CMOS RF: No Longer an Oxymoron

Prof. Thomas H. Lee

Stanford University Center for Integrated Systems

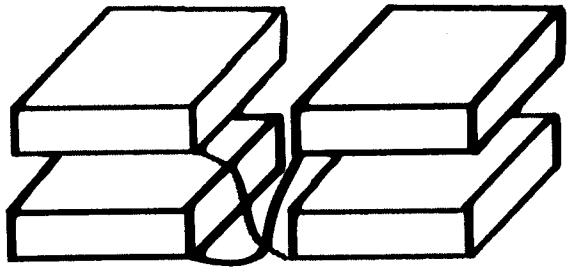
Introduction

- Physics of propagation guarantee little terrestrial wireless activity much above ~5 GHz.
- Continued improvements in Si technology have made GHz circuits feasible, even in CMOS.
- Bad News: Traditional RF designs have large ratio of passive to active devices.
  - Infamous planar spiral inductors have poor Q, and don't scale in size as transistors do.
  - Traditional parallel-plate capacitors (needed for linearity) do not scale either.

Lateral Flux Capacitors

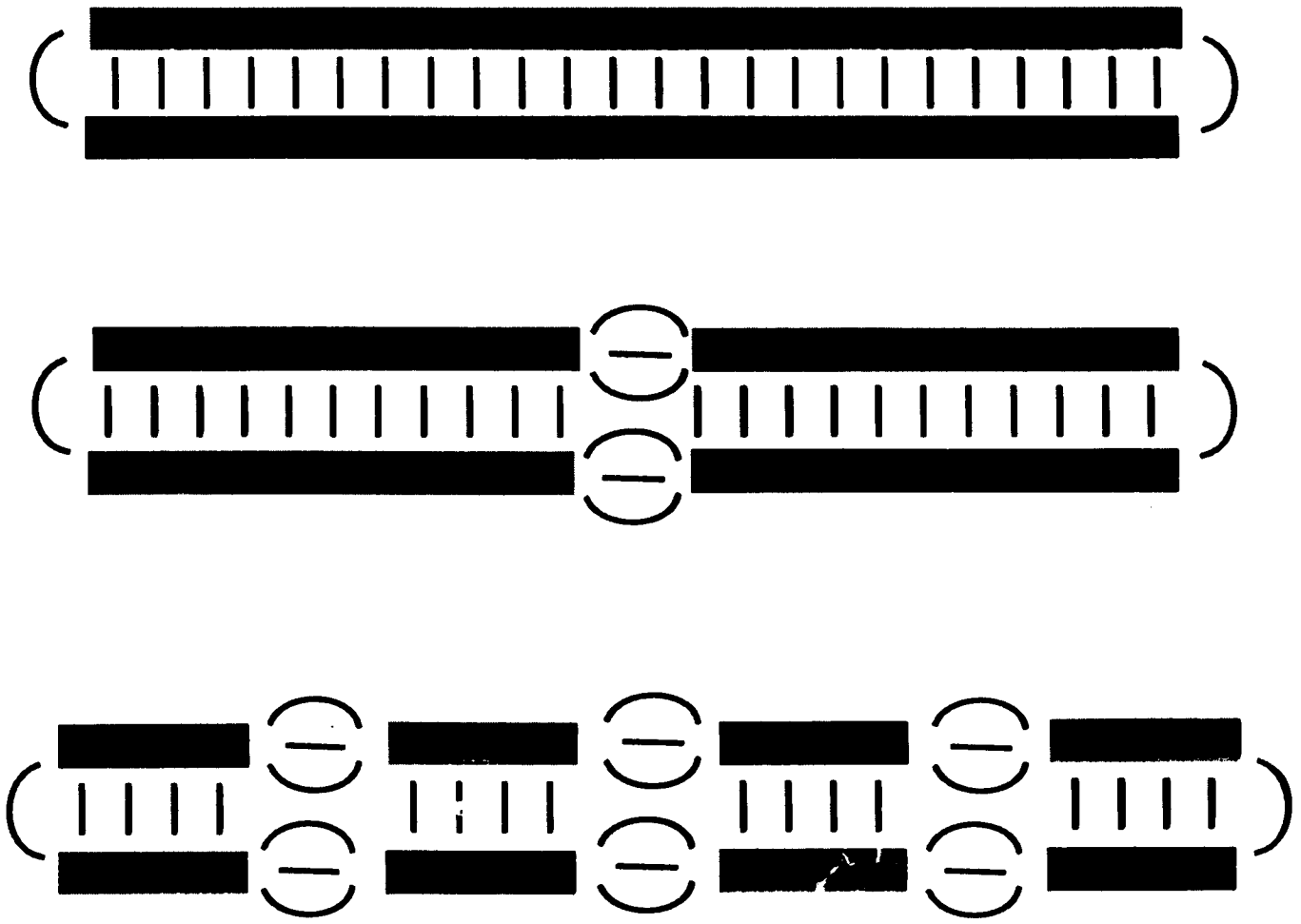
- Linear, dense capacitors are available in deep submicron technologies if lateral flux is used to augment vertical flux
  - Structure is MIM capacitor: no special processing required
  - 3-10x boost in capacitance per unit area
  - Parasitic bottom plate capacitance reduced by same factor
  - Substrate loss and noise coupling attenuated by large factors
  - Pretty die photos

Layout Issues

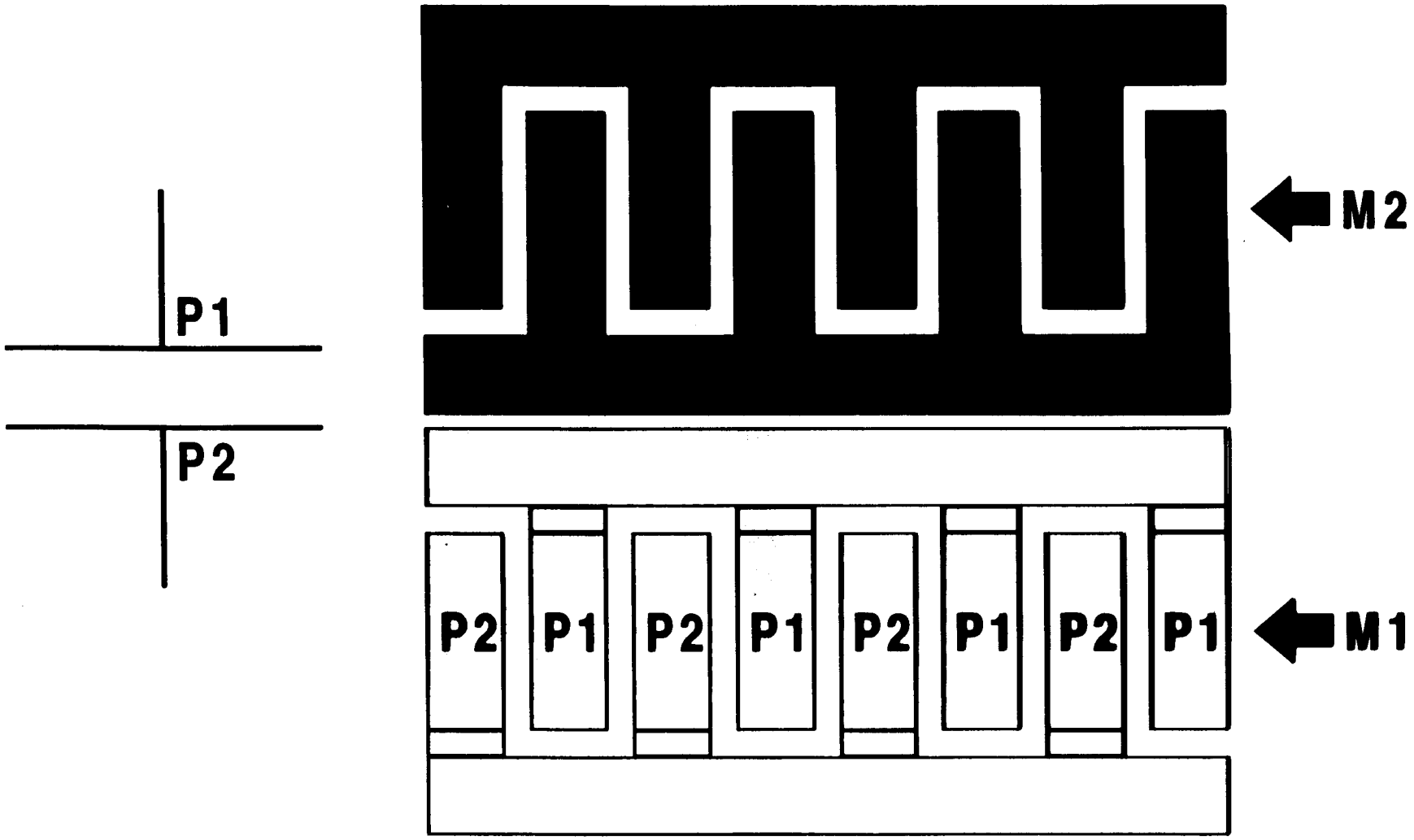


- Cross connected
- Control over the final shape
  - -Correct choice of the initiator and the generator
- Capacitance per unit area vs. series resistance
  - Fractal dimension
  - Average width of metal

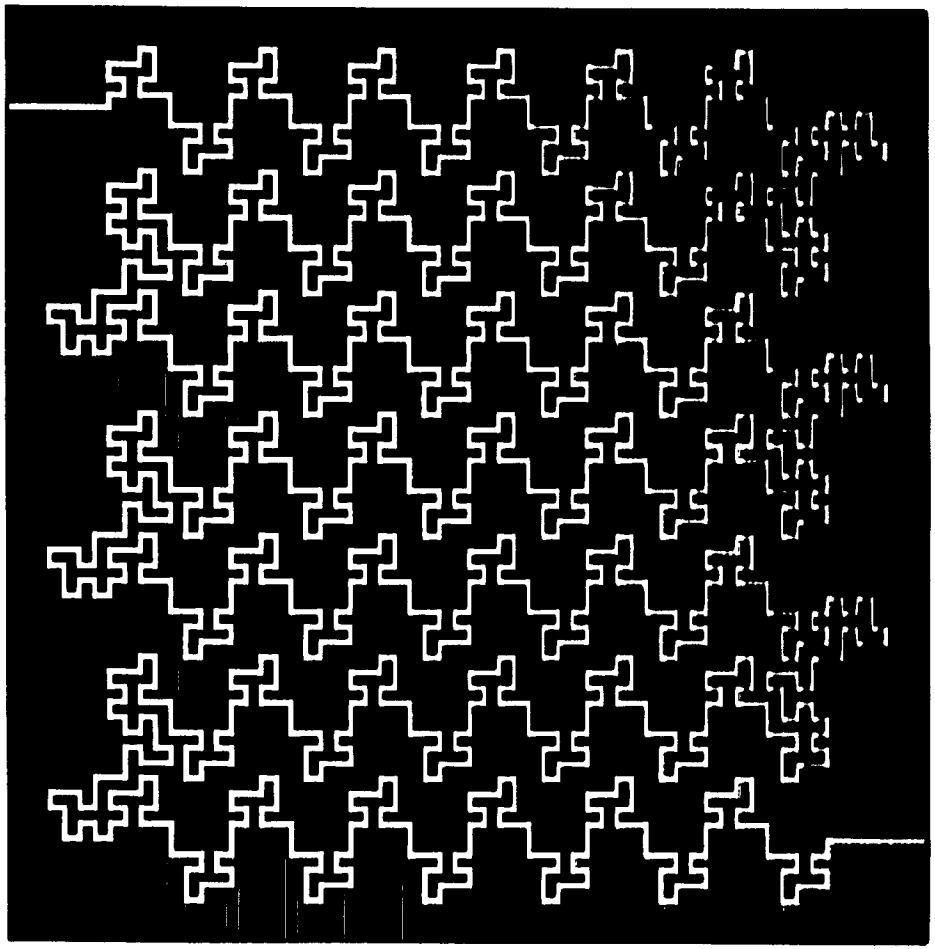
Vertical Flux vs. Lateral Flux



Lateral-Flux Capacitor (LFC)



