A CMOS RF Power Amplifier With Parallel Amplification for Efficient Power Control

Alireza Shirvani, *Student Member, IEEE*, David K. Su, *Member, IEEE*, and Bruce A. Wooley, *Fellow, IEEE*

***Abstract*—This paper introduces a CMOS radio-frequency (RF) power amplifier that uses parallel amplification to provide high efficiency over a broad range of output power. Three binary-weighted class-F unit amplifiers act in conjunction with an efficient power-combination network to provide a digital-to-analog conversion between a 3-b control signal and the amplitude of the output RF signal. The power-combination network is based on quarter-wavelength transmission lines that also serve as class-F harmonic terminations. A pMOS switch to the positive supply rail is used to avoid power dissipation when a unit amplifier is shut down. The parallel-amplifier architecture, integrated in a 0.25-μm CMOS technology, occupies an active die area of 0.43 mm^2, operates at 1.4 GHz from a 1.5-V supply, and provides an output power adjustment range of 7–304 mW. The amplifier achieves a maximum power-added efficiency (PAE) of 49% and maintains a PAE of greater than 43% over the range of 100–300 mW.**

***Index Terms*—Class F, CMOS radio frequency, parallel architecture, power amplifier, power combining, power control, transmission line.**

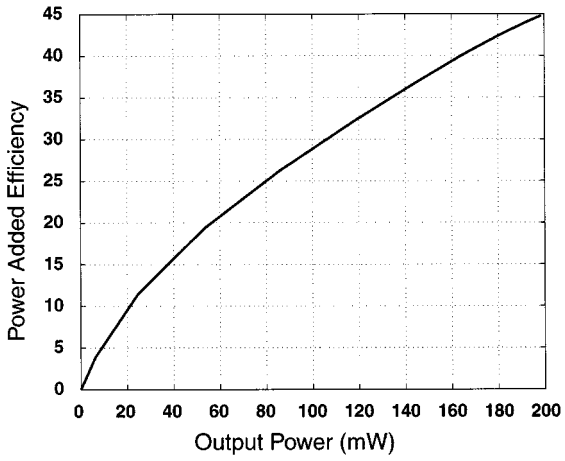

Fig. 1. PAE versus output power for a typical RF power amplifier.

I. Introduction

EMERGING wireless communications infrastructures employ increasingly complex functionalities to optimize the use of bandwidth, minimize the cost, and enhance the portability of wireless personal communication systems. A majority of modern telecommunication protocols require the transmitter power to be adjusted over a wide range. This feature, commonly referred to as power control, ensures that adequate power is received by the base station, while saving power by reducing the transmitted power when the maximum transmitted power is not required and reducing the potential interference in other channels. Many wireless networks require that the transmitted power be adjustable over a range as wide as 20 dB.

The traditional approach to power control is to adjust the output power by varying the bias conditions in the output stage of the power amplifier, adjusting the bias levels in the stage that drives the output stage, or varying the supply voltage. However, if a fixed matching network is used at the output, the power amplifier can only be optimized for a specific operating condition in terms of the bias conditions and the supply voltage. Any deviation from this optimum results in degraded power efficiency [1].

A typical power amplifier only achieves its peak efficiency at a specific output power level, usually the maximum. As illustrated in Fig. 1, the efficiency usually drops rapidly as the power is reduced below the maximum. Since, statistically, the transmitted power in a wireless system is below the maximum for most of the transmission time, the average power efficiency is well below its maximum. To enhance the average power efficiency, it is therefore essential to maintain a high efficiency over a wide range of output power [2].

This paper introduces a power amplifier architecture that mitigates the efficiency drop at lower power levels through parallel amplification [3]. The proposed architecture employs three unit amplifiers, the outputs of which are binary weighted and combined in a power-combination network implemented using quarter-wavelength transmission lines. The capability to turn off each individual unit amplifier, with no penalty on power dissipation, is provided by the addition of pMOS shorting switches.

An experimental prototype of the proposed amplifier has been integrated in 0.25-mm CMOS technology and provides a power adjustment range of 7–304 mW. Tested with a continuous-wave (CW) signal, the amplifier achieves a maximum power-added efficiency (PAE) of 49% and maintains a PAE greater than 43% over 70% of the output power range at 1.4 GHz. The amplifier operates from a 1.5-V supply and provides eight distinct levels of output power.

Section II describes the architecture of the proposed amplifier. An overview of various power amplifier topologies and the merits of class-F amplifiers are first presented, followed by an introduction of the power-combination and parallel-amplifica-

Manuscript received September 21, 2001; revised March 1, 2002. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) under Contract N65326-96-C-8608 and in part by National Semiconductor Corporation.

A. Shirvani is with the Center for Integrated Systems, Stanford University, Stanford, CA 94305-0407 USA, and also with Tavanza Inc., Sunnyvale, CA 94085 USA (e-mail: alireza@par.stanford.edu).

D. K. Su is with Atheros Communications, Sunnyvale, CA 94085-3512 USA.

B. A. Wooley is with the Center for Integrated Systems, Stanford University, Stanford, CA 94305-0407 USA (e-mail: wooley@ee.stanford.edu).

Publisher Item Identifier S 0018-9200(02)04936-3.

tion architectures. Section III describes the design and implementation of the proposed architecture. Experimental results for the integrated prototype are presented in Section IV, and Section V considers the possibility of integrating the power-combining network on the same die as the parallel amplifiers.

II. Architecture

A. Class-F Power Amplifiers

Power amplifiers are often classified as either linear or nonlinear amplifiers. In linear power amplifier topologies, comprising class-A, class-B, class-C, and class-AB amplifiers, the output transistor operates as a current source. In these classes of amplifier, the current through the output device and the voltage across it have sinusoidal or partially sinusoidal waveforms [4]. Amplifiers in this category benefit from relatively good linearity and are traditionally used for applications where the modulated signal has a variable envelope. Examples of systems employing variable envelope modulation are systems operating under the IS-95 code division multiple access (CDMA) standard, using spread-spectrum modulation and the North American Digital Cellular (NADC) standard, using $\pi/4$-QPSK modulation. Unfortunately, linear amplifiers suffer from low power efficiency, due to power loss resulting from the current–voltage overlap in the output transistor [5].

Nonlinear amplifiers mitigate the efficiency degradation by operating the output device as a switch. The switching, under the control of a large driver voltage swing, results in sharp transitions that reduce the current–voltage overlap and, hence, increase the power efficiency. Under such operating conditions, referred to as deep compression, the output voltage is determined by the load impedance and the supply voltage, rather than the amplitude of the driving voltage [5]. Hence, most of the information in the amplitude of the input signal is not transferred to the output and the amplifier's response is extremely nonlinear. Because of this nonlinear operation, such switching amplifiers are best suited to applications that employ a constant envelope modulation scheme in which the signal information is contained in the phase of the carrier. Examples of such systems include those complying with the Global System for Mobile communications (GSM) standard, employing Gaussian minimum-shift keying (GMSK) modulation and Bluetooth, which uses frequency-shift keying (FSK) modulation. These systems can take advantage of the high efficiency achievable with nonlinear amplification.

