

# RF CMOS Technology Scaling in High-k/Metal Gate Era for RF SoC (System-on-Chip) Applications

C.-H. Jan, M. Agostinelli, H. Deshpande, M. A. El-Tanani, W. Hafez, U. Jalan, L. Janbay, M. Kang, H. Lakdawala\*, J. Lin, Y.-L. Lu, S. Mudanai, J. Park, A. Rahman, J. Rizk, W.-K. Shin, K. Soumyanath\*, H. Tashiro, C. Tsai, P. VanDerVoorn, J.-Y. Yeh, P. Bai

Logic Technology Development (LTD), Intel Corporation, Hillsboro, Oregon, USA

\*Integrated Platform Research/Radio Integration Research (RIR), Intel Labs, Intel Corporation, Hillsboro, Oregon, USA

Contact: e-mail chia-hong.jan@intel.com

## Abstract

The impact of silicon technology scaling trends and the associated technological innovations on RF CMOS device characteristics are examined. The application of novel strained silicon and high-k/metal gate technologies not only benefits digital systems, but significantly improves RF performance. The peak cut-off frequency ( $f_T$ ) doubles from 209 GHz in the 90 nm node to 445 GHz at the 32 nm node.  $1/f$  flicker noise reduces by an order of magnitude from the 0.13  $\mu\text{m}$  node to the 32 nm node. Transistor noise figure, high voltage tolerance, and quality factors of RF passives all show similar benefits from technology scaling.

## Introduction

As silicon technology scaling progresses to the 32 nm node, single chip integration of RF and communication designs with the micro-processor cores on a common CMOS system-on-chip (SOC) platform has become increasingly appealing. This increased attention to RF SOC is driven by both the huge improvements in device performance afforded by Moore's Law and by the advantage of higher vertical integration and lower manufacturing cost of mainstream CMOS technology.

Modern technology scaling, however, is no longer a simple matter of shrinking device dimensions. Today's technology scaling is enabled by introducing disruptive, innovative materials and novel device structures. Examples of these innovations include strained silicon and high-k/metal gates for transistors and low-k ILD and Cu metallization in the backend. The impact of these new inventions to mixed signal/RF designs needs to be examined with the intent to identify the true promise and challenges of radio integration into a general SOC platform.

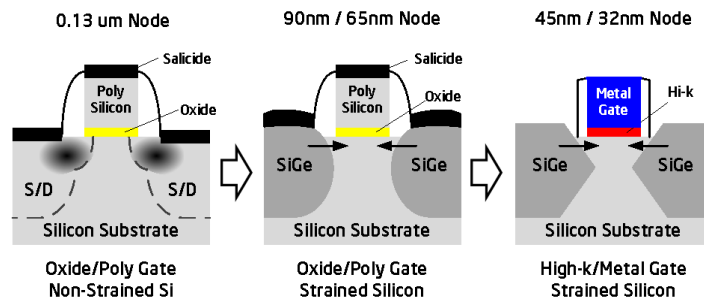


Fig. 1 Evolution of CMOS transistor architecture in the last decade, from a non-strained oxide/poly gate structure to a high-k/metal gate strained silicon transistor at the 32 nm node.

## CMOS Technology Scaling Trend

Fig. 1 summarizes the main transistor architecture evolutions in the last decade - traditional oxide/poly gate architec-

tures (<0.13  $\mu\text{m}$ ), oxide /poly gate structures with strained silicon (90 nm/65 nm), and high-k/metal gate technology with strained silicon (45/32 nm, Fig. 2) [1-11]. The corresponding transistor pitch scaling has maintained a steady 0.7x per 2 year trend as shown in Fig. 3. The electrical  $\text{Tox}$  trend in Fig. 4 also followed a 0.7x per generation scaling trend until the 90 and 65 nm nodes, where scaling stalled as further thinning of the silicon dioxide resulted in unacceptably high gate leakage. Performance enhancements at these technology nodes were achieved by the incorporation of novel strained silicon technology. High-k/metal gate technology was brought into production starting at the 45 nm node to enable further dielectric scaling along with drive current and  $g_m$  improvements as shown in Fig. 5.

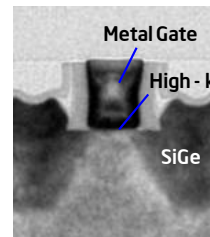


Fig. 2 State of the art 32 nm PMOS high-k/metal gate transistor.

