Introduction to RF Simulation and Its Application

Kenneth S. Kundert

***Abstract*—Radio-frequency (RF) circuits exhibit several distinguishing characteristics that make them difficult to simulate using traditional SPICE transient analysis. The various extensions to the harmonic balance and shooting method simulation algorithms are able to exploit these characteristics to provide rapid and accurate simulation for these circuits.**

This paper is an introduction to RF simulation methods and how they are applied to make common RF measurements. It describes the unique characteristics of RF circuits, the methods developed to simulate these circuits, and the application of these methods.

***Index Terms*— Circuit simulation, cyclostationary noise, envelope methods, harmonic balance, interchannel interference, intermodulation distortion, jitter, mixer noise, mixers, nonlinear oscillators, phase noise, quasi-periodic methods, shooting methods, SPICE.**

I. The RF Interface

WIRELESS transmitters and receivers can be conceptually separated into baseband and radio-frequency (RF) sections. Baseband is the range of frequencies over which transmitters take their input and receivers produce their output. The width of the baseband determines the underlying rate at which data can flow through the system. There is a considerable amount of signal processing that occurs at baseband designed to improve the fidelity of the data stream being communicated and to reduce the load the transmitter places on the transmission medium for a particular data rate. The RF section of the transmitter is responsible for converting the processed baseband signal up to the assigned channel and injecting the signal into the medium. Conversely, the RF section of the receiver is responsible for taking the signal from the medium and converting it back down to baseband.

With transmitters there are two primary design goals. First, they must transmit a specified amount of power while consuming as little power as possible. Second, they must not interfere with transceivers operating on adjacent channels. For receivers, there are three primary design goals. First, they must faithfully recover small signals. Second, they must reject interference outside the desired channel. And third, like transmitters, they must be frugal power consumers.

A. Small Desired Signals

Receivers must be very sensitive to detect small input signals. Typically, receivers are expected to operate with as little as 1 μV at the input. The sensitivity of a receiver is limited by the noise generated in the input circuitry of the receiver. Thus, noise is an important concern in receivers, and the ability to predict noise by simulation is very important. As shown in Fig. 1, a typical superheterodyne receiver first filters and then amplifies its input with a low-noise amplifier (LNA). It then translates the signal to the intermediate frequency (IF) by mixing it with the first local oscillator (LO). The noise performance of the front end is determined mainly by the LNA, the mixer, and the LO. While it is possible to use traditional SPICE noise analysis to find the noise of the LNA, it is useless on the mixer and the LO because the noise in these blocks is strongly influenced by the large LO signal.

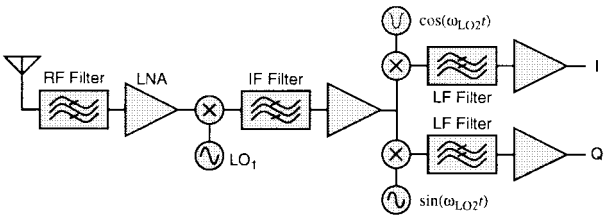

Fig. 1. A coherent superheterodyne receiver's RF interface.

The small input signal level requires that receivers must be capable of a tremendous amount of amplification. Often as much as 120 dB of gain is needed. With such high gain, any coupling from the output back to the input can cause problems. One important reason why the superheterodyne receiver architecture is used is to spread that gain over several frequencies to reduce the chance of coupling. It also results in the first LO's being at a different frequency than the input, which prevents this large signal from contaminating the small input signal. For various reasons, the direct conversion or homodyne architecture is a candidate to replace the superheterodyne architecture in some wireless communication systems [1], [16], [44], [45]. In this architecture, the RF input signal is directly converted to baseband in one step. Thus, most of the gain will be at baseband and the LO will be at the same frequency as the input signal. In this case, the ability to determine the impact of small amounts of coupling is quite important and will require careful modeling of the stray signal paths, such as coupling through the substrate, between package pins and bondwires, and through the supply lines.

B. Large Interfering Signals

Receivers must be sensitive to small signals even in the presence of large interfering signals, often known as blockers. This situation arises when trying to receive a weak or distant transmitter with a strong nearby transmitter broadcasting in an adjacent channel. The interfering signal can be 60–70-dB

Manuscript received January 7, 1999; revised April 21, 1999.

The author is with Cadence Design Systems, San Jose, CA 95134 USA (e-mail: kundert@cadence.com).

Publisher Item Identifier S 0018-9200(99)06492-6.

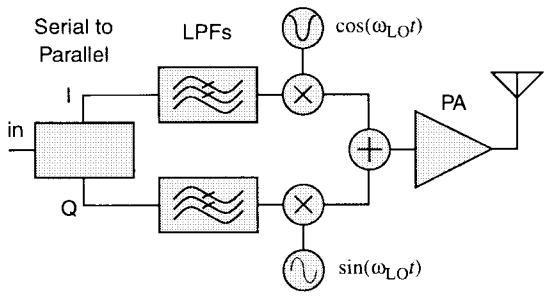

Fig. 2. A digital direct conversion transmitter's RF interface.

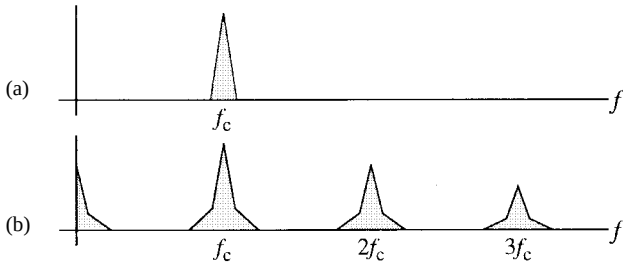

Fig. 3. Spectrum of a narrow-band signal centered at a carrier frequency f_c (a) before and (b) after passing though a nonlinear circuit. The nonlinearity causes the signal to be replicated at multiples of the carrier, an effect referred to as harmonic distortion, and adds a skirt to the signal that increases its bandwidth, an effect referred to as intermodulation distortion. It is possible to eliminate the effect of harmonic distortion with a bandpass filter; however, the frequency of the intermodulation distortion products overlaps the frequency of the desired signal and so cannot be completely removed with filtering.

larger than the desired signal and can act to block its reception by overloading the input stages of the receiver or by increasing the amount of noise generated in the input stage. Both of these problems result if the input stage is driven into a nonlinear region by the interferer. To avoid these problems, the front end of a receiver must be very linear. Thus, linearity is also an important concern in receivers. Receivers are narrow-band circuits and so the nonlinearity is quantified by measuring the intermodulation distortion. This involves driving the input with two sinusoids that are in band and close to each other in frequency and then measuring the intermodulation products. This is generally an expensive simulation with SPICE because many cycles must be computed in order to have the frequency resolution necessary to see the distortion products.

C. Adjacent Channel Interference

Distortion also plays an important role in the transmitter where nonlinearity in the output stages can cause the bandwidth of the transmitted signal to spread out into adjacent channels. This is referred to as spectral regrowth because, as shown in Figs. 2 and 3, the bandwidth of the signal is limited before it reaches the transmitter's power amplifier (PA), and intermodulation distortion in the PA causes the bandwidth to increase again. If it increases too much, the transmitter will not meet its adjacent channel power requirements. When transmitting digitally modulated signals, spectral regrowth is virtually impossible to predict with SPICE. The transmission of around 1000 digital symbols must be simulated to get a representative spectrum, and this combined with the high carrier frequency makes use of transient analysis impractical.

II. Characteristics of RF Circuits

RF circuits have several unique characteristics that are barriers to the application of traditional circuit simulation techniques. Over the last decade, researchers have developed many special-purpose algorithms that overcome these barriers to provide practical simulation for RF circuits, often by exploiting the very characteristic that represented the barrier to traditional methods [28].

A. Narrow-Band Signals

RF circuits process narrow-band signals in the form of modulated carriers. Modulated carriers are characterized as having a periodic high-frequency carrier signal and a low-frequency modulation signal that acts on the amplitude, phase, or frequency of the carrier. For example, a typical cellular telephone transmission has a 10–30-kHz modulation bandwidth riding on a 1–2-GHz carrier. In general, the modulation is arbitrary, though it is common to use a sinusoid or a simple combination of sinusoids as test signals.

The ratio between the lowest frequency present in the modulation and the frequency of the carrier is a measure of the relative frequency resolution required of the simulation. General-purpose circuit simulators, such as SPICE, use transient analysis to predict the nonlinear behavior of a circuit. Transient analysis is expensive when it is necessary to resolve low modulation frequencies in the presence of a high carrier frequency because the high-frequency carrier forces a small timestep while a low-frequency modulation forces a long simulation interval.

Passing a narrow-band signal though a nonlinear circuit results in a broad-band signal whose spectrum is relatively sparse, as shown in Fig. 3. In general, this spectrum consists of clusters of frequencies near the harmonics of the carrier. These clusters take the form of a discrete set of frequencies if the modulation is periodic or quasi-periodic, and a continuous distribution of frequencies otherwise.

RF simulators exploit the sparse nature of this spectrum in various ways and with varying degrees of success. Steady-state methods (Section IV-A) are used when the spectrum is discrete, and transient methods (Section IV-C) are used when the spectrum is continuous.

B. Time-Varying Linear Nature of the RF Signal Path

Another important but less appreciated aspect of RF circuits is that they are generally designed to be as linear as possible from input to output to prevent distortion of the modulation or information signal. Some circuits, such as mixers, are designed to translate signals from one frequency to another. To do so, they are driven by an additional signal, the LO, a large periodic signal the frequency of which equals the amount of frequency translation desired. For best performance, mixers are designed to respond in a strongly nonlinear fashion to the LO. Thus, mixers behave both near linearly (to the input) and strongly nonlinearly (to the LO).

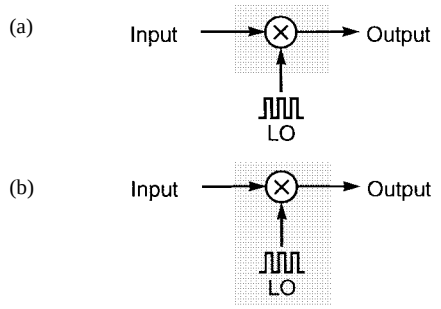

Fig. 4. One can often approximate (a) a nonlinear clocked or periodically driven circuit with (b) a linear periodically varying circuit.

A timing or clock signal such as the LO is independent of the information signal, and so they may be considered to be part of the circuit rather than an input to the circuit, as shown in Fig. 4. This simple change of perspective allows the mixer to be treated as having a single input and a near-linear, though periodically time-varying, transfer function. As an example, consider a mixer made from an ideal multiplier and followed by a low-pass filter. A multiplier is nonlinear and has two inputs. Applying an LO signal of $\cos(\omega_{\rm LO}t)$ consumes one input and results in a transfer function of

$$v_{\rm out}(t) = {\rm LPF}\{\cos(\omega_{\rm LO}t)v_{\rm in}(t)\} \tag{1}$$

which is clearly time varying and is easily shown to be linear with respect to $v_{\rm in}$. If the input signal is

$$v_{\rm in}(t) = m(t)\cos(\omega_c t) \tag{2}$$

then

$$v_{\rm out}(t) = {\rm LPF}\{m(t)\cos(\omega_c t)\cos(\omega_{\rm LO}t)\} \tag{3}$$

and

$$v_{\rm out}(t) = m(t)\cos\left((\omega_c - \omega_{\rm LO})t\right). \tag{4}$$

This demonstrates that a linear periodically varying transfer function implements frequency translation.

Often we can assume that the information signal is small enough to allow the use of a linear approximation of the circuit from its input to its output. Thus, a small-signal analysis can be performed, as long as it accounts for the periodically varying nature of the signal path, which is done by linearizing about the periodic operating point. This is the idea behind the small-signal analyses of Section IV-B. Traditional simulators such as SPICE provide several small-signal analyses, such as the ac and noise analyses, that are considered essential when analyzing amplifiers and filters. However, they start by linearizing a nonlinear time-invariant circuit about a constant operating point and so generate a linear time-invariant representation that cannot exhibit frequency translation. Linearizing a nonlinear circuit about a periodically varying operating point extends small-signal analysis to clocked circuits, or circuits that must have a periodic clock signal present to operate properly, such as mixers, switched filters, samplers, and oscillators (oscillators are self-clocked, so the clock signal is the desired output of the oscillator and the information signal is generally an undesired signal, such as the noise). In doing so, a periodically varying linear representation results, which does exhibit frequency translation.

All of the traditional small-signal analyses can be extended in this manner, enabling a wide variety of applications (some of which are described in [59]). In particular, a noise analysis that accounts for noise folding and cyclostationary noise sources can be implemented [40], [52], which fills a critically important need for RF circuits. When applied to oscillators, it also accounts for phase noise [8], [9], [21], [22].

C. Linear Passive Components

At the high frequencies present in RF circuits, the passive components, such as transmission lines, spiral inductors, packages (including bond wires), and substrates, often play a significant role in the behavior of the circuit. The nature of such components often makes them difficult to include in the simulation.

Generally, the passive components are linear and are modeled with phasors in the frequency domain, using either analytical expressions or tables of S-parameters. This greatly simplifies the modeling of distributed components such as transmission lines. Large distributed structures, such as packages, spirals, and substrates, often interface with the rest of the circuit through a small number of ports. Thus, they can be easily replaced by an N-port macromodel that consists of the N^2 transfer functions. These transfer functions are found by reducing the large systems of equations that describe these structures using Gaussian elimination, leaving only the equations that relate the signals at their ports. The relatively expensive reduction step is done once for each frequency as a preprocessing step. The resulting model is one that is efficient to evaluate in a frequency-domain simulator if N is small. This is usually true for transmission lines and spirals and less true for packages and substrates.

Time-domain simulators are formulated to solve sets of first-order ordinary-differential equations (ODE's). However, distributed components, such as transmission lines, are described with partial-differential equations (PDE's) and are problematic for time-domain simulators. Generally, the PDE's are converted to a set of ODE's using some form of discretization [6], [35]. Such approaches suffer from bandwidth limits. An alternative approach is to compute the impulse response for a distributed component from a frequency-domain description and use convolution to determine the response of the component in the circuit [20], [54], [56]. Evaluating lossy or dispersive transmission-line models or tables of S-parameters with this approach is generally expensive and error prone. Packages, substrates, and spirals can be modeled with large lumped networks, but such systems can be too large to be efficiently incorporated into a time-domain simulation, and so some form of reduction is necessary [11], [42].

