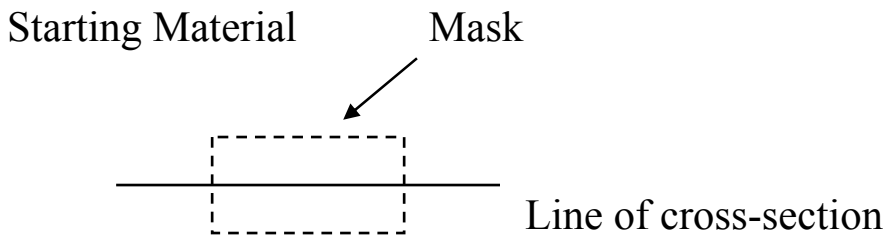
Process Sequence with Mask

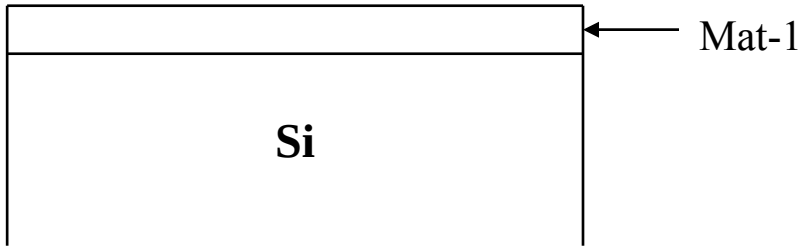
- A common process sequence is normally done for each mask transfer to Si. This sequence is given here so that as designers, you are aware of what masks do. Fabrication of whole inverter is progressively developed with a full mask set that is used in the proper sequence.

- As process development is not the main focus of this module, we will not go through all the process steps. We will only look at one important process step that determines gate mask and final cross section of the inverter.

Process Sequence with Mask

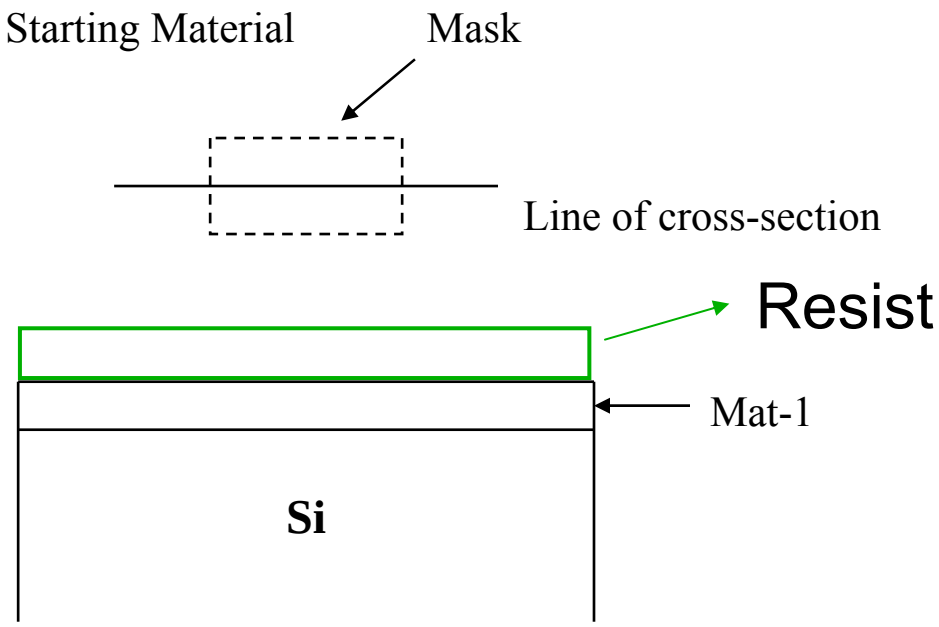
- Process Step: Selective Removal of material – 1 (Mat-1) to dope Si with Boron (p type).
- Process Sequence:





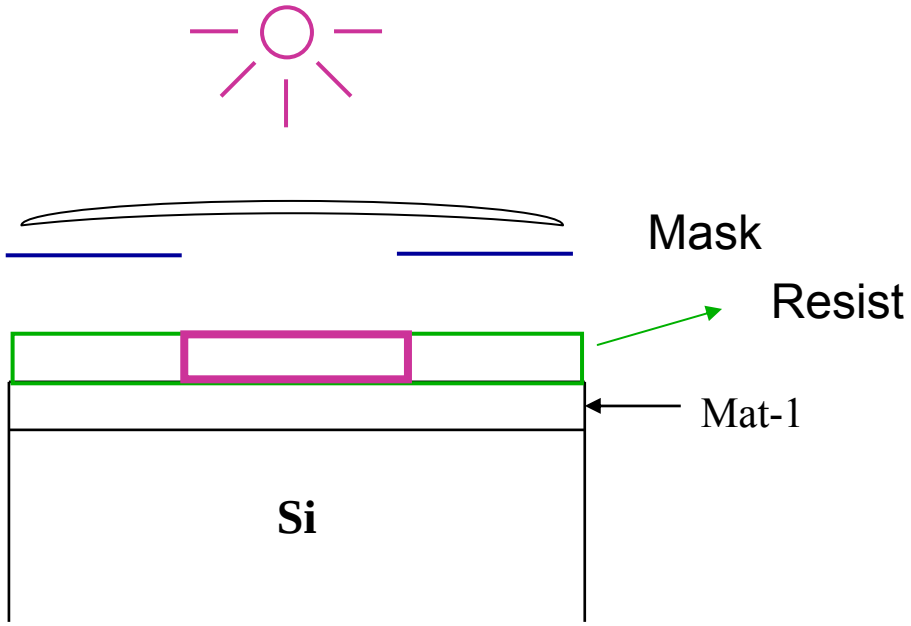
Process Sequence with Mask

- i) Deposit Resist



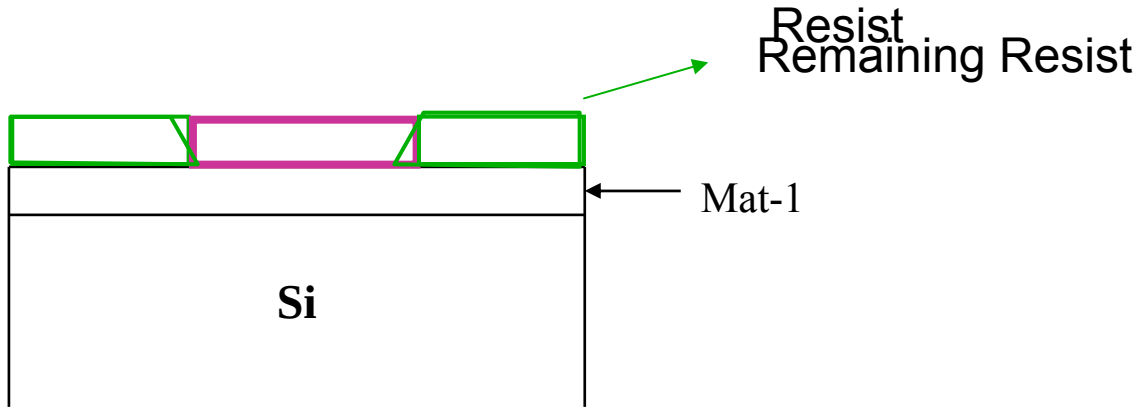
Process Sequence with Mask

- ii) Softbake Resist ~ 60^0c, 1-3 minutes
- iii) Exposure – Expose Resist through the mask. Either the mask or its complement used depending on the effect desired.



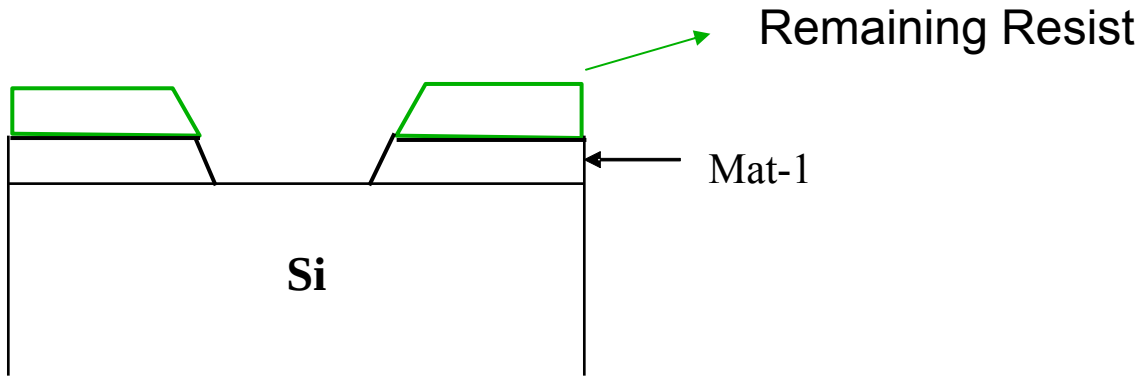
Process Sequence with Mask

- iv) Hard Bake Resist ~ 80^0c, 1 - 6 minutes
- v) Develop Resist: Dissolves the exposed resist



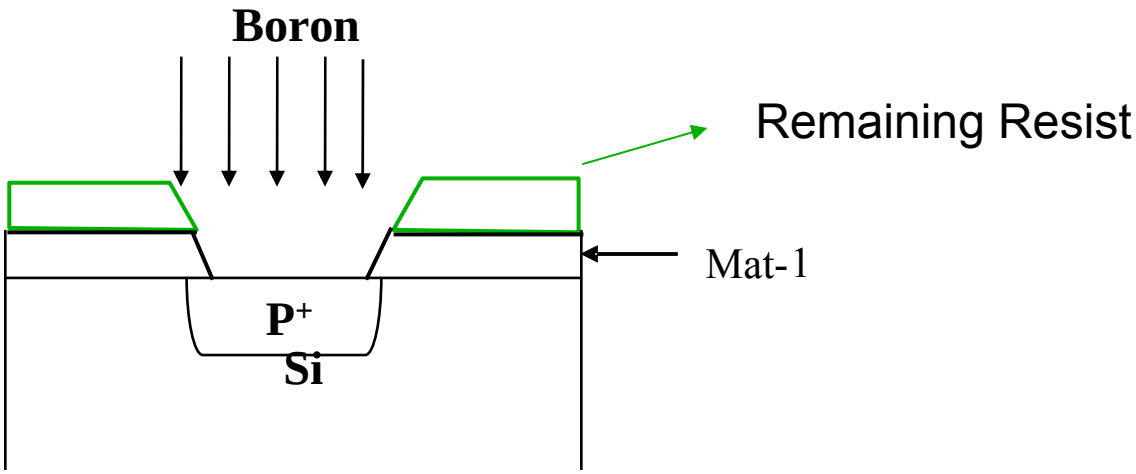
Process Sequence with Mask

- vi) Etch Material-1 selecting a chemical which does not attack resist and silicon or do a plasma etch.



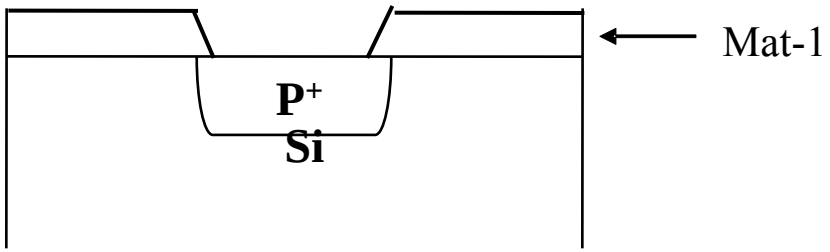
Process Sequence with Mask

- vii) If Boron is to be implanted, it can be done now with resist present.If diffusion or oxidation is involved, the resist must be removed at this stage. In this case, material-1 must block oxidation (material-1 -> Nitride) and diffusion (material-1-> Oxide). Here, implant Boron.



Process Sequence with Mask

- viii) Remove the resist as selective job is finished
- ix) Wafer ready for the next step. The cross-section after (viii) will be normally shown.

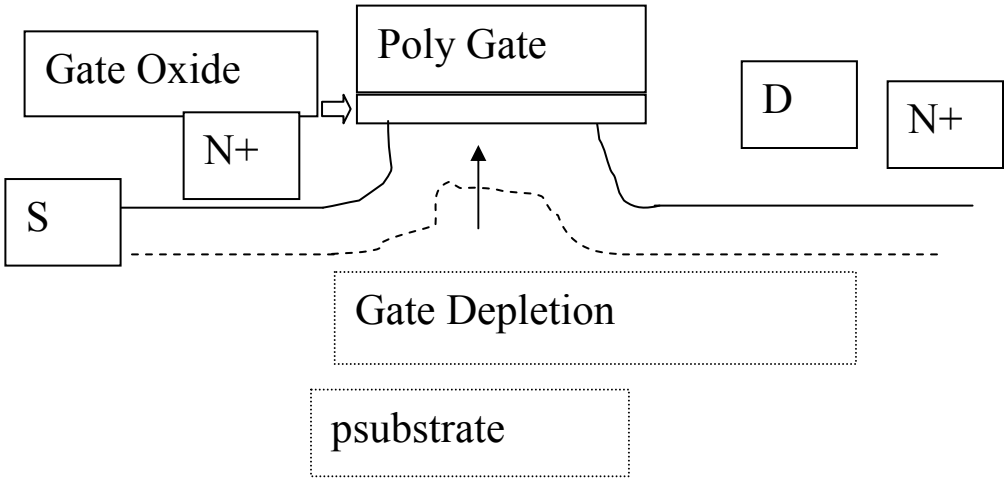


Layout of Stand-alone CMOS Inverter

● Earlier, we drew a circuit of CMOS inverter without worrying about how to make these devices and connect them on Silicon.

● Let us now discuss how many different layers will be needed on the layout at the minimum. Normally, many more layers are needed as there are additional steps and more than one metal layer.

● We need to achieve the device structure below for both n and p devices.



Layout of Stand-alone CMOS Inverter

● The starting substrate will normally have one type of doping. Here we use p-type substrate. Hence to make both p- and n-channel devices, we need to create an n-type regions using implant/diffusion in p-type substrate. This level is labeled as NWEL colored in brown.

● There should be a mask that defines gate regions using selective etching. This will normally be called as GATE or POLY mask named after polysilicon material normally used to form the gate.

● There should be a mask layer to form the N+ doped regions around the gate to make n-channel devices using implant/diffusion.

● There should be a mask layer to form the P+ doped regions around the gate to make p-channel devices using implant/diffusion.

● We need to connect these Source, Drain, Gate regions depending on the circuits to be implemented using metal lines. These metal lines are isolated from each other by an insulator such as oxide. For this step, oxide is deposited over whole silicon substrate wafer and holes are made in it over S/D/G to make actual connections by filling these holes with the metal. This is achieved using contact mask.

Layout of Stand-alone CMOS Inverter

● We still need to connect these Source, Drain, Gate regions depending on the circuits to be implemented using metal lines. For this step, metal is deposited over whole silicon substrate wafer and the metal is etched from unwanted areas. This is achieved using metal mask.

● There can be several layers of contacts and metal mask pairs – up to 8 in advanced technologies. However, for simplicity, we will just confine to maximum 2 layers of metal.

● The mask is determined by process limitations and fabrication. Some key issues that drive this are covered here.

● As many devices are integrated on any chip, isolation region with thick oxide is formed wherever there are no channel/S/D regions. This adds an extra process step and mask.

● S/D or some other regions need to connect to metal. For this, a contact is placed in S/D areas. As despite of misalignment due to lithography, the contact must fall within this region, S/D areas have to be larger than the contact by a certain amount. This determines poly/S/D contact design rule.

● Similarly, as metal must cover full contact despite of misalignment, metal mask is larger than the contact by the same amount. This determines metal contact design rule.

Layout of Stand-alone CMOS Inverter

● Consider forming source-drain-gate structure. Let us say we form S/D first by B or P or As implant followed by post implant anneal. This was the case when Al was used as a gate material which cannot withstand high temperature.