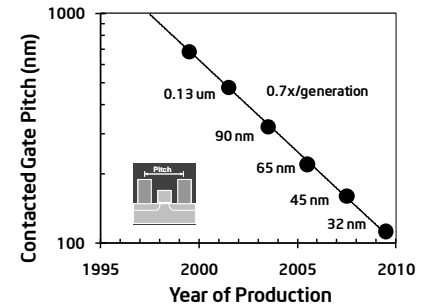


Fig. 3 Intel CMOS transistor pitch scaling trend

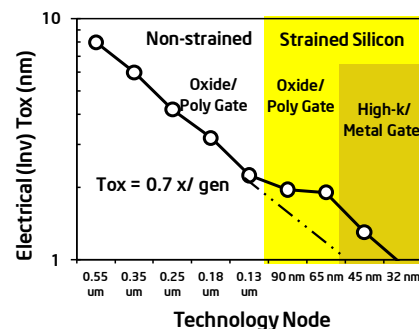


Fig. 4 CMOS transistor gate dielectric thickness scaling trend.

While digital systems are continuously benefiting by general CMOS scaling per Moore's Law, CMOS implementation for RF and mixed signals presents different challenges. Fig. 6 shows a typical fully integrated RF architecture, including key sub-systems of switches, PAs, LNAs, mixers, frequency synthesizers, ADCs, DACs and baseband processors. It can be seen from Table I that the key RF devices and their characteristics, including device matching, linearity, cutoff frequencies, flicker

noise, thermal noise, on-state and output resistance, and quality factors of RF passives, emphasized in such analog subsystems are very different from digital system requirements, and necessitate distinct optimization of process and design methodologies. Furthermore, the monolithic integration of noise-sensitive analog blocks alongside rapidly switching digital circuits places demanding requirements on the noise isolation capabilities of a process; while the implementation of higher resistivity substrates, triple-well architectures, and guard-rings can reduce the impact of digital crosstalk, alternative analog layout techniques, such as clean power rails and differential circuit designs, are often required.

### RF Transistor Scaling Trend

RF transistor matching performance has been improved with the advance of technology scaling (Fig. 7). The transistor matching coefficient  $c_2$  has been improved 40% since the 0.18  $\mu\text{m}$  node, noting that  $c_2$  reduction slowed in the 90 and 65 nm nodes due to minimal oxide scaling. Through the implementation of high-k dielectrics,  $T_{ox}$  scaling resumed, enabling  $c_2$  scaling beyond 45 nm as explained by Eq. 1 [7]. Intra-die process variations due to oxide thickness, doping profile, and work-function variation represent another challenge for RF CMOS scaling. Such variations increase with scaling, and are responsible for local transistor mismatches. The  $\sigma V_T$  in Fig. 7 shows excellent control, remaining flat over the last three nodes.

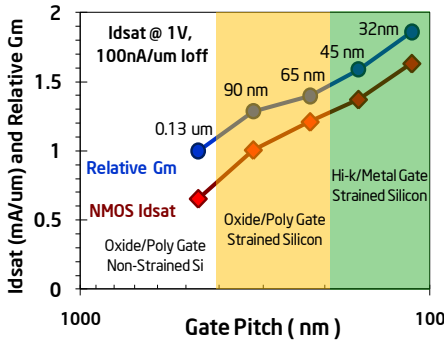


Fig. 5  $I_{dsat}$  and  $G_m$  scaling trend from 0.13  $\mu\text{m}$  to 32 nm technology.

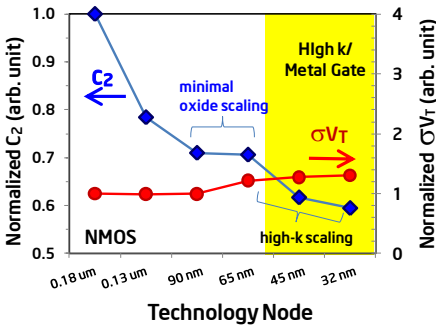


Fig. 7 Transistor matching metrics,  $C_2$  and  $\sigma V_T$ , scaling trends from 0.18  $\mu\text{m}$  to 32 nm.