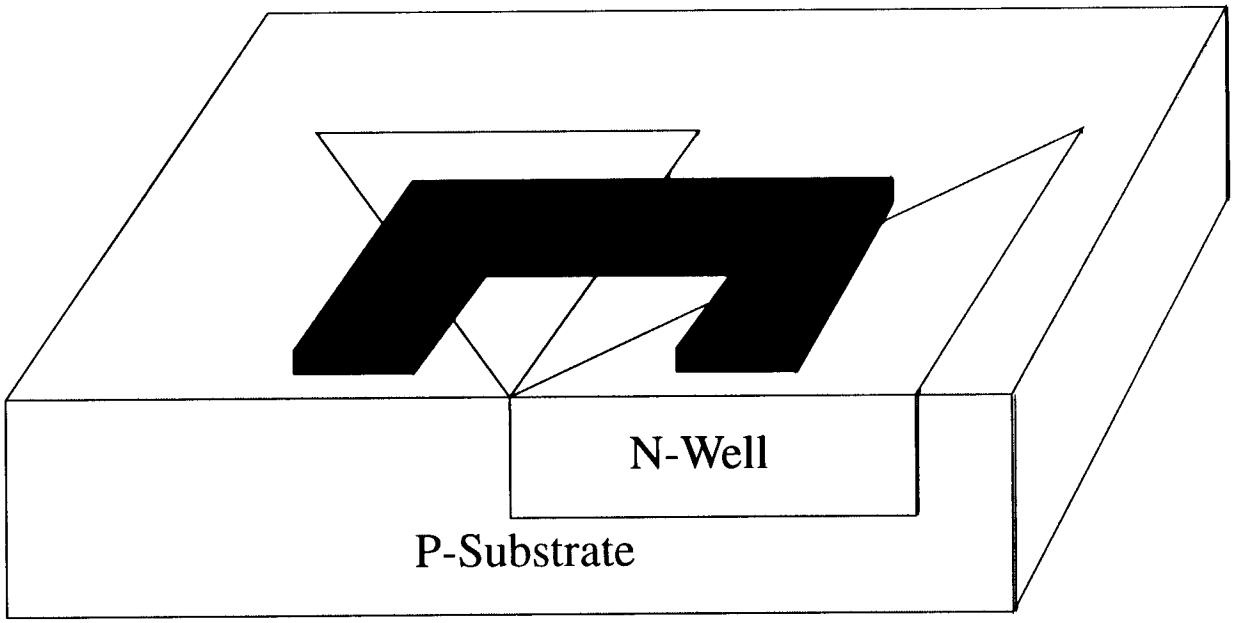
Fractal Capacitor



Improved Inductors in Si Technology

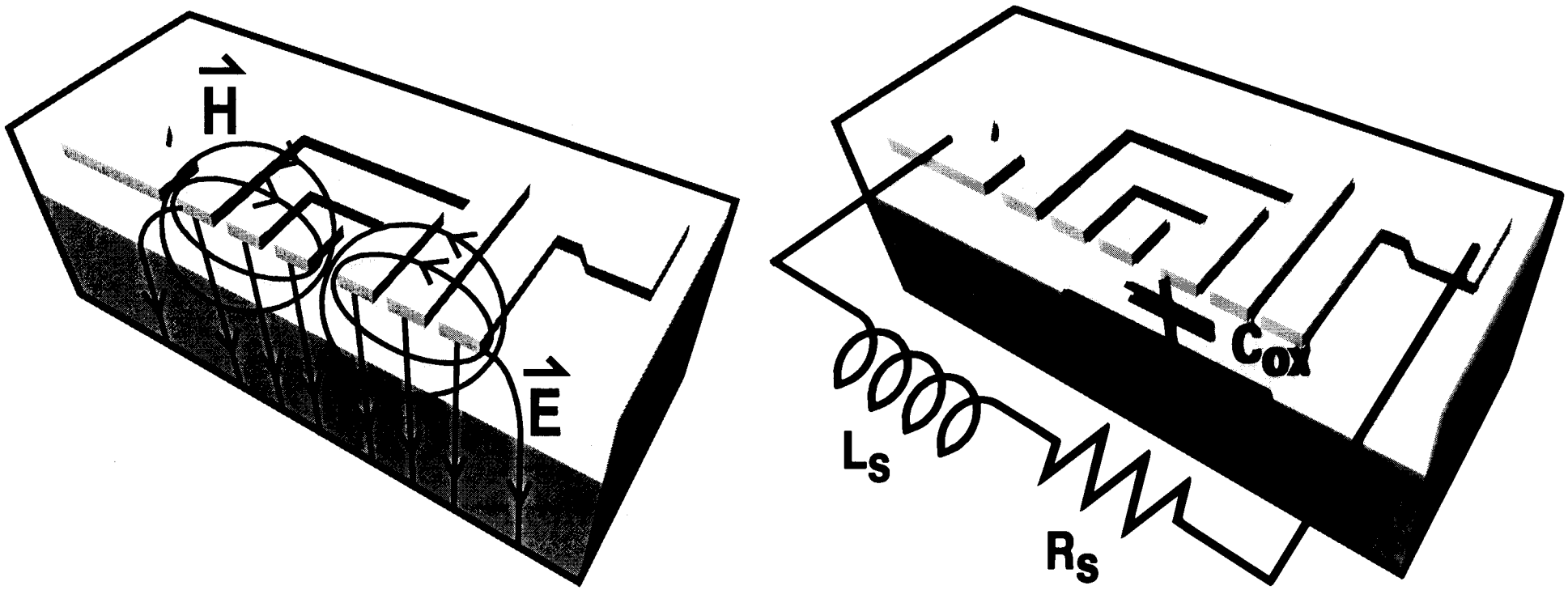
- Scaling problem is fundamentally intractable, unfortunately, because its solution requires 3D structures
  - Increasing number of interconnect layers helps, however.
- Can significantly improve inductor Q, without departing from ordinary technology.
  - Skin effect/proximity effect conductor losses are ~fixed.
  - Eddy (image) currents magnetically induced in substrate are relatively unimportant. Can build inductor over alternating well/substrate wedges to reduce this loss further in any case.
  - Displacement currents flowing into lossy substrate through parasitic capacitance between inductor and substrate can be suppressed through the use of a *patterned ground shield* (Yue and Wong, 1997), nearly doubling the Q of resonators.

Improved Inductors

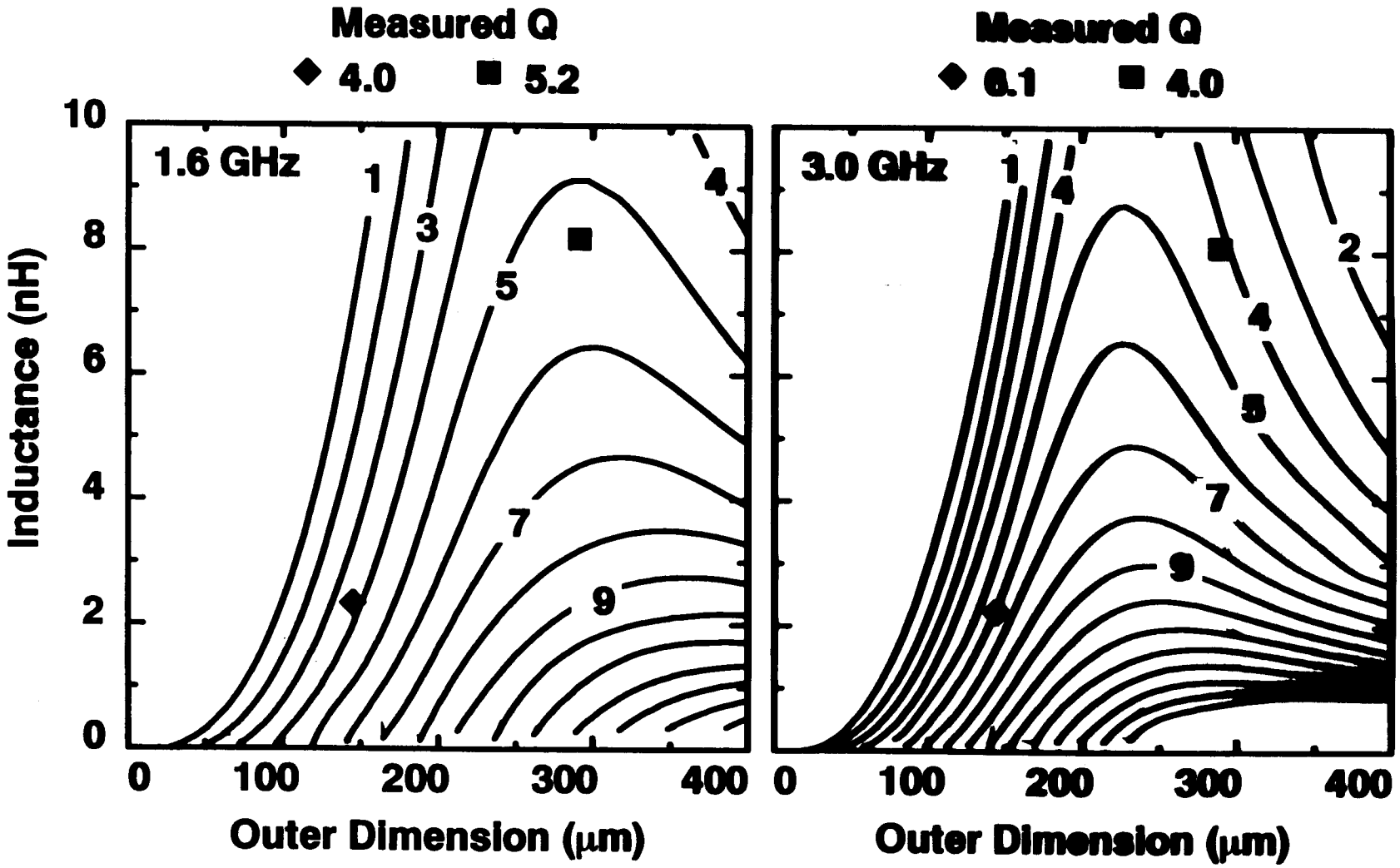


- Alternating wedges of N-well and substrate underneath the inductor force image currents to flow farther away.
  - Circulating currents to depth of well are inhibited by back-biased PN junctions.

