# Libre Silicon process specification

David Lanzendörfer

February 8, 2018

**Abstract**

Copyright © 2017 LANCEVILLE TECHNOLOGY GROUP CO., LIMITED. All rights reserved.

This process is licensed under the Libre Silicon public license; you can redistribute it and/or modify it under the terms of the Libre Silicon public license as published by the Libre Silicon alliance, either version 1 of the License, or (at your option) any later version.

This design is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the Libre Silicon Public License for more details.

This is the specification of the free silicon manufacturing standard for manufacturing the LibreSilicon standard logic cells[1] and related free technology nodes from the LibreSilicon project.

---

[1] https://github.com/chipforge/StdCellLib

# Contents

# 1   CMOS in a nutshell

This basic initial project is dedicated to the CMOS Technology only and for this reason two types of metal–oxide–semiconductor field-effect transistors (MOSFET) are required.

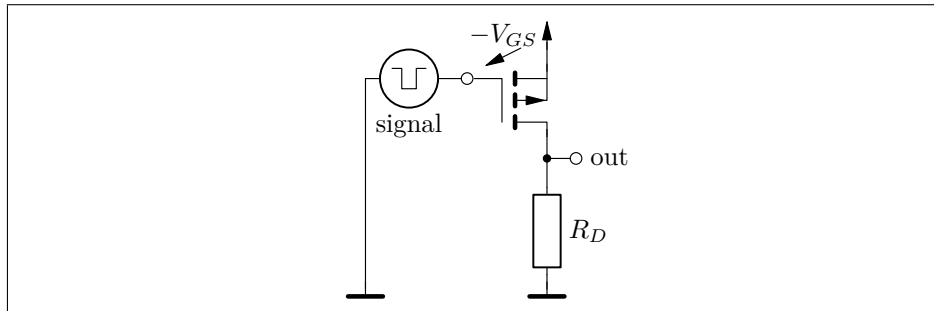Historicaly, the first chips with MOSFETs on the mass market were p-channel MOSFETs in enhancement-mode.



**Figure 1:** enhancement-mode PMOS transistor use-case

The sectional view of a PMOS transistor in silicon is being shown below
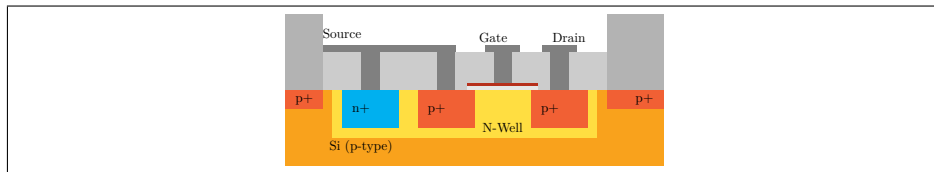


**Figure 2:** Sectional view of a PMOS transistor

Historicaly later, faster chips with MOSFETs on the mass market were marked as n-channel MOSFETs in enhancement mode also.
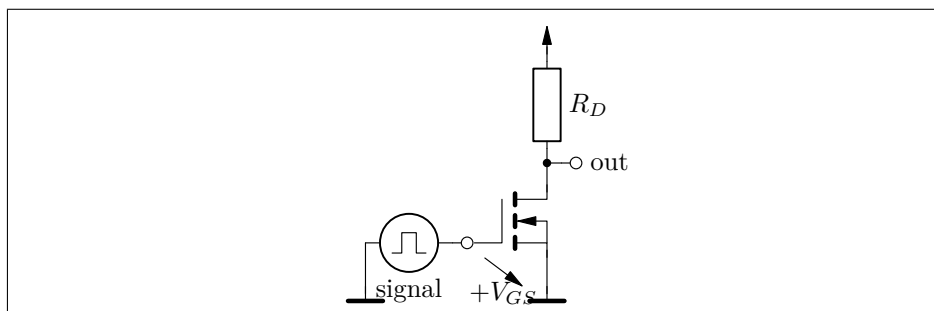


**Figure 3:** enhancement-mode NMOS transistor use-case

The sectional view of a NMOS transistor in silicon is being shown here also.
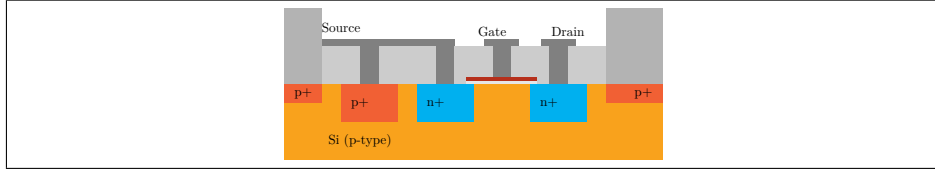
**Figure 4:** Sectional view of a NMOS transistor

