

Transformer-Based Tunable Matching Network Design Techniques in 40-nm CMOS

Apsara Ravish Suvarna, Venumadhav Bhagavatula, *Member, IEEE*, and Jacques C. Rudell, *Senior Member, IEEE*

Abstract—A fully integrated transformer-based tunable impedance-matching network is described. The tuning network independently tunes the real and imaginary parts of the impedance. A test chip implemented in a 40-nm CMOS process achieves the resistive tuning range of one octave, making it suitable for shared Bluetooth/Wi-Fi power amplifier (PA) applications. The main sources of insertion loss are identified, and strategies to minimize the insertion loss while maximizing the tuning range of the real and imaginary parts are discussed in this brief.

Index Terms—CMOS technology, passive circuits, radio-frequency integrated circuits, tunable circuits and devices.

I. INTRODUCTION

THE expanding market for small form-factor low-power mobile consumer electronics requiring numerous wireless communication platforms on the same device has driven the demand for multimode, multistandard, and software-defined radios. However, while highly programmable digital back ends are easily implemented for use in multistandard applications, the analog front end traditionally supports a much narrower range of applications because of the need to tune over such a broad range of conditions, such as the carrier frequency (ω_C) and output power. Analog circuits are tuned to a specific carrier frequency, ω_C , with the help of matching networks realized using integrated passive components to both maximize the power gain at a particular carrier frequency and improve the overall efficiency. Thus, single transceiver solutions for multiple standards, which entail processing signals over a large range of carrier frequencies, bandwidths, and power levels, would address one of several challenges associated with reconfigurable radios.

A matching network is characterized by two key parameters, namely, the center frequency and the impedance transformation ratio. In this brief, the design of a fixed-frequency *impedance-tunable* matching network is described. While design examples are provided for a variable impedance-matching network for use as a high-power PA–antenna interface, the concepts described in this brief are generic and can be applied to a broad class of matching networks used in integrated radios.

Manuscript received October 8, 2015; revised December 12, 2015; accepted February 8, 2016. Date of publication February 15, 2016; date of current version June 23, 2016. This brief was recommended by Associate Editor J. Kim.

A. Ravish Suvarna is with Maxim Integrated Products, San Jose, CA 95134 USA (e-mail: apsara@uw.edu).

V. Bhagavatula is with Samsung Semiconductors, Inc., San Jose, CA 95134 USA.

J. C. Rudell is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA.

Color versions of one or more of the figures in this brief are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSII.2016.2530258

There are several advantages in PA applications of a reconfigurable matching network in which the impedance transformation ratio between the PA load and antenna can be calibrated to vary the effective load resistance. One obvious application is to address the variation in a PA voltage standing-wave ratio (VSWR). The antenna impedance is a strong factor of many environmental conditions and changes continuously during the operation of any transceiver [1]. The challenge of modulating the PA output resistance, to minimize the VSWR, is maintaining a constant resonance at ω_C while modulating the load resistance. As such, any tuning or calibration must allow for the modulation of both the real and the imaginary PA load impedance to adequately address VSWR performance issues. The benefits of an integrated tuned load extend to multimode applications where a single PA could be reconfigured for use in multiple standards, operating at the same frequency but at different output powers, for example, Wi-Fi/ Bluetooth power amplifiers.

To date, there has been research on methods for both off-chip and on-chip tunable matching networks. These efforts rely on LC-ladder-based topologies, such as a two-stage ladder network [2], a T-network [3], and a combination of multiple Pi-networks [4]. While LC-ladder-based circuits are well suited for low transformation ratios, their insertion loss is prohibitively high for transformation ratios on the order of 10–12. Transformers, on the other hand, are well suited for such applications because the insertion loss is independent of the transformation ratio [5]. In addition, transformer-based networks may also serve as a balun to interface a differential PA with a single-ended antenna. The center tap of the transformer often serves as the supply bias node, thus eliminating the need for a large choke inductor. Therefore, the work presented in this brief assumes the use of a transformer with a fixed impedance transformation. This is supplemented with a Pi-network that converts a fixed-valued impedance that results from the transformer to variable input impedance.