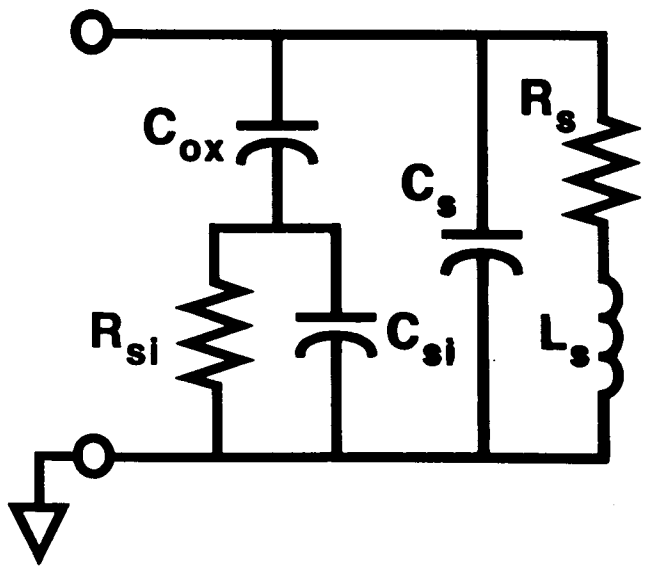
Electromagnetic Fields of Conventional On-Chip Inductors



Contour Plots of Q

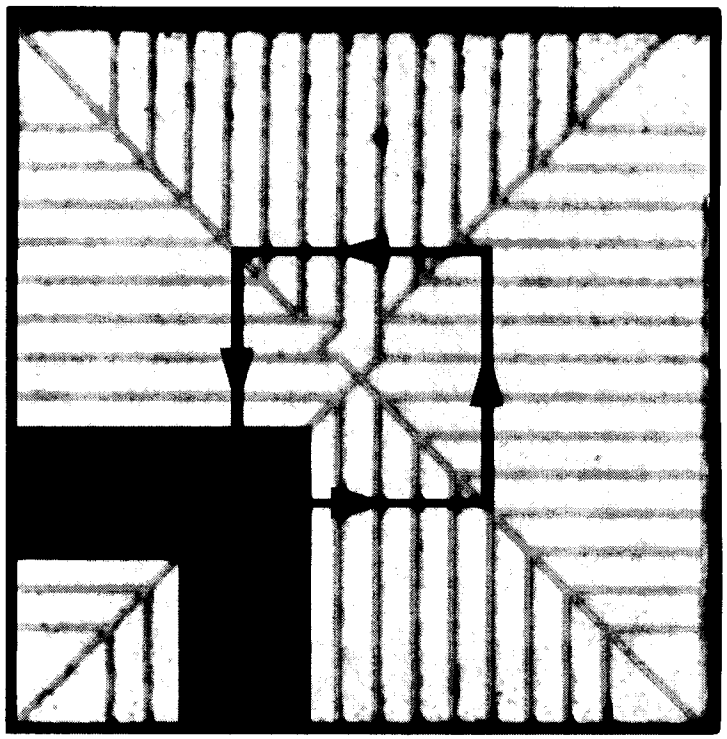


Model Description

Physical Model of an Inductor on Silicon		Physical Phenomena
	L_s: Greenhouse Method	Mutual Coupling
	$R_s = \frac{\rho \cdot l}{w \cdot \delta \cdot (1 - e^{-t/\delta})}$	Eddy Current
	$C_s = n \cdot w^2 \cdot \frac{\varepsilon_{ox}}{t_{ox\ M1\text{-}M2}}$	Feed-Through Capacitance
	$C_{ox} = \frac{1}{2} \cdot l \cdot w \cdot \frac{\varepsilon_{ox}}{t_{ox}}$	Oxide Capacitance
	$C_{si} = \frac{1}{2} \cdot l \cdot w \cdot C_{Sub}$	Si Substrate Capacitance
	$R_{si} = \frac{2}{l \cdot w \cdot G_{Sub}}$	Si Substrate Ohmic Loss

Patterned Ground Shield Design

- Pattern
  - Orthogonal to spiral (induced loop current)
- Resistance
  - Low for termination of the electric field
  - Avoid attenuation of the magnetic field

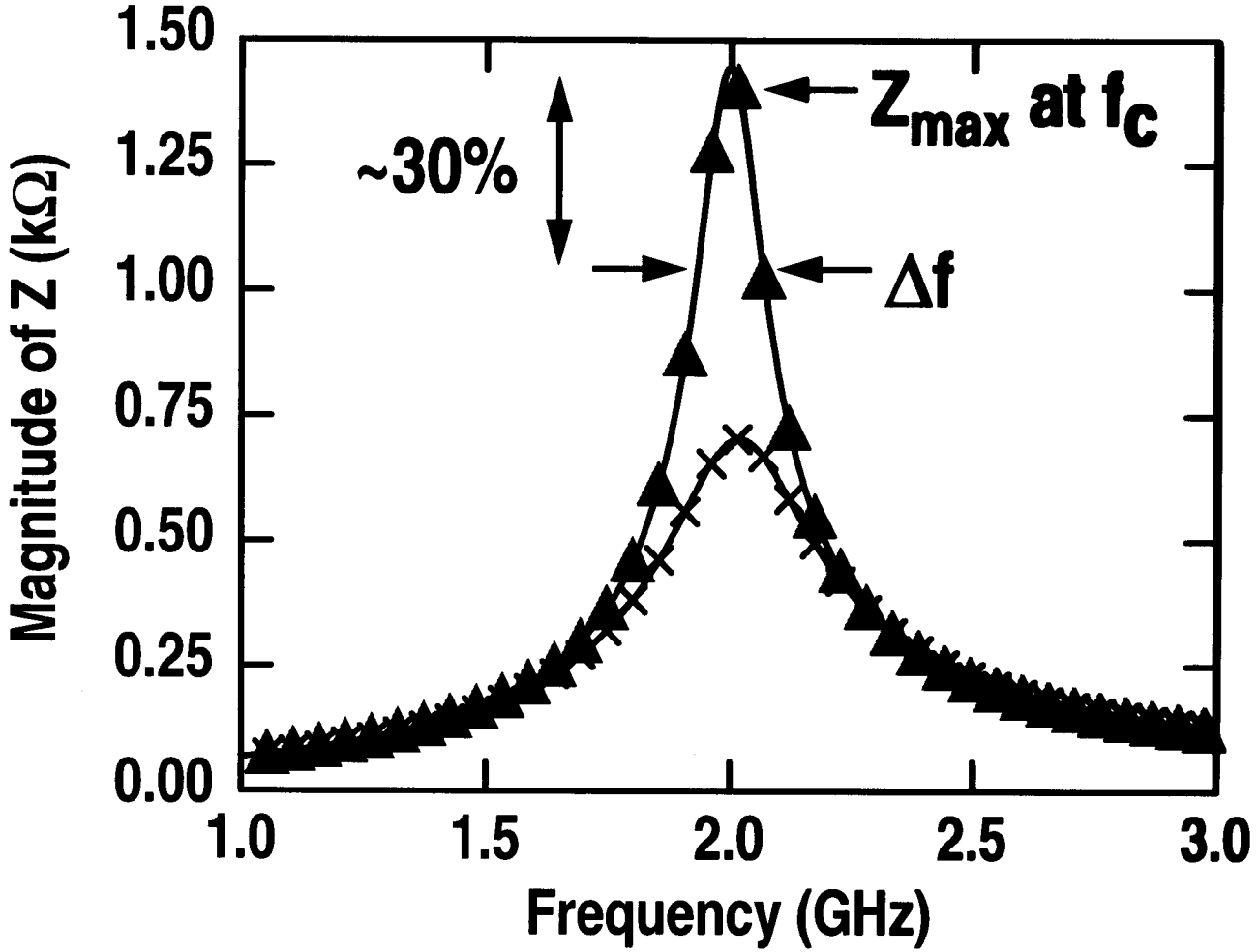


Ground Strips

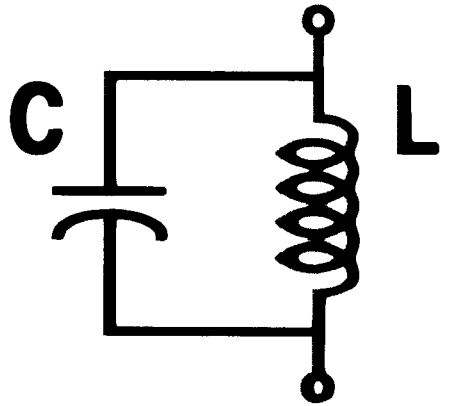
Slot between Strips

Induced Loop Current

Parallel LC Resonator at 2 GHz



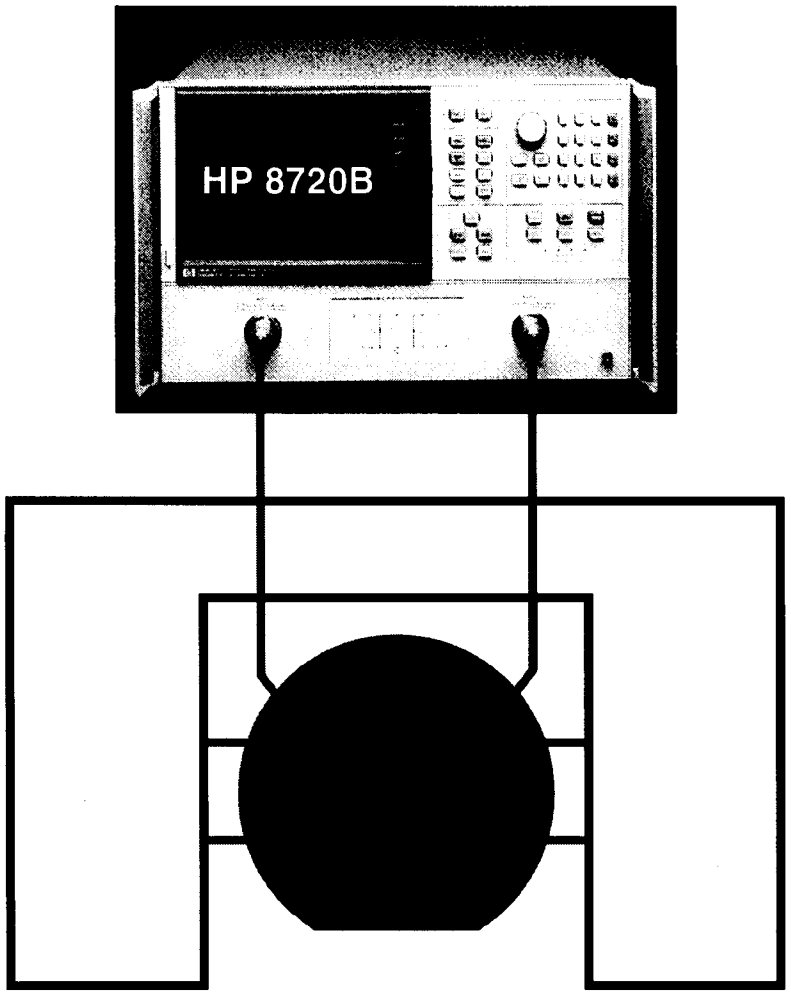
$$Q_{RESONATOR} = \frac{f_c}{\Delta f}$$



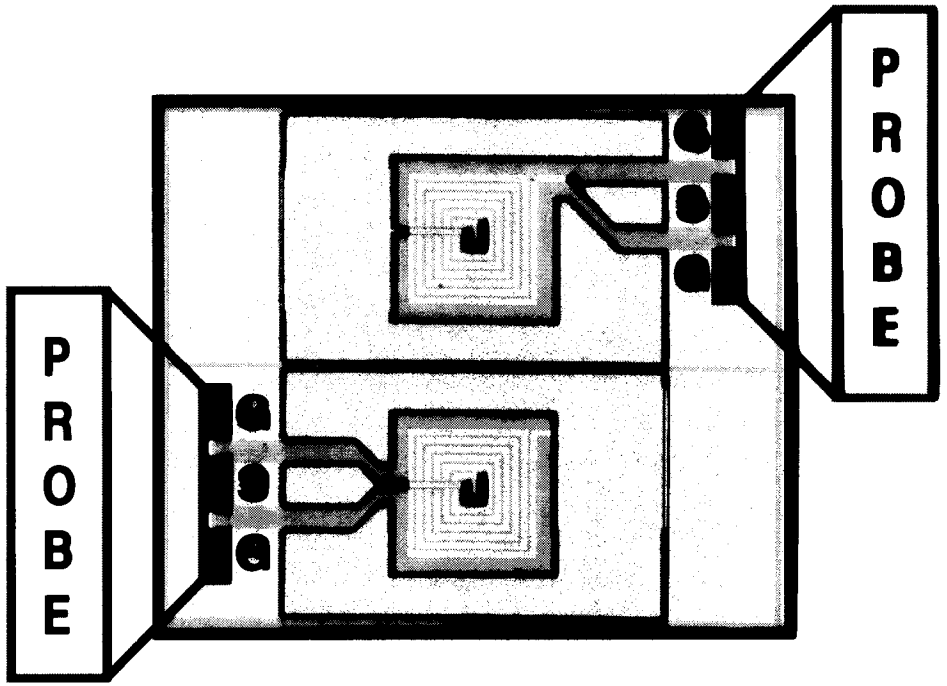
Patterned Polysilicon

None (11 Ω-cm)