Both technologies, the older NMOS as the newer PMOS, have the same disadvantage. Every time, the transistor is switched on, the current between Drain and Source of the transistor is limited by the Resistor on Drain only. Higher currents here meaning also higher power consumption for the chip where the transistors are integrated also. If the transistors are switched off, now currents flows between Drain and Source anymore, the power consumption of the chip also goes low. Et violà, the US-Patent with Number 3356858[2] changed the world and combines both technologies to the new complementary metal-oxide-semiconductor (CMOS) technology. Instead of every transistor is working against a weak resistor, the transistor works against a complementary switched-off transistor. With the Eyes of our antecessor CMOS doubles the transistor count, but contemporary chips all are build in CMOS.
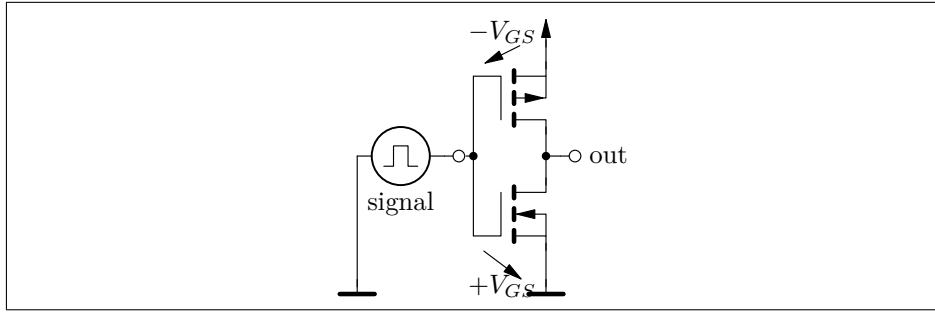


**Figure 5:** complementary PMOS and NMOS transistor couple use-case

Below the sectional view of the inverter circuitry can be seen. For the run through of this process we will use this cross section diagram as reference.
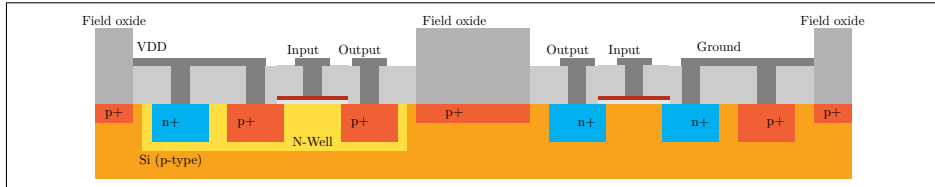


**Figure 6:** Sectional view of a NMOS-PMOS transistor circuit

[2]https://www.google.com/patents/US3356858

# 2 Physics

## 2.1 Infusion

The redistribution process depends on the ratio of the solubility of the doping material in silicon and $SiO_2$. At the $Si/SiO_2$ interface the dopants are redistributed by segregation until the ratio of their concentration at the interface is the same as the ratio of their solubility in both materials. The ratio of dopant solubility is expressed by the segregation coefficient $m$ which is

$$m = \frac{\text{solubility in silicon}}{\text{solubility in } SiO_2} \qquad (1)$$

As listed in Table 1 below there are dopant species which solubilize better in $SiO_2$ than in silicon ($m < 1$) and species which have a reversed behavior ($m > 1$). In case of $m < 1$, as for Boron, the dopant concentration is enhanced at the $SiO_2$ side, whereas beneath the interface, there is a dopant depletion at the silicon surface. For reversed solubility ratios ($m > 1$, like Phosphorus), only few dopant atoms penetrate the interface. In order to obtain the by $m$ determined concentration ratio at the interface, dopant atoms from deeper silicon zones diffuse back to the surface zone. Therefore, the dopant concentration at the silicon surface is enhanced, as illustrated in Figure 7b. In Figure 7 $C_c$ denotes the dopant concentration in the silicon surface zone before oxidation. $x$ is the distance from the silicon surface.
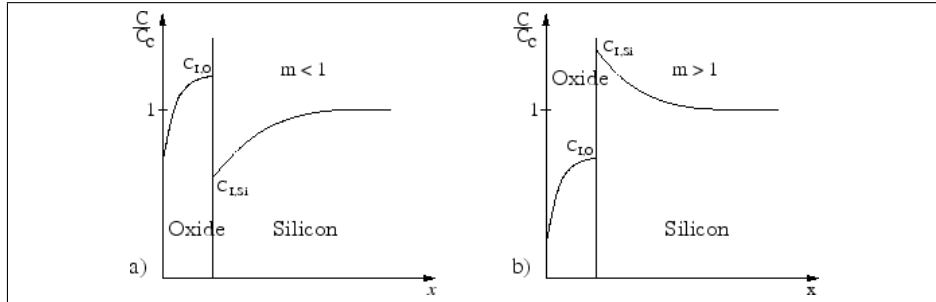


**Figure 7:** Schematic illustration of dopant redistribution

| Dopant species | Boron | Phosphor | Antimon | Arsen | Gallium |
|---|---|---|---|---|---|
| $m$ | 0.1-0.3 | 10 | 10 | 10 | 20 |

Table 1: Segregation coefficients $m$ for important dopant species in silicon

# 3 Chemistry

## 3.1 Etching silicon dioxide

A very "selective" chemical for SiO2 - i.e. does not etch silicon at all - is hydrofluoric acid (HF). If used directly such etchant has a too fast and aggresive action on the oxide, making very difficult the undercut and the linewidth control. For such reason, HF is universally used as a "buffered" solution, which can keep the etch rate low and constant, by moderating the PH level of the bath. This allows the etching time to be reliably correlated to the etching depth.