D. Semiconductor Models

The semiconductor models used by RF simulators must accurately model the high-frequency, small-signal behavior of the devices to accurately predict the behavior of RF circuits. BJT's have long been used in high-frequency analog

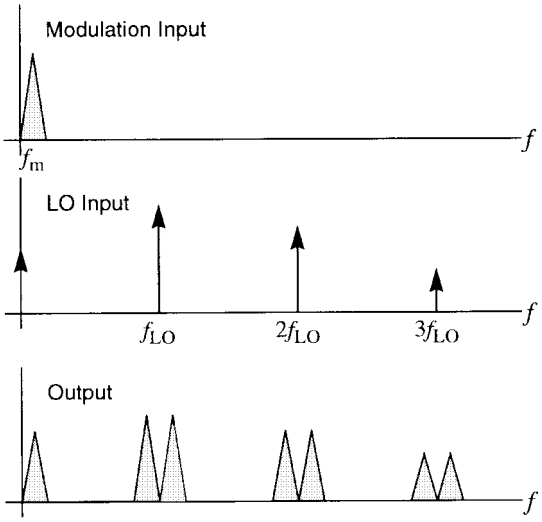

Fig. 5. Signals at the inputs and outputs of an up-conversion mixer. The modulation signal is mixed up to the upper and lower sidebands of the LO and its harmonics.

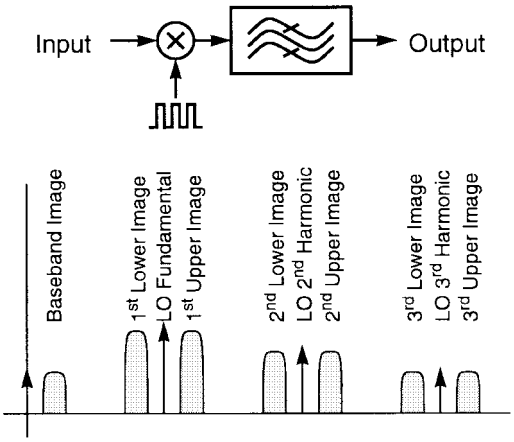

Fig. 6. Images at the input of the first mixing stage of a typical receiver. The images are frequency bands where the output is sensitive to signals at the input.

circuits, and their models are well suited for RF circuits. With the advent of submicrometer technologies, RF circuits are now being realized in standard CMOS processes [1], [16]; however, existing MOS models are inadequate for RF applications. In particular, the distributed resistance in the gate and substrate are not well modeled, which affects the driving point immittances, the transfer functions, and, perhaps most important, the noise [19]. In addition, flicker noise is not well modeled, which plays a large roll in oscillator phase noise, and is particularly important for CMOS oscillators because of the large amount of flicker noise produced by MOS devices [32].

III. Basic RF Building Blocks

RF systems are constructed primarily using four basic building blocks—amplifiers, filters, mixers, and oscillators. Amplifiers and filters are common analog blocks and are well handled by SPICE. However, mixers and oscillators are not heavily used in analog circuits, and SPICE has limited ability to analyze them. What makes these blocks unique is presented next.

A. Mixers

Mixers translate signals from one frequency range to another. They have two inputs and one output. One input is for the information signal and the other is for the clock signal, the LO. Ideally, the signal at the output is the same as that at the information signal input, except shifted in frequency by an amount equal to the frequency of the LO. As shown in Section II-B, a multiplier can act as a mixer. In fact, a multiplier is a reasonable model for a mixer except that the LO is passed through a limiter, which is usually an integral part of the mixer, to make the output less sensitive to noise on the LO.

The input and output signals of a mixer used for up-conversion (as in a transmitter) are shown in Fig. 5. The LO is shown after passing through the limiter so that the output in the time domain is simply the product of the inputs, or the convolution of the two inputs in the frequency domain. The information signal, here a modulation signal, is replicated at the output above and below each harmonic of the LO. These bands of signal above and below each harmonic are referred to as *sidebands.* There are two sidebands associated with each harmonic of the LO. The ones immediately above the harmonics are referred to as the *upper sidebands* and the ones below are referred to as the *lower sidebands.* The sideband at dc is referred to as the *baseband.*

When the LO has a rich harmonic content, an input signal at any sideband will be replicated to each of the sidebands at the output. Usually, only one sideband is of interest, and the others must be eliminated. If the desired sideband is the baseband, then the undesired sidebands are eliminated with a low-pass filter. Otherwise, the undesired sidebands are removed with a bandpass filter. This works well for sidebands of harmonics different from that of the desired sideband. However, special techniques are then required to eliminate the remaining undesired sideband [44].

Consider a down-conversion mixer (as in a receiver) and assume that the mixer is followed by a filter. This filter is used to remove all but the desired channel. The output of the mixer/filter pair is sensitive to signals in each sideband of the LO. Associated with each sideband is a transfer function from that sideband to the output. The shape of the transfer function is determined largely by the filter. Thus, the bandwidth of the passband is that of the filter. If the filter is a bandpass, then the passband of the transfer function will be offset from the LO or its harmonic by the center frequency of the filter. These passbands are referred to as the *images* of the filter and are shown in Fig. 6. Generally, only one image is desired, and the rest are undesired. The most troubling is usually the one that shares the same harmonic as the desired image. Image-reject mixers are designed to reduce the gain associated with this undesired image [44].

Sidebands and images are related but are not the same. Sidebands are frequency bands in the signal actually produced at the output of a mixer, whereas images are frequency bands at the input of a mixer that have the potential to produce a response at the output.

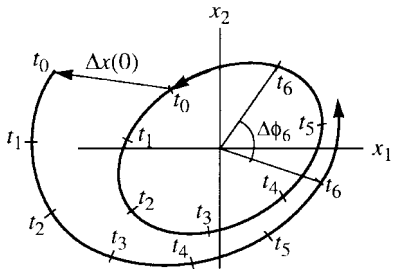

Fig. 7. The trajectory of an oscillator shown in state space with and without a perturbation Δx. By observing the time stamps $(t_0, \ldots, t_6)$, one can see that the deviation in amplitude dissipates while the deviation in phase does not.

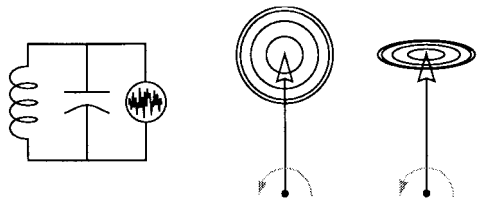

Fig. 8. (a) A linear oscillator along with (b) its response to noise and (c) the response of a nonlinear oscillator to a noise. The arrows are phasors that represents the unperturbed oscillator output, the carriers, and the circles represent the response to perturbations in the form of noise. With a linear oscillator, the noise simply adds to the carrier. In a nonlinear oscillator, the nonlinearities act to control the amplitude of the oscillator and so to suppress variations in amplitude, thereby radially compressing the noise ball and converting it into predominantly a variation in phase.

B. Oscillators

Oscillators generate a reference signal at a particular frequency. In some oscillators, referred to as voltage-controlled oscillators (VCO's), the frequency of the output varies proportionally to some input signal. Compared to mixers, oscillators seem quite simple. That is an illusion.

Oscillators are generally used in RF circuits to generate the LO signal for mixers. The noise performance of the mixer is strongly affected by noise on the LO signal. The LO is always passed through a limiter, which is generally built into the mixer, to make the mixer less sensitive to small variations in the amplitude of the LO. However, the mixer is still sensitive to variations in the phase of the LO. Thus, it is important to minimize the phase noise produced by the oscillator.

Nonlinear oscillators naturally produce high levels of phase noise. To see why, consider the trajectory of an oscillator's stable periodic orbit in state space. Furthermore, consider disturbing the oscillator by applying an impulse $u(t) = \delta(t)$. The oscillator responds by following a perturbed trajectory $x(t) + \Delta x(t)$, as shown in Fig. 7, where $x(t)$ represents the unperturbed solution and $\Delta x(t)$ is the perturbation in the response.

Decompose the perturbed response into amplitude and phase variations

$$v(t) = x(t) + \Delta x(t) = (1 + \alpha(t))x\left(t + \frac{\phi(t)}{2\pi f_c}\right) \qquad (5)$$

where $v(t)$ represents the noisy output voltage of the oscillator, $\alpha(t)$ represents the variation in amplitude, $\phi(t)$ is the variation in phase, and f_c is the oscillation frequency.

Since the oscillator is stable and the duration of the disturbance is finite, the deviation in amplitude eventually decays and the oscillator returns to its stable orbit. In effect, there is a restoring force that tends to act against amplitude noise. This restoring force is a natural consequence of the nonlinear nature of the oscillator and at least partially suppresses amplitude variations, as shown in Fig. 8. With linear oscillators, there is no restoring force, and so the amplitude is arbitrary (i.e., they do not have stable orbits). As such, linear oscillators exhibit equal amounts of amplitude and phase noise because the amplitude noise is not suppressed.

Since the oscillator is autonomous, any time-shifted version of the solution is also a solution. Once the phase has shifted due to a perturbation, the oscillator continues on as if never disturbed except for the shift in the phase of the oscillation. There is no restoring force on the phase, and so phase deviations accumulate. This is true for both linear and nonlinear oscillators. Notice that there is only one degree of freedom—the phase of the oscillator as a whole. There is no restoring force when the phase of all signals associated with the oscillator shift together; however, there would be a restoring force if the phase of signals shifted relative to each other. This is important in oscillators with multiple outputs, such as quadrature oscillators or ring oscillators. The dominant phase variations appear identically in all outputs, whereas relative phase variations between the outputs are naturally suppressed by the oscillator or added by subsequent circuitry and so tend to be much smaller [8].

After being disturbed by an impulse, the asymptotic response of the amplitude deviation is $\alpha(t) \to 0$ as $t \to \infty$. However, the asymptotic response of the phase deviation is $\phi(t) \to \Delta\phi$. If responses that decay are neglected, then the impulse response of the phase deviation $\phi(t)$ can be approximated with a unit step $s(t)$. Thus, the phase shift over time for an arbitrary input disturbance u is

$$\phi(t) \sim \int_{-\infty}^{\infty} s(t-\tau)u(\tau)\,d\tau = \int_{-\infty}^{t} u(\tau)\,d\tau \qquad (6)$$

or the power spectral density (PSD) of the phase is

$$S_\phi(f) \sim \frac{S_u(f)}{(2\pi f)^2}. \qquad (7)$$

The disturbance u may be either deterministic or random in character and may result from extraneous signals coupling into the oscillator or from variations in the components that make up the oscillator, such as thermal, shot, and flicker noise.

If $S_u(f)$ is white noise, then $S_\phi(f)$ is proportional to $1/(2\pi f)^2$. This result has been shown here to apply at low frequencies, but with a more detailed derivation it can also be shown to be true over a broad range of frequencies [21]. Assume that u is white and define a such that

$$S_\phi(f) = a\,\frac{f_c^2}{f^2} \qquad (8)$$

where $f_c = 1/T$ is the oscillation or carrier frequency. S_ϕ is the PSD of the phase variable in (5). Phase cannot easily be observed directly, so instead one is often interested in S_v, the

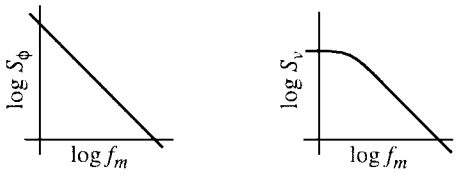

Fig. 9. Two different ways of characterizing noise in the same oscillator. S_ϕ is the spectral density of the phase and S_v is the spectral density of the voltage. S_v contains both amplitude and phase-noise components, but with oscillators the phase noise dominates except at frequencies far from the carrier and its harmonics. S_v is directly observable on a spectrum analyzer, whereas S_ϕ is only observable if the signal is first passed through a phase detector. Another measure of oscillator noise is L, which is simply S_v normalized to the power in the fundamental.

PSD of v. Near the fundamental [9], [21], [23], [57]

$$S_v(f_c + f_m) = |X_1|^2 \frac{af_c^2}{a^2\pi^2 f_c^4 + f_m^2} \tag{9}$$

where f_m is the frequency offset from the fundamental and X_1 is the first Fourier coefficient for x

$$x(t) = \sum_{k=-\infty}^{\infty} X_k \, e^{j2\pi k f_c t}. \tag{10}$$

This spectrum is a Lorentzian, as shown in Fig. 9. The corner frequency f_Δ is known as the linewidth of the oscillator and is given by $f_\Delta = a\pi f_c^2$, with

$$S_v(f_c + f_m) = \frac{|X_1|^2}{\pi} \frac{f_\Delta}{f_\Delta^2 + f_m^2}. \tag{11}$$

As $t \to \infty$, the phase of the oscillator drifts without bound, so $S_\phi(f_m) \to \infty$ as $f_m \to 0$. However, even as the phase drifts without bound, the excursion in the voltage is limited by the diameter of the limit cycle of the oscillator [represented by the periodic function x in (5)]. Therefore, as $f_m \to 0$, the PSD of v flattens out and $S_v(f_c + f_m) \to |X_1|^2/(\pi f_\Delta)$, which is inversely proportional to a. Thus, the larger a, the more phase noise, the broader the linewidth, and the lower the signal amplitude within the linewidth. This happens because the phase noise does not affect the total power in the signal; it only affects its distribution. Without phase noise, $S_v(f)$ is a series of impulse functions at the harmonics of f_c. With phase noise, the impulse functions spread, becoming fatter and shorter but retaining the same total power [9].

It is more common to report oscillator phase noise as $\mathcal{L}$, the ratio of the single-sideband (SSB) phase-noise power to the power in the fundamental (in dBc/Hz)

$$\mathcal{L}(f_m) = \frac{S_v(f_c + f_m)}{|X_1|^2} = \frac{1}{\pi} \frac{f_\Delta}{f_\Delta^2 + f_m^2}. \tag{12}$$

At frequencies where the oscillator phase noise dominates over the amplitude noise and that are also outside the linewidth ($f_m > f_\Delta$), the phase noise is approximated with[1]

$$\mathcal{L}(f_m) = \frac{f_\Delta}{\pi f_m^2} = \frac{af_c^2}{f_m^2} = S_\phi(f_m) \quad \text{for} \quad f_\Delta \ll f \ll f_c. \tag{13}$$

[1]Other references report that $\mathcal{L}(f_m) = S_\phi(f_m)/2$, which is true when S_ϕ is the single-sided PSD [48], [61]. Here, S_ϕ is the double-sided PSD.

The rolloff in $S_v(f_c + f_m)$ and $\mathcal{L}(f_m)$ as $f_m \to 0$ is a result of the circuit's responding in a nonlinear fashion to the noise itself. As such, it cannot be anticipated by the small-signal noise analyses that will be presented in Section V-B. However, as can be seen from Fig. 9, S_ϕ does not roll off at low frequencies, so these analyses along with (13) can be used to compute S_ϕ at low frequencies.

Phase noise acts to vary the period of the oscillation, a phenomenon known as jitter. Assume that u is a white stationary process. Then its variance is constant, and from (6) the variance of ϕ increases linearly with time. Demir [9] shows that the variance of a single period is aT, where a is defined in (8) and $T = 1/f_c$. The jitter J_k is the standard deviation of the length of k periods, so

$$J_k = \sqrt{kaT}. \tag{14}$$

In the case where u represents flicker noise, $S_u(f)$ is generally pink or proportional to $1/f$. Then $S_\phi(f)$ would be proportional to $1/f^3$ at low frequencies [22]. In this case, there are no explicit formulas for f_Δ and J_k or S_v and $\mathcal{L}$ at low offset frequencies.

