WEAK INVERSION IN ANALOG AND DIGITAL CIRCUITS

Eric A. Vittoz

CSEM, Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, CH 2007 Neuchâtel, Switzerland
eric.vittoz@csem.ch

- Behaviour and model of MOS transistors in weak inversion [1,2,3].
- Examples of analog circuits.
- Exploratory analysis of weak inversion logic [4,5].
MOS TRANSISTOR : DEFINITIONS

n-channel

symbols:

\[ W, L \] width, length of the channel
\[ C_{ox} \] gate capacitance per unit area
\[ U_T = kT/q \] ( = 26 mV at 300°C)
\[ V \] local non-equilibrium voltage in channel: channel voltage (quasi-Fermi potential of electrons)
  - at source end of channel: \( V = V_S \)
  - at drain end of channel: \( V = V_D \)
\[ Q_i \] local mobile inversion charge in channel (electrons)
\[ V_{T0} \] gate threshold voltage for \( V=0 \).
Weak inversion

- Given by: \( I_D = \beta \int_{V_S}^{V_D} - \frac{Q_i}{C_{ox}} \, dV \) with \( \beta = \mu C_{ox} \frac{W}{L} \) (\( \mu \)=mobility)

DRAIN CURRENT

- **Strong inversion, slope factor** \( n = 1.2 \text{ to } 1.6 \)
- **Weak inversion:** \( \frac{-Q_i}{C_{ox}} = 2nU_T \exp \left( \frac{V_P - V}{U_T} \right) \) exponential

Pinch-off voltage \( V_P \equiv \frac{V_G - V_{T0}}{n} \)

- Weak inversion already possible for \( V_S = 0 \) if \( V_G < V_{T0} \) ("subthreshold")
DRAIN CURRENT IN WEAK INVERSION

(vertical axis magnified)

\(-Q_i/C_{ox}\)

\(V_G-V_T0 > 0\)

\(V_P > 0\)

\(V_S\)

\(V_D\)

\(I_D/\beta\)

\(V_P < 0\)

\(V_S\)

\(V_D\)

\(V_G-V_T0 < 0\)
FORWARD AND REVERSE CURRENTS

\[ I_D(V_G, V_S, V_D) = F(V_G, V_S) - F(V_G, V_D) = I_F - I_R \]

- Drain current is the **superposition** of independent and symmetrical effects of source and drain voltages.
  - **basic property** of long-channel transistors, independent of current [6].
- Transistor saturated if \( I_R \ll I_F \), then \( I_D = I_F \).
**DRAIN CURRENT EXPRESSION IN WEAK INVERSION**

\[ V_P-V = \frac{V_P-V_{S,D}}{U_T} \]

- \( -Q_i/C_{ox} = 2nU_T e^{-\frac{V}{U_T}} \ll 2nU_T \) thus: \( I_{F,R} = I_S e^{-\frac{V}{U_T}} \)

- Definition: **specific current** of the transistor: \( I_S = 2n\beta U_T^2 \) (10 to 300 nA for \( W=L \))

- Introducing \( V_P \equiv (V_G-V_{T0})/n \) and \( I_D = I_F - I_R \), this yields:

\[
I_D = I_S e^{\frac{V_G-V_{T0}}{nU_T}} \left( e^{-\frac{V_S}{U_T}} - e^{-\frac{V_D}{U_T}} \right)
\]

for \( I_F \) and \( I_R \ll I_S \)
FORWARD CHARACTERISTICS IN WEAK INVERSION

\[ I_D = I_{D0} e^{\frac{V_G}{nU_T}} \left( e^{\frac{V_S}{U_T}} - e^{\frac{V_D}{U_T}} \right) \]

where \( I_{D0} = I_S e^{-\frac{V_{T0}}{nU_T}} \)

- output \( V_G, V_S = \text{const.} \)
- transfer from gate \( V_S, V_D = \text{const.} \)
- transfer from source \( V_G, V_D = \text{const.} \)

\( I_D \sim 1 - e^{-\frac{V_D - V_S}{U_T}} \)

minimum \( V_{DSsat} \)

exponential, slope \( 1/n \)

exponential, slope 1
CONTINUOUS MODELS WEAK-STRONG INVERSION

a. From charge analysis [7,8]:

\[
\frac{V_{P-V_{S,D}}}{U_T} = \sqrt{1 + 4 \frac{I_{F,R}}{I_S}} - 1 + \ln \frac{1 + 4 \frac{I_{F,R}}{I_S}}{2}
\]