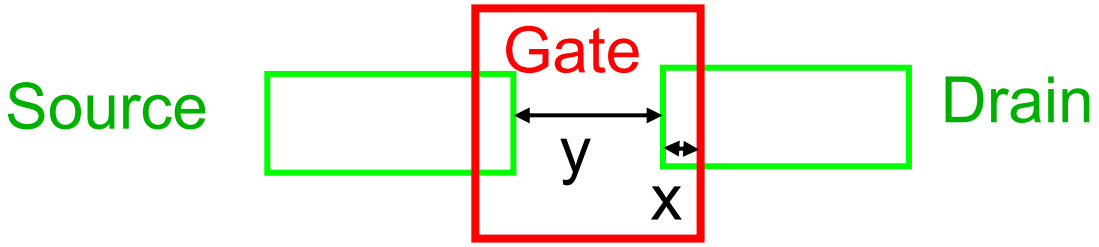
● Then we grow gate oxide and put gate on the top.

● For MOSFET, gate must overlap S/D.

● As these are different masks, that will mean gate should be wider by a large amount so that even if lithography does not align the gateproperly, it will still overlap S/D.

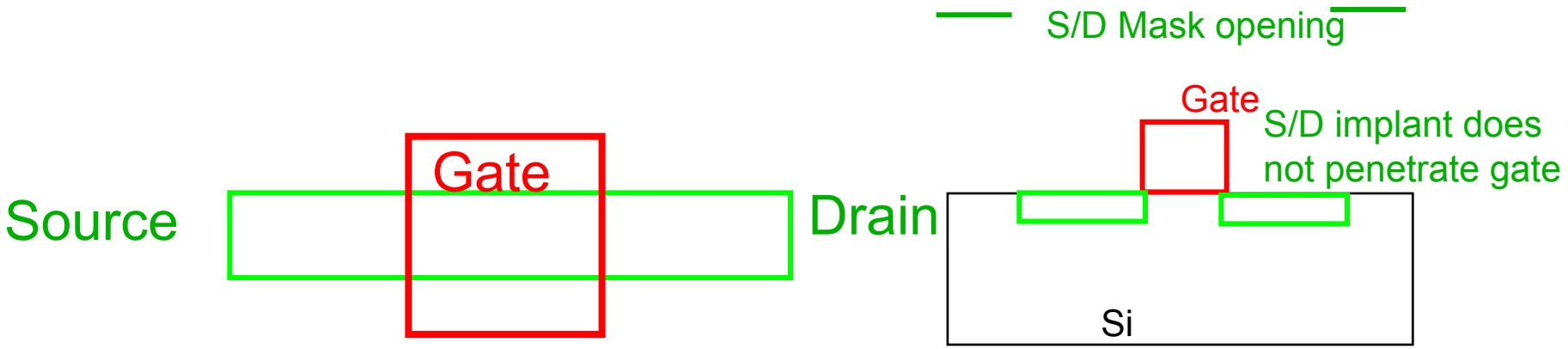
● Hence if x is lithography misalignment possible, y is separation between S and D junctions under gate, then gate would need to be at least (y+2x) wide.

● Of this, only distance y induces channel, the rest just gives extra load to slow circuit down. Hence a very clever trick is applied in processing.



Layout of Stand-alone CMOS Inverter

- In this method, the gate is first formed as shown using red Gate mask.
- S/D green mask is a continuous rectangle through which S/D implant is done. S/D implant energy is chosen so that it does not go through the polysilicon gate.
- Hence only the exposed silicon regions on the two sides of the gate receive implant. Clearly, S/D mask is a sum of source, channel and drain area.
- The source drain regions automatically attach to the gate as shown. This is known as self-alignment and has made 45nm technology today feasible.



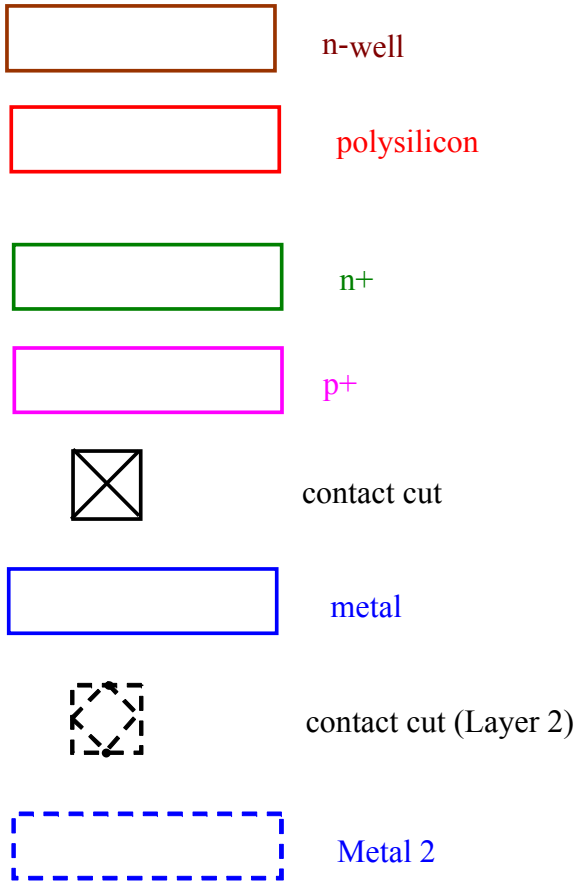
Layout of Stand-alone CMOS Inverter

- Note that it is possible for polysilicon, which forms the gate material, to be used as a short interconnect if many inputs are connected together. Also source or drain regions can be used as short interconnects if they are connected together for the same channel devices.
- Also, note that many other considerations may add some process complexities or masks to the above basic set.
- The color convention is now given in the table. The convention for poly, N type active area, contact and metal is universal. Others may vary a lot.

Layer	Color
NWEL	Brown
POLY (Gate)	Red
N+ (NSD)	Green
P+ (PSD)	Pink/Purple
CONT	Black
METAL	Blue

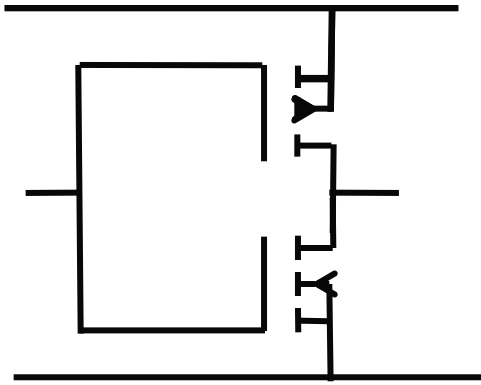
Layout of Stand-alone CMOS Inverter

Layout diagram colour code in the process sequence order

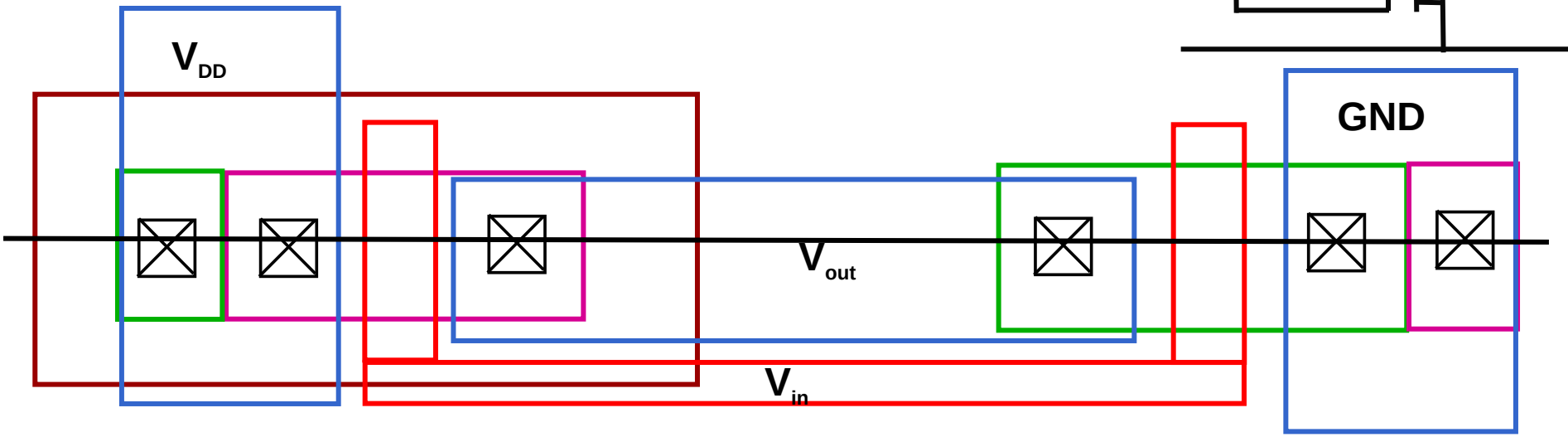


Layout of Stand-alone CMOS Inverter

- Let us now evolve a primitive, stand alone, layout of the inverter.
- First form n-well which gives n-type substrate within p-type wafer. This has to accommodate all of p-type MOSFETs.
- Now form S/D regions for both devices. They should be wide enough to accommodate connections for the inverter.
- Now draw gate mask slightly wider so that it separate S/D.



- Now put six contacts for the inverter connections. The gates are already connected by polysilicon. Contacting layers are always larger and S/D/channel regions merged.
- Finally, the metal connections as shown.



Layout of Stand-alone CMOS Inverter

- Notice that an n+ (Also denoted by NSD – n type source drain) area is added which is butting to p+ (PSD) and connected to V_{DD}. This is extra geometry added for reliability. Similarly p+ area is also attached to n+ which is grounded.

- The sizes of these geometries and spacing are based on design rules, which now follow. The design rules are largely determined by the processes.

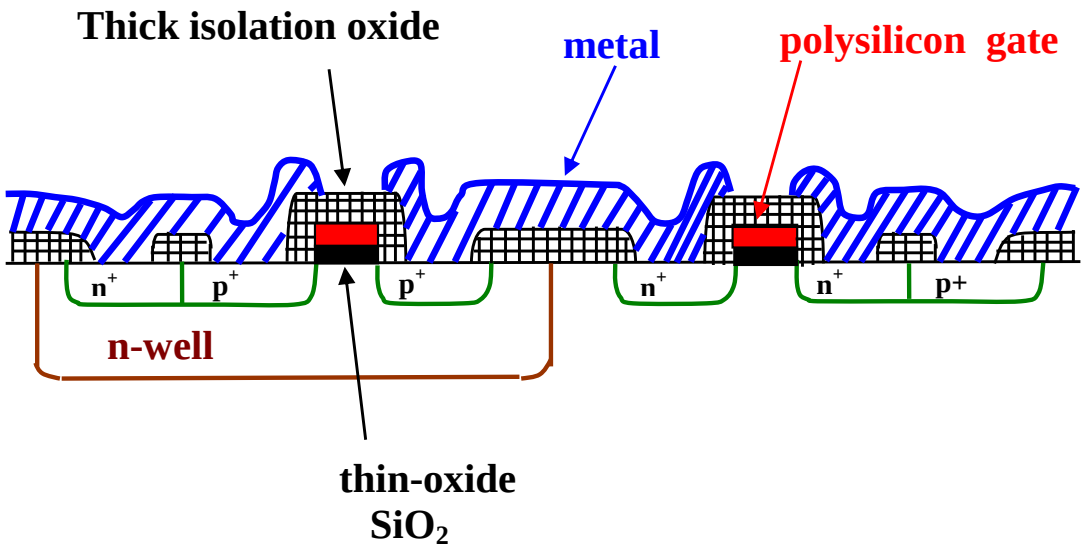
- Cross-section within silicon of this inverter along a line at the centre of the layout is as shown. For one process, the additional contacts are not included for simplicity.

- Cross-section can be drawn using the mask sequence and evolution using process steps that were explained in initial slides. Can you notice anything important?

Cross-section of CMOS Inverter

- Uses simple isolation process

n-well process



Other processes :

p-well, twin-tub,

SOI (silicon on insulator)

Oxide or sapphire or magnesium aluminate

p-substrate

moderately doped

Cross-section of CMOS Inverter

● The following are the process steps that yield the above cross section.

(1) Starting p-type wafer – deep Nwell formation to create p-channel MOSFETs.

(2) Selective thick isolation oxide where there are no N+ or P+ regions.

(3) Clear P+ and N+ areas and grow thin gate oxide.

(4) Deposit polysilicon and patern poly/gate mask to form gates of MOSFETs.

(5) Open only N+ areas in the resist and implant As. As stops in resist in P+ areas, isolation and gate giving self alighnment.

(6) Open only P+ areas in the resist and implant B. B stops in resist in N+ areas, isolation and gate giving self alighnment.

(7) Deposit thick BPSG oxide. This provides inslulation between metal that will fill contact. All channel poly is covered by this thick oxide. Hence metal can run over polysilicon without connecting to it.

Cross-section of CMOS Inverter

- The following are the process steps that yield the above cross section.

(8) Open contacts using contact mask.

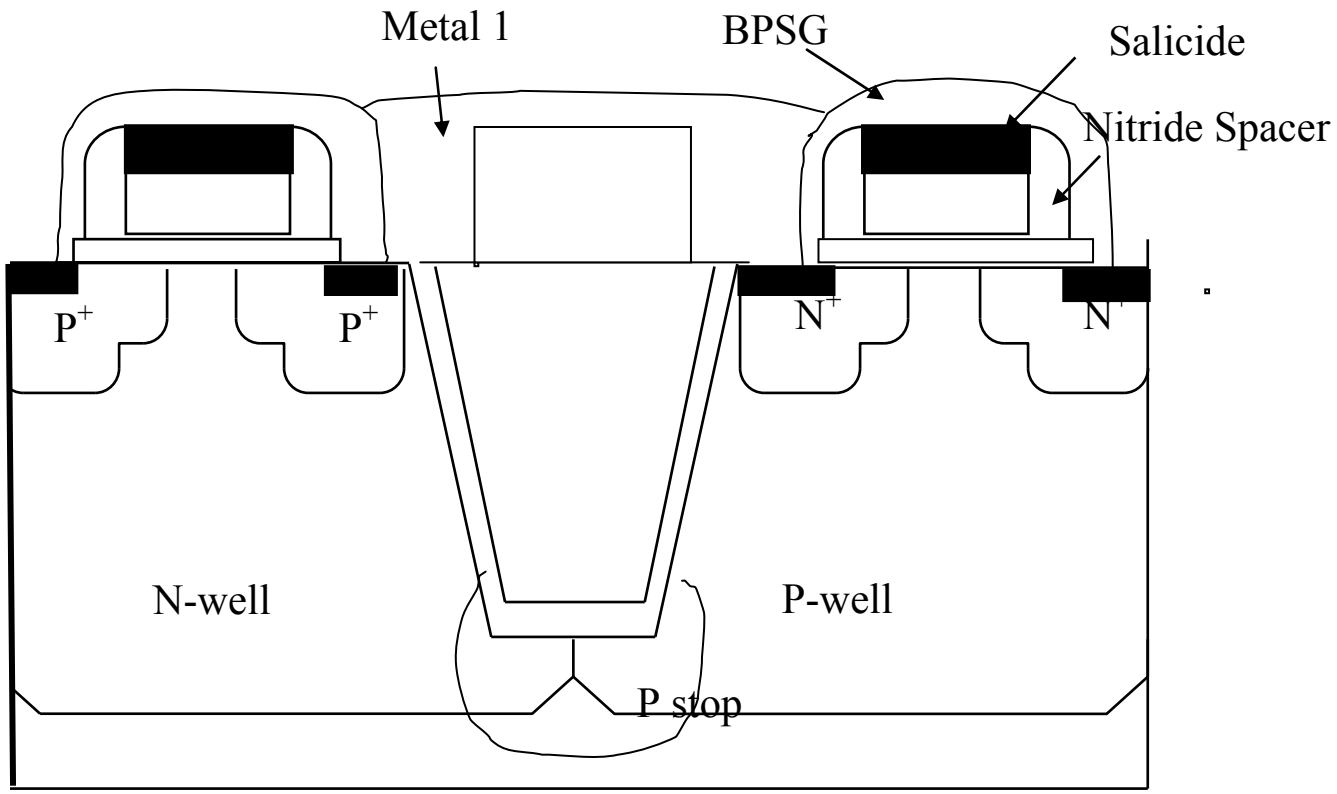
(9) Deposit metal 1 and pattern metal 1 mask to remove unwanted metal.