IV. RF Analyses

SPICE provides several different types of analyses that have proven themselves essential to designers of baseband circuits. These same analyses are also needed by RF designers, except they must extended to address the issues described in Section II and the circuits of Section III. The basic SPICE analyses include dc, ac, noise, and transient. RF versions of each have been developed in recent years based on two different foundations: harmonic balance and shooting methods. Both harmonic balance and shooting methods started off as methods for computing the periodic steady-state solution of a circuit but have been generalized to provide all the functionality needed by RF designers. In their original forms, they were limited to relatively small circuits. Recently, Krylov subspace methods have been applied to accelerate both harmonic balance and the shooting methods, which allows them to be applied to much larger circuits [13], [30], [33], [58], [60], [61].

A. Periodic and Quasi-Periodic Analysis

Periodic and quasi-periodic analyses can be thought of as RF extensions of SPICE's dc analysis. In dc analysis, one applies constant signals to the circuit and it computes the steady-state solution, which is the dc operating point about which subsequent small-signal analyses are performed. Sometimes, the level of one of the input signals is swept over a range, and the dc analysis is used to determine the large-signal dc transfer curves of the circuit.

With periodic and quasi-periodic analyses, the circuit is driven with one or more periodic waveforms and the steady-state response is computed. This solution point is used as a periodic or quasi-periodic operating point for subsequent small-signal analyses. In addition, the level of one of the input signals may be swept over a range to determine the power transfer curves of the circuit.

Periodic and quasi-periodic analyses are generally used to predict the distortion of RF circuits and to compute the operating point about which small-signal analyses are performed (presented later). When applied to oscillators, periodic analysis is used to predict the operating frequency and power and can also be used to determine how changes in the load affect these characteristics (load pull).

Quasi-periodic steady-state (QPSS) analyses compute the steady-state response of a circuit driven by one or more large periodic signals. The steady-state or eventual response is the one that results after any transient effects have dissipated. Such circuits respond in steady state with signals that have a discrete spectrum with frequency components at the drive frequencies, at their harmonics, and at the sum and difference frequencies of the drive frequencies and their harmonics. Such signals are called quasi-periodic and can be represented with a generalized Fourier series

$$v(t) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} V_{kl}\, e^{j2\pi(kf_1+lf_2)t} \tag{15}$$

where V_{kl} are Fourier coefficients and f_1 and f_2 are fundamental frequencies. For simplicity, a 2-fundamental quasi-periodic waveform is shown in (15), though quasi-periodic signals can have any finite number of fundamental frequencies. If there is only one fundamental, the waveform is simply periodic. f_1 and f_2 are assumed to be noncommensurate, which means that there exists no frequency f_0 such that both f_1 and f_2 are exact integer multiples of f_0. If f_1 and f_2 are commensurate, then $v(t)$ is simply periodic.

The choice of the fundamental frequencies is not unique. Consider a down-conversion mixer that is driven with two periodic signals at f_{RF} and f_{LO}, with the desired output at $f_{\mathrm{IF}} = f_{\mathrm{RF}} - f_{\mathrm{LO}}$. The circuit responds with a 2-fundamental quasi-periodic steady-state response where the fundamental frequencies can be f_{RF} and f_{LO}, f_{LO} and f_{IF}, or f_{IF} and f_{RF}. Typically, the drive frequencies are taken to be the fundamentals, which in this case are f_{RF} and f_{LO}. With an up-conversion mixer, the fundamentals would likely be chosen to be f_{IF} and f_{LO}.

As discussed in Section II-A, computing signals that have the form of (15) with traditional transient analysis would be very expensive if f_1 and f_2 are widely spaced so that $\min(f_1, f_2)/\max(f_1, f_2) \ll 1$ or if they are closely spaced so that $\min(f_1, f_2)/\max(f_1, f_2) \ll 1$. Large-signal steady-state analyses directly compute the quasi-periodic solution without having to simulate through long time constants or long beat tones (the beat tone is the lowest frequency present excluding dc). The methods generally work by directly computing the Fourier coefficients V_{kl}. To make the computation tractable, it is necessary for all but a small number of Fourier coefficients to be negligible. These coefficients would be ignored. Generally, we can assume that all but the first K_i harmonics and associated cross terms of each fundamental i are negligible. With this assumption, $K = \Pi_i(2K_i + 1)$ coefficients remain to be calculated, which is still a large number if the number of fundamentals is large. In practice, these methods are typically limited to a maximum of three or four fundamental frequencies.

1) Harmonic Balance: Harmonic balance [27], [30], [36], [47] formulates the circuit equations and their solution in the frequency domain. The solution is written as a Fourier series that cannot represent transient behavior, and so harmonic balance directly finds the steady-state solution. Consider

$$f(v(t), t) = i(v(t)) + \frac{dq(v(t))}{dt} + u(t) = 0. \tag{16}$$

This equation is capable of modeling any lumped time-invariant nonlinear system; however, it is convenient to think of it as being generated from nodal analysis, and so representing a statement of Kirchhoff's current law for a circuit containing nonlinear conductors, nonlinear capacitors, and current sources. In this case, $v(t) \in R^N$ is the vector of node voltages, $i(v(t)) \in R^N$ represents the current out of the node from the conductors, $q(v(t))$ represents the charge out of the node from the capacitors, and $u(t)$ represents the current out of the node from the sources. To formulate the harmonic balance equations, assume that $v(t)$ and $u(t)$ are T-periodic and reformulate the terms of (16) as a Fourier series

$$\sum_{k=-\infty}^{\infty} F_k(V) e^{j2\pi kft} = 0 \tag{17}$$

where $f = 1/T$ is the fundamental frequency, and

$$F_k(V) = j2\pi kfQ_k(V) + I_k(V) + U_k. \tag{18}$$

Since $e^{j2\pi kft}$ and $e^{j2\pi lft}$ are linearly independent over a period T if $k \neq l$ then $F_k(V) = 0$ for each k individually, and so (17) can be reformulated as a system of equations, one for each harmonic k. To make the problem numerically tractable, it is necessary to consider only the first K harmonics. The result is a set of K complex equations $(F_k(V) = 0)$ and K complex unknowns (V_k) that are typically solved using Newton's method [27].

It is, in general, impossible to directly formulate models for nonlinear components in the frequency domain. To overcome this problem, nonlinear components are usually evaluated in the time domain. Thus, the frequency-domain voltage is converted into the time domain using the inverse Fourier transform, the nonlinear component (i and q) is evaluated in the time domain, and the current or charge is converted back into the frequency domain using the Fourier transform.

2) Autonomous Harmonic Balance: An extremely important application of harmonic balance is determining the steady-state behavior of oscillators. However, as presented, harmonic balance is not suitable for autonomous circuits such as oscillators. The method was derived assuming the circuit was driven, which made it possible to know the operating frequency in advance. Instead, it is necessary to modify harmonic balance to directly compute the operating frequency by adding the oscillation frequency to the list of unknowns and adding an additional equation that constrains the phase of the computed solution [27].

3) Quasi-Periodic Harmonic Balance: A two-fundamental quasi-periodic signal takes the form

$$x(t) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} X_{kl}\, e^{j2\pi(kf_1+lf_2)t} \tag{19}$$

where f_1 and f_2 are the fundamental frequencies. Rearranging (19) shows this to be equivalent to constructing the waveform as a conventional Fourier series where the frequency of each term is an integer multiple of f_1, except that the Fourier coefficients themselves are time varying. In particular, the coefficient $\tilde{X}_k(t)$ is periodic with period $T_2 = 1/f_2$ and can itself be represented as a Fourier series

$$x(t) = \sum_{k=-\infty}^{\infty} \underbrace{\sum_{l=-\infty}^{\infty} X_{kl}\, e^{j2\pi l f_2 t}}_{\tilde{X}_k(t)}\, e^{j2\pi k f_1 t}. \tag{20}$$

Define $\hat{x}(t_1, t_2)$ such that $x(t) = \hat{x}(t, t)$ with $\hat{x}(t_1, t_2)$ being T_1 periodic in t_1 and T_2 periodic in t_2. In this way a two-dimensional version of x is created where temporal dimensions are associated with the time scales of each of the fundamental frequencies. Then

$$\hat{x}(t_1, t_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} X_{kl}\, e^{j2\pi l f_2 t_2}\, e^{j2\pi k f_1 t_1}. \tag{21}$$

This is a two-dimensional Fourier series, and so $\hat{x}$ and X are related by a two-dimensional Fourier transform.

Using these ideas, we can reformulate (16) in terms of t_1 and t_2

$$\frac{\partial q(\hat{v}(t_1, t_2))}{\partial t_1} + \frac{\partial q(\hat{v}(t_1, t_2))}{\partial t_2} + i(\hat{v}(t_1, t_2)) + \hat{u}(t_1, t_2) = 0 \tag{22}$$

or

$$f(\hat{v}(t_1, t_2), t_1, t_2) = 0. \tag{23}$$

Assuming that $\hat{v}$ and f of (23) take the form of (21)

$$\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_{kl}(V)\, e^{j2\pi(k f_1 t_1 + l f_2 t_2)} = 0 \tag{24}$$

where

$$F_{kl}(V) = j2\pi(k f_1 + l f_2) Q_{kl}(V) + I_{kl}(V) + U_{kl}. \tag{25}$$

The terms in (24) are linearly independent over all t assuming that f_1 and f_2 are noncommensurate (share no common period). So $F_{kl}(V) = 0$ for each k, l. This becomes finite-dimensional by bounding $k < K$ and $l < L$. When evaluating I and Q, the multidimensional discrete Fourier transform is used.

Using a multidimensional Fourier transform is just one way of formulating harmonic balance for quasi-periodic problems [49], [66]. It is used here because of its simple derivation and because it introduces ideas that will be used later in Section IV-C. An alternate approach that is generally preferred in practice is the false frequency method, which is based on a one-dimensional Fourier transform [18], [27].

4) Shooting Methods: Traditional SPICE transient analysis solves initial-value problems. A shooting method is an iterative procedure layered on top of transient analysis that is designed to solve boundary-value problems. Boundary-value problems play an important role in RF simulation. For example, assume that (16) is driven with a nonconstant T-periodic stimulus. The T-periodic steady-state solution is the one that also satisfies the two-point boundary constraint

$$v(T) - v(0) = 0. \tag{26}$$

Define the state transition function $\phi_T(v_0, t_0)$ as the solution to (16) at $t_0 + T$ given that it starts at the initial state v_0 at t_0. In general, one writes

$$v(t_0 + T) = \phi_T(v(t_0), t_0). \tag{27}$$

Shooting methods combine (26) and (27) into

$$\phi_T(v(0), 0) = v(0) \tag{28}$$

which is a nonlinear algebraic problem, and so Newton methods can be used to solve for $v(0)$. The combination of the Newton and shooting methods is referred to as the shooting-Newton algorithm.

When applying Newton's method to (28), it is necessary to compute both the response of the circuit over one period and the sensitivity of the final state $v(T)$ with respect to changes in the initial state $v(0)$. The sensitivity is used to determine how to correct the initial state to reduce the difference between the initial and final state [2], [58].

5) Autonomous Shooting Methods: As with harmonic balance, it is extremely important to be able to determine the steady-state behavior of oscillators. To do so it is necessary to modify shooting methods to directly compute the period of the oscillator. The period is added as an extra unknown, and an additional equation is added that constrains the phase of the computed solution [27].

6) Quasi-Periodic Shooting Methods: As shown in (20), a two-fundamental quasi-periodic signal can be interpreted as a periodically modulated periodic signal. Designate the high-frequency signal as the *carrier* and the low-frequency signal as the *modulation.* If the carrier is much higher in frequency than the modulation, then the carrier will appear to vary only slightly from cycle to cycle. In this case, the complete waveform can be inferred from knowledge of a small number of cycles of the carrier appropriately distributed over one period of the modulation. The number of cycles needed can be determined from the bandwidth of the modulation signal. If the modulation signal can be represented using K harmonics, then the entire quasi-periodic signal can be recovered by knowing the waveform over $2K+1$ cycles of the carrier that are evenly distributed over the period of the modulation. This is the basic idea behind the mixed frequency-time (MFT) method [13], [26], [27].

Consider a circuit driven by two periodic signals that responds in steady-state by producing two-fundamental quasi-periodic waveforms as in (15). Designate the fundamental frequencies as f_1 and f_2 and consider the case where $f_1 \gg$

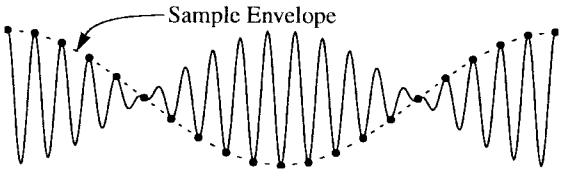

Fig. 10. The sample envelope is the waveform that results from sampling a signal with a period equal to that of the carrier.

f_2.[2] This may be because one input is a high-frequency signal and the other is a low-frequency signal, as would be the case with an up-conversion mixer. Or it may be that both inputs are high-frequency signals but their frequencies are close to each other and so they generate a low beat frequency, as with a down-conversion mixer. Designate f_1 as the carrier frequency and f_2 as the modulation frequency. Then $v(t)$ is the quasi-periodic response, where

$$v(t) = \sum_k \sum_l V_{kl}\, e^{j2\pi(kf_1 + lf_2)t}. \tag{29}$$

Consider sampling the signal $v(t)$ at the carrier frequency. The sampled signal is referred to as the sample envelope and is related to the continuous signal by $\hat{v}_n = v(nT_1)$, where $T_1 = 1/f_1$. $\hat{v}_n$ represents a sampled and perhaps scaled version of the modulation signal.

The MFT method works by computing the discrete sequence $\hat{v}$ instead of the continuous waveform v. Notice that if every $\hat{v}_n$ is related to the subsequent sample point $\hat{v}_{n+1}$ by

$$\hat{v}_{n+1} = \phi_{T_1}(\hat{v}_n, nT_1) \tag{30}$$

then all the $\hat{v}_n$ will satisfy the circuit equations. The transition function in (30) can be computed by standard SPICE transient analysis and serves to translate between the continuous signal and the discrete representation. The key to the MFT method is to require that the samples $\hat{v}_n$ represent a sampled quasi-periodic signal. This requirement is easily enforced because, as shown in Fig. 10, sampling a two-fundamental quasi-periodic signal at the carrier frequency results in a sampled waveform that is one-fundamental quasi-periodic, or simply periodic, at the modulation frequency. In other words, the sampled waveform can be written as a Fourier series with the carrier removed

$$\hat{v}_n = v(nT_1) = \sum_{l=-\infty}^{\infty} \hat{V}_k\, e^{j2\pi lnf_2T_1}. \tag{31}$$

Alternatively, one can write

$$\hat{v} = \mathcal{F}^{-1}\hat{V} \tag{32}$$

which states that $\hat{v}$ is the inverse Fourier transform of $\hat{V}$. Consider the nth sample interval and let $x_n = \hat{v}_n$ be the solution at the start of the interval and $y_n = \hat{v}_{n+1} = x_{n+1}$ be the solution at the end. Then, (30) uses the circuit equations to relate the solution at both ends of the interval

$$y_n = \phi_{T_1}(x_n, nT_1). \tag{33}$$

[2]The MFT method does not require that $f_1 \gg f_2$. However, if true, using MFT gives significant performance advantages over traditional transient analysis.