\text{cannot} be inverted to express } I_{F,R}(V_P, V_{S,D})

b. Interpolation formula:

\[
\frac{I_{F,R}}{I_S} = \ln^2 (1 + e^{\frac{V_{P-V_{S,D}}}{2U_T}}) \tag{9}
\]

Both converge asymptotically towards:

\[
\frac{I_{F,R}}{I_S} = e^{\frac{V_{P-V_{S,D}}}{U_T}} \quad \text{for } V_{P-V_{S,D}} \ll U_T \text{ (weak inversion)}
\]

\[
\frac{I_{F,R}}{I_S} = \left(\frac{V_{P-V_{S,D}}}{2U_T}\right)^2 \quad \text{for } V_{P-V_{S,D}} \gg U_T \text{ (strong inversion)}
\]

• Only 3 parameters: \( V_{T0}, n \) (inside \( V_P \)) and \( I_S \) (or \( \beta \)) to model the current from weak to strong inversion.
**Definition**: Inversion coefficient: \( IC = \) the larger of \( IF/IS \) and \( IR/IS \)

- **Weak inversion**: \( IC \ll 1 \)
- **Moderate inversion**: \( IC \approx 1 \)
- **Strong inversion**: \( IC = \left( \frac{V_{DSsat}}{2UT} \right)^2 \gg 1 \)

CONTINUOUS MODELS WEAK-STRONG INVERSION

\[
\frac{IF, R}{IS} \quad \text{current} \quad \frac{VP - VS,D}{UT} \quad \text{voltage}
\]

with:

\[
VP = (VG - V_{T0})/n \\
I_D = IF - IR
\]
TRANCONDUCTANCE FROM WEAK TO STRONG INVERSION

- Transconductance $g_m$ from gate in **saturation**

\[
\frac{g_m}{I_D} = \frac{I_D}{(nU_T)}
\]

**Weak inversion asymptote:**

\[
\frac{g_m}{I_D} = \frac{I_D}{(nU_T)}
\]

**Strong inversion asymptote:**

\[
\frac{g_m}{I_D} = \sqrt{2\beta} \frac{I_D}{n}
\]

- $g_m/I_D$ decreases with increasing inversion coefficient $IC$.
- $g_m/I_D$ is maximum in **weak** inversion.
SUMMARY OF FEATURES OF WEAK INVERSION

- Large-signal DC model:

\[
I_D = l_s e^{nU_T} \left( e^{-\frac{V_S}{U_T}} - e^{-\frac{V_D}{U_T}} \right)
\]

+ exponential
+ min. \( V_DSSat \)
+ min. gate voltage
+ min. gate capacitance
+ max. \( g_m/I_D \)
+ \( g_m(I_D) \) linear
+ \( g_m \) independent of \( \beta \)

- Low speed: \( f_T = \frac{\mu U_T}{2\pi L^2} \)

+ max. intrinsic voltage gain
+ min. input noise density for given \( I_D \)
+ max. bandwidth for given \( kT/C \) and \( I_D \)
+ min. input offset voltage
- max. output noise current for given \( I_D \)
- max. current mismatch: dominated by \( VT \)-mismatch: \( \frac{\Delta I_D}{I_D} = \frac{\Delta V_{T0}}{nU_T} \)

+ max. \( l_on/l_{off} \) for given voltage swing
- intermodulation in RF front ends

⇒ translinear circuits and log domain filters

CSEM, E. Vittoz, 2003
EVOLUTION OF $IC$ WITH SCALED-DOWN PROCESSES

- Scaling-down of process:
  - dimension scaling by factor $k$
  - all voltages decreased by $k$, except $U_T$:
    - analog circuits: $V_{DSsat}$ must be decreased by $k$, thus
      \[
      IC = \left( \frac{V_{DSsat}}{2U_T} \right)^2 \text{ decreased by } k^2
      \]
    - digital circuits: $V_B$ decreased by $k$, thus
      \[
      I_{C_{on}} = \left( \frac{V_B - V_{T0}}{2nU_T} \right)^2 \text{ decreased by } k^2
      \]

- Weak inversion approached for constant temperature $T$.
- Transition frequency: $f_T = \frac{\mu V_{DSsat}}{2\pi L^2}$ increased by $k$
  - weak inversion with $L=100$nm : $f_T > 4$ GHz
LOW-VOLTAGE CASCODE IN WEAK INVERSION

\[ V_{DSSat} = 4 \text{ to } 6 U_T \text{ per transistor} \]

\[ V_{DS1} = U_T \ln \left[ P \left( 1 + 2M \right) \right] \]

for \( P = M = 8 : V_{DS1} = 5 U_T, \)

thus \( V_{D2} = 10 U_T \text{ sufficient} \) to saturate \( T_1 \) and \( T_2 \)
For $T_1$ and $T_2$ in weak inversion:

$$V_R = RL_2 = UT \ln K$$

Self-starting if leakage of $T_2$ larger than that of $T_1$. 

