

DC Stability Analysis of High-Order, Lowpass $\Sigma\Delta$ Modulators With Distinct Unit Circle NTF Zeros

Ngai Wong, *Member, IEEE* and Tung-Sang Ng, *Fellow, IEEE*

Abstract—This paper presents an analytical approach to the investigation of the dc stability of high-order (order > 2), low-pass (LP) $\Sigma\Delta$ modulators with distinct noise transfer function (NTF) zeros on the unit circle. The techniques of state-space diagonalization and decomposition, continuous-time embedding and Poincaré map analysis are combined and extended. It is revealed that high-order $\Sigma\Delta$ modulators can be transformed and decomposed into second- and first-order subsystems. The investigation, coupled with efficient numerical methods, generalizes itself to different types of transition flow and provides theoretical insight into the state trajectory and limit cycle behavior. It is shown that estimation of dc input bounds based solely on the boundary transition flow is inadequate. A procedure utilizing the information from different transition flow assumptions and the discrete nature of a modulator is introduced for locating the stable dc input bounds of practical, discrete-time $\Sigma\Delta$ modulators.

Index Terms—dc stability, $\Delta\Sigma$, embedding, nonlinear dynamical systems, Poincaré map, $\Sigma\Delta$, state space.

I. INTRODUCTION

DESPITE the widespread deployment of $\Sigma\Delta$ modulators [1], [2] in modern electronic products, the theoretical understandings of high-order (order > 2) $\Sigma\Delta$ modulators have been lagging behind. Much of the modulator design work still relies heavily on extensive computer simulation [3]. Among the longstanding unresolved issues, modulator stability remains a major topic which is especially critical for the conditionally-stable high-order modulators. Numerous attempts to tackle the stability problem are either limited to second-order modulators or bound to be too conservative for higher order systems. This paper presents an analytical, nonlinear dynamical approach to the dc stability study of a class of high-order, lowpass (LP) $\Sigma\Delta$ modulators with distinct noise transfer function (NTF) zeros on the complex unit circle. Special attention is paid to the derivation of stable dc input bounds.

The incentive for pursuing high-order $\Sigma\Delta$ modulators, even though they are more prone to instability, is the high signal-to-noise ratio (SNR) or “resolution” achievable without an overly high oversampling ratio (OSR) [2]. Single-bit $\Sigma\Delta$ modulators are usually preferred over their multibit counterparts due to their inherent linearity and higher tolerance of analog component im-

perfections. The assumption of distinct unit circle NTF zeros encompasses a large class of $\Sigma\Delta$ modulators including those in the popular cascade-of-resonators structure [4], [5] with optimized, distributed NTF zeros to boost the SNR by several bits [3]. Baseband or LP modulators with NTF zeros about dc are studied in this paper as dc stability is only meaningful in the context of these modulators. Stable bandpass modulators that accept inputs centered about some carrier frequencies can, however, be obtained from stable LP prototypes through appropriate transformations [2].

Current attempts to analyze the stability of $\Sigma\Delta$ modulators can mainly be classified into linear and nonlinear approaches. Linear methods [6]–[9] suffer from their approximation nature and deficiency to explain phenomena such as chaos and limit cycles present in the practical, discrete-time modulators. Nonlinear methods [10]–[20], though capable of providing rigorous theoretical treatment and explaining some of the intricate phenomena, are largely constrained by their complexity and poor scalability. A recent computational approach of Schreier *et al.* [21]–[23] that proves stability by locating a positively invariant set (PIS) also faces the barrier of exponential growth in its computational complexity and the results are too conservative when modulator order grows beyond three.

By generalizing the matrix diagonalization framework of Steiner and Yang [24] and its subsequent application to stability analysis [25]–[27], this paper addresses the case of state-space transition matrix with complex-conjugate eigenvalues that are present in many practical modulators utilizing distinct NTF zeros on the unit circle [2], [3]. It is shown that these modulators can be transformed into second- and first-order subsystems coupled through the only nonlinearity, namely, the quantizer function. With this simplification, it is possible to apply the continuous-time embedding technique of Wang [10], [12] (which used to give complicated formulas even for third-order modulators) to obtain straightforward continuous-time trajectories for modulators of arbitrary orders. The boundary transition flow assumption in the Poincaré map analysis of Wang is also waived and all possible limit cycles corresponding to different types of transition flow are obtained via efficient numerical methods [28 Ch. 10]. Fixed-point stability is analyzed with closed-form tangent linear manifolds. Criteria for dc stability of practical, discrete-time $\Sigma\Delta$ modulators are derived by utilizing the properties of the decomposed subsystems and the concept of overlapped domains of attraction (DOAs, also basins of attraction [18]). The stable dc input bound or overload limit [23] thus obtained provides a solid proof, which is lacking so far, for the dc stability of the modulators under consideration. It is also a crucial parameter for LP $\Sigma\Delta$ modulator design

Manuscript received October 25, 2001; revised December 12, 2002. This work is supported by the Hong Kong Research Grants Council and by the University Research Committee of The University of Hong Kong. This paper was recommended by Associate Editor G. Gielen.

The authors are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong (e-mail: nwong@eee.hku.hk; tsg@eee.hku.hk).

Digital Object Identifier 10.1109/TCSII.2003.808901

because it specifies the maximum dc input for which the modulator can correctly “modulate” into its output bitstream. As nominal inputs into LP $\Sigma\Delta$ modulators are slow time-varying signals relative to the oversampling rate, and in the limiting case dc signal, dc input bound gives a meaningful figure for determining the safe input range. Furthermore, dc bound is shown to be tighter than that for conventional ac signals [3] and, therefore, can also be sensibly regarded as the worst-case bound.

This paper is organized as follows. Section II begins with a state-space formulation and transforms a $\Sigma\Delta$ modulator into a representation with the least amount of coupling. The diagonalization and decomposition methods of Steiner and Yang [24] are extended to transition matrices having complex conjugate eigenvalues as a result of the distinct NTF zeros. Continuous-time embedding [12] is then applied to obtain expressions for the time evolution of state trajectories. Section III employs piecewise linear analysis to investigate the dynamics of decomposed second-order subsystem(s) and a first-order subsystem present in odd-order modulators only. Systems of nonlinear equations for locating the fixed-points under different transition flow assumptions are established in Section IV. Such equations allows for the application of efficient numerical algorithms based on a class of quasi-Newton methods. Section V gives the closed-form tangent linear manifold that characterizes the stability of a fixed-point and its associated limit cycle. Section VI discusses the DOA associated with a stable limit cycle. This concept, along with the properties of the coupled subsystems and the discrete-time constraint, is then utilized to devise a testing procedure for finding the stable dc input bound of a practical $\Sigma\Delta$ modulator. Numerical results of fifteen $\Sigma\Delta$ modulators are also given. Finally, Section VII draws the conclusion.

II. STATE-SPACE TRANSFORMATION AND EMBEDDING

This section describes two important techniques for the state-space manipulation and continuous-time modeling of a discrete-time $\Sigma\Delta$ modulator. They represent extension of the diagonalization and decomposition technique of Steiner and Yang [24] and the embedding process of Wang [12]. For brevity, only a fourth- and a fifth-order LP $\Sigma\Delta$ modulators with distinct

unit circle NTF zeros are exemplified but the procedures are readily extendable to modulators of arbitrary orders. The cascade-of-resonators architecture is chosen for illustration due to its popularity and favorable properties like unit circle NTF zeros and better robustness than other architectures [4], [5].