(10) Deposit thick oxide and open vias to connect to Metal 1.

(11) Deposit metal 2 and pattern metal 2 mask to remove unwanted metal.

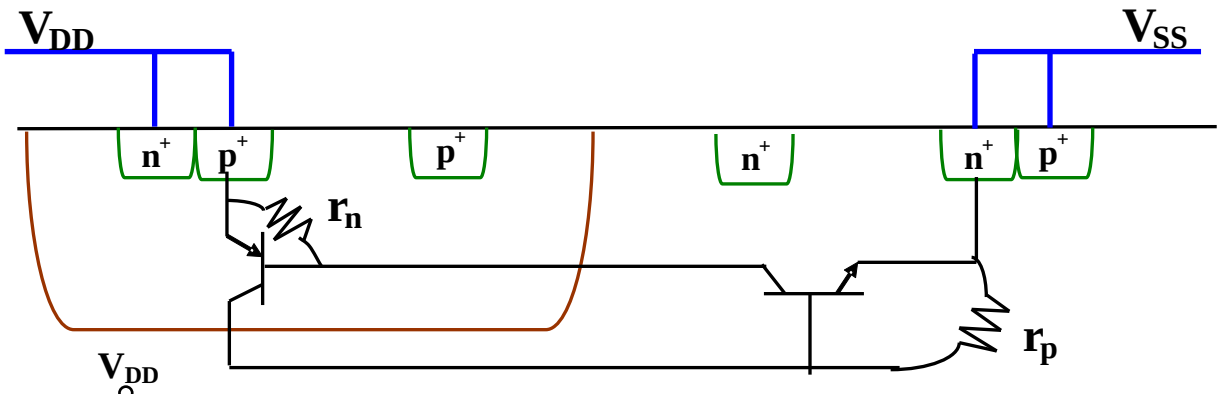
Cross-section of CMOS Inverter

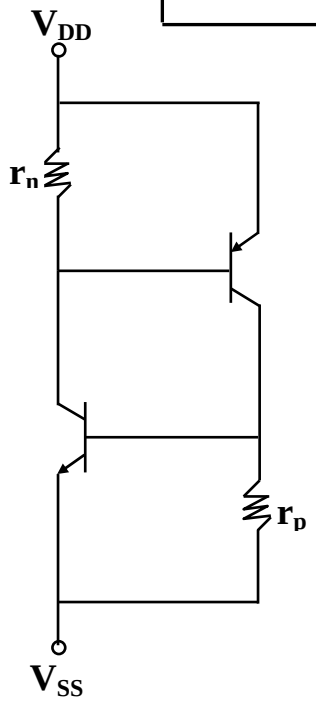
- Uses advanced shallow trench isolation process



CMOS Latch-up

● It is clear that we have got many parasitic devices along with the MOSFET in the cross-section. Here is one such parasitic p-n-p-n device which is responsible for latch-up reliability issue in CMOS. It has one npn and one pnp bipolar transistor and 2 resistors.





● The pnpn parasitic circuit above can be re-arranged as shown on the left.

● Due to positive feedback, this circuit can conduct large current between two power voltages and gate is immaterial.

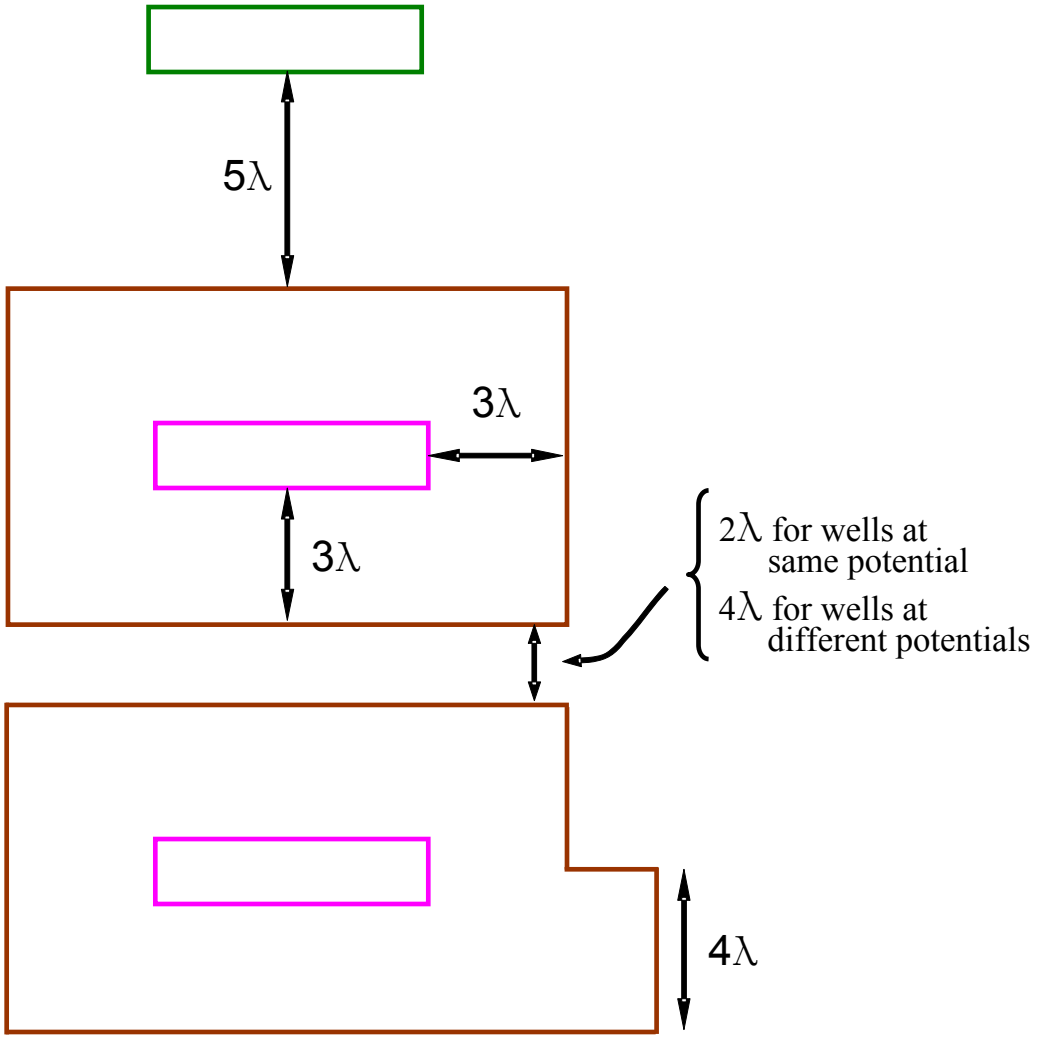
● What can we do to suppress this?

● Gives rise to large design rules and extra features. The design rules are technology independent using a parameter λ which shrinks as technology scales.

Design Rules

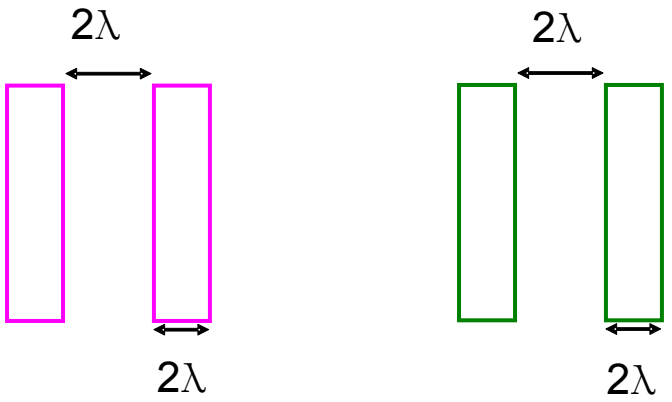
- n-well spacing.

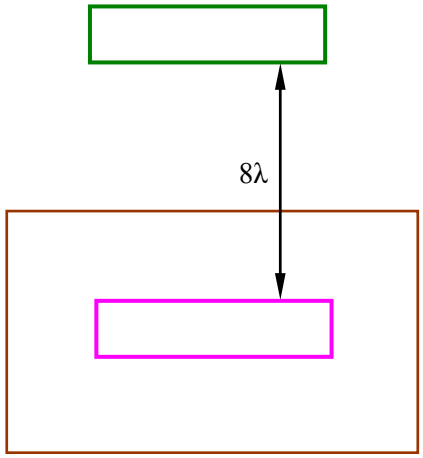
- All design rules are specified in terms of λ which accounts for change in a feature on layout when is transferred to silicon. As technology progresses, λ reduces. Hence this provides a simple way to deal with design rules.



Design Rules

- P+(PSD)/N+(NSD) dimension and spacing.

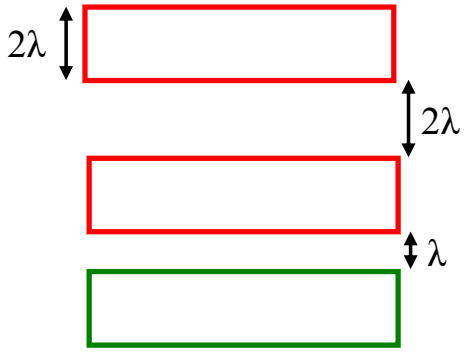




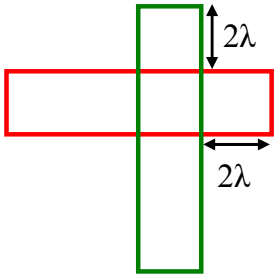
Design Rules

- Polysilicon-n+ spacing.

- MinimumPoly dimension and spacing is 2λ. Parallel n+ and poly can be λ apart.



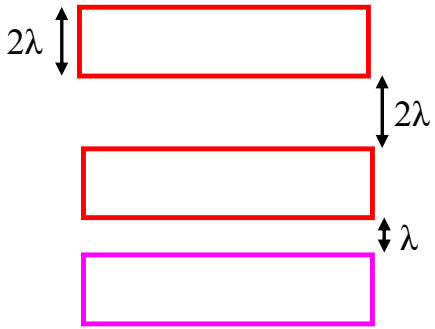
- n+-poly must extend 2λ beyond edge in device areas where poly separates S/D.

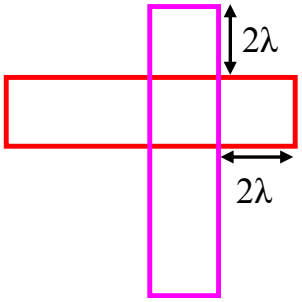


Design Rules

- Polysilicon-p+ spacing.

- Minimum Poly dimension and spacing is 2λ. Parallel p+ and poly can be λ apart.
- p+-poly must extend 2λ beyond edge in device areas where poly separates S/D.

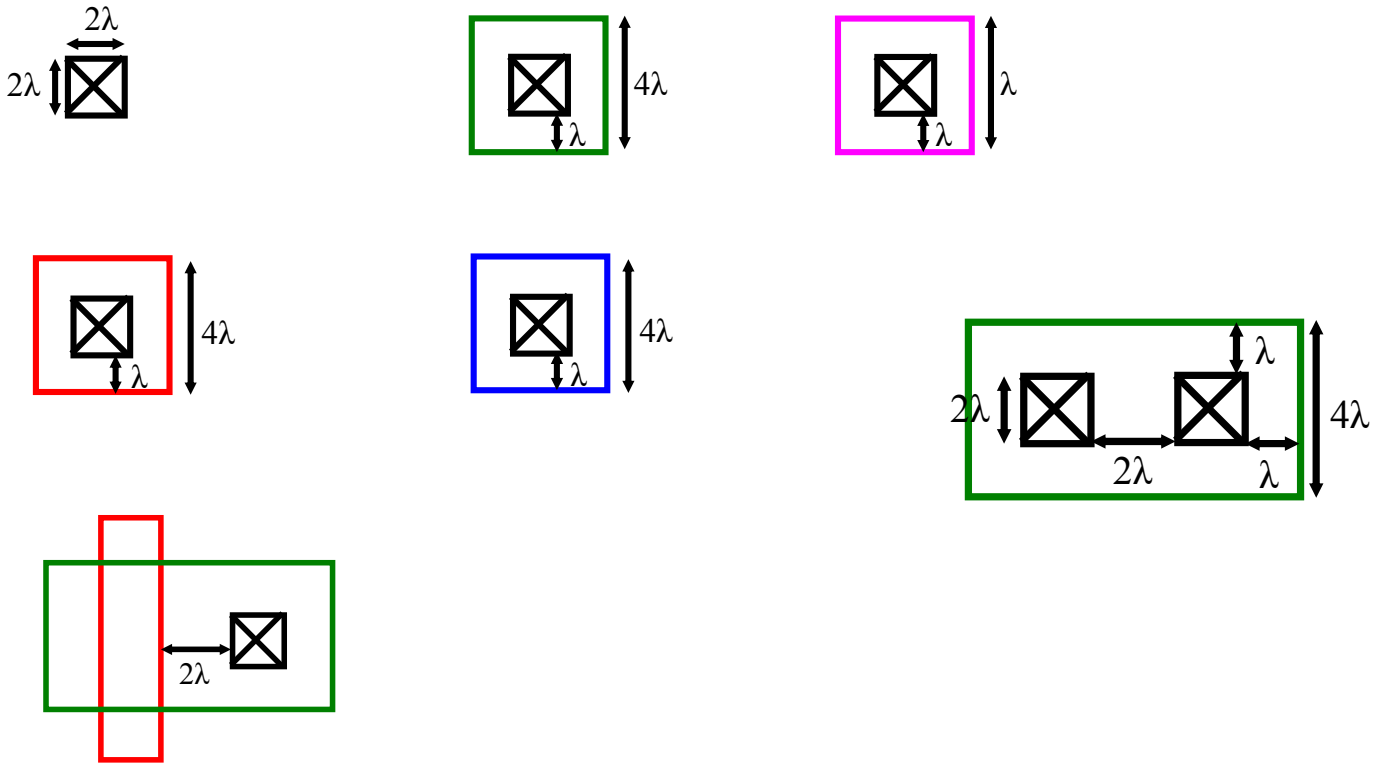




Design Rules

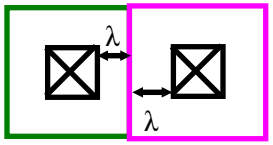
- contact dimension and spacing.

- Minimum contact dimension and spacing is 2λ. All contacting layers must cover contact by λ on all sides.
- p+/n+ contact-poly spacing must be at least 2λ.

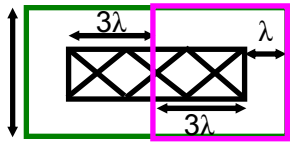


Design Rules

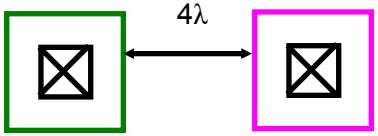
- Butting n+-p+ contact.



- Long n+-p+ contact.



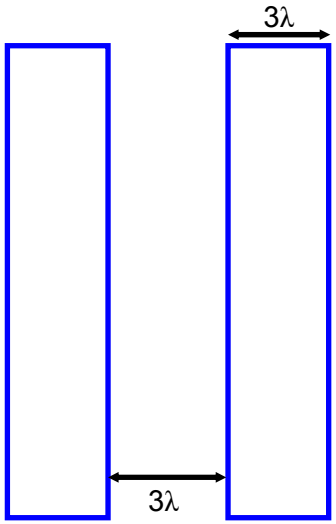
- Separated n+-p+ contact.



