An intermodulation distortion estimation method for linear CMOS circuits

Dimitrios Baxevanakis | Vassilis Alimisis | Paul P. Sotiriadis

School of Electrical and Computer Engineering, National Technical University of Athens, 9, Iroon Polytechniou Str., Athens, 157 80, Greece

Correspondence
Dimitrios Baxevanakis, School of Electrical and Computer Engineering, National Technical University of Athens, 9, Iroon Polytechniou Str., Athens, 157 80, Greece.
Email: dimbaxev@central.ntua.gr

Funding information
State Scholarships Foundation (IKY), European Social Fund, Grant/Award Number: MIS-5000432

Summary
This article presents a general, time-domain method for intermodulation distortion estimation that offers a fast, systematic, and compact formulation. It can be applied to any linear CMOS circuit architecture, with any number of stages, ranging from amplifiers and transconductors to filters, that are characterized by weakly nonlinear behavior. Each CMOS stage of the circuit is modeled as a $G_m$-stage with an output current expressed as a more involved function of its input and output voltages, taking into account both powers and cross-product terms necessary to accurately capture the nonlinear behavior. The proposed method is easily implemented in numerical computing environments like MATLAB or Python and results in a very fast distortion estimation. A number of example topologies simulated in Cadence Spectre illustrate the application of the method and demonstrate its accuracy.

KEYWORDS
CMOS, estimation, $G_m$-stage, intermodulation distortion (IMD), linear circuits, weak nonlinearities

1 | INTRODUCTION

Intermodulation distortion is a quantity of major significance in circuit design; intermodulation and harmonic distortion characterize the linearity of a circuit, and thus their behavior is considered crucial in the performance of power amplifiers, radio-frequency amplifiers,1 low-noise amplifiers, filters, and more.

The most popular test to measure intermodulation distortion is the two-tone test,2,3 where two sinusoidal signals at frequencies $\omega_1$ and $\omega_2$ drive the circuit under consideration. The result is the generation of intermodulation products at frequencies $\pm m\omega_1 \pm n\omega_2$, where $m,n=1,2,3,...$; the sum $m+n$ is the order of the corresponding intermodulation product.

Of particular interest is the behavior of the third-order intermodulation products at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, since these frequencies are very close to the original input signals, should $\omega_1 = \omega - \delta\omega$ and $\omega_2 = \omega + \delta\omega$, with $\delta\omega \ll \omega$. As a consequence, the third-order intermodulation distortion IM3, defined as the ratio of the power of the aforementioned third-order intermodulation products to the power of the input signals, is one of the most common intermodulation distortion metrics and the estimation subject of this article.

Usually, the estimation of intermodulation distortion is performed by simulation methods like harmonic balance or shooting.4,5 However, these methods can be time-consuming and computationally expensive6,7; this has favored the development of dedicated methods for distortion estimation. The use of the Volterra series2,8,9 remains a popular approach, yielding very accurate distortion results. Its drawback is that the mathematical expressions of the involved
operators become very complicated as the number of circuit elements rises, making their manipulation practically unmanageable.

Other existing methods include the use of linear-centric circuit models to account for individual distortion contributions\(^{10}\); the method is based on the Volterra analysis framework and gives very accurate results. However, it requires a steady-state analysis to be initially performed. For weakly nonlinear fully-differential \(G_m - C\) filters of any order, a systematic state-space approach has been proposed\(^{11}\); the approach results in a fast distortion estimation, but it is developed for this particular circuit family. A distortion contribution analysis by means of the best linear approximation has been recently proposed\(^{12}\); it distinguishes itself from classic distortion methods by adopting a noise-like analysis, being able to handle complex excitation signals and strong nonlinearities. Its trade-off is an increased simulation time. Various harmonic distortion estimation methods that rely on algebraic manipulation of simplified amplifier models\(^{13-21}\) offer circuit intuition and could be modified to account for intermodulation distortion, but they are usually tailored for specific amplifier topologies, requiring extensive algebraic manipulation in order to be used for more general cases.

This article uses the same principles as our work on harmonic distortion estimation\(^{22}\) and presents its standalone counterpart for the more complex problem of estimating the intermodulation distortion in CMOS circuits that exhibit weakly nonlinear behavior. The proposed method offers a compact and systematic way applicable to general circuit structures with any number of stages, ranging from simple transconductor stages to filters and cascaded amplifiers. The estimation of \(\text{IM}_3\) is characterized by a high level of accuracy and is performed on a \(G_m\)-stage equivalent model of the circuit under consideration. Each \(G_m\)-stage captures the current-characteristic of its corresponding stage accurately by employing a more involved current model and curve-fitting. All the necessary mathematical expressions are provided to the potential reader, enabling the immediate application of the proposed method. The method is easily implemented in numerical computing environments like MATLAB or Python and gives a very fast distortion estimation due to its formulation.

The remainder of this article is organized as follows. Section 2 presents the modeling of CMOS stages and the derivation of the model coefficients, while Section 3 introduces the proposed method for the estimation of intermodulation distortion. Section 4 validates the method’s accuracy by comparison with simulation, and Section 5 concludes the article.

2 | MODELING CMOS STAGES AS \(G_m\)-STAGES

CMOS circuits are eventually interconnections of distinct CMOS stages. A CMOS stage that produces an output current as a response to an input voltage can be considered and modeled as a \(G_m\)-stage, like the one presented in Figure 1. Generally, the output current of a CMOS stage is a nonlinear function of its input and output voltages. This nonlinear nature causes distortion generation; thus, it comprises the vital behavior to be captured by the stage’s corresponding \(G_m\)-stage model.

