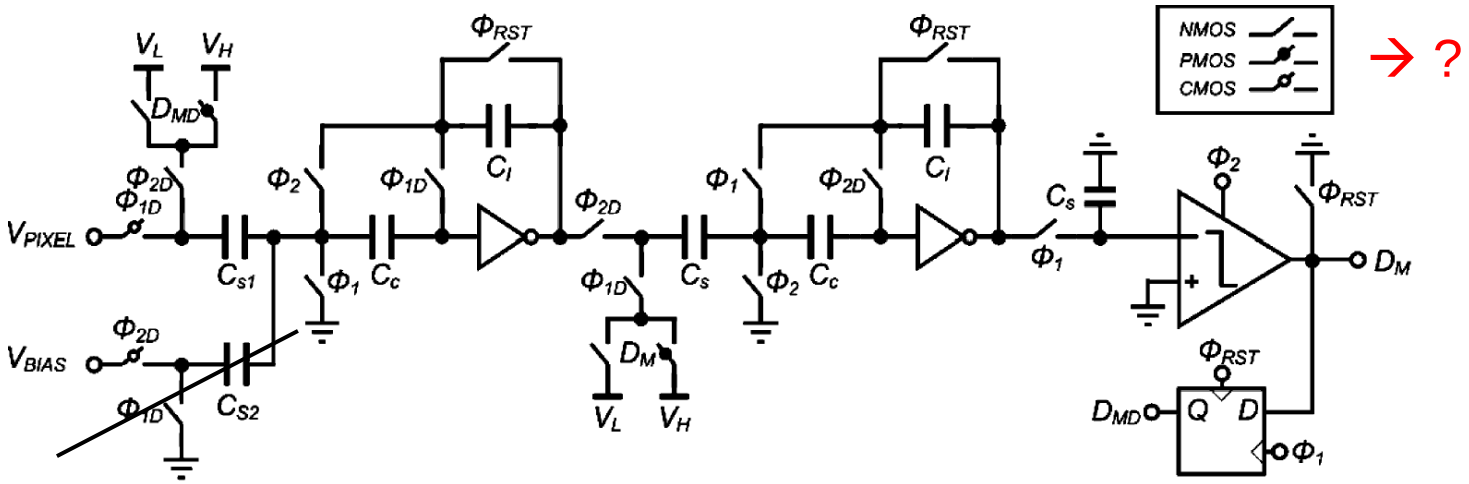
Switches

Boris Murmann

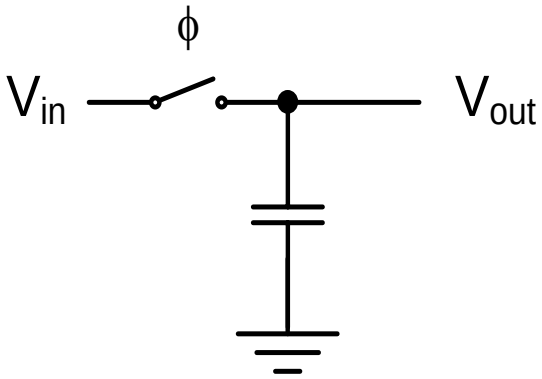
bmurmann@hawaii.edu

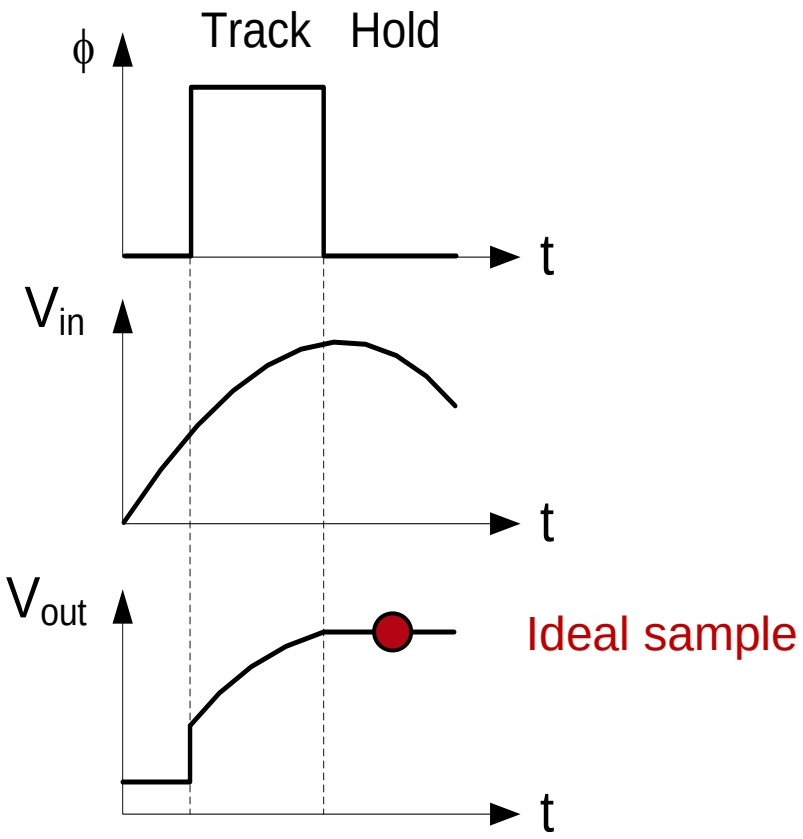
Motivation



- The last piece of our circuit that needs to be transistorized are the switches and the clock generator

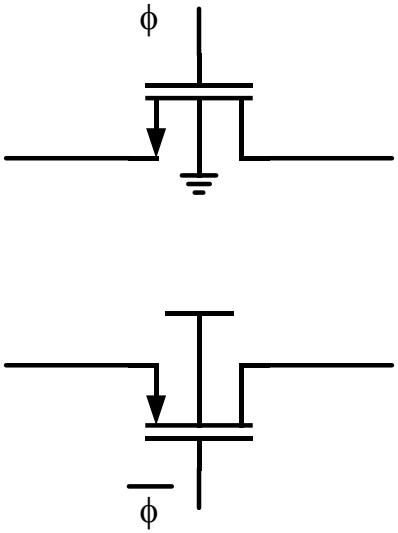
Ideal Switch Operation



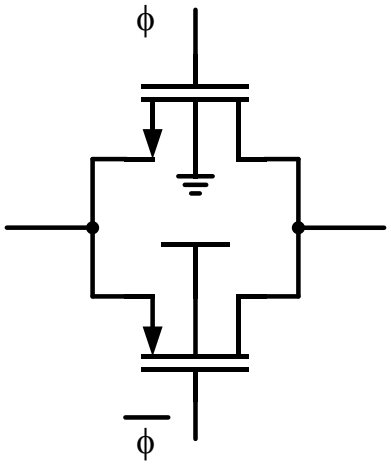


Common Switch Implementation Options

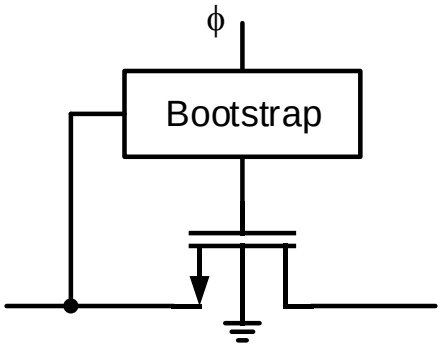
NMOS or PMOS



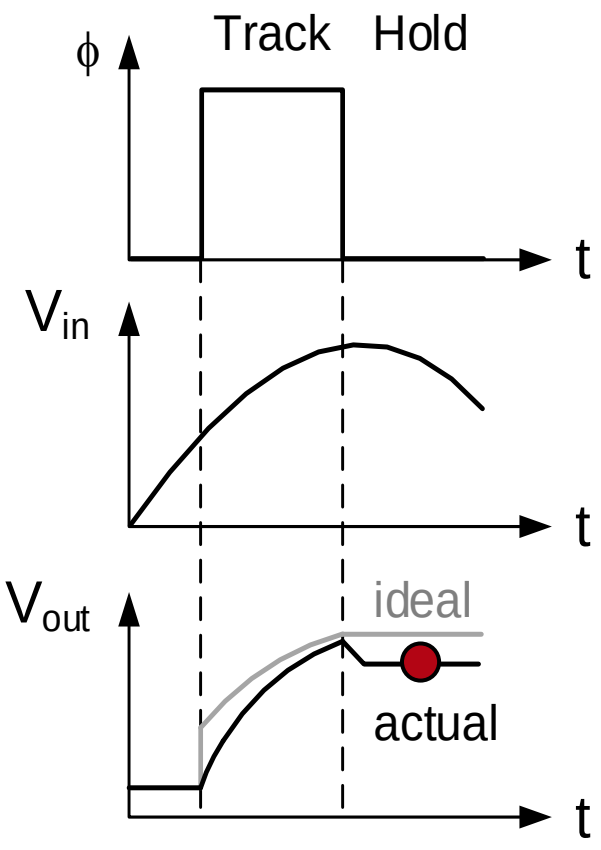
CMOS
(transmission gate)



Bootstrapped NMOS
(advanced concept)



Long List of Nonidealities



- Sampling noise (kT/C)
- Finite on resistance
- Charge injection and clock feedthrough
- Clock jitter
- Tracking nonlinearity
- Hold-mode feedthrough and leakage

