

Simulation Methods for RF Integrated Circuits

Ken Kundert

Cadence Design Systems, San Jose, California, USA

Abstract — The principles employed in the development of modern RF simulators are introduced and the various techniques currently in use, or expected to be in use in the next few years, are surveyed. Frequency- and time-domain techniques are presented and contrasted, as are steady-state and envelope techniques and large- and small-signal techniques.

I. RF CIRCUITS

The increasing demand for low-cost mobile communication systems has greatly expanded the need for simulation algorithms that are both efficient and accurate when applied to RF communication 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.

Despite dramatic progress, the average design cycle of an RFIC is still twice the length of that for other ICs found in a communication system, such as a cellular phone. This represents a significant practical barrier to integration of the RF and baseband sections of a transceiver onto a single chip. Clearly, more progress is necessary.

This paper is an overview of RF simulation methods that seeks to provide an understanding of how the various methods address the RF simulation problem, and how they relate to each other.

It begins by describing the unique characteristics of RF circuits. The basic solution methods of transient analysis, harmonic balance, and shooting methods are presented and contrasted. Small-signal analysis versions of both harmonic balance and shooting methods are covered. Composite methods are next. These methods apply the base

Published in the *Proceedings of ICCAD'97*, November 9-13, 1997 in San Jose, California. Manuscript received August 11, 1997. Updated since publication; updated last on November 20, 1998.

This work was supported by the Defense Advanced Research Projects Agency under the MAFET program.

K. Kundert can be reached via e-mail at kundert@cadence.com.

methods in layers to provide dramatic new capabilities. The composite methods are divided into two groups. The first is based on a multidimensional representation of time. The second is based on sampling the RF signal and finding its envelope. The paper concludes with comparisons of the methods.

A. Narrowband Signals

RF circuits process narrowband 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 either 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 simple periodic or quasiperiodic modulations constructed from a small number of sinusoids for test signals.

The ratio between 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 inefficient 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 time step while a low-frequency modulation forces a long simulation interval.

Passing a narrowband signal through a nonlinear circuit results in a broadband signal whose spectrum is relatively sparse, as shown in Figure 1. In general, this spectrum

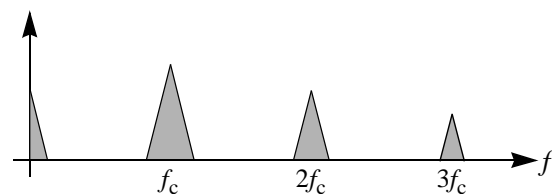


Fig. 1. Spectrum of a narrowband signal centered at a carrier frequency f_c after passing through a nonlinear circuit.

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 quasiperiodic, 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 are used when the spectrum is discrete, and transient methods 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).

Since *timing* or *synchronization signals*, such as the LO or the clock, are not part of the path of the information signal, they may be considered to be part of the circuit rather than an input to the circuit. 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 $\cos(\omega_{LO}t)$ consumes one input and results in a transfer function of

$$v_{out}(t) = \text{LPF}\{\cos(\omega_{LO}t)v_{in}(t)\}, \quad (1)$$

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

$$v_{in}(t) = m(t)\cos(\omega_c t), \quad (2)$$

then

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

and

$$v_{out}(t) = m(t)\cos((\omega_c - \omega_{LO})t). \quad (4)$$

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 III. 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, which cannot exhibit frequency translation. By linearizing a nonlinear circuit about a periodically varying operating point, we extend small-signal analysis to circuits that must have a periodic timing signal present to operate properly, such as mixers, switched filters, samplers, and oscillators (for oscillators the timing signal is the desired output of the oscillator, while 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 [39,45]). In particular, a noise analysis that accounts for noise folding and cyclostationary noise sources can be implemented [26,32], which fills a critically important need for RF circuits. When applied to oscillators, it also accounts for phase noise [3,12,13].

The linearity of the RF signal path can also be exploited to improve the convergence of the methods, as will be presented later.

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 make them difficult to include in the simulation.

Generally the passive components are linear and are modeled with phasors in the frequency-domain. This greatly simplifies the modeling of distributed components such as transmission lines, using either analytical expressions or tables of S-parameters. 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 a 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, 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 if N is small. This is usually true for

transmission lines and spirals, and less often true for packages and substrates.

Time-domain simulators are formulated to solve sets of first-order ordinary-differential equations (ODE). However, distributed components, such as transmission lines, are described with partial-differential equations (PDE) and so are problematic for time-domain simulators. Generally, the PDEs are converted to a set of ODE using some form of discretization [22]. Such approaches suffer from bandwidth limits. A more general 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 [11,34]. Evaluating lossy or dispersive transmission line models or tables of S-parameters with this approach is generally expensive and error-prone [36]. Packages, substrates and spirals can be modeled with large lumped networks, but such systems are too large to be efficiently incorporated into a time domain simulation, and so some form of reduction is necessary [6,28].

II. NUMERICAL SOLUTION OF DAEs

If we ignore distributed components for the moment, the basic idea behind circuit simulation is to solve a nonlinear set of differential-algebraic equations (DAEs). We start by proposing the form of the solution as a linear combination of basis functions. A primary requirement of the basis functions is that we must know their derivatives. This allows us to analytically evaluate the derivatives in the DAE, thereby converting the DAE to a system of nonlinear algebraic equations that can be solved with Newton's method. Examples of possible basis functions are polynomials and sinusoids. The efficiency and accuracy of the simulation is dependent on the choice of basis functions. A careful choice of basis functions can result in tremendous performance advantages.

The choice of basis functions defines the solution space for the method. Note that this space may not contain good approximations to all of the solutions to the original DAE, in which case the method based on that choice of basis functions will not be able to find those solutions. This is used to advantage to avoid particular solutions that are either uninteresting or undesirable. For example, if the basis functions are chosen to be sinusoids, then only steady-state behavior is representable and so the initial transient behavior is avoided. Furthermore, the circuit can be simulated even if it is unstable.

The other way to avoid undesired solutions is to apply constraints on the solution in the form of initial or boundary conditions. In general, DAEs have an infinite continuum of solutions, and so constraints must be applied before they can be solved for a particular solution unless

the solution space of the method contains only isolated solutions of the DAE.

A. Transient Analysis

Transient analysis breaks the time continuum into a series of adjacent short intervals and uses low-order polynomials as the basis functions over each interval (the time step) with the constraint that the solution must be continuous across interval boundaries (the time points). For example, consider the nonlinear DAE

$$f(v(t), t) = i(v(t)) + \frac{dq(v(t))}{dt} + u(t) = 0. \quad (5)$$

While this equation is capable of modeling any lumped time-invariant nonlinear system, 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, $v(t) \in \mathbb{R}^N$ is the node voltage, $i(v(t)) \in \mathbb{R}^N$ 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.