This brief is organized as follows. Section II analyzes integrated transformers while describing a design methodology to minimize the insertion loss. Section III presents the design of a tunable matching network by using a transformer-plus-Pi-matched network (TPMN). Section IV presents the simulation and measured results. This brief concludes in Section V.

II. TRANSFORMER MODELING AND INSERTION LOSS MINIMIZATION

Transformers are capable of providing large impedance transformation ratios; therefore, they are frequently employed in the design of on-chip matching networks for power amplifiers. A simple circuit model for a transformer is shown

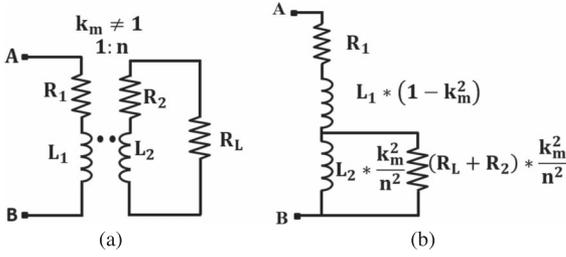


Fig. 1. (a) 1: n transformer with finite Q-factor modeled as a series resistor in the primary and secondary inductors. (b) Simplified transformer model used to estimate the input impedance.

in Fig. 1(a), where the primary and secondary windings have self-inductances L_1 and L_2 , respectively. These inductors are magnetically coupled with coupling coefficient k_m and turns ratio $n = \sqrt{L_2/L_1}$. The effective turns ratio of this 1: n turn transformer is n/k_m [6].

For the TPMN described in this brief, the transformer is a core building block. Integrated transformers fabricated in modern CMOS processes have several loss mechanisms, including the ohmic loss of metal traces, substrate losses, etc. For the sake of simplicity, all these losses can be modeled by the resistors R_1 and R_2 , on the primary and secondary sides, respectively. Several layout-based techniques have been proposed in prior literature to minimize the losses (R_1 and R_2 are minimized) in the transformer. However, a fact not widely recognized is that even with an optimized transformer layout, the effective insertion loss is a function of the load resistance (R_L) and tuning capacitor. An analysis of the relationship between insertion loss/load resistance/tuning capacitor and design guidelines for the optimal choice of tuning capacitor are provided in the following section.

A. Transformer Input Impedance

For a transformer in which the secondary coil is terminated with a load R_L , the impedance at the primary (across nodes A and B) can be simplified to the form shown in Fig. 1(b). Finite magnetic coupling between L_1 and L_2 results in a leakage inductance, $L_1(1-k_m^2)$ [5], [6]. The impedance at the secondary, which consists of L_2 in parallel with (R_L+R_2) , can be reflected to the primary with a scaling factor of $(k_m/n)^2$. The reflected impedance appears in series with $L_1(1-k_m^2)$ and R_1 . This model reveals that the impedance Z_{AB} has both resistive and inductive components. Therefore, to obtain a real Z_{AB} , the inductive component must be resonated with a capacitive element at the frequency of interest. This (tuning) capacitor could be added to either the primary or the secondary side. Tuning capacitors added on the primary coil have minimal effect on the transformer insertion loss. Therefore, the impact of a capacitor added to the secondary terminal, on the insertion loss, is analyzed. The loss analysis follows a sequence of series-to-parallel transformations described in Fig. 2. The analysis begins by assuming that the capacitor C_2 is the tuning capacitance. First, the impact of C_2 on the insertion loss resulting from R_2 is considered. Next, similar concepts are extended to understand the impact of C_2 on R_1 .