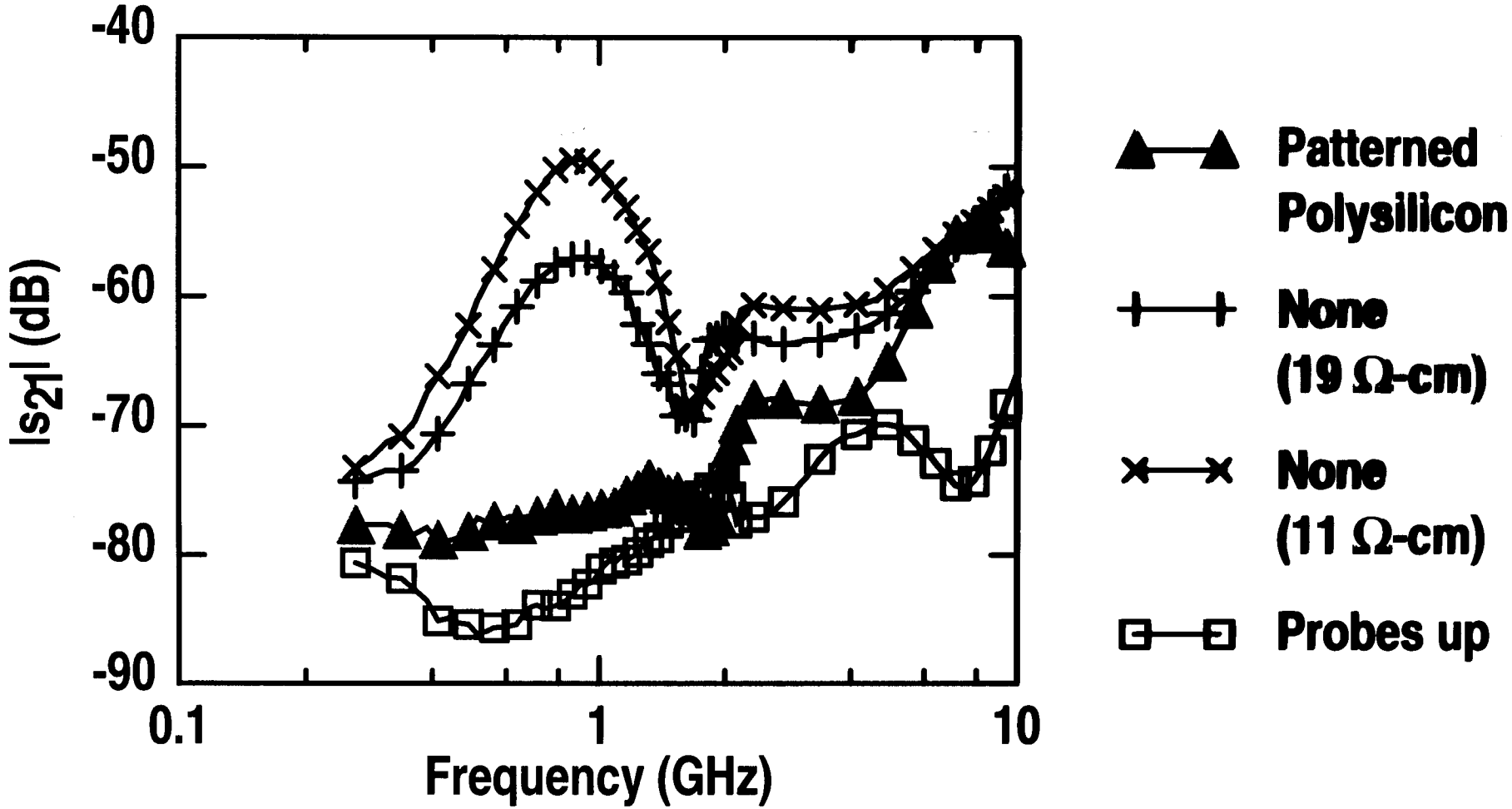
Noise Coupling Measurement



Probe Station



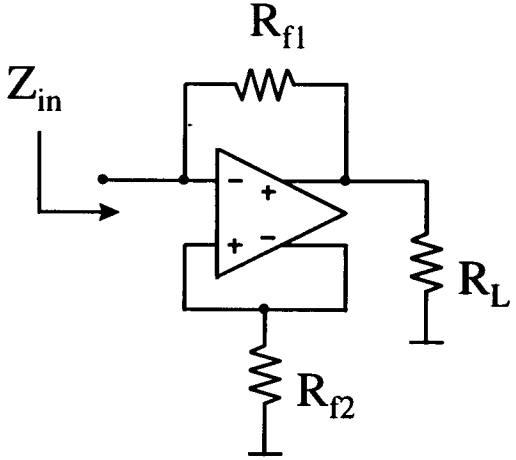
Effect of Polysilicon Ground Shield on Noise Coupling



Low-Noise Amplifiers

- Power- (actually, current-) constrained noise optimization method is key breakthrough.
  - For a specified allowed power budget, method yields best LNA noise figure subject to constraint of perfect input match.
  - Method yields optimum device size, unlike any other noise theory. Value is about 750μm-GHz.
- Noise figures for single-ended LNAs are approximately 1.5dB at 1.6GHz in 0.35μm processes on ~10mW budgets.
  - Improves with scaling.
  - Differential LNAs recommended for on-chip environments to improve noise rejection and ease packaging constraints.
    - Tradeoff is doubling of power for same noise figure as single-ended circuit.

How To Get 50Ω

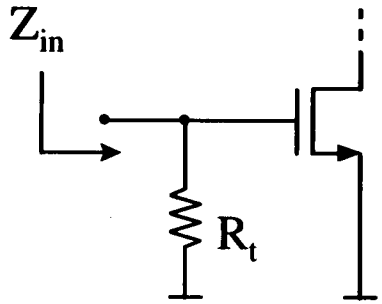


Dual Feedback

$$Z_{in} = \sqrt{R_{f1}R_{f2}}$$

Need high gain.

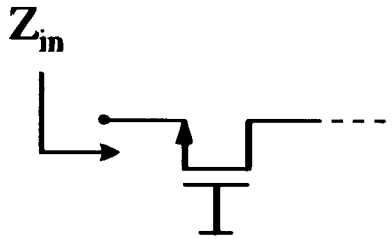
Stability problems.



Resistive Termination

$$Z_{in} = R_t$$

Poor NF.

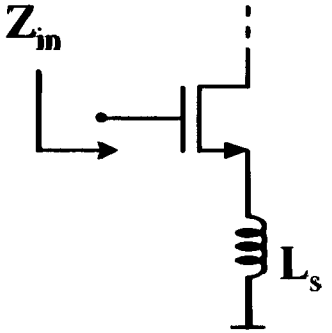


$1/g_m$ Termination

$$Z_{in} = \frac{1}{g_m}$$

NF > 3dB

($\gamma > 1$)



Inductive Degeneration

$$\mathbf{Re}[Z_{in}] = \frac{g_m}{C_{gs}} L_s$$

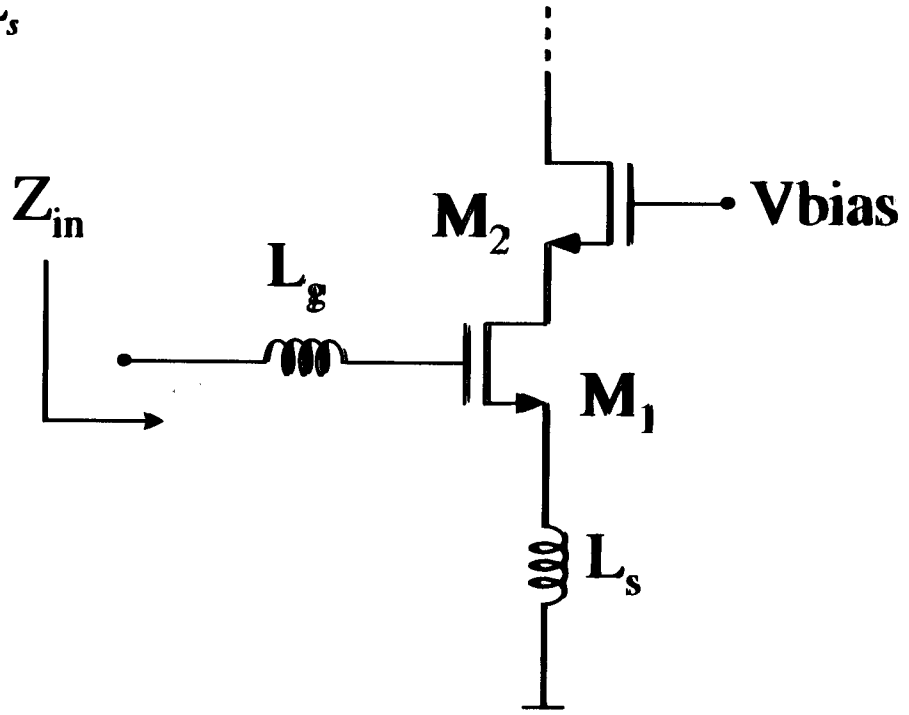
Narrowband.

LNA Input Stage

$$Z_{in} = s\left(L_s + L_g\right) + \frac{1}{sC_{gs}} + \left(\frac{g_m}{C_{gs}}\right)L_s \approx \omega_T L_s$$

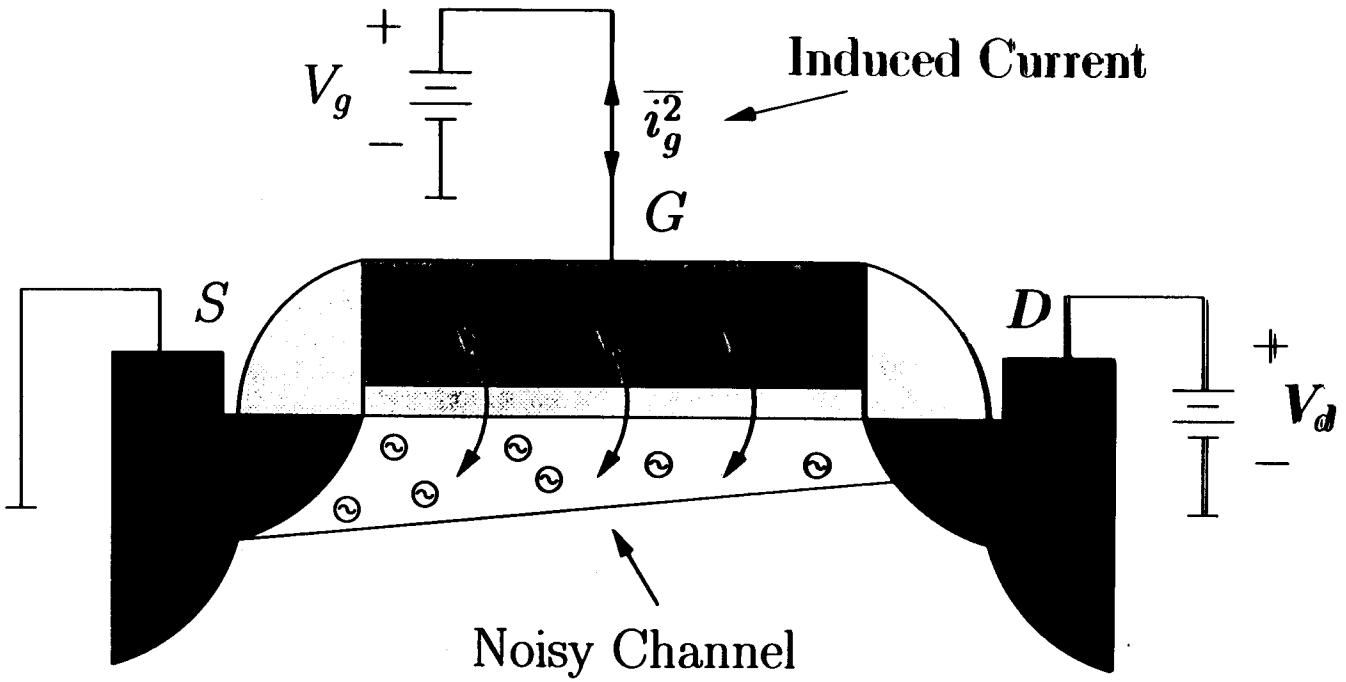
$$G_{m,eff} = g_{m1}Q_{in} = \frac{g_{m1}}{\omega C_{gs}\left(R_s + \omega_T L_s\right)}$$

$$= \frac{\omega_T}{\omega R_s\left(1 + \frac{\omega_T L_s}{R_s}\right)} = \frac{\omega_T}{2\omega R_s}$$