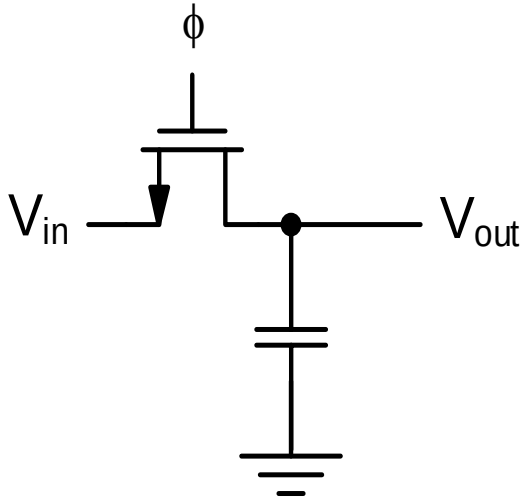
- As a result, samples contain
  - Offset
  - Noise
  - Nonlinear errors (distortion)

On-Resistance of NMOS Switch

$$I_D \approx \mu C_{ox} \frac{W}{L}\left(V_{GS} - V_t - \frac{V_{DS}}{2}\right) V_{DS}$$

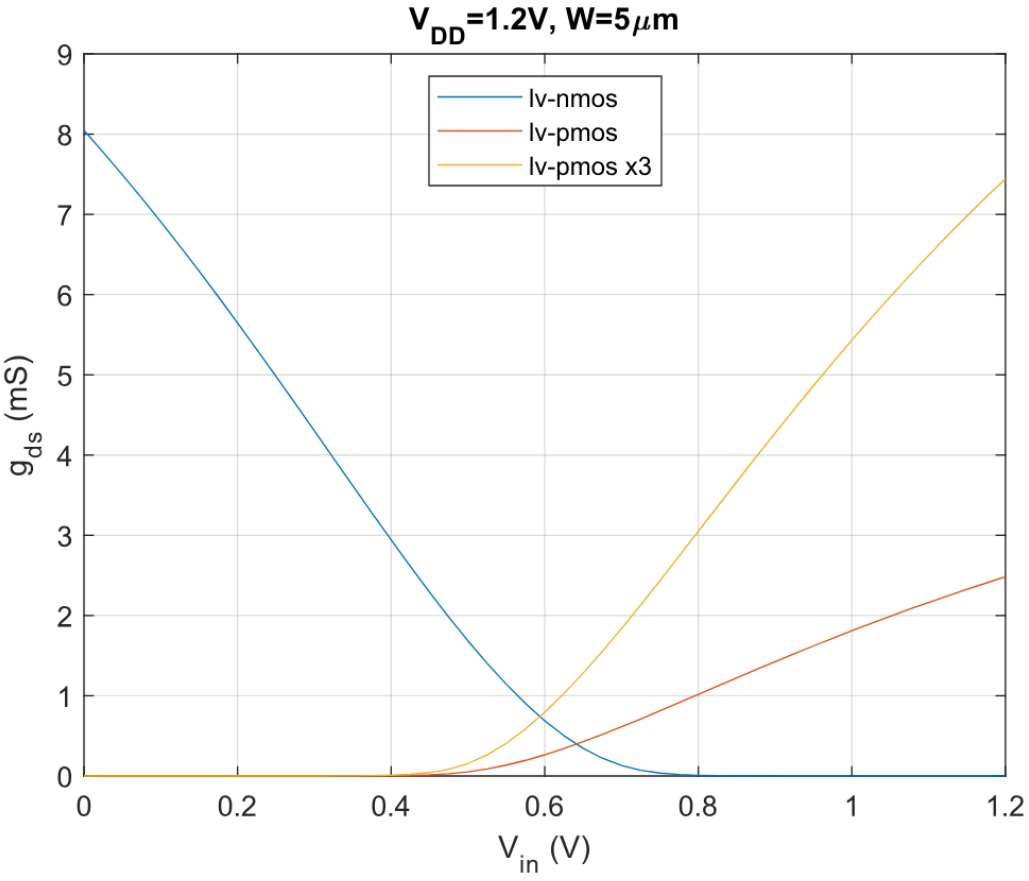
$$r_{ds} \approx \left[\left.\frac{dI_D}{dV_{DS}}\right|_{V_{DS}\to 0}\right]^{-1} = \frac{1}{\mu C_{ox} \frac{W}{L}(V_{GS} - V_t)}$$

$$r_{ds} \approx \frac{1}{\mu C_{ox} \frac{W}{L}(\phi - \boldsymbol{V_{in}} - V_t)} \qquad g_{ds} = \frac{1}{r_{ds}}$$



- The on-resistance varies with V_{in}
  - NMOS switch works well for low V_{in}
  - PMOS switch works well for high V_{in}

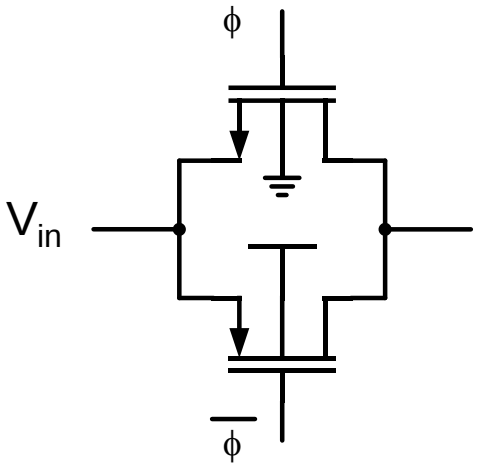
Simulation of NMOS/PMOS Switches in IHP Technology



https://github.com/bmurmann/EE628/tree/main/4_Technology/6_ron_tgate

CMOS Switch (Transmission Gate)

- For increasing V_{in}, NMOS resistance increases, but PMOS resistance decreases → sounds great!



In the on-state:

$$V_{GSn} = V_{DD} - V_{in}$$

$$V_{SGp} = V_{IN}$$

$$r_{on} = \frac{1}{\mu_n C_{ox} \frac{W_n}{L_n}\left(V_{GS_n} - V_{tn}\right)} \Bigg\| \frac{1}{\mu_n C_{ox} \frac{W_n}{L_n}\left(V_{SG_p} - \left|V_{tp}\right|\right)}$$

- Assuming “equal strength” sizing, such that

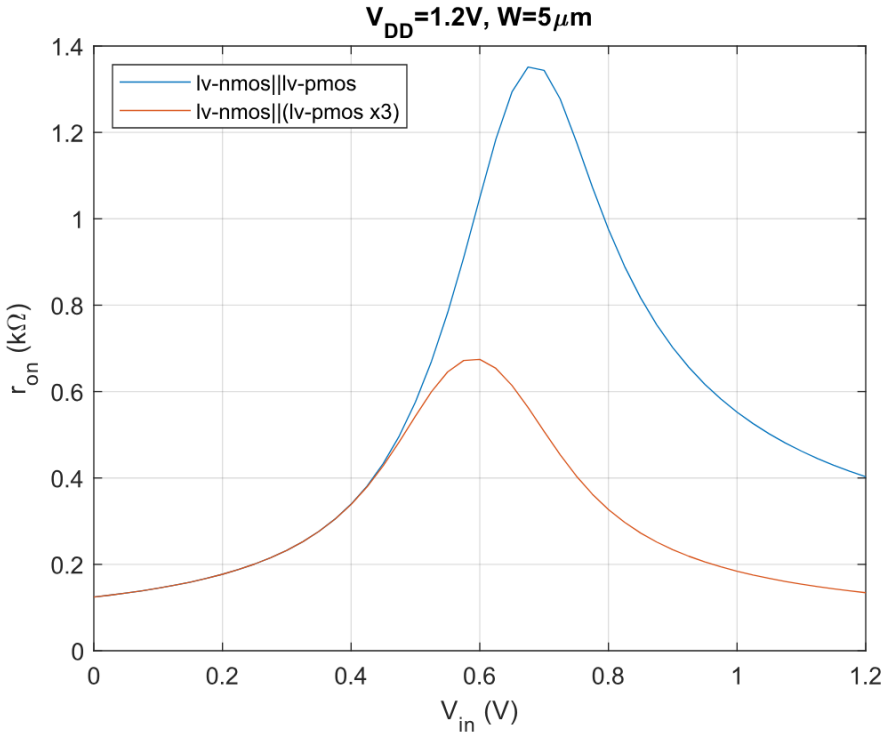
$$\mu_n C_{ox} \frac{W_n}{L_n} = \mu_p C_{ox} \frac{W_p}{L_p}$$

we find

$$r_{on} = \frac{1}{\mu_n C_{ox} \frac{W_n}{L_n} \left(V_{DD} - V_{tn} - |V_{tp}| \right)}$$

- Independent of V_{in} → too good to be true!
- Missing factors
  - Backgate effect
  - Short channel effects

