Abstract—The paper reviews the current status of noise simulation and modeling for semiconductor devices (in particular, transistors) operated in large-signal (forced) conditions. From a practical standpoint, large-signal noise modeling is relevant in the simulation and design of analog components such as mixers and frequency multipliers. The specific features of cyclostationary noise are discussed, and the various modeling techniques are presented, including simulation strategies for evaluating the large-signal steady state within the framework of a physics-based model. Particular attention is given to the issue of microscopic noise source modulation and frequency conversion, still an open problem in the case of 1/f noise not amenable to a superposition of more elementary (e.g. generation-recombination) sources. Starting from physics-based large-signal simulation techniques, the review also covers compact modeling strategies, where noise source modulation is even more involved and no general, device-independent strategy seems to provide correct results, but ad-hoc solutions have to be tailored on specific device classes.

I. INTRODUCTION

Noise modeling and simulation is a classical topic in device analysis [1], [2], both because of the fundamental information that fluctuations can provide, e.g. concerning the quality of materials and interfaces, and of the importance of noise assessment within the framework of the circuit and system applications of the device. From the standpoint of circuit applications, an understanding of the noise mechanisms taking place in an active device, typically a transistor, is a fundamental step in devising and designing low-noise subsystems (e.g. amplifiers, mixers and oscillators), while the capability to correctly predict the fluctuation performance as a function of the device electrical operation is a prerequisite for the identification of device noise models in terms of topologies and component parameters, that are in turn indispensable for the CAD design and optimization of low noise circuits.

Therefore, great importance has been given in the literature to the development of transistor noise models. These can be divided basically into two large families: physics-based (PB) models, and compact models. PB models derive the device noise characteristics starting, in the vast majority of cases, from the numerical solution of physically sound descriptions of carrier transport and of the fluctuation causes (the so-called microscopic noise sources) distributed within the device volume [3], [4]. Compact models, on the other hand, are characterized by a numerically very efficient description of device electrical characteristics (including noise) obtained at the cost of (often severe) simplifications of the physical device structure and operation, and of the introduction of empirical corrections to account for specific features, e.g. connected to device downscaling: a recent review can be found in [5]. In compact models, fluctuation features correspond to the statistical characterization of the electrical noise sources connected to the device terminals, expressed in term of either voltage or current fluctuations and of their correlation. Such noise sources are the macroscopic, terminal-level manifestation of the microscopic noise sources inside the device: PB noise models allow to estimate the link between the two quantities, thus providing the foundation for the development of compact noise modeling.

From the application standpoint, a fundamental impact on the very statistical nature of noise is provided by the device operating conditions [3]: if fluctuations are calculated assuming that the device is working in a DC bias point, the stochastic processes representing noise are stationary random processes and the condition is referred to as stationary noise. On the other hand, in many cases devices are driven by large amplitude time varying signals: when this happens, noise processes are no longer stationary. A particular, but practically important, case of large-signal drive is that of time-periodic (forced) operation (we call this LS regime), wherein fluctuations are represented by cyclostationary processes [4], [6], [7]. The LS regime has several practical applications, ranging from power amplifiers (where, however, noise is not typically an issue) to mixers and frequency multipliers [8]. In practical terms, large-signal operation involves the phenomenon of frequency conversion, i.e. signal and noise spectral components are distributed around the harmonics of the LS fundamental frequency by means of the nonlinear (or linear time-varying) behavior of the device [3], [4], [7]: this effect is of paramount importance in case of the so-called stationary low frequency noise (e.g. 1/f or flicker noise); since the spectral density of such a noise is large at low frequency (before conversion), it becomes the main noise component around the LS harmonics after frequency upconversion. Notice that LS operation is also relevant for oscillator noise analysis. This case, however, is still a matter of open research: an example of PB device noise simulation in this specific operating regime is [9]. Most
of the literature, however, is focused on the development of compact noise models to be exploited in the oscillator circuit optimization.