The category of nonlinear switching power amplifiers is subdivided into several classes, often based on the type of the matching network placed between the load impedance and the output device [5]. Among these, class-E and class-F architectures have been the most widely considered for the implementation of RF power amplifiers [6]–[9]. Although some successful implementations of class-E amplifiers have been reported [8], [9], this class of amplifiers suffers from two major drawbacks. First, theoretical analyses show that the voltage at the output of the device (drain of a MOSFET/MESFET or collector of a bipolar device) during the class-E operation reaches a peak value of 3.6 times the supply voltage. Such a large voltage swing is of increasing concern since the breakdown voltage of a transistor scales along with its dimensions. This constraint forces the supply voltage to values lower than 1 V in an implementation of a class-E architecture in deep-submicron CMOS technologies. Second, class-E operation is based on the elimination of current–voltage overlap region at the device turn-on transient, but it does not provide any solution for the large overlap region at the device turn-off [10].

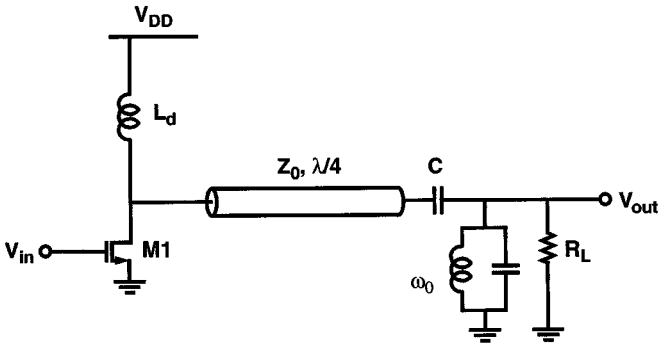

Fig. 2. Class-F power amplifier with multiharmonic termination.

Class-F amplification is based on shaping the waveforms of the current and voltage through and across the output device by adding the higher harmonics of the carrier frequency to the fundamental sinusoidal waveform. Fig. 2 shows a class-F architecture with multiharmonic termination. The bandpass filter at the output establishes a low impedance in shunt with R_L at all of the harmonics of the carrier frequency. At all of the odd harmonics, this low impedance at the load is converted by the quarter-wavelength transmission line to a high impedance at the drain of the output transistor, while at all even harmonics a low impedance is presented at the drain. The addition of odd harmonics to the drain voltage of the output transistor shapes it into an approximate square wave. This results in sharp transitions that reduce the current–voltage overlap and hence increase the drain efficiency of the output stage [11].

The class-F architecture has been adopted in this work. Within the constraints of the design proposed here, it provides superior low-voltage operation. In addition, the power-combination network inherently provides the termination needed for class-F operation, as described in Section II-D.

B. Power Control Through Parallel Amplification

Fig. 1 shows the PAE versus output power for a typical RF power amplifier, as the bias current of the output stage is swept from zero to its maximum. As noted previously, the amplifier achieves high efficiency only at the maximum output power level, and the efficiency falls off rapidly at lower power levels. In systems featuring power control, this efficiency drop results in low average efficiency, since the output power is statistically distributed across the entire power range.

The solution proposed here to address the problem of low average efficiency is to employ several amplifiers in parallel, with different output power levels for amplification in different regions of the power range, as illustrated in Fig. 3. As the output power is reduced below the maximum, the task of amplification

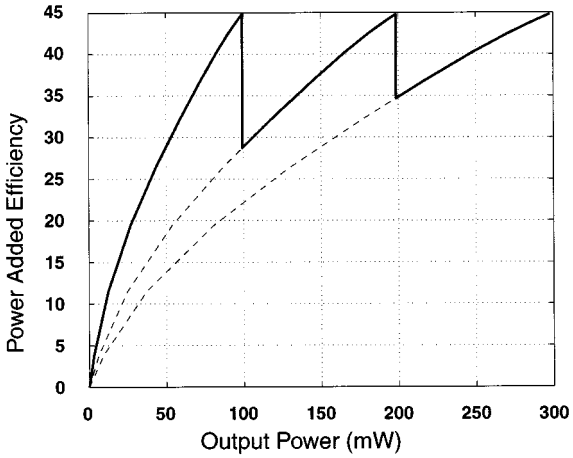

Fig. 3. Power range subdivision.

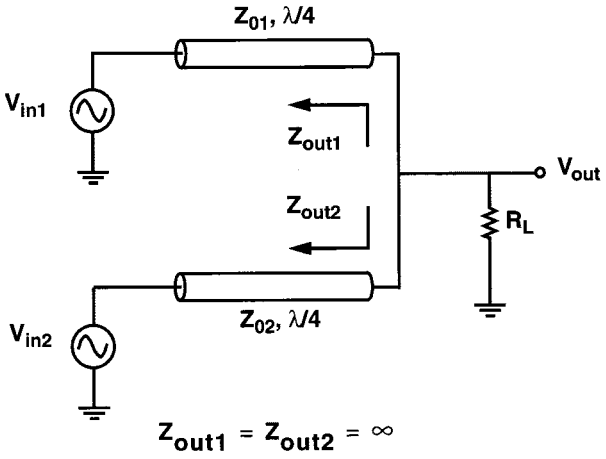

Fig. 4. Power-combination using quarter-wavelength transmission lines.

is provided by another amplifier that maintains a higher efficiency within the new region. To implement such an amplification scheme, it must be possible to switch between different unit amplifiers and combine their outputs efficiently.

C. Power Combination

Various methods have been historically used to split and combine RF signals. Transformers have been widely used as means for combining RF power [5]. However, a typical RF transformer has a minimum insertion loss of 2–3 dB, which is equivalent to a 40%–50% loss of power [12]. Since the loss of power at the output of the amplifier translates directly to the same fraction of loss in power efficiency, such a loss negates the benefit gained from a parallel architecture. Transmission lines, on the other hand, represent a viable approach. Combining architectures based on transmission lines have been employed to achieve higher maximum output power [13], as well as enhanced linearity [14].

This work presents a power-combination scheme based on quarter-wavelength transmission lines. An ideal model for the power-combination architecture is shown in Fig. 4. Two quarter-wavelength transmission lines, driven by ideal voltage sources $V_{\rm in1}$ and $V_{\rm in2}$, are connected together at a resistive load R_L. A quarter-wavelength transmission line with characteristic

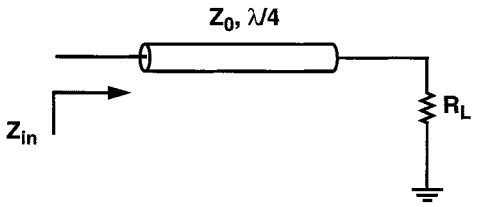

Fig. 5. Impedance conversion by a $\lambda/4$ transmission line.

impedance of Z_0, as shown in Fig. 5, converts a load resistance R_L into an impedance given by

$$Z_{\rm in} = \frac{Z_0^2}{R_L}. \tag{1}$$

Hence, the output impedance of each of the branches in Fig. 4, calculated after shorting the corresponding source, is infinite since the quarter-wavelength transmission lines convert the zero source impedance into an infinite impedance.