A. Discrete System Similarity Transforms

With respect to Fig. 1, and following a similar approach as in [29], we define the state-vector $\mathbf{x}^{(0)} = [x_1^{(0)} \ x_2^{(0)} \ \dots \ x_N^{(0)}]$ and $\rho = z - 1$. The state-space representation of a parameterized n th-order modulator is

$$\begin{cases} \rho\mathbf{x}^{(0)} = \mathbf{A}_z^{(0)}\mathbf{x}^{(0)} + \mathbf{b}_z^{(0)}u + \mathbf{a}_z^{(0)}v \\ y = b_0u + \mathbf{d}_z^{(0)}\mathbf{x}^{(0)} \\ v = \text{sgn}(y). \end{cases} \quad (1)$$

Here, $\text{sgn}(y)$ is the sign function whose output is $+1$ when $y \geq 0$ and -1 when $y < 0$. The bracketed superscripts denote the number of similarity transforms the state vectors or matrices have undergone. Specifically, for our example in Fig. 1 where $N = 4$ and 5 , see (2) at the bottom of the page. Note, that (2) depicts the case of an odd-order system ($N = 5$). For an even-order modulator ($N = 4$), the elements outside the delimiting lines are simply dropped, for example, the rightmost column and bottommost row in $\mathbf{A}_z^{(0)}$, the last element in $\mathbf{x}^{(0)}$, $\mathbf{b}_z^{(0)}$ etc. This convention of combined representation of even- and odd-order systems will be followed throughout this paper. Next, by substituting $\rho = z - 1$ and considering the special structure of \mathbf{A}_2 , the first equation of (1) is rewritten as

$$z\mathbf{x}^{(0)} = (\mathbf{I} + \mathbf{A}_2)(\mathbf{I} + \mathbf{A}_1)\mathbf{x}^{(0)} + (\mathbf{I} + \mathbf{A}_2)\mathbf{b}_z^{(0)}u + (\mathbf{I} + \mathbf{A}_2)\mathbf{a}_z^{(0)}v. \quad (3)$$

Examining Fig. 1, the NTF zeros are just the poles of the transfer function from v to y with zero input ($u = 0$). In other words, $(\mathbf{I} + \mathbf{A}_2)(\mathbf{I} + \mathbf{A}_1)$ must have eigenvalues corresponding to the distinct NTF zeros and is diagonalizable by a nonsingular matrix \mathbf{T}_d whose columns consist of the eigenvectors of $(\mathbf{I} + \mathbf{A}_2)(\mathbf{I} + \mathbf{A}_1)$, resulting in

$$\begin{cases} z\mathbf{x}^{(1)} = \mathbf{A}_z^{(1)}\mathbf{x}^{(1)} + \mathbf{b}_z^{(1)}u + \mathbf{a}_z^{(1)}v \\ y = b_0u + \mathbf{d}_z^{(1)}\mathbf{x}^{(1)} \end{cases} \quad (4)$$

$$\begin{aligned} \mathbf{x}^{(0)} &= [x_1^{(0)} \ x_2^{(0)} \ x_3^{(0)} \ x_4^{(0)} \ |x_5^{(0)}]^T \\ \mathbf{b}_z^{(0)} &= [b_1k_1 \ b_2k_1k_2 \ b_3k_1k_2k_3 \ b_4k_1 \dots k_4 \ |b_5k_1 \dots k_5]^T \\ \mathbf{a}_z^{(0)} &= -[a_1k_1 \ a_2k_1k_2 \ a_3k_1k_2k_3 \ a_4k_1 \dots k_4 \ |a_5k_1 \dots k_5]^T \\ \mathbf{d}_z^{(0)} &= [k_1^{-1} \ 0 \ 0 \ 0 \ |0] \\ \mathbf{A}_z^{(0)} &= \mathbf{A}_1 + z\mathbf{A}_2 \\ &= \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ -r_1k_2 & 0 & k_3^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -r_2k_4 & 0 & k_5^{-1} \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] + z \left[\begin{array}{cccc|c} 0 & k_2^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_4^{-1} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{aligned} \quad (2)$$

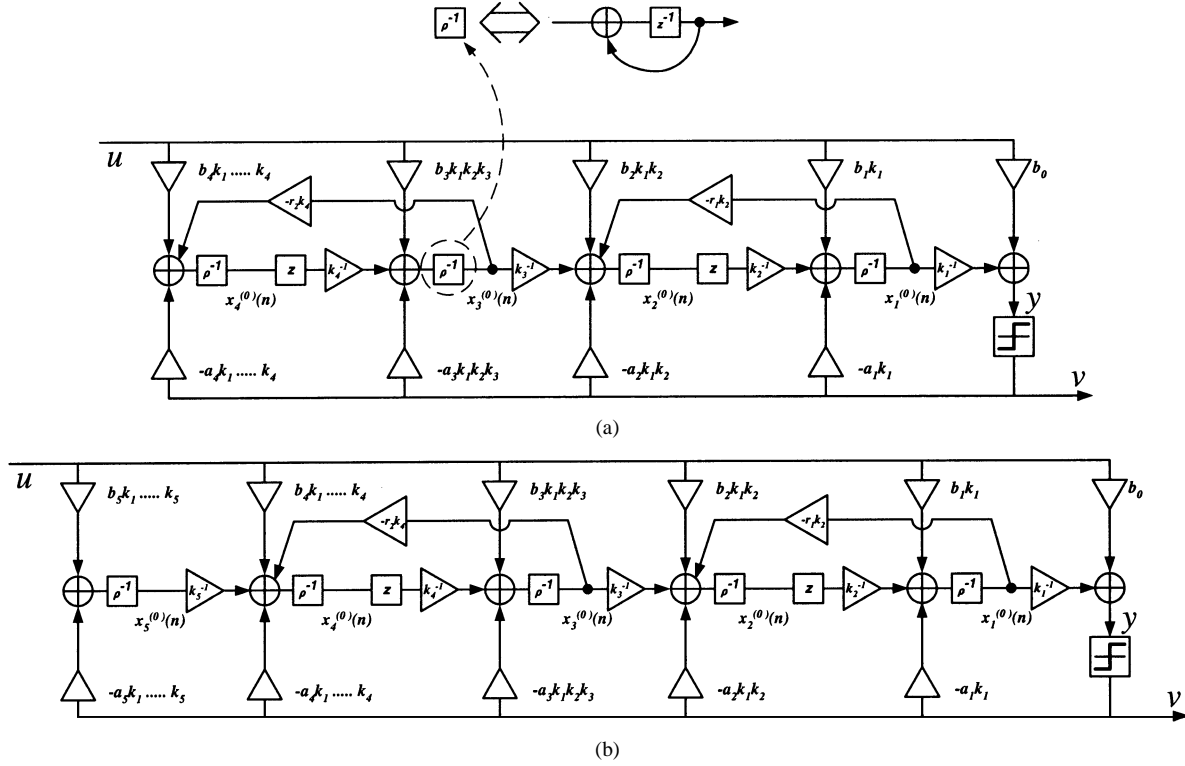


Fig. 1. Two example $\Sigma\Delta$ modulators in parameterized cascade-of-resonators structures: (a) fourth-order, (b) fifth-order.

where

$$\begin{aligned} \mathbf{x}^{(1)} &= \mathbf{T}_d^{-1} \mathbf{x}^{(0)}, \quad \mathbf{b}_z^{(1)} = \mathbf{T}_d^{-1} (\mathbf{I} + \mathbf{A}_2) \mathbf{b}_z^{(0)} \\ \mathbf{a}_z^{(1)} &= \mathbf{T}_d^{-1} (\mathbf{I} + \mathbf{A}_2) \mathbf{a}_z^{(0)}, \quad \mathbf{d}_z^{(1)} = \mathbf{d}_z^{(0)} \mathbf{T}_d \\ \mathbf{A}_z^{(1)} &= \mathbf{T}_d^{-1} (\mathbf{I} + \mathbf{A}_2) (\mathbf{I} + \mathbf{A}_1) \mathbf{T}_d \\ &= \begin{bmatrix} \mathbf{e}_1 & \mathbf{0} & \mathbf{0} \\ & \ddots & \vdots \\ \mathbf{0} & \mathbf{e}_{N_2} & \mathbf{0} \\ 0 & \dots & 0 & 1 \end{bmatrix} \\ \mathbf{e}_i &= \begin{bmatrix} e^{j\omega_i} & 0 \\ 0 & e^{-j\omega_i} \end{bmatrix}, \quad N_2 = \left\lfloor \frac{N}{2} \right\rfloor. \end{aligned} \quad (5)$$

Here, N_2 is the number of second-order subsystems. The bolded $\mathbf{0}$'s stand for zeros apart from the diagonal. The ω_i 's (by convention $\omega_i > 0$) correspond to the distinct NTF zero frequencies. It can be verified that the elements in other state matrices are also complex conjugate pairs. To visualize the dynamics (graphically) it is desirable to work with real state variables. This is achieved by a similarity transform that separates the real and imaginary parts of the conjugate elements, namely

$$\mathbf{x}^{(1)} = \mathbf{T}_{\Re} \mathbf{x}^{(2)}, \quad \mathbf{T}_{\Re} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{0} & \mathbf{0} \\ & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{\Omega} \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

$$\mathbf{\Omega} = \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}.$$

Such transform produces

$$\begin{cases} z\mathbf{x}^{(2)} = \mathbf{A}_z^{(2)} \mathbf{x}^{(2)} + \mathbf{b}_z^{(2)} u + \mathbf{a}_z^{(2)} v \\ y = b_0 u + \mathbf{d}_z^{(2)} \mathbf{x}^{(2)} \end{cases} \quad (7)$$

with

$$\begin{aligned} \mathbf{x}^{(2)} &= \mathbf{T}_{\Re}^{-1} \mathbf{x}^{(1)}, \quad \mathbf{b}_z^{(2)} = \mathbf{T}_{\Re}^{-1} \mathbf{b}_z^{(1)} \\ \mathbf{a}_z^{(2)} &= \mathbf{T}_{\Re}^{-1} \mathbf{a}_z^{(1)}, \quad \mathbf{d}_z^{(2)} = \mathbf{d}_z^{(1)} \mathbf{T}_{\Re} \end{aligned}$$

$$\begin{aligned} \mathbf{A}_z^{(2)} &= \mathbf{T}_{\Re}^{-1} \mathbf{A}_z^{(1)} \mathbf{T}_{\Re} = \begin{bmatrix} \Theta_1 & & \mathbf{0} & 0 \\ & \ddots & & \vdots \\ \mathbf{0} & & \Theta_{N_2} & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \\ \Theta_i &= \begin{bmatrix} \cos \omega_i & -\sin \omega_i \\ \sin \omega_i & \cos \omega_i \end{bmatrix}. \end{aligned} \quad (8)$$

