Reliability Analysis of Nanoscale CMOS Device with HCI Effect

Sandeep Lalawat and Prof. Y.S. Thakur

Abstract: The paper focuses on the reliability due to hot carrier effects of CMOS and the HC effect under high gate voltage stress. How these stress factors vary with operating conditions, the required foundry qualification test data, and how to alter circuit topologies and device geometries to mitigate their effect are discussed. Also where device reliability stress is high, a matter of serious concern is how many devices are being operated under such conditions. As the number of devices operated under high stress.

Keywords: CMOS; HCI; CHE; CHISEL;

NOMENCLATURE

HCI: Hot carrier injection: Carrier injection into the channel or gate insulator produced by impact ionization near the drain end of the channel

CHE: Channel hot electron: Impact ionization produced electrons resulting from the vector sum of lateral and transverse electric fields near the drain end of the channel. In the CHE mode of hot electron damage, the source-bulk voltage is zero.

CHISEL: Channel-initiated secondary-electron generation: secondary hot electrons produced when carriers resulting from a first impact ionization near the drain are accelerated toward the bulk, generating hole-electron pairs in a secondary impact ionization event. In the CHISEL mode of hot electron

Introduction

HCI Effects: Operation of bulk CMOS high dc gain, wide-dynamic-range output stages within the < \text{1 V H P}} technology VCC_{MAX} limits at the 45 nm technology node and beyond is a significant design challenge. It is therefore useful to understand both the specified limits of the process and how rapidly reliability stress factors increase as a limit is reached, or exceeded. All reliability stress factors do not impose equal device lifetime limits when supply voltages reach VCC_{MAX}.

Foundry-prescribed OD voltages VCC_{OD} in excess of VCC_{MAX} are usually specified so that no single reliability factor will cause unacceptable reliability lifetime degradation. For example, under certain conditions, HCI stress may, or may not, substantially degrade device lifetime at Vd, voltages (VCC_{EOD}) in excess of specified OD limits. Low HCI-stress factors include low drain current, long device L, certain drain voltage waveform shapes, etc. Other limits, such as the V_{ps} limit based on TDDB stress, have fewer low-stress factors. Reliability stress factors do not cause step changes in device lifetimes at the specified process voltage limits.

HCI in the case where V_{ab} = 0 is commonly known as the CHE mode of device degradation. In CHE, impact ionization results in carrier pairs generated near the drain end of the channel causing interface trap (N_{it}) and N_{int} buildup. The buildup of N_{it} and N_{int} near the gate-channel interface degrades the I-V characteristics and eventually results in device failure.

Cascade circuits where V_{ab} > 0 have been used for a long time in analog, voltage multiplier, and current-mode logic digital designs, to overcome V_{ds} limitations. Operation at V_{ab} > 0 has been used by EEPROM circuit designers to obtain higher electron programming energies and currents. As V_{ab} is increased, the drain-bulk field increases, and one carrier from the first carrier pair generated by impact ionization near the drain may be accelerated toward the bulk with sufficient energy to generate a secondary carrier pair. One member of that pair is consequently accelerated upward, toward the gate acquiring high energy and capability for creating damage. In an NMOS device, this process is known as channel-initiated secondary electron generation, or CHISEL mode of HCI stress. The process can also occur in PMOS devices at nearly the same voltage stress levels, with equally serious damage consequences.

a) HCI stress in CHE mode: In CHE, the buildup of N_{it} is directly proportional to transverse field strength (E_y) and to the ratio of bulk-to-drain current (Ib/Id)^2, where n is slightly lower than 0.5. At low Vg, Ib/Id is low, but E_y is high and proportional to Vd — Vg. Because E_y is high, impact ionization multiplication factors are high. However, very low drain current results in low Ib, even in the presence of a high avalanche multiplication factor. Conversely, for high Vg, Vd — Vg is low, E_y is low, and low impact ionization rates result in reducing Ib to a small value. Most process qualification for HCI damage is usually with V_{ab} = 0 at
maximum substrate current, where \( V_{gs} \) (or \( V_d - V_g \)) is approximately \( V_{ds}/2 \). Testing at maximum substrate current is not the condition that corresponds to the maximum HCI damage rate when \( V_{sb} \) is more than a small fraction of \( V_{ds} \). At high \( V_{sb} \), gate current is the best indicator for damage in an NMOS device.

The relative magnitudes of the device terminal voltages determine the location of the hot carrier injection along the channel and have a substantial effect on the type and degree of degradation. With \( V_{sb} = 0 \), lateral field \( E_y \) is relatively independent of \( V_g \) and generally peaks between the drain junction and the edge of the lightly doped drain region. High transverse \( E_y \) fields result from \( V_d - V_g \). \( E_y \) and \( E_z \) work together to accelerate carriers in the channel and create HCI. As \( V_g \) approaches \( V_d \), \( E_y \) is substantially reduced but \( E_z \) fields are still high. Substrate current decreases significantly as \( N_d \) and \( N_{oa} \) peak values diminish but widen, and move toward the center of the channel. There, the \( N_d \) and \( N_{oa} \) peak values become more efficient in degrading analog \( g_m \) and \( V_t \) characteristics. Increasing device length by 25% to 50% over minimum reduces maximum \( V_d - V_g \) reduces only \( E_y \). While bulk current is reduced to a small value as \( V_g \) approaches \( V_d \), the damage rate for analog \( g_m \) and \( V_t \) may be reduced by only 20% because \( E_z \) remains relatively unchanged.