Define Φ as the function that maps the sequence x to the sequence y by repeated application of (33)

$$y = \Phi(x). \tag{34}$$

Let $X = \mathcal{F}x$ and $Y = \mathcal{F}y$ (X and Y are the Fourier transforms of x and y). Then, from (31) and since $y_n = x_{n+1}$

$$X_l = e^{-j2\pi lf_2T_1}\, Y_l \tag{35}$$

or

$$X = D_{T_1}Y \tag{36}$$

where D_{T_1} is referred to as the delay matrix. It is a diagonal matrix with $e^{-j2\pi lf_2T_1}$ being the lth diagonal element. Equation (36) is written in the time domain as

$$x = \mathcal{F}^{-1}D_{T_1}\mathcal{F}y. \tag{37}$$

Together, (34) and (37) make up the MFT method, where (34) stems solely from the circuit equations and (37) solely from the requirement that $\hat{v}$ represent a sampled quasi-periodic waveform. They can be combined into

$$x = \mathcal{F}^{-1}D_{T_1}\mathcal{F}\Phi(x) \tag{38}$$

or

$$\hat{v} = \mathcal{F}^{-1}D_{T_1}\mathcal{F}\Phi(\hat{v}). \tag{39}$$

Equation (39) is an implicit nonlinear equation that can be solved for $\hat{v}$ using Newton's method.

In practice, the modulation signals in the circuit are band-limited, and so only a finite number of harmonics is needed. Thus, the envelope shown in Fig. 10 can be completely specified by only a few of the sample points $\hat{v}_n$. With only K harmonics needed, (39) is solved over $2K+1$ distinct intervals using shooting methods. In particular, if the circuit is driven with one large high-frequency periodic signal at f_1 and one moderately sized sinusoid at f_2, then the number of harmonics needed K is small and the method is efficient. The total simulation time is proportional to the number of harmonics needed to represent the sampled modulation waveform and is independent of the period of the low-frequency beat tone or the harmonics needed to represent the carrier.

B. Small-Signal Analyses

The ac and noise analyses in SPICE are referred to as small-signal analyses. They assume that a small signal is applied to a circuit that is otherwise at its dc operating point. Since the input signal is small, the response can be computed by linearizing the circuit about its dc operating point (apply a Taylor series expansion about the dc equilibrium point and discard all but the first-order term). Superposition holds, so the response at each frequency can be computed independently. Such analyses are useful for computing the characteristics of circuits that are expected to respond in a near-linear fashion to an input signal and that operate about a dc operating point. This describes most "linear" amplifiers and continuous-time filters.

The assumption that the circuit operates about a dc operating point makes these analyses unsuitable for circuits that are

expected to respond in a near-linear fashion to an input signal but that require some type of clock signal to operate. Mixers fit this description, and if one considers noise to be the input, oscillators also fit. However, there is a wide variety of other circuits for which these assumptions also apply, such as samplers and sample-and-holds, switched-capacitor and switched-current filters, chopper-stabilized and parametric amplifiers, frequency multipliers and dividers, and phase detectors. These circuits, which are referred to as a group as clocked circuits, require the traditional small-signal analyses to be extended such that the circuit is linearized about a periodically varying operating point. Such analyses are referred to as linear periodically varying (LPV) analyses.

A great deal of useful information can be acquired by performing a small-signal analysis about the time-varying operating point of the circuit. LPV analyses start by performing a periodic analysis to compute the periodic operating point with only the large clock signal applied (the LO, clock, carrier, etc.). The circuit is then linearized about this time-varying operating point (expand about the periodic equilibrium point with a Taylor series and discard all but the first-order term) and the small information signal is applied. The response is calculated using linear time-varying analysis.

Consider a circuit whose input is the sum of two periodic signals, $u(t) = u_{\rm L}(t) + u_{\rm s}(t)$, where $u_{\rm L}(t)$ is an arbitrary periodic waveform with period $T_{\rm L}$ and $u_{\rm s}(t)$ is a sinusoidal waveform of frequency $f_{\rm s}$ whose amplitude is small. In this case, $u_{\rm L}(t)$ represents the large clock signal and $u_{\rm s}(t)$ represents the small information signal.

Let $v_{\rm L}(t)$ be the steady-state solution waveform when $u_{\rm s}(t)$ is zero. Then allow $u_{\rm s}(t)$ to be nonzero but small. We can consider the new solution $v(t)$ to be a perturbation $v_{\rm s}(t)$ on $v_{\rm L}(t)$, as in $v(t) = v_{\rm L}(t) + v_{\rm s}(t)$. The small-signal solution $v_{\rm s}(t)$ is computed by linearizing the circuit about $v_{\rm L}(t)$, applying $u_{\rm s}(t)$, and then finding the steady-state solution. Given that

$$u_{\rm s}(t) = U_{\rm s}\, e^{j2\pi f_{\rm s} t} \tag{40}$$

the perturbation in steady-state response is given by

$$v_{\rm s}(t) = \sum_{k=-\infty}^{\infty} V_{sk}\, e^{j2\pi(f_s + k f_L)t} \tag{41}$$

where $f_{\rm L} = 1/T_{\rm L}$ is the large-signal fundamental frequency [39], [61]. V_{sk} represents the sideband for the kth harmonic of $V_{\rm L}$. In this situation, shown in Fig. 11, there is only one sideband per harmonic because $U_{\rm s}$ is a single frequency complex exponential and the circuit has been linearized. This representation has terms at negative frequencies. If these terms are mapped to positive frequencies, then the sidebands with $k < 0$ become lower sidebands of the harmonics of $v_{\rm L}$ and those with $k > 0$ become upper sidebands.

$V_{sk}/U_{\rm s}$ is the transfer function for the input at $f_{\rm s}$ to the output at $f_{\rm s} + k f_{\rm L}$. Notice that with periodically varying linear systems there are an infinite number of transfer functions between any particular input and output. Each represents a different frequency translation.

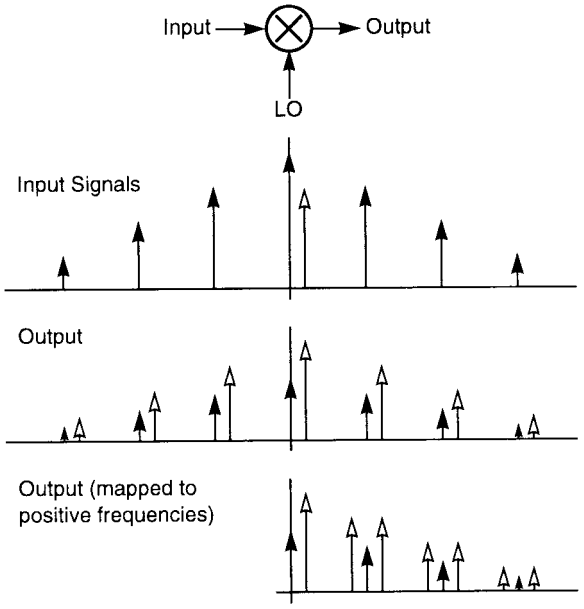

Fig. 11. The steady-state response of a linear periodically varying system to a small complex exponential stimulus. The large signals are represented with solid arrows and the small signals with hollow arrows.

Versions of this type of small-signal analysis exist for both harmonic balance [17], [24], [31] and shooting methods [39], [40], [61].

There are two different ways of formulating a small-signal analysis that computes transfer functions [59], [61]. The first is akin to traditional ac analysis and is referred to here as a periodic ac (PAC) analysis. In this case, a small-signal is applied to a particular point in the circuit at a particular frequency, and the response at all points in the circuit and at all frequencies is computed. Thus, in one step one can compute the transfer function from one input to any output. It is also possible to do the reverse, compute the transfer functions from any input to a single output in one step using an "adjoint" analysis. This is referred to as a periodic transfer function (PXF) analysis. PAC is useful for predicting the output sidebands produced by a particular input signal, whereas PXF is best at predicting the input images for a particular output.

Small-signal analysis is also used to perform cyclostationary noise analysis [8], [40], [52], which is an extremely important capability for RF designers [57]. It is referred to as a periodic noise (PNoise) analysis and is used to predict the noise figure of mixers. PNoise analysis is also used to predict the phase noise of oscillators. However, this is a numerically ill-conditioned problem that requires special techniques in order to overcome the ill conditioning and accurately compute close-in phase noise [21], [22].

LPV analyses provide significant advantages over trying to get the same information from equivalent large-signal analyses. First, they can be much faster. Second, a wider variety of analyses are available. For example, noise analysis is much easier to implement as a small-signal analysis. Last, they can be more accurate if the small signals are very small relative to the large signals. Small signals applied in a large-signal

analysis can be overwhelmed by errors that stem from the large signals. In a small-signal analysis, the large and small signals are applied in different phases of the analysis. Small errors in the large-signal phase typically have only a minor effect on the linearization and hence the accuracy of the small-signal results.

All of the small-signal analyses are extensible to the case where the operating point is quasi-periodic. This is important when predicting the effect of large interferers or blockers and is discussed further in Section VI-C. Such analyses are referred to as linear quasi-periodically varying (LQPV) analyses as a group, or individually as QPAC, QPXF, QPNoise, etc.

C. Transient-Envelope Analyses

Transient-envelope analyses are applied to simulate modulated carrier systems when the modulation is something other than a simple sinusoid or combination of sinusoids. They do so by performing a series of linked large-signal pseudoperiodic analyses, which are periodic analyses that have been modified to account for slow variations in the envelope over the course of each period of the carrier as a result of the modulation. The pseudoperiodic analyses must be performed often enough to follow the changes in the envelope. In effect, transient-envelope methods wrap a conventional transient analysis algorithm around a modified version of a periodic analysis. Thus, the time required for the analysis is roughly equal to the time for a single periodic analysis multiplied by the number of time points needed to represent the envelope. If the envelope changes slowly relative to the period of the carrier, then transient-envelope simulation can be very efficient relative to traditional transient analysis.

Transient-envelope methods have two primary applications. The first is predicting the response of a circuit when it is driven with a complicated digital modulation. An important problem is to determine the interchannel interference that results from intermodulation distortion. Simple intermodulation tests involving a small number of sinusoids as can be performed with quasi-periodic analysis are not a good indicator of how the nonlinearity of the circuit couples digitally modulated signals between adjacent channels. Instead, one must apply the digital modulation, simulate with transient-envelope methods, and then determine how the modulation spectrum spreads into adjacent channels.

The second important application of transient-envelope methods is to predict the long-term transient behavior of certain RF circuits. Examples include the turn-on behavior of oscillators, power supply droop or thermal transients in power amplifiers, and the capture and lock behavior of phase-locked loops. Another important example is determining the turn-on and turn-off behavior of time-division multiple access transmitters, which broadcast during a narrow slice of time. During that interval the transmitter must power up, stabilize, send the message, and then power down. If it powers up and down too slowly, the transmitter does not work properly. If it powers up and down too quickly, the resulting spectrum will be too wide to fit in the allotted channel. Simulating with traditional transient analysis would be prohibitively expensive because the time slice lasts on the order of 10–100 ms and the carrier frequency is typically at 1 GHz or greater.

1) Fourier Envelope Method: With the Fourier envelope method, the envelope is represented by slowly varying Fourier coefficients. First developed by Sharrit and referred to as circuit envelope [55], the Fourier envelope method is a transient-envelope method based on harmonic balance. In Section V-A, the concept of harmonic balance with time-varying Fourier coefficients was introduced. In that case, the Fourier coefficients were assumed to be periodic, with the result that signals themselves were quasi-periodic. With the Fourier envelope method [12], [38], [48], the Fourier coefficients in (20) are time varying but are not necessarily periodic. Instead, the Fourier coefficients $\tilde{X}_k(t)$ are taken to be slowly varying transient waveforms. Thus, signals take the form

$$x(t) = \sum_{k=-\infty}^{\infty} \tilde{X}_k(t)\, e^{j2\pi k f_t} \tag{42}$$

where f is the fundamental frequency of the base Fourier series. $\tilde{X}_k(t)$ represents the complex modulation of the kth harmonic. $\tilde{X}_k(t)$ must vary slowly relative to f because if the bandwidth of $\tilde{X}_k$ becomes greater than $f/2$, then the sidebands of adjacent harmonics begin to overlap and the representation is not unique.

Now, rewrite (16) assuming that v and f take the form of (42)

$$\sum_{k=-\infty}^{\infty} \tilde{F}_k\Big(\tilde{V}(t),\, t\Big)\, e^{j2\pi k f t} = 0 \tag{43}$$

where

$$F_k\Big(\tilde{V}(t),\, t\Big) = \frac{d\tilde{Q}_k\Big(\tilde{V}(t)\Big)}{dt} + j2\pi k f \tilde{Q}_k\Big(\tilde{V}(t)\Big) + \tilde{I}_k\Big(\tilde{V}(t)\Big) + \tilde{U}_k. \tag{44}$$

Assume that the variations in $\tilde{V}(t)$ are slow enough so that the bandwidth of each term in (44) is much less than $f/2$; then the terms associated with each harmonic k will sum to zero individually. Then $\tilde{F}_k(\tilde{V}(t),\, t) = 0$ for each k, or in vector form

$$\tilde{F}\Big(\tilde{V}(t),\, t\Big) = \frac{d\tilde{Q}\Big(\tilde{V}(t)\Big)}{dt} + \Omega\tilde{Q}\Big(\tilde{V}(t)\Big) + \tilde{I}\Big(\tilde{V}(t)\Big) + \tilde{U}(t) = 0 \tag{45}$$

where Ω is a diagonal matrix with $j2\pi kf$ on the kth diagonal. As with transient analysis, discretization methods such as trapezoidal rule or the backward difference formulas replace dQ/dt with a finite-difference approximation, converting (45) to a system of nonlinear algebraic equations that is solved with Newton's method. For example, applying backward Euler converts (45) to

$$\frac{\tilde{Q}\Big(\tilde{V}(t_m)\Big) - \tilde{Q}\Big(\tilde{V}(t_{s-1})\Big)}{t_m - t_{s-1}} + \Omega\tilde{Q}\Big(\tilde{V}(t_m)\Big) + \tilde{I}\Big(\tilde{V}(t_m)\Big) + \tilde{U}(t_m) = 0. \tag{46}$$

$\tilde{I}(\tilde{V}(t_m))$ and $\tilde{Q}(\tilde{V}(t_m))$ are evaluated at t_s by converting $\tilde{V}(t_m)$ into the time domain using the inverse Fourier transform, passing the time-domain voltage waveform through $i(\cdot)$ and $q(\cdot)$, and converting the resulting current and charge waveforms back into the frequency domain using the forward Fourier transform. This procedure relies on the envelope's being essentially constant over the length of a cycle of the carrier.