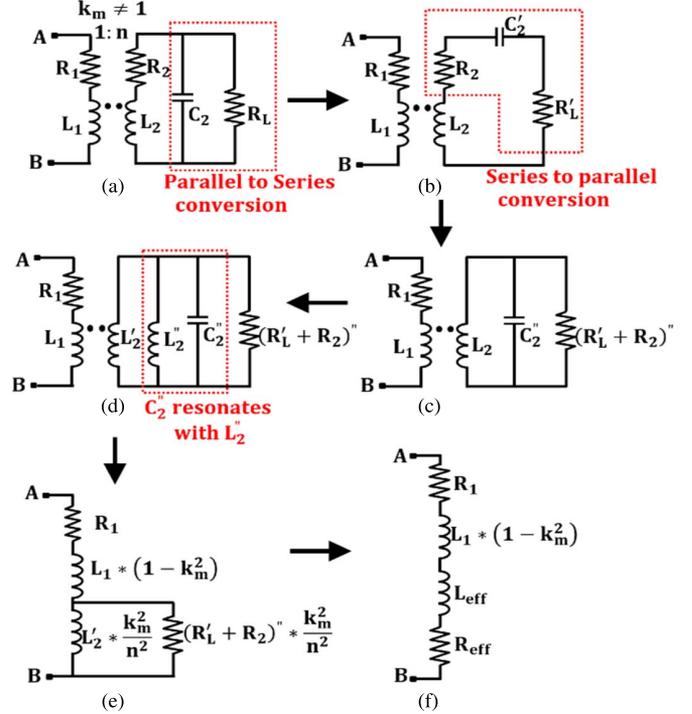


Fig. 2. Steps involved in simplifying a transformer network with a parallel secondary capacitor.

B. Minimizing Loss From R_2 Alone

At the resonant frequency ω_0 , the parallel circuit consisting of C_2 and R_L can be transformed to a narrow-band equivalent series network, R_L' in series with C_2' , as shown in Fig. 2(a) and (b), where

$$R_L' = \frac{R_L}{(1 + Q_p^2)} \quad \text{and} \quad C_2' = C_2 * (1 + Q_p^{-2}). \quad (1)$$

The parameter $Q_p = \omega_0 C_2 R_L$. Resistors R_L' and R_2 form a voltage divider. Therefore, in order to minimize the losses due to the secondary winding resistance R_2 , the component R_L' must be large compared to R_2 . From (1), one observes that to maximize R_L' , the capacitor C_2 should be minimized.

C. Minimizing R_1 Losses

The choice of C_2 also determines the extent to which the primary side resistance R_1 contributes to the overall insertion loss. To explain the impact of losses introduced by R_1 , commonly accepted network transformations described in Fig. 2(c)–(f) are first described.

The series network consisting of C_2' and $(R_L' + R_2)$ can be transformed to an equivalent parallel network, as shown in Fig. 2(b) and (c) with

$$(R_L' + R_2)'' = (R_L' + R_2) * (1 + Q_s^2) \quad (2a)$$

$$C_2'' = \frac{C_2'}{(1 + Q_s^2)} \quad (2b)$$

where Q_s in (2a) and (2b) is $1/(\omega_0 C_2' (R_L' + R_2))$. The effective shunt capacitance resonates with part of L_2 at ω_0 . Inductance L_2 can be represented as a parallel combination of

L'_2 and L''_2 , where L''_2 and C''_2 resonate at ω_0 , satisfying (3a). The residual inductance L'_2 is given by (3b)

$$L''_2 = \frac{1}{\omega_0^2 C''_2} \quad (3a)$$

$$L'_2 = \frac{L_2 L''_2}{L''_2 - L_2}. \quad (3b)$$

Next, as shown in Fig. 2(d) and (e), L'_2 in parallel with $(R'_L + R_2)''$ is referred to the primary of the transformer. As a result, the impedances are scaled down by a factor of $(n/k_m)^2$. On the primary side, the parallel-to-series conversion of $L'_2 * (k_m^2/n^2)$ and $(R'_L + R_2)'' * (k_m^2/n^2)$ gives L_{eff} and R_{eff} , as indicated in Fig. 2(f). Resistance R_{eff} represents the loading on the transformer primary and can be obtained using

$$Q_{\text{pri}} = \frac{(R'_L + R_2)''}{\omega_0 L'_2} \quad (4a)$$

$$R_{\text{eff}} = (R'_L + R_2)'' * \frac{\frac{k_m^2}{n^2}}{(1 + Q_{\text{pri}}^2)} = R_{L,\text{eff}} + R_{2,\text{eff}}. \quad (4b)$$