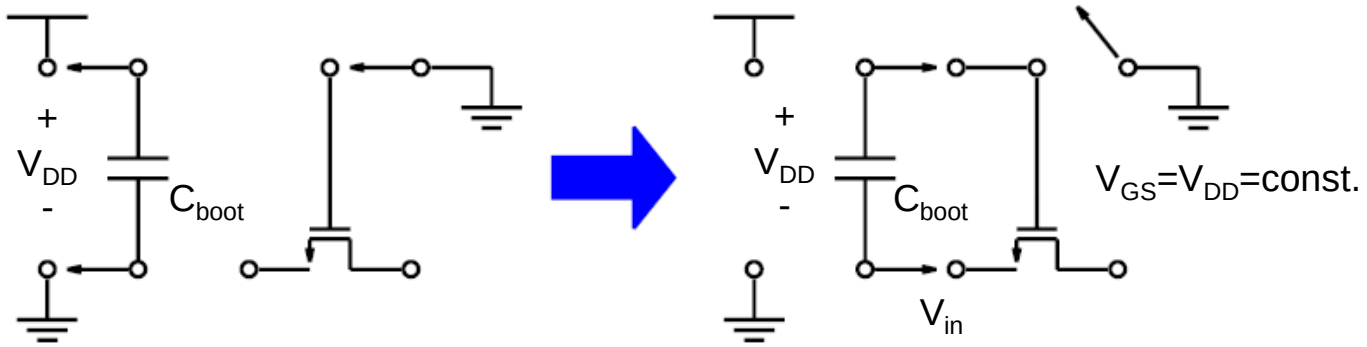
Simulation CMOS Switch in IHP Technology



https://github.com/bmurmann/EE628/tree/main/4_Technology/6_ron_tgate

- Design
  - Size P/N ratio to minimize curvature
  - Size P and N simultaneously to meet r_{on} requirement

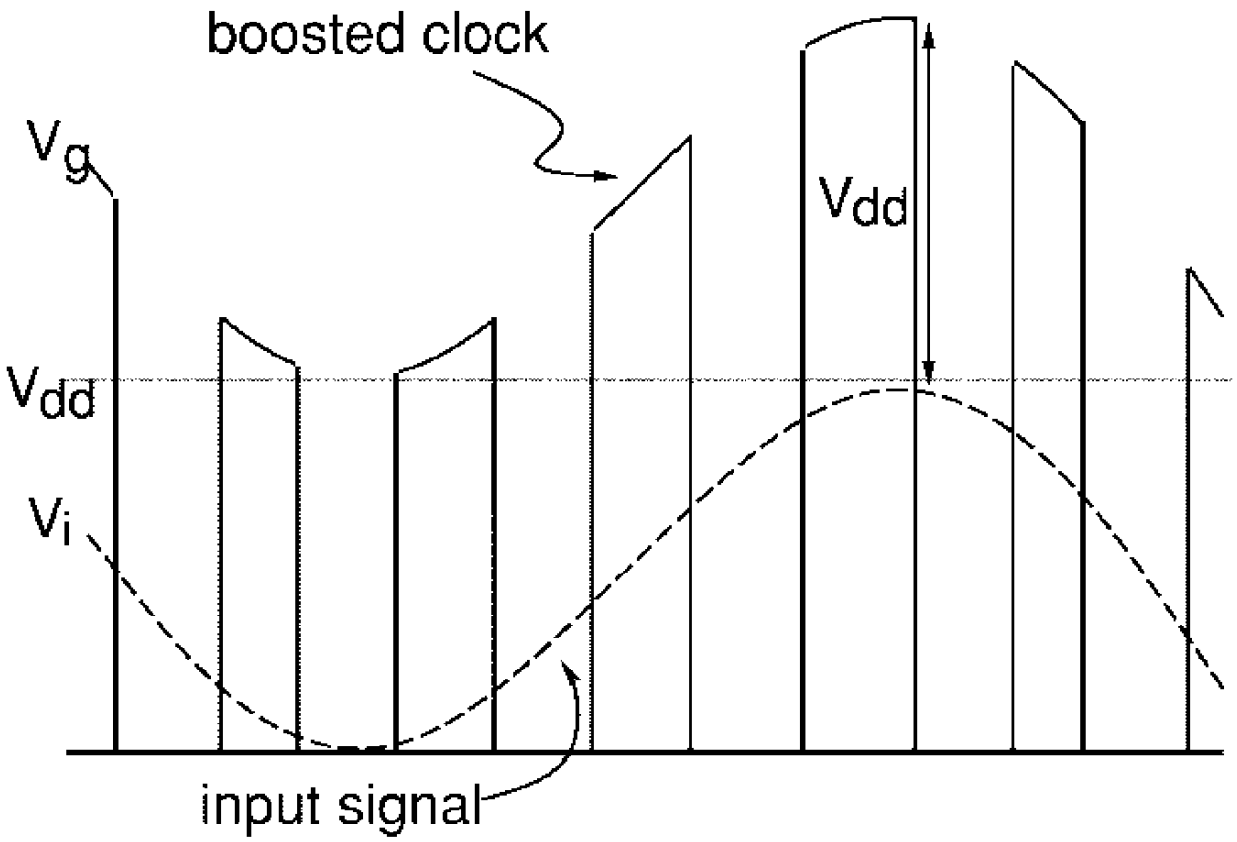
Aside: Clock Bootstrapping



A. Abo, "Design for Reliability of Low-voltage, Switched-capacitor Circuits," PhD Thesis, UC Berkeley, 1999.

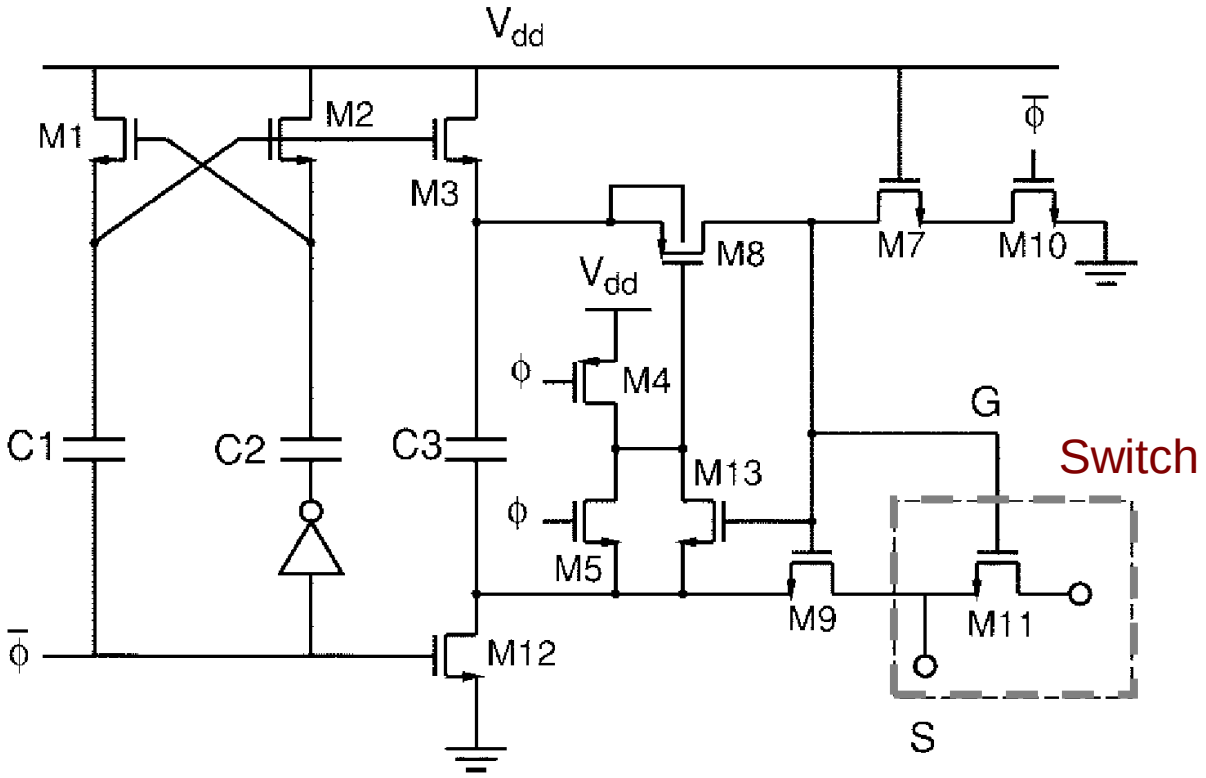
- Phase 1
  - C_{boot} is precharged to V_{DD}
  - Sampling switch is off
- Phase 2
  - Sampling switch is on with $V_{GS}=V_{DD}$=const.
  - To first order, r_{on} is signal independent

Waveforms



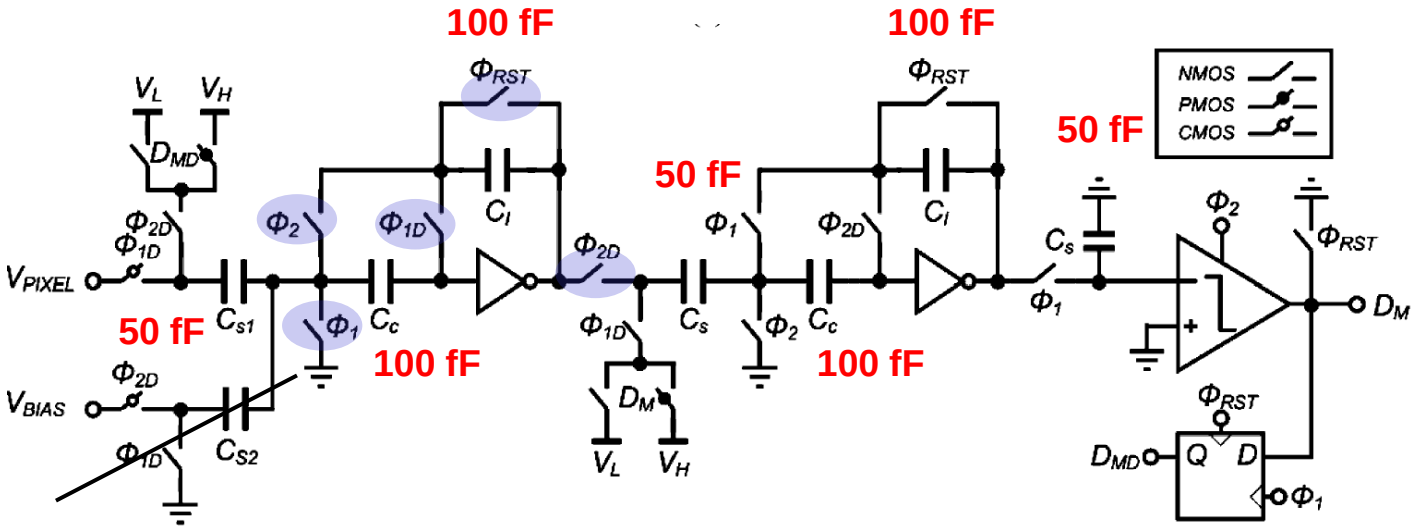
A. Abo, "Design for Reliability of Low-voltage, Switched-capacitor Circuits," PhD Thesis, UC Berkeley, 1999.