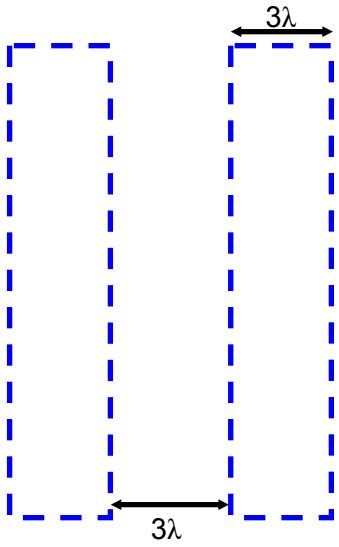
- As butting contact is the most compact, it is generally used.

Design Rules

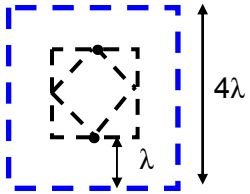
- Metal 1 spacing.



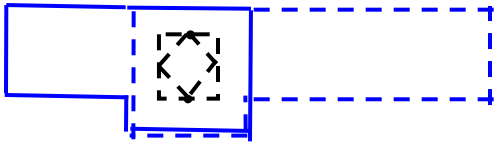
- Metal 2 spacing. Larger for reliability.



- Metal layer 2 contact.



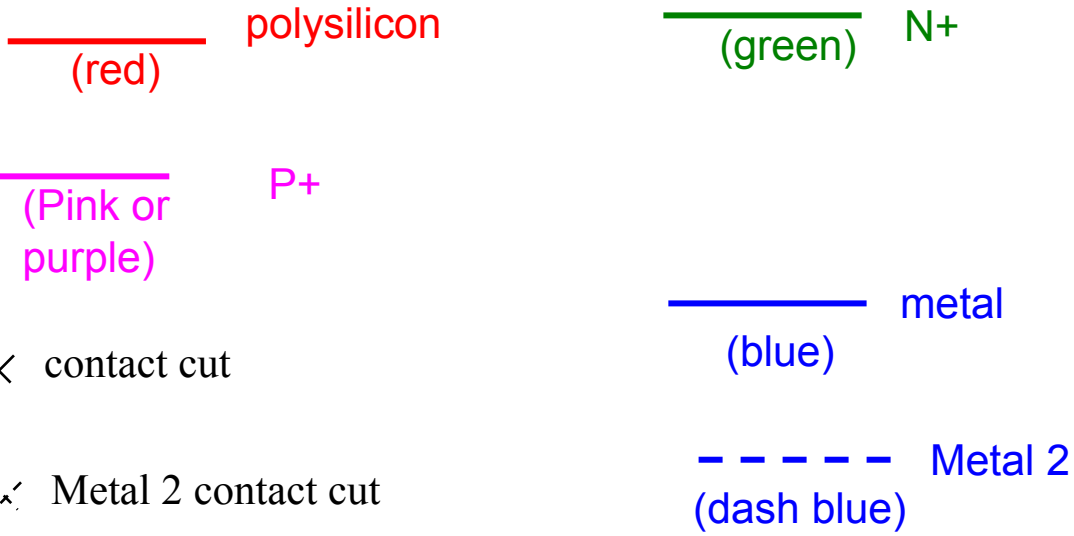
- Metal 1 to Metal 2 contact. This is the only allowed contact with standard dimensions. Metal 2 can directly connect to only metal 1 in these rules. Not universal.

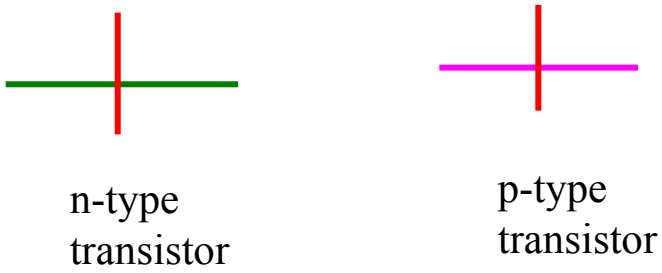


Stick diagram

- This is a simple representation of layout as a plan. All the layers are represented as lines without widths. A cross at a junction of two layers indicates a connection using contact.

- Useful to decide how to do actual physical layout by assigning dimensions to the lines with due consideration to design rules.

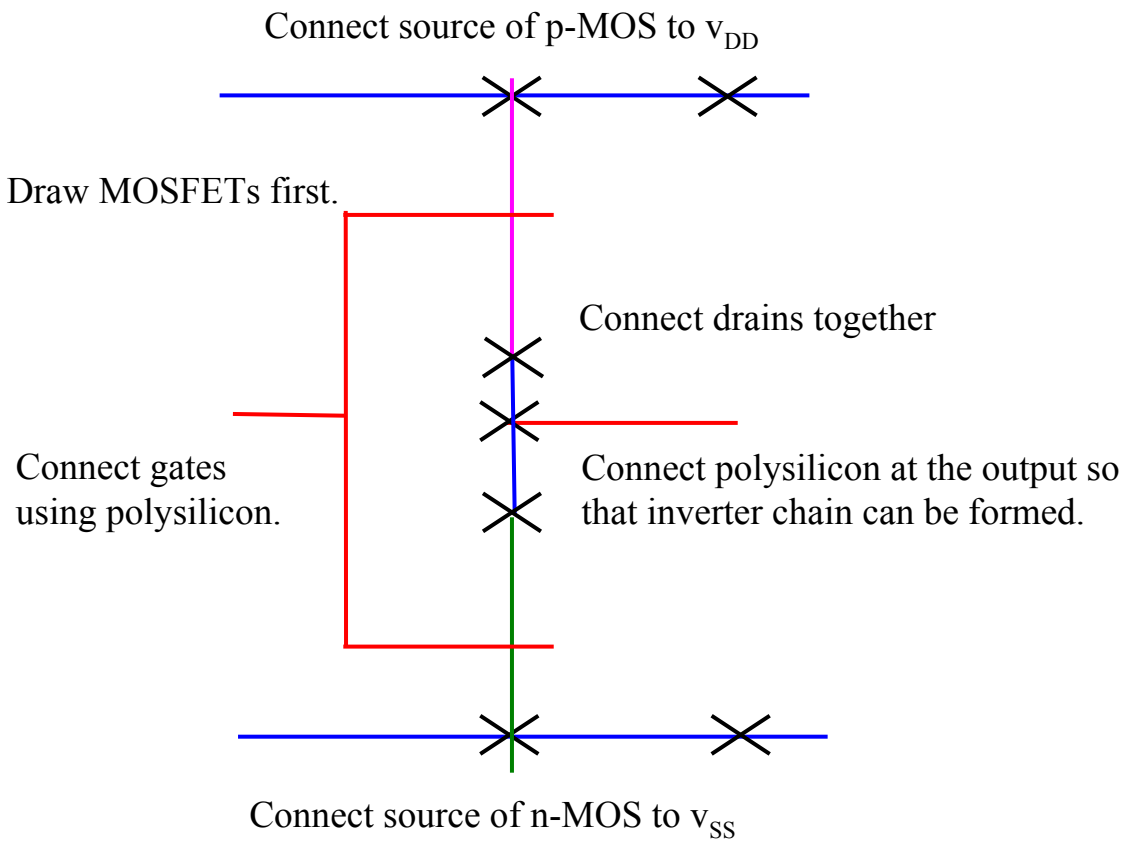


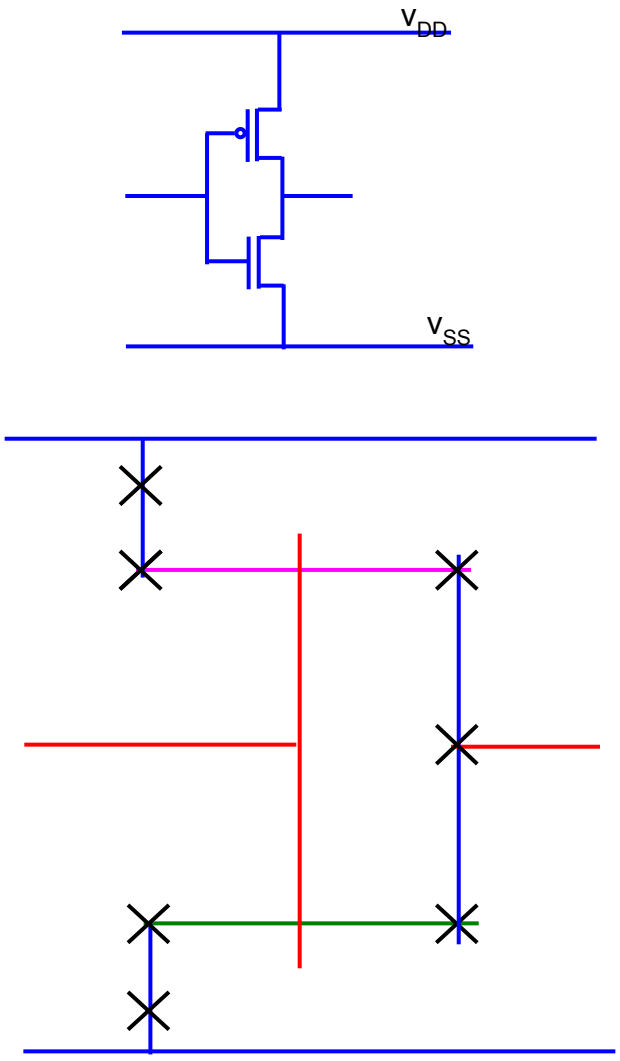


- A MOSFET is formed whenever a polysilicon wire crosses an N+ or P+ wire.

Inverter Layout

- Start with a stick diagram of an inverter. There are two possible versions. Full description given only once. Later only animations will be given. Any observations?





Inverter Layout

Place nwell and do final adjustments.

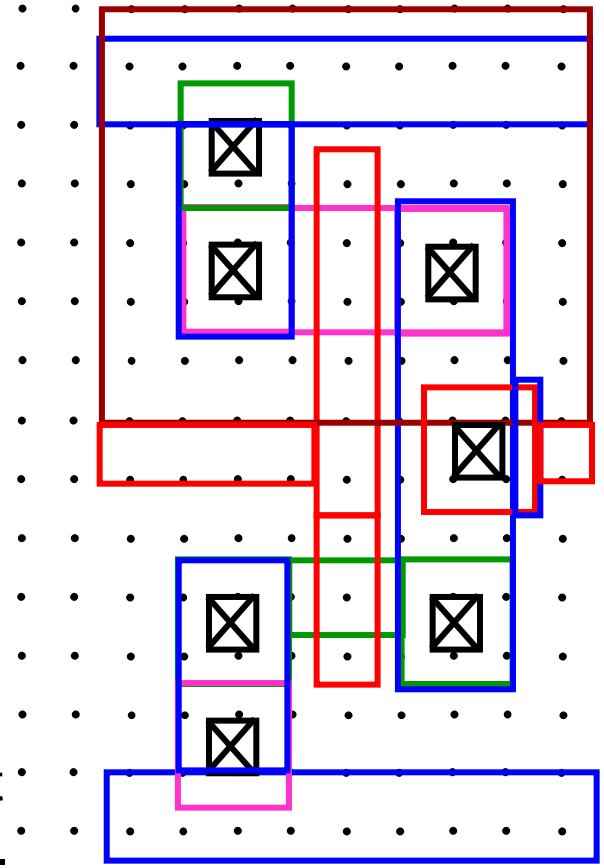
Draw p-channel 8λ away and complete gate.

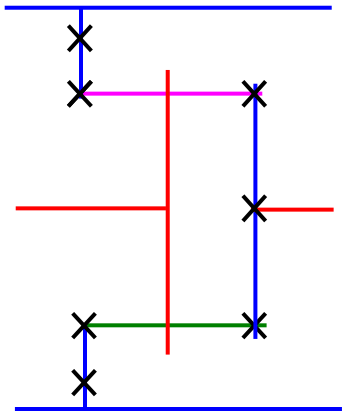
Draw n-MOSFET gate extending 2λ.

Draw drain contact on the right.

Stack two contacts on the left.

Start with lowest ground (0V) line.





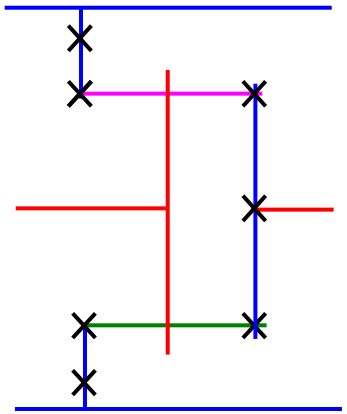
Complete p- and n-channel connections.

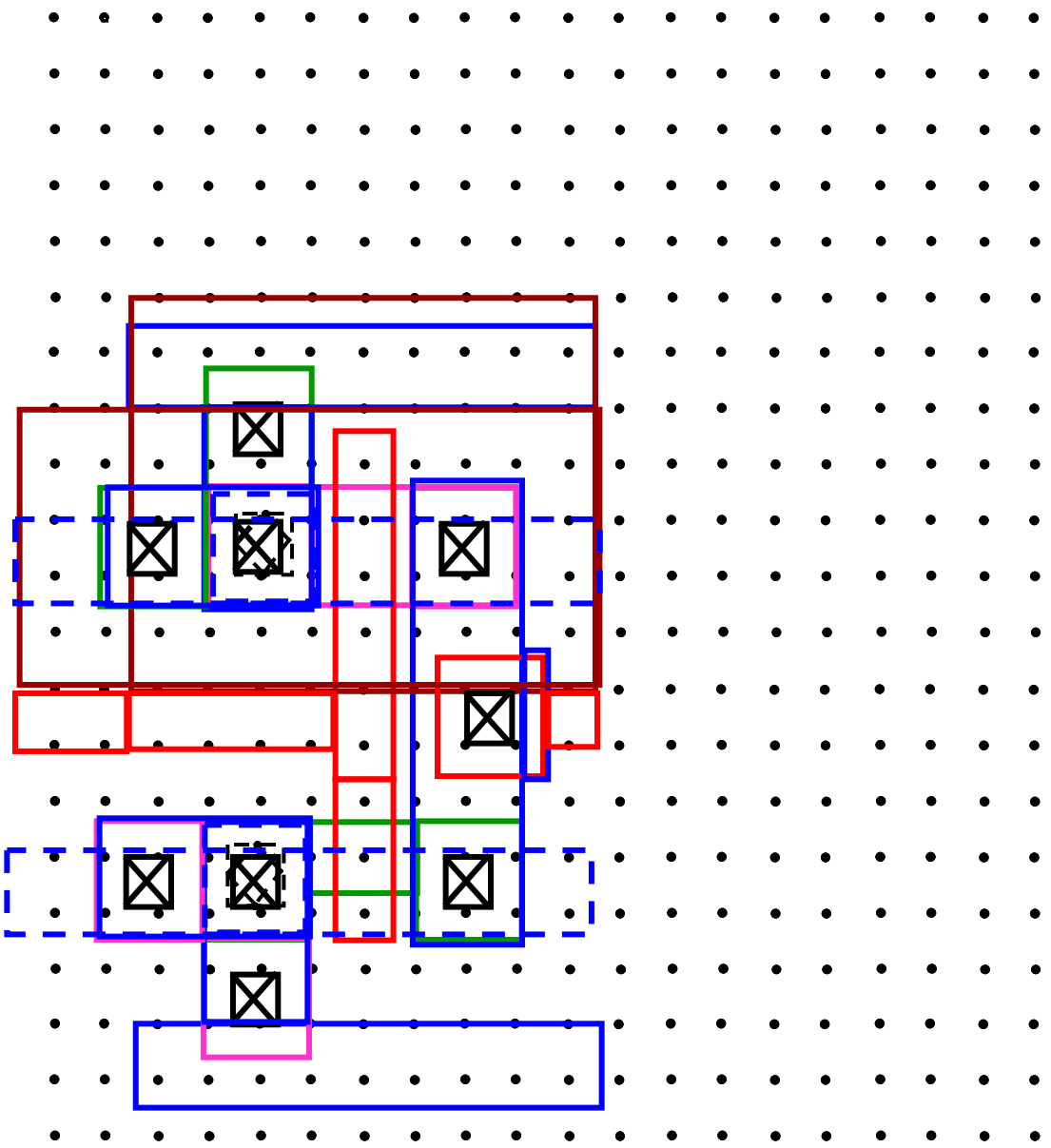
As poly is 2λ from contact, and W is 2λ, draw device.