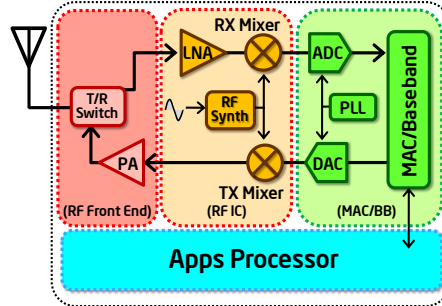


Fig. 6 Typical fully integrated radio-on-chip architecture, including sub-systems of T/R switch, PA, LNA, Mixer, RF synthesizer, ADC, DAC and baseband.

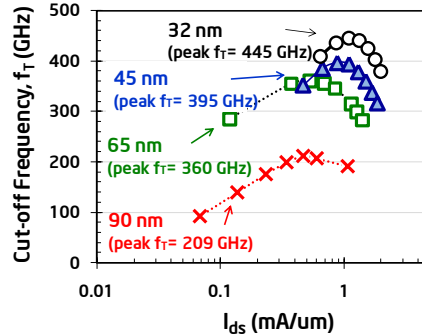


Fig. 8 State of the art cut-off frequency  $f_T$  scaling trend vs.  $I_{ds}$  for 90 nm to 32 nm nodes.

$$\sigma V_T = \left( \frac{\sqrt[4]{4q^3 \epsilon_{Si} \phi_B}}{2} \right) \cdot \frac{T_{ox}}{\epsilon_{ox}} \cdot \left( \frac{\sqrt[4]{N}}{\sqrt{W_{eff} \cdot L_{eff}}} \right) = \frac{1}{\sqrt{2}} \left( \frac{C_2}{\sqrt{W_{eff} \cdot L_{eff}}} \right) \quad (\text{Eq. 1})$$

Cut-off frequency  $f_T$  has also steadily improved with drive current enhancements as shown in Fig. 8. The peak  $f_T$  of the state of the art 32 nm NMOS is 445 GHz, twice as fast as the then-record 209 GHz for 90 nm. The benchmark of  $f_T$  vs. gate length in Fig. 9 shows that strained silicon and high-k/metal gate implementation has enabled better cut-off frequency performance over the ITRS roadmap. This performance improvement can be explained by the simple analytical form of  $f_T$  in Eq. 2, namely that the transconductance improvement due to transistor enhancements is outpacing the increase of the parasitic capacitance due to device geometry and dielectric scaling. Achieving higher cutoff frequencies via the scaling of RF CMOS technologies will be increasingly challenging to maintain the transconductance gain outpacing parasitic increases due to shrinking geometry and pitch. Fringing capacitance between the gate and the source/drain contacts will negatively impact  $C_{gg}$ , and gate resistance due to poly line-width scaling will present an obstacle to higher frequency operation. Alternative layout strategies and transistor architectures will be required to mitigate such parasitic increases and enable cutoff frequencies to continue to scale.

$$f_T = \frac{g_m}{2\pi C_{gg}} \quad (\text{Eq. 2})$$

Table I Key RF device characteristics of primary RF circuit blocks

| RF Devices                 | RF Circuits                 | Key Device Characteristics  |
|----------------------------|-----------------------------|---|
| Logic Transistor           | MAC/BB, ADC, DAC            | $I_{dsat}$ , $I_{dlin}$ , $V_t$ , $I_{off}$                         |
| Analog Transistor          | ADC, DAC, MAC/BB            | $G_m$ , $R_{out}$ , Matching, Linearity, Noise, $NF_{min}$          |
| RF Transistor              | PA, Mixer, T/R Switch       | $f_T$ , $f_{max}$ , $1/f$ Noise, $NF_{min}$                         |
| PA Transistors             | PA                          | $R_{on}$ , Linearity, $f_T$ , $f_{MAX}$ , Efficiency, Breakdown $V$ |
| Precision Resistors        | ADC, DAC, BB Filter, others | $R$ , $\sigma R/R$ , Matching                                       |
| Linear Capacitors          | PLL, VCO                    | $C$ , $Q$ , Matching  |
| Varactors                  | PLL, VCO                    | Tuning Ratio, $Q$ , $KV_{CO}$                                       |
| Inductor/Transformer/Balun | PA, LNA, Mixer              | $L$ , $Q$   |