Circuit Implementation



A. Abo et al., "A 1.5-V, 10-bit, 14.3-MS/s CMOS Pipeline Analog-to-Digital Converter," IEEE J. Solid-State Circuits, pp. 599, May 1999

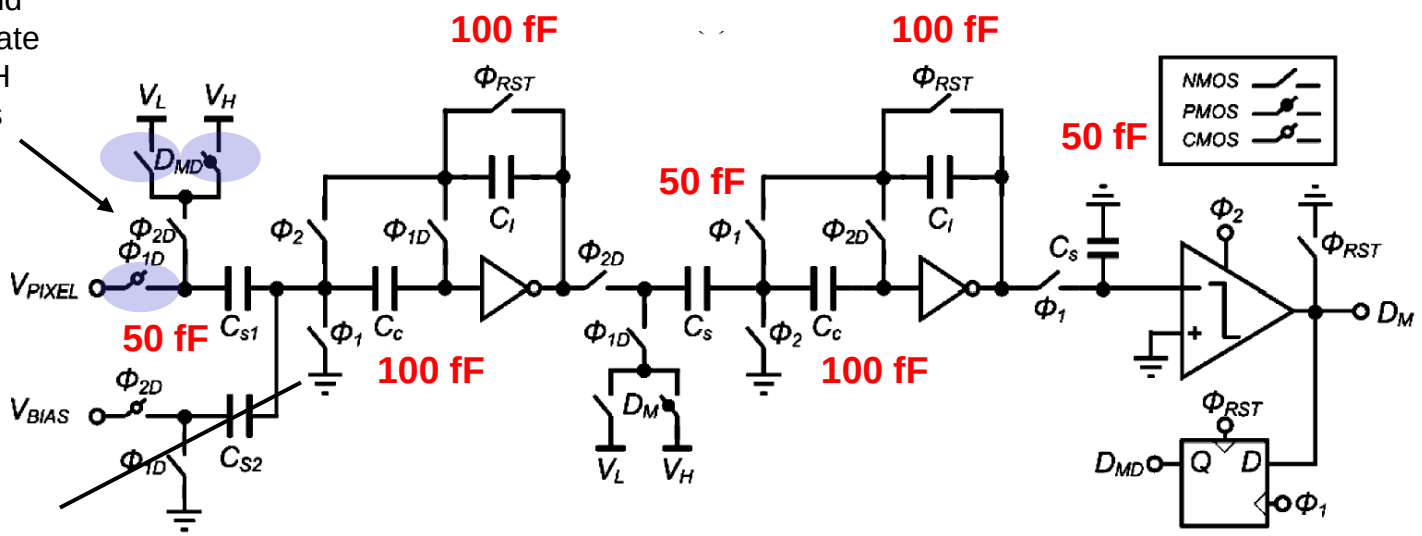
How Large Should the Switch Devices Be?



- Highlighted above are NMOS switches operating around 0.6 V
- For a 5-um device, g_{ds} ~0.6 mS, r_{on} ~ 1.5 kΩ
- With C ~ 150 fF, the time constant is 0.225 ns
- Half clock period at ~50 MHz is 10 ns, i.e., ~44 time constants
- We can reduce the switch size by about 4x, so W > 1.25 um
  - W = 2um is a safe choice (process variation, parasitics, …)

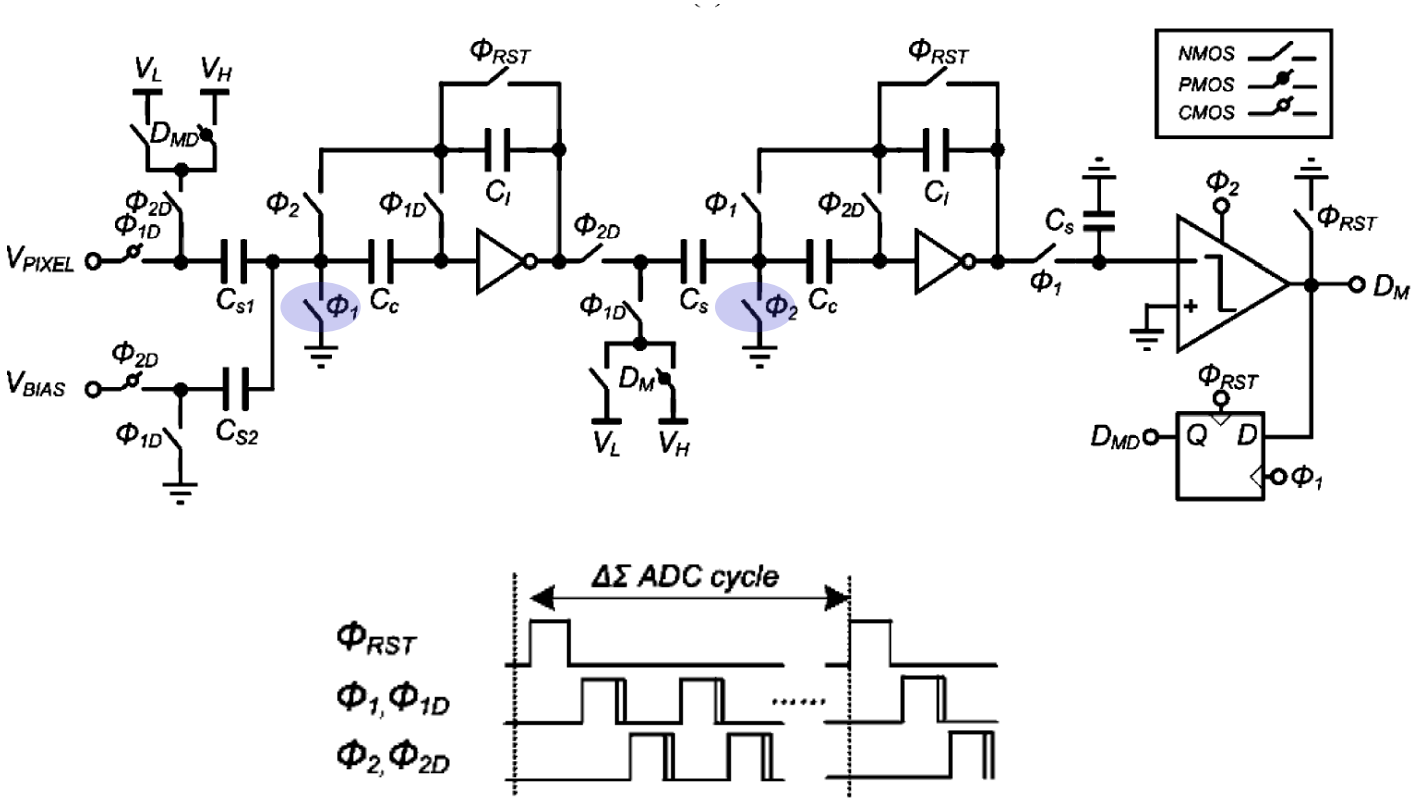
How Large Should the Switch Devices Be?

Don't need this; should use AND gate for VL/VH switches



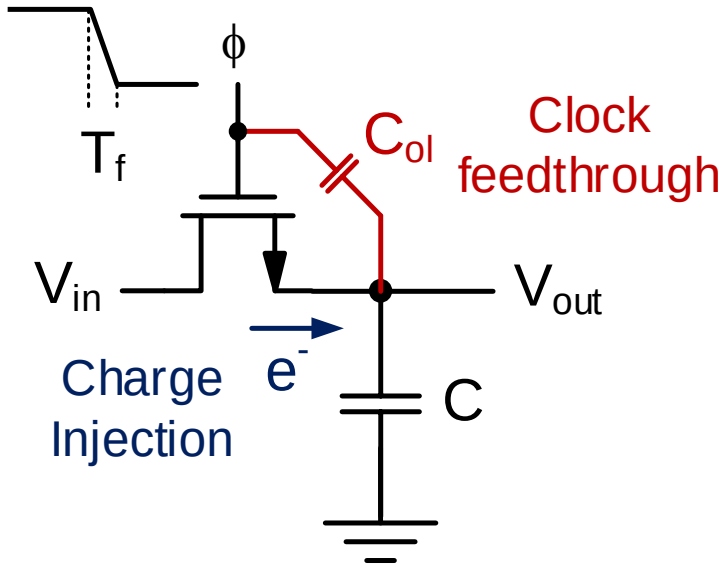
- VL switch can be smaller since g_{ds} at ~0.3 V is 4 mS
  - W = 0.5 um should be fine
- VH switch should be about 3x wider (low PMOS mobility)
- CMOS input switch can be sized like switch on previous slide
  - NMOS ~ 2 um, PMOS ~ 6 um