[1] Extraction of $U_T$ and current reference
CURRENT GENERATION WITHOUT RESISTOR

- Resistor replaced by transistor $T_8$ in conduction [10]:

  - $T_6 = T_3 = T_4 = T_7$ and $T_5 = T_1$
  - $T_8$ and $T_9$ in strong inversion with $\beta_8 = A\beta_9$ ($A \gg 1$ to have $T_8$ in conduction)
  - $T_2$ and $T_1$ in weak inversion with $\beta_2 = K\beta_1$

  yields: $I = 2n\beta_8 U_T^2 A \ln^2 K = l_{S8} A \ln^2 K$

- Reference current $I$ proportional to specific current $l_{S8}$
- Useful to bias transistors at inversion coef. $IC$ independently of process.
- If mobility $\sim T^{-2}$, then compensation by $U_T^2 : I \sim I_S$ independent of $T$
MOS TRANSISTOR OPERATED AS A PSEUDO-RESISTOR

[11,12,13,6]

Consequence of basic property \( I_D = F(V_S) - F(V_D) \): 

- Networks of transistors with same gate voltage are 
  - linear with respect to currents 
  - thus equiv. for currents to a resistive prototype, with \( G_i = 1/R_i \sim I_{Si} \) 
  - ground in res. prototype correspond to saturated transistors. 
  - example of application: current-mode linear attenuator (e.g. \( R-2R \)). 

- In weak inversion: 
  - linearity of currents even for different gate voltages 

\[
G_i = 1/R_i \sim I_{Si} \exp \frac{V_{Gi}}{nU_T}
\]
simple example of pseudo-R network in weak inversion:

**CALCULATION OF HARMONIC MEAN**  \([14,13]\)

- Series combination of \(G_i\): \(G = \frac{1}{\sum 1/G_i}\)
- Same voltage across \(G\) and \(G_i\), thus \(I = \frac{1}{\sum 1/I_i} = \frac{I_{hm}}{N}\)
- Can be used as a fuzzy AND gate.
With bipolar transistors:

\[
\sum V_{BEi} = \sum V_{BEi}
\]

\[V_{BEi} = U_T \ln \frac{I_i}{I_{Si}}\]

\[
\prod I_i \prod I_{Si} = \prod I_i \prod I_{Si} = \lambda
\]

With MOS transistors in weak inversion: \[16,17\]

\[
\sum (V_{Gi} - V_{Si}) = \sum (V_{Gi} - V_{Si})
\]

\[
\frac{V_{Gi}}{n} - V_{Si} = U_T \ln \frac{I_i}{I_{D0i}}
\]

- If + and - are **alternated** then: pairs of equal \(V_{Gi}\) both sides of equation:

\[V_{Gi} \Rightarrow V_{Gi}/n\] for each pair,

and then

- Otherwise: **separate wells** connected to sources to impose \(V_{Si} = 0\)

- Precision degraded by \(V_{T0}\) mismatch
BASIC CONSIDERATIONS FOR WEAK INVERSION LOGIC

- Dynamic power consumption: \( P_{\text{dyn}} = f C \Delta V V_B \)

- Weak inversion model can be rewritten as

\[
I_D = I_0 e^{\frac{V_{GS}}{nU_T} (1 - e^{-\frac{V_{DS}}{nU_T}})}
\]

- exponential in \( V_{GS} \), with maximum \( g_m/I_D \), thus:
  - minimum swing \( \Delta V \) for given \( I_{on}/I_{off} \), hence
  - minimum \( P_{\text{dyn}} \) for given \( I_{off} \)

- with: \( I_0 = I_S e^{-\frac{V_{T0}+(n-1)V_S}{nU_T}} \) adjustable by \( V_S \).

