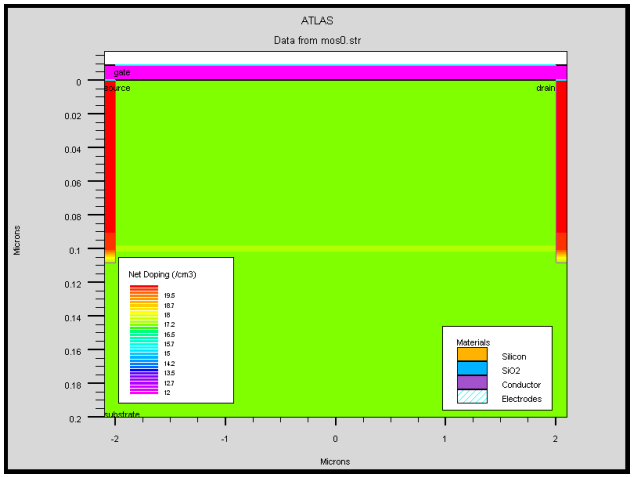
MOSFET Short Channel Effects

Vishal Saxena
ECE, Boise State University

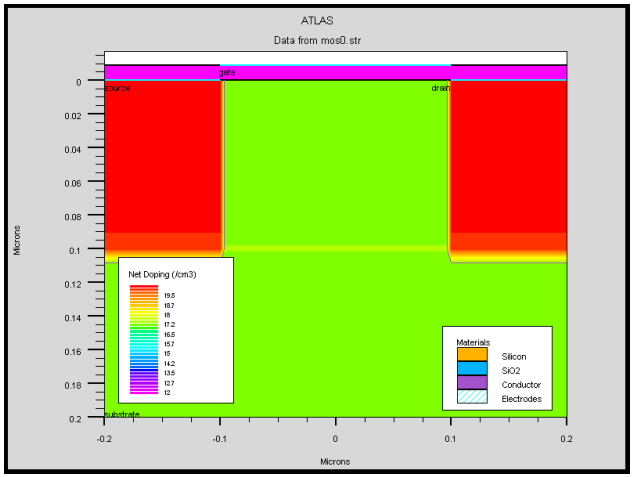
Oct 10, 2010

Short Channel MOSFETs

- MOSFETs with channel length (L) of the same order of magnitude as the drain/source region depths.
- Submicron MOSFETs (with $L < 1\,\mu m$) exhibit worsening short channel effects (SCE) as L is scaled down.



A long-channel NMOS (L=4μm)



A short-channel NMOS (L=0.2μm)

SC MOSFET Characteristics

- Due to the high electric fields in the channel region, the carriers in SC MOSFETs are velocity saturated (v_{sat})
- MOSFET current (I_D) in saturation exhibites somewhat linear relation to V_{GS}