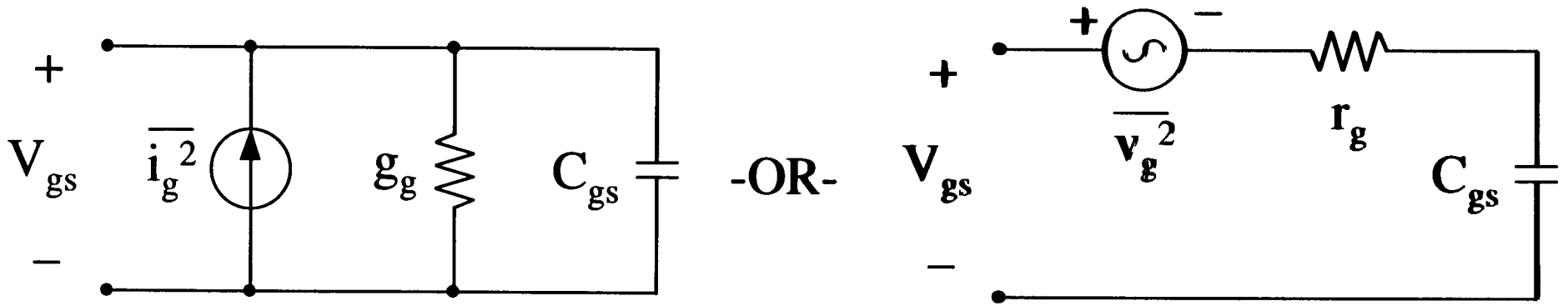


Note: $G_{m,eff}$ is <u>independent</u> of g_{m1}!

Induced Gate Effects



Equivalent Gate Circuit



$$\overline{i_g^2} = 4kTB\delta g_g \qquad g_g = \frac{1}{5}\frac{\omega^2 C_{gs}^2}{g_{d0}}$$

"Blue" Noise

$$\overline{v_g^2} = 4kTB\delta r_g \qquad r_g = \frac{1}{5g_{d0}}$$

"White" Noise

- δ (~ 4/3) modified by hot electron effects
- $\overline{i_g^2}$ partially correlated with $\overline{i_d^2}$ (c = 0.395j)
- $\overline{i_g^2}$ and g_g not modeled in HSPICE

Epi Noise

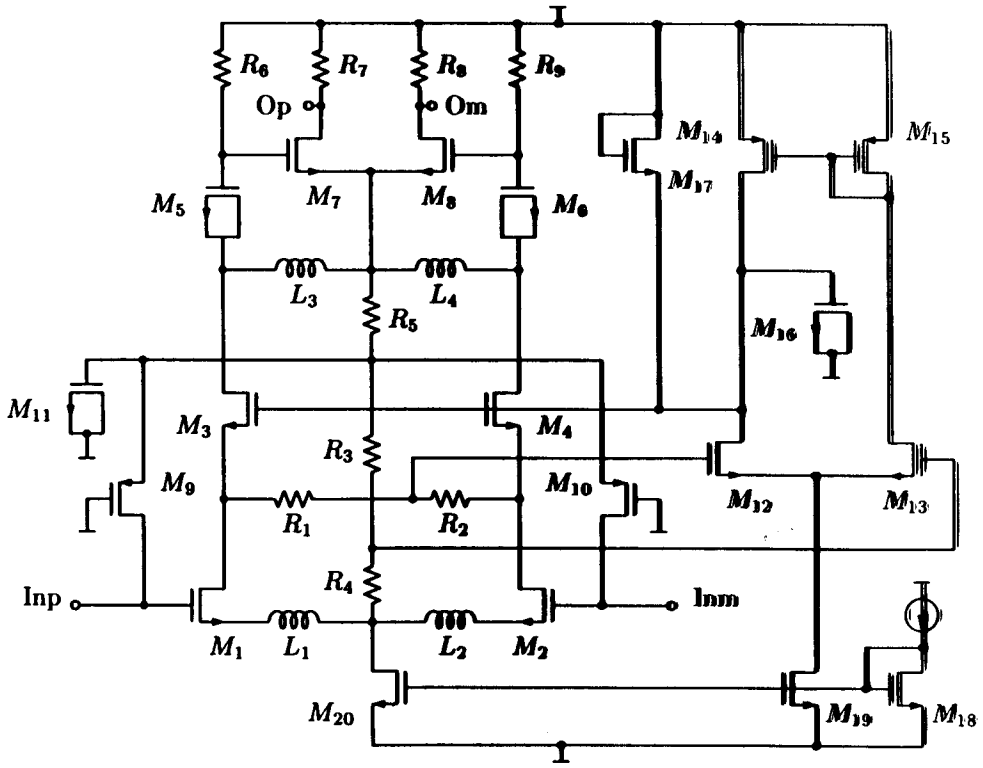
- Additional broadband noise source is the modulation of the back gate by the substrate resistance's thermal noise
- A more rigorous analysis reveals that this additional source may be neglected as long as the effective resistance from the back gate to an incremental ground is less than about $2R_S$.
  - Generally trivially satisfied in bulk technologies
  - In stubborn cases, some epi processes may require the use of subdivided input devices with liberal substrate contacts distributed throughout

LNA Design Procedure

- Select device width roughly equal to (750μm-GHz)/f_0 (for a 50Ω system)
- Adjust bias to obtain desired power dissipation
  - Keep V_{DS}–V_{DSAT} as small as practical to minimize hot-electron effects (say, under half a volt or so)
- Select source degeneration inductance (assuming equal-sized cascoding and main devices) according to:

$$L_S \approx \frac{R_S \cdot [1 + 2(C_{gd}/C_{gs})]}{\omega_T}$$

- Add enough gate inductance to bring input to resonance
- Noise factor bound is $1 + 2.4(\gamma/\alpha)(\omega/\omega_T)$, so scaling continues to help directly

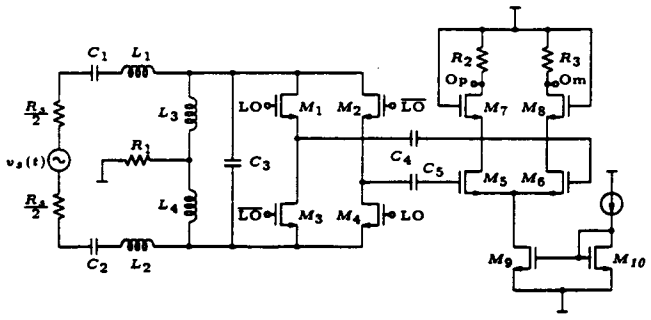

I
R6
R7
R8
R9
Op
Om
M14
M15
M17
M5
M7
M8
M6
L3
L4
R5
M16
M11
M3
M4
R3
M9
M10
M12
M13
R1
R2
Inp
R4
Inm
M1
L1
L2
M2
M20
M19
M18
I

CMOS LNAs

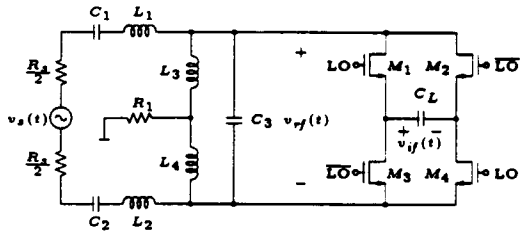
- Scaling continues to help. For single-ended LNAs, one can expect the following:
  - < 2 dB NF at 5 mW achievable in 0.35 μm @ 1.5 GHz.
  - < 1.5 dB NF at 5 mW achievable in 0.35 μm @ 900 MHz.
- These values assume negligible second-stage contributions, as well as relatively low loss in input tuning inductors.

Wide Dynamic Range Passive Mixers

- Gilbert cells used in bipolar technology because switching currents is easier there than switching voltages
  - Tradeoff is degraded linearity/power because of V-I conversion step
- CMOS has good switches so rationale for Gilbert cell as mixer is not as strong
- Have achieved conversion loss of < 4 dB at 1.6GHz.
- Measured SSB NF of < 10 dB, input IP3 >10 dBm, at "zero" power consumption
  - ~200μW power consumption dominated by LO drive power, and can be reduced further if gate capacitance is resonated out.
- Lack of DC bias implies no 1/f noise (important for direct-conversion and very-low-IF architectures)

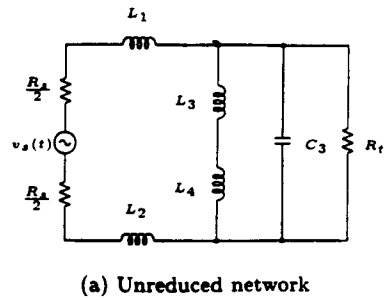


(a) Mixer with probe buffer

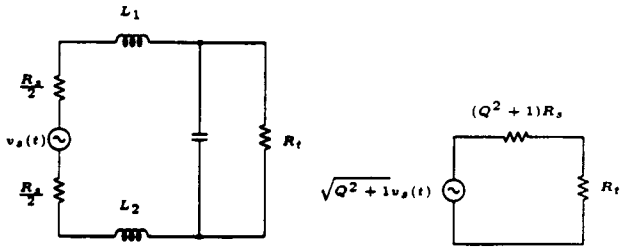


(b) Mixer circuit used in analysis

Fig. 5. Mixer circuit diagram



(a) Unreduced network



(b) RF tank reduced (c) L-match reduced

Fig. 6. Illustration of passive network reduction at resonance

lated. Figure 4 plots noise figure as a function of input device width and clearly shows an optimum width of about 500μm, corresponding to a noise figure of 1.8dB. Note that this curve represents the theoretical noise contribution of the input devices only.

The implemented width is only 290μm, because the detailed nature of the gate noise was unknown to the authors when this amplifier was designed. However, the curve has a broad minimum, so the achievable noise figure is little affected by using transistors of this width, at least in principle. The discrepancy between the theoretical minimum of 1.8dB and the measured noise figure of 3.8dB will be addressed in Section IV.

III. Mixer

A. Mixer Description

The mixer consists of the four transistors, M_1 through M_4, in Figure 5(a). These four transistors are grouped together into two pairs of two transistors each. Transistors M_1 and M_4 work together and are controlled by the local oscillator signal, while transistors M_2 and M_3 form a unit controlled by the inverse of the LO signal. Each pair serves the function of connecting the intermediate frequency (IF) port to the RF port of the mixer. The difference between the two pairs is the polarity with which they connect the IF port to the RF port. When M_1 and M_4 are on, the IF port is connected with a positive polarity to the RF port. But when M_2 and M_3 are on, the IF port is connected with a negative polarity to the RF port.