It is reported that parasitic capacitances in MOS transistors do not substantially contribute to the generation of distortion; mainly, they reduce the magnitude of a stage’s output impedance at high frequencies.\(^{23,24}\) This makes possible the formation of a \(G_m\)-stage model that is based solely on the DC-characteristics of its corresponding CMOS stage. A way of obtaining such a model is by approximating the stage’s output current with a power-series expansion around its DC-operating point.

2.1 | Proposed \(G_m\)-stage model

Throughout this article, \(G_{m,ij}^{\text{th}}\) marks the \(G_m\)-stage that has positive input, \(u_i\), at node \(i\); negative input, \(k_{ij}u_t\), from node \(t\), with \(k_{ij}\) a real feedback factor\(^*\); and produces output current, \(i_{ij}\), at node \(j\). The stage’s differential input voltage is

\(^*\)This particular notation is adopted to capture a range of popular topologies.
\( \tilde{u}_{ij} = u_i - k_{ij} u_j. \)  \hspace{1cm} (1)

The \( itj \)-triplet of notation is included in all of the characteristics of a particular \( G_m \)-stage, whereas the AC-ground is marked as \( r \) (reference potential). The tilde symbol \( (\sim) \) placed above a quantity is used to highlight that that specific quantity is input-related. It is also used to mark coupling capacitances, as will be stated later.

Most \( G_m \)-stage models express the stage’s output current as a power series of its input and output voltages, where input- and output-related terms are independent of one another.\(^{14,15,17,18,20}\) The absence of cross-terms between the two voltages can however lead to significant deviations and errors.\(^{22,25,26}\)

An accurate model must include cross-products of the input and output voltages of the stage. This has been done at transistor level,\(^{1,23,24,27,28}\) where the MOS device acting as amplifier is supposed to admit a 2D or 3D Taylor series expansion, and the coefficients of the approximation are calculated by means of the partial derivatives of the transistor’s current relationship.

In this article, it is assumed that each \( G_m \)-stage’s operation is characterized by weakly nonlinear behavior and that its output current is considered to have a time-domain\(^{11,29-32}\) power-series expression\(^{22,26}\)

\[
i_{ij} = \sum_{k, \ell \geq 0} g_{ij}^{k, \ell} \tilde{u}_{ij}^k u_j^{\ell}.
\]  \hspace{1cm} (2)

In order to combine good accuracy with reasonable complexity, it is set \( k + \ell = 1, 2, 3 \), and (2) becomes

\[
i_{ij} = g_{00}^{10} \tilde{u}_{ij} + g_{00}^{20} \tilde{u}_{ij}^2 + g_{02}^{30} \tilde{u}_{ij}^3 + g_{01}^{11} u_j + g_{02}^{02} u_j^2 + g_{03}^{13} u_j^3 + g_{11}^{11} \tilde{u}_{ij} u_j + g_{12}^{21} \tilde{u}_{ij}^2 u_j + g_{13}^{12} \tilde{u}_{ij} u_j^2.
\]  \hspace{1cm} (3)

The \( g_{ij}^{k, \ell} \)-coefficients of (3) characterize each \( G_m \)-stage.

2.2 \ G_m -stage equivalent circuit representation

An equivalent \( G_m \)-stage representation of a circuit can be constructed by identifying the CMOS stages that the circuit is composed of, and their interconnections. Standard CMOS stages include the differential pair, the common-source, common-gate, and source-follower stages; stages that result from combinations of them, such as a cascode stage, can either be represented by the individual \( G_m \)-stages of their basic blocks, or even be treated as a single, individual stage. Fully-differential structures can be easily modeled by single-ended \( G_m \)-stages. More details on the formulation of CMOS stages as \( G_m \)-stages and the construction of a circuit’s \( G_m \)-stage equivalent can be found in our work on harmonic distortion estimation.\(^{22}\)

A specific procedure for identifying the CMOS stages and decomposing a circuit automatically is out of the scope of this article. A possible, semi-automatic option would require the designer to mark each distinct stage during the design process, in the same way different subcircuits are represented and connected by their higher-level symbols. Then, the process of deriving the \( g_{ij}^{k, \ell} \)-coefficients of each stage and the mapping of the circuit’s \( G_m \)-stage equivalent could be automated.

2.3 \ G_m -stage model coefficients

Most of the works in the literature and various analyses derive the model’s coefficients by a Taylor series expansion at the DC-operating point of the stage. Such an approach, however, yields accurate results only locally and for small signal
amplitudes. In order to capture in detail the amplitude-adjusted nonlinearities of each stage, the $g_{ij}^{k\ell}$-coefficients of (3) of each model are instead derived by curve-fitting, through a three-step derivation procedure that involves the following.

### 2.3.1 | AC-analysis for estimation of amplitude levels

Given the decomposition of the complete circuit into its distinct CMOS stages, an initial AC-analysis is performed in order to estimate the maximum expected signal amplitude at the input and the output of each stage, in the frequency range of interest.

### 2.3.2 | 2D DC-sweeps for model coefficients derivation

Each stage of the circuit is set to its input and output DC-operating points, and a 2D DC-sweep of its unloaded output current is performed. The sweeping ranges for the stage’s input and output voltages are the maximum input and output amplitudes derived in the AC-analysis, with $r_1$ and $r_2$ number of steps, respectively.