Consider trapezoidal rule. This is a second-order method that assumes that the solution and its first derivative are continuous at the interval boundaries. As such, they would be known at the start of the interval from the solution of the previous interval. With a small amount of algebra, one can show that if t_{s-1} is the initial point in the interval, t_s is the final point, and $h_s = t_s - t_{s-1}$, then for a second-order polynomial,

$$\frac{dq(v(t_s))}{dt} = \frac{2}{h_s}(q(v(t_s)) - q(v(t_{s-1}))) - \frac{dq(v(t_{s-1}))}{dt}. \quad (6)$$

Substituting (6) into (5) converts it into a sequence of nonlinear algebraic systems of equations that can be solved with Newton's method to build up a pointwise approximation to the solution from some initial state.

Under certain mild conditions, it is possible to show that Newton's method will converge at every time point of a transient analysis. Newton's method is an iterative procedure that converges to the solution of a nonlinear system of equations, *if* the initial starting point is close to the final solution. As long as the solution trajectory for the DAE is continuous, which it must be except at a set of distinct points (in particular, at jump or catastrophe points, which are rare), and if the starting point used for the Newton iteration is an extrapolation from the values at previous time points, then one can always take a time step small enough for Newton's method to converge.

B. Harmonic Balance

Harmonic balance [17,23,29] uses harmonically related sinusoids for basis functions. As noted earlier, such basis functions cannot represent the initial transient, so harmonic balance directly computes the steady-state behavior of the circuit.

For simplicity, we will use phasors (complex exponentials) as basis functions rather than sines and cosines. We will also treat the time-domain signals as being complex valued and simply ignore the imaginary parts. While this is not as efficient as if we assumed real-valued signals, it results in a simpler presentation. For more efficient implementations, see [17,20,30].

With phasors as basis functions, harmonic balance has a natural ability to incorporate the frequency-domain descriptions of distributed components. Thus, we formulate a new test problem that is similar to (5), except that the signals are complex valued and another term is added to model distributed components. In addition, we assume that the circuit is driven (i.e. that $u(t)$ is not constant).

$$i(v(t)) + \frac{dq(v(t))}{dt} + \int_{-\infty}^t y(t-\tau)v(\tau)d\tau + u(t) = 0 \quad (7)$$

or

$$f(v, t) = 0, \quad (8)$$

where $v(t)$, $u(t)$, $y(t)$, $i(v(t))$, $q(v(t))$, and $f(v, t) \in \mathbb{C}^N$, $v(t)$ and $u(t)$ are assumed to be T -periodic, $y(t)$ is the impulse response of the linear components, while i and q now represent only the nonlinear components.

Harmonic balance assumes that both v and f of (8) are formulated as Fourier series,

$$x(t) = \sum_{k=-\infty}^{\infty} X(k)e^{j\omega_k t}, \quad (9)$$

where $\omega_k = k\lambda$ and $\lambda = 2\pi/T$ is the fundamental frequency. Now (8) is rewritten as

$$\sum_{k=-\infty}^{\infty} F(V, k)e^{j\omega_k t} = 0, \quad (10)$$

where

$$F(V, k) = j\omega_k Q(V, k) + I(V, k) + Y(k)V(k) + U(k) \quad (11)$$

Since $e^{j\omega_{k_1}}$ and $e^{j\omega_{k_2}}$ are linearly independent if $k_1 \neq k_2$, $F(V, k) = 0$ for each k individually. To make the problem numerically tractable, it is necessary to consider only the first K harmonics. Then, in vector form,

$$F(V) = I(V) + \Omega Q(V) + YV + U = 0, \quad (12)$$

where $F(V)$, $I(V)$, $Q(V)$, V , $U \in \mathbb{C}^{(2K+1)N}$ are vectors of vectors. Each is composed of N vectors that represent the spectrum at each node. Ω , Y are block matrices. Ω is block diagonal and

$$\Omega = \text{diag}\{-j\omega_K, \dots, 0, \dots, j\omega_K\} \quad (13)$$

is the frequency-domain differentiation operator and $j = \sqrt{-1}$. Y_{mn} is the Laplace transform of y_{mn} evaluated at $j\omega_k$ for each k . It is diagonal because the components modeled by y are time-invariant.

While it is possible to model some types of nonlinearities directly in the frequency domain [37], this is not practical or desirable in all cases. Instead, a procedure is employed where V is first converted into the time domain at evenly spaced sample points using the Inverse Discrete Fourier Transform (IDFT),

$$x(s) = \frac{1}{K} \sum_{k=-K}^K X(k)e^{j2\pi ks/S} \quad (14)$$

where $s = 0, 1, \dots, S-1$ and $S \geq S_{\min} = 2K + 1$. While S may be set equal to S_{\min} , typically it is chosen in the range $2S_{\min} < S < 10S_{\min}$ to reduce aliasing. For each time point, both $i(v(t_s))$ and $q(v(t_s))$ are evaluated where $t_s = sT/S$, and the result converted back into the frequency domain using the forward Discrete Fourier Transform (DFT),

$$X(k) = \sum_{s=0}^{S-1} x(s)e^{-j2\pi ks/S} \quad (15)$$

Both the DFT and the IDFT can be written as matrix operations [17]:

$$X = Fx, \quad (16)$$

$$x = F^{-1}X, \quad (17)$$

where F represents the DFT and F^{-1} represents the IDFT. (16) is a restatement of (15) and (17) is a restatement of (14). If $S > S_{\min}$, these matrices are not square. Now

$$v = [v_n] = [F^{-1}V_n], \quad (18)$$

$$I_m(V) = Fi_m(v), \quad (19)$$

$$Q_m(V) = Fq_m(v). \quad (20)$$

Applying Newton-Raphson to solve (12) results in the iteration

$$J(V^{(r)})(V^{(r+1)} - V^{(r)}) = -F(V^{(r)}), \quad (21)$$

where r is the iteration number and

$$J(V) = \frac{\partial F(V)}{\partial V} = \frac{\partial I(V)}{\partial V} + \Omega \frac{\partial Q(V)}{\partial V} + Y \quad (22)$$

is the *harmonic Jacobian*.

$$J(V) = [J_{mn}(V)] = \left[\frac{\partial F_m(V)}{\partial V_n} \right], \quad (23)$$

where

$$\frac{\partial F_m(V)}{\partial V_n} = \frac{\partial I_m(V)}{\partial V_n} + \Omega_{mm} \frac{\partial Q_m(V)}{\partial V_n} + Y_{mn} \quad (24)$$

is a *conversion matrix*. The derivation of $\partial I_m(V)/\partial V_n$ follows from the chain rule.

$$I_m(V) = F i_m(v) \quad (25)$$

$$\frac{\partial I_m(V)}{\partial V_n} = F \frac{\partial i_m(v)}{\partial v_n} \frac{\partial v_n}{\partial V_n} \quad (26)$$

Since $i(v)$ is algebraic, $\partial i_m/\partial v_n$ is diagonal.

$$\frac{\partial i_m(v)}{\partial v_n} = \text{diag}\{g_{mn}(v(t_s))\}, \quad (27)$$

where $g_{mn}(v(t)) = \partial i_m(v(t))/\partial v_n(t)$ is the conductance waveform for the nonlinear resistors.