Although the power-combining structure is a distributed network, it is a linear system and the voltage at the output terminal can be calculated using superposition. If one of the voltage sources, say, $V_{\rm in2}$, is shorted to ground, the output impedance of its corresponding branch is infinite and the voltage generated by the other source at the load terminal is

$$V_{\rm out}|_{V_{\rm in2}=0} = \frac{R_L}{Z_{01}} V_{\rm in1} \tag{2}$$

where Z_{01} is the characteristic impedance of the first transmission line. A similar relation can be obtained for the output voltage contributed by the second source, when the first source is shorted. It then follows from superposition that the voltage developed at the output equals the sum of the voltages that each sources would generate in the absence of the other, which is the sum of the input voltages, scaled by the ratio of the load impedance to the characteristic impedance of the corresponding line:

$$V_{\rm out} = \frac{R_L}{Z_{01}} V_{\rm in1} + \frac{R_L}{Z_{02}} V_{\rm in2}. \tag{3}$$

The power-combination network performs summation without weighting when lines with equal characteristic impedances are utilized. It also allows the shorting of one of the sources without interfering with the operation of the other branch. However, in practice, the voltage sources driving the transmission lines will have a finite impedance. As the analysis presented in the Appendix indicates, when taking the finite source impedance into account, as illustrated in Fig. 6, the output voltage of the network is given by

$$V_{\rm out} = \frac{Z_{01} Z_{02} R_L}{Z_{01}^2 r_{s2} R_L + Z_{02}^2 r_{s1} R_L + Z_{01}^2 Z_{02}^2} \cdot (Z_{02} V_{\rm in1} + Z_{01} V_{\rm in2}). \tag{4}$$

Shown in Fig. 7 is the loss of power in the combining network plotted as a function of the characteristic impedance of the transmission lines and the source resistance, both normalized to the load resistance, under the assumptions that $Z_{01} = Z_{02} = Z_0$, $r_{s1} = r_{s2} = r_s$, and $V_{\rm in1} = V_{\rm in2}$.

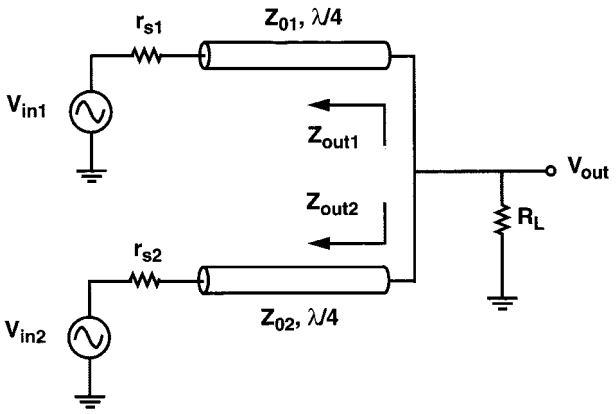

Fig. 6. Source impedance effects on the power-combination network.

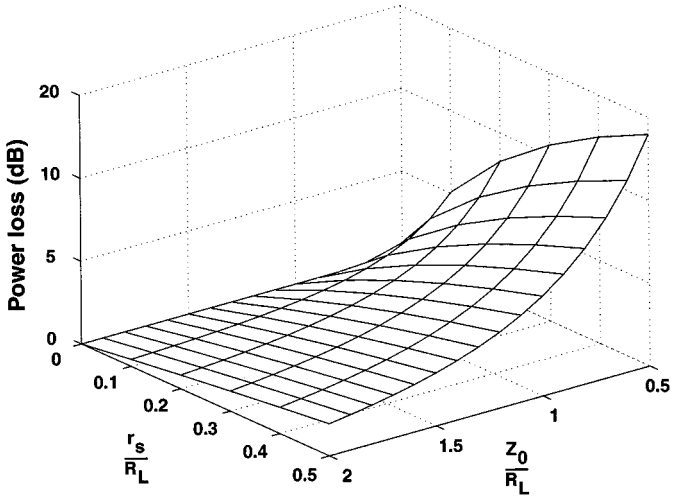

Fig. 7. Loss in power-combination network due to source resistance.

D. Parallel-Amplifier Architecture

Fig. 8 illustrates the proposed parallel-amplifier architecture. Switched nMOS output transistors are used as low-impedance voltage sources driving quarter-wavelength transmission lines. When driven with a large gate-to-source input voltage, these switch transistors present a low average output impedance, also referred to as the effective on-resistance. The required on-resistance is determined using the analytical results of the preceding section and the output transistors are sized accordingly to achieve this resistance.

Each individual amplifier branch in the architecture of Fig. 8 operates as a class-F amplifier since the loading by the other branches is minimal if their corresponding source impedances are sufficiently small. The transmission lines thus serve a dual purpose. They provide class-F harmonic termination as well as a means for power combination. This architecture enables the weighed summation of the output voltages contributed by each individual amplifier. Any of the amplifiers can be turned off without interfering with the operation of the other branches, as long as a low output impedance is maintained by the off amplifiers. This can be achieved by disconnecting the signal drive to the output transistor, but maintaining a high dc voltage on its gate to keep it turned on. However, such an approach results in significant dc loss in the output transistors in the amplifier branches that are off. This problem can be overcome by introducing a pMOS switch that pulls the output of the nonamplifying branches to the positive supply rail, as described in Section III.

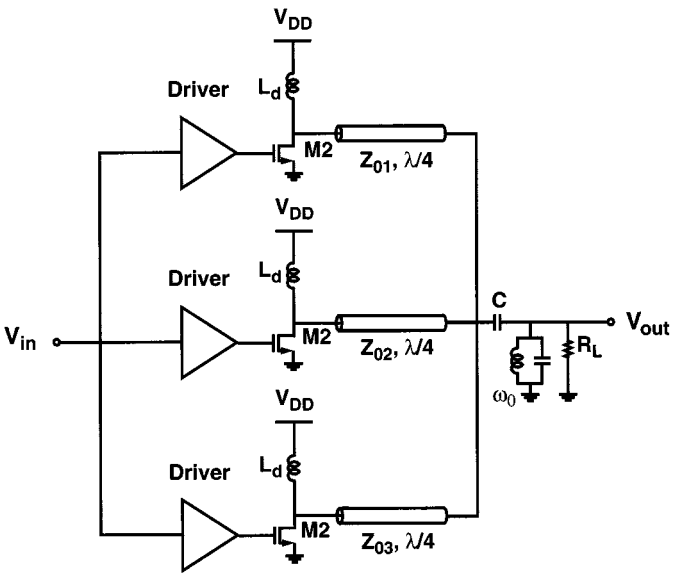

Fig. 8. Parallel-amplifier architecture.