### 2.3.3 | Model coefficients extraction via linear regression

The acquired data from the 2D DC-sweeps are processed in a linear regression fashion to derive each stage’s $g_{ij}^{k\ell}$-coefficients. More specifically, a linear least-squares problem is formed

$$I_g = UG,$$

where $I_g \in \mathbb{R}^{(r_1 \times r_2) \times 1}$ is a column matrix with the values of the stage’s output current, and the matrices $U$ and $G$ are defined as

$$U = \begin{bmatrix} \tilde{u}_{ij}, \tilde{u}_{ij}^2, \tilde{u}_{ij!}^3, u_j, u_j^2, \tilde{u}_{ij!}u_j, \tilde{u}_{ij!}u_j^2, \tilde{u}_{ij!}u_j^2 \end{bmatrix} \in \mathbb{R}^{(r_1 \times r_2) \times 9},$$

$$G = \begin{bmatrix} g_{10}, g_{10}, g_{10}, g_{10}, g_{10}, g_{10}, g_{10}, g_{10}, g_{10} \end{bmatrix}^T \in \mathbb{R}^{9 \times 1}.$$

The solution of (4) gives the $g_{ij}^{k\ell}$-coefficients of the corresponding stage.

### 3 | INTERMODULATION DISTORTION ESTIMATION

The proposed intermodulation estimation method can be performed in general circuit structures, like the one of Figure 2. The circuit may be composed of $G_m$-stages, resistors, and capacitors. Each circuit node $j$, $j = 0, 1, \ldots, n$, can have a resistor, $R_j$, and a capacitor, $C_j$, connected to ground, while coupling between nodes $\ell$ and $j$ can be provided by capacitor $\tilde{C}_{\ell j}$. Each node can also have an excitation signal as an independent current source, $\hat{i}_j$.

For each node $j$, it is

$$\hat{i}_j + \sum_{\ell \neq j} i_{\ell j} + \sum_{\ell} i_{\tilde{C}_{\ell j}} = \frac{u_j}{R_j} + C_j \hat{u}_j,$$

where current $i_{\tilde{C}_{\ell j}}$ of coupling capacitor $\tilde{C}_{\ell j}$ is given by

---

8A more detailed description of the proposed derivation procedure can be found in our work on harmonic distortion estimation.
Thus, (7) is reorganized as

$$i_{\tau j} = \tilde{C}_{\tau j}(\dot{u}_\tau - \ddot{u}_j) = -i_{\tau j}.$$  \hspace{1cm} (8)

Thus, (7) is reorganized as

$$i + \sum_{i \ell} i_{\tau j} + \sum_{\tau} \tilde{C}_{\tau j} \dot{u}_\tau = \frac{u_j}{R_j} + \left(C_j + \sum_{\tau} \tilde{C}_{\tau j}\right) \ddot{u}_j.$$  \hspace{1cm} (9)

In order to estimate the intermodulation distortion $I_{M_3}$, it is assumed that the circuit operates in steady-state and that the voltage of node $j$ is of the form

$$u_j = u_j^f + u_j^m,$$  \hspace{1cm} (10a)

$$u_j^f = \theta^f S_j^f,$$  \hspace{1cm} (10b)

$$u_j^m = \theta^m S_j^m.$$  \hspace{1cm} (10c)

That is, it is assumed that the voltage of each node $j$ has two components. The first component, $u_j^f$, includes the fundamental tones at frequencies $\omega_1$ and $\omega_2$. The component of $u_j^m$ represents the desired intermodulation products at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, and also includes the dominant harmonic and intermodulation tones that are involved with their generation; due to the nature of intermodulation distortion, including only $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ in $u_j^m$ would result in a poor estimation of $I_{M_3}$. The additional tones interact with the fundamental ones and significantly contribute to the power levels at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$.

For a combination of good accuracy and reasonable complexity, the additional terms that are taken into account are products of up to the third-order. Given the assumption of weakly nonlinear stage behavior, products of higher order than that will have much smaller power, and so, negligible contribution to the desired intermodulation products. Thus, the additional tones considered are $\omega_2 - \omega_1$, $2\omega_1$, $2\omega_2$, $2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$, $3\omega_1$, and $3\omega_2$, where it is assumed that $\omega_2 > \omega_1$.

Returning to (10b) and (10c), vectors $\theta^f$ and $\theta^m$ contain the sin and cos terms for the two components, while vectors $S_j^f$ and $S_j^m$ feature the corresponding amplitude coefficients. So, it is

$$S_j^f = [a_{j,1}, b_{j,1}, c_{j,1}, d_{j,1}]^T \in \mathbb{R}^{4 \times 1},$$  \hspace{1cm} (11)

$^8$ Vectors $\theta$ and $\theta^m$ are functions of time, so a more accurate notation would be that of $\theta(t)$ and $\theta^m(t)$; time is omitted for simplicity.
The excitation current source of node \( j \), \( \hat{i}_j \), is described in the same way by

\[
\hat{i}_j = \hat{i}_j^f + \hat{i}_j^m,
\]

\[
\hat{i}_j^f = \theta^f P_j^f,
\]

\[
\hat{i}_j^m = \theta^m P_j^m,
\]

where

\[
P_j^f = [\hat{a}_{j,1}, \hat{b}_{j,1}, \hat{c}_{j,1}, \hat{d}_{j,1}]^T \in \mathbb{R}^{4 \times 1},
\]

\[
P_j^m = [\hat{e}_j, \hat{m}_j, \hat{a}_{j,2}, \hat{b}_{j,2}, \hat{c}_{j,2}, \hat{d}_{j,2}, \hat{e}_j, \tilde{f}_j, \hat{h}_j, \hat{r}_j, \hat{p}_j, \hat{v}_j, \hat{x}_j, \hat{y}_j, \hat{a}_{j,3}, \hat{b}_{j,3}, \hat{c}_{j,3}, \hat{d}_{j,3}]^T \in \mathbb{R}^{18 \times 1}.
\]

For convenience, it is assumed that there is only one excitation signal in the circuit under consideration, and it is placed at node 0. If more than one excitation signals exist, they are included as independent current sources at the corresponding nodes.

