

A.6 The MOS capacitor

The MOS capacitor consists of a metal-oxide-semiconductor layer structure which forms a voltage dependent capacitor. This particular structure has been studied extensively because it is present in all MOS transistors. Its analysis provides details related to the threshold voltage of the transistor and the quality of the oxide-semiconductor interface. In addition it is frequently used to measure device parameters such as substrate doping concentration, oxide thickness and oxide charge. This section describes the approximate solution based on the full depletion model first, followed by a more detailed analysis.

A.6.1 Analytical solution based on the full depletion approximate

The MOS structure is shown in Fig.A6.1. It consists of a metal-oxide-semiconductor layer structure. The most common MOS structures contain silicon as the semiconductor, silicon dioxide as the oxide and aluminum as the metal. Since the metal also acts as the gate of a MOS transistor, the metal is also referred to as the gate metal. Heavily doped polysilicon layers are often used instead of metal as gate electrode. Contacts are made to the metal and to the semiconductor. A voltage is applied to the metal while the semiconductor is grounded.

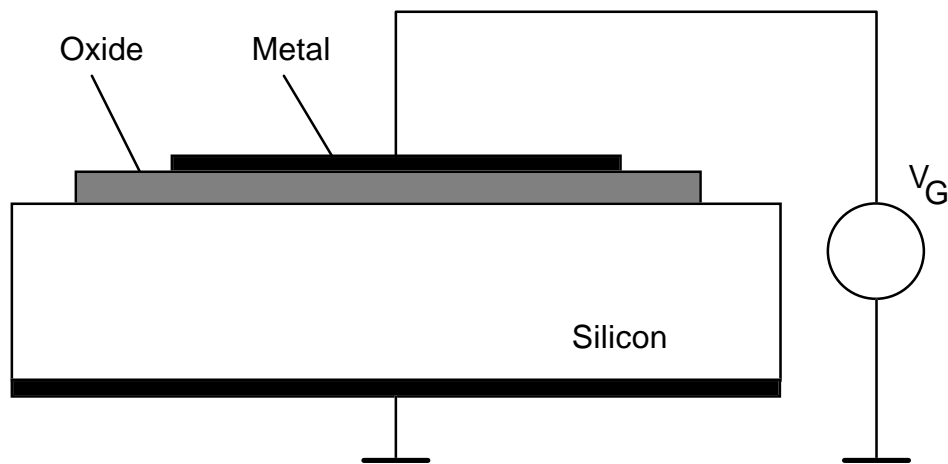


Fig.A6.1 MOS capacitor structure

The capacitance of a MOS structure is voltage dependent since the semiconductor region under the oxide can be either depleted of carriers, can accumulate carriers or an inversion layer can be formed. For instance for a p-type substrate one finds that for a large negative voltage, V_G , applied to the metal, holes are attracted to the interface, causing accumulation. A positive voltage on the other hand repels the holes which are present in the p-type material and thereby creates a depletion layer. A larger positive voltage at the gate causes sufficient bending of the energy bands in the semiconductor at the interface so that "inversion" occurs, i.e. a change from holes being the majority carrier type to electrons. An even larger positive voltage causes "strong inversion" where

the carrier density at the oxide-semiconductor interface exceeds that of the opposite carrier type in the substrate.

The capacitance of the structure changes accordingly, but before any discussion about the capacitance can be started, one has to distinguish between "low frequency" and "high frequency" capacitance measurements. The low frequency capacitance is measured when the semiconductor is in thermal equilibrium at any time during which the AC voltage is applied. This type of measurement is sometimes called the quasi-static capacitance measurement. For a high frequency capacitance measurement one uses a frequency which is so high that the minority carrier concentration can not follow the applied AC voltage and therefore maintains its value as determined by the DC bias voltage. Both capacitance measurements are performed as a function of the bias voltage which is slowly swept from accumulation, through depletion into inversion. The high frequency capacitance measurement is more common and therefore the focus of the approximate analysis.

The high frequency capacitance under accumulation conditions is the same as the oxide capacitance. Since carriers can easily be moved to and from the interface, the charges build up at both sides of the oxide as in a parallel plate capacitor. This changes as the applied bias voltage becomes positive, creating a depletion layer in the semiconductor. This depletion layer prevents carriers from moving towards the semiconductor-oxide interface. The variation of the charge therefore occurs at the edge of the depletion region so that the measured capacitance is the series connection of the oxide capacitance and the depletion layer capacitance.