- Assumptions on process:
  1. Threshold \( V_{T0} \) close to 0 (\( V_S \) cannot be too negative).
  2. Triple well (true twin well): separate local p and n substrates
     - adjustment of \( I_0 \) by \( V_S \) for n- and p-channel.
STABLE STATES OF CMOS FLIP-FLOP

- Simplifying assumptions: \( n_n = n_p = n \), \( I_{0n} = I_{0p} = I_0 \)
- Normalized voltages \( v_k = \frac{V_k}{U_T} \)

\[
\begin{align*}
    n & = 1.6 \\
    V_B & > 1.91 U_T \\
    95\% \text{ swing for } V_B & = 4 U_T
\end{align*}
\]
• Since $V_L = V_B - V_H > 0$, static current $I_{stat}$ at each state is larger than $I_0$

![Graph showing $I_{stat}/I_0$ vs normalized supply voltage $v_B$]

- $I_{stat} < 4\%$ above $I_0$ for $v_B \geq 4$: the difference can be neglected thus:

- Static power: $P_{stat} \approx I_0 V_B$
STANDARD TRANSITIONS IN HOMOGENEOUS SYSTEM

- Chain of inverters

\[ V_H \quad V_{o2} \quad V_{o4} \quad V_{o6} \quad V_{o8} \]

\[ V_L \]

\[ v_{on} = v_{in+1} \]

\[ v_{o1} \quad v_{o3} \quad v_{o5} \quad v_{o7} \]

\[ v_{o2} \quad v_{o4} \quad v_{o6} \quad v_{o8} \]

\[ n = 1.6 \]

\[ v_B = 4 \]

- Characteristic time: \[ T_0 = C U_T / I_0 \]

- Transitions become standard after a few stages

- Normalized delay time \[ T_d / T_0 \] only depends on \( V_B \) and \( n \).
**DELAY TIME FOR STANDARD TRANSITIONS**

- Approximation:
  
  \[ T_d \approx \frac{CV_B}{I_{on}} \approx \frac{CV_B}{I_0 e^{V_B/nU_T}} \]

  or
  
  \[ I_0 \approx \frac{CV_B}{T_d} e^{-V_B/nU_T} \]

  (for calcul. of \( P_{stat} \))

- \( T_d \) decreases approximately **exponentially** with increasing \( V_B \).

\( n = 1.6 \)
- Short-circuit charge $Q_{sc} < 1.4\%$ capacitor charge $Q_C$: negligible, thus:
  - dynamic power $P_{dyn} \cong fQCV_B \cong fCV_B^2$
  - with static power $P_{stat} = I_{stat}V_B \cong I_0V_B$
POWER-DELAY PRODUCT

- Definition: **duty factor** $\alpha = 2f \frac{T_d}{\alpha} \leq 1$
- proportion of time during which the gate is in transition.

- Then, total power $P = P_{dyn} + P_{stat} \Rightarrow P = \frac{CU_f^2}{T_d} v_B^2 (\frac{\alpha}{2} + e^{-v_B/n})$

\[ PT_d = \frac{CU_f^2}{v_B^2} \]
\[ normalized \ power-delay \ product \]
\[ normalized \ supply \ voltage \ v_B \]

- $P_{dyn}$ dominates for large $\alpha \Rightarrow$ min. $V_B$ for min. $PT_d$
- $P_{stat}$ dominates for small $\alpha \Rightarrow$ increase $V_B$ to increase $I_{on}/I_{off}$

CSEM, E. Vittoz, 2003
• By re-using $\alpha = 2f T_d$:

$$\frac{P/f}{C(nU_T)^2} = \frac{P_{dyn}}{f} = C(nU_T)^2 \left( \frac{v_B}{n} \right)^2 \left( 1 + \frac{\alpha}{2} e^{-v_B/n} \right)$$

- $V_{B_{opt}}$ and $P_{min}$ increase for decreasing $\alpha$
- At $P_{min}$: $P_{dyn} \gg P_{stat}$
  - Increasing $I_0$ does not allow to reduce $V_B$ significantly for $T_d$ const.
- For $\alpha > 5\%$, power reduction by $>20$ compared to $P_{dyn}$ at 1V.
Weak inversion

MAXIMUM SPEED

- Since $T_d \equiv \frac{CV_B}{I_{on}}$ and $I_{onmax} \equiv I_{Con} I_S$ (inv. coeff* spec. current), thus:

\[
T_{dmin} \equiv \frac{V_B}{I_{Con}} \frac{C}{I_S}
\]

- Limit of weak inversion: $I_{Con} \equiv 1$, thus

\[
T_{dmin(weak)} \equiv V_B \frac{C}{I_S}
\]

- Higher speed can only be obtained by entering moderate or strong inv.
EFFECT OF ENTERING MODERATE AND STRONG INVERSION

(using continuous model of $I_D$)

- More **voltage swing** needed
to obtain $I_{on}/I_{off}$

- from continuous current model:

- **Degeneration** of logic states:
  - reduction of logic swing
  - large increase of static current $I_{stat}$
  - loss of bistability
  - more supply voltage needed.
NUMERICAL RESULTS