The industry standard buffered hydrofluoric acid solution (BHF) has the following formulation:

- 6 volumes of ammonium floride (NH4F, 40

This can be prepared, for example, by mixing 113 g of NH4F in 170 ml of H2O, and adding 28 ml of HF. The etch rate at room temperature can range from 1000 to 2500 Å/min. This depends on the actual density of the oxide which, as an amorphous layer, can have a more compact structure (if thermally grown in is oxygen) or less compact (if grown by CVD). The following etching reaction holds:

$$SiO_2 + 6HF \rightarrow H_2SiF_6 + H_2O \qquad (2)$$
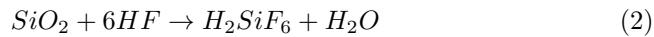
where $H_2SiF_6$ is water soluble.

Sometimes the BHF is prepared and stored according to the formulation given above, but it is 7:1 diluted in water just before being used. This allows an even better control of the etching rate. It is good practice to use the diluited BHF only once and then discard it, in order to assure process repeatability. If a diluited BHF bath at 35°C is used, the etching rate for thermal oxide is around 800Å/min.

Another popular etching formulation is the P-etch:

60 volumes of $H_2O$ + 3 vol. of HF + 2 vol. of $HNO_3$, that is: 300 ml of $H_2O$ + 15 ml of HF + 10 ml of $HNO_3$.

The P-etch action is strongly dependent on oxide density, as it results from the growth technique. An example is reported in the literature[3], indicating 120 Å/min for thermal oxide and 250-700 Å/min for sputtered oxide.

A slow etching bath is preferred for opening mask windows for a silicon substrate. However, the etching process could be used just for removing the oxide film from the whole surface. In this case the etching speed is not critical, and a fast solution can be used, such as HF diluited 1:10 in water. The etching time can be easily evaluated by visually inspecting the surface. Once the oxide film is removed, the metal-grey color of the silicon surface appears.

Sometimes a very light etch is required, for removing just a few atomic layers. This is the case of surface cleaning and decontamination. HF diluited 1 : 50 in water can be used. The etching speed will be around 70 Å / min. For example, a typical 50 Å "native" oxide on silicon can be removed with a 45 - 50 sec light etch.

---

[3]A. Pliskin, J.Vac.Sci Technol., vol. 14, p.1064, 1977

## 3.2 Etching silicon nitride

Thin films made of amorphous silicon nitride ($Si_3N_4$) are usually deposited by chemical vapour deposition from silane ($SiH_4$) and ammonia ($NH_3$). Since they act as a barrier for water and sodium, they have a major role as passivation layers in microchip fabrication. Patterned nitride layers are also used as a mask for spatially selective silicon oxide growth, and as an etch mask when $SiO_2$ masks cannot be used.

One example of the latter situation is given by the anisotropic etching of silicon in KOH. The etching rate of $SiO_2$ in KOH is nearly 1000 times slower than the etching rate of silicon, and in most cases a $SiO_2$ mask can be used successfully. However, a very deep selective etch may require a long etching time, and the 1000:1 etching rate ratio may result still too small to prevent the $SiO_2$ mask from being etched off before the process is completed. In this circumstance $Si_3N_4$, thanks to its reduced etched rate, can successfully replace the oxide mask layer.

The wet etching of nitride films is often performed in concentrated hot orthophosphoric acid ($H_3PO_4$). The bath temperature can range from 150°C to 180°C (boiling point) with a corresponding etch rate between 10 and 100 Å/min. It is good practice to bring the vapours into contact with a cold surface and to drive the condensed liquid back into the etching bath. This technique is referred to as "reflux".

The etching rates of silicon nitride, silicon oxide, and silicon in $H_3PO_4$ are respectively in the 50 : 5 : 1 ratio.

# 4 Process

Below the general flow chart of the overall process flow can be seen in Figure 8. These process steps will be discussed within the following sections.
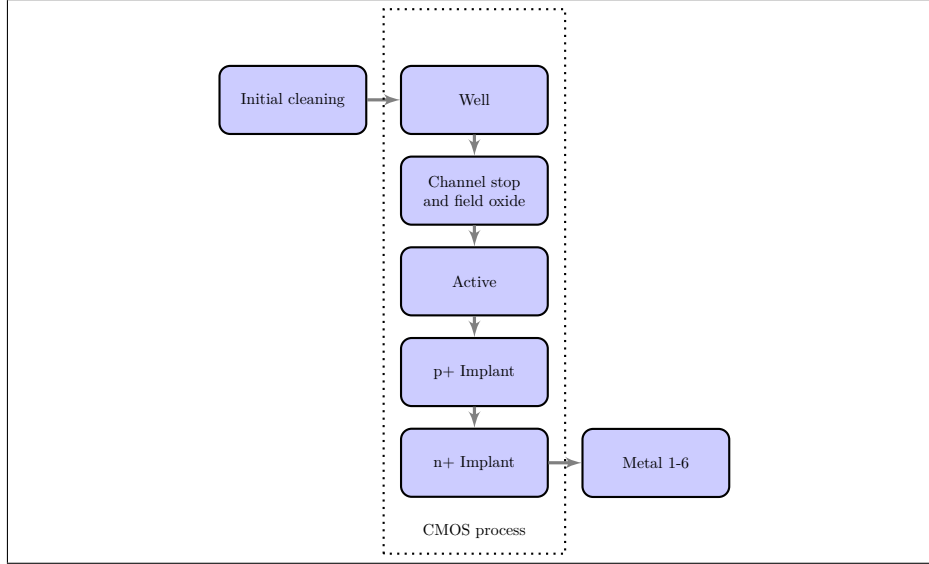