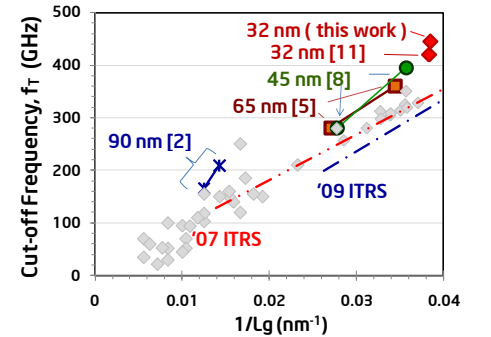


Fig. 9 Cut-off frequency  $f_T$  scaling trend vs.  $1/L_g$  for 90 nm to 32 nm nodes. [2,5,8,11]

The scaling of the core  $V_{CC}$  from 5V on the 0.8  $\mu\text{m}$  node to less than 1V for the 32 nm node presents another deleterious aspect impacting RF designs. The voltage scaling of I/O and platform peripherals typically lags behind the CPU core voltage, necessitating additional high voltage 1.8V-3.3V native oxide transistors to be supported. PA integration in a monolithic SOC requires devices that can sustain high voltage swings; a special high drain-voltage ( $> 5\text{ V}$ ) tolerant device has been developed which possess performance similar to the native 3.3 V I/O transistor but with  $> 50\%$  higher breakdown voltage (Fig. 10). The scaling trend for well architecture is transitioning from the traditional twin well to triple well/deep n-well structures. The deep n-well, Fig. 11, enables improved substrate noise isolation at low to moderate frequencies. In addition, parasitic NPN BJTs (Fig. 12) and p-ch JFETs (Fig. 13) can be formed for applications which require improved mixed signal/RF circuit noise performance.

### RF Noise Scaling Trend – Flicker Noise and $NF_{min}$

Drain current noises,  $S_{id}$  (Eq. 3), and input referred gate noises,  $S_{vg}$  (Eq. 4), are two key transistor metrics characterizing flicker noise. Figs. 14 and 15 show that normalized  $S_{vg}$  is monotonically decreasing with each successive technology node, resulting in a 10x reduction from the 0.18  $\mu\text{m}$  node to the 32 nm node. This benefit from oxide scaling, enabled by high-k/metal

gate technologies, can be quantified by the dependency of  $C_{ox}$  in Eq.4.

$$S_{id} = \frac{K}{f} \frac{g_m^2}{W \cdot L \cdot C_{ox}^2} \quad (\text{Eq. 3})$$

$$S_{vg} \cdot W \cdot L = \frac{S_{id}}{g_m^2} \cdot W \cdot L = \frac{K}{f} \frac{1}{C_{ox}^2} \quad (\text{Eq. 4})$$

The thermal noise trend of the minimum noise figure,  $NF_{min}$ , is shown in Figs. 16 and 17 where a steady improvement with scaling is observed from 0.18 $\mu\text{m}$  to the 45nm nodes, but saturates at the latest technology node. Fukui's equation (Eq. 5) describes that  $NF_{min}$  improves with increasing  $g_m$  driven by scaling gate length and oxide thickness, but is negatively impacted by increasing parasitic components (i.e.  $C_{gg}$  and  $R_g$ ) due to higher fringe capacitances and resistances due to the tighter pitch.

$$NF_{min} = 1 + K \frac{f}{f_T} \sqrt{g_m \cdot (R_g + R_s)} = 1 + K \frac{2\pi f \cdot C_{gg}}{\sqrt{g_m}} \sqrt{(R_g + R_s)} \quad (\text{Eq. 5})$$

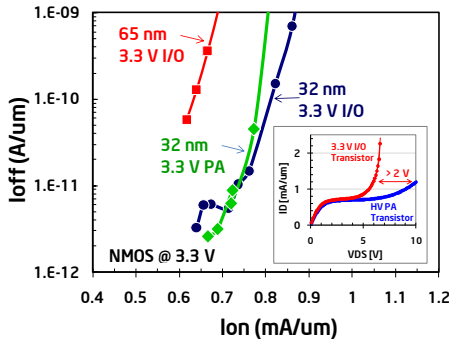


Fig. 10  $I_{on}/I_{off}$  and breakdown voltage (inset) comparison between native 3.3V I/O and HV PA transistors.