The elements in all state matrices are made real by this transformation. Now the state-space representation is transformed into parallel second-order subsystems plus an additional first-order subsystem for odd-order modulators. In this block-diagonal form, the state variables of every subsystem become almost decoupled, interacting only through the quantizer function $v = \text{sgn}(y)$. By detailing the second equation in (7)

$$\begin{aligned} y &= b_0 u + [d_1 \dots d_{2N_2} \mid d_{2N_2+1}] \\ &\quad \times [x_1^{(2)} \dots x_{2N_2}^{(2)} \mid x_{2N_2+1}^{(2)}]^T \\ &= b_0 u + [d_1 \quad d_2] [x_1^{(2)} \quad x_2^{(2)}]^T + \dots \\ &\quad + [d_{2N_2-1} \quad d_{2N_2}] [x_{2N_2-1}^{(2)} \quad x_{2N_2}^{(2)}]^T + d_{2N_2+1} x_{2N_2+1}^{(2)}. \end{aligned} \quad (9)$$

The second line of (9) is a sum of dot products which represent the projections of the state variables of every second-order subsystem onto the vectors formed by every element pair in $\mathbf{d}_z^{(2)}$, namely $[d_{2i-1} \quad d_{2i}]^T$, $i = 1, 2, \dots, N_2$. A further transformation and simplification is to rotate and scale the coordinate system of each second-order subsystem so that these vectors

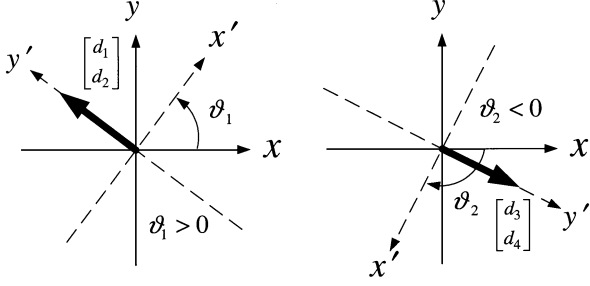


Fig. 2. Normalization process by rotating to the new coordinate system x', y' .

become the unit vectors in the conventional upward sense (see Fig. 2), and to scale d_{2N_2+1} to unity for odd-order modulators. This process, denoted as the *normalization* process, is done by the normalization matrix \mathbf{T}_n defined as

$$\mathbf{T}_n = \begin{bmatrix} \mathbf{n}_1 & \mathbf{0} & 0 \\ & \ddots & \vdots \\ \mathbf{0} & \mathbf{n}_{N_2} & 0 \\ 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ d_{2N_2+1}^{-1} \end{bmatrix} \quad (10)$$

$$\mathbf{n}_i = (d_{2i-1}^2 + d_{2i}^2)^{-1/2} \begin{bmatrix} \cos \vartheta_i & -\sin \vartheta_i \\ \sin \vartheta_i & \cos \vartheta_i \end{bmatrix}$$

where ϑ_i 's are as determined in Fig. 2. An interesting property of the transformation by \mathbf{T}_n is that it has no effect on $\mathbf{A}_z^{(2)}$. This is sound because only the coordinate system of every second-order subsystem is changed but not its nature. Summarizing

$$\begin{cases} z\mathbf{x}^{(3)} = \mathbf{A}_z^{(3)}\mathbf{x}^{(3)} + \mathbf{b}_z^{(3)}u + \mathbf{a}_z^{(3)}v \\ y = b_0u + \mathbf{d}_z^{(3)}\mathbf{x}^{(3)} \end{cases} \quad (11)$$

where

$$\begin{aligned} \mathbf{x}^{(3)} &= \mathbf{T}_n^{-1}\mathbf{x}^{(2)}, & \mathbf{b}_z^{(3)} &= \mathbf{T}_n^{-1}\mathbf{b}_z^{(2)} \\ \mathbf{a}_z^{(3)} &= \mathbf{T}_n^{-1}\mathbf{a}_z^{(2)}, & \mathbf{A}_z^{(3)} &= \mathbf{T}_n^{-1}\mathbf{A}_z^{(2)}\mathbf{T}_n = \mathbf{A}_z^{(2)} \\ \mathbf{d}_z^{(3)} &= \mathbf{d}_z^{(2)}\mathbf{T}_n = [0 \ 1 \ \dots \ 0 \ 1 \ | \ 1]. \end{aligned} \quad (12)$$

Equations (1), (4), (7), and (11) all describe the same $\Sigma\Delta$ modulator because y (and, therefore, the modulator output v) is invariant. The only distinction lies in the different choices of the state vectors. The second equation in (11) corresponds to the hyperplane \mathbf{P} that divides the state space into two regions, called the positive half-plane \mathbf{PHP} and the negative half-plane \mathbf{NHP} , wherein the quantizer outputs, denoted by v , are $+1$ and -1 , respectively.

The global dynamics of a practical $\Sigma\Delta$ modulator is reviewed in Fig. 3. Employing the convention of Wang [12], mode- \wedge and mode- \vee distinguish the positive and negative states of the quantizer output. First, a practical $\Sigma\Delta$ modulator (usually implemented with switched-capacitors) is a sampled-data system and its state trajectory undergoes discrete mapping as shown in Fig. 3(a). The dynamics is affine within each half-plane and the trajectory, being solvable by linear techniques, follows certain rules until it crosses \mathbf{P} . If continuous trajectory is assumed, the quantizer output changes sign (and so do the dynamics) when the trajectory hits exactly on \mathbf{P} . However, the discrete constraint in practice will carry the trajectory somewhere beyond (and inclusive of) \mathbf{P} where it stays for one and only one time instance, known as the *transition point*, before changing dynamics. The dynamics is then governed by another set of rules in the opposite half-plane and the cycle repeats. The two shaded wedges

in Fig. 3(b), bounded by \mathbf{P} and the far end boundaries $\hat{\mathbf{P}}_1$ and $\check{\mathbf{P}}_1$, denote the possible regions for transition point occurrence. These wedges are called the positive and negative transition wedges $\hat{\mathbf{T}}$ and $\check{\mathbf{T}}$. In reality the transition may take place anywhere on $\hat{\mathbf{P}}_\phi$ and $\check{\mathbf{P}}_\eta$ (within $\hat{\mathbf{T}}$ and $\check{\mathbf{T}}$) where $\phi, \eta \in [0, 1]$ are called the $\hat{\mathbf{T}}$ and $\check{\mathbf{T}}$ *tilt factors* for obvious reasons. Both $\hat{\mathbf{P}}_\phi$ and $\check{\mathbf{P}}_\eta$ reduce to \mathbf{P} when ϕ and η approach zero.