The estimation of IM3 requires the estimation of the fundamental tones coefficients and the coefficients of the intermodulation products; that is, it is required to know vector \( S_j^f \) and the elements \( e_j, f_j, h_j \), and \( r_j \) of \( S_j^m \), for all \( j \); thus, to compute \( S_j^f \) and \( S_j^m \), for all \( j \).

### 3.1 Fundamental tones estimation

The fundamental tones at \( \omega_1 \) and \( \omega_2 \) are expected to be unaltered by involvement of intermodulation and harmonic products since all \( G_m \)-stages are assumed to have weakly nonlinear behavior. As such, the component \((10b)\) of each node \( j \) should satisfy the linear part of \((9)\). Thus, the output current of stage \( G_{L_j} \), \( I_{ij}^f \), is considered to be the linear part of \((3)\)

\[
i_{ij}^f = \theta^f \left( g_{ij}^{10} S_i^f - g_{ij}^{10} k_{ij} S_i^f + g_{ij}^{01} S_i^f \right).
\]

Relying on \((9)\), the above expression of \( i_{ij}^f \) can be exploited in order to form a system of equations for the coefficients of the fundamental tones of the complete circuit. Let
\[ L'_{\omega} = \text{diag}(\omega_1, \omega_2) \otimes \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 4}, \]  

(19)

where \( \text{diag}(\ldots) \) is a diagonal matrix and \( \otimes \) denotes the Kronecker’s product.\(^{35}\) Then, the derivative of (10b) can be expressed as

\[ \hat{u}'_j = \theta' \hat{S}'_j = \theta' L'_{\omega} \hat{S}'_j. \]  

(20)

The above results are valid for any \( t \in \mathbb{R} \). Since \( \theta' \) consists of four linearly independent functions of time and embodies four independent equations with respect to the coefficients of sin and cos, it can be eliminated. Thus, the combination of (9), (18), and (20), followed by the elimination of \( \theta' \), yields for node \( j \)

\[ P'_j + \sum_{l,t} g_{ij}^{10} S'_j - \sum_{l,t} g_{ij}^{10} k_{ij} S'_j + \sum_{l} \hat{C}_{ij} L'_{\omega} \hat{S}'_j = \left( \frac{1}{R_j} - \sum_{l,t} g_{ij}^{01} \right) S'_j + \left( C_j + \sum_{l} \hat{C}_{ij} \right) L'_{\omega} \hat{S}'_j. \]  

(21)

Let the vector of the coefficients of the fundamental tones of all voltages be

\[ S' = \left[ \left( S'_0 \right)^T, \left( S'_1 \right)^T, ..., \left( S'_n \right)^T \right]^T \in \mathbb{R}^{4(n+1)\times 1}, \]  

(22)

and, accordingly, consider the vector of the coefficients of the excitation current source\(^{a}\)

\[ P' = \left[ \left( P'_0 \right)^T, (0, 0, 0, 0), ..., (0, 0, 0, 0) \right]^T \in \mathbb{R}^{4(n+1)\times 1}. \]  

(23)

Then, the equivalent of (21) for all nodes can be constructed in block-matrix form

\[ P' + (G' + K' + F') S' = (T' + W') S'. \]  

(24)

The matrices involved in (24) are defined as

\[ G' = \left[ \sum_{l,t} g_{ij}^{10} \right]_{j=0}^{n} \otimes I_4 \in \mathbb{R}^{4(n+1)\times 4(n+1)}, \]  

(25)

\[ K' = \left[ -\sum_{l,t} g_{ij}^{10} k_{ij} \right]_{j=0}^{n} \otimes I_4 \in \mathbb{R}^{4(n+1)\times 4(n+1)}, \]  

(26)

\[ F' = \left[ \hat{C}_{ij} \right]_{j'=0}^{n} \otimes L'_{\omega} \in \mathbb{R}^{4(n+1)\times 4(n+1)}, \]  

(27)

\[ T' = \text{diag}\left( \left[ \frac{1}{R_j} - \sum_{l,t} g_{ij}^{01} \right]_{j=0}^{n} \right) \otimes I_4 \in \mathbb{R}^{4(n+1)\times 4(n+1)}, \]  

(28)

\(^{a}\)If more than one excitation signals are present, they should be included at the corresponding entries of the vector.
\[ W^f = \text{diag}\left( \left[ C_j + \sum_{\ell} \tilde{C}_{\ell j} \right]_{j=0}^n \right) \otimes L^f_{\infty} \in \mathbb{R}^{4(n+1) \times 4(n+1)}, \]  

(29)

where \( I_n \) is the \( n \times n \) identity matrix. The coefficients of the fundamental tones are immediately obtained by

\[ S^f = [T^f - G^f - K^f + W^f - F^f]^{-1} P^f. \]  

(30)

### 3.2 Intermodulation products estimation

The nonlinear terms of (3) are the cause of the generation of the intermodulation and harmonic products, (10c), in each node \( j \). The estimation of the distortion terms by the inclusion of all generated terms in (3) after substitution of all voltages in the form of (10a) results in a nonlinear problem that is challenging and not computationally efficient.