From $v_n = F^{-1}V_n$,

$$\frac{\partial I_m(V)}{\partial V_n} = F \frac{\partial i_m(v)}{\partial v_n} F^{-1}. \quad (28)$$

$\partial Q_m/\partial V_n$ is derived and constructed from c_{mn} in a similar manner, where $c_{mn}(v(t)) = \partial q_m(v(t))/\partial v_n(t)$ is the capacitance waveform for the nonlinear capacitors.

Efficiency: $J(V)$ is big and relatively dense, and so expensive to store and factor. To reduce the cost of solving (21), matrix-implicit methods are used. In these methods, only the component pieces of the Jacobian (g , c , and Y) are stored and iterative linear equation solvers such as the Krylov-subspace methods [33] are employed. These methods solve this linear system of equations by evaluating a sequence of matrix-vector products (MVP) that involve the Jacobian. The MVPs can be evaluated with nearly linear time and storage by exploiting the structure of (28) and by using fast algorithms such as the FFT to implement F and F^{-1} [21]. Unfortunately, these iterative linear solvers are not guaranteed to converge and require the use of a preconditioner to improve their convergence. The original linear system of equations

$$Ax = b \quad (29)$$

is preconditioned by multiplying both sides by \hat{A}^{-1} . One then applies the iterative solver to the preconditioned system

$$\hat{A}^{-1}Ax = \hat{A}^{-1}b. \quad (30)$$

Generally, the preconditioner \hat{A} is chosen to be a close approximation to A that is also easy to invert. For mildly nonlinear problems, constructing $\hat{A} = J(V_{DC})$ by linear-

izing about the DC operating point and performing a simple AC analysis at each ω_k is an effective and efficient choice [9,21,40]. In this case, \hat{A} is the Jacobian generated by the Gauss-Jacobi-Newton harmonic relaxation algorithm [17]. This preconditioner is not sufficient for strongly nonlinear problems. To handle these problems, it is necessary to adaptively prune the full harmonic Jacobian as in the harmonic-relaxation Newton algorithm [17,20].

When using the matrix-implicit methods, harmonic balance requires roughly $O(NK)$ operations, where N is the number of circuit equations and K is the number of frequencies required to accurately represent both v and f . This does not include the operations required to precondition the system of equations, which on strongly nonlinear problem may be far from negligible.

Extensions: 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 [17].

Applications: Harmonic balance is generally used to predict the distortion of RF circuits. It is also used to compute the operating point about which small-signal analyses are performed (presented later). When applied to oscillators, it 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).

Its space of application is similar to that of shooting methods (presented next). It is preferred over shooting methods when the circuit includes distributed components.

C. Shooting Methods

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 (5) is driven with a non-constant T -periodic stimulus. The T -periodic steady state solution is the one that also satisfies the two-point boundary constraint,

$$v(T) - v(0) = 0. \quad (31)$$

If the initial state $v(t_0)$ is known, then transient analysis can solve (5) and compute the state at some later time t_1 . In general, one writes

$$v(t_1) = v(t_0) + \phi(v(t_0), t_0, t_1) \quad (32)$$

where ϕ is the state transition function for the differential equation. Shooting methods reformulate (5) and (31) as

$$\phi_T(v(0), 0) = 0, \quad (33)$$

where $\phi_T(v(t_0), t_0) = \phi(v(t_0), t_0, T + t_0)$, which is a nonlinear algebraic problem and so standard Newton methods can be used to solve for $v(0)$. We refer to the combination of the Newton and shooting methods as the shooting-Newton algorithm.

When applying Newton's method directly to (33), 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.

Applying Newton's method to (33) results in the iteration $v_0^{(r)} = v_0^{(r-1)} - [J_{\phi}(v_0^{(r-1)}) - 1]^{-1} [\phi_T(v_0^{(r-1)}, 0) - v_0^{(r-1)}]$ (34)

where r is the iteration number, $v_0 = v(0)$, 1 is the identity matrix, and

$$J_{\phi}(v_0) = \frac{d\phi_T(v_0, 0)}{dv_0} = \frac{dv(T)}{dv_0}. \quad (35)$$

There are two important pieces to the computation of the Newton iteration given in (34): evaluating the state-transition function $\phi_T(v_0, 0)$, and forming and factoring the matrix $J_{\phi}(v_0)$, which is a dense matrix in general.

The state-transition function is computed by integrating (5) numerically over the shooting interval. The derivative of the state-transition function, referred to as the sensitivity matrix, is computed simultaneously because there are several quantities that are common to both computations [1,17,38,40].

Efficiency: Forming J_{ϕ} requires $O(N^2S)$ operations where S is the number of time points used to evaluate ϕ_T . Factoring $J_{\phi} - 1$ requires $O(N^3)$ operations. As a result, forming and factoring $J_{\phi} - 1$ becomes intractable when N exceeds several hundred. As with harmonic balance, matrix-implicit Krylov-subspace methods are used to avoid forming and factoring $J_{\phi} - 1$ [38]. Again, the component pieces of $J_{\phi} - 1$ are saved and the matrix-vector products are performed on the fly. The component pieces are

$$J_f(v(t_s)) = G(v(t_s)) + \frac{C(v(t_s))}{h_s} \quad (36)$$

and $C(v(t_s))$ at each time point s , where

$$G(v(t)) = \frac{di(v(t))}{dv(t)}, \quad (37)$$

$$C(v(t)) = \frac{dq(v(t))}{dv(t)}, \quad (38)$$

and $h_s = t_s - t_{s-1}$. A natural preconditioner is applied simply by saving and applying J_f in LU factored form [40]. This preconditioner has proven itself to be extremely robust.

Convergence: Newton's method is applied both in the outer loop to solve (33) and in the inner loop to solve (5) at each time point, making this a multi-level Newton method. As described before, the ability to adjust the time step during transient analysis results in Newton's method being extremely reliable in the inner loop. The outer loop is also quite reliable because ϕ_T is generally near linear as a direct result of RF circuits having near linear signal paths. Thus, shooting-Newton represents a well-designed multi-level Newton method where the inner loop is robust and shields the outer loop from the nonlinearity inherent in the problem.

Extensions: 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 [17].

Applications: Shooting methods are applied in the same situations as harmonic balance. It is generally preferred if the circuit is driven with strongly discontinuous signals (pulses as opposed to sinusoids). As such, shooting methods are well suited for simulating switching mixers, switched filters, samplers, frequency dividers, and relaxation oscillators as long as the circuits do not contain distributed components.