**Figure 8:** Frontend and backend process flow

The five overall process steps are part of an active part of the technology, while the final metal (respectively contact) layers will be used for making a contact between the logic gates and macro cells and making them available to the exterior world. For this process p-substrate is the required basic substrate, but forks and modifications will be very well possible based on a Graphene substrate or alike, still under the LSPL. The starting material is a p-type, <100> oriented silicon with a doping concentration of $\approx 9 \times 10^{14} cm^{-3}$.

**Reasons for using p-substrate**:

- We can't use two different substrates for our design because in the design both PMOS and NMOS is present. We have to choose which is more beneficial from fabrication point of view. In general or say it's true that NMOS devices are always more in the Semiconductor Industry in comparison to PMOS devices. For your reference-SRAM requires 6 transistors (4 NMOS, 2 PMOS).

- Another reason for more number of NMOS is because of difference of mobility of electron and holes. Electron mobility is almost twice of holes mobility and because of this ON-RESISTANCE of n-channel device is half of p-channel device with the same geometry and under the same operating conditions. That means to achieve same impedance size of n-channel transistors is almost half of p-channel devices. Same thing I can say in the different way that for same size of wafer, we can have more number of NMOS (means can perform more logical operation) in comparison to PMOS.

## 4.1 Initial cleaning

In order to remove the initial naturally grown silicon dioxide from the wafer, acid is being applied to the wafer which leads to a pure silicon substrate wafer as in the process illustration shown below.



**Figure 9:** Initial cleaning

This needs to be done because the naturally grown initially existing silicon oxide is not pure and may contain contamination which may render the final product unusable.

### 4.1.1 Sulfuric Cleaning

The sulfuric acid mixture, $H_2SO_4 + H_2O_2$ is being applied to the wafer for 10 minutes at a temperature of 120 °C.

### 4.1.2 HF dip

After the sulfuric cleaning a HF (HF:$H_2O$ , 1:50) dip is being performed for one minute.
Hydrofluoric acid (HF) is used to remove native silicon dioxide from wafers. Since it acts quickly, one needs to only expose the wafer for a short time ("dip"). After that the wafer needs to be dried and quickly processed further before new uncontrolled natural oxide can build up on the wafer through the contact with air.

## 4.2 Well

In order to build CMOS (Complementary metal–oxide–semiconductor/P and N MOS) on the same substrate an n-well is required for building the complementary P-channel transistor for a n-p-channel logic circuitry as shown above in the example section. The n-well will serve us as an island of n-doped substrate within the p-doped basis substrate. The cross section as well as the top view of the targeted geometry are shown below.



### 4.2.1 Dioxide layer

In order to selectively inject charge carrying atoms into the crystalline structure a protective dioxide ($SiO_2$) layer needs to be grown on top of a p-type substrate.



The industrial best practice is a layer of around (500nm≈5000Å) thickness or more. For this purpose the wafer is being oxidized for at least 90 minutes at 1000°C using wet oxidation which results in a dioxide layer at least 500nm(≈5000Å) in thickness.

### 4.2.2 Patterning

The resist is being deposited using spin coating and then baked depending on the baking time for the specific resist. The layout for being exposed onto the resist is being extracted from the "nwell" layer within the GDS2 file.



11

The thickness of the resist layer and the backing duration will variate depending on the specific equipment for which this process will be implemented with.

### 4.2.3 Etching
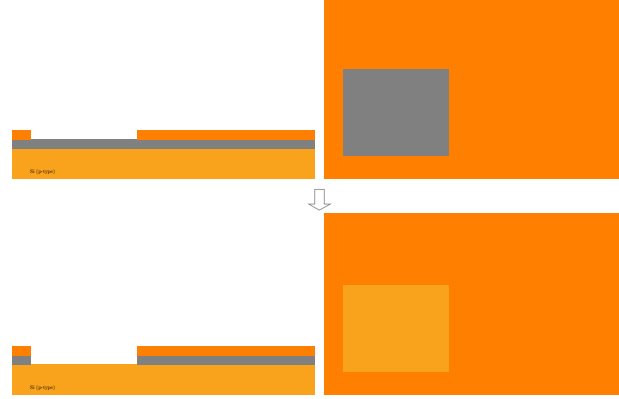


### 4.2.4 Cleaning



### 4.2.5 Injection
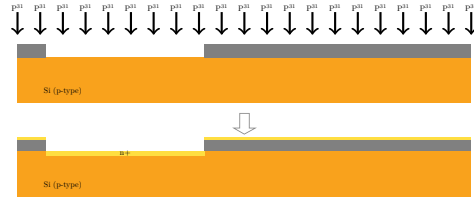


The n-well is implanted with a Phosphorus ($P^{31}$) dose of $2.5 \times 10^{12} cm^{-2}$ at an energy of 100 KeV. The n-well is then annealed.