A probe buffer, included only for testing purposes, follows the mixer, and presents a high impedance load to the mixer while interfacing to off-chip test equipment. A passive filter network precedes the mixer, and can be separated into two parts for convenient analysis: an L-match and a RF tank. The L-match is formed by inductors L_1 and L_2 with part of the tank capacitance, C_3, while the RF tank is formed by inductors L_3 and L_4 with the remainder of C_3. L_1 through L_4 are implemented with bondwires, and C_3 is a metal to metal capacitor that incorporates lateral flux, as well as vertical flux. The purpose of the L-match is to boost the signal voltage across the mixer's RF port via an impedance transformation, while the RF tank is used to filter out of band noise at the RF port of the mixer. As will be discussed later, this filtering is important because multiple frequencies at the mixer's RF port are converted to the intermediate frequency at the mixer's IF port.

B. Mixer Conversion Gain

Figure 5(b) shows a simplified mixer circuit that is used in the following analysis.

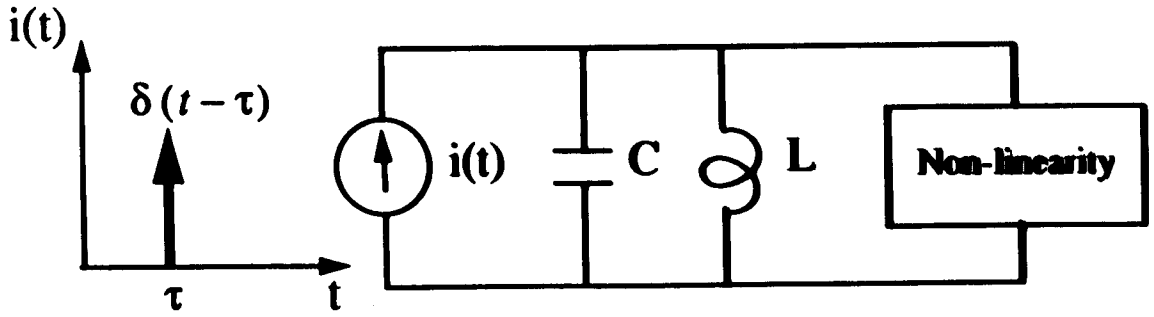
B.1 Definition

The voltage conversion gain for this mixer is found by exciting the circuit with a RF sinusoid, $v_s(t) = cos(2\pi f_{RF}t + \phi_{RF})$, and determining the IF signal amplitude at the IF port. The task of determining voltage conversion gain is broken into two steps. First, the voltage gain between the source and the RF port is computed. Second, the voltage conversion gain between the RF port and the IF port is computed. The mixer's voltage conversion gain is the product of the two steps.

Time-Varying Phase Noise Theory

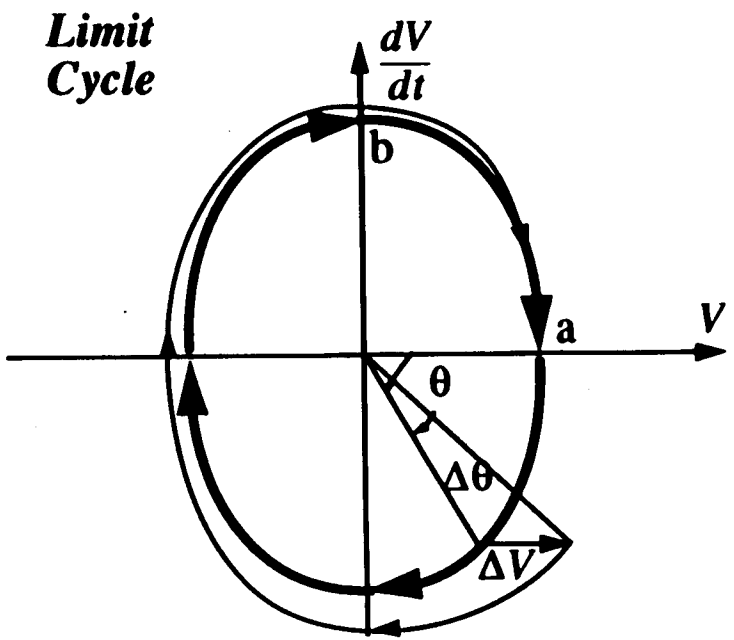
- Have developed quantitative phase noise model.
  - Acknowledges that oscillators are periodically time-varying linear systems.
- Re-interprets Leeson model in some areas, refutes it in others.
- Key insights:
  - Ideal oscillator is one which returns energy to resonator in impulsive fashion, timed precisely to coincide with voltage peak.
  - Device 1/f noise need not upconvert into close-in phase noise; can suppress such upconversion if certain previously unappreciated symmetry criteria are satisfied.
    - Good news for technologies with notoriously poor 1/f performance, such as MOSFETs and MESFETs.

Current Impulse Input

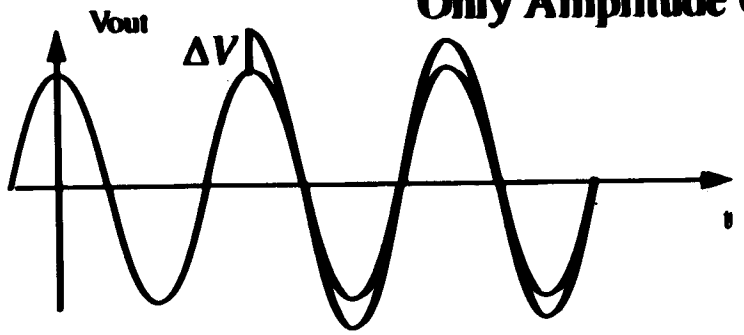


Phase error persists over time.

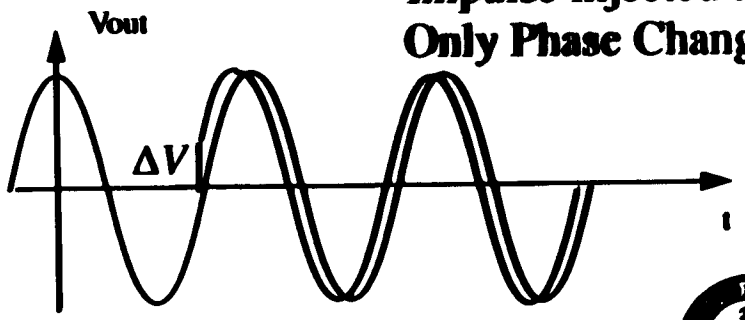
Non-linearity quenches amplitude changes over time.



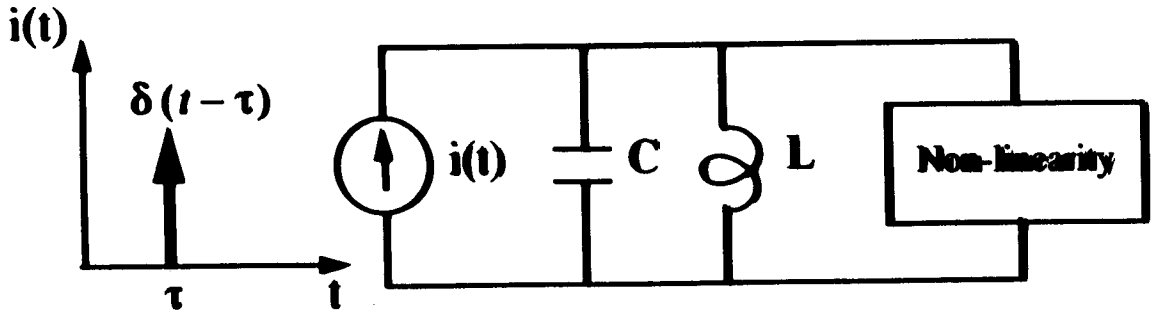
Impulse injected at *a*
Only Amplitude Change



Impulse injected at *b*
Only Phase Change

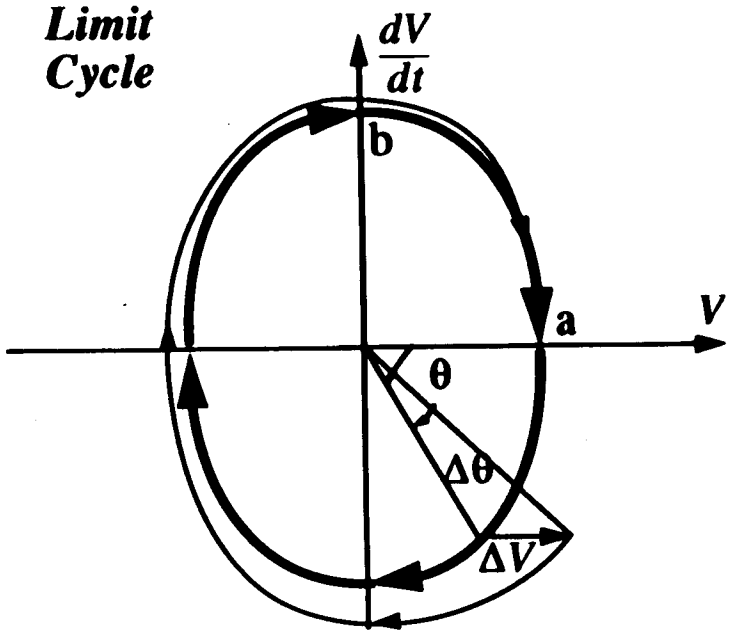


Current Impulse Input

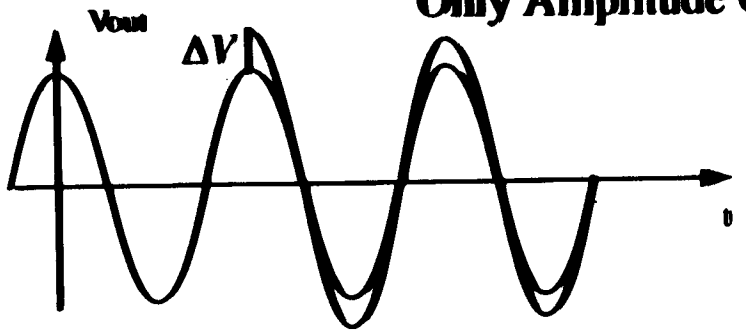


Phase error persists over time.

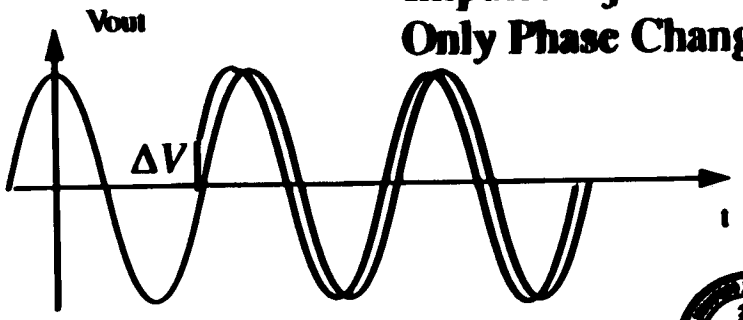
Non-linearity quenches amplitude changes over time.