B. Embedding and Transition Flow

Embedding [12] refers to the construction of a set of continuous differential equations whose solution, in the form of a continuous-time trajectory as in Fig. 3(c), contains every point along the discrete trajectory. Such formulation results in continuous-time functions that are easier to handle analytically. To embed an arbitrary order discrete-time system given by (11), we consider a continuous-time counterpart

$$\begin{cases} \dot{\mathbf{x}}_c = \mathbf{A}_c\mathbf{x}_c + \mathbf{b}_c u + \mathbf{a}_c v \\ y = b_0 u + \mathbf{d}_c \mathbf{x}_c \end{cases} \quad (13)$$

where the subscript c stands for continuous-time. Assuming $t \geq k$ and $\mathbf{x}_c(k)$ is known, the exact solution of the first equation (13), e.g., see [30]

$$\begin{aligned} \mathbf{x}_c(t) &= \exp(\mathbf{A}_c(t-k))\mathbf{x}_c(k) \\ &+ \int_k^t \exp(\mathbf{A}_c(t-\tau))[\mathbf{b}_c u(\tau) + \mathbf{a}_c v(\tau)]d\tau \end{aligned} \quad (14)$$

where $\exp(\circ)$ stands for matrix exponential. Suppose u and v remain constant during the time interval $k \leq t \leq k+1$, then at the next time instance

$$\begin{aligned} \mathbf{x}_c(k+1) &= \exp(\mathbf{A}_c)\mathbf{x}_c(k) + \int_k^{k+1} \exp(\mathbf{A}_c(k+1-\tau)) \\ &\times [\mathbf{b}_c u(\tau) + \mathbf{a}_c v(\tau)]d\tau \\ &= \exp(\mathbf{A}_c)\mathbf{x}_c(k) + \left[\int_0^1 \exp(\mathbf{A}_c\tau)d\tau \right] \\ &\times [\mathbf{b}_c u(k) + \mathbf{a}_c v(k)]. \end{aligned} \quad (15)$$

Mapping (15) to (11) and noting the identity $\mathbf{T}^{-1}\wedge\mathbf{T} = \ln(\mathbf{T}^{-1}\exp(\wedge)\mathbf{T})$ where $\ln(\circ)$ denotes matrix logarithm and \wedge is a diagonal matrix, the embedded, continuous-time matrices are

$$\begin{aligned} \mathbf{A}_c &= \ln(\mathbf{A}_z^{(3)}) = \begin{bmatrix} \mathbf{w}_1 & \mathbf{0} & 0 \\ & \ddots & \vdots \\ \mathbf{0} & \mathbf{w}_{N_2} & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \\ \mathbf{w}_i &= \begin{bmatrix} 0 & -\omega_i \\ \omega_i & 0 \end{bmatrix} \\ \mathbf{b}_c &= \left(\int_0^1 \exp(\mathbf{A}_c\tau)d\tau \right)^{-1} \mathbf{b}_z^{(3)} \\ \mathbf{a}_c &= \left(\int_0^1 \exp(\mathbf{A}_c\tau)d\tau \right)^{-1} \mathbf{a}_z^{(3)}, \quad \mathbf{d}_c = \mathbf{d}_z^{(3)}. \end{aligned} \quad (16)$$

According to (15), a unit time step advancement (i.e., increasing t by 1) in the continuous-time trajectory corresponds to a hop to the next state in the discrete-time trajectory.

Next, consider the extra first-order subsystem in an odd-order modulator, and assuming $\mathbf{x}_c = [x_{c1}x_{c2}\dots x_{c2N_2}|x_{c2N_2+1}]$ (the subscript c is omitted from x_{2N_2+1} for reasons that will become

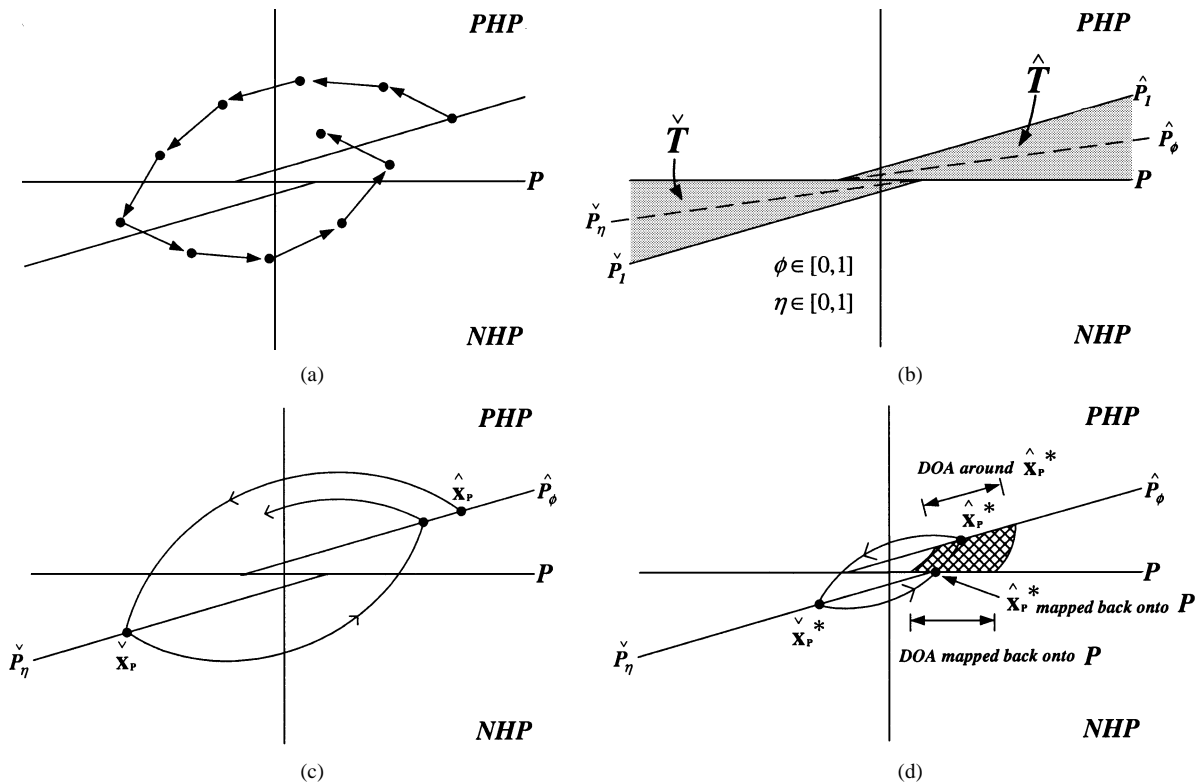


Fig. 3. Conceptual drawings of the nonlinear dynamics: (a) a discrete trajectory, (b) transition wedges and Poincaré sections, (c) continuous trajectory, and (d) limit cycle, fixed-points, and an associated domain of attraction (DOA).

clear later), $\mathbf{b}_c = [\beta_1 \ \beta_2 \ \dots \ \beta_{2N_2} \ |\beta_{2N_2+1}]^T$ and $\mathbf{a}_c = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_{2N_2} \ |\alpha_{2N_2+1}]^T$, then from (13) and (16)

$$\begin{aligned} \dot{x}_{2N_2+1} &= \beta_{2N_2+1}u + \alpha_{2N_2+1}v \\ &= \begin{cases} \gamma_+ \triangleq \beta_{2N_2+1}u + \alpha_{2N_2+1}, & \text{for } y \geq 0 \\ \gamma_- \triangleq \beta_{2N_2+1}u - \alpha_{2N_2+1}, & \text{for } y < 0. \end{cases} \quad (17) \end{aligned}$$

For properly designed modulators, γ_+ and γ_- are of opposite signs and, therefore, x_{2N_2+1} represents an “oscillating” quantity under normal operation of the odd-order modulator. For convenience, we denote a matrix or vector after dropping the terms outside the delimiters (for the case of odd-order modulators) by a tilde sign, [see (16)], $\tilde{\mathbf{A}}_c = \text{diag}[\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_{N_2}]$ is the block-diagonal matrix after omitting the last row and column of \mathbf{A}_c , and $\tilde{\mathbf{b}}_c$ is \mathbf{b}_c without β_{2N_2+1} etc. It should be stressed that this tilde sign is immaterial for even-order modulators wherein $\mathbf{A}_c = \tilde{\mathbf{A}}_c$, $\mathbf{b}_c = \tilde{\mathbf{b}}_c$ etc. As pointed out in Section I the analysis assumes a constant input u . This enables further simplification of (13) through assigning another continuous-time state-vector $\mathbf{x}(t)$ which is a displacement of $\mathbf{x}_c(t)$, namely

$$\mathbf{x} \left(= \begin{bmatrix} \tilde{\mathbf{x}} \\ x_{2N_2+1} \end{bmatrix} = [x_1 \ \dots \ x_{2N_2} \ | \ x_{2N_2+1}]^T \right) \triangleq \begin{bmatrix} \tilde{\mathbf{x}}_c \\ x_{2N_2+1} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{b}}_c u \\ 0 \end{bmatrix}. \quad (18)$$