Furthermore, R_{eff} can be split as $R_{L,\text{eff}}$ and $R_{2,\text{eff}}$ as given in (5a) and (5b). $R_{L,\text{eff}}$ represents the effective load resistor, and $R_{2,\text{eff}}$ represents the component of R_2 referred to the primary

$$R_{L,\text{eff}} = R_L * \left(\frac{1 + Q_s^2}{1 + Q_p^2} \right) * \frac{\frac{k_m^2}{n^2}}{(1 + Q_{\text{pri}}^2)} \quad (5a)$$

$$R_{2,\text{eff}} = R_2 * (1 + Q_s^2) * \frac{\frac{k_m^2}{n^2}}{(1 + Q_{\text{pri}}^2)}. \quad (5b)$$

Finally, to minimize the insertion loss, the value of $R_{L,\text{eff}}$, which represents the effective load resistance at the primary, should be much larger than $R_{2,\text{eff}}$ and R_1 . Since all three resistances appear in series, the IL is minimized when the factor M is maximized, where

$$M \triangleq \frac{R_{L,\text{eff}}}{R_{2,\text{eff}} + R_1}. \quad (6)$$

Several key insights can be obtained from (6). First, if C_2 is minimized, then, from (1), R'_L is maximized, and the loss associated with R_2 is minimized. However, small values of C_2 imply lower values of $R_{L,\text{eff}}$ in (5a), thus increasing the loss from R_1 . Therefore, an optimum value of C_2 exists for which the loss due to both R_1 and R_2 is minimized. This optimum value of C_2 is found by finding the maxima in M . The insertion loss can be estimated by using (7) as

$$IL = 10 * \log \left[\frac{R_{L,\text{eff}}}{R_{L,\text{eff}} + R_{2,\text{eff}} + R_1} \right]. \quad (7)$$

To verify the aforementioned model, the insertion-loss variation as a function of C_2 was characterized using two methods: first, by using the analytic expression derived in (7) and, second, by running Spectre circuit-level s -parameter simulations. As a test case, a transformer with a turns ratio of 2 was designed, and High-Frequency Structural Simulator electromagnetic simulations were run to generate s -parameters for the structure. A lumped-element model ($L_1 = 362$ pH, $R_1 = 0.7$ Ω , $L_2 = 1.25$ nH, $R_2 = 1.87$ Ω , and $k_m =$

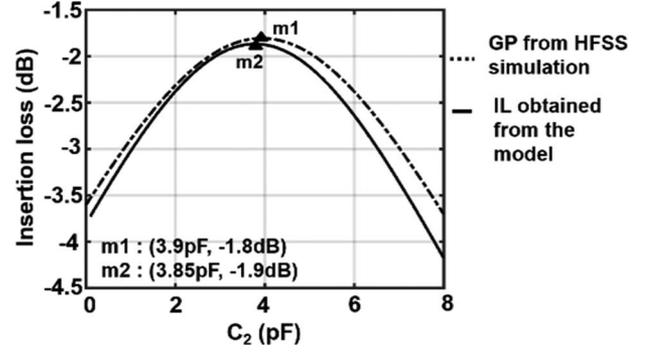


Fig. 3. Plot of transformer insertion loss versus C_2 .