In this application of the parallel-amplifier architecture, a 3-b control signal is provided to the amplifier by the system and is determined based on the feedback from the base station. This is the main distinction between power control and linearization in RF power amplifiers. Although the techniques described herein are, for the most part, suitable for both applications, linearization architectures generally include built-in sensing and control circuitry and the control bandwidth is much higher.

To calculate the output voltage of the architecture shown in Fig. 8, note that the amplitude of the output voltage of a single class-F power amplifier (Fig. 2) is given by [10]

$$V_{\text{out}} = \alpha \cdot \frac{R_L}{Z_0} \cdot V_{DD} \tag{5}$$

where α is the wave-shaping factor and varies between 1 for a purely sinusoidal voltage waveform at the drain and $4/\pi$ for an ideal square waveform. Superposition then leads to the general equation for the output voltage of the parallel architecture:

$$V_{\text{out}} = \alpha \cdot R_L \cdot V_{DD} \cdot \sum_{i=1}^{n} \frac{D_i}{Z_{0i}} \tag{6}$$

where D_i is 1 for an amplifying unit and 0 for a nonamplifying unit. It is apparent from (6) that different output levels can be established by adjusting the characteristic impedances of the different transmission lines.

The architecture of Fig. 8 enables each of the unit amplifiers to operate close to their optimum operating point from a constant supply voltage. Hence, each individual amplifier can be optimized to operate at the maximum PAE. Also, for a specific output power, the use of parallel amplifiers results in lower current levels in each individual amplifier relative to current that would be needed in a single amplifier. Since the losses in the output matching network are proportional to the current

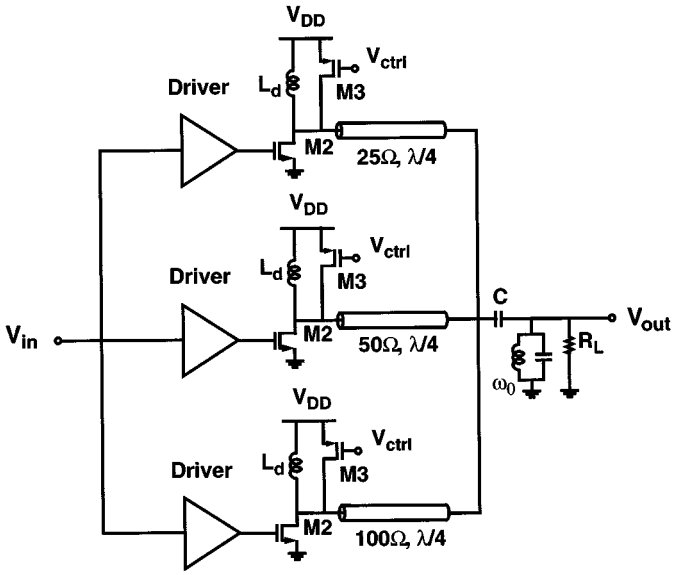

Fig. 9. Implemented parallel-power amplifier.

squared, the total loss is reduced by splitting the current into several branches. The lower current levels also relax the requirement on the individual device sizes. Although the total device size might be the same for the parallel and single amplifiers, the requirements on the driver stage of each unit amplifier in the parallel architecture is less stringent. These features make the architecture specifically attractive for low-voltage applications, where scaled-down voltages demand higher current levels to deliver a specified output power.

III. Implementation

Fig. 9 shows a simplified schematic of the parallel-amplifier architecture implemented in 0.25-μm CMOS technology. Power adjustment is achieved by combining the outputs of three nominally identical unit amplifiers through transmission lines with binary-weighted characteristic impedances, which are set to 25, 50, and 100 Ω. The binary-weighted characteristic impedances result in binary-weighted summation of the output voltages contributed by each individual amplifier, as indicated by (6). This binary weighting provides for eight equally spaced voltage levels at the output of the amplifier. In effect, the system functions as a digital-to-analog converter (DAC) between the 3-b control signal and the envelope of the output RF signal. The values used for the binary-weighted characteristic impedances of the lines are determined by the maximum output power required. In the event that the characteristic impedance values for a target maximum output power cannot be fabricated, it is apparent from (6) that the line impedances can be scaled and additional impedance matching circuitry used to accommodate the maximum power requirement.

As noted in the preceding section, to avoid loading by the amplifiers in the nonamplifying mode, a low impedance must be maintained at the output of these amplifiers. A straightforward means of achieving this is to maintain a large gate bias voltage on the output transistor to keep the device in the triode region as an on switch. However, in this configuration, the transistor draws large dc current, which degrades the overall efficiency of the amplifier. To avoid this problem, a pMOS switch, M3 in Fig. 9, is added at the output of each unit amplifier. When the amplifier is in the amplification mode, the gate of this pMOS transistor is tied to V_{DD} and the device is turned off. When the amplifier switches to nonamplifying mode, M3 is turned on by tying its gate to ground. Thus, a low small-signal impedance is maintained at the output of the unit amplifier without a dc path for current to flow from V_{DD} to ground.

The only drawback to addition of the pMOS output switch is the increase in the parasitic capacitance at the output node. The total capacitance at this node is resonated out by the inductive load of the unit amplifier. For this work, the inductance required to resonate the total parasitic capacitance is implemented using bondwires.

A. Output Stage

A schematic for the unit amplifiers is shown in Fig. 10. The output stage of each unit amplifier is based on a class-F architecture with multiharmonic termination. The output transistor, M2 in Fig. 10, is sized based on the loss analysis described previously and the required power-handling capability. Simulations show that in order to meet the target power objectives for this work, the output device of the MSB amplifier, the amplifier with the highest output power, must have a width of at least 4000 μm with a channel length of 0.25 μm. Although increasing the device size beyond the minimum required reduces its effective on-resistance and improves the drain efficiency of the output stage, the incremental improvement that can be obtained becomes minimal at large device sizes. In addition, a larger output device increases the demands on the driver stage due to increased gate capacitance.

In the experimental prototype, the three unit amplifiers have been designed to be nominally identical. Smaller device sizes could be used in the LSB amplifiers, which would in fact result in improved efficiency for that amplifier, although the efficiency of the LSB amplifier does not have a substantial effect on the overall efficiency. The effective on-resistance of the output transistor, M2, is simulated to be approximately 1.5 Ω. The standby switch, M3, is also sized at 4000 μm, which results in an effective on-resistance of almost twice that of M2. The size of M3 must be comparable to that of M2 in order to ensure that an adequately low output impedance is maintained in the nonamplifying branches.

Although in general it is preferable not to use a resonant circuit at the output of a class-F amplifier because it acts to suppress the higher harmonics, in CMOS implementations the overall performance is usually improved by resonating out the large parasitic capacitance at the drain of the output transistor with a load inductance. The parasitic capacitance at the drain of the output transistor, contributed by both M2 and M3, is resonated out with a bondwire inductor. Due to the low quality factor of the resonant circuit at the drain, the design is tolerant of $\pm 20\%$ variation on the 0.8-nH target inductance of the bondwires.