As a result the second equation in (13) can be verified to be equal to

$$\begin{aligned} y &= b_0 u + \tilde{\mathbf{d}}_c \left(\tilde{\mathbf{x}} - \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{b}}_c u \right) + x_{2N_2+1} \\ &= (b_0 - \tilde{\mathbf{d}}_c \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{b}}_c) u + \tilde{\mathbf{d}}_c \tilde{\mathbf{x}} + x_{2N_2+1} \\ &= (b_0 + \tilde{\mathbf{d}}_c (\tilde{\mathbf{I}} - \tilde{\mathbf{A}}_c)^{-1} \tilde{\mathbf{b}}_c) u + \tilde{\mathbf{d}}_c \tilde{\mathbf{x}} + x_{2N_2+1} \\ &= K u + \tilde{\mathbf{d}}_c \tilde{\mathbf{x}} + x_{2N_2+1} \quad (19) \end{aligned}$$

where $\tilde{\mathbf{I}}$ is the identity matrix of appropriate dimension. In the third line of (19) the superscripts for the discrete system matrices are omitted because the bracketed value is invariant under similarity transforms. Referring to Fig. 1, this value, denoted by a constant K , is the forward-path resonator dc gain [ignoring the last branch in an odd-order modulator which is not part of a resonator, i.e., by putting $b_5 k_1 \dots k_5 = 0$ in Fig. 1(b)] from u to y with $v = 0$. Therefore, we end up with a much simplified system of equations as in (20), shown at the bottom of the page. It can be seen that the state equations are piecewise linear with respect to the sign of the quantizer output $v = \text{sgn}(y)$, which changes only when the trajectory $\mathbf{x}(t)$ crosses P ($P: K u + \sum_{i=1}^{N_2} x_{2i} + x_{2N_2+1} = 0$).

$$\begin{cases} \dot{\tilde{\mathbf{x}}} = \tilde{\mathbf{A}}_c \tilde{\mathbf{x}} + \tilde{\mathbf{a}}_c v \\ \dot{x}_{2N_2+1} = \begin{cases} \gamma_+ & \text{for } y \geq 0 \\ \gamma_- & \text{for } y < 0 \end{cases} \\ y = K u + \tilde{\mathbf{d}}_c \tilde{\mathbf{x}} + x_{2N_2+1} = K u + \sum_{i=1}^{N_2} x_{2i} + x_{2N_2+1} \end{cases} \quad \text{(for odd-order modulators only).} \quad (20)$$

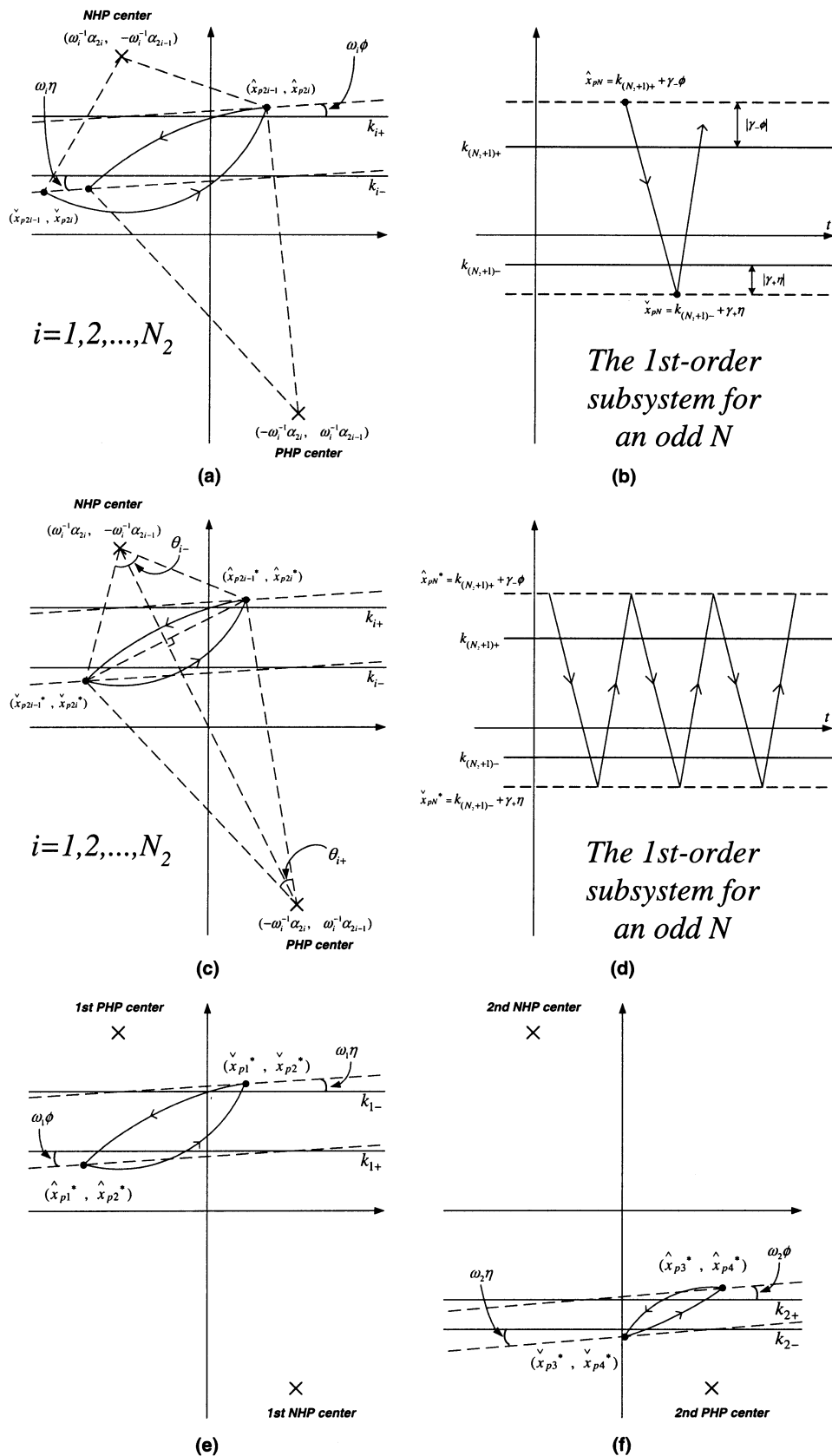


Fig. 4. Decomposed subsystems: (a) second-order; (b) first-order; (c), (d) limit cycles in subsystems; (e), (f) two interacting second-order subsystems in a fourth-order $\Sigma\Delta$ modulator.

where $\lfloor (N + 1)/2 \rfloor$ gives the total number of decomposed subsystems inclusive of second-order and first-order ones. The positive

threshold lines denote the thresholds across which a trajectory enters the **PHP** and the negative threshold lines denote the

thresholds for entering the **NHP**. After crossing these lines, a trajectory go on for ϕ and η more time step (i.e., rotate for $\omega_i\phi$ and $\omega_i\eta$ more radians in a second-order subsystem or travel for $\gamma_-\phi$ and $\gamma_+\eta$ more distance in a first-order subsystem) onto \hat{P}_ϕ and \check{P}_η , denoted as $\hat{\mathbf{x}}_p = [\hat{x}_{p1} \dots \hat{x}_{pN}]$ and $\check{\mathbf{x}}_p = [\check{x}_{p1} \dots \check{x}_{pN}]$, respectively. Strictly speaking, the parameters ϕ and η should be included in the notations of $\hat{\mathbf{x}}_p$ and $\check{\mathbf{x}}_p$ but for simplicity they are omitted and should be clear from the context. Referring to Fig. 4(a), it can be verified that

$$\hat{x}_{p2i} = (\hat{x}_{p2i-1} - \omega_i^{-1}\alpha_{2i}) \tan \omega_i\phi + \frac{k_{i+} + \omega_i^{-1}\alpha_{2i-1}}{\cos \omega_i\phi} - \omega_i^{-1}\alpha_{2i-1} \quad (25)$$

and

$$\check{x}_{p2i} = (\check{x}_{p2i-1} + \omega_i^{-1}\alpha_{2i}) \tan \omega_i\eta + \frac{k_{i-} - \omega_i^{-1}\alpha_{2i-1}}{\cos \omega_i\eta} + \omega_i^{-1}\alpha_{2i-1} \quad (26)$$

for $i = 1, 2, \dots, N_2$. And for the first-order subsystem when N is odd

$$\hat{x}_{pN} = k_{(N_2+1)+} + \gamma_-\phi \text{ and } \check{x}_{pN} = k_{(N_2+1)-} + \gamma_+\eta. \quad (27)$$

Here, another expression for $N_2 + 1$ is the floor function $\lfloor (N + 1)/2 \rfloor$.

Since the boundaries \hat{P}_ϕ and \check{P}_η inside \hat{T} and \check{T} are obtained through mapping P by ϕ and η time step forward about appropriate half-plane centers, their equations can be derived by recognizing that when states on \hat{P}_ϕ or \check{P}_η (i.e., $\hat{\mathbf{x}}_p$ and $\check{\mathbf{x}}_p$) are mapped ϕ and η time step backward, they will land on $P : Ku + \mathbf{d}_c\mathbf{x} = 0$. This yields (28) and (29) shown at the bottom of the page.