Leave 3λ space for another metal.

Inverter Layout

Use metal 2 to reduce area.





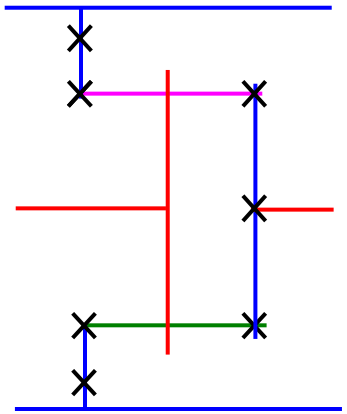
8λ design rule leaves no flexibility for output contact.

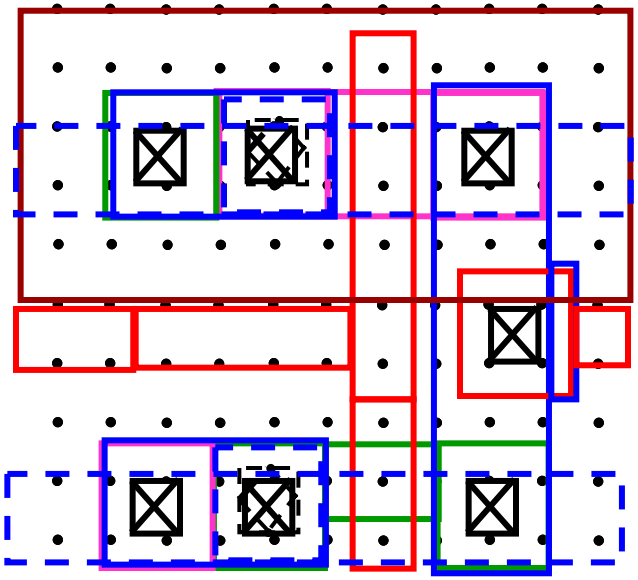
We can move substrate contacts up and route power and ground lines in metal 2 to save area.

Inverter Layout

Use metal 2 to reduce area.

It is noted here that only M2 can only be connected to M1 through a via, and not directly to p+/n+/poly .





8λ design rule leaves no flexibility for output contact.

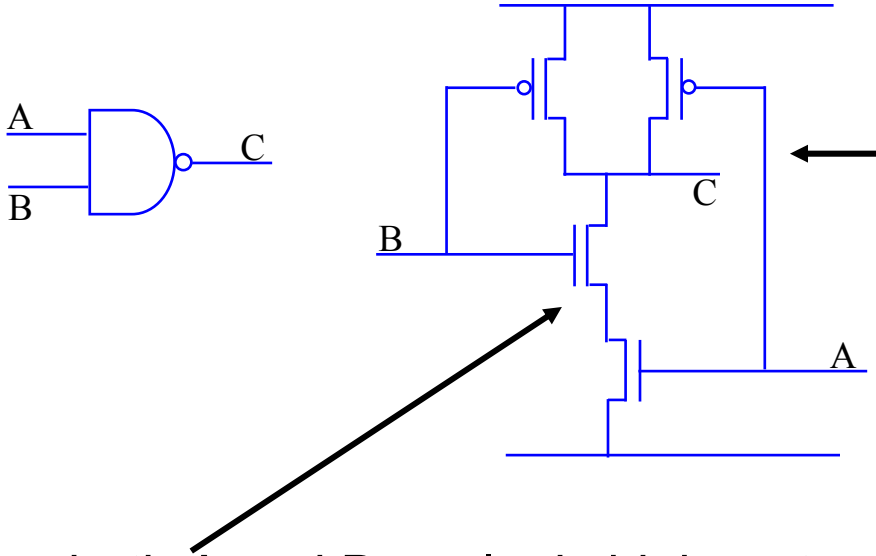
We can move substrate contacts up and route power and ground lines in metal 2 to save area.

CMOS Logic Implementations and Layout

- After discussing inverter and layout representations, we now move to general CMOS logic circuits.
- The design is carried out by deciding n-type pull down path that sets logic zero.
- P-type devices are complement of n-type network as will be clear when we look at designs.
- All power and delay aspects apply to these CMOS logic implementations also.

CMOS Logic Implementations and Layout

2-input NAND Gate

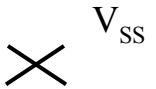


• p-MOS block is complementary to properly designed n-MOS block i.e. p-MOS devices are in parallel as n-MOS devices are in series.

• When both A and B are logic high, output is pulled down to zero as needed for NAND gate. Otherwise it remains high.

• If NAND gate above needs to have the same delay as an inverter driving identical load, the n-channel device width has to be doubled as two n-channel devices are in series effectively double device resistance and the delay. Similar considerations can be applied to other gates. But it is the longest series connected path that determines delay and sizing..

CMOS Logic Implementations and Layout

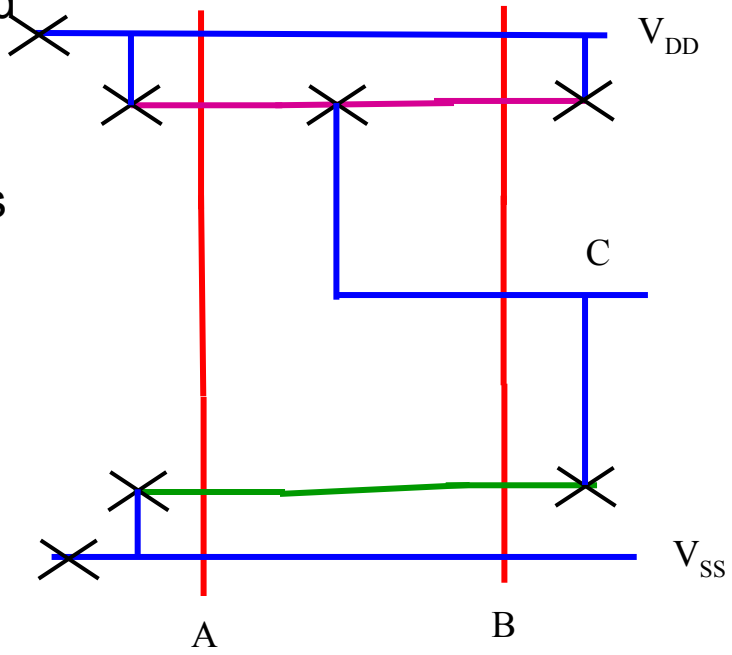


- Stick diagram of 2 input NAND gate.

Start with two n- and p-channel devices needed.

Connect both gates for inputs together.

Do substrate and n-MOS Ground connection.



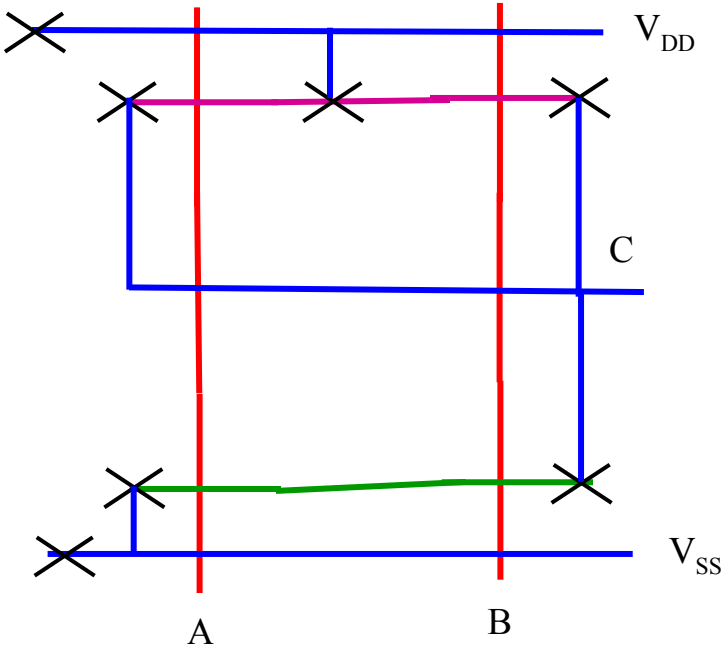
Do nwell and p-MOS Power supply connection.

Finally, do output voltage connection using common shared connection of p-MOS and right n-MOS connection of input B. This completes the stick diagram.

Connect common shared areas for n-channel and p-channel so as to save area for connections. There is no contact required for common n-MOS as it is not connected anywhere. for inputs together.

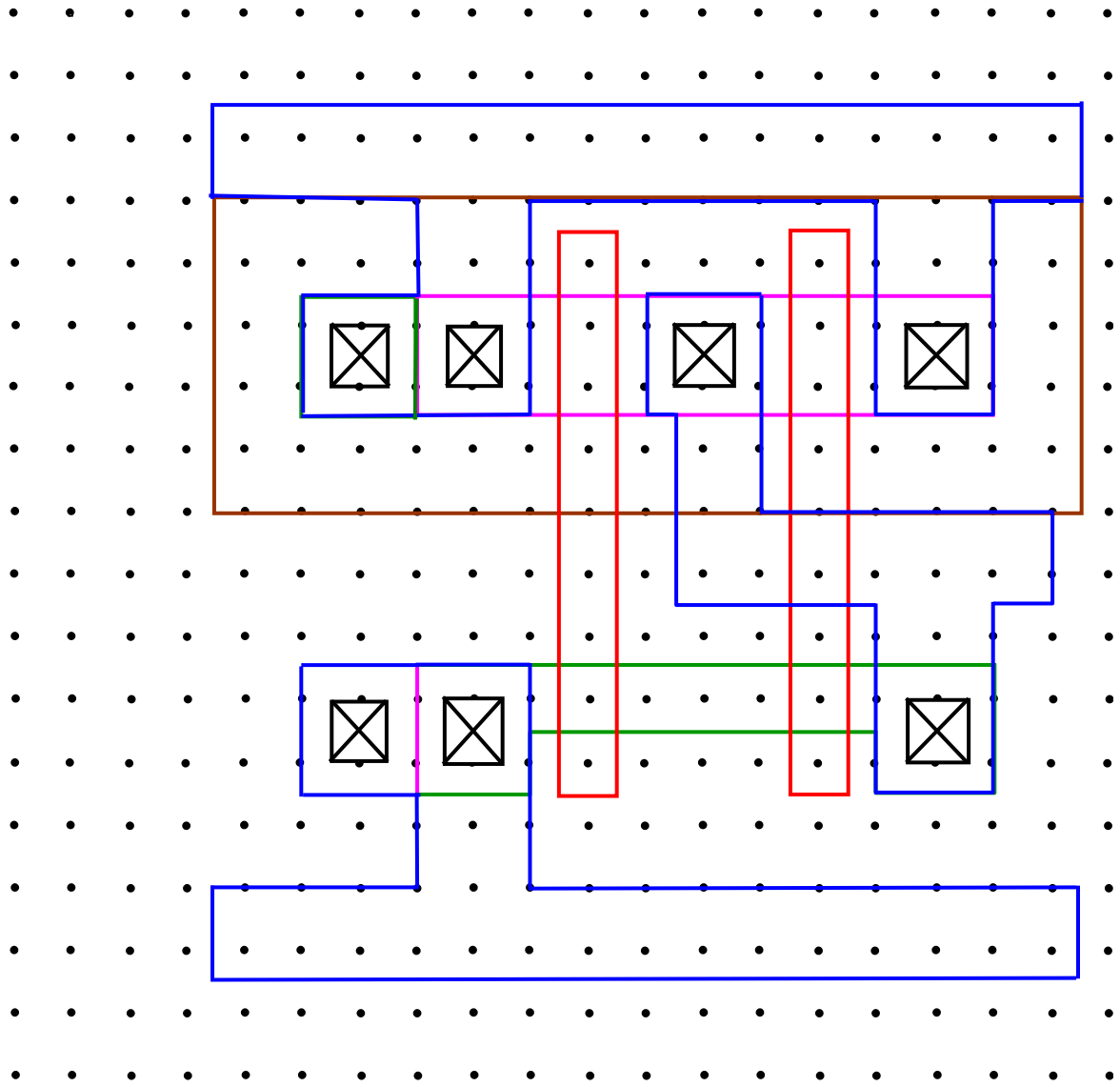
CMOS Logic Implementations and Layout

● Alternate Stick diagram of 2 input NAND gate results by using common p-MOS connection to power supply and left/right connections to output.



CMOS Logic Implementations and Layout

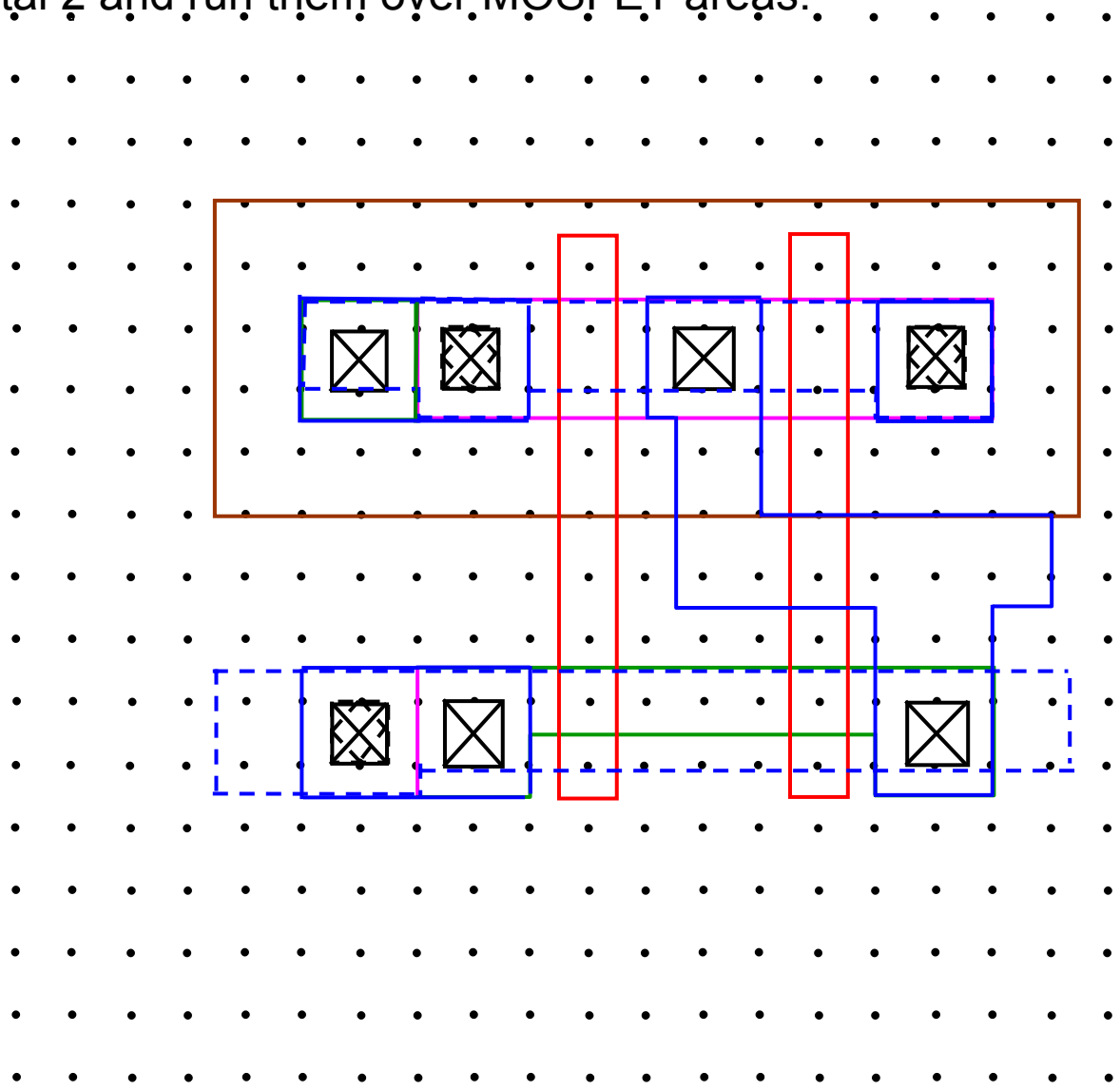
- Full layout of 2 input NAND gate of the first stick diagram is given below.