Finally as a larger positive voltage is applied, inversion occurs at the interface. The presence of the inversion layer makes the depletion layer width almost independent of the applied voltage, yielding a constant (and minimum) capacitance which can be obtained from a series connection of the oxide capacitance and the minimum capacitance of the depletion layer. The approximate value of the capacitance can then be calculated from the following equations:

	accumulation	depletion	inversion
p-type substrate	$\phi_s < 0$	$0 < \phi_s < 2\phi_F$	$\phi_s > 2\phi_F$
n-type substrate	$\phi_s > 0$	$0 > \phi_s > -2\phi_F$	$\phi_s < -2\phi_F$
	$C = C_{ox}$	$C = \frac{1}{\frac{1}{C_{ox}} + \sqrt{\frac{2 \phi_s }{q(N_A - N_D)\epsilon_s}}}$	$C = \frac{1}{\frac{1}{C_{ox}} + \sqrt{\frac{4\phi_F}{q(N_A - N_D)\epsilon_s}}}$
		$V_G = V_{FB} + \phi_s + \frac{N_A - N_D}{C_{ox}} \sqrt{\frac{2q\epsilon_s \phi_s }{N_A - N_D}}$	

where ϕ_s is the potential across the semiconductor and V_{FB} is the flat band voltage at which no net charge exists in the semiconductor. ϕ_F is the reference potential as given by equation [A.6.4] Finally there is the condition of deep depletion which occurs when the bias voltage is swept rapidly while performing a high-frequency capacitance measurement. As the capacitor is bias from depletion into inversion the inversion layer needs to be formed. In the absence of light this inversion layer is formed by thermal generation of minority carriers. Since the inversion layer density is much larger than the thermal equilibrium value of the minority carrier density in the substrate, the time required to generate the inversion layer is orders of magnitude larger than the minority carrier life time, typically of the order of seconds in high quality silicon. The absence of the inversion layer causes the depletion layer to be larger than its thermal equilibrium value, thereby lowering the capacitance of the MOS structure. When sweeping from inversion to depletion one typically does not observe deep depletion since the inversion layer is already formed and the minority carriers recombine rapidly when changing the bias voltage. This leads to a typical hysteresis in the C-V measurement. While deep depletion affects the interpretation of a C-V measurement it confirms the long minority carrier life time which is desired to obtain low leakage currents in FETs. One might add that the deep depletion effect does not occur in FETs since the minority carriers can easily be supplied from the source or drain regions of the transistor.

A.6.1.a Extraction of device parameters from a high-frequency capacitance voltage measurement

The oxide capacitance is obtained from the MOS capacitance bias in accumulation. The capacitance should be independent of the bias voltage and equals the oxide capacitance. The substrate doping is obtained from the MOS capacitance in inversion. Again the capacitance should be independent of the applied voltage, but now equals the series connection of the oxide capacitance and the capacitance of the depletion region. The potential across the depletion region is $2|\phi_F|$, so that the doping concentration can be obtained from the depletion layer capacitance. Using the result from the exact solution described in the next section one can then calculate the flat-band capacitance using [A.6.12]. This enables to identify the flatband voltage of the MOS for which $\phi_F = 0$ Volt. The difference between the measured flatband voltage and the expected value based on the position of the Fermi level in the gate and the substrate is typically attributed to charge in the oxide or charge located at the oxide-semiconductor interface.

A.6.2 Exact solution to the MOS capacitor

We now derive the exact solution of the MOS capacitor. Whereas most of the derivation is applicable for both n and p-type substrates, the equations are written in a form which is more convenient for p-type substrates, but can easily be rewritten for n-type substrates.

The total charge density, ρ , in the semiconductor is given by:

$$\rho = q (p + N_d^+ - n - N_a^-) \quad [A.6.1]$$

Under thermal equilibrium, p and n can be expressed as a function of the potential ϕ and a reference potential ϕ_F .

$$p = n_i e^{(\phi_F - \phi(x))/V_t} \quad [A.6.2]$$

$$n = n_i^2/p = n_i e^{(\phi(x) - \phi_F)/V_t} \quad [A.6.3]$$

far away from the oxide-semiconductor interface, the charge density is zero and we define the potential, ϕ , to be zero there also, so that

$$N_d^+ - N_a^- = -2 n_i \sinh(\phi_F/V_t) \quad [A.6.4]$$

Poisson's equation then takes the following form:

$$\frac{d^2\phi}{dx^2} = \frac{2q n_i}{\epsilon_s} \left\{ \sinh\left[\frac{\phi - \phi_F}{V_t}\right] + \sinh\left[\frac{\phi_F}{V_t}\right] \right\} \quad [A.6.5]$$

multiplying both sides of the equation with $2 d\phi/dx$ and integrating while replacing $-d\phi/dx$ by the electric field \mathcal{E} , one obtains:

$$\mathcal{E}(\phi) = \text{sign}(\phi) \sqrt{\frac{4q n_i V_t}{\epsilon_s} \left\{ \cosh\left[\frac{\phi - \phi_F}{V_t}\right] + \frac{\phi}{V_t} \sinh\left[\frac{\phi_F}{V_t}\right] + K \right\}} \quad [A.6.6]$$

the constant K can be determined from the boundary condition at $x = \infty$ where $\phi = \mathcal{E} = 0$ yielding

$$K = -\cosh\left[\frac{\phi_F}{V_t}\right] \quad [A.6.7]$$

the electric field has the same sign as the potential as described with the sign function.

The relation between the field and the potential at the surface under thermal equilibrium is then:

$$\mathcal{E}_{\text{seq}} = \text{sign}(\phi_s) 2 \sqrt{\frac{q n_i V_t}{\epsilon_s} \left\{ \cosh\left[\frac{\phi_s - \phi_F}{V_t}\right] + \frac{\phi_s}{V_t} \sinh\left[\frac{\phi_F}{V_t}\right] - \cosh\left[\frac{\phi_F}{V_t}\right] \right\}} \quad [A.6.8]$$

The gate voltage can be expressed as a function of the flatband voltage, the voltage across the oxide and the potential across the semiconductor:

$$V_G = V_{\text{FB}} + \phi_s + V_{\text{OX}}, \text{ with } V_{\text{OX}} = x_{\text{OX}} \mathcal{E}_{\text{seq}}(\phi_s) \epsilon_s/\epsilon_{\text{OX}} \quad [A.6.9]$$

A.6.2a Low frequency capacitance

The low frequency capacitance of the MOS structure per unit area can then be calculated from:

$$C_{\text{LF}} = |dQ_s/dV_G| = \epsilon_s d\mathcal{E}_{\text{seq}}/d\phi_s d\phi_s/dV_G = \left[\frac{1}{C_{\text{OX}}} + \frac{1}{C_{\text{sLF}}} \right]^{-1} \quad [A.6.10]$$

where $C_{\text{OX}} = \epsilon_{\text{OX}}/x_{\text{OX}}$ and

$$\begin{aligned}
C_{s,LF} &= \epsilon_s d \mathcal{E}_{seq} / d\phi_s = \epsilon_s \sqrt{\frac{q n_i}{\epsilon_s V_t}} \frac{\sinh[\frac{\phi_s - \phi_F}{V_t}] + \sinh\frac{\phi_F}{V_t}}{\sqrt{\cosh[\frac{\phi_s - \phi_F}{V_t}] + \frac{\phi_s}{V_t} \sinh\frac{\phi_F}{V_t} - \cosh\frac{\phi_F}{V_t}}} \\
&= 2 \frac{q n_i \text{sign}(\phi_s)}{\mathcal{E}_{seq}} \left[\sinh[\frac{\phi_s - \phi_F}{V_t}] + \sinh\frac{\phi_F}{V_t} \right] \quad [A.6.11]
\end{aligned}$$

This result is often referred to as the low frequency capacitance of a MOS capacitor since we calculated the change in charge between two equilibrium situations. The result can be interpreted as a series connection of the oxide capacitance and the low frequency capacitance of the semiconductor $C_{s,LF}$. By starting from a series of values for ϕ_s , one can use the above equations to first calculate the electric field, the gate voltage and the capacitance. This enables to plot the low frequency capacitance as a function of the gate voltage as shown in Fig.A6.2.