Cyclostationarity was first recognized as a fundamental marker of noise in (time-periodic) large-signal operation in the seminal paper [10], where amplitude modulation of a stationary process by means of the time-periodic instantaneous device working point is recognized as the fundamental element leading to the loss of stationarity. Amplitude modulation is also the basis for the transformation of the microscopic noise sources exploited in PB noise modeling according to the simulation techniques developed independently in [7], [11], in [12], [13] and in [14], [15], as far as PDE-based transport models are concerned. Large-signal PB noise simulations were performed based on the direct solution of the Boltzmann transport equation (BTE) in [16], where a Monte Carlo approach is exploited. A deterministic solution of the BTE based on spherical harmonics expansion, including noise simulation capabilities both in small- and large-signal, was also discussed in [17], [18].

In the rest of the contribution, we first discuss in Section II the statistical characterization of cyclostationary noise. A system-level interpretation of amplitude modulation is presented in Section III, while Section IV provides a discussion on physics-based LS noise modeling strategies. The very important topic of compact LS noise modeling is treated in Section V, while some conclusions are finally drawn in Section VI.

II. STATISTICAL CHARACTERIZATION OF CYCLOSTATIONARY NOISE

Let us consider a (real) cyclostationary stochastic process \( x(t) \), characterized by an average \( m_x(t) = \mathbb{E}\{x(t)\} \) and a (auto-) correlation function\(^1\) \( R_{x,x}(t_1,t_2) = \mathbb{E}\{x(t_1)x(t_2)\} \) periodic in time with period \( T \) [6], [19]

\[
m_x(t) = m_x(t+T) = R_{x,x}(t_1,t_2) = \tilde{R}_{x,x}(t_1+t,T, t_2+T) \quad (1)
\]

where \( \mathbb{E}\{\cdot\} \) denotes the statistical average operator. From (1), we can derive the modified correlation function \( R_{x,x}(t,\tau) = \tilde{R}_{x,x}(t+\tau/2, t-\tau/2) \) which is \( T \)-periodic as a function of \( t \). Therefore

\[
R_{x,x}(t,\tau) = \sum_n R^{(n)}_{x,x}(\tau) e^{im\omega_0 t} \quad (2)
\]

where \( i = \sqrt{-1} \) and \( \omega_0 = 2\pi/T \). \( R^{(n)}_{x,x}(\tau) \) is called harmonic [20] or cyclic correlation function [19] of \( x(t) \). The Fourier transform of \( R^{(n)}_{x,x}(\tau) \), denoted as \( S^{(n)}_{x,x}(\omega) \), is the corresponding harmonic [20] or cyclic correlation spectrum [19].

Calculating the double Fourier transform of (2) with respect to \( t \) and \( \tau \) we find

\[
G_{x,x}(\omega_1, \omega_2) = 2\pi \sum_n S^{(n)}_{x,x}(\omega_1 - \frac{n\omega_0}{2}) \delta(\omega_1 - \omega_2 + n\omega_0) \quad (3)
\]

thus proving that the correlation spectrum is zero unless \( \omega_1 = \omega_2 = n\omega_0 \), meaning that the condition for two frequency components to be correlated is that their distance is an integer multiple of the fundamental frequency \( \omega_0 \). In other words, we have correlation only between frequency components characterized as \( \omega_1 = \omega + k\omega_0 \) and \( \omega_2 = \omega + l\omega_0 \), where \( k \) and \( l \) are integers. The neighborhood of any fundamental frequency \( \omega_0 = i\omega_0 \), characterized by a local distance \( \omega \) (see Fig. 1), is called the upper (lower) sideband of \( \omega_0 \) if \( \omega > 0 \) \((\omega < 0)\). Usually, sidebands are defined as non-overlapping intervals, i.e. \( \omega \leq \omega_0/2 \). According to this description, the autocorrelation spectrum of \( x(t) \) is completely determined by a sideband correlation matrix (SCM) whose elements, dependent on \( \omega \) only, express the correlation spectra between the amplitude of the various sidebands:

\[
(S_{x,x}(\omega))_{k,l} = S^{(k-l)}_{x,x}(\omega + \frac{k+l}{2}\omega_0) . \quad (4)
\]