CMOS Logic Implementations and Layout

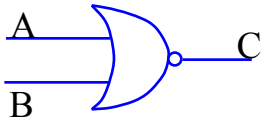
- Use of metal 2 saves area as before. The strategy again is to move power and ground lines to metal 2 and run them over MOSFET areas.

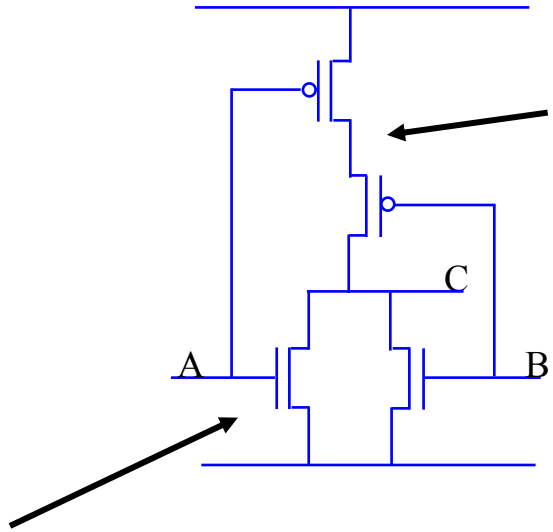
- Note that there is no power connection to common p-MOS point as there is no metal 1 to metal 2 contact there. Same way, right n-MOS contact is not connected to ground.



CMOS Logic Implementations and Layout

- The schematic diagram of 2 input NOR gate is given below.





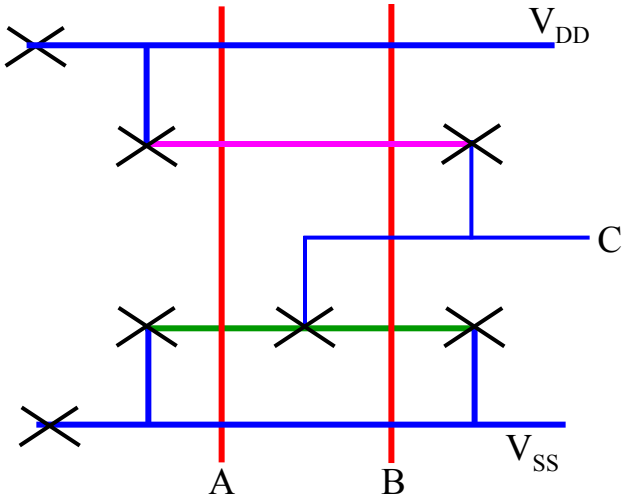
- Two p-MOS devices are complementary i.e. in series. Hence both A and B need to be low to make output high.

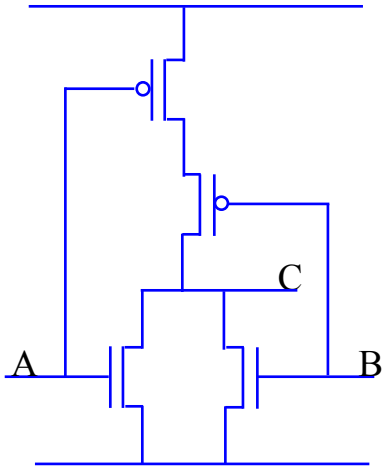
- Two n-MOS devices in parallel ensure that the output is pulled low even when on one of the inputs A or B is high as needed for NOR.

CMOS Logic Implementations and Layout

- Stick diagram of 2 input NOR gate is given below.

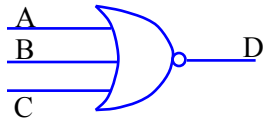
$C = \overline{A + B}$

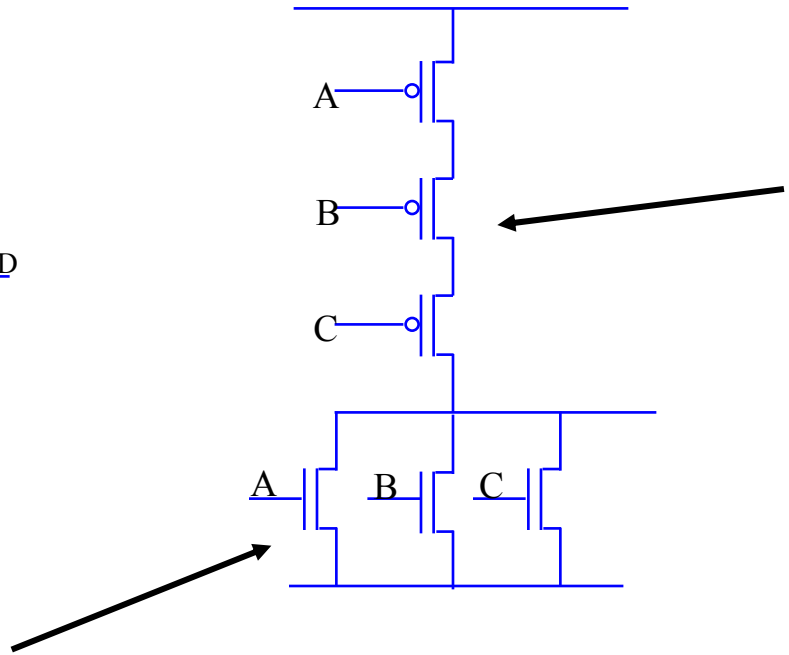




CMOS Logic Implementations and Layout

- The schematic diagram of 3 input NOR gate is given below.





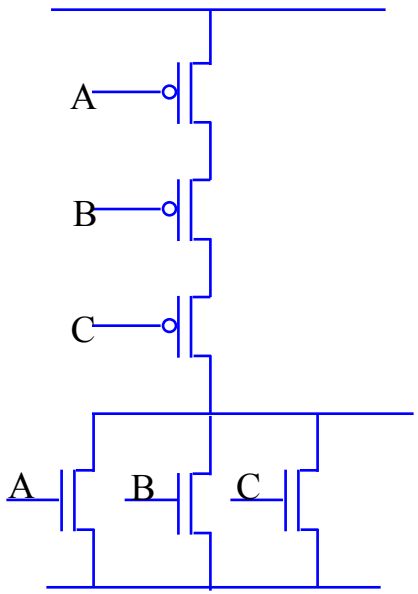
- Three p-MOS devices are complementary i.e. in series. Hence all A, B and C need to be low to make output high.

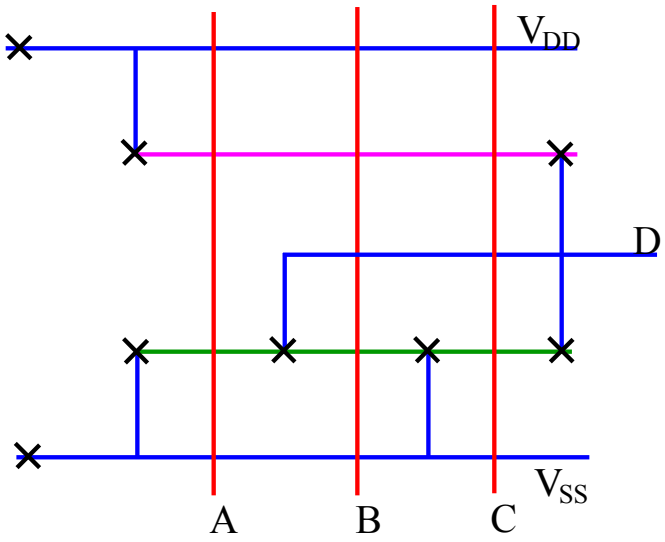
- Three n-MOS devices in parallel ensure that the output is pulled low even when on one of the inputs A or B or C is high as needed for NOR.

CMOS Logic Implementations and Layout

- Here is stick diagram of the 3 input NOR. Use single N+ or P+ region as far as possible.

$$D = \overline{A + B + C}$$

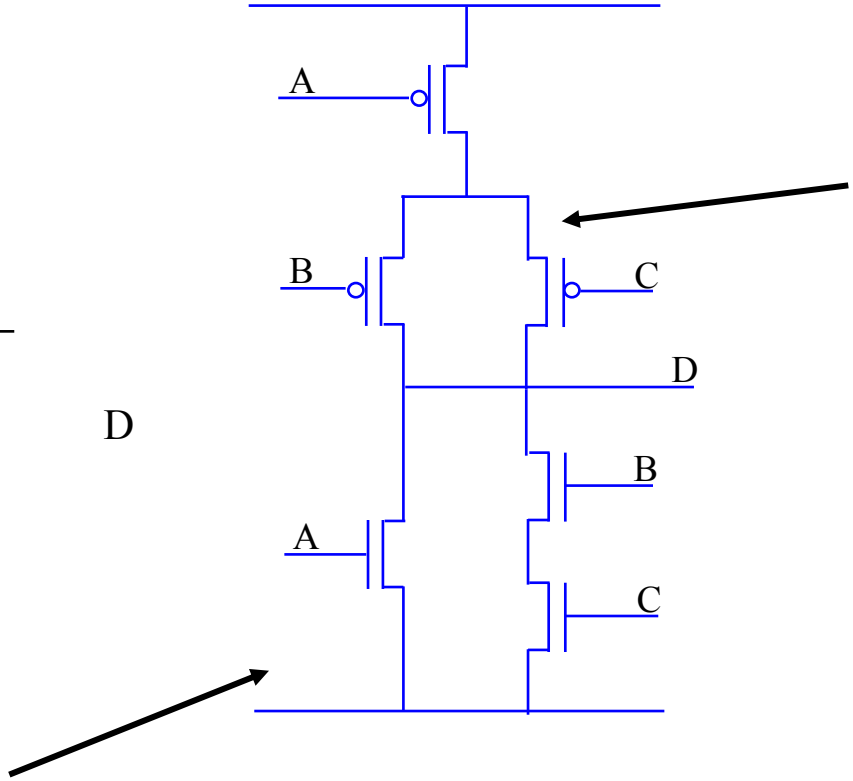




CMOS Logic Implementations and Layout

- Here is implementation of arbitrary combinational logic in CMOS.

$$D = \overline{A + B \cdot C}$$

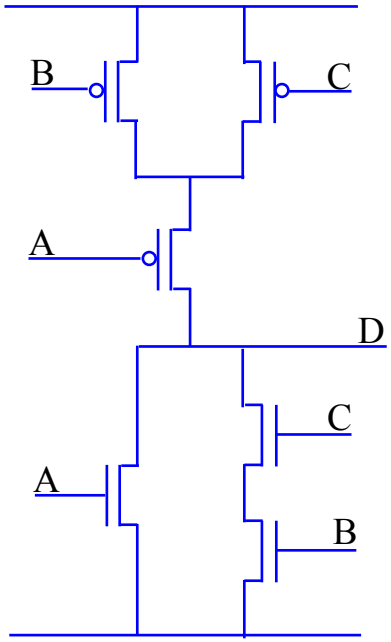


- Three p-MOS devices are complementary to n-MOS i.e. B and C p-MOS are in parallel and this combination is in series with p-MOS for A.

- Three n-MOS devices was arranged so that high A alone can pull the output D to low. Both B and C need to be high to pull D low if A is low.

CMOS Logic Implementations and Layout

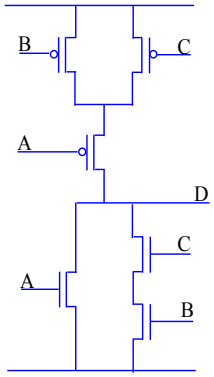
● Here is the same implementation of arbitrary combinational logic in CMOS with some rearrangement. This flexibility sometimes helps in connectivity and achieving better layout.

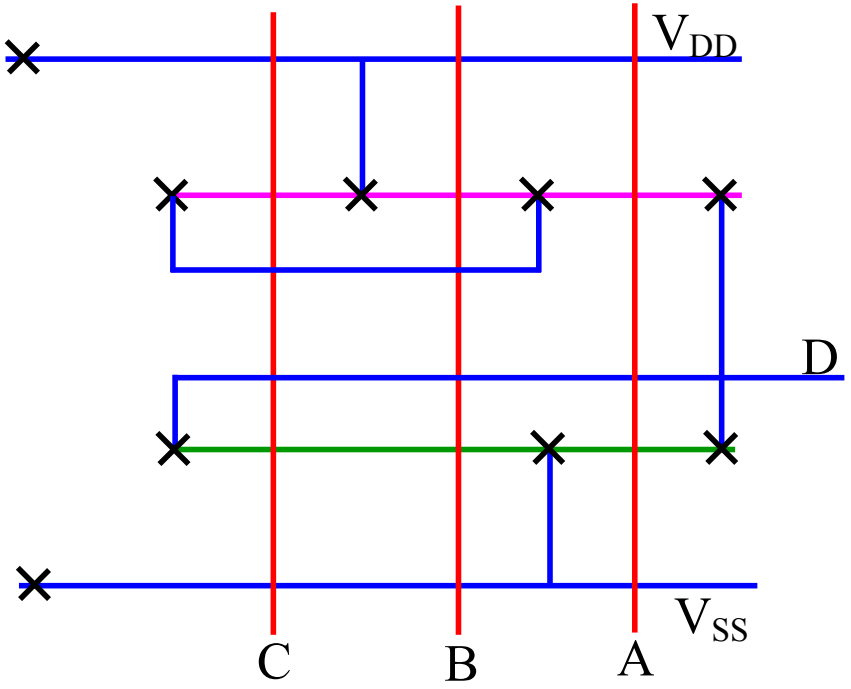


CMOS Logic Implementations and Layout

● Here is the stick diagram of the logic implemented. In this implementation, N+ to P+ spacing will be larger than 8λ as two independent metal lines have to be accommodated between them. How much is the spacing?

$$D = \overline{A + B \cdot C}$$

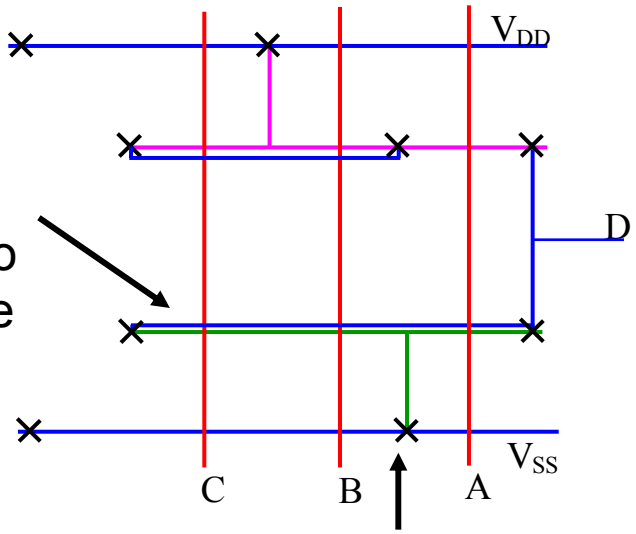




CMOS Logic Implementations and Layout

- This two metal line issue can be mitigated by running a metal line over N+ or P+. In the previous slides, all contacts were placed in a row and metal lines were pulled out to make connections. In the following style, part of the connections are implemented by pulling out P+ or N+ regions and then making contacts as shown. This does not require two metal lines between N+ and P+.