0.65) is extracted from the s -parameters. Using the lumped-element parameters and (1)–(7), the IL is computed as a function of C_2 . Next, using the transformer's s -parameters in circuit simulations, the power gain (G_p) is plotted as a function of C_2 . The G_p [5] of a two-port network does not include the loss due to source mismatch and therefore is an accurate estimate of the inherent insertion loss of the network. The two resulting graphs are shown in Fig. 3. This analysis was carried out at a center frequency of $f_0 = 2$ GHz with $R_L = 50$ Ω . From the analytical model, the optimum C_2 was found to be 3.85 pF and resulted in the minimum insertion loss of 1.9 dB. Based on circuit simulations, the optimum C_2 is 3.9 pF with a minimum insertion loss of 1.8 dB. Thus, the analytical model and simulations show good agreement.

In a transformer, the parameters k_m and n , are lossless; however, they influence the IL as seen by (6) and (7). One may observe that the IL is dependent on the ratio n/k_m (as opposed to the individual terms). Therefore, a low k_m does not necessarily result in higher IL. However, a low k_m results in a large primary leakage inductance $L_1(1 - k_m^2)$, which increases the Q of the primary branch, resulting in a lower bandwidth.

III. TUNABLE MATCHING NETWORK DESIGN

At radio frequencies, a higher Q can be achieved by implementing a variable capacitance as compared to varying the inductance. Therefore, traditionally, “tuning” techniques have focused on matching network with varactors or switched-capacitor banks. In a transformer-based matching network, C_2 is the only capacitor available for tuning. However, as described in Section II, C_2 should be optimized to reduce the transformer insertion loss. (For the transformer described in Section II, if C_2 is a variable capacitance with a range from 0 to 7 pF, the input impedance can be tuned from 4 to 17 Ω . However, the total insertion loss would be greater than 3.5 dB, even with an ideal switch.) Therefore, for impedance tuning, we propose a two-stage tunable matching network realized with a transformer and a C-L-C Pi-network.

The circuit of the proposed TPMN is shown in Fig. 4(a). The inductor L_3 , along with the two variable capacitors C_{var} and C_{par} , makes up the Pi-network. The TPMN is designed to absorb a fixed output capacitance associated with the PA into the matching network while presenting the PA with a load impedance Z_{in} , which is real and variable. Two sets of variable-capacitor banks are used in this design: C_{var} tunes the $\text{Re}(Z_{\text{in}})$, while the C_{par} capacitor bank is used to maintain resonance at the carrier frequency.

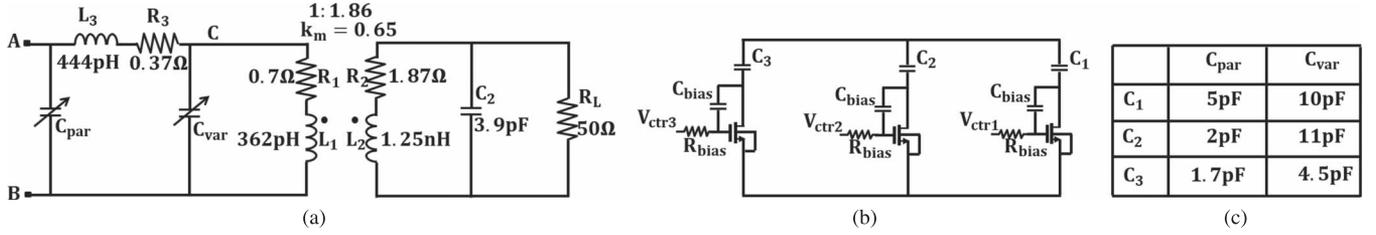


Fig. 4. (a) Circuit implementation of TPMN. (b) Variable capacitor implemented as a switched-capacitor bank. (c) Values of C_1 , C_2 , and C_3 that implement capacitors C_{var} and C_{par} .