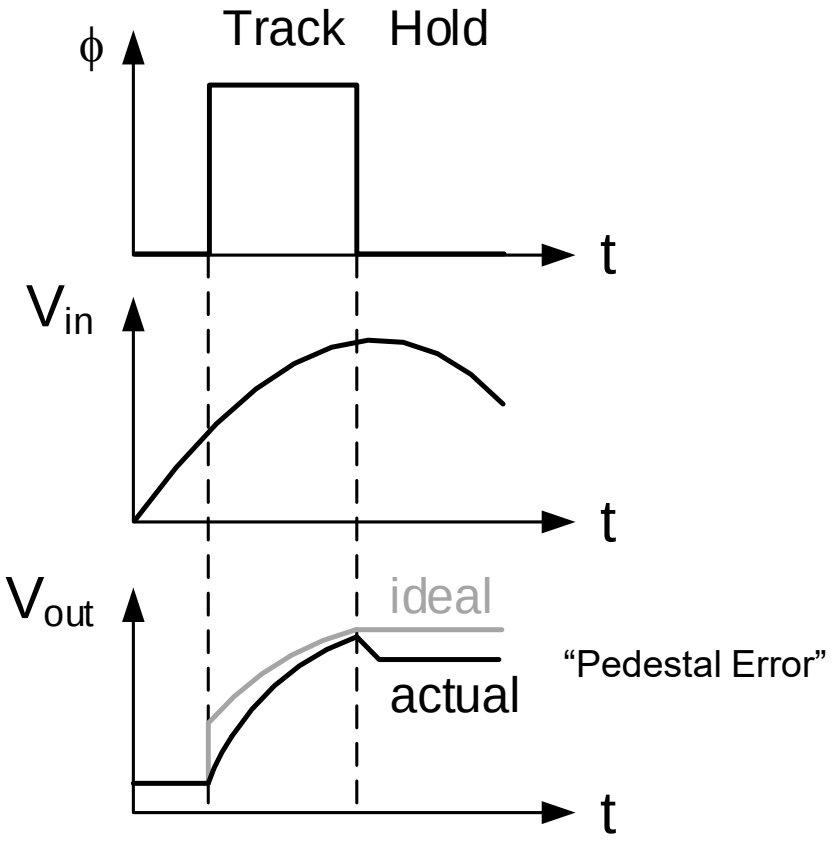
Switch Timing



- Highlighted switches turn off early. Why?

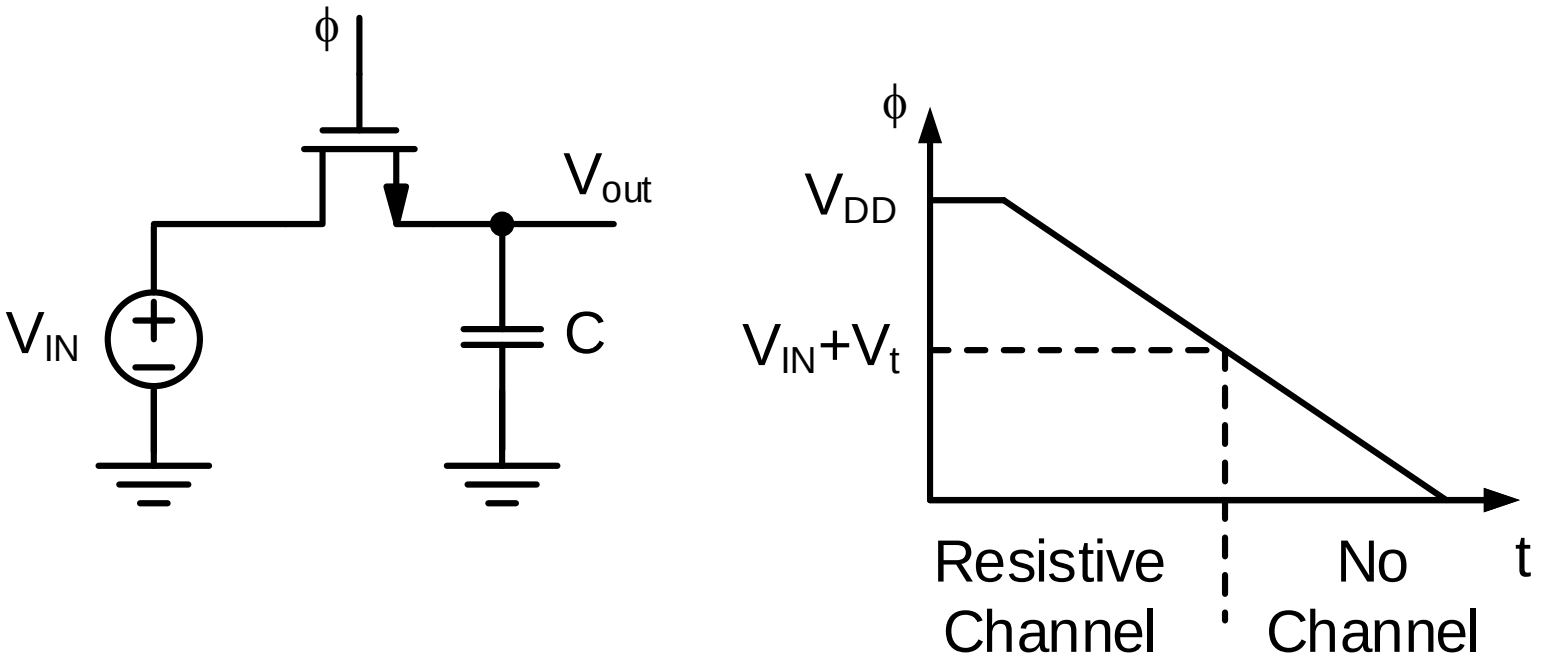
Charge Injection and Clock Feedthrough





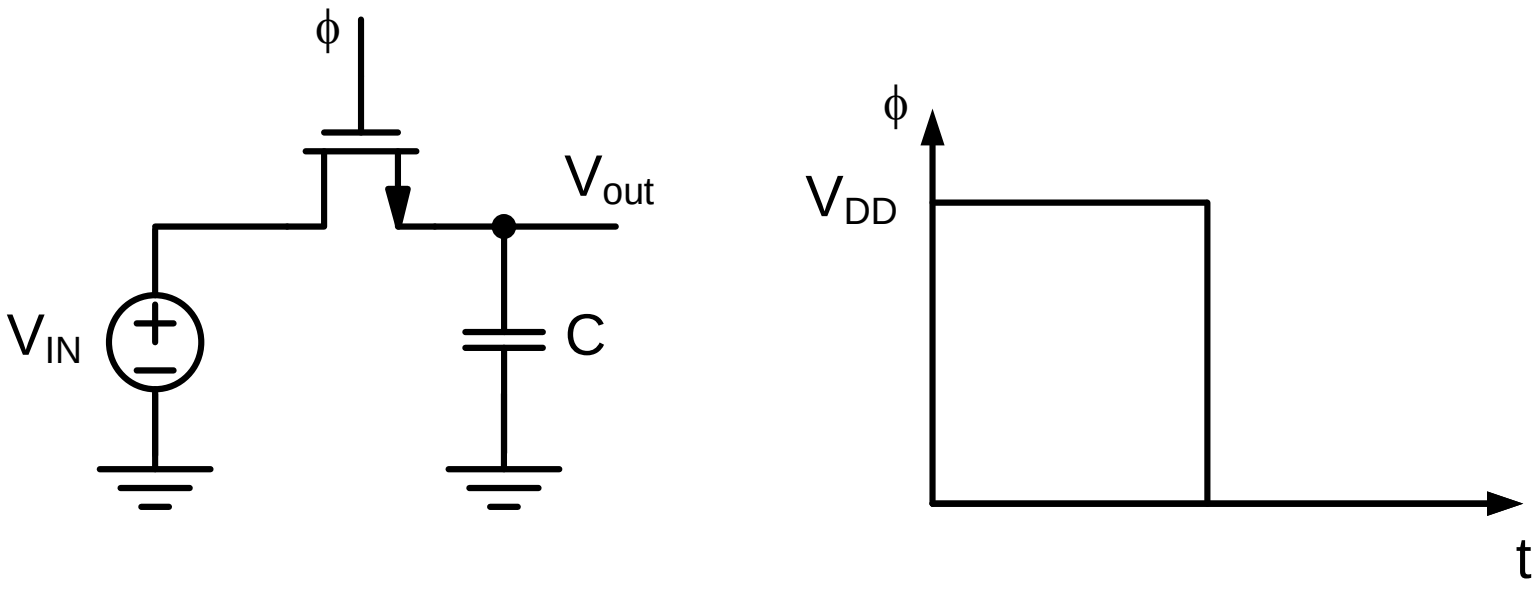
- Focus on charge injection first and analyze two extreme cases
  - Very large T_f ("slow gating")
  - Very small T_f ("fast gating")

Slow Gating



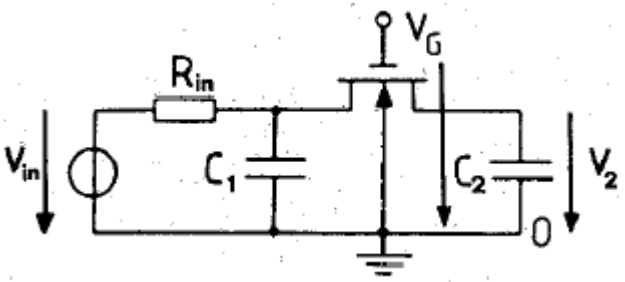
- All channel charge is absorbed by the input source, none of it goes to C
- This is because the MOSFET can manage to equalize V_{IN} and V_{out} during the slow turn-off transient