The SCM of a cyclostationary process can be given a spectral interpretation as follows. The Fourier transform of \( x(t) \) can be interpreted as an array of band-limited functions \( \tilde{x}_n(\omega) \), where \( -\omega_0/2 < \omega < \omega_0/2 \) is the sideband angular frequency and \( n \) denotes the \( n \)-th sideband. Then, the SCM element \( (k,l) \) simply is:

\[
(S_{x,x}(\omega))_{k,l} = \langle \tilde{x}_k(\omega)\tilde{x}^*_l(\omega) \rangle . \quad (5)
\]

where * is the complex conjugate. Because of this interpretation, which directly extends the usual concept of stationary noise spectrum to the LS case, we shall adopt the SCM as a representation of noise in cyclostationary operation.

Notice that the previous discussion can be directly extended to the case of quasi-periodic device excitation, meaning that the driving signal is made of a combination of incommensurable input tones (i.e., tones whose frequency ratio is not rational); the corresponding noise processes become quasi-cyclostationary [21].

III. SYSTEM LEVEL STATIONARY NOISE MODULATION APPROACHES

The key element in the comprehension of LS noise analysis is the amplitude modulation effect corresponding to the time-varying nature of the device working point. The issue was thoroughly examined in [6] at the circuit level (electrical noise sources), and the same approach was extended to the analysis of the microscopic noise sources in [4], [7]. To introduce the problem, we consider a stationary process \( \gamma(t) \) representing a fluctuation characterized in small signal operation by a spectrum \( S_{\gamma,\gamma}(\omega) = f^2 |\overline{h}(\omega)|^2 \), where \( f \) is a (constant, in

![Fig. 1: Definition of upper (lower) sideband of a fundamental frequency \( \omega_0 = k\omega_0 \) \((\omega_1 = l\omega_0)\).](image-url)
small-signal operation) factor containing the spectrum dependence on the DC working point and $h(t)$ (whose Fourier transform is $\hat{h}(\omega)$) is the impulse response of a linear, time-invariant system collecting the frequency-dependent part of the spectrum. Although this is somehow arbitrary, in the vast majority of practical cases the stationary noise sources can actually be factorized as assumed. The direct consequence of the factorization is the interpretation of $\gamma(t)$ as the output of the block system represented in Fig. 2 (a), where $\eta(t)$ is a unit, white Gaussian random process providing the stochastic nature of $\gamma$.

In the LS regime, the working point becomes time-varying and therefore the $f$ factor becomes a $T$-periodic modulating function $f(t)$, thus providing the amplitude modulation by means of a time-periodic function which makes the fluctuations cyclostationary. The modulation, however, can be carried out at least in two ways [4], [6], [22], [23]:

- the input process is first transformed into a cyclostationary process $y_{\text{MF}}(t)$, and then filtered by $h(t)$ (see Fig. 2 (b)). We call this “MF modulation”, and the SCM of the resulting output $\gamma_{\text{MF}}(t)$ is given by

$$(S_{\gamma_{\text{MF}},\gamma_{\text{MF}}}(\omega))_{m,n} = \hat{h}(\omega_m^+)G_{m-n}\hat{h}^*(\omega_n^+),$$

where $\omega_m^+ = \omega_m + \omega$ and $G_k$ is the $k$-th Fourier component of the $T$-periodic function $g(t) = f^2(t)$;

- linear filtering is performed first (see Fig. 2 (c)), yielding the stationary process $y_{\text{FM}}(t)$ which is then modulated by $f(t)$. We call this “FM modulation”, whose SCM is characterized by elements

$$(S_{\gamma_{\text{FM}},\gamma_{\text{FM}}}(\omega))_{m,n} = \sum_k F_{m-k}F^*_{n-k} |\hat{h}(\omega_k^+)|^2,$$

where $F_k$ is the $k$-th harmonic amplitude of the periodic function $f(t)$.