III. SMALL-SIGNAL ANALYSIS

Consider a circuit whose input is the sum of two periodic signals, $u(t) = u_L(t) + u_s(t)$, where $u_L(t)$ is an arbitrary periodic waveform with period T_L and $u_s(t)$ is a sinusoidal waveform of radial frequency ω_s whose amplitude is small.

Let $v_L(t)$ be the steady-state solution waveform when $u_s(t)$ is zero. Then allow $u_s(t)$ to be small, but nonzero. We can consider the new solution $v(t)$ to be a perturbation $v_s(t)$ on $v_L(t)$, as in $v(t) = v_L(t) + v_s(t)$. The small-signal solution $v_s(t)$ is computed by linearizing the circuit about $v_L(t)$ and applying one of the methods for finding the steady-state solution already described. From the theory of periodically time-varying systems [5,43], it is known that for

$$u_s(t) = U_s e^{j\omega_s t} \quad (39)$$

the steady-state response is given by

$$v_s(t) = \sum_{k=-\infty}^{\infty} V_s(k) e^{j(\omega_s + k\lambda)t}. \quad (40)$$

where $\lambda = 2\pi/T_L$ is the large signal fundamental frequency. $V_s(k)$ represents the sideband for the k^{th} harmonic of V_L . In this situation, there is only one sideband per harmonic because U_s is a single frequency complex exponential and the circuit is linear. 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_L and those with $k > 0$ become upper sidebands.

$V_s(k)/U_s$ is the transfer function for the input at ω_s to the output at $\omega_s + k\lambda$. 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.

Extensions: In the next few sections, the basic techniques are introduced that are used to compute the small-signal steady state response of a circuit linearized about a time varying operating point. This is sufficient for performing a time-varying AC analysis and can be extended to other types of small-signal analyses, such as computing the S-parameters of the circuit. These small-signal analyses are also extendable to cyclostationary noise analysis [3,7], which is an extremely important capability for RF designers [39].

These methods can also be extended so as to allow a small-signal analysis about a quasiperiodic operating point or small-signal analysis of autonomous circuits.

Applications: Small-signal analyses are tremendously useful for computing transfer functions (such as conversion gain and supply rejection) and predicting noise performance. In this way, they are similar to the AC and noise analyses in SPICE, but they can be applied to all kinds of circuits that the traditional small-signal analyses cannot, such as mixers, switched-filters, and samplers. Because they compute transfer functions in the presence of large signals, they can be used to determine the degradation of gain and noise when there is a large interfering signal on an adjacent channel. They can also be used to estimate intermodulation distortion [39].

When performing noise analysis, they accurately capture noise folding present in all circuits that have time-varying operating points, but which is especially important in mixers, samplers, and oscillators. In addition, they can be used to predict the phase noise of oscillators.

A. Transient Small-Signal Analysis

Though beyond the scope of this paper, there has been recent progress on performing small-signal noise analysis about a transient operating point [3]. However, in most cases, it is the small-signal behavior when the circuit is in steady state that is of interest. Using these methods generally requires integrating through any initial transient.

B. Harmonic Balance Small-Signal Analysis

Consider (12) where $U = U_L + U_s$ and $V = V_L + V_s$. Initially set $U_s = 0$ and solve (12) for V_L such that

$$F(V_L) = I(V_L) + \Omega Q(V_L) + YV_L + U_L = 0. \quad (41)$$

Since U_s is expected to be small, we can expand about V_L using a Taylor series truncated to first order.

$$\frac{\partial I(V_L)}{\partial V} V_s + \Omega_s \frac{\partial Q(V)}{\partial V} V_s + YV_s + U_s \approx 0 \quad (42)$$

or

$$J(V_L) V_s \approx -U_s \quad (43)$$

where $J(V_L)$ is computed as in (22) and Ω_s is as in (13) except with $\omega_k = \omega_s + k\lambda$.

Solving (43) for V_s gives the sidebands generated by U_s .

These ideas are extended to cyclostationary noise analysis in [10,12,32].

C. Shooting Method Small-Signal Analysis

Consider the circuit described by (5). Assume $u_s(t) = 0$ and let $v_L(t)$ be the resulting solution that also satisfies the two-point constraint (31). Linearizing (5) about $v_L(t)$ yields a time-varying linear system

$$\frac{dC(v_L(t))}{dt} v_s(t) + G(v_L(t)) v_s(t) + u_s(t) = 0 \quad (44)$$

that can be solved for $v_s(t)$ if $u_s(t)$ is small.

From (40) it follows that

$$v_s(t + T_L) = v_s(t) e^{j\omega_s T_L}. \quad (45)$$

Equation (45) in the periodically time-varying linear steady-state problem is analogous to (31) in the standard steady-state problem. It is solved using a modified shooting method [25,41]. Note that (45) implies that the entire small-signal steady-state response of the periodic time-varying system is determined by the behavior of $v_s(t)$ on any interval of length T_L .

Since the solution is computed by performing a time-domain simulation, there is an upper bound on the analysis frequency that is imposed by the largest time step used. The period of analysis frequency must be much larger than the largest time step. Typically, the time steps are the same

ones used when computing the large signal operating point.

Standard shooting methods can be used if we form $x_s(t) = v_s(t)e^{-j\omega_s t}$. Then, (45) becomes

$$x_s(t + T_L) = x_s(t), \quad (46)$$

which is the same form as (31) if $t = 0$. Now conventional shooting methods are applied to find x_s . Once known, Fourier analysis [18] is applied to find X_k , from which V_k is trivially computed. This is a different, but equivalent, way of deriving the method given in [25,41].

This procedure is applied to cyclostationary noise analysis in [26,41].

IV. MULTI-VARIATE METHODS

An interesting approach to solving narrowband problems uses a *multi-variate* representation of the signals and reformulates (5) as a partial-differential equation. Consider a simple two-tone quasiperiodic signal

$$x(t) = \cos(\omega_1 t) \cos(\omega_2 t), \quad (47)$$

where ω_1 is much smaller than ω_2 . A large number of time points would be needed to accurately approximate such a signal. For example, if 15 points per period were needed to accurately represent the highest frequency, and if $\omega_2 = 1000 \omega_1$, then 15,000 points are needed to accurately represent $x(t)$. Now consider a bi-variate representation obtained by replacing t with t_1 in the slowly varying parts of the expression, and t with t_2 in the rapidly varying parts. Then

$$\hat{x}(t_1, t_2) = \cos(\omega_1 t_1) \cos(\omega_2 t_2). \quad (48)$$

$x(t)$ is easily recovered from $\hat{x}(t_1, t_2)$ simply by setting $t = t_1 = t_2$. In this case, a grid of $15 \times 15 = 225$ points are needed to accurately represent $\hat{x}(t_1, t_2)$.