One of the important strengths of harmonic balance is its ability to easily incorporate frequency-domain models for the linear components such as lossy or dispersive transmission lines. Unfortunately, this is not true with the Fourier envelope method. The transient nature of the modulations $\tilde{V}_k(t)$ introduces the same difficulties that are present with distributed components in transient analysis, which are addressed using similar techniques. In particular, one can use convolution [20], [54], or the model for the distributed component can be separated into delay and dispersion, with the dispersion being replaced by a lumped approximation [56].

As the distributed components are linear, the sidebands for each harmonic k can be treated individually. Thus, a separate model is generated for each harmonic k, which greatly reduces the bandwidth requirements on the models. The model for each harmonic must only be valid over the bandwidth of the sidebands associated with that harmonic. In RF circuits, the bandwidths of the sidebands are usually small relative to the carrier frequency, and so generating models of distributed components for use in the Fourier envelope method is much easier than for conventional transient analysis. In fact, it is not uncommon for the bandwidth of an RF circuit to be so small that the transfer function of a distributed component does not change appreciably over the bandwidth of the sidebands. In this case, the transfer function is taken to be constant.

2) Sample Envelope Method: With the sample envelope method, the envelope is represented by a slowly varying sampled version of the waveform, as shown in Fig. 10. First developed by Petzold and referred to as envelope following [41], it is a transient-envelope method based on shooting methods. It approximates the sample envelope of Fig. 10 as a piecewise polynomial [25], [41] in a manner that is analogous to conventional transient analysis. This approach is efficient if the sequence formed by sampling the state at the beginning of each clock cycle, $v(0)$, $v(T)$, $v(2T)$, ..., $v(mT)$, changes slowly as a function of m. A "differential-like" equation is formed from (27)

$$\Delta v(mT) = \phi_T(v(mT), mT) \tag{47}$$

where $\Delta v(mT) = v((m+1)T) - v(mT)$ is a measure of the time-derivative of the sample envelope at mT. We can apply traditional integration methods to compute an approximation to the solution using a procedure that involves solving (47) at isolated time points. If the sample envelope is accurately approximated by a low-order polynomial, then this procedure should allow us to skip many cycles and so find the solution over a vast number of cycles in an efficient manner. For RF circuits, (47) is stiff and so requires implicit integration methods such as backward Euler, which can be written as

$$\Delta v(mT) \approx \frac{v(mT) - v((m-l)T)}{l} \tag{48}$$

where l is the timestep, which is measured in terms of cycles. This equation represents a two-point boundary constraint on (47), and so together they can be solved with shooting-Newton methods to find $v(mT)$. If desired, other integration methods can be used, such as the backward-difference formula.

As with transient analysis, once $v(mT)$ is computed, it is necessary to check that the trajectory is following the low-order polynomial as assumed. If not, the point should be discarded and the step l should be reduced. If rapid changes in the envelope are encountered, envelope following can reduce its stepsize down to the point where no cycles are skipped, in which case envelope following reduces to simple transient analysis. Thus, envelope following does not suffer the accuracy problems of the Fourier envelope method when small timesteps are taken to resolve a rapidly changing envelope.

D. Other Methods

The methods described above are either currently available or expected to be available soon in the mainstream commercial RF simulators. However, there other methods that have the potential to become significant to RF designers. In particular, two families of methods seem promising: the Volterra methods [31] and the multirate partial differential equation (MPDE) methods [4], [51].

The Volterra methods are similar in concept to the small-signal analyses in that they represent the circuit using a Taylor series expansion, except they take into account more than just the first term in the expansion. In this way, Volterra methods can efficiently compute the response of circuits exhibiting a small amount of distortion.

MPDE represents a family of methods based on the idea of replacing the single time variable with a sum of time variables, one for each of the time scales in the circuit. Consider a mixer with a 1-GHz LO and a 100-MHz IF. Then t would be replaced with $t = t_1 + t_2$, where t_1 is associated with the LO and t_2 is associated with the IF. The underlying ordinary differential equations that describe the circuit are reformulated as partial differential equations in t_1 and t_2. The various MPDE methods are formulated by applying particular boundary conditions and numerical methods to the t_1 and t_2 dimensions. For example, quasi-periodic harmonic balance from Section IV-A is an MPDE method that applies periodic boundary conditions and harmonic balance to both dimensions. Similarly, the Fourier envelope method from Section IV-C is an MPDE method that applies a periodic boundary condition and harmonic balance to the t_1 dimension and an initial condition and transient analysis to the t_2 dimension. Many other variations are possible.

V. Comparing the Methods

All of the methods presented can be grouped into two broad families: those based on harmonic balance and those based on shooting methods. Most of the differences between the methods emanate from the attributes of the base methods:

harmonic balance and shooting methods. So only the base methods will be compared.

A. Linear Passive Component Models

The main strength of harmonic balance is its natural support for linear frequency-domain models. Distributed components such as lossy and dispersive transmission lines and interpolated tables of S-parameters from either measurements or electromagnetic simulators are examples of linear models that are handled easily and efficiently with harmonic balance.

The difficulty with which shooting methods handle distributed component models contrasts sharply with harmonic balance. The problem is that the state vector associated with distributed components is infinite-dimensional. The state vector must somehow be discretized before shooting methods can be applied. However, even then shooting methods will be expensive if the state vector is large [27], [28]. This disadvantage explains why shooting-method-based RF simulation techniques have mainly been applied to radio-frequency integrated circuits (RFIC's). Most RFIC's can be modeled completely with lumped components. New approaches for generating lumped equivalent models for distributed components [35], [42], and components described in the frequency domain such as with tables of S-parameters [6], are becoming available that are more reliable and effective than existing methods, allowing shooting methods to be applied to circuits that contain a small number of distributed components.

B. Nonlinearity

Harmonic balance is very accurate and very efficient if the circuit is near linear and the voltage and current waveforms are near sinusoidal. In fact, assuming that the component models are correct, harmonic balance becomes exact in the limit where the circuit is linear and the stimulus are sinusoidal. This is not true for shooting methods. However, this feature is generally only significant when trying to determine the distortion of low-distortion amplifiers and filters. It does not help when analyzing mixers, oscillators, and sampling circuits because these circuits contain signals that are far from sinusoidal.

Harmonic balance can struggle on strongly nonlinear circuits or circuits that contain signals with abrupt transitions. Such signals are common in RF circuits. For example, mixers are driven with an LO that resembles a square wave, and even sinusoidal oscillators contain current waveforms that are narrow pulses. In this case, many frequencies are needed to accurately represent the signal, which increases the expense of harmonic balance. In addition, the magnitude of the harmonics drops slowly for signals with sharp transitions, making it difficult to know how many harmonics must be computed by harmonic balance. If too few harmonics are included, the results are inaccurate; if too many are included, the simulations can be impractical.

Harmonic balance is also susceptible to convergence problems when applied to strongly nonlinear circuits. Convergence can be improved by employing continuation or homotopy methods [3], [27]. These methods initially reduce the power of the input signal until convergence is achieved. Then the power is stepped up in a sequence of harmonic balance analyses, where the result computed at one step is used as the starting point for the next to improve convergence. In this way, harmonic balance can be made robust. However, because continuation methods end up calling harmonic balance tens, or perhaps hundreds, of times, they can be slow.

In contrast to harmonic balance, the ability of shooting methods to handle circuits strongly is quite good. The strengths of shooting methods stem from the properties of its underlying transient analysis. In particular, it chooses nonuniform timesteps in order to control error, and it has excellent convergence properties.

The ability of transient analysis, and so shooting methods, to place time points in a nonuniform manner allows it to accurately and efficiently follow abruptly discontinuous waveforms. Small timesteps can be used to accurately resolve rapid transitions without taking small steps everywhere. This is important for circuits such as mixers, relaxation oscillators, switched-capacitor filters, and sample-and-holds. In addition, the timestep is automatically chosen to control error. With harmonic balance, the timestep is constrained to be uniform by the fast Fourier transform (FFT); however, there is work that explores the possibility of using new FFT algorithms that do not require equally spaced points with harmonic balance [10], [37].

The strong convergence properties of shooting methods result from its implementation as a multilevel Newton method and not from the fact that it is a time-domain method. Indeed, it is possible to formulate harmonic balance as a time-domain method [27], [60], yet its convergence properties do not fundamentally change. As described in Section V-A, shooting methods apply Newton's method to solve

$$\phi_T(v(0), 0) = v(0) \tag{49}$$

for $v(0)$. $\phi_T(v(0), 0)$ relates the initial state of the circuit $v(0)$ to the state one period later. Newton's method is applied to solve (49) and is both efficient and reliable if ϕ_T is a near-linear function. This is usually the case even when the underlying circuit is behaving in a strongly nonlinear fashion because ϕ_T is evaluated over exactly a period of the large periodic clock signal, the signal that is driving the circuit to behave nonlinearly. Evaluating ϕ_T itself still involves solving strongly nonlinear sets of equations; however, that is done using transient analysis, a natural continuation method, and so is quite robust.

The ability of shooting methods to converge on a large class of strongly nonlinear circuits without the need for continuation methods or other convergence aids represents a significant advantage in efficiency over harmonic balance.

With shooting methods, it is natural to perform transient analysis for a while before starting the shooting iteration in order to generate a good starting point. This is usually sufficient to get convergence even on troublesome circuits except when the time constants in the circuit are much larger than the period of the signal. If this is not sufficient, one can also use continuation methods with shooting methods. The initial transient analysis has the side benefit that it helps to identify circuits that are unexpectedly unstable.

VI. RF Measurements

Simulators are used to predict the performance of RF circuits before they are actually constructed. This section introduces several of the most common RF measurements used to verify performance, with a description of how these measurements are made using an RF simulator. The measurements presented are representative of the most important and common measurements being made on RF circuits.

A. Transfer Functions

1) Conversion Gain: Conversion gain is the generalization of gain to periodically varying circuits such as mixers. It is simply the small-signal gain through a mixer as a function of frequency. Typically, conversion gain refers to the transfer function from the desired input to the desired output. But there are many other transfer functions of interest, such as the gain from an undesired image or from an undesired input such as the LO, power, and bias supplies.

Remember that the output signal for a periodically varying circuit such as a mixer may be at a different frequency than the input signal. The transfer functions must account for this frequency conversion. As described earlier, these circuits may have many images, and so for a single output frequency there may be many transfer functions from each input.

One measures a transfer function of a mixer by applying the LO, computing the steady-state response to the LO alone, linearizing the circuit about the LO, applying a small sinusoid, and performing one of the LPV analyses described in Section IV-B, such as PAC or PXF. One might also want to measure the transfer function with a large interferer present. If the interferer is assumed periodic, then the circuit would be linearized about the quasi-periodic response to both the LO and the interferer, and a LQPV analysis such as QPAC or QPXF is performed.

Actual measurements on mixers have shown that it is possible to predict conversion gain to within 0.25 dB [7].

2) AM and PM Conversion: As shown in Fig. 11, when a small sinusoid is applied to a periodically driven or clocked circuit, the circuit responds by generating both the upper and lower sidebands for each harmonic. The sidebands act to modulate the harmonics, or carrier, and the relationship between the sidebands and the carrier determines the character of the modulation. In Fig. 12, both the carrier and its sidebands are shown as phasors [50]. Assume that the sidebands are small relative to the carrier and that the circuit is driven at baseband with a small sinusoid with a frequency of f_m. The sideband phasors rotate around the end of the carrier phasor at a rate of f_m, with the upper sideband rotating one way and the lower rotating the other. The composite of the sideband phasors traces out an ellipse, as shown in Fig. 12(b). However, if the two sidebands have identical amplitudes and their phase is such that they align when parallel to the carrier, the phase variations from each sideband cancel, with the result being pure amplitude modulation (AM), as shown in Fig. 12(c). If instead the amplitudes are identical but the phases align when perpendicular to the carrier, then the amplitude variations cancel and the result is almost purely a phase modulation (PM),

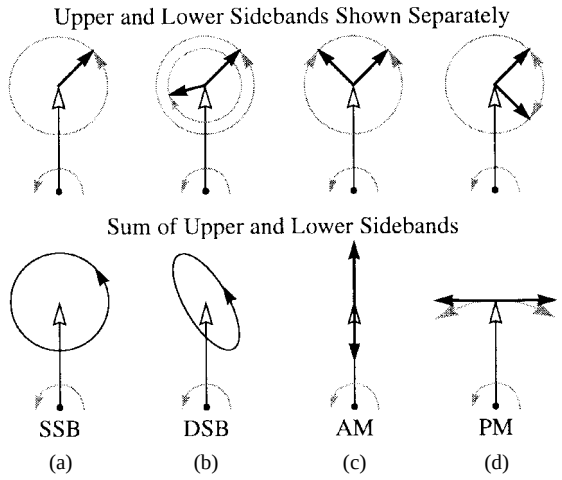

Fig. 12. How the amplitude and phase relationship between sidebands causes AM and PM variations in a carrier. The phasors with the hollow tips represent the carrier, and those with the solid tips represent the sidebands. The upper sideband rotates in the clockwise direction and the lower in the counterclockwise direction. The composite trajectory is shown below the individual components: (a) single-sideband modulation (only upper sideband), (b) arbitrary double-sideband modulation where there is no special relationship between the sidebands, (c) amplitude modulation (identical magnitudes and phase such that phasors point in the same direction when parallel to carrier), and (d) phase modulation (identical magnitudes and phase such that phasors point in the same direction when perpendicular to carrier).

as shown in Fig. 12(d) (assuming the sidebands are small). The double-sideband (DSB) modulation shown in Fig. 12(b) can be considered a combination of both AM and PM modulation.

AM and PM conversion occurs either when a tone is injected at either baseband or at a sideband. The former is referred to as baseband-to-AM/PM conversion and the latter is SSB-to-AM/PM conversion. Both cases were demonstrated in the case of an oscillator by Razavi [43].

A PAC analysis directly computes the transfer function from some small input signal to the upper and lower sidebands components of a modulated carrier. It is also possible, using a change of basis, to recast these transfer functions in terms of the AM and PM components of the modulation [53]. To show this, consider a circuit that is generating a sinusoidal carrier. Assume that the carrier is both amplitude and phase modulated by small complex exponentials at the same frequency f_m. The resulting signal would take the form

$$v_m(t) = A_c(1+\alpha(t))\cos(2\pi f_c t + \phi_c + \phi(t)) \tag{50}$$

where $\alpha(t) = Ae^{j2\pi f_m t}$ is the amplitude modulation and $\phi(t) = \Phi e^{j2\pi f_m t}$ is the phase modulation. Both A and Φ are complex coefficients. Using the narrow-band angle modulation approximation [66], (50) can be expanded into a sum of complex exponentials in order to identify the upper and lower sidebands

$$l(t) = Le^{j2\pi(f_m - f_c)t} \tag{51}$$

$$u(t) = Ue^{j2\pi(f_m + f_c)t} \tag{52}$$

where $L = (A - j\Phi)e^{-j\phi_c}$ and $U = (A + j\Phi)e^{j\phi_c}$. These can be rearranged to give A and Φ in terms of L and U

$$A = \left(Le^{j\phi_c} + Ue^{-j\phi_c}\right)/2 \tag{53}$$

$$\Phi = j\left(Le^{j\phi_c} - Ue^{-j\phi_c}\right)/2. \tag{54}$$

Thus, given the phase of the carrier ϕ_c, which can be computed with a periodic steady-state analysis, and the transfer functions from the input to the upper and lower sidebands L and U, which can be computed with a periodic ac analysis, one can compute the to-AM (A) and to-PM (Φ) transfer functions.