Clearly, (6) and (7) yield markedly different results. In fact, since $\hat{h}(\omega)$ is generally a low-pass function, assuming that its cutoff frequency is much lower than the working point frequency $\omega_0$ (6) implies that the MF SCM elements are negligible unless $m = n = 0$, i.e. the baseband sideband. On the other hand, the FM scheme yields non zero elements even if $m, n \neq 0$ provided that the modulating function has large enough harmonic components (i.e., the device is driven in nonlinear operation). In other words, only the FM approach provides frequency conversion effects: this explains why FM is the modulation scheme most commonly used in circuit simulators [24]–[26], although for the case of flicker noise sources the MF scheme has also been used [20], [27]. More details on this will be provided in Section V.

On the other hand, for white noise sources (i.e., if $\hat{h}(\omega) = 1$), both modulation schemes yield the same result

$$(S_{\gamma_{\text{MF}},\gamma_{\text{FM}}}(\omega))_{m,n} = (S_{\gamma_{\text{FM}},\gamma_{\text{FM}}}(\omega))_{m,n} = G_{m-n} = \sum_k F_{m-k}F^*_{n-k},$$

thus suggesting that white noise sources are uniquely modulated in LS operations. Notice that (8) yields the result derived in [10] with reference to the cyclostationary noise in a diode represented by a memoryless model (i.e., neglecting capacitive effects).

IV. PHYSICS-BASED LS NOISE MODELING

We discuss in this section the available simulation tools developed for the implementation of cyclostationary noise analysis exploiting physics-based models.

We consider first the case of PB transport models based on partial-differential equations (PDE). Among these, the still most widely used tool is the drift-diffusion (DD) model, despite its relative simplicity in the charge transport description and, therefore, the limitations to which it is subject. Nevertheless, the comparatively low numerical intensity and the availability of reliable material parameter models still makes DD widely used [28], in particular for 3D simulations. The following discussion, therefore, is based assuming the DD model. Extension, at least formal, to more advanced PDE approaches is easy, although the difficulty is moved to the determination of the corresponding microscopic noise sources [29]–[31].

Within the limits of PDE models, the microscopic noise sources are added as stochastic forcing terms to the transport equations, and their effect in terms of terminal electrical...
fluctuations (the device electrical noise sources) is calculated assuming they represent a linear perturbation of the noiseless device working point [3], [32]. Due to the linearity assumption, the propagation from the microscopic noise sources to the output fluctuations is characterized by an impulse response (as a function of time and spatial position, besides the input equation and the output variable) which corresponds to the Green’s function of the model (see [3] for details). For stationary noise, the linearized system is time-invariant and therefore can be easily represented in the frequency domain by a (frequency-dependent) Green’s function. Since the microscopic noise sources are assumed known from basic physics (or more fundamental numerical simulations), the Green’s functions evaluation is the heaviest step from a computational standpoint. Efficient numerical techniques for the frequency domain evaluation of the Green’s functions were proposed in [34], with reference to monopolar modeling, and later extended to the bipolar (and, in general, to any PDE-based model) in [35]: such techniques are currently exploited by many simulation codes, both commercial and academic [36]–[39].

In the cyclostationary case, on the other hand, the noiseless linearized system becomes periodically time-varying [7], and therefore provides frequency conversion effects among the various sidebands. The theoretical foundation of noise analysis in LS operation is based on the following main features (see [3], [4], [7] for details):

- the most natural representation (in terms of generalization of the small-signal case) is in the frequency domain, i.e. exploiting the Fourier series to represent the time-periodic signals [7];
- each perturbation \( x(t) \) is represented by the collection of amplitudes of the corresponding sidebands \( \mathbf{X}^+ \), whose size \( 2N_S+1 \) depends in turn on the number of harmonics \( 2N_i+1 \) used in the Fourier representation of the noiseless working point. A simple calculation allows to prove that, for a truncated spectrum, \( N_{iL} = 2N_S \) [4], [33];
- the linear time-varying system is easily represented in the frequency domain by a rank-2 tensor operator, the conversion matrix, linking the sideband representation of the input to the sideband representation of the output. Therefore, the Green’s functions become Conversion Green’s Functions (CGFs) [3], [7], which can again be efficiently estimated using an extension of the numerical technique originally developed for the stationary case [7].