The transition flow occurring on \hat{P}_ϕ and \check{P}_η is denoted as $\mathbf{F}(\hat{P}_\phi, \check{P}_\eta)$. Wang [10], [12] has investigated the special case of the intuitively farthest stretched $\mathbf{F}(\hat{P}_1, \check{P}_1)$, called the *boundary transition flow* corresponding to $\phi = \eta = 1$. In nonlinear context the two planes, \hat{P}_ϕ and \check{P}_η , $\phi, \eta \in [0, 1]$, are chosen to be the *Poincaré sections* (see [31]). The successive mapping of states, namely, a two-step return map

$$\left(\mathcal{R} : \hat{P}_\phi \rightarrow \hat{P}_\phi \right) = \left(\mathcal{R} : \hat{P}_\phi \rightarrow \check{P}_\eta \right) \circ \left(\mathcal{R} : \check{P}_\eta \rightarrow \hat{P}_\phi \right) \quad (30)$$

is known as the first return or Poincaré map. This mapping reduces the system dimension by one because only states on the Poincaré sections are of concern. A (global) limit cycle of $\mathbf{F}(\hat{P}_\phi, \check{P}_\eta)$ exists if a continuous closed-loop trajectory can be found resulting in two fixed-points, $\hat{\mathbf{x}}_p^*$ and $\check{\mathbf{x}}_p^*$, on the two chosen Poincaré sections \hat{P}_ϕ and \check{P}_η as in Fig. 3(d).

This requires finding closed-loop trajectories in all subsystems whose state variables satisfy (24)–(27), as is shown in Fig. 4(c) and (d). The stability of these fixed-points and their associated limit cycles is then investigated through the tangent linear manifolds about them.

For an N th-order modulator, there are N_2 second-order subsystems interacting with each other wherein different subsystems have different arrangements of half-plane centers. The trajectories of all second-order subsystems will be traveling in a sense that they counteract each other. Fig. 4(e) and (f) show the two second-order subsystems resulting from a fourth-order modulator. For example, in **PHP** the trajectory of the first subsystem is traveling upward to k_{1-} while that of the second subsystem is going downward to k_{2-} till they cross the threshold lines and penetrate for η more time step into **NHP**, whereupon the dynamics changes and proceeds by the rules of **NHP** accordingly. For an odd-order modulator, there are $N_2 + 1$ subsystems with the last one being a first-order subsystem that interacts with the other second-order subsystems in a similar manner. Fig. 5 shows a practical, discrete-time fifth-order modulator under stable and unstable operations. The state variables are first transformed into and viewed by the subsystem framework. In an unstable modulator, the integrator outputs all evolve toward infinity but in Fig. 5(e) and (f) one subsystem remains bounded. This is so because the state variables are in the decoupled, transformed state space. When the state variables are transformed back to the integrator outputs (coupled subsystems), all of them will go to infinity as is observed in real cases. A major difference and difficulty in practical $\Sigma\Delta$ modulators is that the trajectory is discrete and never falls exactly onto any continuous-time limit cycle, making the exact transitions intractable. Wang's work on third-order $\Sigma\Delta$ modulators argues that the discrete trajectories are close to the neighborhood and mostly within the space encompassed by the intuitively farthest stretched $\mathbf{F}(\hat{P}_1, \check{P}_1)$ limit cycle. The modulator is predicted to be stable when $\mathbf{F}(\hat{P}_1, \check{P}_1)$ is stable. Excellent match with experimental results are observed for the third-order examples.

Nonetheless, uncertainty remains because there is no proof as to the existence of stable fixed-points and $\mathbf{F}(\hat{P}_1, \check{P}_1)$ limit cycle will guarantee stable discrete trajectories. In real modulators, the transition may take place anywhere within \hat{T} and \check{T} with intractable values of ϕ and η instead of exactly on \hat{P}_1 and \check{P}_1 . Simulation of some $\Sigma\Delta$ modulators (order > 2) shows that though the stable limit cycles produced by $\mathbf{F}(\hat{P}_1, \check{P}_1)$, if they

$$\begin{aligned} & \hat{P}_\phi : Ku + \tilde{\mathbf{d}}_c \left(\mathbf{M}_t^{-\phi} \left(\hat{\mathbf{x}}_p - \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{a}}_c \right) + \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{a}}_c \right) \Big| + \hat{x}_{pN} - \gamma_-\phi = 0 \\ \text{or } \hat{P}_\phi : & \left(\sum_{i=1}^{N_2} - (\hat{x}_{p2i-1} - \omega_i^{-1}\alpha_{2i}) \sin \omega_i\phi + (\hat{x}_{p2i} + \omega_i^{-1}\alpha_{2i-1}) \cos \omega_i\phi - \omega_i^{-1}\alpha_{2i-1} \right) \Big| + \hat{x}_{pN} - \gamma_-\phi = -Ku \quad (28) \end{aligned}$$

$$\begin{aligned} & \check{P}_\eta : Ku + \tilde{\mathbf{d}}_c \left(\mathbf{M}_t^{-\eta} \left(\check{\mathbf{x}}_p + \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{a}}_c \right) - \tilde{\mathbf{A}}_c^{-1} \tilde{\mathbf{a}}_c \right) \Big| + \check{x}_{pN} - \gamma_+\eta = 0 \\ \text{or } \check{P}_\eta : & \left(\sum_{i=1}^{N_2} - (\check{x}_{p2i-1} + \omega_i^{-1}\alpha_{2i}) \sin \omega_i\eta + (\check{x}_{p2i} - \omega_i^{-1}\alpha_{2i-1}) \cos \omega_i\eta + \omega_i^{-1}\alpha_{2i-1} \right) \Big| + \check{x}_{pN} - \gamma_+\eta = -Ku. \quad (29) \end{aligned}$$