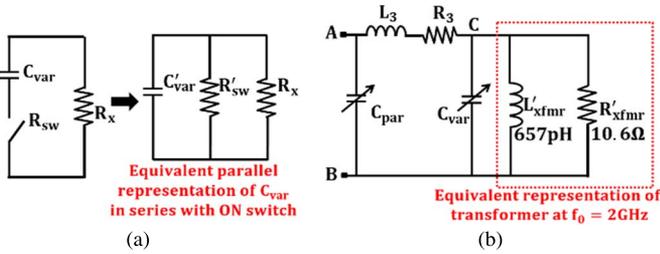


Fig. 5. (a) Transformation to intuitively understand the impact of switch resistance on the loss in Pi-network. (b) Simplified representation of the TPMN.

To provide a framework of discussion for the TPMN design methodology, a single shared Bluetooth/Wi-Fi PA load is explored. Wi-Fi and Bluetooth operate in the same frequency band; however, they have significantly different P_{OUT} requirements. For a 24-dBm Wi-Fi PA (operating off of a 1.8-V supply) and a 12-dBm Bluetooth PA (operating from a 1-V supply), the r_{opt} is 20 and 4 Ω , respectively, assuming a 2-dB matching network loss and class-A operation. Assuming a 50- Ω antenna, the matching network must provide a maximum transformation ratio of 50:4 while supporting a tuning range that spans 4 to 20 Ω .

For a transformation ratio as large as 10 \times , an important question is how to distribute this transformation between the transformer and the Pi-match. For example, should the transformer step down the 50- Ω antenna impedance to 20 Ω , 4 Ω , or some value in between? This question is explored in the context of the impact on the overall TPMN insertion loss from the transformation ratio distribution between the transformer and the Pi-match; this is analyzed in the next section.

A. Minimizing Switched-Capacitor Bank Losses

Any tuning capacitor bank requires a series switch that has a finite on-resistance, R_{SW} . A convenient and intuitive way to model the impact of R_{SW} on the total matching network insertion loss is by transforming the series resistance to an equivalent shunt resistance R_{SW}' , shown in Fig. 5(a). The signal loss through R_{SW}' depends on the R_X presented by the transformer at the loading terminal of the Pi-network. A large transformer turns ratio will result in a low value of R_X . The key observation is that, by increasing the ratio R_{SW}'/R_X , the loss due to R_{SW}' will be proportionally reduced. In summary, minimizing the insertion loss associated with the capacitor bank switches in series with C_{var} requires a large turns ratio in the transformer.

B. Minimizing Loss Due to the Series Inductor L_3

The loss introduced by the finite Q of L_3 can be modeled with a series resistance R_3 . Unlike the loss in the switched-capacitor bank of C_{var} , R_3 appears in series with R_X' (where

R_X' is the resistance resulting from the parallel-to-series conversion of C_{var} and R_X). Resistors R_3 and R_X' effectively form a voltage divider; therefore, to minimize the loss through R_3 , R_X' must be made as large as possible. Alternatively, for a given R_X' , R_3 needs to be smaller. Assuming that the inductor Q cannot be improved, R_3 can be reduced by using a lower value for L_3 . However, reducing the value of L_3 minimizes the impedance tuning range of the Pi-matching network. Therefore, increasing R_X' is preferable, which is equivalent to maximizing the value of R_x .

To reduce the loss due to the low quality factor of C_{var} , resistance R_x of the Pi-network needs to be minimized, and to reduce the loss due to series resistance R_3 , R_x needs to be maximized. Given these two conflicting requirements, an empirical choice for the “optimal” R_x is the geometric mean of the high and the low end of the desired impedance tuning range

$$R_X = \sqrt{R_{IN_MAX} * R_{IN_MIN}}. \quad (8)$$

The transformer in the TPMN converts the 50- Ω antenna resistance to the optimal R_x to minimize the Pi-network loss. For the Wi-Fi/Bluetooth PA output matching network considered in this brief, the desired range of load resistance presented to the PA is 4 to 20 Ω . Therefore, from (8), the transformer should convert the 50- Ω antenna impedance to 9 Ω at the Pi-network input.