For the sake of definiteness, let us assume to consider as the output of the physical simulation the short-circuit current noise generators connected to the \( i \)-th device terminal, represented by the corresponding sideband amplitudes \( \Gamma_{n,i}^+(\omega) \). Time-varying linearity allows to express

\[
\Gamma_{n,i}^+(\omega) = \sum_{\alpha=\varphi,n,p} \int_{\Omega} G_{\alpha,i}(r,\omega) \cdot \Gamma_{\alpha}^+(r,\omega) \, dr,
\]

where \( \Gamma_{\alpha}^+(r,\omega) \) is the vector of sideband amplitudes of the microscopic noise source added to equation \( \alpha \), and \( G_{\alpha,i}(r,\omega) \) is the CGF linking a unit injection in equation \( \alpha \) and point \( r \) to the \( i \)-th terminal current variation. The SCM of the short circuit noise generators connected to terminals \( i \) and \( j \) can therefore be derived as [7] (see (5)):

\[
S_{i_k,i_j,j_j}(\omega) = \sum_{\alpha,\beta=\varphi,n,p} \int_{\Omega} G_{\alpha,i}(r,\omega) \cdot K_{\alpha,\alpha}(r,\omega) \cdot G_{\beta,j}(r,\omega) \, dr,
\]

where \( K \) is the SCM of the local noise source for spatially uncorrelated microscopic fluctuations.

Since the CGFs can be computed exploiting a comparatively efficient numerical approach, the bottleneck in LS noise simulations is in fact the determination of the device noiseless working point. This problem can be tackled either in the time- or frequency-domain. In the first case, standard transient simulation is in general quite inefficient because of the widely spaced time constants present in the device dynamics, which require a combination of small time steps, to accurately define the fast device dynamics, and a large simulated time, to allow the device to reach steady-state conditions. Special numerical approaches are available to directly estimate the steady-state solution in time domain [9], [40], [41], thus avoiding the computation of the transient solution: they are based on the shooting technique originally developed for circuit LS simulation [42]–[44]. The other approach, again based on the direct determination of the steady-state periodic solution, is the Harmonic Balance (HB) technique, which transforms the dynamic nonlinear differential equations into nonlinear algebraic equations using as unknowns the harmonic amplitudes of the required variables [3], [7], [20], [24], [25], [33], [43]. HB was applied to multi-dimensional PB device modeling first in [45]–[47], and extended to cyclostationary noise simulation in [7], [11]–[15]. The shooting technique and HB were compared in [41]: basically, HB involves larger nonlinear systems to be solved, but on the other hand shooting needs to solve a long sequence of smaller problems. From the standpoint of simulation time, no technique appears clearly superior (at least in the strictly periodic case). As a general remark, in our experience shooting is less sensitive to initial data, and therefore somehow more robust, while HB yields a faster convergence once iterations have led near enough to the solution.

A. The microscopic noise source issue

From a physical standpoint, the main open issue in PB cyclostationary noise simulation still is the identification of reliable microscopic noise sources. Besides the requirement of physical consistency of such sources, already present in the stationary case for specific devices such as the MOSFET where conduction takes place very near to the Si/SiO\(_2\) interface, LS operation calls for the transformation of the sources into cyclostationary processes. As discussed in Section III, the modulation approach leads to a unique result for stationary white noise sources only [23]. This condition
is met (at least approximately) by diffusion and generation-recombination (GR) noise mechanisms [3], [48], [49] (in the latter case, however, trap-assisted GR phenomena should be added to the PB model including the rate equation for each trap level [49]). Unfortunately, the same does not always apply to the microscopic noise source for flicker noise, which still is a matter of open research: only $1/f$ fluctuations deriving from a superposition of non-interacting GR sources can be traced back to white stationary fluctuations [52].