B. Driver Stage

The driver stage in each of the amplifiers, included in Fig. 10, utilizes three tuned circuits. The impedance formed by these cir-

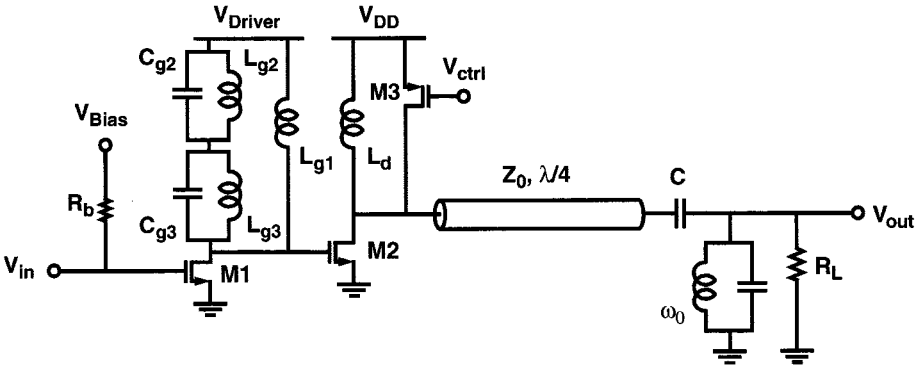

Fig. 10. Unit RF power amplifiers.

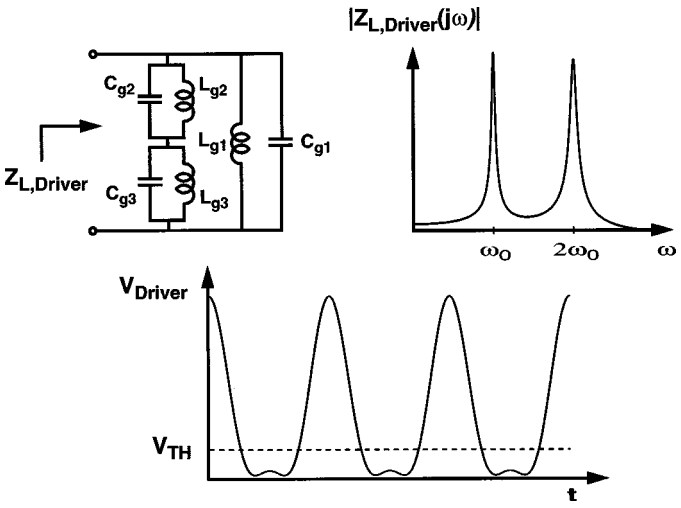

Fig. 11. Second harmonic peaking at the driver stage.

cuits has two resonant frequencies that are distinct from the resonant frequencies of the individual tuned circuits. Although the voltage at the output of the driver swings between ground and $2V_{DD}$, the output transistor M2 switches at its threshold. The efficiency degradation due to the resulting duty cycle, which is larger than 50%, is overcome by setting the two resonant frequencies at the driver output to the fundamental and the second harmonic of the RF signal, as illustrated in Fig. 11. The addition of the second harmonic flattens the lower half of the sinusoidal waveform at the driver output and brings the duty cycle closer to 50%.

The inductors L_{g1} and L_{g2} in the driver are bondwire inductors tuned with external sliding capacitors, while L_{g3} is an on-chip spiral inductor. The spiral inductor is optimized using the analysis and optimization software introduced in [15] and [16]. To reduce the loss due to substrate coupling, patterned ground shields are used underneath the spiral inductors [17].

IV. Experimental Results

The RF power amplifier of Fig. 9 has been integrated in a 0.25-μm CMOS technology. A die micrograph of the prototype amplifier is shown in Fig. 12. Three unit amplifiers are at the top, left, and bottom of the die in this photograph, and the right side is dedicated to ground pads. Several bondwires are used to provide a low-inductance ground connection to the board. An

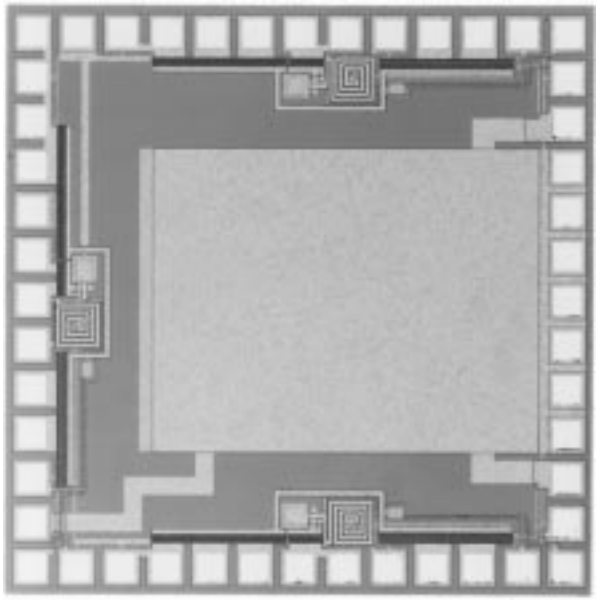

Fig. 12. Die micrograph.

n^+ diffusion shield layer is used underneath all signal routings to reduce the loss due to substrate coupling. The active die area measures 0.43 mm^2.

The die is attached directly to a gold-coated printed circuit board (PCB), using the die-on-board technology, to avoid package parasitics. The transmission lines are implemented using PCB microstrip lines. Dimensions of the transmission lines were determined using electromagnetic field solvers and verified with time division reflectrometry (TDR) measurements on the test structures.

The performance of the design was assessed by driving the amplifier with an RF signal generated using an HP8648C RF signal generator. The output power and spectral characteristics were measured using an HP437B power meter and an HP8563E spectrum analyzer. A custom microstrip power splitter was used to provide three equal and in-phase input signals to the three unit amplifiers.

When operated from a 1.5-V supply, the parallel architecture provides an output power range of 7–304 mW, in eight distinct amplification modes. The measured PAE for different power levels is shown in Fig. 13. The circuit achieves a maximum PAE of 49% at 184-mW output power and maintains a

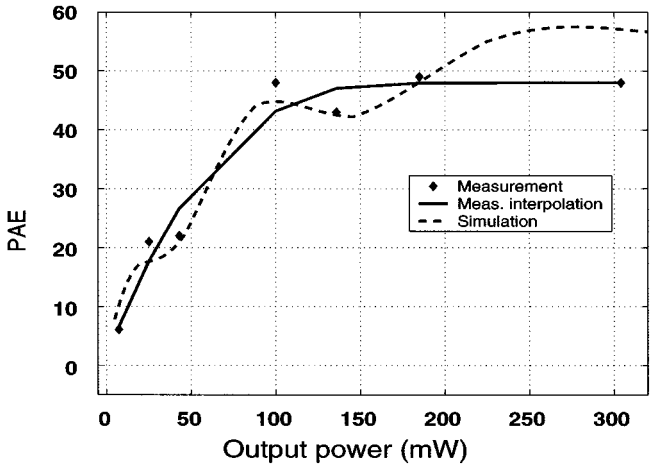

Fig. 13. Measured PAE.