### 4.2.6 Oxide for drive-in



The wafer is being oxidized for 32 minutes at 1000°C in order to achieve a cover silicon layer of 250nm thickness (≈2500Å).

### 4.2.7 Drive-in

In order to drive the carrier atoms deeper into the crystalline structure the wafer needs to be driven in after predeposition.



In this step the wafer is driven-in for 960 minutes at $1150°C$ in an inert ambient.

### 4.2.8 Oxide removal

We want an oxide free wafer with the n-well accessible for the further process steps.



We use hydrofluoric acid, because it doesn't etch silicon at all but is very aggressive towards $SiO_2$ while with non oxidized silicon there is no reaction.

## 4.3 Channel stop and field oxide

The channel-stop is material with the primary function to limit the spread of the channel area and to prevent the formation of parasitic channels. It is a highly doped p+ area. On top of the channel stop impurities is a thick layer of silicon oxide which has the primary function of isolating all the transistors from each other.

Its top view and cross section can be seen below.



It surrounds all the active areas and covers all the non-active area. The channel-stop region is always accompanied by the field oxide layer and shares the same layout mask with it.

### 4.3.1 Dioxide layer



### 4.3.2 Nitride layer
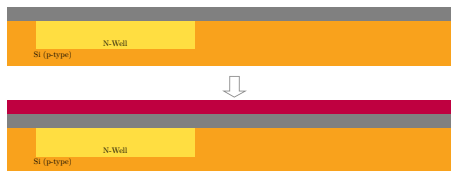


### 4.3.3 Patterning
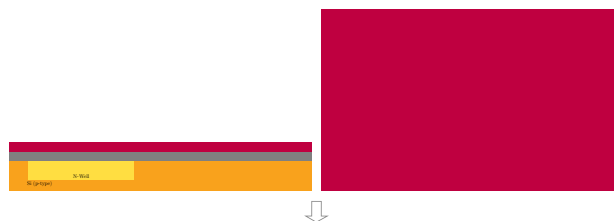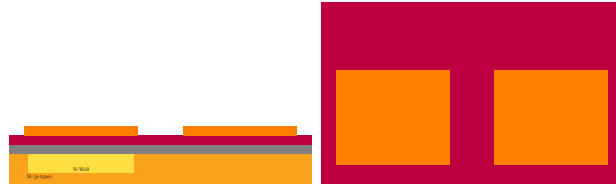
### 4.3.4 Etching
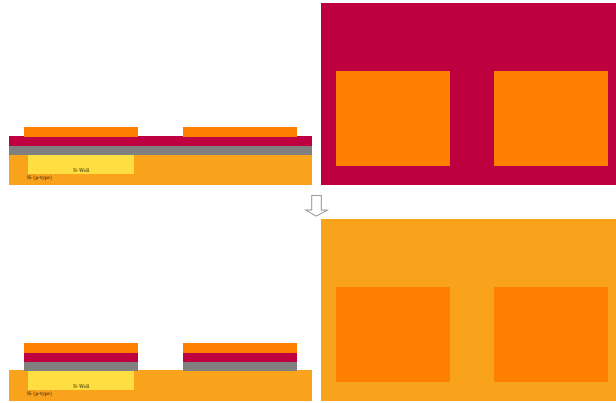


### 4.3.5 Cleaning
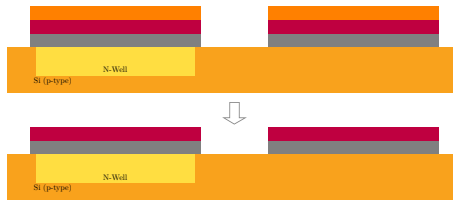


### 4.3.6 Injection
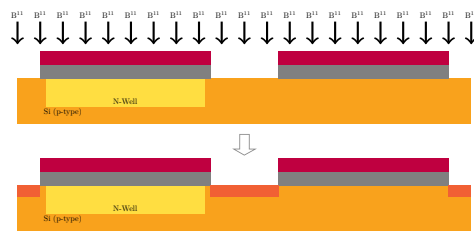


### 4.3.7 Thick oxide layer

The initial thickness of this field region is $\approx 7500$Å, but is later reduced to $\approx 6000$Åby various unmasked $SiO_2$ etches prior to the polysilicon deposition.

The thick field oxide is grown in a pyrogenic steam ambient (partial pressure of $H_2O \approx 0.80\ atmospheres$) for 190 minutes at 1000 °C.