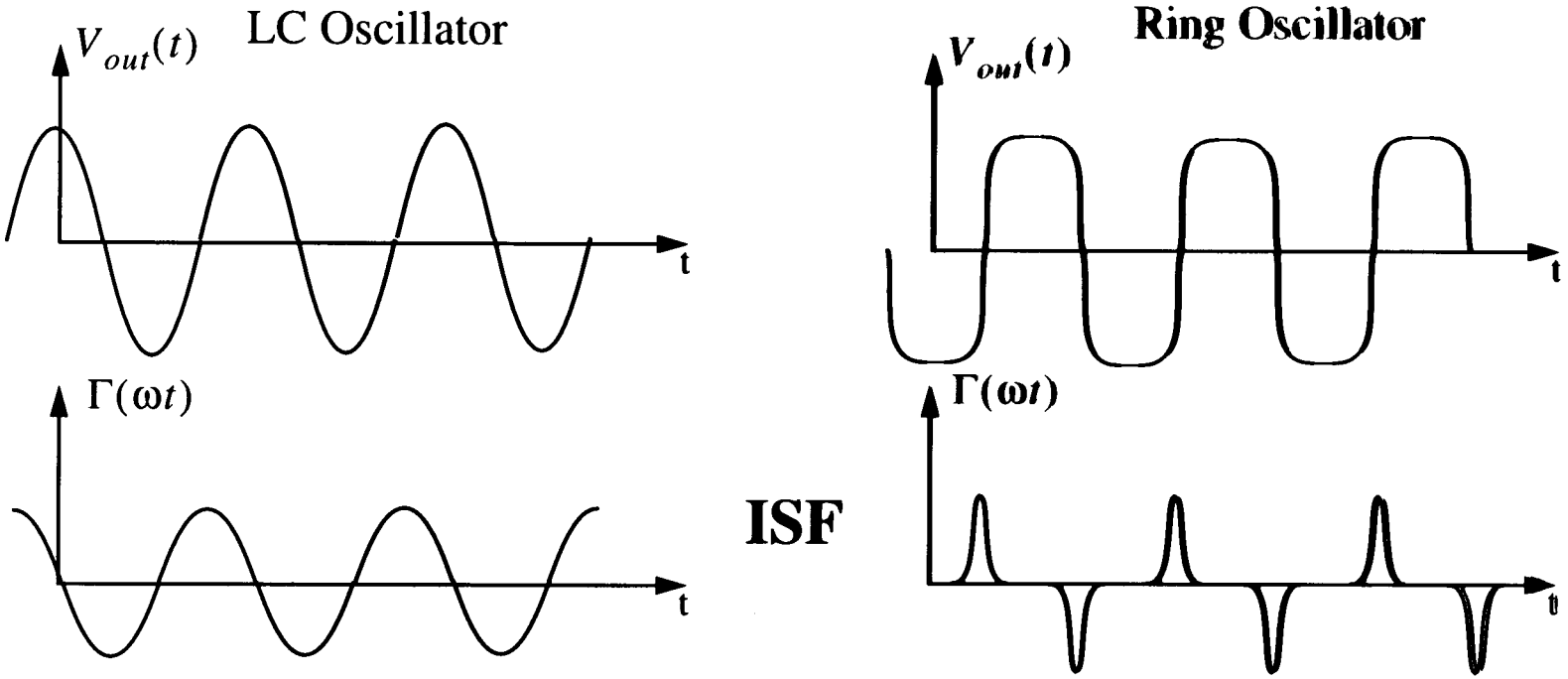
Impulse injected at *a*
Only Amplitude Change



Impulse injected at *b*
Only Phase Change



Impulse Sensitivity Function (ISF)

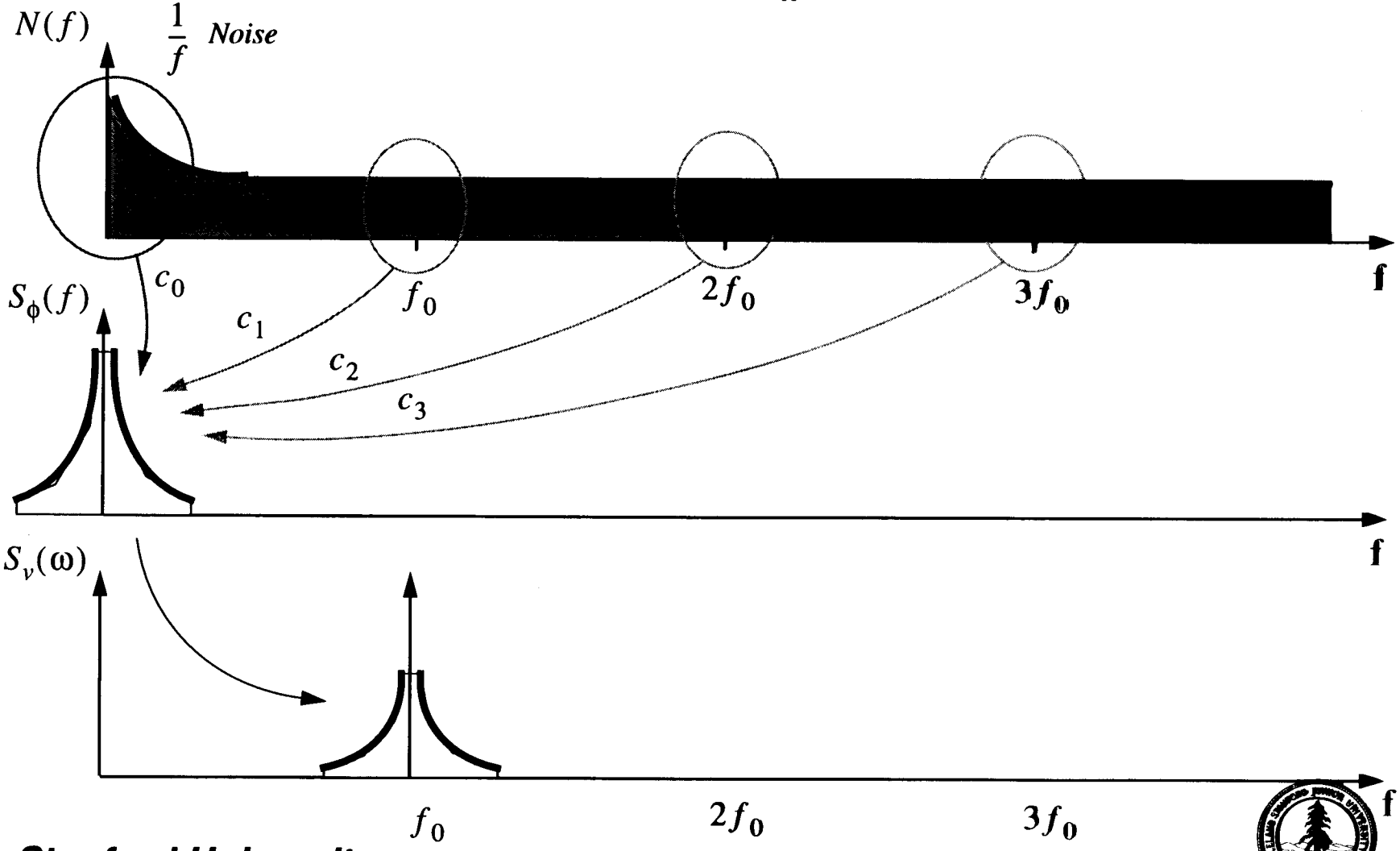


$\Gamma(\omega t)$ ***can be directly calculated from the waveform.***

Note that although the transfer function from $i_n(t)$ to $\phi(t)$ is linearized, the model still takes into account the effect of non-linear elements in the circuit, through the function $\Gamma(\omega t)$

Major Contributors to Phase Noise

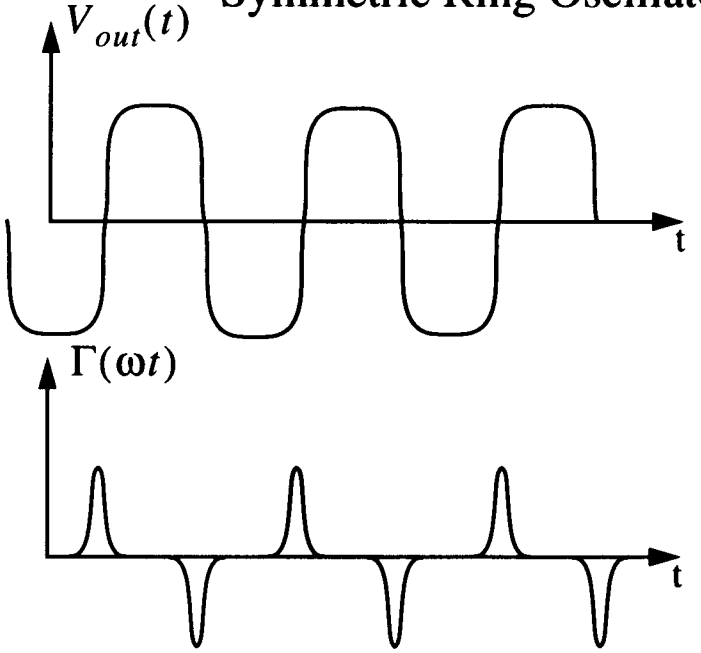
$$\phi(t) = \frac{1}{q_{max}}\left[\frac{c_0}{2}\int_{-\infty}^{t} i(\tau)d\tau + \sum_{n=1}^{\infty} c_n \int_{-\infty}^{t} i_n(\tau)\cos(n\omega\tau)d\tau\right]$$



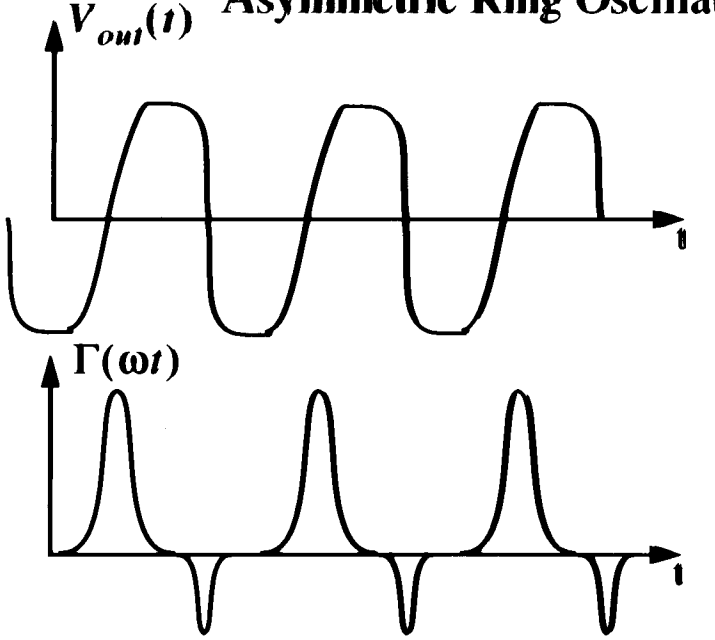
Effect of Symmetry on Low Frequency Noise Upconversion