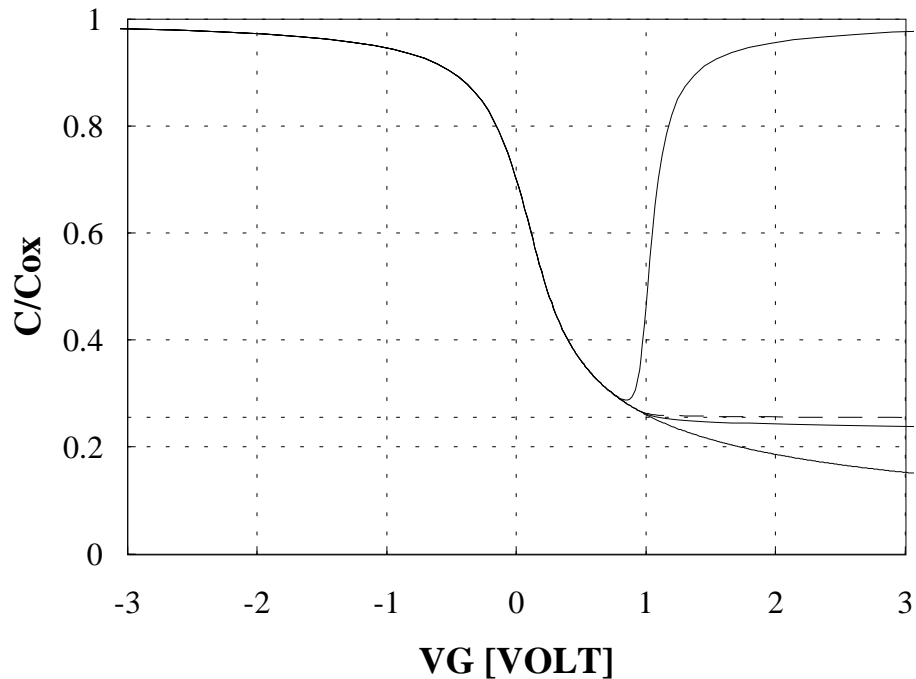


Fig.A6.2 Capacitance versus voltage for a MOS structure with $N_a = 10^{15} \text{ cm}^{-3}$ and $x_{OX} = 0.1 \mu\text{m}$. The curves from top to bottom are: The low frequency capacitance, the approximate high frequency capacitance, the exact high frequency capacitance and the capacitance under deep depletion conditions. The dotted line indicates C_{min} , the capacitance at the onset of strong inversion calculated using the full depletion approximation.

Under flat-band conditions, where $0 = |\phi_s| < V_t$, the capacitance reduces to

$$C_{FB} = [1/C_{OX} + L_D/\epsilon_s]^{-1} \quad [A.6.12]$$

where L_D is the extrinsic Debye length in the semiconductor with doping $|N_a - N_d|$:

$$L_D = \sqrt{\frac{\epsilon_s V_t}{q |N_a - N_d|}} \quad [A.6.13]$$

A.6.2b Deep depletion capacitance

If the gate voltage is changed faster than that electrons can be generated at the oxide-semiconductor interface to obtain the equilibrium density, no inversion layer is generated. In this case the gate voltage will cause the depletion layer in the semiconductor to exceed the maximum depletion layer width as defined at the onset of strong inversion. A typical measurement starts from an equilibrium situation where no inversion layer is present and the gate voltage is swept rapidly while creating a depletion layer in the semiconductor. The capacitance is measured as the change in charge flowing into the structure for a given voltage change. For a p-type substrate, this situation can be modeled by eliminating the charge term due to electrons in Poisson's equation:

$$\frac{d^2\phi}{dx^2} = \frac{q n_i}{\epsilon_s} \{2 \sinh \frac{\phi_F}{V_t} - \exp[\frac{\phi_F - \phi}{V_t}]\} \quad [A.6.14]$$

Using the same procedure as above the relation between surface field and surface potential can be found:

$$\mathcal{E}_{sdd} = \text{sign}(\phi_s) \sqrt{\frac{2 q n_i V_t}{\epsilon_s} \{2 \frac{\phi_s}{V_t} \sinh \frac{\phi_F}{V_t} + \exp \frac{\phi_F}{V_t} (\exp \frac{-\phi_s}{V_t} - 1)\}} \quad [A.6.15]$$

and the capacitance of the semiconductor becomes:

$$C_{sdd} = \frac{q n_i \text{sign}(\phi_s)}{\mathcal{E}_{sdd}} [2 \sinh \frac{\phi_F}{V_t} - \exp[\frac{\phi_F - \phi_s}{V_t}]] \quad [A.6.16]$$

and the corresponding gate voltage is:

$$V_G = V_{FB} + \phi_s + V_{OX}, \text{ with } V_{OX} = x_{OX} \mathcal{E}_{sdd}(\phi_s) \epsilon_s/\epsilon_{OX} \quad [A.6.17]$$

Using a similar procedure as for the low frequency capacitance we can also plot the capacitance under deep depletion conditions.