C. Example Implementation

To demonstrate the concepts described in this brief, a matching network was designed using a transformer with primary and secondary inductances of 360 pH and 1.25 nH, respectively. A stacked transformer was built using the two thick metal layers available in the 40-nm metal stack. The transformer has a turns ratio of 1:1.86 and a magnetic-coupling coefficient of $k_m = 0.65$. Based on the criteria described in Section II, $C_2 = 3.9$ pF. The insertion loss of the transformer is 1.8 dB. As shown in Fig. 5(b), the transformer presents the Pi-network with a load $R_x = 10.6$ Ω .

The TPMN either steps up or steps down this impedance based on the C_{var} capacitor setting. The switches in the variable capacitors C_{var} and C_{par} are implemented using series n-channel MOS (NMOS) transistors. To minimize the insertion loss and have good linearity, the NMOS switches are placed on the ground side of the capacitors. To improve the reliability of the switch in the OFF state, capacitor C_{bias} is used to couple a portion of the ac signal at the drain of the switch onto the gate. Sizing the switch entails a tradeoff between insertion loss and tuning range. Ideally, the switches should be sized as large as permissible to minimize the insertion loss. However, increasing the switch size increases the parasitic capacitance, thus lowering both the tuning range and the minimum value of C_{var} and C_{par} . Based on these tradeoffs, a switch with a

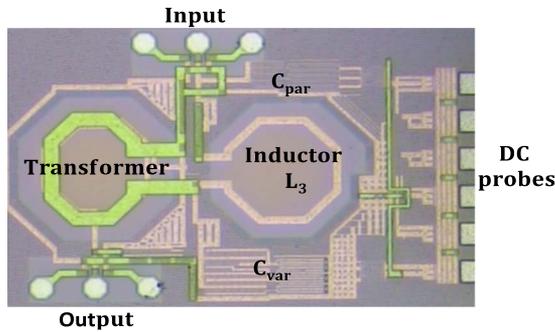


Fig. 6. TSMC 40-nm chip micrograph.

TABLE I
REAL INPUT IMPEDANCE AND INSERTION LOSS VERSUS C_{var} ,
BOTH SIMULATED AND MEASURED, AT $f_0 = 2$ GHz

C_{var} values	Simulation		Measurement	
	Real(Z_{in}) (Ω)	IL (dB)	Real(Z_{in}) (Ω)	IL (dB)
$C_1 + C_2$	3.8	-3	17	-9.8
C_2	9.2	-2.4	7	-7
C_3	15.5	-2.3	10.1	-4.7
None	21	-2.3	20.2	-4.3

minimum length and $W = 2.5$ mm was used. Capacitors C_{var} and C_{par} are implemented as 3-b capacitor banks.

IV. SIMULATION AND MEASUREMENT RESULTS

Based on the techniques described in this brief, a chip was fabricated in a 40-nm CMOS low power (LP) process. The chip micrograph is shown in Fig. 6. The TPMN was characterized using two-way on-chip wafer probing using RF ground-signal-ground probes.

With 6 b of digital control, the TPMN can be set to 64 unique complex impedance values. However, only those capacitor combinations that result in real Z_{in} values, at a frequency of 2 GHz, are listed in Table I. In simulation, Z_{in} can be modulated from 3.8 to 21 Ω while maintaining phase (Z_{in}) ~ 0 over a 100-MHz bandwidth. The corresponding insertion loss and C_{var} settings are also listed in Table I.

A discrepancy between simulated and measured results can be observed in Table I. This was due to a 180- μ m-long on-chip ground trace that introduced 180 pH of parasitic inductance (L_{GND}) in series with C_{var} and C_{par} . To validate this, the TPMN was simulated with the additional parasitics modeled, as shown in Fig. 7(d). Plots comparing the measured and simulated (with L_{GND}) insertion loss, and the magnitude and phase of Z_{in} are shown in Fig. 7(a)–(c). In each of these plots, two cases are considered: all the switches OFF and all the switches ON. The close correlation between measured and simulation (with L_{GND}) data confirms that the parasitic ground inductance was indeed the primary reason for the performance degradation.