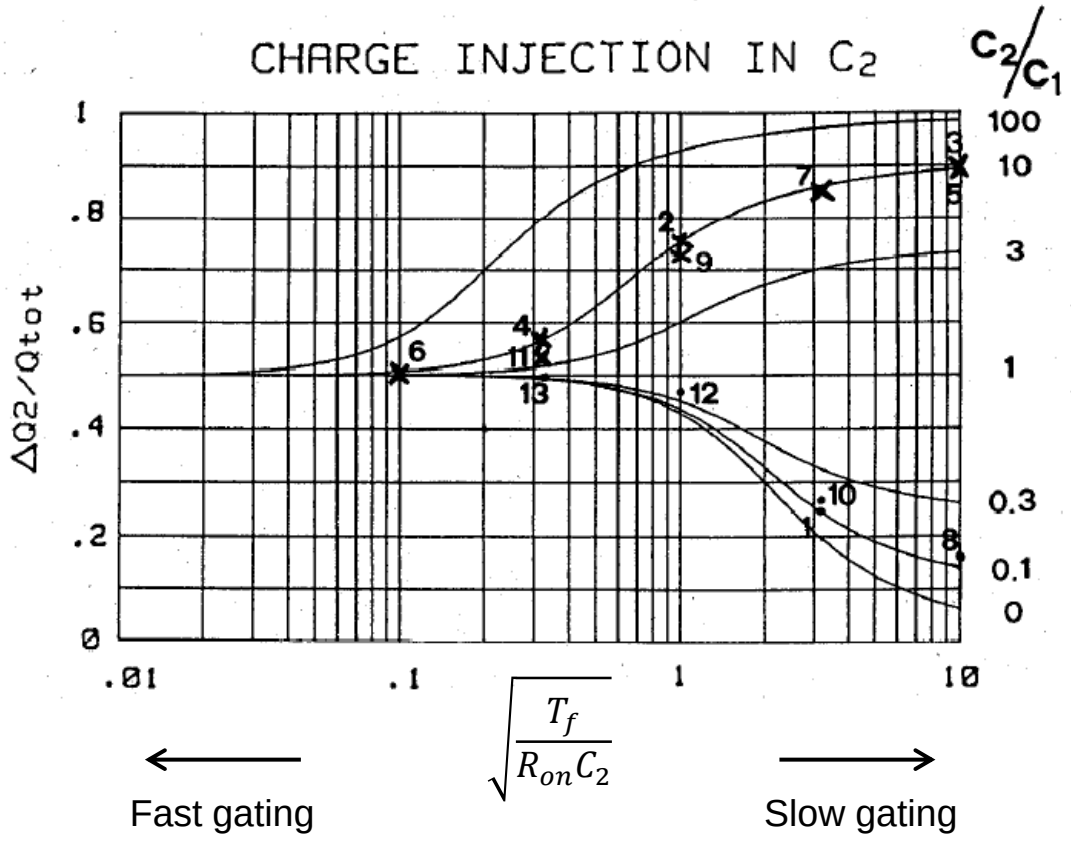
Fast Gating



- Assuming “infinitely” fast gating, there is no time for the transistor’s on-resistance to play a role in the transient
- With $V_{GS,D} < V_t$ at the source and drain side of the channel, pinch-off occurs at <u>both</u> sides
- Roughly speaking, we see a current source looking into both sides of the MOSFET and the charge splits 50/50 (independent of termination impedance)

Charge Split Ratio Data





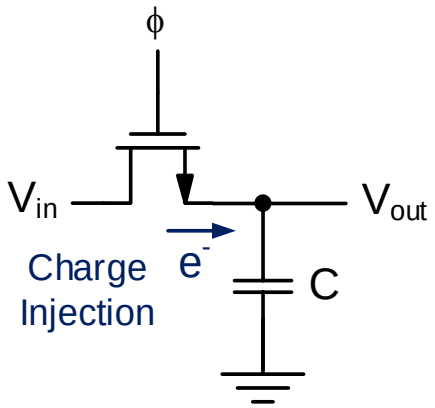
G. Wegmann et al., "Charge injection in analog MOS switches," JSSC, June 1987
Y. Ding and R. Harjani, "A universal analytic charge injection model," Proc. ISCAS, May 2000

What Happens in Practical Circuits?

- We tend to drive the switch gate as fast as the technology allows
- $R_{on}C_2$ and T_f are thus usually comparable (same order of magnitude)
  - This brings us into the x-axis range <1 on the chart by Wegmann
  - Means that the charge split may in practice still have “some” dependence on the termination impedances

- Trying to predict the charge split accurately is somewhat hopeless
- We typically use the fast-gating approximation (50/50 split) to get an estimate of the injected charge
  - But we know that slightly more charge will go to the side with lower impedance (perhaps 60…80%)

Fast Gating Estimate of Charge Injection Error



$$V_{out} = V_{in} + \frac{1}{2}\frac{Q_{ch}}{C} \qquad Q_{ch} \approx -WLC_{ox}(V_{DD} - V_{in} - V_t)$$

$$V_{out} = V_{in} - \frac{1}{2}\frac{WLC_{ox}}{C}(V_{DD} - V_{in} - V_t)$$

$$V_{out} = V_{in}(1 + \epsilon) + V_{os}$$

$$\epsilon = \frac{1}{2}\frac{WLC_{ox}}{C} \qquad V_{os} = \frac{1}{2}\frac{WLC_{ox}}{C}(V_{DD} - V_t)$$

Gain error **Offset error**

Example: $\epsilon = \frac{1}{2}\frac{20\mu m \cdot 2fF/\mu m}{1pF} = +2\%$ $V_{os} = 2\%(1.8V - 0.5V) = -26mV$

Impact of Technology Scaling

$$\Delta V = \frac{1}{2}\frac{Q_{ch}}{C} \qquad Q_{ch} \approx -WLC_{ox}(V_{GS} - V_t)$$

$$R = \frac{1}{\mu C_{ox}\frac{W}{L}(V_{GS} - V_t)} = \frac{L^2}{\mu Q_{ch}}$$

$$NRC = \frac{T_s}{2} = \frac{1}{2f_s}$$

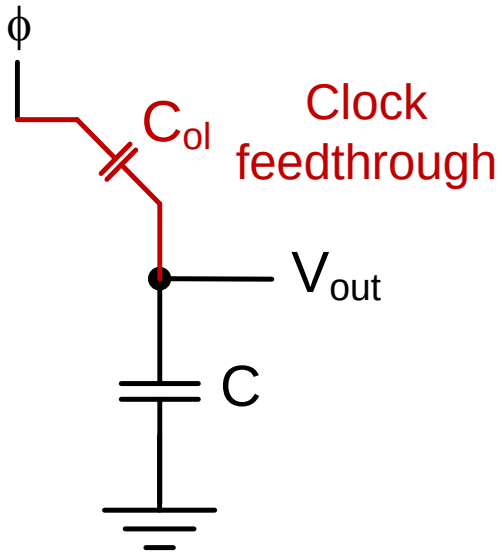
$$\boxed{\frac{\Delta V}{N f_s} \approx \frac{L^2}{\mu}}$$

- Ratio of charge injection error and speed benefits from shorter channels and increases in mobility (e.g., due to strain)

Note On Gain Error

- Our simplistic analysis assumed that the channel charge is linearly related to V_{in} and hence we see a linear gain error
- This is of course true only to first order (consider backgate effect, etc.)
  - There will always be nonlinear errors in the injected charge
- Computing these nonlinear error terms is analytically difficult and tedious
  - Typically need to rely on simulations to get a feel
- Simulation that the nonlinear portion of the injected charge can be on the order of 0.05…0.1%
- As a rule of thumb, one can assume that signal-dependent charge injection errors limit the sampling linearity to ~10-11 bits

Clock Feedthrough (Assuming Fast Gating)



$$V_{out} = V_{in} - V_{DD}\frac{C_{ol}}{C_{ol}+C} = V_{in} + V_{os}$$

$$V_{os} = -V_{DD}\frac{C_{ol}}{C_{ol}+C}$$

Offset error

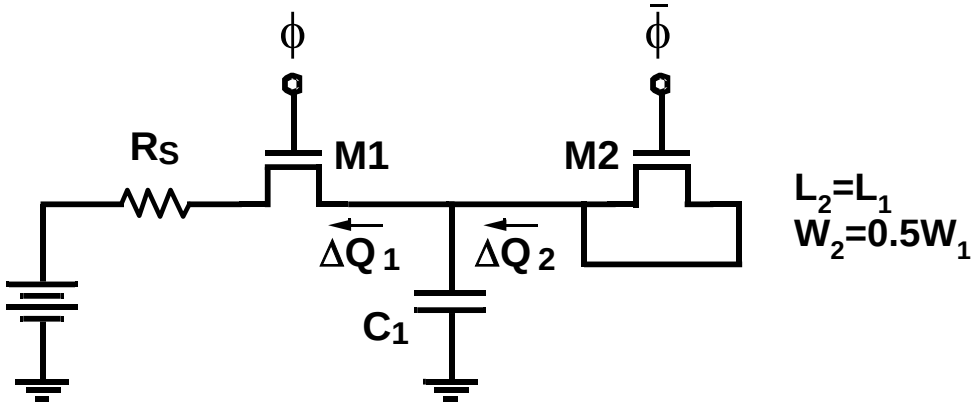
Example: $C_{ol} = 20\mu m \cdot 0.2\frac{fF}{\mu m} = 4fF$ $V_{os} = -1.8V\frac{4fF}{4fF+1pF} = -7.1mV$