### 4.3.8 Patterning
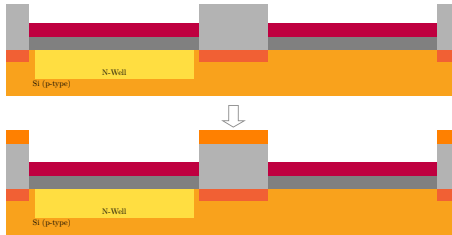


### 4.3.9 Etching
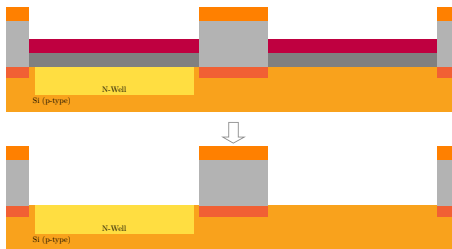
## 4.4 Active
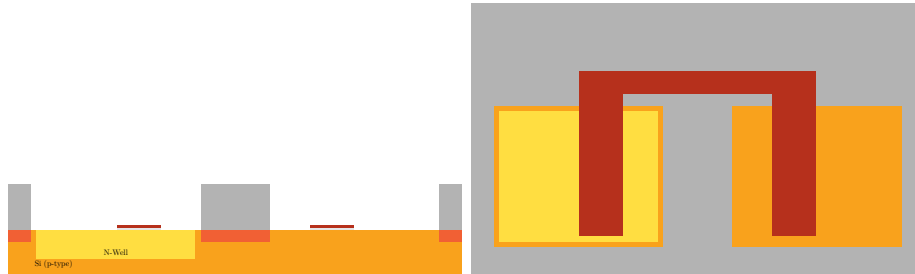


### 4.4.1 Gate oxide growth

The thickness of the gate oxide depends the capacity (slew rate) of the transistor. The thinner the layer is, the steeper the edges of the CMOS circuitry will be, however also the threshold voltage will be reduced the thinner the gate oxide gets.
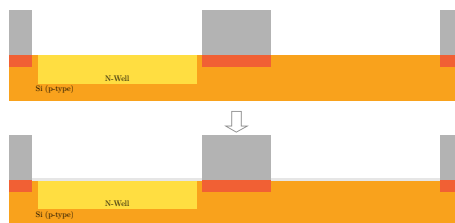


### 4.4.2 Polysilicon growth
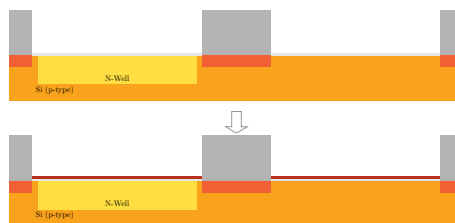


### 4.4.3 Patterning
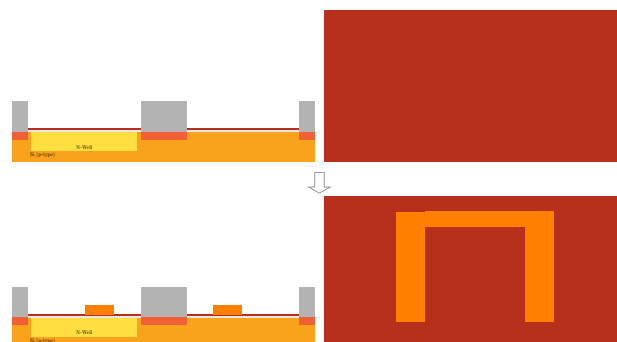
### 4.4.4   Etching
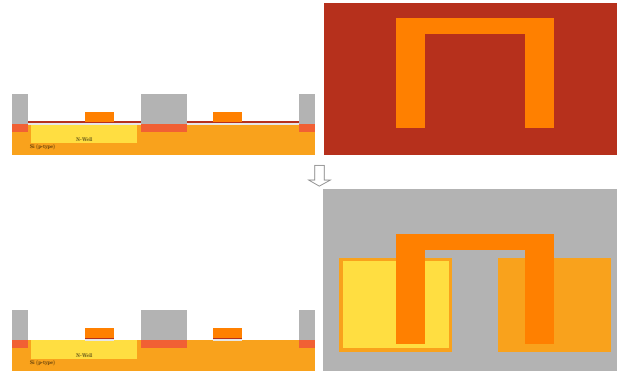


Because the exact shape of the gate contact is required for a reproducible property characterization of the transistor geometry, dry etching is being used for etching the poly-oxide layer stack.