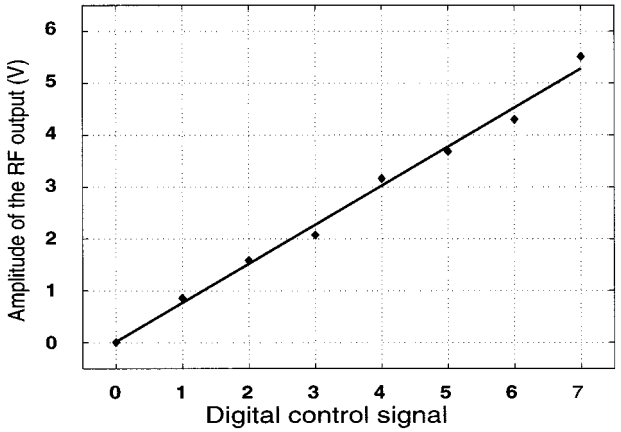

Fig. 14. Measured amplitude of the output RF signal.

PAE higher than 43% for output power in the 100–304 mW range. The seven measurement points correspond to peak output power at seven different power settings. The solid line is an interpolation of the measured points to illustrate the wide high-efficiency region of the PAE versus the output power curve. The dashed line shows the simulation results. As evident from the plot, there is close agreement between the simulation results and the measurements.

The lower PAE at the power levels corresponding to the 001, 011, and 101 values of the binary control signal is a consequence of the lower PAE of the LSB power amplifier. This is the result of using three identical unit amplifiers. The efficiency of the LSB amplifier could be improved by scaling its bias currents and device sizes with the output power level.

Fig. 14 shows the amplitude of the output voltage as a function of the control signal, illustrating a digital-to-analog conversion with a differential nonlinearity (DNL) of 0.6 LSB. The source of nonlinearity can be traced back to nonidealities of the amplifiers, inaccuracies in the characteristic impedance of the transmission lines and different loading, in parallel with the load resistor R_L, seen by different unit amplifiers. One means of improving the integral and differential nonlinearity would be to use seven equally weighted, rather than three binary weighted, parallel amplifiers with a thermometric input code. However, implementing seven parallel amplifiers would severely complicate the layout and packaging.

TABLE I
PERFORMANCE SUMMARY

Output power range	7mW - 304mW
Maximum PAE	49%
Technology	0.25μ CMOS
Supply voltage	1.5V
Frequency	1.4GHz
Active die area	0.43mm^2

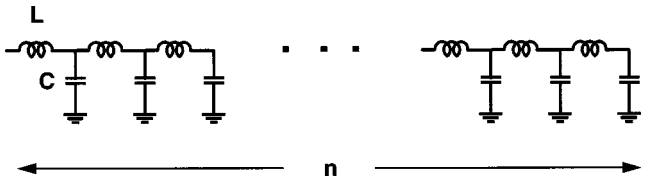

Fig. 15. Lumped model for the transmission lines.

The measured performance of the prototype amplifier is summarized in Table I.

V. On-Chip Power Combination

This section considers the feasibility and tradeoffs of integrating power-combining quarter-wavelength transmission lines on the same die as the amplifiers. Although the length needed to implement such lines on a PCB decreases at the increasingly higher carrier frequencies employed in wireless communications systems, integration of the lines with the rest of the RF circuitry remains of interest.

On-chip transmission lines and waveguides have been fabricated in CMOS technology [18]. However, a quarter-wavelength transmission line on a silicon substrate is 27-mm long at 1.4 GHz. Such a long line is not only practically infeasible to fabricate, but also suffers from an insertion loss as large as 6–7 dB. To alleviate these problems, a multisection line composed of discrete components is proposed.

As illustrated in Fig. 15, a transmission line can be approximated with an LC ladder [10]. To emulate a quarter-wavelength transmission line at carrier frequency of f_0 and with characteristic impedance of Z_0 using the network shown in Fig. 15

$$\tau_D = n\sqrt{LC} = \frac{1}{4f_0} \quad \text{and} \quad Z_0 = \sqrt{\frac{L}{C}} \tag{7}$$

where n is the number of segments. Solving for the component values of the segments in Fig. 15

$$L = \frac{Z_0}{4nf_0} \quad \text{and} \quad C = \frac{1}{4nf_0Z_0}. \tag{8}$$

Simulations indicate that the response of a line consisting of ten sections approximates that of an actual transmission line to within 5%. For a 50-Ω line at 1.4 GHz, a model consisting of ten sections requires inductor and capacitor values of 0.893 nH and 357 fF, respectively. Some final adjustments must be made to inductance values to account for mutual coupling of adjacent

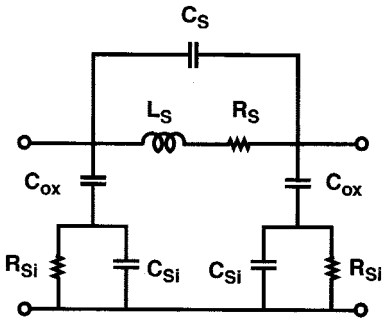

Fig. 16. Lumped physical model of a spiral inductor on silicon.

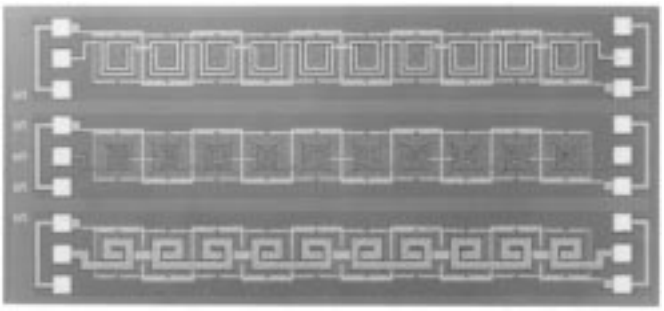

Fig. 17. Die micrograph of the test structures.

inductors. However, the coupling factor is usually small for reasonably distanced inductors. Inductors and capacitors with the aforementioned values are quite feasible in present VLSI fabrication technologies [10]. In fact, a spiral inductor is sufficient to implement both elements. The capacitance can be obtained with the bottom-plate parasitic capacitance of the spiral, as indicated in the inductor model shown in Fig. 16.

To test the idea of building on-chip transmission lines using *LC* ladders, several lines with different characteristic impedances were fabricated in a 0.25-μm CMOS technology. A die micrograph of the test structures appears in Fig. 17. The spiral inductors with the inductance and bottom-plate capacitance specified above have been optimized for the least amount of loss, using the analysis and optimization software described in [15] and [16], and patterned ground shields are used to reduce the substrate loss [17]. Each fabricated line occupies a die area of 200 μm $\times$ 2000 μm and is approximately 14 times shorter than its distributed counterpart.