Inspecting (3) alongside (10a), it follows that the generated coefficients of the intermodulation and harmonic products will ultimately be a sum of products of

a. only coefficients of fundamental tones,

b. a single intermodulation or harmonic coefficient to the power of one and one or more fundamental tone coefficients, and

c. higher orders or products of intermodulation and harmonic coefficients.

Products (c) will contribute negligible power and can be safely ignored,\(^{22}\) for the amplitudes of the intermodulation and harmonic products are expected to be much smaller compared to the ones of the fundamental tones. As such, for the estimation of the intermodulation products, the output current of stage \( G_{m_i}^j \),

\[ \begin{aligned} i_{ij}^m &= \theta^m \left[ g_{ij}^{[0]} S_{ij}^m - g_{ij}^{[3]} k_{ij} S_{ij}^m + g_{ij}^{[01]} S_{ij}^m + Z_{ij} + X_{ij}^{[\alpha]} S_{ij}^m + X_{ij}^{[\beta]} S_{ij}^m + X_{ij}^{[\gamma]} S_{ij}^m \right]. \end{aligned} \]  

(31)

Equation (31) takes into account the linear part of (3) and the kept products (a) and (b), included in terms \( \theta^m Z_{ij} \) and \( \theta^m \left( X_{ij}^{[\alpha]} S_{ij}^m + X_{ij}^{[\beta]} S_{ij}^m + X_{ij}^{[\gamma]} S_{ij}^m \right) \), respectively. The above reasoning results in a linear form of \( i_{ij}^m \), making the estimation problem of the intermodulation coefficients linear. The terms introduced in (31) are given by

\[ \begin{aligned} Z_{ij} &= \begin{bmatrix} (Z_{ij}^{[\alpha]} \end{bmatrix}^T, (Z_{ij}^{[\beta]} \end{bmatrix}^T, (Z_{ij}^{[\gamma]} \end{bmatrix}^T \right)^T \in \mathbb{R}^{18 \times 1}, \end{aligned} \]  

(32)

where

\[ \begin{aligned} Z_{ij}^{[\alpha]} &= g_{ij}^{[0]} A_{ij} + g_{ij}^{[02]} A_j + g_{ij}^{[11]} \frac{1}{2} H_{ij} \in \mathbb{R}^{6 \times 1}, \end{aligned} \]  

(33)

\[ \begin{aligned} Z_{ij}^{[\beta]} &= g_{ij}^{[03]} \frac{3}{2} Y_{ij} + g_{ij}^{[03]} \frac{3}{2} Y_j + g_{ij}^{[1]} \frac{1}{2} V_{ij} + g_{ij}^{[11]} \frac{1}{2} V_{ij} \in \mathbb{R}^{8 \times 1}, \end{aligned} \]  

(34)

\[ \begin{aligned} Z_{ij}^{[\gamma]} &= g_{ij}^{[03]} \frac{1}{4} R_{ij} \bigcirc S_{ij}^f + g_{ij}^{[03]} \frac{1}{4} R_j \bigcirc S_j^f + g_{ij}^{[2]} \frac{1}{2} E_{ij} S_{ij}^f + g_{ij}^{[12]} \frac{1}{2} E_j S_{ij}^f \in \mathbb{R}^{4 \times 1}, \end{aligned} \]  

(35)

and \( \bigcirc \) denotes the Hadamard’s product.\(^{36}\) Matrices \( X_{ij}^{[\alpha]} \), \( X_{ij}^{[\beta]} \), and \( X_{ij}^{[\gamma]} \) are defined as

\[ \begin{aligned} X_{ij}^{[\alpha]} &= g_{ij}^{[0]} N_{ij} + g_{ij}^{[3]} \frac{3}{2} M_{ij} + g_{ij}^{[1]} \frac{1}{2} N_j + g_{ij}^{[2]} \frac{1}{2} Q_{ij} + g_{ij}^{[1]} \frac{1}{2} M_j \in \mathbb{R}^{18 \times 18}, \end{aligned} \]  

(36)
\[ X_{ij}^\beta = -k_{ij}X_{ij}^\alpha \in \mathbb{R}^{18 \times 18}, \tag{37} \]

\[ X_{ij}^\gamma = g_{ij}^{02}N_j + g_{ij}^{03}3\hat{M}_j + g_{ij}^{11}1\hat{N}_{ij} + g_{ij}^{21}1\hat{M}_{ij} + g_{ij}^{12}1Q_{ij} \in \mathbb{R}^{18 \times 18}. \tag{38} \]

The quantities that form the matrices of (33)–(38) are omitted for reading comprehension purposes. They can be found in the Appendices A and B of this article.