Using these ideas, we can replace (5) with

$$\frac{\partial q(\hat{v}(t_1, t_2))}{\partial t_1} + \frac{\partial q(\hat{v}(t_1, t_2))}{\partial t_2} + i(\hat{v}(t_1, t_2)) + \hat{u}(t_1, t_2) \quad (49)$$

or

$$f(\hat{v}(t_1, t_2), t_1, t_2) = 0. \quad (50)$$

A wide variety of methods are formulated from this equation by proposing different forms of the solution and different boundary conditions along each time axis. One can apply transient, harmonic balance, and shooting methods in layers and customize methods to particular classes of problems [2,31].

A. Quasiperiodic Harmonic Balance

If a circuit is driven with two signals at unrelated fundamental frequencies, it generally responds in steady state by

generating quasiperiodic (also known as polyperiodic) signals. Quasiperiodic signals have the form of a Fourier series that is generalized in that the frequencies of the sinusoids are not just integer multiples of a single fundamental frequency. Instead, they are the linear combination of integer multiples of several fundamentals. For example, if there are two fundamentals, then

$$x(t) = \sum_{k_1 = -\infty}^{\infty} \sum_{k_2 = -\infty}^{\infty} X(k_1, k_2) e^{j(k_1 \lambda_1 + k_2 \lambda_2)t}, \quad (51)$$

where λ_1 and λ_2 are the fundamental frequencies.

The response contains components at multiples of each fundamental frequency as well as at the sums and differences of these frequencies. Rearranging (51) allows us to see this as being equivalent to constructing the waveform as a conventional Fourier series where the sinusoidal frequencies are at integer multiples of λ_1 , except that the Fourier coefficients themselves are time-varying. In particular, the coefficient $\tilde{X}(t, k_1)$ is periodic with period $T_2 = 2\pi/\lambda_2$ and can itself be represented as a Fourier series in t_2 .

$$x(t) = \sum_{k_1 = -\infty}^{\infty} \underbrace{\sum_{k_2 = -\infty}^{\infty} X(k_1, k_2) e^{jk_2 \lambda_2 t}}_{\tilde{X}(t, k_1)} e^{jk_1 \lambda_1 t} \quad (52)$$

Convert to a bi-variate representation by associating t_1 with λ_1 and t_2 with λ_2 .

$$\hat{x}(t_1, t_2) = \sum_{k_1 = -\infty}^{\infty} \sum_{k_2 = -\infty}^{\infty} X(k_1, k_2) e^{jk_2 \lambda_2 t_2} e^{jk_1 \lambda_1 t_1} \quad (53)$$

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

Assuming v and f of (50) take the form of (53) (a linear combination of periodically AM modulated sinusoids) results in

$$\sum_{k_1 = -\infty}^{\infty} \sum_{k_2 = -\infty}^{\infty} F(V, k_1, k_2) e^{jk_1 \lambda_1 t_1 + jk_2 \lambda_2 t_2} = 0, \quad (54)$$

where

$$F(V, k_1, k_2) = j(k_1 \lambda_1 + k_2 \lambda_2)Q(V, k_1, k_2) + I(V, k_1, k_2) + U(k_1, k_2) \quad (55)$$

As with periodic harmonic balance, the terms in (54) are linearly independent, so $F(V, k_1, k_2) = 0$ for each k_1, k_2 . In vector form,

$$F(V) = I(V) + \Omega Q(V) + U = 0, \quad (56)$$

where Ω is generalized such that $\omega_k = k_1 \lambda_1 + k_2 \lambda_2$. This becomes finite-dimensional by bounding $|k_j| \leq K_j$. This is similar to (12), except the term for the distributed linear

components (YV) is missing. This term can be included in (56) without difficulty because the spectrum of V is discrete. 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 quasiperiodic problems. It is used here because of its simple derivation and because it represents an early application of multivariate methods [30,46]. An alternate approach that is generally preferred in practice is the false frequency method, which is based on a one-dimensional Fourier transform [8,17].

Parametric Harmonic Balance: Harmonic balance as described above exploits the sparse spectrum of modulated carrier signals and the linear nature of the passive components. However, it only incidentally exploits the nature of RF circuits to respond in a near-linear manner to the input signal (more effective preconditioner, fewer Newton iterations). Parametric harmonic balance [42] is a variation that retains the advantages of standard harmonic balance, but also exploits the near-linear nature of quasiperiodic RF circuits in a more deliberate manner. It splits harmonic balance into two phases, one to compute the response to the large periodic drive signal alone (ex., the LO in a mixer), and a second phase computes the response from the input as a perturbation of the solution computed in the first phase. The advantage of this method is that it solves for many fewer frequencies in the first phase when convergence is a struggle. In addition, information generated in the first phase can be used to accelerate the second phase. In this regard, parametric harmonic balance is similar to several of the methods presented later. Unfortunately, further description of this method is beyond the scope of this article.

Extensions: One obvious extension is to allow more than two fundamentals, though the method becomes expensive fast as the number of fundamentals increases, so in practice only two or three fundamentals are used. Another important extension would be to support autonomous or semi-autonomous circuits [46]. An example of an autonomous quasiperiodic circuit is one that consists of two coupled oscillators. An example of an semi-autonomous circuit is a free-running oscillator connected to a mixer with a driven input.

Applications: Quasiperiodic harmonic balance is used predict intermodulation distortion of narrowband circuits and both harmonic and intermodulation distortion of mixers. Both it and the mixed frequency-time method are used in similar situations. Quasiperiodic harmonic balance is preferred when the circuit includes distributed components.

B. Circuit Envelope Method

In the last section, 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 quasiperiodic. With circuit envelope [24,35], the Fourier coefficients are again time varying, but are not necessarily periodic. Instead, the Fourier coefficients $\tilde{X}(t, k)$ are taken to be transient waveforms. Thus, signals take the form

$$x(t) = \sum_{k=-\infty}^{\infty} \tilde{X}(t_1, k) e^{j\omega_k t_2}, \quad (57)$$

where $\omega_k = k\lambda$ and λ is the fundamental frequency of the base Fourier series. $\tilde{X}(t_1, k)$ represents the complex modulation of the k^{th} harmonic.