### 4.4.5   Cleaning
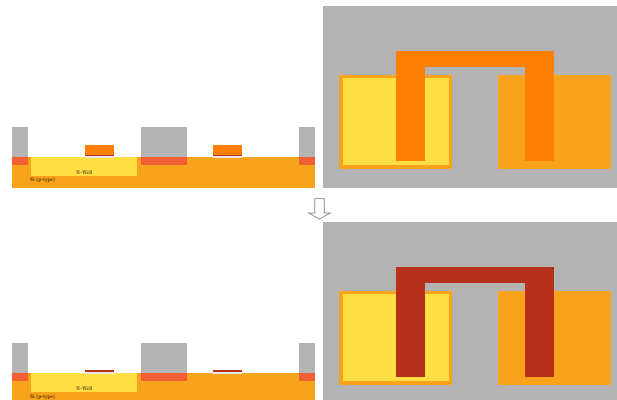
## 4.5   p+ Implant
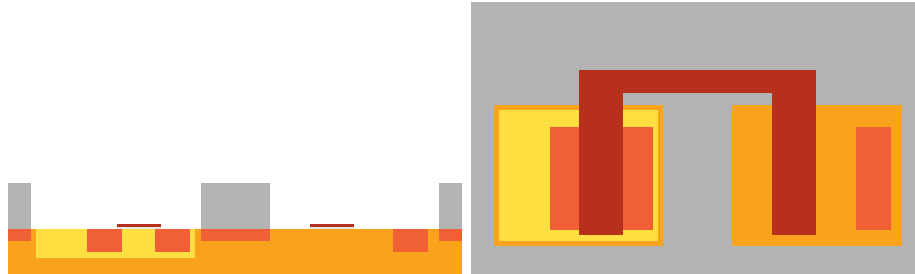


For the bulk of the NMOS transistors and for the source and drain of the PMOS transistors highly doped p+ areas are required. In this step we're going to build these.

### 4.5.1   Dioxide layer
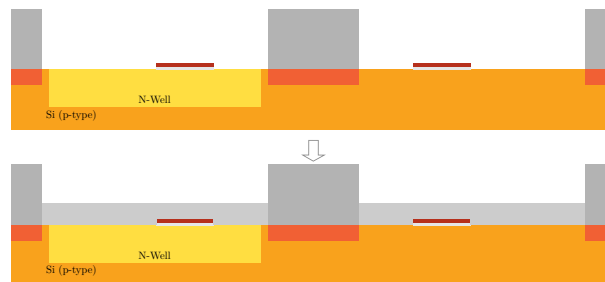


### 4.5.2   Pattering
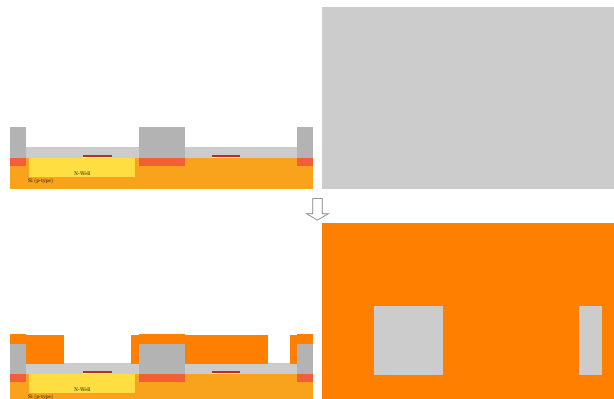


### 4.5.3   Etching
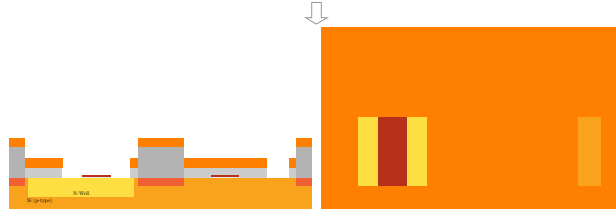
### 4.5.4 Cleaning
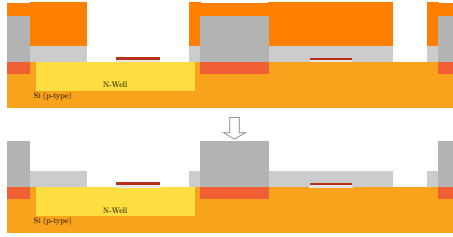


### 4.5.5 Predeposition
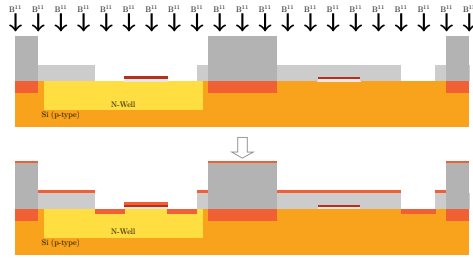


The p+ islands are implanted with a Boron ($B^{11}$) dose of $4 \times 10^{11} cm^{-2}$ at an energy of 35 KeV.
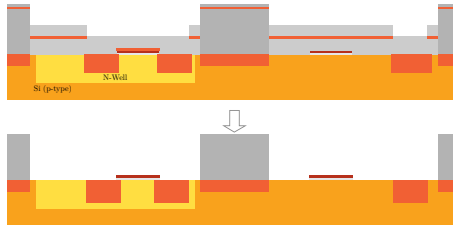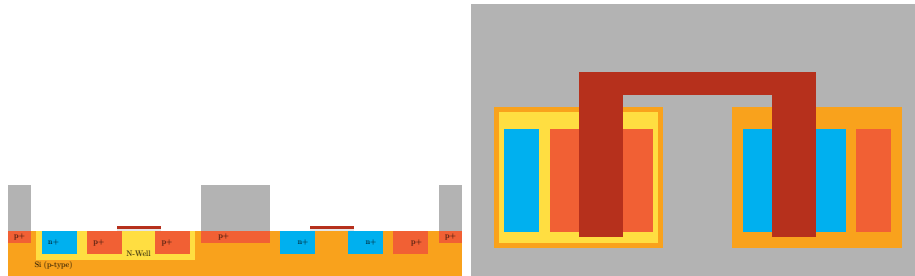
### 4.5.6 Sacrificial oxide
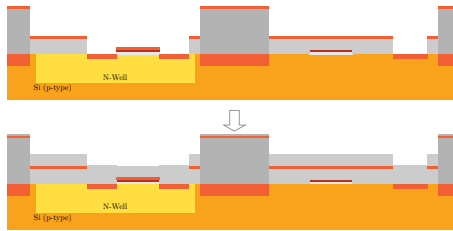