$$I_D = W \cdot v_{sat} \cdot C'_{ox}(V_{GS} - V_{THN} - V_{DS,sat})$$

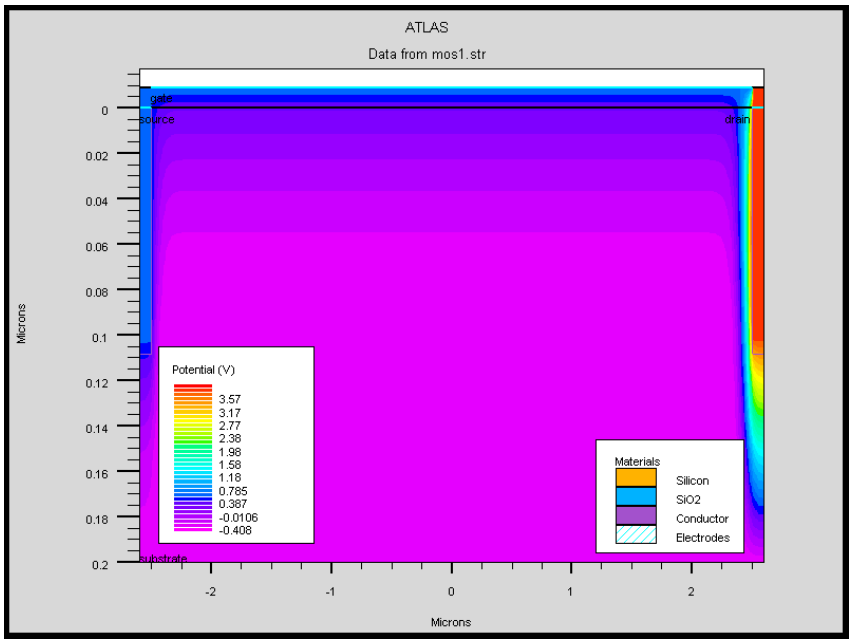
SC MOSFET Characteristics

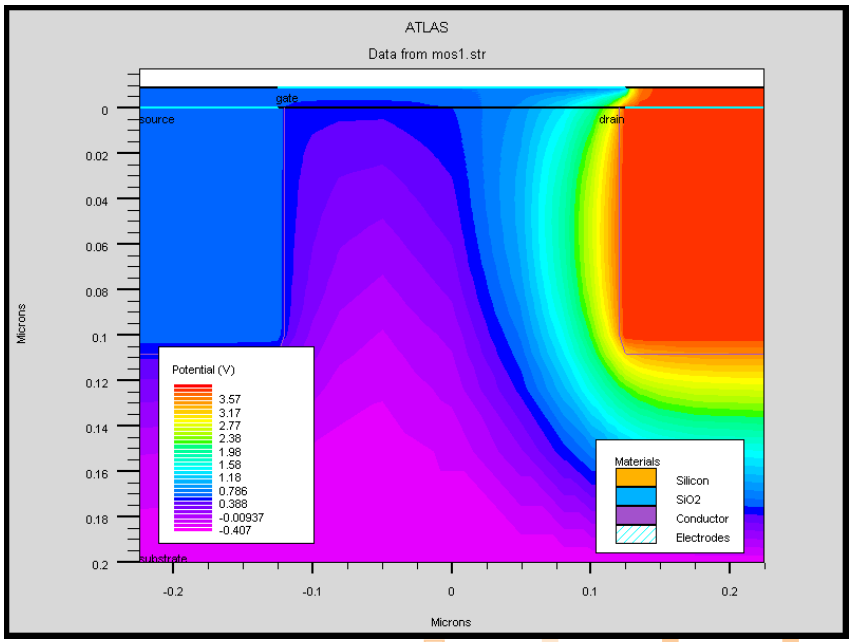
- Due to the high electric fields in the channel region, the carriers in SC MOSFETs are velocity saturated (v_{sat})
- MOSFET current (I_D) in saturation exhibites somewhat linear relation to V_{GS}