If the to-FM transfer function is desired instead, let $\omega(t) = \Omega e^{j2\pi f_m t}$ be the modulation signal where in (50) ϕ becomes

$$\phi(t) = \int \omega(t)\, dt. \tag{55}$$

Then the to-FM transfer function is $\Phi = j\omega_m \Phi$, or

$$\Omega = 2\pi f_m \left(U e^{-j\phi_c} - L e^{j\phi_c}\right)/2. \tag{56}$$

3) Oscillator Load Pull: Load pull refers to shifts in the frequency of an oscillator as a function of changes in load impedance, supply voltage, substrate, bias lines, etc. A change in load impedance represents a parametric change in the circuit and so requires a full periodic steady-state analysis to compute the response in the oscillation frequency. However, as long as changes in the signal levels on supplies, substrates, and bias lines are small, sensitivity of the oscillator frequency to perturbations of this type can be computed using the technique given above for computing the to-FM transfer function.

B. Cyclostationary Noise

With clocked systems, there are two effects that act to complicate noise analysis. First, for noise sources that are bias dependent, such as shot-noise sources in bipolar junction transistors or the thermal noise of MOSFET's, the time-varying operating point acts to modulate the noise sources. Such noise sources are referred to as being cyclostationary. Second, the transfer function from the noise source to the output is also periodically varying and so acts to modulate the contribution of the noise source to the output. In this case, even if the noise source were stationary, as it would be for thermal noise of a linear time-invariant resistor, the noise at the output is cyclostationary.

Modulation is a multiplication of signals in the time domain and so in the frequency domain, the spectrum of the noise source is convolved with the spectrum of the transfer function [68]. The transfer function is periodic or quasi-periodic and so has a discrete line spectrum. Convolution with a discrete spectrum involves a series of scale, shift, and sum operations, as shown for a mixer in Fig. 13. The final result is the sum of the noise contributions from each source both up-converted and down-converted by the harmonics of the LO to the desired output frequency. This is referred to as noise folding.

Periodic modulation of a stationary noise source, either from a periodic bias or from a periodically varying signal path from the source to the output, results in cyclostationary noise at the output. In stationary noise, there is no correlation between noise at different frequencies. As can be seen from Fig. 13, at frequencies separated by kf cyclostationary noise is correlated, where f is the modulation frequency and k is an integer [15].

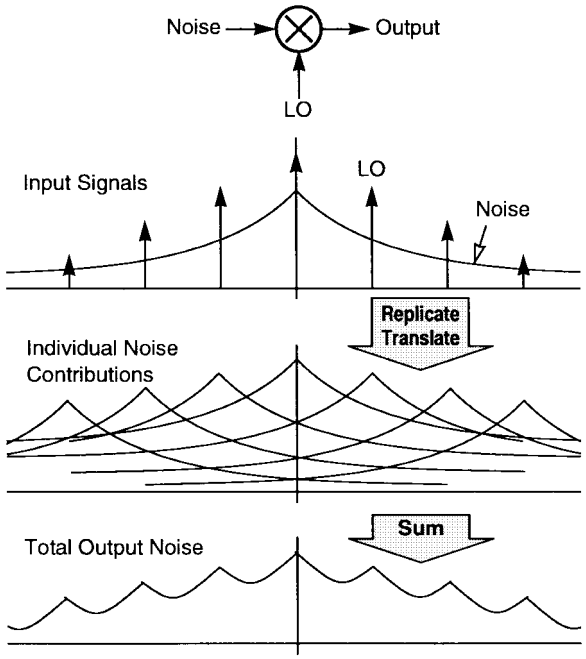

Fig. 13. How noise is moved around by a mixer. The noise is replicated and translated by each harmonic of the LO, resulting in correlations at frequencies separated by kf_{LO}.

1) Noise Figure: Noise is a critical concern in receivers because of the small input signals. Typically, designers characterize the noise of individual blocks using the noise figure (NF) of the block because it is relatively simple to combine the noise figure of cascaded blocks to determine the noise figure of the entire receiver [44], [65]. The NF of a block is a measure of how much the signal-to-noise ratio (SNR) degrades as the signal passes through the block. It is defined as

$$\mathrm{NF} = 10 \log \frac{\mathrm{SNR_{in}}}{\mathrm{SNR_{out}}}. \tag{57}$$

At the input of a receiver, the SNR is defined as the signal power relative to the background noise power picked up by the antenna. Fig. 6 shows that a receiver is sensitive to noise at the input at each of its images. $\mathrm{SNR_{in}}$ only includes the noise power in the images where the input signal is found. In heterodyne receivers, the input signal is found in a single sideband, and so SSB NF is employed. In this case, $\mathrm{SNR_{in}}$ includes only the noise power in the image associated with the input. In homodyne, or direct-conversion, receivers, the signal is found in both sidebands of the carrier, and so DSB NF is employed. In this case, the images associated with the carrier overlap, and $\mathrm{SNR_{in}}$ includes the noise power from both. In both cases, $\mathrm{SNR_{out}}$ includes the effect of the input noise from all images; however, it excludes the noise generated in the load at the output frequency.

In (57), the signal power is both in the numerator and the denominator and so cancels out. Thus, (57) can be rewritten

$$\mathrm{NF} = 10 \log \frac{P_O - P_L}{P_S} \tag{58}$$

where P_O is the total output noise power, P_{L} is the output noise power that results from noise generated by the load

at the output frequency, and P_{S} is the output noise power that results from noise generated by the source at the input frequency. Generally, the P_{S} used when computing DSB NF is twice as large as the one used with computing SSB NF, whereas P_O and P_{L} are both the same. Thus the DSB NF is usually 3 dB smaller than the SSB NF for the same circuit.

One computes the noise of a mixer by applying the LO, computing the steady-state response to the LO alone, linearizing the circuit about the LO, and applying the PNoise analysis of Section IV-B. Actual measurements on bipolar mixers have shown that noise figure can be predicted to within 0.25 dB [7], [34]. Results are not expected to be as accurate on CMOS mixers because the noise model for MOS transistors is not as accurate as the one for bipolar transistors.

2) Impact of Cyclostationarity on Subsequent Stages: In general, clocked circuits such as oscillators and mixers produce cyclostationary noise, which implies that the noise statistics, such as the PSD, denoted $S(f, t)$, varies as a periodic or quasi-periodic function of time. If a spectrum analyzer is used to observe this noise, and if the frequency of the noise variation is much faster than the analyzer can track, then the spectrum analyzer will measure the time-average PSD. Of course, the time-average PSD is not a complete characterization of the noise, but often it is sufficient. An important question is: When is the time average PSD sufficient to characterize the output noise of a circuit, and when is it not?

This question can be answered with the help of the following observation. If an uncertainty in time is introduced into the cyclostationary process—a uniformly distributed random variable from zero to T is added to t—the resulting process is stationary and its statistics are the time average of the statistics of the cyclostationary process [15]. Similarly, if the cyclostationary process is input to a system that does not track the variation of the PSD with time, then the phase of the variation is unknown to the system. In the absence of information about the phase of the variation, the process becomes stationary, with the PSD equal to the time average of $S(f, t)$ [62].

There are two common situations that would cause a subsequent stage to track the variations of a cyclostationary process. The first occurs if the signal driving the subsequent stage is large enough to generate a nonlinear response. This happens, for example, when an oscillator drives a limiter. The oscillator signal is large enough to drive the limiter into a nonlinear region, causing the characteristics of the limiter to track the variations in the cyclostationary noise produced by the oscillator. The same is true when an oscillator drives a mixer. The second situation is when both circuits are being driven by large signals derived from the same reference. This would occur if, for example, the output of one mixer were fed to the input of another and both were driven by the same LO. Because they are driven by the same LO, the second mixer is synchronous with, and tracks the variations in the cyclostationary noise of, the first mixer. In both of these cases, just knowing the time average of $S(f, t)$ is not sufficient to predict the noise performance of the entire system. In particular, knowing the time-averaged noise figure of each of the two mixers does not give sufficient information to predict the noise figure of the cascaded pair.

However, in the case of the two mixers, if the second mixer were driven with an independent LO, even one that was close in frequency to the first LO, then the phase drift between the two LO's would cause the synchronism between the two mixers to be lost, with the result that using the time-average noise statistics is adequate when predicting the noise performance of the whole system. Thus, in this case, if one knows the time-averaged noise figures for the two mixers, then one can predict the noise figure of the combination using standard formulas [44], [65].

It is common to combine two mixers in cascade in a superheterodyne receiver architecture and to generate the LO signals for the two mixers by using two phase-locked loops (PLL's) with a common reference frequency. However, it is interesting to note that even in this case, if the ratio of the LO frequencies is m/n, where m and n are relatively large integers (often $m, n \geq 4$ is sufficient) with no common factors, then using the time average of the noise at the output of the first mixer will usually introduce little error when estimating the noise at the output of the second mixer [62]. In addition, interstage filtering also acts to reduce the chance of error. Remember that noise sidebands must be correlated for the noise to be cyclostationary. Filtering can convert a cyclostationary signal to a stationary signal if the filter's bandwidth and center frequency are such that it eliminates all but one sideband.

3) AM and PM Noise: As shown in Fig. 13, clocked circuits generate noise with correlated sidebands. And as shown in Fig. 12, depending on the magnitude and phase of the correlation, the noise at the output of the circuit can be AM noise, PM noise, or some combination of the two. For example, oscillators almost exclusively generate PM noise near the carrier, whereas noise on the control input to a variable-gain amplifier results almost completely in AM noise at the output of the amplifier.

This ability to emphasize one type of noise over another is a characteristic of clocked circuits and cyclostationary noise. Linear time-invariant circuits driven by stationary noise sources can only produce additive noise, which can be decomposed into AM and PM noise, but there will always be equal amounts of both.

To find the AM or PM noise of a carrier, one must perform PNoise analysis and output the noise at both the upper and lower sidebands of the carrier along with their complex correlation. The AM and PM components of the noise can then be computed using (53) and (54).

4) Oscillator Phase Noise: One can apply the PNoise analyses of Section IV-B to oscillators to compute their phase noise. Or one can apply the PXF analysis to determine the sensitivity of the output to small interfering signals such as those on the power supply. As indicated earlier in this section, these analyses are able to properly account for frequency conversions and for the fact that the noise in the output manifests itself largely as changes in the phase of the output.

These are small-signal analyses that assume that the circuit being analyzed does not respond in a nonlinear way to the small-signal inputs. However, (7) indicates that even small inputs can generate large changes in the phase if they are close

in frequency to the fundamental or to one of its harmonics. While the deviation in oscillator phase is generally a linear function of the input, the output voltage or current is a linear function of the phase only when the deviations in phase are small. If the phase changes by a significant fraction of a period, the small-signal assumption is violated and the response becomes a nonlinear function of the input. It is this nonlinear response that causes the linewidth of S_v and $\mathcal{L}$ [the rolloff at very low offset frequencies that is given in (12)]. As a result, the small-signal analysis results do not predict the linewidth and so are inaccurate for computing S_v and $\mathcal{L}$ at frequencies very close to the carrier or its harmonics. Thus, the small-signal PNoise analysis can only be used to compute S_v and $\mathcal{L}$ for f_m well above f_Δ. Usually, f_Δ is quite small, and so this is not considered an issue. PNoise analysis can be used to compute S_ϕ for all frequencies.

Actual measurements on oscillators have shown that it is possible to predict oscillator phase noise on bipolar resonant oscillators to within 2 dB [67].

5) Jitter and Phase Noise of PLL's: Oscillators are frequently encapsulated in phase-locked loops in order to stabilize their output frequency and reduce their phase noise and jitter by locking them to a more stable reference. The reference is usually a fixed frequency, whereas the oscillator may need to change its frequency, perhaps to allow the receiver to tune over a range of channels. Complex feedback schemes are often necessary to satisfy the often competing requirements of frequency tuning resolution and noise performance [14]. Predicting the phase noise and jitter of such circuits can be quite difficult. They rarely operate with simple periodic or quasi-periodic signals, and so the PNoise and QPNoise analyses cannot be directly applied. Even if they are periodic, the frequency ratio between the oscillator and the reference can be so large as to make these techniques impractical. Transient-envelope methods could be applied, but like simple transient analyses, the transient-envelope methods are not set up to include the effect of component noise sources, which play a large role in the noise and jitter of most PLL's. Attempts to include device noise in generic transient and transient-envelope analyses generally fail because of the small amplitude and wide bandwidth of device noise. Device noise is usually much smaller than the simulator error that masquerades as noise. To avoid this problem, tolerances must be set very tight. To accurately model wide-band noise without excessively coloring it, generally a very small and usually fixed timestep is needed. Both of these combine to make this approach impracticable in most cases.

PLL's simulations often occur at the behavioral level rather than at the transistor level because the transistor-level simulations are very expensive and because behavioral simulation of PLL's can be quite effective. If the behavioral models can be made to include the effect of component noise sources, they can be used to efficiently predict the noise and jitter of PLL's [8], [29]. To do so, it is first necessary to individually characterize the noise behavior of the blocks that make up the PLL using transistor-level simulation. For each block, representative periodic signals are applied and a PNoise analysis is performed. Then the jitter is extracted and provided as a parameter to behavioral models for inclusion in a high-level simulation of the entire PLL. The noise and jitter of the PLL can be obtained by observing the transition times in simulated output.

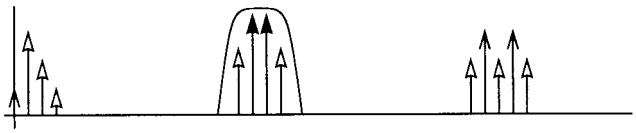

Fig. 14. A narrow-band circuit driven with two closely spaced sinusoidal tones (solid-tipped arrows) responds by generating harmonics (open-tipped arrows) and intermodulation (hollow-tipped arrows) tones. Distortion of the output signal results because several of the odd-order intermodulation tones fall within the bandwidth of the circuit.

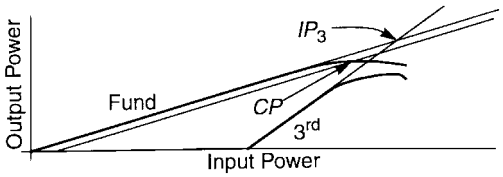

Fig. 15. The 1-dB compression point (CP_1) is the point where the output power of the fundamental crosses the line that represents the output power extrapolated from small-signal conditions minus 1 dB. The third-order intercept point (IP_3) is the point where the third-order term as extrapolated from small-signal conditions crosses the extrapolated power of the fundamental.