Now, (50) can be rewritten in the form of (57),

$$\sum_{k=-\infty}^{\infty} \tilde{F}(\tilde{V}(t_1), t_1, k) e^{j\omega_k t_2} = 0, \quad (58)$$

where

$$\tilde{F}(\tilde{V}(t_1), t_1, k) = \frac{d\tilde{Q}(\tilde{V}(t_1), k)}{dt} + j\omega_k \tilde{Q}(\tilde{V}(t_1), k) + \tilde{I}(\tilde{V}(t_1), k) + \tilde{U}(t_1, k) \quad (59)$$

Because $e^{j\omega_{k_1} t_2}$ is linearly independent of $e^{j\omega_{k_2} t_2}$ if $k_1 \neq k_2$, and because $\tilde{F}(\tilde{V}(t_1), t_1, k)$ is independent of t_2 , each term in (58) is zero independently of the others. In other words, $\tilde{F}(\tilde{V}(t), t, k) = 0$ for each k , or in vector form

$$\tilde{F}(\tilde{V}(t), t) = \frac{d\tilde{Q}(\tilde{V}(t))}{dt} + \Omega \tilde{Q}(\tilde{V}(t)) + \tilde{I}(\tilde{V}(t)) + \tilde{U}(t) = 0 \quad (60)$$

As with transient analysis, discretization methods such as trapezoidal rule or the backward difference formulae replace dQ/dt with a finite-difference approximation, converting (60) to a system of nonlinear algebraic equations that is solved with Newton's method. For example, applying backward Euler converts (60) to

$$\frac{\tilde{Q}(\tilde{V}(t_s)) - \tilde{Q}(\tilde{V}(t_{s-1}))}{t_s - t_{s-1}} + \Omega \tilde{Q}(\tilde{V}(t_s)) + \tilde{I}(\tilde{V}(t_s)) + \tilde{U}(t_s) = 0 \quad (61)$$

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 circuit envelope method. The transient nature of the modulations $\tilde{V}(t, k)$ 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 [11,34], or the model for

the distributed component can be separated into delay and dispersion, with the dispersion being replaced by a lumped approximation [36].

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 circuit 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. In other words, $Y_{mn}(\omega, k)$ is simply replaced with $Y_{mn}(\omega_k, k)$ where $\omega = \omega_k + \delta\omega$.

Extensions: One can build circuit envelope on top of a quasiperiodic harmonic balance rather than on top of periodic harmonic balance as has been done here. One can also build it on top of autonomous harmonic balance.

One can also wrap shooting around circuit envelope to find a periodic or quasiperiodic modulation waveform, which results the multi-variate mixed frequency-time method [31].

Applications: Circuit envelope has two primary applications. The first is predicting the response of a circuit when it is driven with a complex 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 quasiperiodic harmonic balance is 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 circuit envelope, and then determine how the modulation spectrum spreads into adjacent channels.

Another important application of circuit envelope 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 TDMA transmitters. In TDMA (time-division multiple access), transmitters 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 100 ms and the carrier frequency is typically at 1 GHz or greater.

The capabilities of circuit envelope are similar to envelope following (presented next). Circuit envelope is preferred when the circuit contains distributed components.

V. SAMPLING METHODS

RF circuits are generally influenced by one periodic timing signal, often referred to as the LO or the clock, and one or more information signals. For oscillators, the timing signal is the oscillation signal itself and the information signal is generally noise. With sampling methods, we designate the one timing signal as the *clock*. If there is more than one timing signal, then usually the largest and fastest is chosen to be the clock. The discrete sample-envelope for a signal x is defined as signal x_{env} that results when x is sampled with a period equal to that of the clock, as shown in Figure 2. The continuous sample-envelope is the trajectory that is traced out if the phase of both the clock and the sampling is allowed to drift relative to the other signals present. The sampling is assumed to always occur at the same phase of the clock.

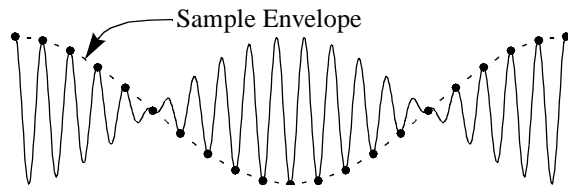


Fig. 2. Sample envelope is the waveform traced out when signal is sampled with a period equal to that of the clock.

The sample-envelope methods apply shooting methods over one or more clock cycles using boundary conditions that are formulated based on assumptions about the sample envelope. One can apply a wide variety of constraints on the sampled envelope, which results in a plethora of methods. For example, assuming the sampled envelope is constant results in simple shooting with a periodic boundary constraint. Assuming the circuit is linear and the sample envelope is sinusoidal results in periodic small-signal analysis. However, more interesting methods result when we make different assumptions about the sample envelope.

A. Envelope Following

Envelope Following approximates the sample envelope as a piecewise polynomial [15,27] 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 (32)

$$\Delta v(mT) = \Phi_T(v(mT), mT), \quad (62)$$

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 (62) 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, (62) is stiff and so requires implicit integration methods such as trapezoidal rule, which can be written as

$$\Delta v(mT) \approx \frac{2}{l}(v(mT) - v((m-l)T)) - \Delta v((m-l)T), \quad (63)$$

where l is the time step, which is measured in terms of cycles. This equation represents a 2-point boundary constraint on (62), and so together they can be solved with shooting-Newton methods to find $v(mT)$. 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 desired, other integration methods can be used, such as the backward-difference formulae.

Extensions: Envelope following can be extended to support autonomous circuits such as oscillators.

One can wrap shooting around envelope following to find a periodic or quasiperiodic envelope, which results the hierarchical shooting method [31].

Applications: The applications of envelope following are the same as circuit envelope. Envelope following is preferred when the carrier signal is strongly discontinuous (consists of pulses rather than sinusoids). As such, envelope following is suitable for simulating switched filters and switching power supplies in addition to the traditional RF circuits.

B. Mixed Frequency-Time Method

The Mixed Frequency-Time (or MFT) Method [16,17] makes the assumption that the sample envelope can be accurately approximated by a Fourier series with K terms (excluding DC), where K is presumed to be small. If true, then once the value of $S = 2K+1$ distinct points along the sample envelope are known, then all points can be found. Specifically, once the S points are known, then the $2K+1$

Fourier coefficients can be determined using the DFT, and then the resulting Fourier series can be evaluated to determine the value of any point. In particular, let

$$V_0 = [v(t_1), v(t_2), \dots, v(t_S)] \quad (64)$$

be the value of S points, then we can find the value of the S points that follow them by one cycle,

$$V_T = [v(t_1 + T), v(t_2 + T), \dots, v(t_S + T)] \quad (65)$$

Since the DFT is a linear operator F , there exists a linear operator $D_T = F^{-1}e^{j\omega_k T}F$ that maps V_0 to V_T .

$$V_T = D_T V_0 \quad (66)$$

This is SN equations with $2SN$ unknowns. They were formulated purely from the constraints on the sample envelope. It represents a boundary condition on solution to (5) in a way analogous to (31) and (45). Designate Φ_T as the collection of S state transition functions from t_s to t_s+T . Then

$$V_T = \Phi_T(V_0) \quad (67)$$

Applying (66) to (67) gives

$$D_T V_0 - \Phi_T(V_0) = 0, \quad (68)$$

which can be solved using Newton’s method for V_0 . As pointed out earlier, from V_0 one can calculate any point on the sample envelope. Any point $v(t)$ on the original waveform is found by integrating (5) from the appropriate point on the sample envelope.