$$I_D = W \cdot v_{sat} \cdot C'_{ox}(V_{GS} - V_{THN} - V_{DS,sat})$$

Potential Distribution in Short Channel MOSFETs

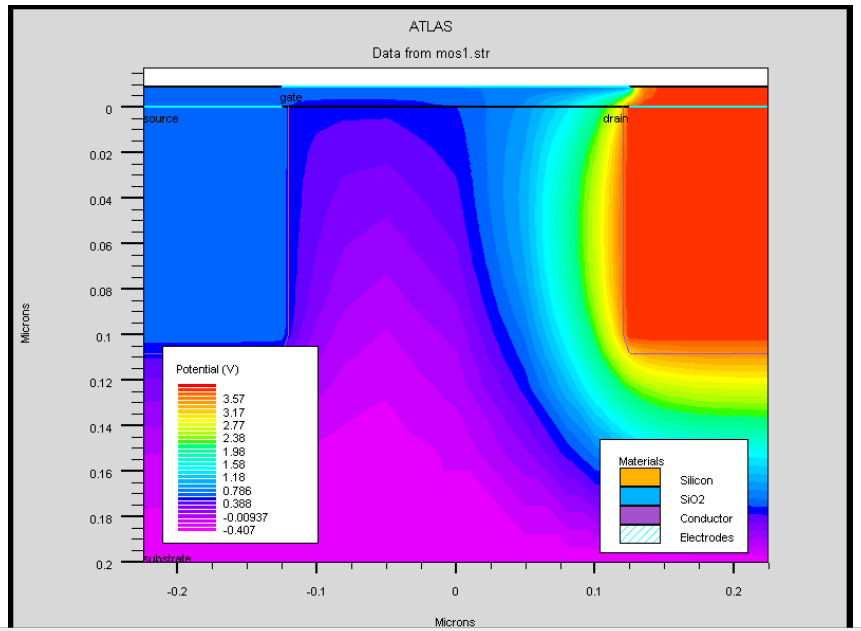
- In the short-channel device, the field and potential contours have a 2D spread while the long-channel device has relatively 1D potential distribution.
- Here, $V_{GS} = 0.1V$ and $V_{DS} = 3.3V$.





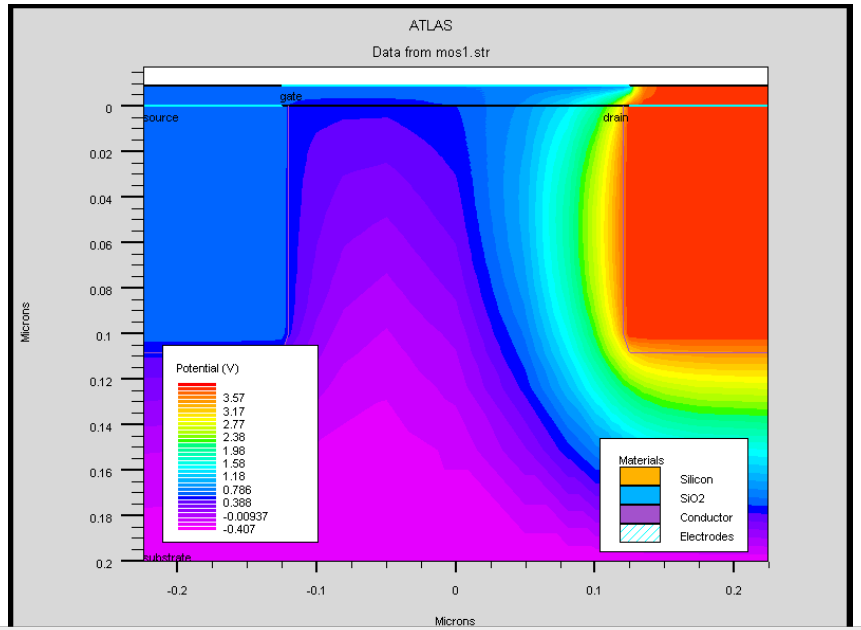
Drain Induced Barrier Lowering (DIBL)

- As seen in the constant potential contour plots, there is more surface potential (ψ_s) in the short channel MOSFET
  - more band bending near drain due to the drain field encroaching into the channel region
  - reduces V_{THN} when drain is biased at higher potential
- This is called drain induced barrier lowering (DIBL)