In a similar manner to the estimation procedure of the coefficients of the fundamental tones, the expression (31) of \( i_{ij} \) is used to form the corresponding system of equations for the intermodulation coefficients, relying again on (9). Let

\[ L_m = \text{diag}(\omega_2 - \omega_1, 2\omega_1, 2\omega_2, 2\omega_1 - \omega_2, 2\omega_2 - \omega_1, 2\omega_1 + \omega_2, 2\omega_2 + \omega_1, 3\omega_1, 3\omega_2) \otimes \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbb{R}^{18 \times 18} \tag{39} \]

to express the derivative of (10c)

\[ \dot{i}_{ij}^m = \dot{\theta}^m L_m S_j^m. \tag{40} \]

By similar reasoning as earlier, vector function \( \theta^m \) can be eliminated. Combining (9), (31), and (40), while eliminating \( \theta^m \), results in the following equation for node \( j \):

\[ \begin{aligned}
& \left( P_j^m + \sum_{l,l}Z_{lj} \right) + \sum_{l,l}g_{lj}^{10}S_l^m - \sum_{l,l}g_{lj}^{10}k_{lj}S_l^m + \sum_{l,l} \left( X_{lj}^\alpha S_l^m + X_{lj}^\beta S_l^m + X_{lj}^\gamma S_l^m \right) + \sum_\ell \tilde{C}_{\ell j}L_{\ell m}S_j^m \\
= & \left( \frac{1}{R_j} - \sum_{l,l}g_{lj}^{01} \right) S_j^m + \left( C_j + \sum_\ell \tilde{C}_{\ell j} \right) L_{\ell m}S_j^m. \end{aligned} \tag{41} \]

Consider the vector of the coefficients of the intermodulation products of all voltages of the circuit

\[ S^m = \left[ \left( S_0^m \right)^T, \left( S_1^m \right)^T, ..., \left( S_n^m \right)^T \right]^T \in \mathbb{R}^{18(n+1) \times 1} \tag{42} \]

and the vector of the excitation current source\(^\perp\)

\[ p^m = \left[ \left( p_0^m \right)^T, \left( 0^1 \right)^T, ..., \left( 0^1 \right)^T \right]^T \in \mathbb{R}^{18(n+1) \times 1}. \tag{43} \]

Moreover, define

\[ B^m = p^m + Z_m \in \mathbb{R}^{18(n+1) \times 1}, \tag{44} \]

where

\(^\perp\)In the case of more than one excitation signals being present, they should be added at the corresponding entries of the vector.
\[ Z^m = \left[ \sum_{i,j} Z_{ij} \right]_{j=0}^n \in \mathbb{R}^{18(n+1) \times 1}. \]  

(45)

As with the case of the fundamental tones, the equivalent of (21) for the complete circuit is grouped in block-matrix form

\[ B^m + (G^m + K^m + X^m + F^m)S^m = (T^m + W^m)S^m, \]

(46)

where

\[ G^m = \left[ \sum_{i,j} g_{ij}^{10} \right]_{j=0}^n \otimes I_{18} \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(47)

\[ K^m = \left[ -\sum_{i,j} k_{ij}^{10} \right]_{j=0}^n \otimes I_{18} \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(48)

\[ F^m = \left[ \tilde{C}_{ij} \right]_{j=0}^n \otimes L^m \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(49)

\[ T^m = \text{diag} \left( \frac{1}{R} - \sum_{i,j} t_{ij}^{01} \right)_{j=0}^n \otimes I_{18} \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(50)

\[ W^m = \text{diag} \left( C + \sum_{j} \tilde{C}_{j} \right)_{j=0}^n \otimes L^m \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(51)

\[ X^m = \left[ \sum_{j} X_{ij}^\alpha \right]_{j=0}^n + \left[ \sum_{j} X_{ij}^\beta \right]_{j=0}^n + \bigoplus_{j=0}^n \left[ \sum_{i} X_{ij}^\gamma \right] \in \mathbb{R}^{18(n+1) \times 18(n+1)}, \]

(52)

and \( \bigoplus \) denotes the direct sum of matrices.36 The desired coefficients of the intermodulation products are given by

\[ S^m = [T^m - G^m - K^m - X^m + W^m - F^m]^{-1} B^m, \]

(53)

and the IM3 can be estimated at any node \( j \) as

\[ \text{IM}_3^j = 10\log_{10} \left( \frac{e_j^2 + f_j^2 + h_j^2 + r_j^2}{a_{j,1}^2 + b_{j,1}^2 + c_{j,1}^2 + d_{j,1}^2} \right). \]

(54)

An advantage of (54) over the classic approach of estimating IM3 by the power ratio of the tone at \( 2\omega_1 - \omega_2 \) over that at \( \omega_1 \) is that (54) takes into account the behavior of both input signals and the corresponding intermodulation products with respect to frequency. Moreover, it allows the input signals to feature different amplitude values.

The method offers a very fast distortion estimation due to its formulation; the estimation of IM3 is acquired by the solution of two linear problems. Thus, the proposed method gives a closed-form solution, while methods like harmonic balance or shooting are iterative and converge to a solution subject to a specific tolerance. Another advantage of the proposed method over methods like the aforementioned ones is that it is unaffected by the value of the beat frequency that may significantly slow down the latter.

### 4 | SIMULATION RESULTS

Two simulation cases validate the proposed intermodulation distortion estimation method. The two circuit examples are implemented in TSMC 90nm technology, and the distortion results are obtained by Cadence Spectre parametric PSS-analysis in harmonic balance mode. In both cases, supply rails are set to \( \pm 0.9 \) V. The proposed method is
implemented in MATLAB. The time of distortion estimation dropped from the order of minutes required during the parametric PSS-analysis, to the order of seconds when using the proposed method.

4.1 Two-stage feedback amplifier

As a first example, the two-stage feedback amplifier of Figure 3A is tested. The amplifier has a feedback factor of $\beta = 0.10$, a DC-gain of 19.07 dB, and a unity-gain frequency of 7.03 MHz, under a load of 10 kΩ $\parallel$ 2 pF.