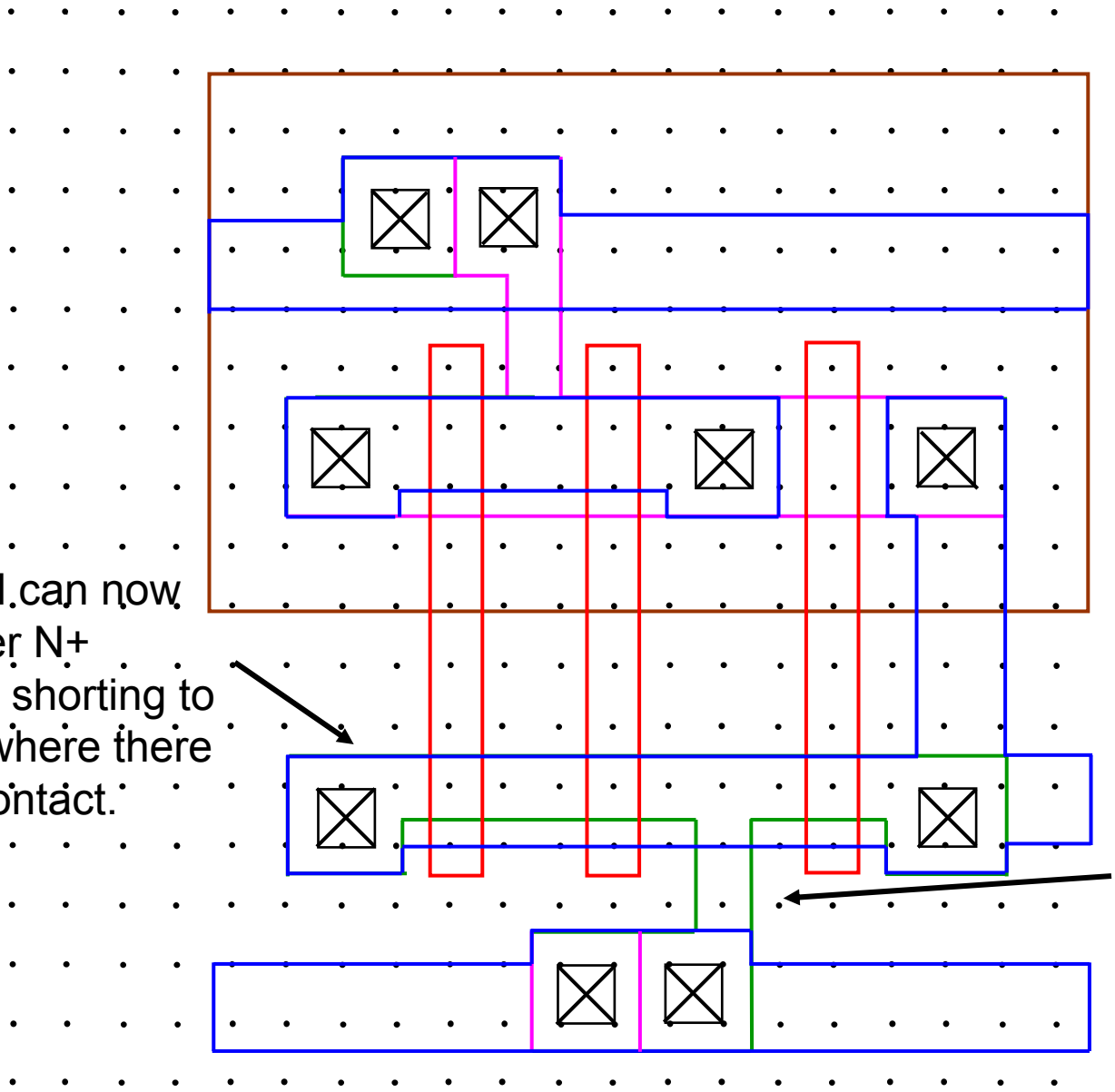
- Metal can now run over N+ without shorting to areas where there is no contact.



- Green N+ area pulled out to make desired connection to ground metal line.

CMOS Logic Implementations and Layout

- Full layout of this stick diagram is given as it is a new layout style.

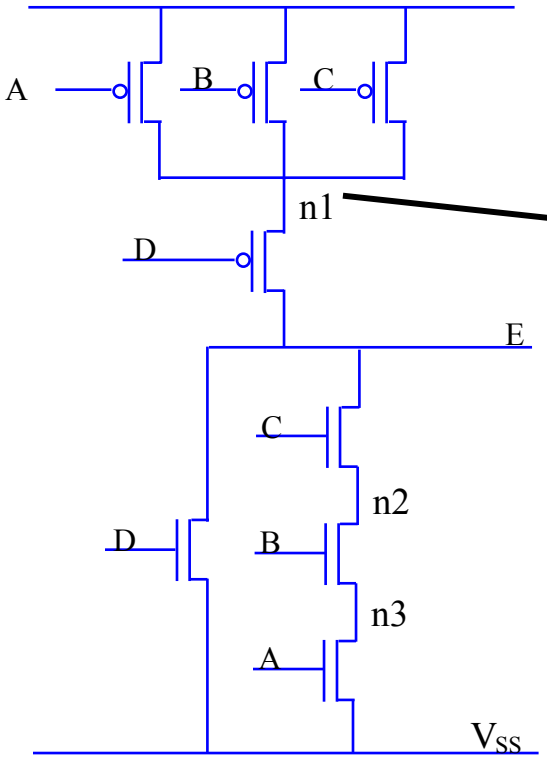


- Metal can now run over N+ without shorting to areas where there is no contact.

- Green N+ area pulled out to make desired connection to ground metal line.

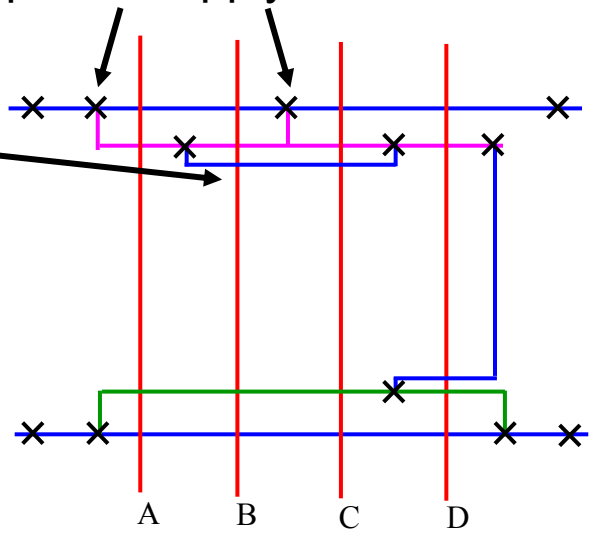
CMOS Logic Implementations and Layout

● Here is another arbitrary logic implemented in CMOS and corresponding stick diagram. What is the logic function implemented?



● This metal line connects 4 P+ regions at n1.

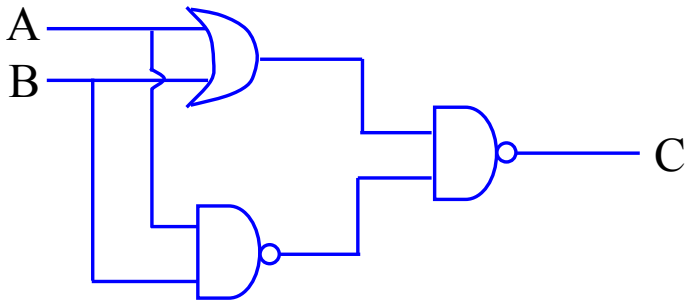
● These two contacts implement 3 source connections of p-MOS to power supply.



CMOS Logic Implementations and Layout

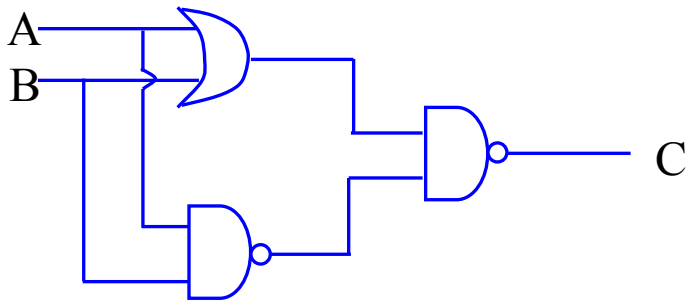
● Here is one of the most complex XNOR logic which will be implemented in CMOS. The OR gate is implemented as NOR gate and an inverter together. Two different stick diagrams are given with different layout strategy.

$$\text{XNOR}: \overline{(A + B) \cdot \overline{AB}}$$



CMOS Logic Implementations and Layout

● Here is one of the most complex XNOR logic which will be implemented in CMOS. The OR gate is implemented as NOR gate and an inverter together. Two different stick diagrams are given with different layout strategy.

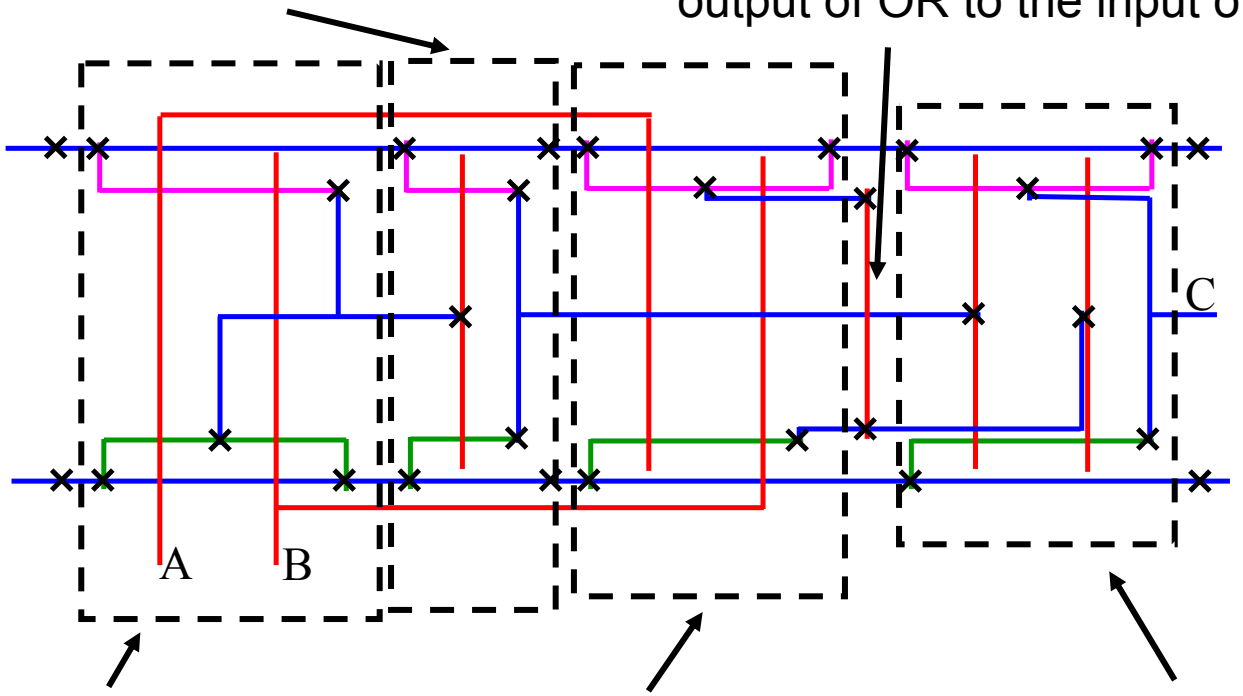


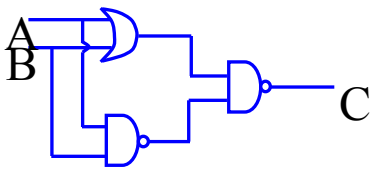
CMOS Logic Implementations and Layout

- Here is first stick diagram where all P+ and N+ regions align and connections are carried out using both metal and polysilicon.

- 2-input NOR output connects to inverter to give needed 2 input OR.

- Need to use a polysilicon jumper here as there is already a metal line that connects output of OR to the input of output stage





- Here is 2-input NOR gate with n-MOS in paprallel.

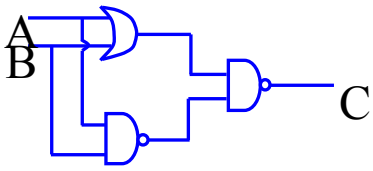
- Here is 2-input NAND connected to inputs A/B whose output is connected to input of output stage NAND.

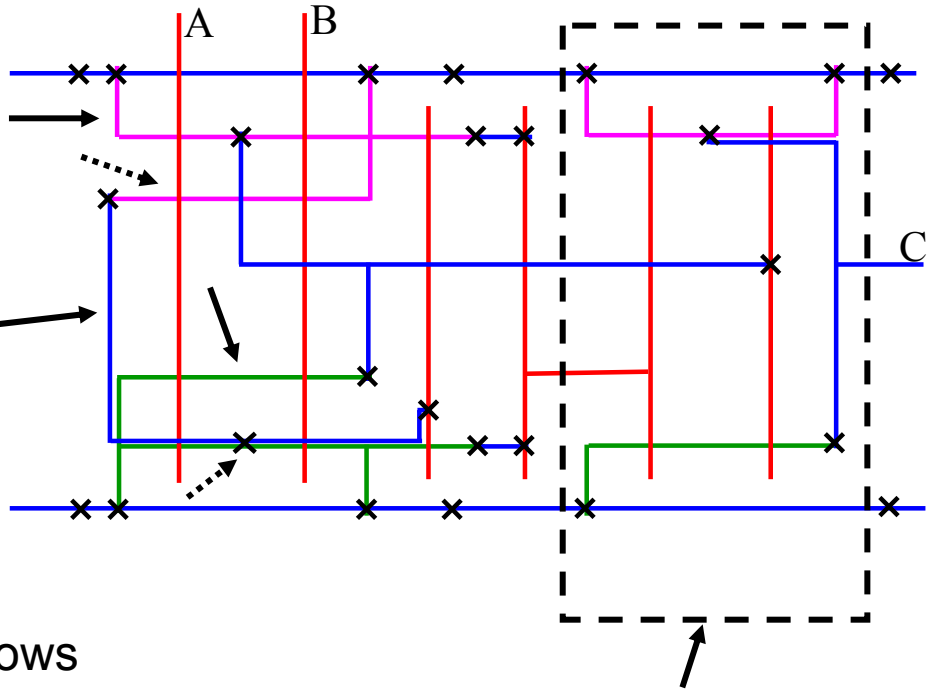
- The output stage 2-input NAND which has outputs of previous stages as inputs.

CMOS Logic Implementations and Layout

- Here is another stick layout of 2 input XNOR where stacked devices are used on the iput side NAND and NOR. What are expected changes?

- The two solid arrows show P+ and N+ regions that implement input side NAND.