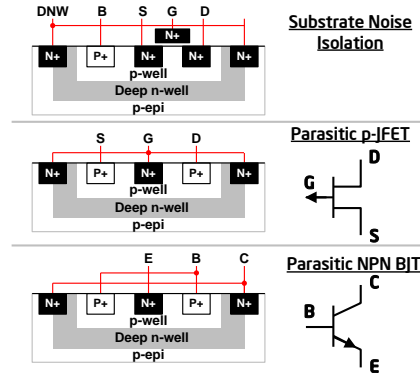


Fig.11 The applications of deep n-well architectures for substrate noise isolation and additional parasitic devices.

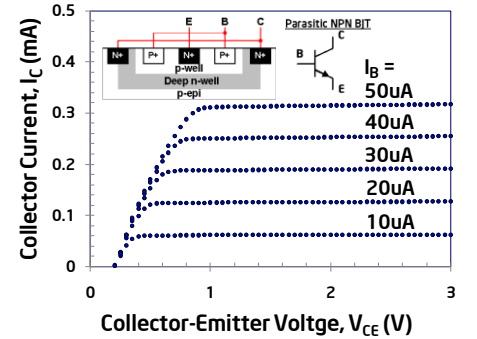


Fig. 12 Typical electrical characteristics of the parasitic NPN BJT fabricated with deep n-well.

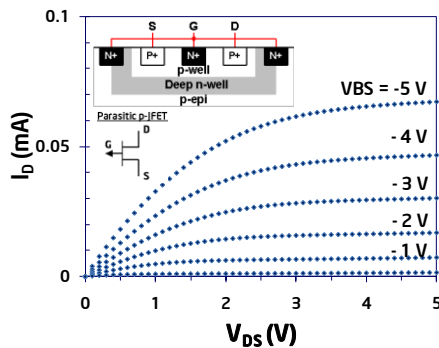


Fig. 13 Typical electrical characteristics of the parasitic p-ch JFET fabricated with deep n-well.

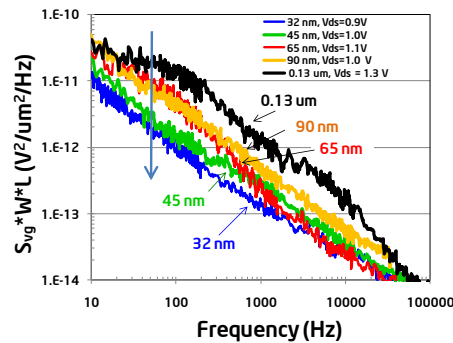


Fig. 14 Flicker noise scaling trend of 0.13  $\mu\text{m}$ , 90, 65, 45 and 32 nm technologies.

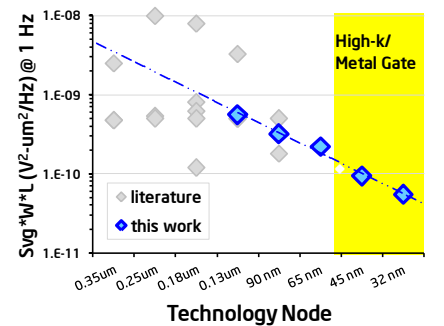


Fig. 15 Normalized input-referred flicker noise ( $S_{vg} \cdot W \cdot L$  extrapolated to @1 Hz) scaling trend.

## RF Interconnects and Passives Scaling Trend

Interconnect pitch and thickness scaling is part of the continuous dimensional scaling as shown in Figs. 18 and 19. The increased capacitance and resistance associated with interconnect scaling can compromise RF designs due to quality factor degradation of the spiral inductors built on the upper metal layers of the interconnect stack. At the 45 nm node and beyond, a thick metal (TM) layer was added on top of the standard interconnect layers to serve as power gating devices and to construct lower loss inductors to compensate for such interconnect scaling effects (Fig. 20). Inductors fabricated on the TM layer exhibit excellent performance when benchmarked with literature (Fig. 21).

## Summary

RF CMOS technology is shown to benefit from the general CMOS technology scaling. The majority of RF device performance metrics have steadily improved by the introduction of innovative transistor and interconnect technologies, including strained silicon and high-k/metal gates in the past decade.