C. Intermodulation Distortion

Distortion is commonly measured in narrow-band circuits by applying two pure sinusoids with frequencies well within the bandwidth of the circuit (call these frequencies f_1 and f_2). The harmonics of these two frequencies would be outside the bandwidth of the circuit; however, there are distortion products that fall at the frequencies $2f_1 - f_2$, $2f_2 - f_1$, $3f_1 - 2f_2$, $3f_2 - 2f_1$, etc. As shown in Fig. 14 these frequencies should also be well within the bandwidth of the circuit and so can be used to measure accurately the intermodulation distortion (IMD) produced by the circuit.

1) Compression and Intercept Points: At low frequencies, it is common to describe the distortion of a circuit by indicating the distortion in the output signal when driven by a sinusoid to achieve a certain output level. At high frequencies, it is more common to characterize the distortion produced by a circuit in terms of a compression point or an intercept point. These metrics characterize the circuit rather than the signal, and as such it is not necessary to specify the signal level at which the circuit was characterized.

To measure the compression point of a circuit, apply a sinusoid to its input and plot the output power as a function of the power of the input. The 1-dB compression point is the point where the gain of the circuit has dropped 1 dB from its small-signal asymptotic value. This is illustrated in Fig. 15. $iCP_{1\,\mathrm{dB}}$ is the input power and $oCP_{1\,\mathrm{dB}}$ is the output power that corresponds to the 1-dB compression point. iCP is normally used for receivers and oCP for transmitters. The compression point is typically measured in dBm, which is decibels relative to one milliwatt.

To measure a two-tone nth-order IMD intercept point IP_n, apply two sinusoids to the circuit's input at f_1 and f_2. Make

the amplitude of the two sinusoids the same and increase their power while plotting the power at the output in a fundamental (either f_1 or f_2) and in an nth-order intermodulation product (for IP_3 use either $2f_1 - f_2$ or $2f_2 - f_1$). This is illustrated for IP_3 in Fig. 15. The nth-order intercept point IP_n is defined in terms of the power levels of the fundamentals and the nth-order products as extrapolated from their asymptotic small-signal behavior. When the input signal is small, a doubling of the input power doubles the fundamental output power and multiplies the output power of the nth-order products by 2^n. Thus, the asymptotic slope of the fundamental is 1 dB/dB, and the asymptotic slope of the nth-order products is n dB/dB. The nth-order intercept point (IP_n) is where the asymptotes for the nth-order intermodulation product and the fundamental cross. iIP_n is the input power and oIP_n is the output power corresponding to the intercept point. They are generally measured in dBm.

In practice, it is only necessary to sweep the input power to determine an appropriate input power P_{in} for an accurate extrapolation. P_{in} should be chosen small enough that P_1, the output power of the first-order term, and P_n, the output power of the nth-order term, are both in their asymptotic ranges. Additionally, P_{in} should be chosen large enough so that P_1 and P_n are computed accurately by the simulator. Once an appropriate P_{in} is applied and the corresponding P_1 and P_n found, the output intercept point is computed with

$$IP_n = P + \frac{\Delta P}{n-1} \tag{59}$$

where IP_n is the nth-order intercept point (dBm), P is the power in the fundamental in dBm, and ΔP is the difference between the desired output signal (P_1) and the undesired nth-order output product (P_n) in dB. P is the input power P_{in} if iIP_n is desired, and it is the output power P_1 if oIP_n is desired [65].

IP_3 is the most commonly used intercept point but others, including IP_2, IP_5, and IP_7, are also of interest.

Actual measurements on receivers have shown that it is possible to predict a 1-dB compression point and IP_3 to within 0.5 dB [7].

2) Rapid IP_2 and IP_3 Prediction: The traditional approach to measuring intermodulation distortion is to apply two large closely spaced test tones to the input and measure the intermodulation products. Both tones are large enough to generate a nonlinear response from the circuit. Thus, at a minimum, a two-tone QPSS analysis is required. If the circuit is clocked, such as a mixer, a third tone is also needed, in this case the LO. As such, it can be an expensive simulation. However, in those cases where the intermodulation products of interest fall at the frequencies $(n-1)f_1 \pm f_2$, it is not necessary for both test tones to be large to measure the intercept point. One tone (f_1) must be large to drive the circuit into a nonlinear region. The second tone (f_2) is needed only to produce an in-band response by mixing with the distortion products generated by the first tone. Thus, it is possible to measure IP_2 or IP_3 using a small-signal analysis [58], and doing so is considerably more efficient in terms of both time and memory than using a full large-signal quasi-periodic analysis.

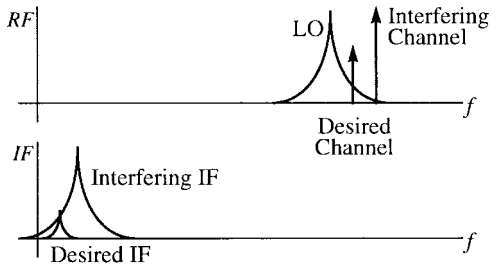

Fig. 16. In a receiver, the phase noise of the LO can mix with a large interfering signal from a neighboring channel and swamp out the signal from the desired channel even though most of the power in the interfering IF is removed by the IF filter. This process is referred to as *reciprocal mixing.*

To compute IP_2 or IP_3 of an amplifier using this approach, apply the first tone with power P (in dBm) and perform a PSS analysis to compute P_1 (in dBm), the power at the output in the fundamental frequency. Then apply the second tone as a small complex exponential at a sideband of the first tone and use a PAC analysis to compute ΔP (in dB), the ratio between the power at the output in the fundamental frequency of the small tone and the power at the output in the nth-order intermodulation product. IP_n is then computed with (59).

The intermodulation distortion of a clocked circuit such as a mixer is measured in a similar manner, except that both the LO and the first tone are applied together and the response computed with a QPSS analysis. The response to the second tone is computed using a QPAC analysis.

In practice, this approach is limited to computing IP_2 and IP_3, but given that restriction, it can be used in the same situations as the traditional approach and gets the same answer, but is considerably faster.

3) Blockers: A blocker is a large interfering signal in a nearby channel that acts to degrade the reception of the desired signal. It does so in two ways: by reducing the gain and increasing the noise floor of the receiver. A typical receiver, shown in Fig. 1, has both an RF and IF filter. The RF filter is always broader than the IF filter. Thus, a blocker is often not eliminated until it has passed through the front end of the receiver. If it is large enough to drive the front end into compression, the effective gain for the desired signal is reduced, which reduces the sensitivity of the receiver. The effect of a blocker on gain can be determined by applying the interferer, using a PSS or QPSS analysis to compute the time-varying operating point, and then performing either a PAC or QPAC analysis to compute the gain for the desired signal.

To see how a blocker acts to increase the noise floor, consider its effect when it interacts with the phase noise from the local oscillator of the mixer used in the front end of the receiver, as shown in Fig. 16. The phase noise of the LO is directly translated onto the mixer products in a process referred to as reciprocal mixing. Although the IF filtering at the output of the mixer may be sufficient to remove the blocker's main mixing product, the desired signal is masked by the phase noise of the LO that was down-converted by the blocker. Thus, the blocker acts to increase the noise floor of the receiver.

Reciprocal mixing is currently a difficult simulation because it is a semiautonomous analysis. In other words, it is a

TABLE I
VERILOG-A OSCILLATOR MODEL THAT INCLUDES PHASE NOISE

```
// Oscillator with Phase Noise

'include "discipline.h"
'include "constants.h"

module osc (out);

output out; electrical out;

parameter real freq=1 from (0:inf);
parameter real ampl=1;
parameter real pn=0;

real phase, phase_noise;

    analog begin
        phase = 2*'M_PI*freq*$realtime;
        phase_noise = flicker_noise(ph,2);
        V(out) <+ ampl*cos(phase + phase_noise);
    end
endmodule
```

quasi-periodic analysis where one fundamental, the blocker, is driven and one fundamental, the LO, is autonomous. While conceptually possible, no simulator is currently available that provides a semiautonomous analysis. So instead, one must replace the oscillator with a nonautonomous equivalent model and then perform a QPNoise analysis. The model must mimic the phase noise produced by the oscillator. A Verilog-A [64] model that includes phase noise is given in Table I. This particular module does not include flicker noise, even though it uses the *flicker_noise* function. Notice that the slope (the second argument) is two. This causes the function to generate noise whose power is proportional to f^{-2}, which matches the characteristic of oscillator noise that stems from white noise sources. It is not difficult to generalize this model to include the effect of flicker noise; simply add in another *flicker_noise* function, this time with a slope of three.

4) Spurious Responses: Spurious responses are undesired responses to signals at the input of the receiver that occur at frequencies that are different from the desired receive frequency. The undesired images of Section III-A are an example of spurious responses. However, there are additional frequency ranges that can result in spurious responses if the interfering signals encountered are at a sufficiently high level. Any RF signal with a frequency that satisfies the following relationship for any integers m and n could result in a spurious response

$$\pm m f_{\rm in} \pm n f_{\rm LO} = \pm f_{\rm IF} \tag{60}$$

where $f_{\rm in}$ is any input frequency, $f_{\rm LO}$ is the LO frequency, and $f_{\rm IF}$ is the desired IF frequency. Of particular importance is the half-IF spurious response in which (m, n) is either (2, -2) or (-2, 2). This occurs at $f_{\rm in} = (f_{\rm RF} + f_{\rm LO})/2$, midway between the desired input frequency $f_{\rm RF}$ and that of the LO. There are two possible causes for a half-IF response [44]. First is if the interfering input signal is subject to significant amounts of second-order distortion and the LO contains a significant second harmonic. In this case, the second harmonics of the interferer and the LO will mix and generate a response at the IF. The second possible cause is if the fundamental of the interferer mixes with the fundamental of the LO and the product, at $f_{\rm IF}/2$, is subject to significant second-order distortion. The half-IF response can be predicted from IP_2 measurements [65].

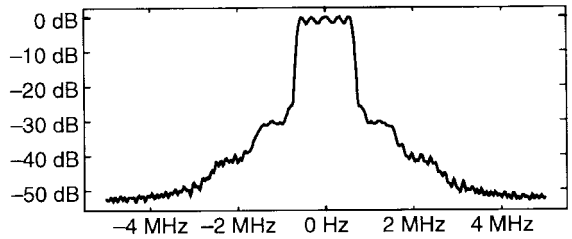

Fig. 17. Intermodulation distortion in a power amplifier spreads the bandwidth of the transmitted signal to the point where it can interfere with adjacent channels.

5) Spectral Regrowth and ACPR: A very important issue when transmitting digitally modulated signals is adjacent channel power (ACP). A transmitter should only emit power in its designated channel. Any power emitted in adjacent channels can interfere with the proper operation of nearby receivers that are attempting to receive signals from distant transmitters. As such, transmitters have strict adjacent channel interference or ACP requirements that they must satisfy. The low-pass filters in the transmitter of Fig. 2 are designed to limit the bandwidth of the transmitted sequence. However, if the mixers or power amplifier are nonlinear, intermodulation distortion can cause the bandwidth to grow back, as shown in Figs. 3 and 17. This effect is referred to as spectral regrowth. Unfortunately, this situation is difficult to quantify using simulation. Simple two-tone intermodulation tests are not representative of digitally modulated signals. Instead, the transmission of a long pseudorandom sequence of symbols is simulated. The output spectrum is calculated from a sequence that typically contains between 1- and 4-k symbols. The adjacent channel power is then characterized with the adjacent channel power ratio (ACPR)

$$\text{ACPR} = \frac{P_{\rm adj}}{P_{\rm in}} \tag{61}$$

where $P_{\rm adj}$ is the total power in the adjacent channel and $P_{\rm in}$ is the power in the desired channel.

The carrier frequencies are typically in the 1–5-GHz range and the symbols typically have a rate of 10–300 kHz. Such a simulation is clearly impractical for traditional transient analysis. Instead, the transient-envelope methods of Section IV-C are used. However, simulating a 1–4-k symbol sequence still requires between 10–100-k simulation points, each of which represents a harmonic balance or shooting method solve, and so even the transient envelope methods are very expensive for this type of simulation, particularly for a circuit with more than just a few components.

An alternative to transient-envelope simulations is to extract a behavioral baseband-equivalent table model of the transmitter using one of the large-signal steady-state methods described in Section IV-A [5]. Because the behavioral model abstracts away the carrier and unnecessary circuit details, the ACPR calculation step is fast regardless of circuit size or complexity.

In transmitter circuits, the input baseband signal is usually well within the transmitter's bandwidth, so a memoryless model often suffices. This is important because it makes generating the behavioral models fairly easy. A table model is constructed by simply exercising the circuit over a range of input amplitudes and phases and computing the steady-state responses. This approach is accurate if the bandwidth of the circuit is wide compared with the bandwidth of the signal. Given the high carrier frequencies (1 GHz and above) and the low signal bandwidths (1.23 MHz for the relatively broad-band IS-95 CDMA standard), this assumption is relatively safe. There is, however, one situation that will cause this assumption to be violated: if low-frequency dynamics in either the power-supply or the bias-supply lines significantly affect the characteristics of the transmitter's signal path. It is easy to check for this situation using a periodic transfer function analysis. Simply apply a large periodic signal to represent an unmodulated carrier, and then perform the PXF analysis over a range of frequencies centered about the carrier. If the transfer functions are relatively constant over the signal bandwidth, it is appropriate to use a memoryless table model. Otherwise, a more sophisticated model or a transient-envelope simulation is needed.

6) Triple Beats: The term *triple beat* refers to three-tone, third-order intermodulation distortion. Three-tone IMD involves terms of the form $f_1 \pm f_2 \pm f_3$, whereas the two-tone IMD discussed earlier involves terms of the form $2f_1 \pm f_2$ and $f_1 \pm 2f_2$. Interference caused by triple beats is 6 dB higher than that caused by two-tone IMD [65]. Thus, if there are only three tones, one can determine the size of the triple beats rather simply from a two-tone IMD measurement. However, if the number of tones becomes large, the number of triple beats becomes very large. The number of triple beats generated is $n(n-1)(n-2)/2$, where n is the number of tones. Even for a moderate number of tones, it becomes an arduous task to track the triple beats because of their numbers, which makes estimation of triple-beat interference using this approach difficult if not impossible when there are a large number of tones.

7) Multichannel Systems: In modern cellular phone systems, the handset transmitters are power controlled, so the power received at the basestation for each channel is roughly the same. The receivers used in cable tuners and repeaters and satellite communications also must operate properly in the presence of a multitude of large carriers. In this setting, distortion performance is often difficult to quantify because of the large number of intermodulation distortion products. If we consider only third-order nonlinearities, there are two types of intermodulation distortion of concern: two-tone IMD and three-tone IMD (triple beats). If there are a large number of equally spaced and equally sized carriers, then two-tone IMD is constant over the band while triple-beat distortion peaks in the middle of the band [65]. In addition, there are many more triple-beat products than two-tone products. These, combined with the triple beats' being twice as large as the two-tone products, result in triple-beat IMD limiting the dynamic range of multichannel systems more than two-tone IMD, particularly in the center of the band.