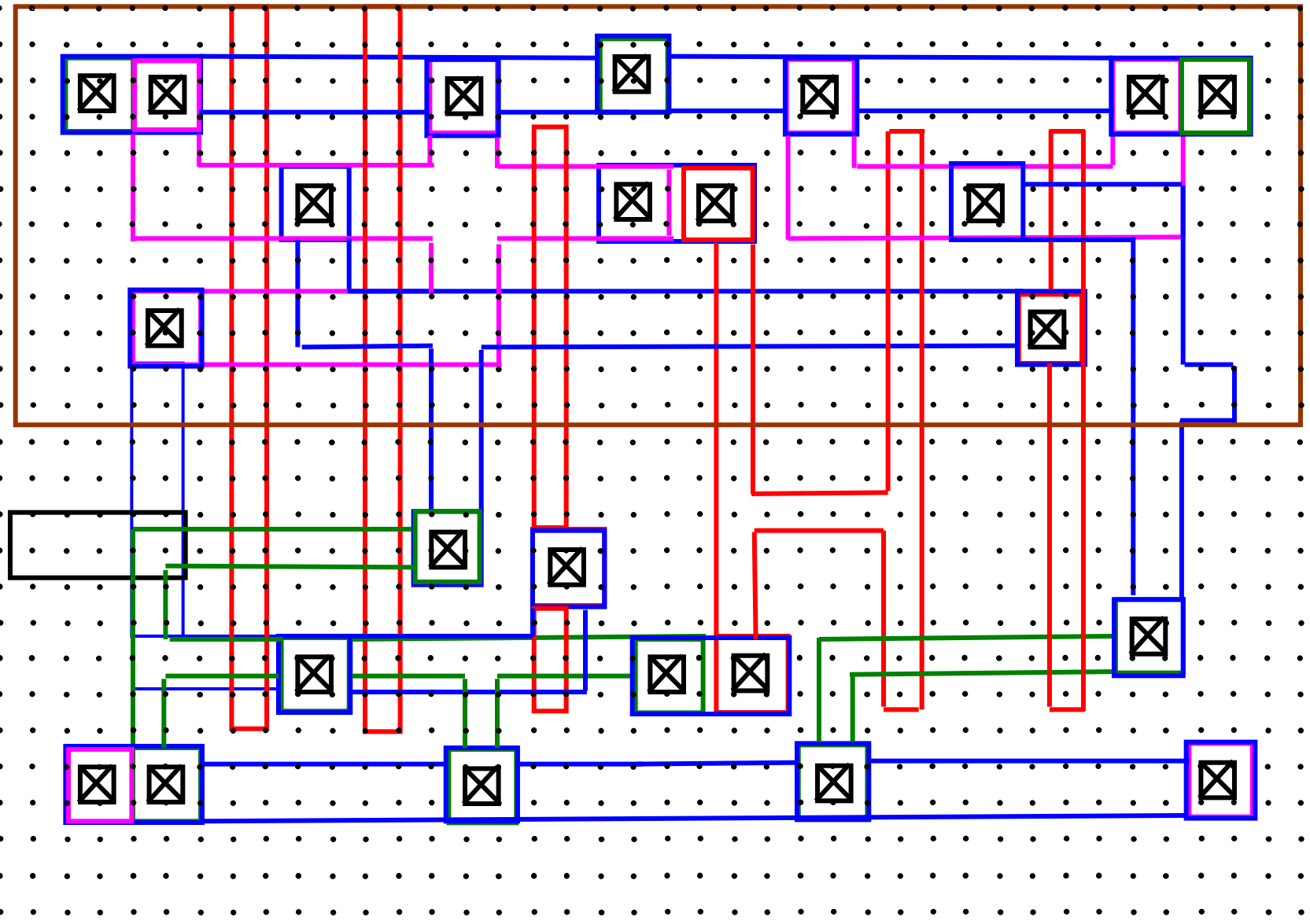


- The output of NOR is connected to the inverter to implement OR gate.

- The two dotted arrows show P+ and N+ regions that implement input side NOR.

- The output stage 2 input NAND with 2 n-MOS in series.

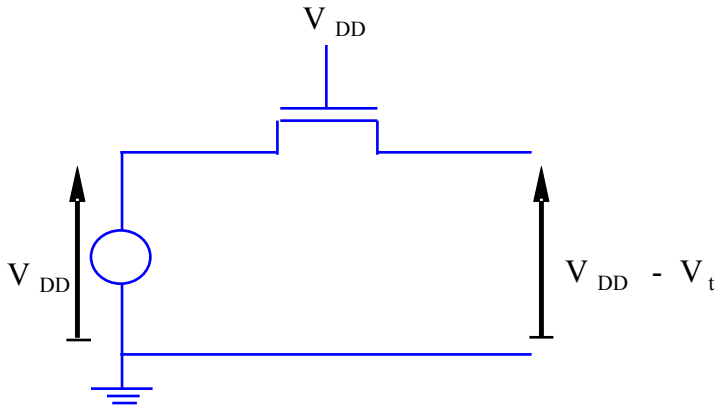
CMOS Logic Implementations and Layout



This shows full layout of XNOR with stacked devices. You may check it against stick diagram.

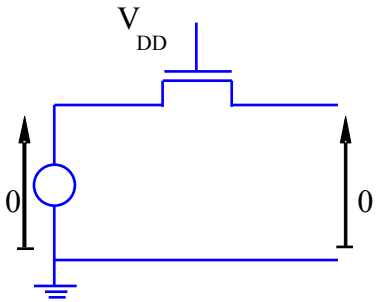
Transmission Gate Logic Implementations and Layout

● Transmission gate based logic provides an alternative to CMOS logic shown earlier. Here is n-MOS transmission gate. It is off when the gate is at 0V. In off state, input and output hold their logic states.



● When the gate is set to power supply voltage V_{DD}, the n-MOS device turns on. Here V_{DD} on the left is held fixed. The charge transfer to output will occur till the device turns off.

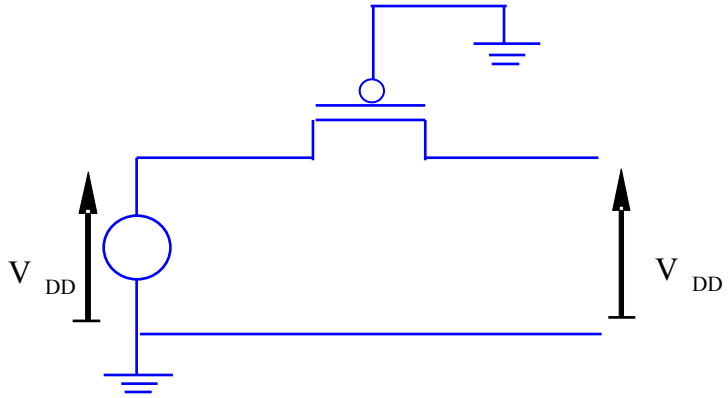
● The n-MOS device turns off when output reaches V_{DD} - V_t. Hence logic “1” is degraded in voltage value but is still logic high.



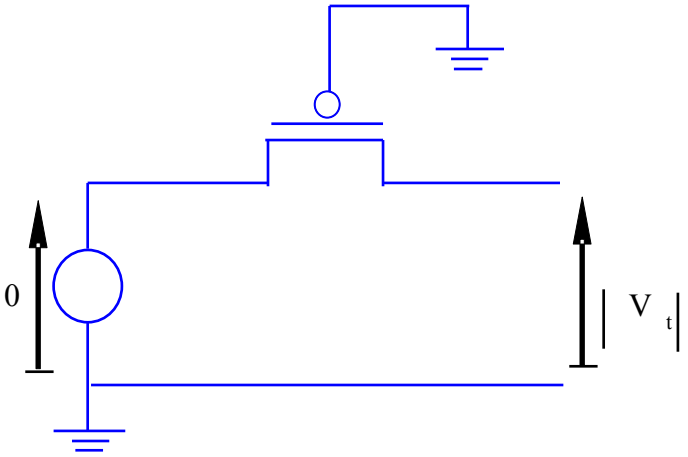
● For the case 0V on the left is the held fixed, again, the charge transfer to output will occur till the device turns off. This happens here only when the output is 0V. Hence logic 0 is good and reaches voltage of 0V.

Transmission Gate Logic Implementations and Layout

● Here is p-MOS transmission gate. It is off when the gate is at V_{DD}. In off state, input and output hold their logic states.



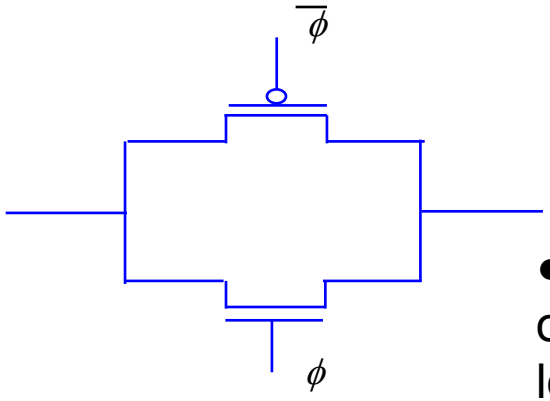
● When the gate is set to 0V, the p-MOS device turns on. Here V_{DD} on the left is the held fixed. The charge transfer to output will occur till the device turns off. Hence the output is charged to logic high and reaches V_{DD} with no degradation for logic high voltage.



● For the case 0V on the left is the held fixed, again, the charge transfer to output will occur till the device turns off. This happens here when the output is - V_t. Hence logic 0 is degraded and does not reach voltage of 0V.

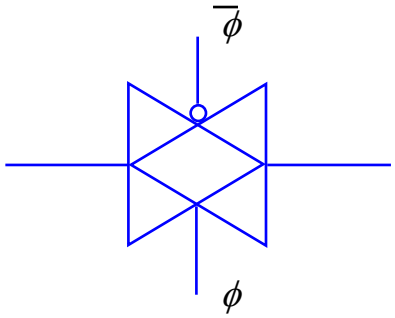
Transmission Gate Logic Implementations and Layout

● It is possible to have an implementation shown below using one n-MOS and p-MOS device each so that none of the logic value degrades.

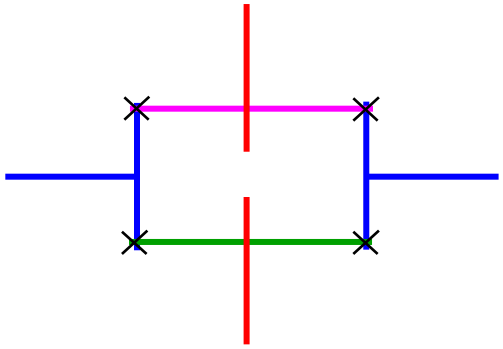


● When Φ is 0V, both devices are off and logic states are held at their values.

● When Φ is V_{DD}, both devices turn on and charge transfer occurs. If input side is held at 0V, the output is first pulled low by both the devices and eventually reduced to zero by n-MOS current discharge. For input side at V_{DD}, p-MOS charges the output to full V_{DD} with initial charging done by both devices.



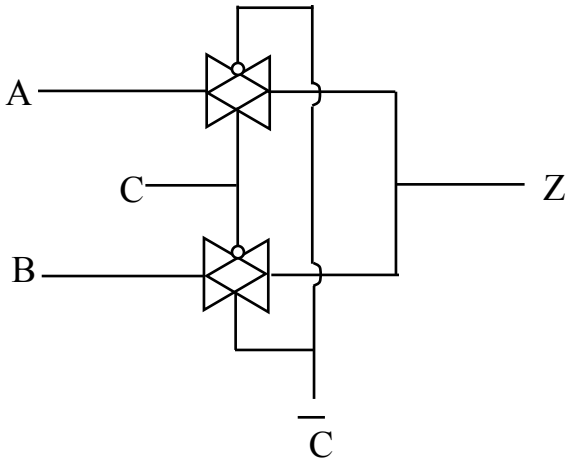
● Figure on the left shows compact notation for this transmission gate free of voltage level degradation and the stick diagram is shown on the right.



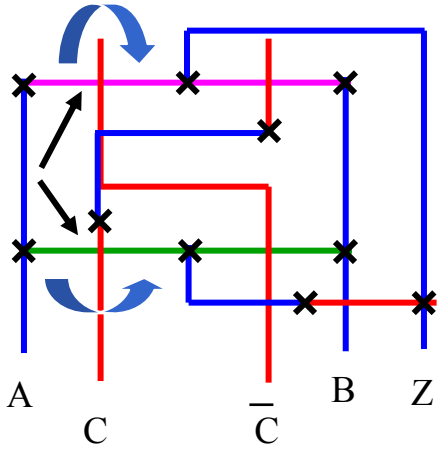
● Note that these gates are bidirectional i.e. either left or right side of the gate could be held logic state. Any side thus can be output side.

Transmission Gate Logic Implementations and Layout

● In some cases, such logic implementation is very efficient. 2-input multiplexer, which transfers an input to output, based on a control logic value, is one such example.



● Here if control C= logic 1, the output Z=A as only the top transmission gate is active. if control C= logic 0, the output Z=B as only the bottom transmission gate is active. Hence this implements a 2-input multiplexer and requires a total of 6 MOSFETs.

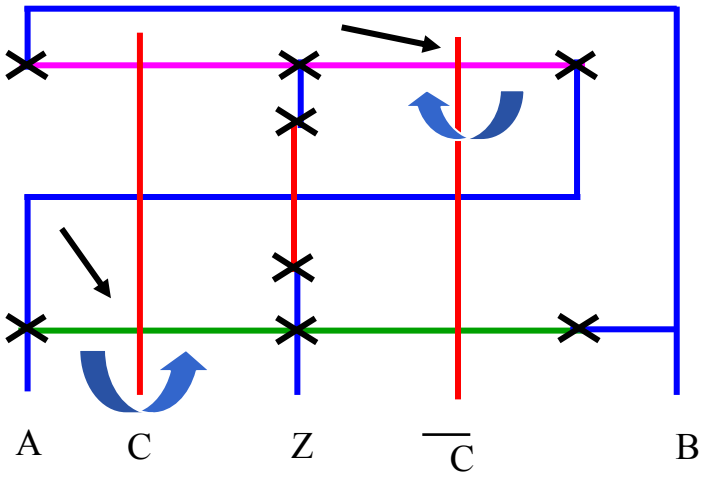


● Here is one possible stick diagram of this circuit which uses aligned input. C and inverted C signals cross, one is carried in polysilicon and the other uses additional metal jumper.

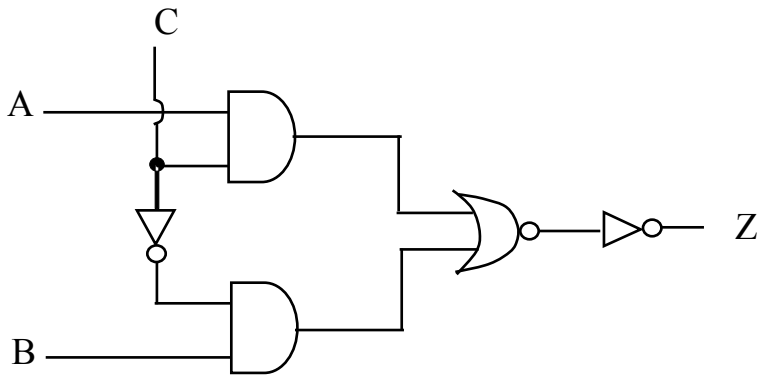
● When C is high, both left MOSFETs pointed by arrow turn on and place A at the output Z as desired.

Transmission Gate Logic Implementations and Layout

● Here is another stick diagram layout plan where the output Z is at the centre and control signal runs straight as gate polysilicon. .



● In this case, when C is high, left n-MOS and right p-MOS pointed by arrow turn on and place A at the output Z as desired.



● The benefits of using transmission gate logic in this case is clear by looking at the CMOS version. This version requires 20 MOSFETs although it can be further simplified.

Transmission Gate Logic Implementations and Layout

- In summary, processing step tolerances, lithography alignment uncertainty and reliability consideration dictate design rule.

- The design rules are technology independent and simple in this module.

- Any CMOS logic can be implemented by first designing logic zero pull down path with n-MOS devices and then construction of p-MOS complementary network.

- Any circuit implementation layout plan is first done with a stick diagram and then full layout can be done starting from one of the edges.