## References

- [1] S. Thompson et al., IEDM Tech. Digest, p. 61 (2002)
- [2] K. Kuhn et al., VLSI Technology Symp., p. 224 (2004)
- [3] P. Bai et al., IEDM Tech. Digest, p. 657 (2004)
- [4] C.-H. Jan et al., IEDM Tech. Digest, p. 60 (2005)
- [5] I. Post et al, IEDM Tech Dig., pp. 1-3 (2006)
- [6] K. Mistry et al, IEDM Tech Dig., pp. 247-250 (2007)
- [7] K. Kuhn et al, IEDM Tech Dig., (2007)
- [8] C.-H. Jan et al, IEDM Tech. Dig., pp. 637-640 (2008)
- [9] S. Natarajan et al, IEDM Tech. Dig., pp. 941-943 (2008)
- [10] C.-H. Jan et al, IEDM Tech. Digest, p. 28.1 (2009)
- [11] P. VanDerVoorn et al, VLSI Tech. Symp., p. 137 (2010)

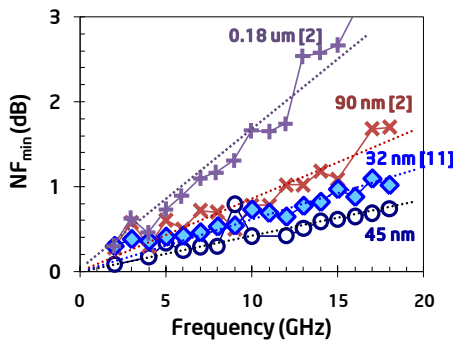


Fig. 16 Minimum noise figure performance vs. frequency scaling trend.

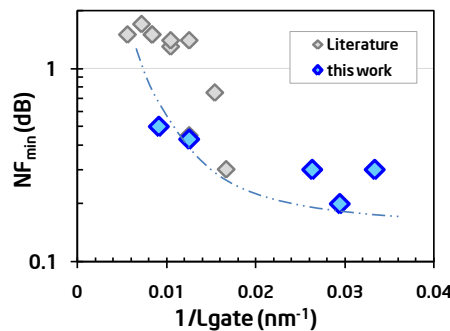


Fig. 17 Minimum noise figure ( $NF_{min}$ ) scaling trend.

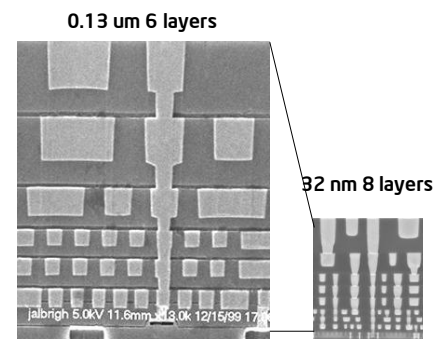


Fig. 18 SEM images of 0.13 um 6 layer metal stack vs. 32 nm [10] 8 layer interconnect stack.

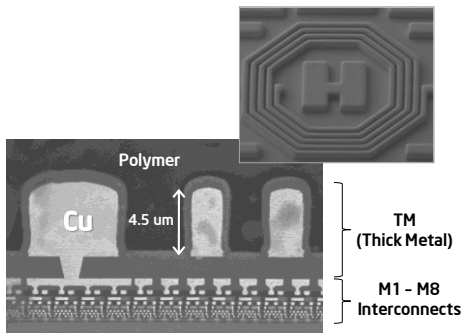


Fig. 19 SEM images of the 32 nm Thick Metal (TM) layer and the inductor made of the TM layer.

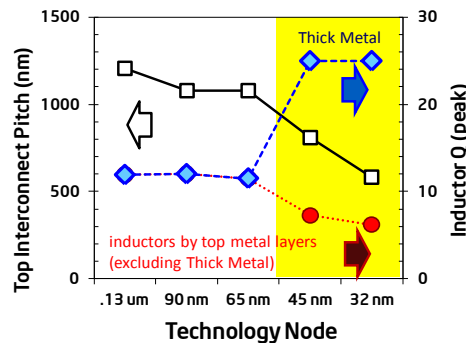


Fig. 20 Interconnect pitch vs. inductor  $Q$  scaling trend.

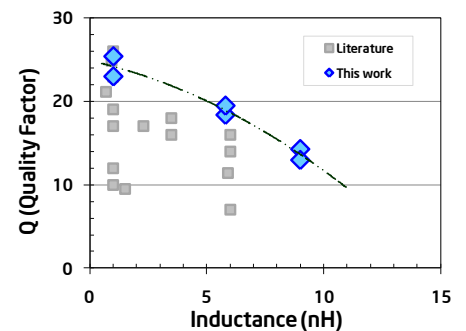


Fig. 21 Inductor  $Q$  benchmark as a function of inductance.