Peak \( V_{ds} \) foundry voltage limits for HCI stress are commonly based on a digital-design requirements where \( I_{sat} \) is degraded at end-of-life by 10%. However, excess device \( g_m \) noise, and \( V_t \) shifts are consequences of increased \( N_d \) and \( N_{oa} \) trap levels prior to device failure. A 20% increase in \( V_d \) beyond \( V_{CCOD} \) limits can easily reduce the lifetime of an NMOS by a factor of ten because of the exponential dependence of damage rate on \( V_d \). The output of an amplifier or voltage-controlled oscillator (VCO) with an inductive load may have peaks above the supply voltage during normal operation. In that case, the use of peak values of \( V_{ds} \), for evaluating HCI stress conditions may be a bit conservative. In circuit designs where an extra 50-100 mV of output peak voltage could provide a valuable increase in dynamic range, it is useful to estimate the effect of the additional peak voltage on the damage rate.

b) Obtaining additional device headroom in CHE mode: HCI damage where \( V_{sb} = 0 \) progresses at a rate proportional to \( e^{E/\Delta V} \cdot f(t) \), where \( C \) is the voltage stress coefficient \( B \cdot VCC_{MAX} \). The value of \( C \) is obtained from foundry reliability data, and typical values of \( B = C/\Delta V_{CC} \) range from 0.1 to 0.2 at the 100 nm node. The second term \( f(t) \) is a slowly varying nonlinear function of time, such as \( t^n \), typically \( n \) ranges from 0.2 to 0.5.

When an inductive load is present, such as in a high-frequency VCO, the drain terminal voltage can easily exceed the supply voltage. Extra headroom can be obtained when peak \( V_{ds} \) values for sinusoidal waveforms are allowed to exceed \( VCC_{MAX} \) by a value up to an excess overdrive limit \( V_{EOD} \). If the relationship between stress and \( V_d - V_g \) were linear, then the HCI damage rate would be independent of the amplitude of the sine wave. In that case \( V_{EOD} \) could exceed \( VCC_{MAX} \) by the zero-to-peak amplitude of the sine wave, \( V_{pk} \). The stress relationship is unfortunately not linear; however, given a value for the coefficient \( C \), the value of \( V_{EOD} \) can be determined.

The damage rate equation can be used to find the voltage difference (AV) between the average sinewave voltage and \( VCC_{MAX} \), where the damage rate for the sinewave is the same as it is for dc. Once AV is determined, then \( V_{EOD} \) is related to \( VCC_{MAX} \) by

\[
V_{EOD} = VCC_{MAX} - AV + V_{O-PK}
\]

Substituting \( VCC_{MAX} - AV + V_{O-PK} \times \sin (\omega t) \) for \( V_{ds} \) in the first term of the steady-state damage rate

\[
e^{-\frac{VCC_{MAX} - AV + V_{O-PK} \sin (\omega t)}{B - VCC_{MAX}}}
\]

is then integrated over one cycle of the sinewave, and the result is equal to the dc damage rate term \( e^{E/\Delta V_{dc}} \) or \( e^{E/\Delta V_{CC} C} \). When the variable changes \( V^{\omega-PK} = A \cdot VCC_{MAX} \), \( C = B \cdot VCC_{MAX} \), and \( AV = E \cdot VCC_{MAX} \) are made, the equality between ac and dc damage rate is

\[
\int_0^{2\pi} \frac{1}{2\pi} e^{1/B - E/B + (A/B) \sin (\omega t)} \, d\omega = \frac{1}{B}
\]

Cancelling the \( e^{1/B} \) terms, factoring out the \( e^{E/\Delta V_{dc}} \) term, and taking the natural log, the last equation reduces to

\[
\frac{E}{B} = \ln \left( \frac{1}{2\pi} \int_0^{2\pi} e^{(A/B) \sin (\theta)} \, d\theta \right)
\]

The value of \( E/B \) can be obtained using numerical integration methods to evaluate the sine integral from zero to \( 2\pi \), and taking the log of the integration result. \( E/B \) as a function of \( A/B \) is shown for a practical range of \( A/B \) (recall \( A/B = V^{\omega-PK}_{dc} / V_{sb} = C \)). Where \( A/B = 2 \) (\( A/B = 0.2/0.1 \)) and \( VCC_{MAX} = 1.0 \) V, an extra 119mV of headroom can be gained; then \( V_{EOD} = 1.119 \) V. For a higher amplitude sinewave, where \( A/B = 4 \), the gain in headroom is 159 mV. Note that the value \( E/A =\)
AV/V°_pk. Therefore, AV = E/A * V°_pk and VEOD = VCC_MAX — AV + V°_pk. As the sinewave amplitude increases, the fraction of amplitude that contributes to increased stress E/A increases, but not fast enough to prevent VEOD from increasing beyond VCC_MAX.