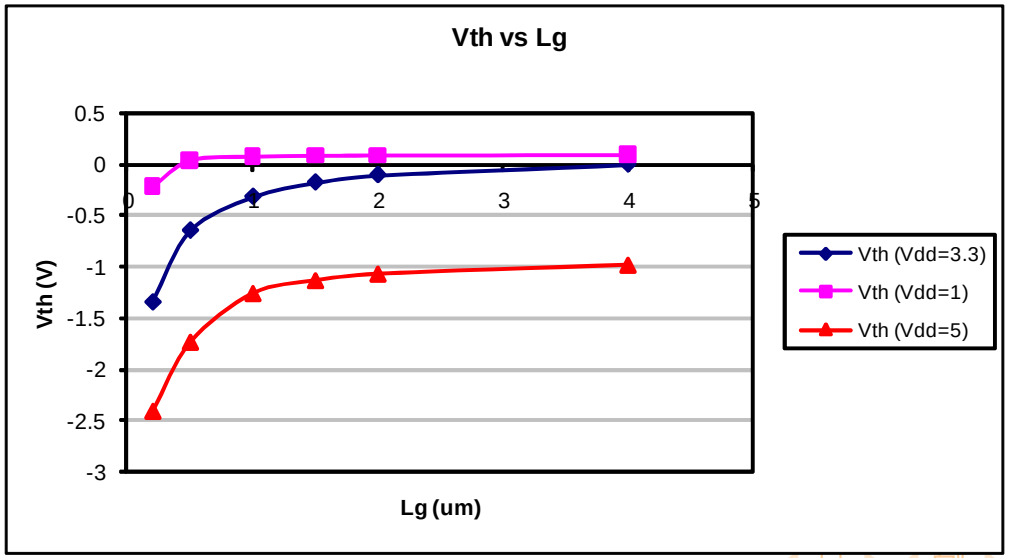
Drain Induced Barrier Lowering (DIBL)

- As seen in the constant potential contour plots, there is more surface potential (ψ_s) in the short channel MOSFET
  - more band bending near drain due to the drain field encroaching into the channel region
  - reduces V_{THN} when drain is biased at higher potential
- This is called drain induced barrier lowering (DIBL)



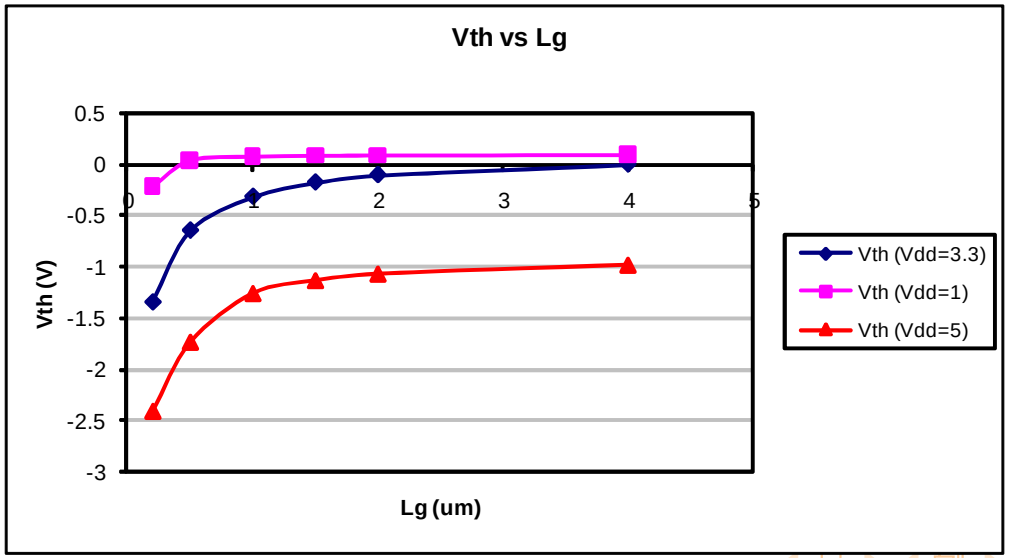
DIBL contd.

- The stronger electric field corresponding to a higher V_{DS}, penetrates deeper into the channel
  - causes more pronounced short channel effects.
- The V_{THN} roll-off slope is higher for a varying L at a higher V_{DS}
  - due to higher 2D electric field in the channel region.



DIBL contd.