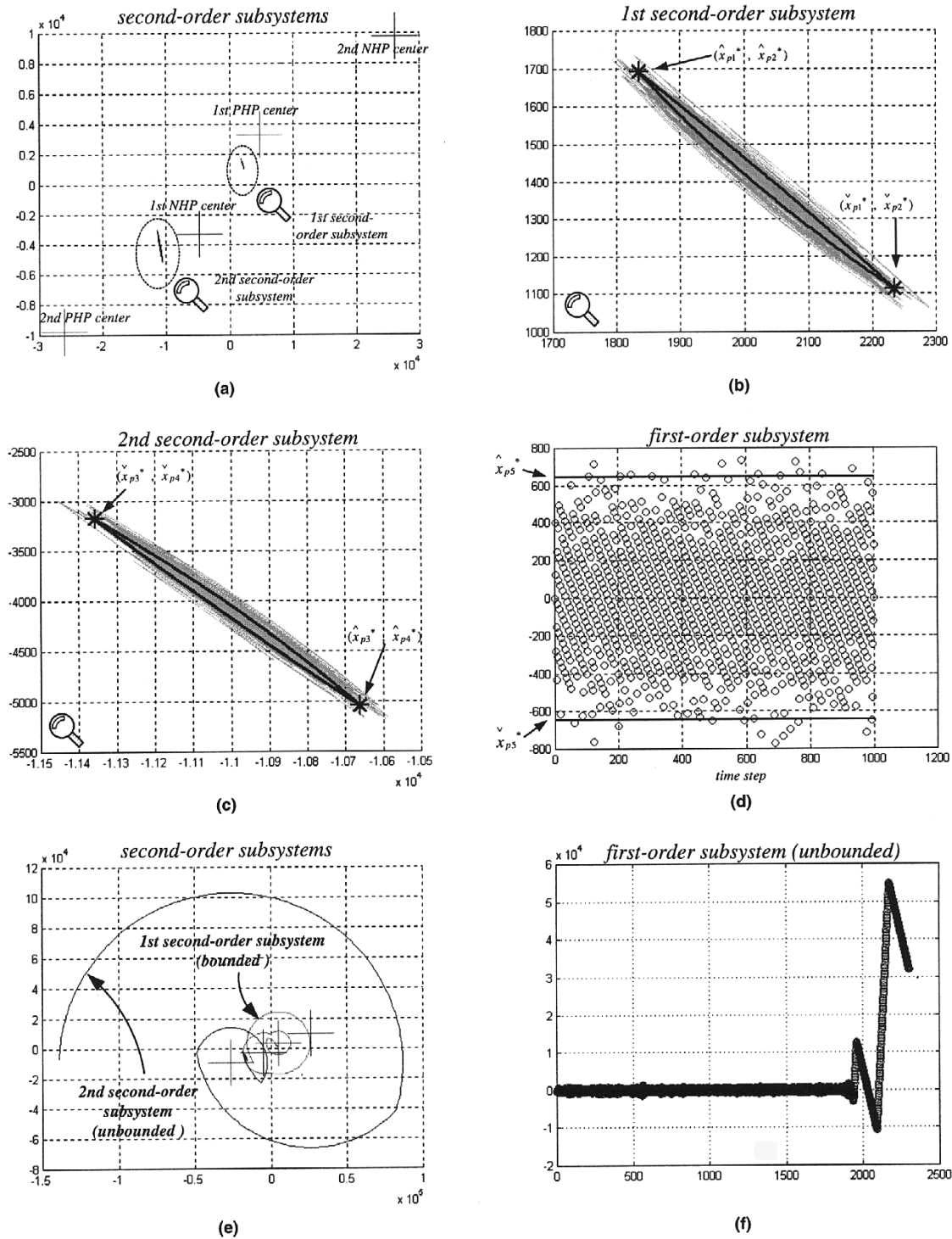


Fig. 5. Simulation of a discrete-time fifth-order $\Sigma\Delta$ modulator (OSR = 64, maximum NTF gain = 1.5, optimized NTF zeros): (a)–(d) discrete trajectories under stable operation ($u = 0.42$, zero-state initial $\mathbf{x}^{(0)}$, 1000 runs) where bold lines represent the stable $\mathbf{F}(\hat{\mathbf{P}}_1, \hat{\mathbf{P}}_1)$ limit cycle; (e), (f) discrete trajectories leading to unstable operation ($u = 0.63$, zero-state initial $\mathbf{x}^{(0)}$, 2300 runs).

ever exist, generally encircle the discrete trajectories, there are frequent occasions that the discrete trajectories go beyond this encirclement [see Fig. 5(b)–(d)]. Counterexamples are easily found where the discrete trajectories eventually go unbounded despite the existence of stable $\mathbf{F}(\hat{\mathbf{P}}_1, \hat{\mathbf{P}}_1)$. It is this intricate nature that renders the stability analysis of $\Sigma\Delta$ modulators highly difficult, especially for high-order modulators since graphical interpretation would become much more complicated. In the

following sections, stability of the general case $\mathbf{F}(\hat{\mathbf{P}}_\phi, \hat{\mathbf{P}}_\eta)$, $\phi, \eta \in [0, 1]$ is investigated. It is shown that in practical modulators, discrete trajectories going beyond the stable limit cycle, if any, of $\mathbf{F}(\hat{\mathbf{P}}_1, \hat{\mathbf{P}}_1)$ is a natural case. And that the existence of stable fixed-points for all types of transition flow serves only as a necessary condition for stable dc operation. Sufficient condition for dc stability and the procedures for finding the stable dc input bounds are presented in Section VI.

IV. FIXED-POINT LOCATION BY NUMERICAL ANALYSIS

Locating the fixed-points and their associated limit cycles under different transition flow assumptions is a critical step toward stability analysis of $\Sigma\Delta$ modulators. The limit cycles approximate the locus of the discrete trajectories exhibited by practical modulators. An unstable (repelling) limit cycle implies unstable operation of the modulator. The inverse is not necessarily true as discrete trajectories cannot fall exactly onto the stable limit cycles and additional criteria for dc stability are needed. The nonlinearity introduced by the quantizer function makes it difficult to solve for the fixed-points algebraically so numerical methods are utilized. The fixed-points are found by recognizing that the positive and negative *fly-times* (which are the time that a trajectory spends in the *PHP* and *NHP*, respectively) of all subsystems, when coupled together, must be equal.

First, for the limit cycle in a second-order subsystem, it turns out that the two threshold lines completely characterize the fixed-points and subsequently the fly-times. Given k_{i+} and k_{i-} , the fixed-points of $\mathbf{F}(\hat{\mathbf{P}}_\phi, \check{\mathbf{P}}_\eta)$ (denoted by asterisks in Figs. 4 and 5) for each second-order subsystem, with the help of (25) and (26) and the fact that the fixed-points and the half-plane centers must form into a symmetrical kite shape [Fig. 4(c)], can be solved by two simultaneous equations as in (31) shown at the bottom of the page, for $i = 1, 2, \dots, N_2$, (note that the parameters ω_i , α_{2i-1} , α_{2i} , ϕ and η are known parameters for a second-order subsystem). Hence the positive and negative fly-times [p_time and n_time in (32)] can be obtained through dividing the angles sustained about each half-plane center [θ_{i+} and θ_{i-} in Fig. 4(c)] by the corresponding ω_i , see (32) at the bottom of the page, where $i = 1, 2, \dots, N_2$ and $\arg(\circ)$ is the angle function of a complex quantity. Combining (31) and (32), it is obvious that the fly-time functions can be expressed as nonlinear functions of the threshold line values. For the first-order subsystem in an odd-order modulator, the same

dependence on threshold line values also applies. From (27), its fixed-points are

$$\hat{x}_{pN}^* = k_{(N_2+1)+} + \gamma_- \phi \text{ and } \check{x}_{pN}^* = k_{(N_2+1)-} + \gamma_+ \eta \quad (33)$$

and the fly-times

$$\begin{cases} p_time_{N_2+1}(\hat{x}_{pN}^*, \check{x}_{pN}^*) = \frac{(\hat{x}_{pN}^* - \check{x}_{pN}^*)}{\gamma_+} \\ n_time_{N_2+1}(\hat{x}_{pN}^*, \check{x}_{pN}^*) = \frac{(\hat{x}_{pN}^* - \check{x}_{pN}^*)}{\gamma_-} \end{cases} \quad (34)$$

Similarly, combining (33) and (34), it is straightforward to see that the fly-time functions can be represented in terms of the threshold line values $k_{(N_2+1)+}$ and $k_{(N_2+1)-}$.

In general, an N th-order modulator can be decomposed into $\lfloor (N+1)/2 \rfloor$ subsystems. Finding the fixed-points or global limit cycle when these individual subsystems are coupled is equivalent to solving for the following set of $2(\lfloor (N+1)/2 \rfloor - 1)$ nonlinear equations in terms of the threshold line values k_{i+} , k_{i-} ($i = 1, 2, \dots, \lfloor (N+1)/2 \rfloor$) (the unknowns) such that all subsystems give the same fly-times: see (35) shown at the bottom of the page. In order words, the objective is to find the vector $\mathbf{k} = [k_{i+} \ k_{i-} \ \dots \ k_{\lfloor (N+1)/2 \rfloor +} \ k_{\lfloor (N+1)/2 \rfloor -}]^T$ that gives $\mathbf{G} = \mathbf{0}$. But from (24), the last two elements in \mathbf{k} are functions of the other elements, namely

$$\begin{cases} k_{\lfloor (N+1)/2 \rfloor +} = -Ku - \sum_{i=1}^{\lfloor (N+1)/2 \rfloor - 1} k_{i+} \\ k_{\lfloor (N+1)/2 \rfloor -} = -Ku - \sum_{i=1}^{\lfloor (N+1)/2 \rfloor - 1} k_{i-} \end{cases} \quad (36)$$

Substituting them back into (35), there are eventually $2(\lfloor (N+1)/2 \rfloor - 1)$ equations with the same number of unknowns that can be solved by standard numerical methods.

The Broyden's method (a quasi-Newton method, Ch. 10, [28]) is chosen for this purpose due to its reduced computational complexity and superlinear convergence. The initial guess of solution to kick off the Broyden's algorithm can be easily estimated from simulation. From the solution of \mathbf{k} (for a global limit cycle), parameters like subsystem fixed-points and fly-times are readily obtainable by back-substituting the

$$\begin{cases} \left(\tan \omega_i \phi + \frac{\alpha_{2i-1}}{\alpha_{2i}} \right) \hat{x}_{p2i-1}^* + \left(\tan \omega_i \eta + \frac{\alpha_{2i-1}}{\alpha_{2i}} \right) \check{x}_{p2i-1}^* \\ = \omega_i^{-1} \alpha_{2i} (\tan \omega_i \phi - \tan \omega_i \eta) - \frac{k_{i+} + \omega_i^{-1} \alpha_{2i-1}}{\cos \omega_i \phi} - \frac{k_{i-} - \omega_i^{-1} \alpha_{2i-1}}{\cos \omega_i \eta} \\ \left(\tan \omega_i \phi - \frac{\alpha_{2i}}{\alpha_{2i-1}} \right) \hat{x}_{p2i-1}^* - \left(\tan \omega_i \eta - \frac{\alpha_{2i}}{\alpha_{2i-1}} \right) \check{x}_{p2i-1}^* \\ = \omega_i^{-1} \alpha_{2i} (\tan \omega_i \phi + \tan \omega_i \eta) - \frac{k_{i+} + \omega_i^{-1} \alpha_{2i-1}}{\cos \omega_i \phi} + \frac{k_{i-} - \omega_i^{-1} \alpha_{2i-1}}{\cos \omega_i \eta} + 2\omega_i^{-1} \alpha_{2i-1} \end{cases} \quad (31)$$