V. COMPACT LS NOISE MODELING

The development of reliable device compact models is the fundamental step in the development of optimized circuit design [8], and thus the optimization of low-noise systems heavily relies on the availability of compact device noise models. There are two approaches to cyclostationary compact modeling: physical compact models are directly derived, basically by means of an analytical implementation of the Green’s function approach (10) obtained through a large enough number of approximations, or cyclostationary noise is estimated starting from stationary models exploiting their working point dependent terms: this ultimately requires to exploit one of the modulation technique previously described. The first approach is in principle to be preferred in terms of physical model consistency, though the implementation is usually confined to very simple devices and/or idealized structures: an example is the $pn$ diode LS noise model described in [50], [51], which we are currently extending to the bipolar transistor case. Other possible system oriented approaches are those exploited for the modeling of the reduction in low-frequency noise experimentally observed in MOSFETs under switched operation [53]–[55].

In practical applications, the modulation technique is the most commonly used. The discussion in Section III shows that the modulation of a stationary compact noise model is trivial if the stationary spectrum is white, which unfortunately quite rarely happens. In fact, in most of the cases non-white stationary spectra are present (e.g., the gate noise spectrum of an FET). Furthermore, in nonlinear applications low-frequency noise plays a fundamental role in limiting the low-noise circuit performance because of the frequency upconversion effect.

The more direct approach to the identification of the correct modulation procedure is of course the comparison with experimental results, which however are quite difficult to obtain, in particular for the cyclostationary case [56]. Furthermore, physical simulations on simple structures showed that an a priori choice between the FM or MF approach is impossible, and that in general none of the two yields fully consistent results [23], [57], [58]. The more general conclusion that can be drawn on the basis of PB simulations and/or experimental results is that a general strategy is not easily identifiable, rather a microscopic noise source- and device-dependent strategy should be pursued, ultimatley to be validated against experiments.

For instance, in case of bipolar devices PB simulations of GR trap noise in a 2D $pn$ junction showed that the terminal cyclostationary noise is the result of a mixed upconversion process: GR noise mainly generated in the depleted region (the device area where strongly nonlinear effects take place) is upconverted according to the FM scheme, while GR noise microscopically generated in a resistive part of the device (e.g., surface trap effects) is transferred to the terminal through the Green’s function of a basically linear device, and therefore behave according to the MF scheme. This forms the theoretical basis for the development of the (mixed modulation) cyclostationary noise HBT models in [59], [60]. Notice also that this demonstrates the importance of PB cyclostationary noise simulations, which, being able to model an almost ideal device structure (thus avoiding the painful de-embedding of device parasitics), allows to obtain a full insight on the inner physical nature of complex noise phenomena.

VI. CONCLUSION

The noise simulation and modeling of semiconductor devices (in particular, transistors) in large-signal forced operation has been reviewed, starting from the fundamental topics of stationary noise modulation into cyclostationary noise and covering both the physics-based approach and the compact modeling strategy, the latter being essential in the domain of circuit design and optimization. A discussion is also provided on large-signal simulation strategies, since this is often the most computationally intensive step in LS noise physics-based simulation. Particular attention has been devoted to the issue of the modulation of low-frequency noise sources, that still represents a difficulty in compact noise modeling since no exact and general strategy exist that allows to derive, at a circuit level, the large-signal noise from stationary noise; as a consequence, ad-hoc solutions have to be tailored on specific device classes. Such a general strategy indeed exists in physics-based modeling, at least whenever the low-frequency noise sources can be traced back to a superposition of local white fluctuations, which, however, is not always possible in the modeling of $1/f$ noise.

REFERENCES