- The total offset error is found by adding the charge injection and clock feedthrough components

Combatting Charge Injection Errors

- Bootstrapped switch
  - Keeps V_{GS} relatively constant, but channel charge is still signal dependent due to backgate effect, etc.
- Charge cancelation
  - Inject a charge packet with opposite sign
- Transmission gate
  - Achieves partial cancellation between NMOS and PMOS
- Bottom plate sampling
  - The ultimate solution for mitigating signal dependent charge injection

Charge Cancellation

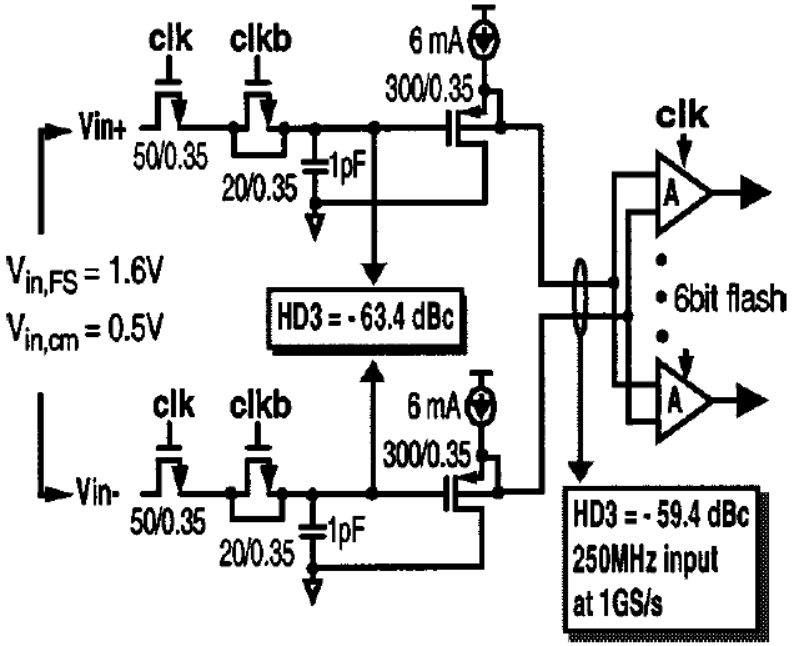


Suarez, Gray and Hodges, JSSC 12/75

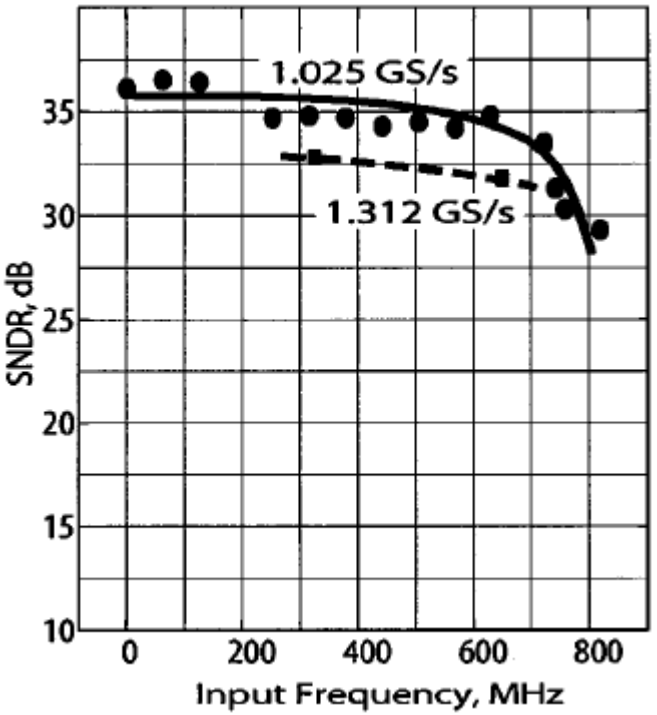
Eichenberger and Guggenbűhl, JSSC 8/89

- Cancellation is not perfect, since the channel charge of M1 will not exactly split 50/50 (a larger fraction tends to go the low-impedance input)
- Not a precision technique, just an attempt to do a partial clean-up
  - For example, mitigate a large common mode bounce

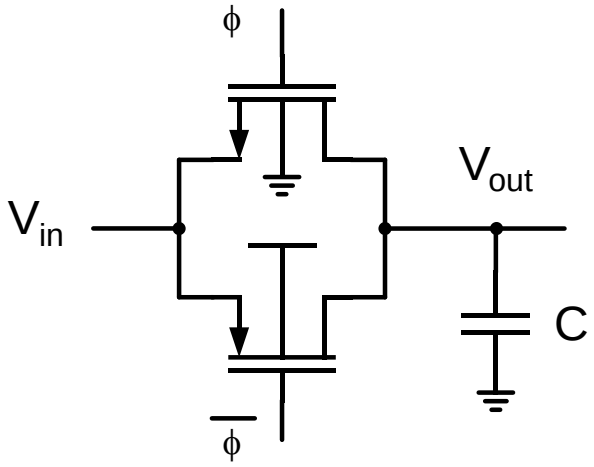
Application Example



Choi and Abidi, JSSC 12/2001



Transmission Gate



$$Q_{chn} \approx -W_n L_n C_{ox}(V_{DD} - V_{in} - V_{tn})$$

$$Q_{chp} \approx +W_p L_p C_{ox}(V_{in} - |V_{tp}|)$$

- Assuming $W_n L_n = W_p L_p$

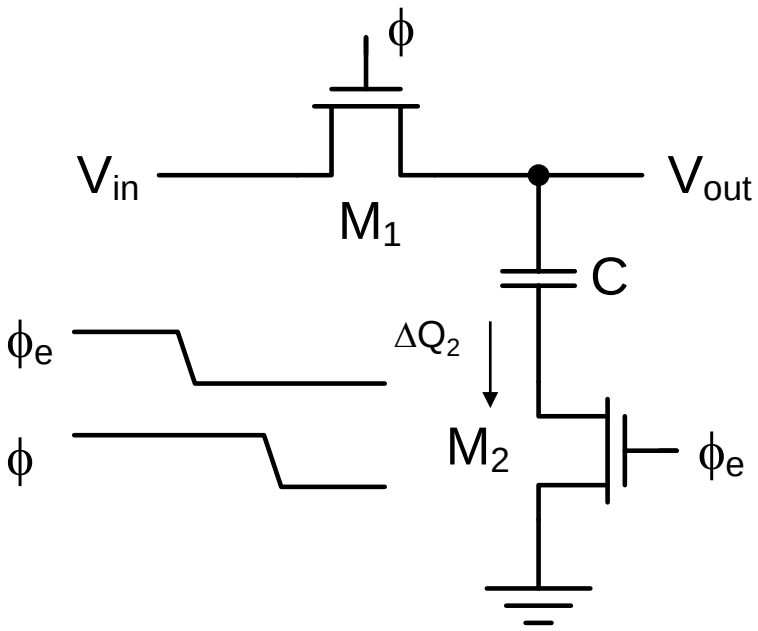
$$\Delta V_{out} \approx \frac{1}{2}\frac{Q_{chn}}{C} + \frac{1}{2}\frac{Q_{chp}}{C} = \frac{W_n L_n C_{ox}}{C}\left(V_{in} - \frac{V_{DD}}{2} + \frac{V_{tn} - |V_{tp}|}{2}\right)$$

- Charges fully cancel e.g. for $V_{in} = V_{DD}/2$, and $V_{tn} = |V_{tp}|$, but there is still signal-dependent residual injection

Bottom Plate Sampling

- How to overcome nonlinear charge injection for high-fidelity designs?
  - For example, DR ~100 dB
- Basic idea
  - Sample signal at the "grounded" side of the capacitor to achieve signal-independent sampling
- Original reference
  - D. G. Haigh and B. Singh, “A switching scheme for SC filters which reduces the effect of parasitic capacitances associated with switch control terminals,” Proc. ISCAS, 1983, pp. 586–589

Bottom Plate Sampling Analysis (1)



- Turn M_2 off slightly before M_1
  - Typically, a few hundred ps delay between falling edges of ϕ_e and ϕ
- During turn off, M_2 injects charge

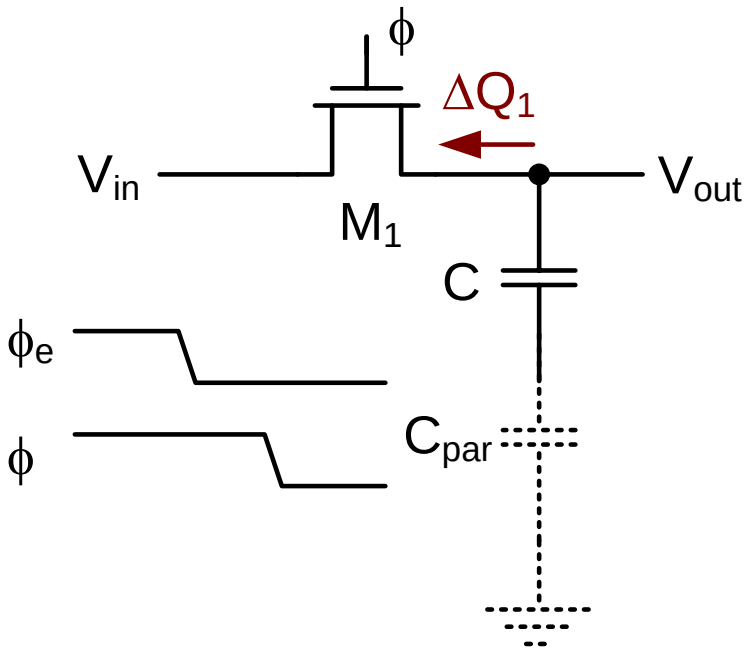
$$\Delta Q_2 \approx \frac{1}{2} WLC_{ox}(V_{DD} - V_t)$$