The characteristic impedances measured for the 25-, 50-, and 100-Ω lines are plotted in Fig. 18 as functions of frequency. The measured impedances are very close to expected values. However, the measured losses of the lines, plotted in Fig. 19, are larger than the expected, although comparable to results for the distributed lines [18]. A close inspection of these results shows that although the loss at low frequencies (up to 200 MHz) matches that of the simulations, the loss increases rapidly as the frequency rises. There appear to be two major sources of loss at high frequencies. First, the skin effect increases the series resistance of each ladder segment, which is modeled by the spiral inductor model shown in Fig. 17 [19]. Although the vertical skin effect was accounted for in simulations, another effect, referred to as current crowding [20], results in the nonuniform

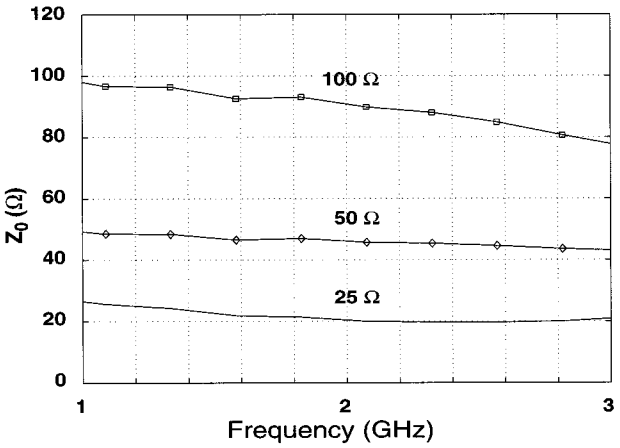

Fig. 18. Measured characteristic impedance of the lines.

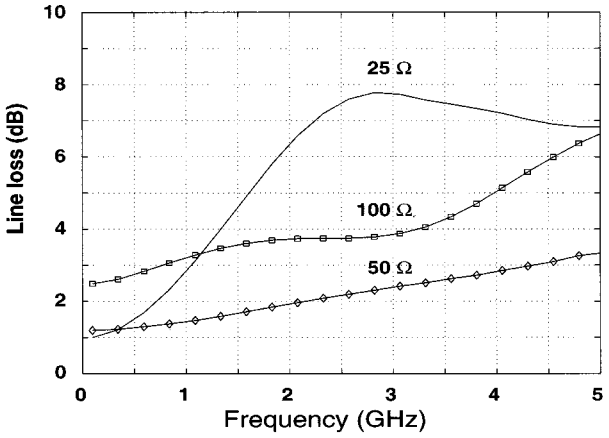

Fig. 19. Measured loss in transmission lines.

distribution of current across the width of the conductor and further increases the series resistance. Second, substrate coupling contributes to significant loss of power, resulting in a relatively small parallel resistance in the model of Fig. 16. The patterned ground shield seems to be less effective at higher frequencies, as the loss is the result of field penetrating the substrate.

Simulations using the experimentally extracted models for the ladder segments indicate that the maximum PAE of the parallel architecture presented in this paper would degrade down to the range of 10%–15% if the power-combining transmission lines were implemented using on-chip *LC* ladders. These results suggest that such an approach may not be suitable for RF power-amplifier applications. However, the good control over characteristic impedance and delay that appears to be achieveable with these on-chip lines might make them attractive in applications where the power loss is not of primary concern. Moreover, the approach might be attractive in a technology that includes a low-loss semi-insulating substrate.

VI. Conclusion

It has been shown that a parallel-amplifier architecture can be used to obtain efficient power control in RF power amplifiers. A parallel architecture provides high efficiency by reducing the current levels in individual branches, which is especially important in low-voltage applications. By subdividing the power

range into separate regions that are handled by parallel amplifiers, these amplifiers can be optimized for different sections of the power range, resulting in high efficiency over a much wider range of output power than can be achieved with a single power amplifier. In an experimental prototype, binary-weighted unit amplifiers have been used to convert a 3-b control signal to the amplitude of the output RF signal, providing a broad power adjustment range. The prototype was implemented in a 0.25-μm digital CMOS technology and is shown to operate from a supply voltage as low as 1.5 V without sacrificing the PAE.

APPENDIX

This Appendix presents a derivation of the result in (4) for the output voltage of a power-combination network, driven by sources with finite impedance, shown in Fig. 6.

The output voltage of the power-combination network can be calculated using the superposition principle. With one of the sources in Fig. 6 set to zero, the output voltage is calculated in presence of a single source. Although the output voltage of the network in Fig. 6 can be determined by analyzing the standing electromagnetic waves in the transmission lines, a simpler analysis can be performed based on the fact that the power is conserved within a lossless transmission line. Equating the input and output power of the first transmission line results in

$$\frac{1}{2Z_{\text{in1}}}\left(\frac{Z_{\text{in1}}V_{\text{in1}}}{Z_{\text{in1}}+r_{s1}}\right)^2=\frac{V_{\text{out}}^2}{2\left(R_L||Z_{\text{out2}}\right)}. \tag{9}$$

The input impedance of the first transmission line Z_{in1} is

$$Z_{\text{in1}}=\frac{Z_{01}^2}{R_L||Z_{\text{out2}}} \tag{10}$$

and the output impedance of the shorted branch can be expressed as

$$Z_{\text{out2}}=\frac{Z_{02}^2}{r_{s2}}. \tag{11}$$

Substituting (10) and (11) in (9) and solving for V_{out} results in

$$V_{\text{out}|_{V_{\text{in2}}=0}}=\frac{Z_{01}Z_{02}^2R_L}{Z_{01}^2r_{s2}R_L+Z_{02}^2r_{s1}R_L+Z_{01}^2Z_{02}^2}\cdot V_{\text{in1}} \tag{12}$$

and by symmetry

$$V_{\text{out}|_{V_{\text{in1}}=0}}=\frac{Z_{02}Z_{01}^2R_L}{Z_{02}^2r_{s1}R_L+Z_{01}^2r_{s2}R_L+Z_{02}^2Z_{01}^2}\cdot V_{\text{in2}}. \tag{13}$$

It follows from (12) and (13), according to superposition, that the output voltage of the network in presence of both sources is

$$V_{\text{out}}=\frac{Z_{01}Z_{02}R_L}{Z_{01}^2r_{s2}R_L+Z_{02}^2r_{s1}R_L+Z_{01}^2Z_{02}^2}\cdot\left(Z_{02}V_{\text{in1}}+Z_{01}V_{\text{in2}}\right). \tag{14}$$

ACKNOWLEDGMENT

The authors would like to thank DARPA for its support of this project, and the National Semiconductor Corporation for fabricating the experimental circuits.