Fig. 18. Spurious-free dynamic range or noise power ratio of a multichannel system.

Dynamic range of multichannel systems is characterized using a spurious-free dynamic range (SFDR) or noise-power ratio (NPR) measurement, which is illustrated in Fig. 18. This measurement mimics the situation where all channels are being received at their maximum level. The transmitter assigned to one channel near the middle of the band (where triple-beat distortion is at its highest) is then turned on and off. SFDR or NPR is the ratio of the power in that channel when the transmitter is on relative to the power when it is off. When the transmitter is off, the power in the channel is due to intermodulation distortion caused by the other channels. SFDR or NPR represents the dynamic range available in that channel.

In cable multichannel systems, simple transient analysis is used to compute SFDR/NPR. However, in wireless multichannel systems, the frequency of the lowest channel is much greater than the channel spacing, and so transient-envelope methods would be used. However, as with ACPR, this can be a very expensive simulation, particularly if the carriers are modulated. And as with ACPR, there are many cases where the multichannel system being tested is broad-band. In this case, it is possible to generate a memoryless table model as before and efficiently evaluate the table model to predict SFDR/NPR. Again, PXF analysis can be used to verify the broad-band assumption.

ACKNOWLEDGMENT

The author would like to thank the following people for the many enlightening conversations about RF simulation and RF circuits that contributed to this article: J. Phillips, J. White, M. Terrovitis, D. Feng, K. Nabors, J. Chen, A. Abidi, F. Carr, K. Gard, G. Dawe, and A. Demir.

REFERENCES

[1] A. Abidi, "Radio frequency integrated circuits for portable communications," in *Proc. 1994 IEEE Custom Integrated Circuits Conf.,* May 1994.

[2] T. Aprille and T. Trick, "Steady-state analysis of nonlinear circuits with periodic inputs," *Proc. IEEE,* vol. 60, pp. 108–114, Jan. 1972.

[3] J. Bolstad and H. Keller, "A multigrid continuation method for elliptic problems with folds," *SIAM J. Sci. Statist. Comput.,* vol. 7, no. 4, pp. 1081–1104, Oct. 1986.

[4] H. Brachtendorf, G. Welsch, R. Laur, and A. Bunse-Gerstner, "Numerical steady state analysis of electronic circuits driven by multi-tone signals," in *Electrical Engineering.* Berlin, Germany: Springer-Verlag, 1996, vol. 79, pp. 103–112.

[5] J. Chen, D. Feng, J. Phillips, and K. Kundert, "Simulation and modeling of intermodulation distortion in communication circuits," in *Proc. 1999 IEEE Custom Integrated Circuits Conf.,* May 1999.

[6] C. Coelho, J. Phillips, and M. Silveira, "Robust rational function approximation algorithm for model generation," in *Proc. 36th Design Automation Conf.,* June 1999.

[7] G. Dawe, J. M. Mourant, and P. Brokaw, "A 2.7V DECT RF transceiver with integrated VCO," in *1997 IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers and Slide Supplement,* Feb. 1997.

[8] A. Demir and A. Sangiovanni-Vincentelli, *Analysis and Simulation of Noise in Nonlinear Electronic Circuits and Systems.* Norwell, MA: Kluwer Academic, 1997.

[9] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," in *Proc. 35th Design Automation Conf.*, June 1998.

[10] A. Dutt and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM J. Sci. Comput.*, vol. 14, no. 6, Nov. 1993.

[11] P. Feldmann and R. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 639–649, May 1995.

[12] P. Feldmann and J. Roychowdhury, "Computation of circuit waveform envelopes using an efficient, matrix-decomposed harmonic balance algorithm," in *IEEE Int. Conf. Computer-Aided Design Dig. Tech. Papers*, Nov. 1996.

[13] D. Feng, J. Phillips, K. Nabors, K. Kundert, and J. White, "Efficient computation of quasi-periodic circuit operating conditions via a mixed frequency/time approach," in *Proc. 36th Design Automation Conf.*, June 1999.

[14] F. Gardner, *Phaselock Techniques.* New York: Wiley, 1979.

[15] W. Gardner, *Introduction to Random Processes: With Applications to Signals and Systems.* New York: McGraw-Hill, 1989.

[16] P. Gray and R. Meyer, "Future directions in silicon IC's for RF personal communications," in *Proc. 1995 IEEE Custom Integrated Circuits Conf.*, May 1995.

[17] D. Held and A. Kerr, "Conversion loss and noise of microwave and millimeter-wave mixers—Part 1: Theory," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-26, pp. 49–55, Feb. 1978.

[18] D. Hente and R. Jansen, "Frequency-domain continuation method for the analysis and stability investigation of nonlinear microwave circuits," *Proc. Inst. Elect. Eng.* vol. 133, no. 5, pt. H, pp. 351–362, Oct. 1986.

[19] X. Jin, J.-J. Ou, C.-H. Chen, W. Liu, M. J. Deen, P. R. Gray, and C. Hu, "An effective gate resistance model for CMOS RF and noise modeling," in *IEDM'98 Tech. Dig.*, Dec. 1998, pp. 961–964.

[20] S. Kapur, D. Long, and J. Roychowdhury, "Efficient time-domain simulation of frequency-dependent elements," in *IEEE/ACM Int. Conf. Computer-Aided Design Dig. Tech. Papers*, Nov. 1996.

[21] F. Käertner, "Determination of the correlation spectrum of oscillators with low noise," *IEEE Trans. Microwave Theory Tech.*, vol. 37, pp. 90–101, Jan. 1989.

[22] ———, "Analysis of white and $f^{-\alpha}$ noise in oscillators," *Int. J. Circuit Theory Applicat.*, vol. 18, pp. 485–519, 1990.

[23] ———, "Noise in oscillating systems," in *Proc. Integrated Nonlinear Microwave and Millimeter Wave Circuits Conf.*, 1992.

[24] A. Kerr, "Noise and loss in balanced and subharmonically pumped mixers—Part 1: Theory," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, pp. 938–950, Dec. 1979.

[25] K. Kundert, J. White, and A. Sangiovanni-Vincentelli, "An envelope-following method for the efficient transient simulation of switching power and filter circuits," in *IEEE Int. Conf. Computer-Aided Design Dig. Tech. Papers*, Nov. 1988.

[26] ———, "A mixed frequency-time approach for distortion analysis of switching filter circuits," *IEEE J. Solid-State Circuits*, vol. 24, pp. 443–451, Apr. 1989.

[27] ———, *Steady-State Methods for Simulating Analog and Microwave Circuits.* Norwell, MA: Kluwer Academic, 1990.

[28] K. Kundert, "Simulation methods for RF integrated circuits," in *Proc. IEEE ICCAD-97 Int. Conf. Computer-Aided Design*, Nov. 1997.

[29] ———, "Modeling and simulation of jitter in phase-locked loops," in *Analog Circuit Design: RF Analog-to-Digital Converters; Sensor and Actuator Interfaces; Low-Noise Oscillators, PLL's and Synthesizers*, R. J. van de Plassche, J. H. Huijsing, and W. M. C. Sansen, Eds. Norwell, MA: Kluwer Academic, Nov. 1997.

[30] D. Long, R. Melville, K. Ashby, and B. Horton, "Full-chip harmonic balance," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1997.

[31] S. Maas, *Nonlinear Microwave Circuits.* Norwood, MA: Artech House, 1988.

[32] S. Martin, G. P. Li, H. Guan, S. D'Souza, M. Matloubian, G. Claudius, and C. Compton, "BSIM3 based RTS and $1/f$ noise models suitable for circuit simulators," in IEDM'98 Tech. Dig., Dec. 1998, pp. 851–884.

[33] R. Melville, P. Feldmann, and J. Roychowdhury, "Efficient multi-tone distortion analysis of analog integrated circuits," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1995.

[34] R. Meyer, W. Mack, and J. Hageraats, "A 2.5GHz BiCMOS transceiver for wireless LAN," in *1997 IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers and Slide Supplement*, Feb. 1997.

[35] K. Nabors, T. Fang, H. Chang, K. Kundert, and J. White, "Lumped interconnect models via Gaussian quadrature," in *Proc. 34th Design Automation Conf.*, June 1997.

[36] M. Nakhla and J. Vlach, "A piecewise harmonic balance technique for determination of periodic response of nonlinear systems," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 85–91, Feb. 1976.

[37] O. Nastov and J. White, "Grid selection strategies for time-mapped harmonic balance simulation of circuits with rapid transitions," in *Proc. 1999 IEEE Custom Integrated Circuits Conf.*, May 1999.

[38] E. Ngoya and R. Larcheveque, "Envelop (sic) transient analysis: A new method for the transient and steady state analysis of microwave communication circuits and systems," in *IEEE Microwave Theory and Techniques Symp. Dig. (MTTS)*, June 1996, pp. 1365–1368.

[39] M. Okumura, T. Sugawara, and H. Tanimoto, "An efficient small signal frequency analysis method for nonlinear circuits with two frequency excitations," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 225–235, Mar. 1990.

[40] M. Okumura, H. Tanimoto, T. Itakura, and T. Sugawara, "Numerical noise analysis for nonlinear circuits with a periodic large signal excitation including cyclostationary noise sources," *IEEE Trans. Circuits Syst. I*, vol. 40, pp. 581–590, Sept. 1993.

[41] L. Petzold, "An efficient numerical method for highly oscillatory ordinary differential equations," *SIAM J. Numer. Anal.*, vol. 18, no. 3, pp. 455–479, June 1981.

[42] J. Phillips, E. Chiprout, and D. Ling, "Efficient full-wave electromagnetic analysis via model-order reduction of fast integral transforms," in *Proc. 33rd Design Automation Conf.*, June 1996.

[43] B. Razavi, "A study of phase noise in CMOS oscillators," *IEEE J. Solid-State Circuits*, vol. 31, pp. 331–343, Mar. 1996.

[44] ———, *RF Microelectronics.* Englewood Cliffs, NJ: Prentice-Hall, 1998.

[45] ———, "RF IC design challenges," in *Proc. 35th Design Automation Conf.*, June 1998.

[46] G. Rhyne, M. Steer, and B. Bates, "Frequency-domain nonlinear circuit analysis using generalized power series," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 717–720, Feb. 1988.

[47] V. Rizzoli and A. Neri, "State of the art and present trends in nonlinear microwave CAD techniques," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 343–365, Feb. 1988.

[48] V. Rizzoli, A. Neri, and F. Mastri, "A modulation-oriented piecewise harmonic-balance technique suitable for transient analysis and digitally modulated signals," in *Proc. 26th Microwave Conf.*, Prague, Czech Republic, Sept. 1996, pp. 546–550.

[49] V. Rizzoli, C. Cechetti, A. Lipparini, and F. Mastri, "General-purpose harmonic balance analysis of nonlinear microwave circuits under multitone excitation," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 1650–1660, Dec. 1988.

[50] W. Robins, *Phase Noise in Signal Sources (Theory and Application)*. London, U.K.: IEE Telecommunications Series, 1996.

[51] J. Roychowdhury, "Efficient methods for simulating highly nonlinear multi-rate circuits," in *Proc. 34th Design Automation Conf.*, June 1997.

[52] J. Roychowdhury, D. Long, and P. Feldmann, "Cyclostationary noise analysis of large RF circuits with multi-tone excitations," *J. Solid-State Circuits*, vol. 33, pp. 324–336, Mar. 1998.

[53] C. Samori, A. Lacaita, F. Villa, and F. Zappa, "Spectrum folding and phase noise in LC tuned oscillators," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 781–790, July 1998.

[54] J. Schutt-Aine and R. Mittra, "Scattering parameter transient analysis of transmission lines loaded with nonlinear terminations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 529–536, Mar. 1988.

[55] D. Sharrit, "New method of analysis of communication systems," in *Proc. MTTS'96 WMFA: Nonlinear CAD Workshop*, June 1996.

[56] M. Silveira, I. El-Fadel, J. White, M. Chilukuri, and K. Kundert, "Efficient frequency-domain modeling and circuit simulation of transmission lines," *IEEE Trans. Comp., Packag., Manufact. Technol. B*, vol. 17, pp. 505–513, Nov. 1994.

[57] J. Stewart, "The power spectrum of a carrier frequency modulated by Gaussian noise," *Proc. IRE*, vol. 42, pp. 1539–1542, Oct. 1954.

[58] R. Telichevesky, K. Kundert, and J. White, "Efficient steady-state analysis based on matrix-free Krylov-subspace methods," in *Proc. 32nd Design Automation Conf.*, June 1995.

[59] ———, "Receiver characterization using periodic small-signal analysis," in *Proc. IEEE Custom Integrated Circuits Conf.*, May 1996.

[60] R. Telichevesky, K. Kundert, I. El-Fadel, and J. White, "Fast simulation algorithms for RF circuits," in *Proc. 1996 IEEE Custom Integrated Circuits Conf.*, May 1996.

[61] R. Telichevesky, K. Kundert, and J. White, "Efficient AC and noise analysis of two-tone RF circuits," in *Proc. 33rd Design Automation Conf.*, June 1996.

[62] M. Terrovitis and R. Meyer, *Cyclostationary noise in communication systems*, Electronics Research Laboratory Publications, Univ. of California, Berkeley, ERL Memo M99/36, 1999.

[63] G. Vendelin, A. Pavio, and U. Rohde, *Microwave Circuit Design Using Linear and Nonlinear Techniques.* New York: Wiley-Interscience, 1990.
[64] *Verilog—A Language Reference Manual: Analog Extensions to Verilog-HDL*, version 1.0, Open Verilog International, 1996.
[65] P. Vizmuller, *RF Design Guide: Systems, Circuits and Equations.* Norwood, MA: Artech House, 1995.
[66] A. Usihda, L. Chua, and T. Sugawara, "A substitution algorithm for solving nonlinear circuits with multi-frequency components," *Int. J. Circuit Theory Applicat.*, vol. 15, pp. 327–355, 1987.
[67] M. Zannoth, B. Kolb, J. Fenk, and R. Weigel, "A fully integrated VCO at 2GHz," in *1998 IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 1998.
[68] R. Ziemer and R. Tranter, *Principles of Communications: Systems, Modulation, and Noise.* New York: Houghton-Miffin, 1976.

Kenneth S. Kundert received the B.S., M.Eng., and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley in 1979, 1983, and 1988, respectively.

He specialized in circuit simulation and analog circuit design. He is a Fellow with Cadence Design Systems, San Jose, CA, and is the principal architect of the Spectre circuit simulation family. As such, he has led the development of Spectre, SpectreHDL, and SpectreRF. He also played a key role in the development of Hewlett-Packard's MNS harmonic balance simulator, as well as the Verilog-AMS and VHDL-AMS analog hardware description languages. He is the author of two books on circuit simulation, *Steady-State Methods for Simulating Analog and Microwave Circuits* (Norwell, MA: Kluwer Academic, 1990) and *The Designer's Guide to SPICE and Spectre* (Norwell, MA: Kluwer Academic, 1995).