- The stronger electric field corresponding to a higher V_{DS}, penetrates deeper into the channel
  - causes more pronounced short channel effects.
- The V_{THN} roll-off slope is higher for a varying L at a higher V_{DS}
  - due to higher 2D electric field in the channel region.



Hot Carrier Generation

- Carriers drifting near the drain can obtain nenergy much higher than the thermal energy of carriers
  - called *hot carriers*
- .The carriers reach velocity saturation (high velocity)
- Hot carriers can tunnel through the gate oxide causing degradation
  - can also cause impact ionization in MOSFETs.
- Concept used in FLASH memory along with a floating gate to trap charges

Hot Carrier Generation

- Carriers drifting near the drain can obtain nenergy much higher than the thermal energy of carriers
  - called *hot carriers*
- .The carriers reach velocity saturation (high velocity)
- Hot carriers can tunnel through the gate oxide causing degradation
  - can also cause impact ionization in MOSFETs.
- Concept used in FLASH memory along with a floating gate to trap charges

Hot Carrier Generation

- Carriers drifting near the drain can obtain nenergy much higher than the thermal energy of carriers
  - called *hot carriers*
- .The carriers reach velocity saturation (high velocity)
- Hot carriers can tunnel through the gate oxide causing degradation
  - can also cause impact ionization in MOSFETs.
- Concept used in FLASH memory along with a floating gate to trap charges

Hot Carrier Generation

- Carriers drifting near the drain can obtain nenergy much higher than the thermal energy of carriers
  - called *hot carriers*
- .The carriers reach velocity saturation (high velocity)
- Hot carriers can tunnel through the gate oxide causing degradation
  - can also cause impact ionization in MOSFETs.
- Concept used in FLASH memory along with a floating gate to trap charges

Lightly-Doped Drain (LDD)

- In CMOS, lightly doped drain (LDD) structure is used to reduce the effect of hot carriers and to reduce short channel effects
  - After Poly patterning an LDD implant is performed in the active regions
  - Afterwards, a nitride spacer is formed by depositing nitride and etching it
  - The nitride spacer acts as a hard mask for the subsequent implants and protects the LDD region
  - The n+ source/drain is formed by a heavy dose of n+ implant

Lightly-Doped Drain (LDD)

- In CMOS, lightly doped drain (LDD) structure is used to reduce the effect of hot carriers and to reduce short channel effects
  - After Poly patterning an LDD implant is performed in the active regions
  - Afterwards, a nitride spacer is formed by depositing nitride and etching it
  - The nitride spacer acts as a hard mask for the subsequent implants and protects the LDD region
  - The n+ source/drain is formed by a heavy dose of n+ implant

Lightly-Doped Drain (LDD)

- In CMOS, lightly doped drain (LDD) structure is used to reduce the effect of hot carriers and to reduce short channel effects
  - After Poly patterning an LDD implant is performed in the active regions
  - Afterwards, a nitride spacer is formed by depositing nitride and etching it
  - The nitride spacer acts as a hard mask for the subsequent implants and protects the LDD region
  - The n+ source/drain is formed by a heavy dose of n+ implant

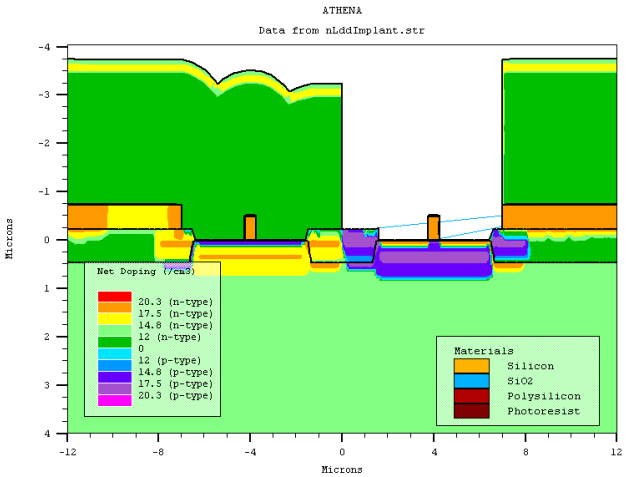
Lightly-Doped Drain (LDD)