Extensions: As with quasiperiodic harmonic balance, the mixed frequency-time method can be extended to an arbitrary number of fundamentals, though in practice is limited to two or three. It can also be extended to handle autonomous and semi-autonomous circuits.

Applications: The applications of the mixed frequency-time method are the same as quasiperiodic harmonic balance. The mixed frequency-time method is preferred when the carrier signal is strongly discontinuous (consists of pulses rather than sinusoids).

VI. COMPARISONS

The methods presented can be grouped into two broad families, those methods based on harmonic balance, and those based on shooting methods. Table I shows how the various methods are related. to baseband methods, and to each other.

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.

TABLE I
RELATING THE RF SIMULATION METHODS

<i>Baseband (SPICE)</i>	<i>HB Family</i>	<i>SM Family</i>
DC	HB	SM
AC, Noise	Sm Sig HB	Sm Sig SM
Transient	Ckt Env	Env Follow
HB, SM	QPHB	MFT

Progress on developing RF simulation methods has been rapid over the last decade and is continuing. As a result, any comparisons based on currently implemented versions of the methods will quickly become dated. So instead, I will try to extrapolate from current trends and compare what is likely to be the eventual attributes of the methods.

A. Strengths of Harmonic Balance

The main strength of harmonic balance is its natural support for frequency-domain models — both linear and nonlinear. 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. Nonlinear frequency domain models are based on Volterra series and are derived either from simulation or from measurement [44]. In addition, it is becoming more common for abstract behavioral models to be written directly in the frequency domain. There is, however, one caveat. It is relatively easy to generate nonphysical models in the frequency domain.

Harmonic balance is extremely efficient and accurate if both v and f in (8) are nearly sinusoidal. However this is not a feature that finds much application in practice. It is generally only significant when trying to determine the distortion of very low distortion amplifiers and filters. It does not help analyzing mixers, amplifiers, and samplers because they contain signals that are far from sinusoidal.

B. Strengths of Shooting Methods

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 time steps can be used to accurately resolve rapid transitions without taking small steps everywhere. This is very important for circuits such as mixers,

relaxation oscillators, switched-capacitor and switched-current filters, samplers, sample-and-holds, and chopper stabilized amplifiers. In addition, the timestep is automatically chosen to control error. Though not an inherent issue, it is a failing of existing harmonic balance simulators that they do not automatically control error.

The strong convergence properties of shooting methods result from its implementation as a multilevel Newton method, and not from the fact it is a time-domain method. Indeed, it is possible to formulate harmonic balance as a time-domain method [17,40], yet its convergence properties do not fundamentally change.

Harmonic balance can be made as robust as shooting methods by incorporating them in a carefully designed continuation or homotopy method [17]. However, because continuation methods end up calling harmonic balance tens, or perhaps hundreds, of times, they can be slow. The ability of shooting methods to converge on a large class of strongly nonlinear circuits without the need for continuation methods 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.

Finally, the preconditioner available in shooting methods seems to be more robust and less burdensome than the preconditioners available with harmonic balance. The preconditioner is used to implement the matrix-implicit linear solvers that allow both shooting methods and harmonic balance to handle large problems. This is currently an area of innovation, and so this situation may change.

A significant disadvantage of shooting methods is that they do not support distributed components. While it is conceivable that shooting methods can be extended to handle distributed components, doing so will likely compromise their strong convergence properties and their preconditioner.

The fundamental strengths of shooting methods and harmonic balance are compared in Table II.

VII. CONCLUSION

There has been a tremendous amount of innovation and progress in RF simulation methods in the past decade, with the result being the wide variety of methods available

TABLE II
STRENGTHS OF SHOOTING METHODS AND HARMONIC BALANCE

<i>Shooting Methods</i>	<i>Harmonic Balance</i>
<ul style="list-style-type: none"> • Convergence • Nonuniform timesteps • Robust preconditioner 	<ul style="list-style-type: none"> • Frequency domain models • No frequency limit on “small” signals

today. Each method carries with it limiting assumptions that it exploits to perform efficiently when those assumptions are satisfied. However, the assumptions also prevent each method from being used in a general setting. There is no universal method, and it is unlikely there will ever be one. This is mainly a problem in that it prevents the whole RF section of a transceiver from being simulated together. While it is starting to be possible to simulate the whole signal path of a receiver or a transmitter, incorporating the frequency synthesizer and any digital signal processing is still beyond reach.

While the existing methods and their obvious extensions do not solve all RF simulation problems for individual blocks, their coverage is pretty good. As such, work in developing new methods is expected to gradually decline. A new area of effort is expected to be in using the existing methods to generate high level models of individual blocks that will allow more general methods to simulate the whole system. Examples of this new trend include using Volterra series to model the nonlinear behavior of RF blocks [44] and using the small-signal time-varying noise analyses on the individual blocks in a phase-locked loop to generate behavioral models that include jitter [3,19], which allows accurate and efficient noise prediction for frequency synthesizers.

ACKNOWLEDGMENTS

I would like to thank Dave Sharrit of HP and Jacob White of MIT for their direct and significant contributions to this paper.

REFERENCES

- [1] T. Aprille and T. Trick. Steady-state analysis of nonlinear circuits with periodic inputs. *Proceedings of the IEEE*, vol. 60, no. 1, pp. 108-114, January 1972.
- [2] H. Brachtendorf, G. Welsch, R. Laur and A. Bunse-Gerstner. Numerical steady state analysis of electronic circuits driven by multi-tone signals. *Electrical Engineering*, Springer-Verlag, vol. 79, pp. 103-112, 1996.
- [3] A. Demir. *Analysis and Simulation of Noise in Nonlinear Electronic Circuits and Systems*. To be published by Kluwer Academic Publishers in 1997.
- [4] C. Dragone. Analysis of thermal and shot noise in pumped resistive diodes. *The Bell System Technical Journal*, vol. 47, no. 9, pp. 1883-1902, November 1968.
- [5] S. Egami. Nonlinear, linear analysis and computer-aided design of resistive mixers. *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-22, no. 3, pp. 270-275, March 1974.
- [6] P. Feldmann and R. Freund. Efficient linear circuit analysis by Padé Approximation via the Lanczos Process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 14, no. 5, pp. 639-649, May 1995.
- [7] W. Gardner. *Introduction to Random Processes: With Applications to Signals and Systems*. McGraw-Hill, 1989.
- [8] D. Hente and R. Jansen. Frequency-domain continuation method for the analysis and stability investigation of nonlinear microwave circuits. *IEE Proceedings*, part H, vol. 133, no. 5, pp. 351-362, October 1986.
- [9] P. Heikkilä. *Object-Oriented Approach to Numerical Circuit Analysis*. Ph. D. dissertation. Helsinki University of Technology, January 1992.
- [10] D. Held and A. Kerr. Conversion loss and noise of microwave and millimeter-wave mixers: part 1 — theory. *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-26, no. 2, pp. 49-55, February 1978.
- [11] S. Kapur, D. Long and J. Roychowdhury. Efficient time-domain simulation of frequency-dependent elements. *IEEE/ACM International Conference on Computer-Aided Design: Digest of Technical Papers*, November 1996.
- [12] F. Kaertner. Determination of the correlation spectrum of oscillators with low noise. *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, no. 1, Jan. 1989, pp. 90-101.
- [13] F. Kaertner. Analysis of white and $f^{-\alpha}$ noise in oscillators. *International Journal of Circuit Theory and Applications*, vol. 18, pp. 485-519, 1990.
- [14] A. Kerr. Noise and loss in balanced and subharmonically pumped mixers: part 1 — theory. *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-27, no. 12, pp. 938-950, December 1979.
- [15] K. Kundert, J. White and A. Sangiovanni-Vincentelli. An envelope-following method for the efficient transient simulation of switching power and filter circuits. *IEEE International Conference on Computer-Aided Design: Digest of Technical Papers*, November 1988.
- [16] K. Kundert, J. White and A. Sangiovanni-Vincentelli. A mixed frequency-time approach for distortion analysis of switching filter circuits. *IEEE Journal of Solid-State Circuits*, April 1989, vol. 24, no. 2, pp. 443-451.
- [17] K. Kundert, J. White and A. Sangiovanni-Vincentelli. *Steady-State Methods for Simulating Analog and Microwave Circuits*. Kluwer Academic Publishers, 1990.
- [18] K. Kundert. Accurate Fourier analysis for circuit simulators. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1994.
- [19] K. Kundert. Modeling and simulation of jitter in phase-locked loops. In *Analog Circuit Design: RF Analog-to-Dig-*