Some pipeline ADC stages and analog sample-and-hold circuits have signal distributions that are substantially uniform over a moderately wide voltage range. For such a stage, A is related to the peak-to-peak amplitude (V_p-p), as in the square wave case by V_p-p_MAX = 2A * VCC_MAX. In a pipeline AD stage with uniform signal distribution, the output is approximated by a square wave with uniform voltage probability density between —A and +A.

$$E = \ln \left( \frac{1}{2A} \int_{-A}^{A} e^B \, dx \right)$$

Using an analytical solution for the integral yields the following result for E/B:

$$E = \ln \left( \frac{B}{2A} \left( e^A - e^{-A} \right) \right)$$

For higher amplitudes, such as A B = 4.208/MV of extra headroom is gained, in the VCC_MAX = 1.0 V example of Excess headroom for the random voltage distribution case is higher than in the sinewave case. This is because the probability of the waveform’s being at any voltage at any time is uniform, while the sinewave spends a larger fraction of time at peak amplitudes.

This general method can be used for various signal distributions and other reliability stress factors with similar exponential dependence on voltage. For an asymmetric digital subscriber line signal-processing case, having typically 2:1 peak-to-average ratios, the signal statistics are even more favorable to excess overdrive voltage allowance. Good analog design procedure, where EOD operating modes are planned, requires extensive consultation with device engineers who understand all of the relevant reliability factors, an analysis of how the EOD will affect total chip reliability, backup by process qualification data, and an awareness that operation too far outside the normal HP limits may require consideration of additional damage modes in the device lifetime predictions.

c) HCI stress in CHISEL mode: Series-connected CMOS devices, as in current-source-differential pair and cascode circuit combinations, reduce V_{ds} drops across the individual devices. However, while V_{ds} is reduced, V_{sb} is increased in one or more of the series-connected devices. When V_{sb} > 0, the hot carriers have higher energy as a result of high transverse field (E_y) acceleration and cause more HCI (CHISEL mode) damage. High E_y also results in more NBTI damage. Voltage ranges of interest circuits with series-connected CMOS include V_{db} > VCC_MAX and V_{sb} > 0 up to V_{sb} = VCC_MAX. For those ranges, accurate data exist for model parameter extraction, but in many cases, long-term reliability data do not. Such data are vital in overcoming headroom limitations that result from future 1/k voltage scaling. Early access to V_{sb} > 0 preliminary reliability projections is required for nanoscale analog CMOS circuit design.

When V_{sb} voltages reach half the supply voltage, the damage rate compared to the condition V_{gs} = 0 can more than double. When V_{sb} approaches VCC_MAX, the damage rate compared to V_{sb} = 0 can increase by an order of magnitude or more. At low values of V_{ds}, the damage is con-centrated near the drain end of the channel and moves toward the center as V_{sb} increases, resulting in the generated traps’ becoming more efficient in degrading V_t, I_{dss}, and I_m.

For NMOS in CHISEL, the HCI damage buildup is proportional to gate current (I_g), not substrate current. The damage buildup rate is proportional to I_g^2, where n is approximately 0.5. The maximum damage for NMOS in CHISEL is not at maximum substrate current, where V_{ps} is approximately V_{ds} 2 and HCI device qualification is usually performed. It occurs where V_{ps} is near V_{ds}. Low values of V_{ds} — V_{gs} do not mitigate high HCI damage in CHISEL mode.

For PMOS in CHISEL, the maximum damage rate does not necessarily correspond to maximum gate current conditions. In a PMOS device, maximum damage can occur at a gate voltage where gate current is near zero. This has been explained as the result of the current cancellation effect from the two carrier currents of opposite polarity’s combining to generate the damage. Damage under these conditions is more closely correlated to the sum of the absolute value of both hole and electron gate currents. Maximum stress conditions may be difficult to determine from external device currents because the hole and electron currents are of opposite sign.

Some series-device circuit configurations using 100 nm core devices operated at VCC = 1.8 V, which could
have potentially high CHE or CHISEL mode stress with $V_{sb} > 0$, are shown in Fig. 8. As in CHE, slight increases in device length will help mitigate high CHISEL mode stress. However, device operation where $V_{db}$ or $V_{gb}$ exceeds the HP VCC_MAX or VCC_OD specifications for the process (EOD conditions) must be subjected to careful reliability analysis.

CONCLUSION

Analog designers favor the use of cascode circuit configurations $V_{sb} > 0$ to mitigate drain-source voltage Vds stress effects, most foundry device qualification are based on $V_{sb} = 0$ test conditions. The condition $V_{sb} > 0$ is playing a stronger role in both hot and cold carrier damage effects for highly scaled devices. Unfortunately, the foundries are providing inadequate reliability models and/or qualification data for analog design using the $V_{sb} > 0$ condition, where devices are subjected to the CHISEL mode of HCI stress. In some cases, it may be helpful to exceed foundry specified drain-source voltage limits by a few hundred millivolts.

REFERENCES