- In CMOS, lightly doped drain (LDD) structure is used to reduce the effect of hot carriers and to reduce short channel effects
  - After Poly patterning an LDD implant is performed in the active regions
  - Afterwards, a nitride spacer is formed by depositing nitride and etching it
  - The nitride spacer acts as a hard mask for the subsequent implants and protects the LDD region
  - The n+ source/drain is formed by a heavy dose of n+ implant

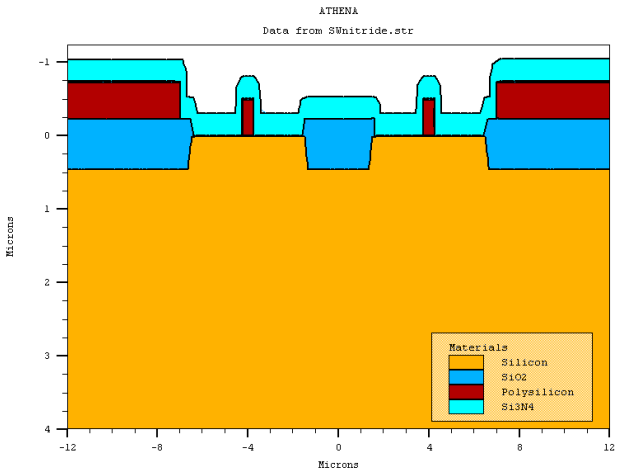
Lightly-Doped Drain (LDD)

- In CMOS, lightly doped drain (LDD) structure is used to reduce the effect of hot carriers and to reduce short channel effects
  - After Poly patterning an LDD implant is performed in the active regions
  - Afterwards, a nitride spacer is formed by depositing nitride and etching it
  - The nitride spacer acts as a hard mask for the subsequent implants and protects the LDD region
  - The n+ source/drain is formed by a heavy dose of n+ implant

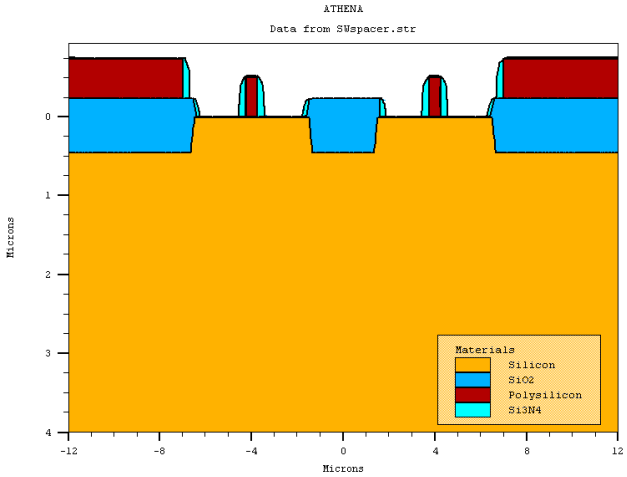
LDD contd.