Its equivalent $G_m$-stage representation is that of Figure 3B; the differential pair of $M_0 - M_4$ is $G_m^{\text{diff}}_0, 2, 1$, and the common-source stage of $M_5 - M_6$ forms $G_m^{\text{cs}}_{1, 2}$. Miller capacitor $C_C$ is represented by $\tilde{C}_{12}$, $R_{L_1} + R_{L_2} = R_L$ by $R_2$, and $C_L$ by $\tilde{C}_C$.

![Two-stage feedback amplifier](image1.png)

**FIGURE 3** Two-stage feedback amplifier and its $G_m$-stage equivalent representation

![IM3 of two-stage feedback amplifier](image2.png)

**FIGURE 4** IM3 of the two-stage feedback amplifier for different $\delta_f$ cases [Colour figure can be viewed at wileyonlinelibrary.com]
Finally, the amplifier’s feedback factor is captured by \( k_{021} = \frac{R_{L2}}{(R_{L1} + R_{L2})} \), and the AC-input signal is realized by the current source \( i_0 = u_{in}/R_0 \) that is injected in \( R_0 = 1\Omega \).

The amplifier is driven by two signals with an amplitude of 20 mV peak, and fundamental frequencies of \( \omega_1 = 2\pi(1 - \delta_f) f \) and \( \omega_2 = 2\pi(1 + \delta_f) f \). Figure 4 depicts the comparison of the IM3 results obtained by simulation to the ones of the proposed method, for \( \delta_f = 0.01 \), \( \delta_f = 0.05 \), and \( \delta_f = 0.10 \). The error in the entire frequency range is found to be less than 0.41 dB for all three cases of \( \delta_f \), indicating a good agreement between the two results.

### 4.2 Fourth-order butterworth low-pass filter

Next, a fourth-order butterworth low-pass filter is simulated. The filter’s architecture\(^{37,38} \) is shown in Figure 5A, and the employed OTA is presented in Figure 5B. The filter has a cut-off frequency of 98.82 kHz, and its \( G_m \)-stage equivalent representation of Figure 5C is immediately derived; each OTA is handled as a single \( G_m \)-stage, and the AC-input signal of the structure is again realized by the current source \( i_0 = u_{in}/R_0 \) acting on \( R_0 = 1\Omega \).

With two signals of 100 mV peak amplitude, and fundamental frequencies of \( \omega_1 = 2\pi(1 - \delta_f) f \) and \( \omega_2 = 2\pi(1 + \delta_f) f \), the obtained IM3 results for \( \delta_f = 0.01 \), \( \delta_f = 0.05 \), and \( \delta_f = 0.10 \) are given in Figure 6. The results of the proposed method are found to be in fine agreement with the simulation ones. In the entire frequency range, the error is less than 0.67 dB for the case of \( \delta_f = 0.01 \), less than 0.58 dB for \( \delta_f = 0.05 \), and less than 0.51 dB for \( \delta_f = 0.10 \).

![Fourth-order butterworth low-pass filter](image)

**Figure 5** Fourth-order butterworth low-pass filter, its employed OTA, and the filter’s \( G_m \)-stage equivalent representation.
In this article, a general, time-domain method for estimation of intermodulation distortion in CMOS circuits is presented that can be systematically applied to circuit topologies with any number of stages. It can be easily implemented in numerical computing environments like MATLAB or Python, and provides fast distortion results that are in good agreement with the ones obtained by Cadence Spectre simulation.

ACKNOWLEDGMENTS

This research is cofinanced by Greece and the European Union (European Social Fund–ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning” in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

ORCID

Dimitrios Baxevanakis © https://orcid.org/0000-0001-5297-1456
Vassilis Alimisis © https://orcid.org/0000-0002-2090-1493
Paul P. Sotiriadis © https://orcid.org/0000-0001-6030-4645

REFERENCES


FIGURE 6 IM3 of the fourth-order butterworth low-pass filter for different \( \delta_f \) cases [Colour figure can be viewed at wileyonlinelibrary.com]


APPENDIX A: QUANTITIES FORMING MATRICES (33)–(35)

Matrices (33)–(35) are formed by

\[
\begin{align*}
\hat{A}_{ij} &= \begin{bmatrix}
-\hat{m}_{ij} & \hat{m}_{ij} & \hat{\eta}_{ij} & \hat{\eta}_{ij} & \hat{f}_{ij} & \hat{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{6 \times 1} \\
\tilde{H}_{ij} &= \begin{bmatrix}
-\tilde{q}_{ij} & \tilde{q}_{ij} & \tilde{\eta}_{ij} & \tilde{\eta}_{ij} & \tilde{f}_{ij} & \tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{6 \times 1} \\
\tilde{V}_{ij} &= \begin{bmatrix}
0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij} & 0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{8 \times 1} \\
\tilde{D}_{ij} &= \begin{bmatrix}
0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij} & 0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{8 \times 1} \\
O_{ij} &= \begin{bmatrix}
0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij} & 0 & 0 & -\tilde{f}_{ij} & \tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{8 \times 4} \\
\tilde{R}_{ij} &= \begin{bmatrix}
-\tilde{a}_{ij,1}^2 + 3\tilde{b}_{ij,1}^2 \\
-3\tilde{a}_{ij,1}^2 + 5\tilde{b}_{ij,1}^2 \\
-\tilde{c}_{ij,1}^2 + 3\tilde{d}_{ij,1}^2 \\
-3\tilde{c}_{ij,1}^2 + 5\tilde{d}_{ij,1}^2
\end{bmatrix} \in \mathbb{R}^{4 \times 1} \\
E_{ij} &= \begin{bmatrix}
\tilde{f}_{ij} & -\tilde{f}_{ij} & \tilde{f}_{ij} & -\tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{4 \times 1} \\
\tilde{f}_{ij} & \tilde{f}_{ij} & \tilde{f}_{ij} & -\tilde{f}_{ij}
\end{bmatrix}^T \in \mathbb{R}^{4 \times 1}
\end{align*}
\]