- Simple inverter replaced by 3-input NAND-gate:
  - approx. equivalent to inverter with
    \[ L = 3 \text{-time that of n-ch transistor} \]
    \[ C = 6 \text{-time that of min. inverter} \]
    (includes \( C_{\text{interconnect}} = \frac{C}{2} \)).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>process A</th>
<th>process B</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>min. channel length, ( L_{\text{min}} )</td>
<td>500</td>
<td>180</td>
<td>nm</td>
</tr>
<tr>
<td>equiv. spec. current, ( I_S )</td>
<td>200</td>
<td>400</td>
<td>nA</td>
</tr>
<tr>
<td>equiv. load capac., ( C )</td>
<td>20</td>
<td>4</td>
<td>fF</td>
</tr>
<tr>
<td>specific energy, ( C(nU_T)^2 )</td>
<td>28</td>
<td>4.2</td>
<td>aJ</td>
</tr>
<tr>
<td>( P/f ) for ( \alpha = 1 ) ( V_B = 4U_T )</td>
<td>228</td>
<td>44</td>
<td>aJ</td>
</tr>
<tr>
<td>( (P/f)<em>{\text{min}} ) for ( \alpha = 0.01 ) ( V</em>{B_{\text{opt}}} = 6nU_T )</td>
<td>1.46</td>
<td>0.22</td>
<td>fJ</td>
</tr>
<tr>
<td>( P_{\text{dyn}}/f ) at ( V_B = 1V )</td>
<td>20</td>
<td>4</td>
<td>fJ</td>
</tr>
<tr>
<td>( f_{\text{max1}} ) for ( \alpha = 1 ) ( V_B = 4U_T )</td>
<td>50</td>
<td>500</td>
<td>MHz</td>
</tr>
<tr>
<td>( f_{\text{max2}} ) for ( \alpha = 0.01 ) ( V_B = V_{B_{\text{opt}}} )</td>
<td>0.22</td>
<td>2.56</td>
<td>MHz</td>
</tr>
<tr>
<td>( P_{\text{min}} ) at ( f_{\text{max2}} )</td>
<td>32.5</td>
<td>56.3</td>
<td>nW</td>
</tr>
</tbody>
</table>
PRACTICAL CONSIDERATIONS AND LIMITATIONS

- Low-voltage power source
  - should be proportional to $U_T$ (PTAT)
  - need for power-efficient adapter from higher supply voltage.
- Asymmetry
  - p/n asymmetry may result in speed reduction.
- Mismatch
  - dominated by threshold mismatch $\delta V_T$
  - may result in speed reduction proportional to $\delta V_T/V_B$.
- Short channel effects: should not drastically degrade the results.
- Gate leakage current: should be alleviated by very low $V_B$.
- Adjustment of $I_0$ or $T_d$ to required value
  - control by $V_S$ with charge pump in loop [18]; $n>1$ needed (no SOI!)
  - corresponds to threshold adjustment unavoidable at very low $V_B$.
- System architectures and applications.
SYSTEM ARCHITECTURE AND APPLICATIONS

• Duty factor $\alpha$ must be maximized to reach minimum $P/f$,
  (where $f$ is the average transition frequency), thus
  • avoid idling gates (contrary to traditional CMOS culture)
  • new architectures needed:
    - maximally active gates of minimum speed (max. delay time $T_d$)
    - particular problem with RAMs (short $T_d$ but sparse activity)
    - how? new constraints should result in novel solutions.
  • partition the system in blocks of comparable $\alpha$ and $T_d$
    - optimum $V_B$ and $I_0$ for each block (separate $I_0$ control).

• Maximum frequency much lower than for strong inversion:
  • best applicable when no high local speed is required
  • $m$-parallelize: $mT_d$ but same power if same $\alpha$ ($m$ units with $P/m$)
    - digital image processing?
CONCLUSION

- Weak inversion permits very low supply voltage $V_B$
  - approached with scaled-down $V_B$: $IC \sim V_B^2$
  - limit for scaled-down $V_B$.

- Analog: $V_B > 10U_T = 250$ mV
  - provides maximum $g_m/I_D$
  - bipolar-like behaviour can be exploited in new schemes.

- Digital: $V_B > 4U_T = 100$ mV
  - transistor not a switch but a current modulator ($I_{on}/I_{off}$)
  - new architectural approaches for max. duty factor $\alpha$.
  - ultimum (asymptotic) limit for low power*delay.

- Low speed, but keeps increasing with $1/L^2$ in scaled down processes.
REFERENCES