A.6.2c High frequency capacitance

The high frequency capacitance of an MOS capacitor is measured by applying a small ac voltage in addition to the DC gate voltage. The capacitance is defined as the ratio of the out-of-phase component of the ac current divided by the amplitude of the ac voltage times the radial frequency. An approximate expression can be obtained by ignoring the change in charge in the inversion layer yielding the expression for the capacitance under deep depletion conditions. However since the gate voltage is changed slowly while measuring the capacitance versus voltage, the gate voltage is

calculated from the surface potential including the charge in the inversion layer under thermal equilibrium. The capacitance is then given by:

$$C_{sHF,app} = \frac{q n_i \text{sign}(\phi_s)}{\mathcal{E}_{sdd}} \left[2 \sinh \frac{\phi_F}{V_t} - \exp\left[\frac{\phi_F - \phi_s}{V_t}\right] \right] \quad [A.6.18]$$

with the electric field, \mathcal{E}_{sdd} , obtained under deep depletion conditions [A.6.15]

This is the same expression as for the capacitance under deep depletion conditions, However the corresponding gate voltage is different, namely:

$$V_G = V_{FB} + \phi_s + V_{OX}, \text{ with } V_{OX} = x_{OX} \mathcal{E}_{seq}(\phi_s) \epsilon_s / \epsilon_{OX} \quad [A.6.19]$$

where the electric field, \mathcal{E}_{seq} , is the thermal equilibrium field

The corresponding capacitance is also included in Fig.A6.2 together with the expected minimum capacitance based on the full depletion approximation corrected for the thermal voltage:

$$\frac{1}{C_{min}} = \frac{1}{C_{OX}} + \sqrt{\frac{2(2\phi_F + V_t)}{qN_a\epsilon_s}} \quad [A.6.20]$$

It should be stressed that this is only an approximate solution. The redistribution of the inversion layer charge with applied gate voltage is ignored in the approximate solution even though it does affect the depletion layer width and with it the capacitance. This approximation therefore introduces an error which was found to be less than 6% at the onset of strong inversion and which increases almost linearly with increasing surface potential.

The exact expression for the high frequency capacitance¹ used in Fig.A6.2. is:

$$C_{sHF,exact} = \frac{q n_i \text{sign}(\phi_s)}{\mathcal{E}_{seq}} \{ e^{U_F} [1 - e^{-U_s}] + e^{-U_F} [e^{U_s} - 1] / (1 + \Delta) \} \quad [A.6.21]$$

Where Δ for a p-type substrate is:

$$\begin{aligned} \Delta &= 0 && \text{for } \phi_s < 0 \text{ and } \phi_F > 0 \\ &= \frac{(e^{U_s} - U_s - 1) / F(U_s | U_F)}{U_s} && \text{for } \phi_s > 0 \text{ and } \phi_F > 0 \\ &\int_0^{U_s} \frac{e^{U_F} (1 - e^{-\xi}) (e^{\xi} - \xi - 1)}{2F^3(\xi | U_F)} d\xi && \end{aligned} \quad [A.6.22]$$

The expression with $\Delta=0$ for all possible surface potentials equals the low frequency capacitance. The function F is related to the equilibrium electric field as:

¹A derivation of this expression can be found in "MOS (Metal Oxide Semiconductor) Physics and Technology", E. H. Nicollian and J. R. Brews, Wiley and Sons, 1982, p 157.

$$F(U|U_F) = \mathcal{E}_{eq} L_{Di} / (2\sqrt{2}V_t) \quad [A.6.23]$$

and the normalized parameters U , U_s and U_F are defined as:

$$U = \phi/V_t, U_s = \phi_s/V_t, U_F = \phi_F/V_t \quad [A.6.24]$$

Where the gate voltage is still given by:

$$V_G = V_{FB} + \phi_s + V_{OX}, \text{ with } V_{OX} = x_{OX} \mathcal{E}_{seq}(\phi_s) \epsilon_s/\epsilon_{OX} \quad [A.6.25]$$

and the electric field, \mathcal{E}_{seq} , is the thermal equilibrium field

As illustrated with Fig.A6.2, the high frequency capacitance at the onset of strong inversion ($\phi_s=2\phi_F$) and beyond is found to be almost constant. Assuming $\phi_F \gg V_t$ one finds

$$C_{sHF} = \sqrt{\frac{qN_a\epsilon_s}{4\phi_F}}, \text{ for } V_G - V_{FB} > 2\phi_F + 1/C_{OX} \sqrt{4qN_a\epsilon_s\phi_F} \quad [A.6.26]$$

this result could also be obtained by calculating the depletion region width in the semiconductor assuming the maximum potential at the surface to be $2\phi_F$ and using the full depletion approximation. The low frequency capacitance at $\phi_s=2\phi_F$, assuming $\phi_F \gg V_t$ is then:

$$C_{sLF} = \sqrt{\frac{2qN_a\epsilon_s}{\phi_F}}, \text{ for } V_G - V_{FB} = 2\phi_F + 1/C_{OX} \sqrt{4qN_a\epsilon_s\phi_F} \quad [A.6.27]$$

also yielding the following relation between both:

$$C_{sLF} = \sqrt{8} C_{sHF} \text{ for } \phi_s=2\phi_F \quad [A.6.28]$$