$$c_0 = \frac{1}{\pi}\int_0^{2\pi} \Gamma(x)dx$$

Symmetric Ring Oscillator



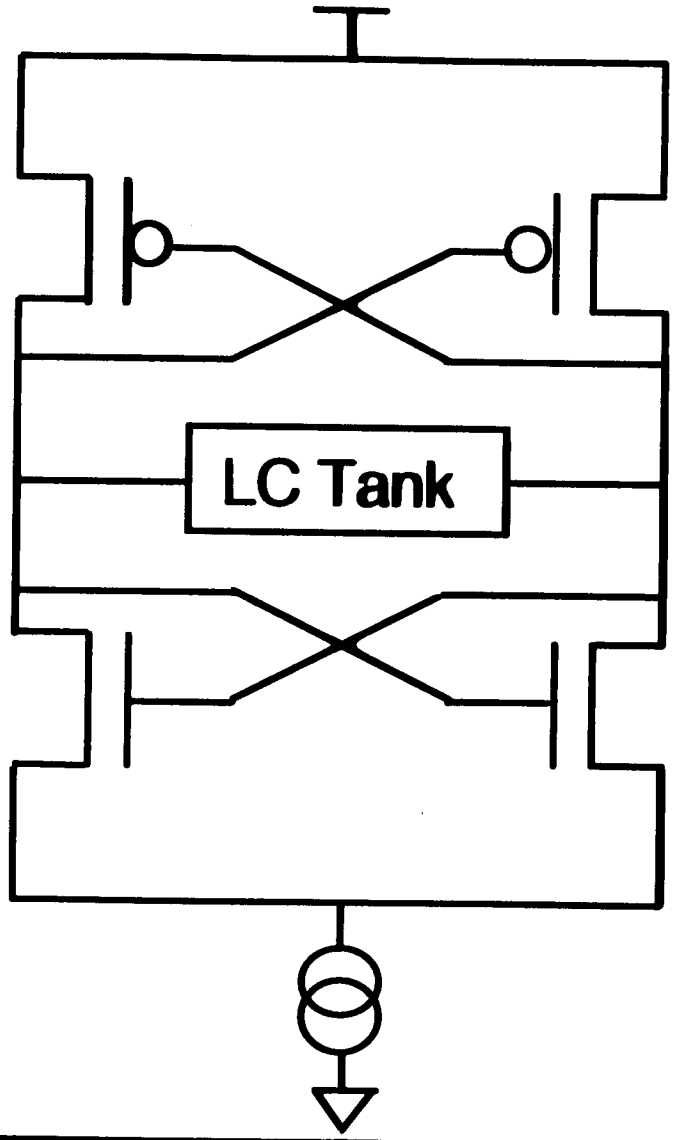
Asymmetric Ring Oscillator



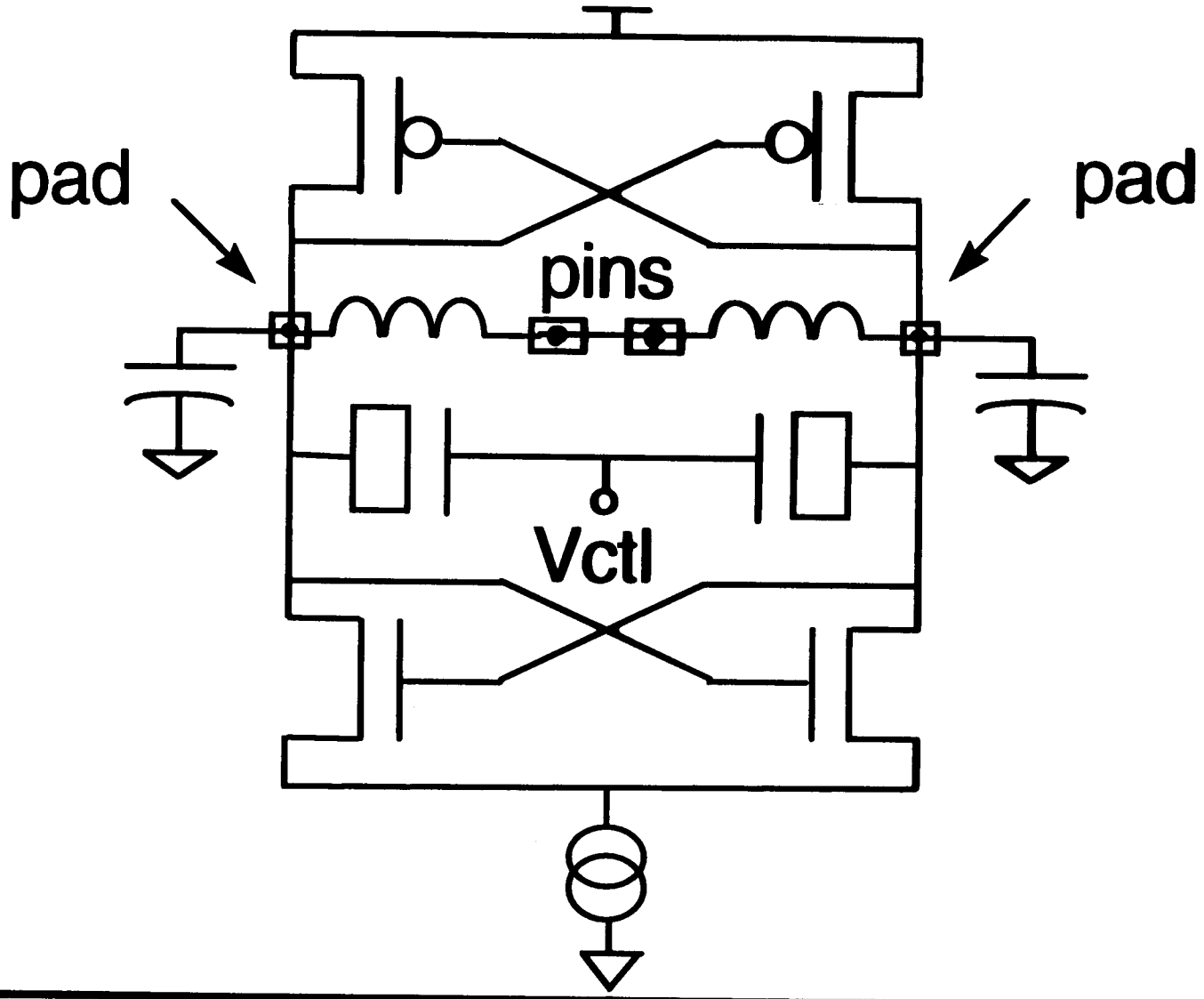
The DC value of ISF, which is governed by symmetry properties of the waveform, controls the contribution of low frequency noise to the phase noise

Basic Design Configuration

$$Wp = 2 \times Wn$$



Circuit Diagram



Results

Frequency	1.6GHz
Power	0.5mW at 2.0V supply
Phase Noise	-95dBc/Hz (100kHz offset)
Tuning Range	130MHz
Process Technology	0.5-µm MOSIS standard CMOS

Figure of Merit at 100kHz offset

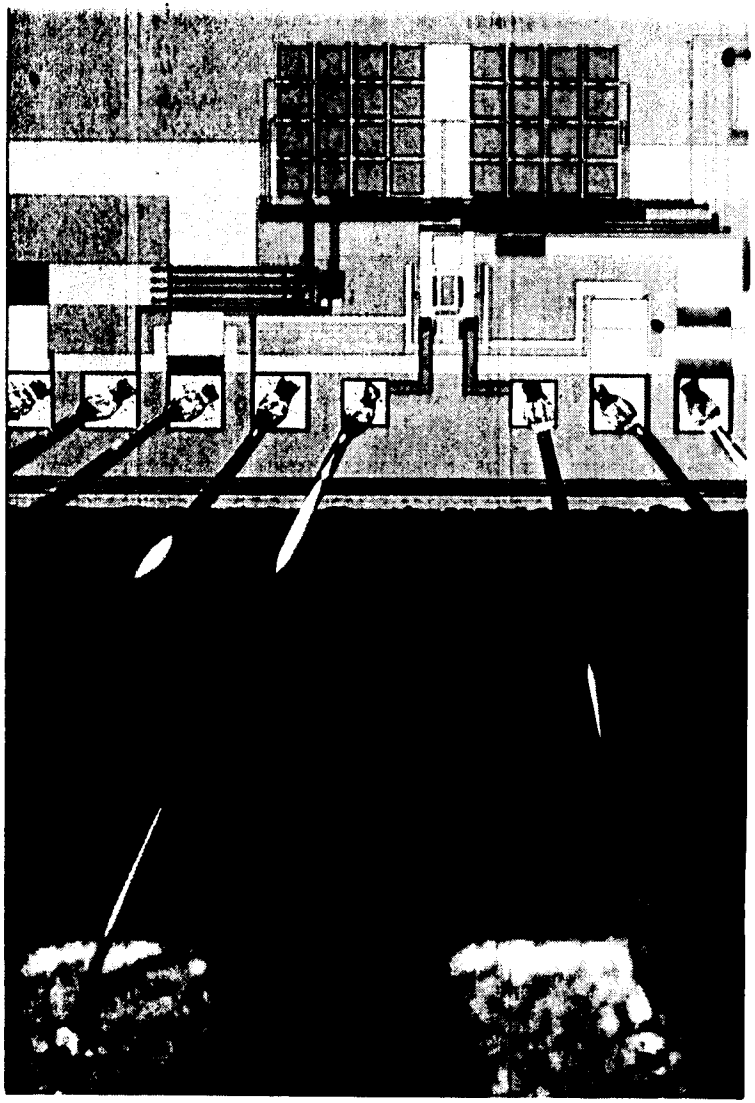
Technology	Frequency	Power	Phase Noise @100kHz	FOM
[7] Bipolar	1.6GHz	3mW	-95.1dBc/Hz	212dBF
CMOS	1.6GHz	0.5mW	-95dBc/Hz	220dBF

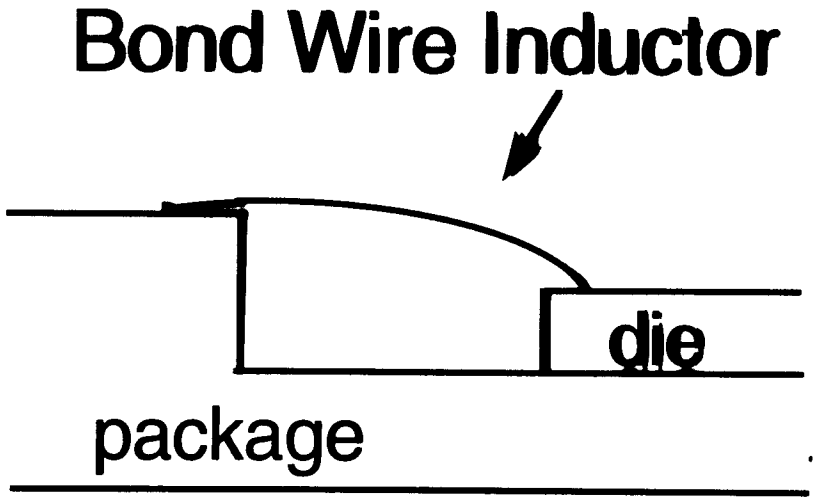
Figure of Merit = 10 log (Freq/(Phase Noise x Power))

CMOS Oscillators

- Ring oscillators are popular CMOS idioms, but suffer from high phase noise-power product.
  - Typical value is -75 dBc/Hz @ 100 kHz offset @ 1 GHz, 1 mW.
  - Improves 10 dB with every 10 dB power increase.
  - Values can be worse because of inferior 1/f noise of CMOS devices.
- At a given level of phase noise, a ring oscillator will consume approximately Q times the power of a tuned oscillator.
  - On-chip Q values limited to < 5-10; power reductions are nevertheless significant.
  - Bondwires have Q > 50; dramatic improvement possible.
    - -90-100 dBc/Hz phase noise at 100 kHz offset theoretically possible with ~1 mW power consumption if tuned oscillator is used.

Die Photo and Packaging





Putting it all Together

- Examples of CMOS receivers/ transmitters accumulate.
- Berkeley (ISSCC '97): DECT (1.9GHz) receiver
  - ~20dB overall NF, 200mW power consumption
  - Synthesizer not included
- UCLA (ISSCC '97): 900MHz frequency-hopped spread-spectrum transceiver
  - ~8-10dB overall NF, 360mW consumption in receive mode
  - Synthesizer not included
- Stanford (ISSCC '98): 1.6GHz GPS receiver
  - 4.9dB measured overall NF, 115mW power consumption
  - 57dB overall SFDR (measured)
  - Includes synthesizer, filters, I/Q downconversion, "A/D"s

Closing Thoughts

- CMOS will soon dominate below 5GHz, except in PAs because:
- CMOS f_T doubling every three years
  - At present, 30GHz max f_T (50GHz f_{max}) is state of the art
  - Experimental 150GHz f_T devices have been demonstrated
- Number of interconnect levels continues to increases slowly
  - Presently, 5 layers are not uncommon
- Substrate problems are not nearly as great as perceived
  - Quality of passive elements now comparable to GaAs
  - Extra interconnect layers continue to help
- Device and oscillator noise continue to improve