REFERENCES

[1] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*. Norwood, MA: Artech, 1999.
[2] D. K. Su and W. J. McFarland, "An IC for linearizing RF power amplifiers using envelope elimination and restoration," *IEEE J. Solid-State Circuits*, vol. 32, pp. 544–550, Apr. 1997.
[3] A. Shirvani, D. K. Su, and B. A. Wooley, "A CMOS RF power amplifier with parallel amplification for efficient power control," in *IEEE Int. Solid State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2001, pp. 157–158.
[4] B. Razavi, *RF Microelectronics*. Upper Saddle River, NJ: Prentice Hall, 1998.
[5] H. L. Kraus, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*. New York: Wiley, 1980.
[6] T. C. Kuo and B. Lusignan, "A 1.5 W class-F RF power amplifier in 0.2 μm CMOS technology," in *IEEE Int. Solid State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2001, pp. 157–158.
[7] F. Fortes and M. J. do Rosario, "A second harmonic class-F power amplifier in standard CMOS technology," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 1216–1220, June 2001.
[8] T. Sowlati, C. Salama, J. Sitch, G. Rabjohn, and D. Smith, "Low-voltage high-efficiency GaAs class-E power amplifier for wireless transmitters," *IEEE J. Solid-State Circuits*, vol. 32, pp. 544–550, Apr. 1997.
[9] K. C. Tsai and P. R. Gray, "A 1.9-GHz 1-W CMOS class-E power amplifier for wireless communications," *IEEE J. Solid-State Circuits*, vol. 34, pp. 962–970, July 1999.
[10] T. H. Lee, *The Design of CMOS Radio Frequency Integrated Circuits*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[11] F. H. Raab, "Maximum efficiency and output of class-F power amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 1162–1166, June 2001.
[12] J. Long, "Monolithic transformers for silicon RF IC design," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1368–1382, Sept. 2000.
[13] I. Yoshida *et al.*, "A 3.6-V 4-W 0.2-cc Si power-MOS-amplifier module for GSM handset phones," in *IEEE Int. Solid State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 1998, pp. 50–51.
[14] W. H. Doherty, "A new high-efficiency power amplifier for modulated waves," *Proc. IRE*, vol. 24, no. 9, pp. 1163–1182, Sept. 1936.
[15] A. M. Niknejad and R. G. Meyer, "Analysis, design and optimization of spiral inductors and transformers for Si RF ICs," *IEEE J. Solid-State Circuits*, vol. 33, pp. 1470–1481, Oct. 1998.
[16] M. Hershenson, S. Mohan, S. Boyd, and T. H. Lee, "Optimization of inductor circuits via geometric programming," in *36th Design Automation Conf.*, 1999, pp. 994–998.
[17] C. P. Yue and S. S. Wong, "On-chip inductors with patterned ground shields for Si-based RF ICs," *IEEE J. Solid-State Circuits*, vol. 33, pp. 743–751, May 1998.
[18] B. Kleveland, T. H. Lee, and S. S. Wong, "50-GHz interconnect design in standard silicon technology," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 1998, pp. 1913–1916.
[19] C. P. Yue, C. Ryu, J. Lau, T. H. Lee, and S. S. Wong, "A physical model for planar spiral inductors on silicon," in *Int. Electron Devices Meeting Tech. Dig.*, Dec. 1996, pp. 155–158.
[20] W. B. Kuhn and N. M. Ibrahim, "Analysis of current crowding effects in multiturn spiral inductors," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 31–38, Jan. 2001.

Alireza Shirvani (S'97) was born in Tehran, Iran, in 1975. He received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1997 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1999, where he is currently working toward the Ph.D. degree in electrical engineering. His doctoral research focuses on design and control of CMOS RF power amplifiers.

During the summer of 1998, he was with Level One Communications, San Francisco, CA, where he designed a sigma–delta modulator for a cordless phone transceiver. In the summer of 1999, he was with Lucent Technologies (Bell Labs), where he investigated linearization techniques for RF power amplifiers. He is currently with Tavanza Inc., Sunnyvale, CA. His interests include RF and mixed-signal integrated circuit design for wireless applications.

David K. Su (S'81–M'94) was born in Kuching, Malaysia, in 1961. He received the B.S. and M.E. degrees in electrical engineering from the University of Tennessee, Knoxville, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1994.

From 1985 to 1989, he worked as an IC Design Engineer with Hewlett-Packard Company, Corvallis, OR, and Singapore, where he designed full-custom and semi-custom application-specific integrated circuits. From 1989 to 1994, he was a Research Assistant with the Center for Integrated Systems, Stanford University. From 1994 to 1999, he was a Member of Technical Staff with the High Speed Electronics Department of Hewlett Packard Laboratories, Palo Alto, CA, where he designed CMOS analog, RF, and mixed-signal ICs for wireless communications. Since February 1999, he has been with Atheros Communications, Sunnyvale, CA, engaging in the design and development of integrated CMOS transceivers for wireless LAN. His research interests include the design of RF, analog, mixed-signal, and data conversion circuits. He has also been a Consulting Assistant Professor at Stanford University since 1997.

Bruce A. Wooley (S'64–M'70–SM'76–F'82) was born in Milwaukee, WI, in 1943. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1966, 1968, and 1970, respectively.

From 1970 to 1984, he was a Member of the Research Staff at Bell Laboratories, Holmdel, NJ. In 1980, he was a Visiting Lecturer at the University of California, Berkeley. In 1984, he joined the faculty at Stanford University, Stanford, CA, where he is the Robert L. and Audrey S. Hancock Professor of Engineering and the Chairman of the Department of Electrical Engineering. At Stanford, he has also served as the Senior Associate Dean of Engineering and the Director of the Integrated Circuits Laboratory. His research is in the field of integrated circuit design, where his interests include oversampling A/D and D/A conversion, low-power mixed-signal circuit design, circuit design techniques for video and image data acquisition, high-speed embedded memory, high-performance packaging and testing, noise in mixed-signal integrated circuits, and circuits for high-speed communications. In 1986, he was a member of the NSF-sponsored JTECH Panel on Telecommunications Technology in Japan. He has published more than 120 technical articles and is a coauthor of *The Design of Low-Voltage, Low-Power Sigma–Delta Modulators* (Norwell, MA: Kluwer, 1998). He is a coeditor of *Analog MOS Integrated Circuits II*, June 2000.

Dr. Wooley has served as the Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS and as the Chairman of both the International Solid-State Circuits Conference (ISSCC) and the Symposium on VLSI Circuits. He is also a past Chairman of the IEEE Solid-State Circuits and Technology Committee, and he has been a member of the IEEE Solid-State Circuits Council, the IEEE Circuits and Systems Society Adcom, the Executive Committee of the ISSCC, and the Executive Committee of the Symposium on VLSI Circuits. He is currently the President of the IEEE Solid-State Circuits Society. He was awarded the University Medal by the University of California, Berkeley, and he was an IEEE Fortescue Fellow. He was also a recipient of the IEEE Third Millennium Medal. He is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu. He received an Outstanding Panelist Award for the 1985 ISSCC and the Winner Editorial Award for papers presented at the 1991 and 1997 ISSCC.