### 4.5.7 Drive-in
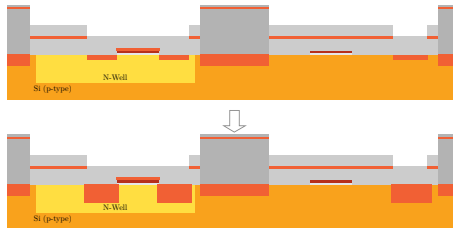
### 4.5.8  Oxide removal

In the final step before the next mask we need to remove the impure oxide layer for further processing of the wafer.



For this purpose we use a wet etching method selectively etching away the before grown, each approximately 250nm thick layers of oxide, together 500nm. Because we don't want to reduce the thick oxide of the field oxide area too much we limit the etching time here, so that we at worst etch maybe a 10Since we have grown it to be around 750nm we can afford to loose around 75nm or so.
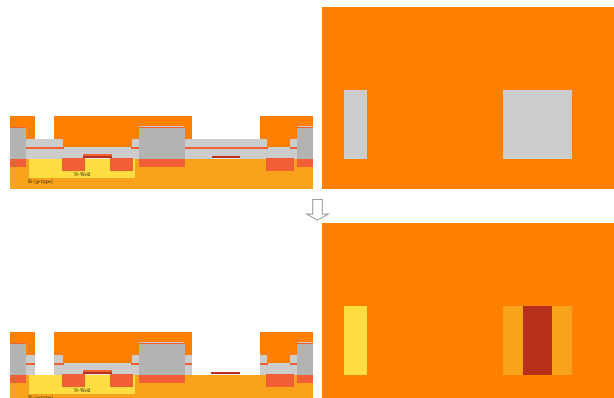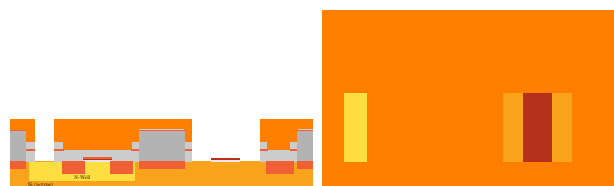
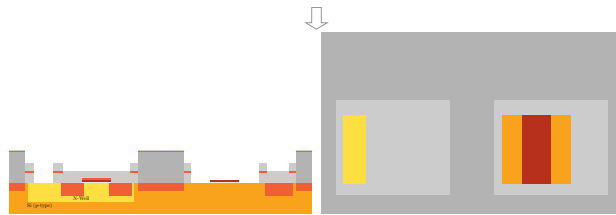## 4.6  n+ Implant
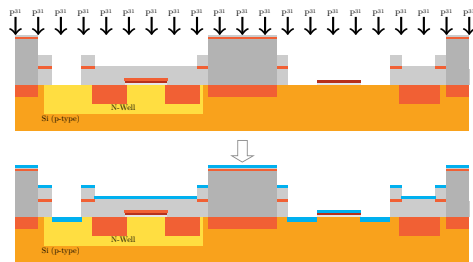

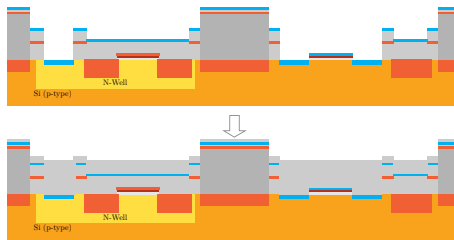
### 4.6.1  Pattering
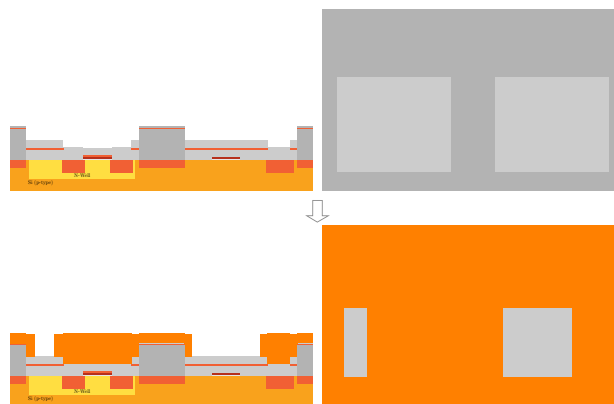


### 4.6.2  Etching



### 4.6.3  Cleaning

### 4.6.4 Predeposition



### 4.6.5 Sacrificial oxide



### 4.6.6 Drive-in