Considering only those measured values that resulted in a smaller insertion loss, it could be said that, in measurement, an impedance variation of 10 to 20 Ω was obtained. With careful layout and parasitic estimation, an impedance tuning range of 1:5 can be achieved.

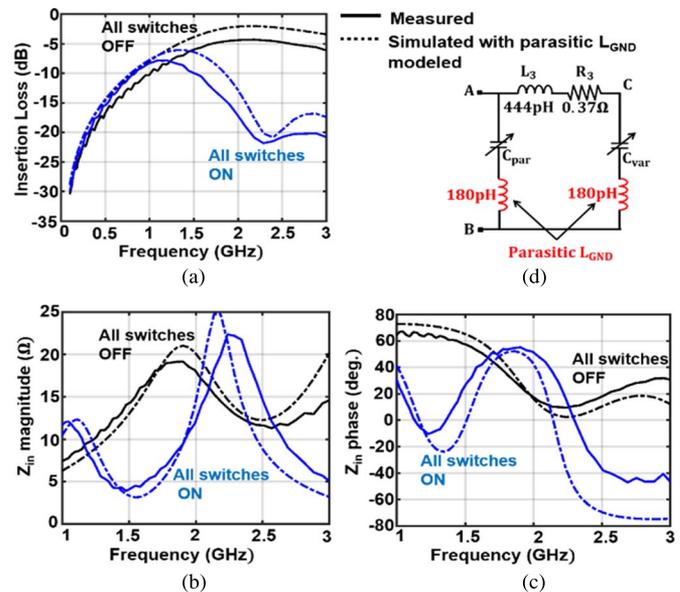


Fig. 7. Comparison of measured versus simulated results using the model with ground parasitic inductance. (a) Insertion loss plot. (b) Magnitude of Z_{in} . (c) Phase of Z_{in} . (d) Model developed to explain the measured results.

V. CONCLUSION

The need for programmable radios and power amplifiers—both in terms of power level and frequency—will continue to grow with the increased demand for portable notebook and smartphone devices. Fully integrated tunable front-end components, such as the TPMN described in this brief, will potentially enable such systems. In prior literature, tuning structures have been proposed using silicon-on-insulator and micro-electromechanical systems technology; however, relatively few implementations have been presented in standard CMOS technologies. The tunable on-chip matching network proposed in this brief was designed in a standard CMOS process and can potentially achieve an impedance tuning range of 1:5 while supporting a total transformation ratio as large as 1:12. In addition, strategies to minimize the insertion loss based on analytical models have been discussed.

REFERENCES

- [1] K. R. Boyle *et al.*, "Analysis of mobile phone antenna impedance variations with user proximity," in *IEEE Trans. Antennas Propag.*, vol. 55, no. 2, pp. 364–372, Feb. 2007.
- [2] W. C. Edmund Neo *et al.*, "Adaptive multi-band multi-mode power amplifier using integrated varactor-based tunable matching networks," *IEEE J. Solid-State Circuits*, vol. 41, no. 9, pp. 2166–2176, Sep. 2006.
- [3] J. Fu *et al.*, "Improving power amplifier efficiency and linearity using a dynamically controlled tunable matching network," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 12, pp. 3239–3244, Dec. 2008.
- [4] P. Sjoblom *et al.*, "An adaptive impedance tuning CMOS circuit for ISM 2.4-GHz band," in *IEEE Trans. Circuits Syst. I Reg. Papers*, vol. 52, no. 6, pp. 1115–1124, Jun. 2005.
- [5] A. M. Niknejad, *Electromagnetics for High-Speed Analog and Digital Communication Circuits*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [6] J. R. Long, "Monolithic transformers for silicon RF IC design," *IEEE J. Solid-State Circuits*, vol. 35, no. 9, pp. 1368–1382, Sep. 2000.