- To first order, the charge injected by M_2 is signal independent
- Voltage across C

$$V_C = V_{in} + \frac{\Delta Q_2}{C}$$

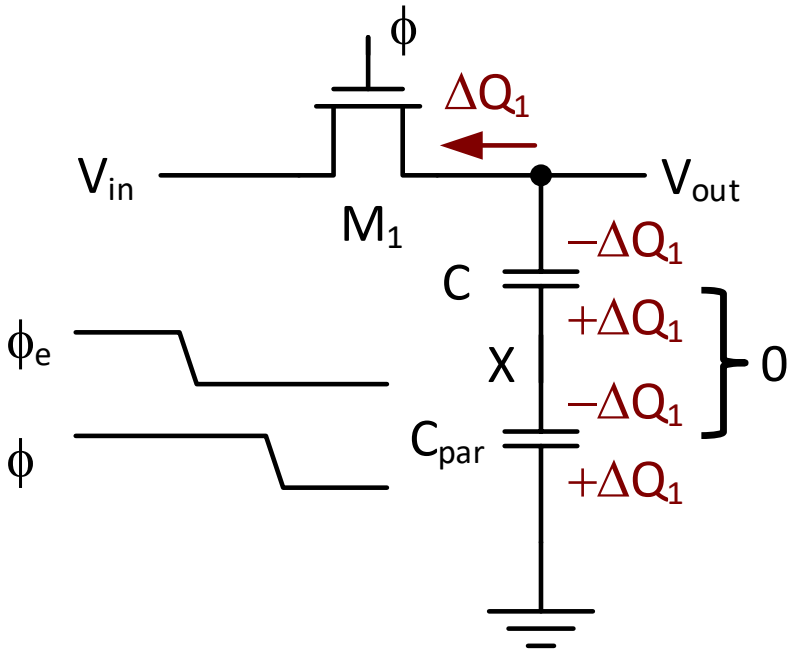
Bottom Plate Sampling Analysis (2)

$$\Delta Q_1 \approx \frac{1}{2} WLC_{ox}(V_{DD} - V_{in} - V_t)$$



- Next, turn off M1
- M1 injects signal-dependent charge into series combination of C and bottom plate parasitic capacitance (C_{par})
- Voltage across C contaminated with error due to signal dependent charge injection
- Looks like, this is not much different from conventional sampling?
  - But wait…

Bottom Plate Sampling Analysis (3)

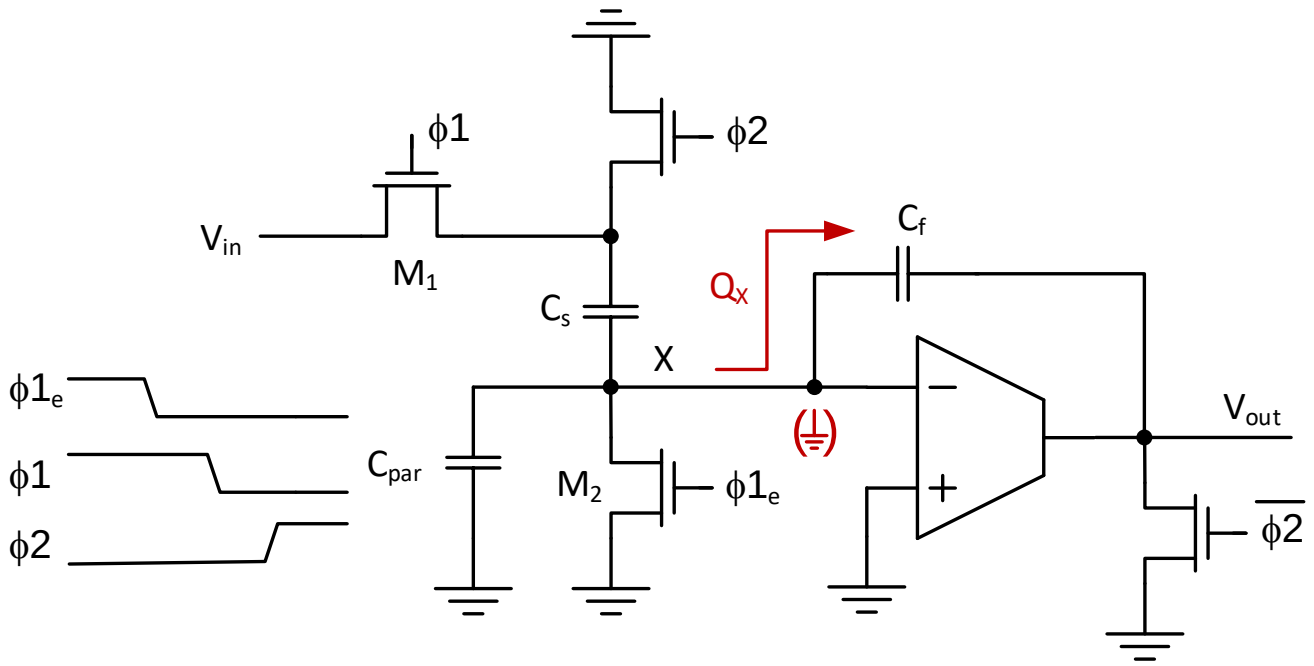


$$Q_X = -CV_{in} - \underbrace{\Delta Q_2}$$

Charge injected by M_2
(signal independent)

- Interesting observation
  - Even though M1 injects charge, total charge at node X cannot change!
  - Gaussian surface around node X
- Idea
  - Process total charge at node X instead of looking at voltage across C

Charge Redistribution



- Amplifier forces voltage at node X to “zero”
  - Charge at node X must redistribute onto feedback capacitor C_f

Charge Conservation Analysis

Charge at node X during ϕ_1: $Q_{X1} = -C_s V_{in} - \Delta Q_2 + 0 \cdot C_f$

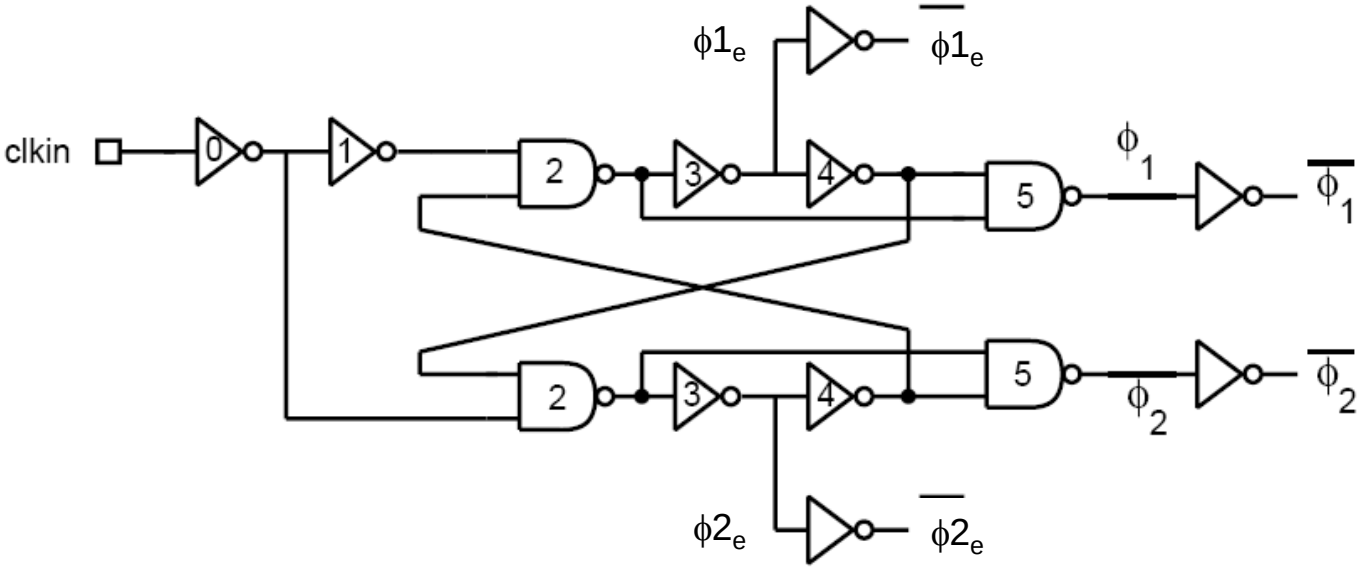
Charge at node X during ϕ_2: $Q_{X2} = -C_f V_{out}$

Charge Conservation: $Q_{X1} = Q_{x2}$

$$\boxed{V_{out} = \frac{C_s}{C_f} V_{in} + \frac{\Delta Q_2}{C_f}}$$

- Key point: ΔQ_2 is signal independent, just a constant offset error
- The offset can be removed using a differential architecture

Clock Generation Example



[A. Abo, "Design for Reliability of Low-voltage, Switched-capacitor Circuits," PhD Thesis, UC Berkeley, 1999]