where
Finally, the forming matrices of (36) are defined as

\[
\begin{align*}
\tilde{m}_{ij}^a &= a_{ij,1} c_{ij,1} - \tilde{b}_{ij,1} \tilde{d}_{ij,1} \\
\tilde{m}_{ij}^\rho &= a_{ij,1} d_{ij,1} + b_{ij,1} c_{ij,1} \\
\tilde{m}_{ij}^\gamma &= a_{ij,1} \tilde{d}_{ij,1} - \tilde{b}_{ij,1} c_{ij,1} \\
\tilde{m}_{ij}^\delta &= a_{ij,1} c_{ij,1} + b_{ij,1} d_{ij,1} \\
\tilde{f}_{ij}^a &= a_{ij,1} b_{ij,1} \\
\tilde{f}_{ij}^\rho &= \frac{-a_{ij,1}^2 + b_{ij,1}^2}{2} \\
\tilde{f}_{ij}^\gamma &= c_{ij,1} d_{ij,1} \\
\tilde{f}_{ij}^\delta &= \frac{-c_{ij,1}^2 + d_{ij,1}^2}{2} \\
\sigma_{ij}^a &= \tilde{a}_{ij,1} a_{ij,1} - \tilde{b}_{ij,1} b_{ij,1} \\
\sigma_{ij}^\rho &= \tilde{a}_{ij,1} b_{ij,1} + \tilde{b}_{ij,1} a_{ij,1} \\
\sigma_{ij}^\gamma &= \tilde{c}_{ij,1} c_{ij,1} - \tilde{d}_{ij,1} d_{ij,1} \\
\sigma_{ij}^\delta &= \tilde{c}_{ij,1} d_{ij,1} + \tilde{d}_{ij,1} c_{ij,1} \\
\end{align*}
\]

and

\[
\tilde{S}_{ij}^f = S_{ij}^f - k_{ij} S_{ij}^f \in \mathbb{R}^{2 \times 1}
\]

that also implies \( \tilde{a}_{ij,1} = a_{ij,1} - k_{ij} a_{ij,1}, \tilde{b}_{ij,1} = b_{ij,1} - k_{ij} b_{ij,1}, \tilde{c}_{ij,1} = c_{ij,1} - k_{ij} c_{ij,1}, \) and \( \tilde{d}_{ij,1} = d_{ij,1} - k_{ij} d_{ij,1}. \) All vectors \( S_{ij}^f, C_{ij}^f, \) and \( S_{ij}^f \) are known from the solution of \( S_f \) by (30).

**APPENDIX B: QUANTITIES FORMING MATRICES (36)–(38)**

Finally, the forming matrices of (36)–(38) are defined as

\[
\begin{align*}
\tilde{N}_{ij} &= \begin{bmatrix}
0_b & \tilde{\Theta}_{ij}^N & \tilde{\Xi}_{ij}^N \\
(\tilde{\Theta}_{ij}^N)^T & 0_b & 0_b \\
(\tilde{\Xi}_{ij}^N)^T & 0_b & 0_b 
\end{bmatrix} \in \mathbb{R}^{18 \times 18} \\
N_j &= \begin{bmatrix}
0_b & \Theta_j^N & \Xi_j^N \\
(\Theta_j^N)^T & 0_b & 0_b \\
(\Xi_j^N)^T & 0_b & 0_b 
\end{bmatrix} \in \mathbb{R}^{18 \times 18} \\
\tilde{M}_{ij} &= (\tilde{S}_{ij}^f)^T \tilde{S}_{ij}^f I_{18} + \begin{bmatrix}
\tilde{\Theta}_{ij}^M & 0_b & 0_b \\
0_b & \tilde{\Xi}_{ij}^M & \Phi_{ij}^M \\
0_b & (\Phi_{ij}^M)^T & \Psi_{ij}^M 
\end{bmatrix} \in \mathbb{R}^{18 \times 18} \\
M_j &= (\tilde{S}_j^f)^T S_f I_{18} + \begin{bmatrix}
\Theta_j^M & 0_b & 0_b \\
0_b & \Xi_j^M & \Phi_j^M \\
0_b & (\Phi_j^M)^T & \Psi_j^M 
\end{bmatrix} \in \mathbb{R}^{18 \times 18} \\
Q_{ij} &= 2(\tilde{S}_{ij}^f)^T S_f I_{18} + \begin{bmatrix}
\Theta_{ij}^Q & 0_b & 0_b \\
0_b & \Xi_{ij}^Q & \Phi_{ij}^Q \\
0_b & (\Phi_{ij}^Q)^T & \Psi_{ij}^Q 
\end{bmatrix} \in \mathbb{R}^{18 \times 18} 
\end{align*}
\]

where \( 0_n \) denotes the \( n \times n \) zero matrix, and