- ital Converters; Sensor and Actuator Interfaces; Low-Noise Oscillators, PLLs and Synthesizers*, Rudy J. van de Plassche, Johan H. Huijsing, Willy M.C. Sansen, Kluwer Academic Publishers, November 1997.
- [20] D. Long, R. Melville, K. Ashby and B. Horton. Full-chip harmonic balance. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1997.
- [21] R. Melville, P. Feldmann and J. Roychowdhury. Efficient multi-tone distortion analysis of analog integrated circuits. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1995.
- [22] K. Nabors, T. Fang, H. Chang, K. Kundert, and J. White. Lumped Interconnect Models Via Gaussian Quadrature. *Proceedings of the 34th Design Automation Conference*, June 1997.
- [23] M. Nakhla and J. Vlach. A piecewise harmonic balance technique for determination of periodic response of nonlinear systems. *IEEE Transactions of Circuits and Systems*. vol. CAS-23, no. 2, pp. 85-91, February 1976.
- [24] E. Ngoya and R. Larcheveque. Envelope transient analysis: a new method for the transient and steady state analysis of microwave communication circuits and systems. *IEEE Microwave Theory and Techniques Symposium Digest (MTTS)*, pp. 1365-1368, June 1996.
- [25] M. Okumura, T. Sugawara and H. Tanimoto. An efficient small signal frequency analysis method for nonlinear circuits with two frequency excitations. *IEEE Transactions of Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, no. 3, pp. 225-235, March 1990.
- [26] 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 Transactions on Circuits and Systems — I. Fundamental Theory and Applications*, vol. 40, no. 9, pp. 581-590, September 1993.
- [27] L. Petzold. An efficient numerical method for highly oscillatory ordinary differential equations. *SIAM Journal of Numerical Analysis*, vol. 18, no. 3, June 1981.
- [28] J. Phillips, E. Chiprout and D. Ling. Efficient full-wave electromagnetic analysis via model-order reduction of fast integral transforms. *Proceedings of the 33rd Design Automation Conference*, June 1996.
- [29] V. Rizzoli and A. Neri. State of the art and present trends in nonlinear microwave CAD techniques. *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 2, pp. 343-365, February 1988.
- [30] V. Rizzoli, C. Cechetti, A. Lipparini and F. Matri. General-purpose harmonic balance analysis of nonlinear microwave circuits under multitone excitation. *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 12, pp. 1650-1660, December 1988.
- [31] J. Roychowdhury. Efficient methods for simulating highly nonlinear multi-rate circuits. *Proceedings of the 34th Design Automation Conference*, June 1997.
- [32] J. Roychowdhury, D. Long and P. Feldmann. Cyclostationary noise analysis of large RF circuits with multi-tone excitations. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1997.
- [33] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing, 1996.
- [34] J. Schutt-Aine and R. Mittra. Scattering parameter transient analysis of transmission lines loaded with nonlinear terminations. *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-36, no. 3, pp. 529-536, March 1988.
- [35] D. Sharrit. New method of analysis of communication systems. *MTTS'96 WMFA: Nonlinear CAD Workshop, June 1996*.
- [36] M. Silveira, I. El-Fadel, J. White, M. Chilukuri, and K. Kundert. Efficient frequency-domain modeling and circuit simulation of transmission lines. *IEEE Transactions on Components, Packaging, and Manufacturing Technology — Part B: Advanced Packaging*, November 1994.
- [37] G. Rhyne, M. Steer and B. Bates. Frequency-domain nonlinear circuit analysis using generalized power series. *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 2, pp. 717-720, February 1988.
- [38] R. Telichevesky, K. Kundert and J. White. Efficient steady-state analysis based on matrix-free Krylov-subspace methods. *Proceedings of the 32nd Design Automation Conference*, June 1995.
- [39] R. Telichevesky, K. Kundert and J. White. Receiver characterization using periodic small-signal analysis. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1996.
- [40] R. Telichevesky, K. Kundert, I. El-Fadel and J. White. Fast simulation algorithms for RF circuits. *Proceedings of the 1996 IEEE Custom Integrated Circuits Conference*, May 1996.
- [41] R. Telichevesky, K. Kundert and J. White. Efficient AC and noise analysis of two-tone RF circuits. *Proceedings of the 33rd Design Automation Conference*, June 1996.
- [42] Y. Thodesen and K. Kundert. Parametric harmonic balance. *IEEE MTT-S International Microwave Symposium Digest*, June 1996.
- [43] H. Torrey and C. Whitmer. *Crystal Rectifiers*. McGraw-Hill, 1948.
- [44] F. Verbeyst and M. Vanden Bossche. VIOMAP, the S-parameter equivalent for weakly nonlinear RF and microwave devices. *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 12, pp. 2531-2535, December 1994.
- [45] N. Verghese and D. Allstot. Verification of RF and mixed-signal integrated circuits for substrate coupling effects. *Proceedings of the IEEE Custom Integrated Circuits Conference*, May 1997.
- [46] A. Usihda, L. Chua and T. Sugawara. A substitution algorithm for solving nonlinear circuits with multi-frequency components. *International Journal on Circuit Theory and Application*, vol. 15, pp. 327-355, 1987.