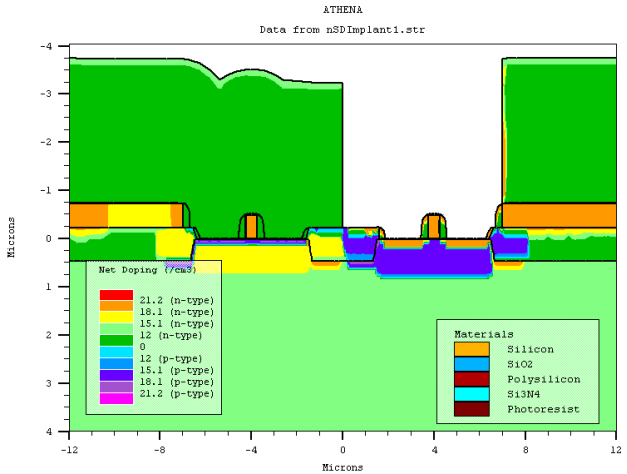
1) NMOS LDD implant



2) Nitride deposition



3) Nitride spacer formation



4) n+ source/drain implant

Oxide Breakdown

- Short channel MOSFETs employ thin gate oxides for faster operation
- The electric field across the gate oxide should be limited to $1V/10\text{Å}$ to avoid oxide breakdown

Oxide Breakdown

- Short channel MOSFETs employ thin gate oxides for faster operation
- The electric field across the gate oxide should be limited to $1V/10\text{Å}$ to avoid oxide breakdown

Gate Tunnel Current

- As the oxide thickness (t_{ox}) scales down, the probability of carriers tunneling through the gate oxide increases
  - leads to gate leakage current
- Gate leakage current increases by an order of magnitude with every node scaling
  - major concern in VLSI design
- High dielectric (Hi-K) gate stacks are employed to increase C'_{ox} while keeping t_{ox} constant
  - scaling without increase in gate leakage

Gate Tunnel Current

- As the oxide thickness (t_{ox}) scales down, the probability of carriers tunneling through the gate oxide increases
  - leads to gate leakage current
- Gate leakage current increases by an order of magnitude with every node scaling
  - major concern in VLSI design
- High dielectric (Hi-K) gate stacks are employed to increase C'_{ox} while keeping t_{ox} constant
  - scaling without increase in gate leakage

Gate Tunnel Current

- As the oxide thickness (t_{ox}) scales down, the probability of carriers tunneling through the gate oxide increases
  - leads to gate leakage current
- Gate leakage current increases by an order of magnitude with every node scaling
  - major concern in VLSI design
- High dielectric (Hi-K) gate stacks are employed to increase C'_{ox} while keeping t_{ox} constant
  - scaling without increase in gate leakage

Substrate Current Induced Body Effect (SCBE)

- Hot carriers can cause impact ionization near the drain
  - hole current is generated and flows through the substrate resistance
- The substrate current increases the substrate potential causing V_{THN} shifts, latchup and degradation of transconductance (g_m) and output resistance (r_o) in short channel devices

Substrate Current Induced Body Effect (SCBE)

- Hot carriers can cause impact ionization near the drain
  - hole current is generated and flows through the substrate resistance
- The substrate current increases the substrate potential causing V_{THN} shifts, latchup and degradation of transconductance (g_m) and output resistance (r_o) in short channel devices

Gate Induced Drain Leakage (GIDL)

- Occurs at higher drain biases in the off state of a transistor.
  - example, a DRAM cell with “1” stored on the capacitor and the gate is “0”
- Band-to-band tunnelling occurs at the drain and gate-oxide overlap region.
  - decreases retention of charge in a DRAM cell.
- In a logic cell, the subthrehold and junction leakages dominate over GIDL.

Gate Induced Drain Leakage (GIDL)

- Occurs at higher drain biases in the off state of a transistor.
  - example, a DRAM cell with “1” stored on the capacitor and the gate is “0”
- Band-to-band tunnelling occurs at the drain and gate-oxide overlap region.
  - decreases retention of charge in a DRAM cell.
- In a logic cell, the subthrehold and junction leakages dominate over GIDL.

Gate Induced Drain Leakage (GIDL)

- Occurs at higher drain biases in the off state of a transistor.
  - example, a DRAM cell with “1” stored on the capacitor and the gate is “0”
- Band-to-band tunnelling occurs at the drain and gate-oxide overlap region.
  - decreases retention of charge in a DRAM cell.
- In a logic cell, the subthrehold and junction leakages dominate over GIDL.