$$\begin{cases} p_time_i(\hat{x}_{p2i-1}^*, \hat{x}_{p2i}^*, \check{x}_{p2i-1}^*, \check{x}_{p2i}^*) = \frac{1}{\omega_i} \arg \left(\frac{(\hat{x}_{p2i}^* + \omega_i^{-1} \alpha_{2i}) + j(\check{x}_{p2i}^* - \omega_i^{-1} \alpha_{2i-1})}{(\hat{x}_{p2i-1}^* + \omega_i^{-1} \alpha_{2i}) + j(\check{x}_{p2i}^* - \omega_i^{-1} \alpha_{2i-1})} \right) \\ n_time_i(\hat{x}_{p2i-1}^*, \hat{x}_{p2i}^*, \check{x}_{p2i-1}^*, \check{x}_{p2i}^*) = \frac{1}{\omega_i} \arg \left(\frac{(\hat{x}_{p2i-1}^* - \omega_i^{-1} \alpha_{2i}) + j(\hat{x}_{p2i}^* + \omega_i^{-1} \alpha_{2i-1})}{(\check{x}_{p2i-1}^* - \omega_i^{-1} \alpha_{2i}) + j(\check{x}_{p2i}^* + \omega_i^{-1} \alpha_{2i-1})} \right) \end{cases} \quad (32)$$

$$\mathbf{G} = [g_1 \ g_2 \ \dots \ g_{2(\lfloor (N+1)/2 \rfloor - 1)} \ g_{2(\lfloor (N+1)/2 \rfloor - 1)}]^T \text{ where} \begin{cases} g_{2i-1} = p_time_1(k_{i+}, k_{i-}) - p_time_{i+1}(k_{(i+1)+}, k_{(i+1)-}) \\ g_{2i} = n_time_1(k_{i+}, k_{i-}) - n_time_{i+1}(k_{(i+1)+}, k_{(i+1)-}), \quad i = 1, 2, \dots, \left\lfloor \frac{(N+1)}{2} \right\rfloor - 1. \end{cases} \quad (35)$$

threshold line values into (31)–(34). Fig. 6(a)–(e) present these numerical results for an example fourth-order modulator against all types of transition flow. Note that by adjusting the initial condition in the Broyden’s algorithm, two sets of fixed-points can be found. Borrowing results from Section V, it is shown that the “inner” set stands for stable fixed-points while the “outer” set represents unstable fixed-points (also, respectively, called foci and saddles in second-order subsystems, [32]). For the first-order subsystem, the fixed-points (in only one dimension) and threshold line values are structured similarly and is not shown here to limit the length of this paper. Loosely speaking, it can be seen that there exists a region, constituting the DOA, between the inner stable limit cycles and the outer unstable limit cycles wherein trajectories, if started with an appropriate initial condition, will be “repelled” back to the stable limit cycles. Trajectories going beyond the unstable fixed-points will swirl away and result in instability (unbounded) or even chaotic behavior (bounded). It can also be seen from Fig. 6(a) and (b) that $\mathbf{F}(\hat{\mathbf{P}}_1, \check{\mathbf{P}}_1)$ has the “least-forgiving” stable limit cycle showing the smallest attractive region.

V. STABILITY VIA TANGENT LINEAR MANIFOLD

A standard technique of evaluating the stability of the fixed-points of $\mathbf{F}(\hat{\mathbf{P}}_\phi, \check{\mathbf{P}}_\eta)$ is to investigate the tangent linear manifolds (lying on $\hat{\mathbf{P}}_\phi$ and $\check{\mathbf{P}}_\eta$) about them. Such approach requires the availability of closed-form, autonomous (time-independent) algebraic functions for the trajectory evolution within each half-plane. Previous expressions derived for high-order $\Sigma\Delta$ modulators are very complicated and are difficult, if not impossible, to formulate into algebraic functions. Even the expressions for third-order modulators contain exceedingly complicated parameters [12]. Nonetheless, the approach introduced in Sections II–IV offers a comprehensible, scalable solution to any modulator order.

Realizing that the Poincaré map of $\mathbf{F}(\hat{\mathbf{P}}_\phi, \check{\mathbf{P}}_\eta)$, denoted by $\mathfrak{R} : \hat{\mathbf{x}}_p \rightarrow \hat{\mathbf{x}}'_p$, is a two-step return map ($\mathfrak{R} : \hat{\mathbf{x}}_p \rightarrow \check{\mathbf{x}}_p$) \circ ($\mathfrak{R} : \check{\mathbf{x}}_p \rightarrow \hat{\mathbf{x}}'_p$), two sets of expressions governing the trajectories in *PHP* and *NHP* are required. Starting with *PHP*, and applying the interpretation of a high-order system as coupled second-order subsystems and an additional first-order sub-

system for odd-order modulators, the following functions for ($\mathfrak{R} : \hat{\mathbf{x}}_p \rightarrow \check{\mathbf{x}}_p$) are defined. The first set of functions is related to the circularity of the trajectories in second-order subsystems,

$$\hat{f}_i(\hat{\mathbf{x}}_p, \check{\mathbf{x}}_p) = (\hat{x}_{p2i-1} + \omega_i^{-1}\alpha_{2i})^2 + (\hat{x}_{p2i} - \omega_i^{-1}\alpha_{2i-1})^2 - (\check{x}_{p2i-1} + \omega_i^{-1}\alpha_{2i})^2 - (\check{x}_{p2i} - \omega_i^{-1}\alpha_{2i-1})^2 \quad (37)$$

for $i = 1, 2, \dots, N_2$. The second set of functions describes the fly-time difference of every two adjacent subsystems

$$\hat{f}_{N_2+i}(\hat{\mathbf{x}}_p, \check{\mathbf{x}}_p) = p\text{-time}_i(\hat{\mathbf{x}}_p, \check{\mathbf{x}}_p) - p\text{-time}_{i+1}(\hat{\mathbf{x}}_p, \check{\mathbf{x}}_p) \quad (38)$$

for $i = 1, 2, \dots, N_2 - 1$ or N_2 for odd-order modulators. Finally there are two equations quantifying the “exactness” of $\hat{\mathbf{x}}_p$ and $\check{\mathbf{x}}_p$ on $\hat{\mathbf{P}}_\phi$ and $\check{\mathbf{P}}_\eta$, namely, (39) and (40) shown at the bottom of the page. Both (39) and (40) become zero when $\hat{\mathbf{x}}_p$ and $\check{\mathbf{x}}_p$ land exactly on $\hat{\mathbf{P}}_\phi$ and $\check{\mathbf{P}}_\eta$. Clearly, about the fixed points $\hat{\mathbf{x}}_p^*$ and $\check{\mathbf{x}}_p^*$, (37)–(40) all go to zero, i.e.

$$\hat{f}_i(\hat{\mathbf{x}}_p^*, \check{\mathbf{x}}_p^*) = 0 \text{ for } i = 1, 2, \dots, N. \quad (41)$$

As mentioned, the Poincaré section analysis reduces the system dimension by one. By setting (39) and (40) to zero it is possible to derive the differentials of $d\hat{x}_{pN}$ and $d\check{x}_{pN}$ in terms of $d\hat{\mathbf{x}}_{p-1} = [d\hat{x}_{p1} \ d\hat{x}_{p2} \ \dots \ d\hat{x}_{pN-1}]^T$ and $d\check{\mathbf{x}}_{p-1} = [d\check{x}_{p1} \ d\check{x}_{p2} \ \dots \ d\check{x}_{pN-1}]^T$, respectively, such that for an even N , see (42) shown at the bottom of the page, and whereas for an odd N , (43) see shown at the bottom of the page. Next, by partially differentiating the rest of the equations in (41) with respect to $\hat{\mathbf{x}}_p$ and $\check{\mathbf{x}}_p$, and substituting back $d\hat{\mathbf{x}}_{pN}$ and $d\check{\mathbf{x}}_{pN}$ by the results in (42) or (43), we arrive at

$$\mathbf{M}_1 d\hat{\mathbf{x}}_{p-1} = \mathbf{M}_2 d\check{\mathbf{x}}_{p-1} \quad (44)$$

where the (i, j) th element of the $(N-1) \times (N-1)$ matrix \mathbf{M}_1 is $\partial \hat{f}_i / \partial \hat{x}_{pj}$ and similarly that of \mathbf{M}_2 is $-\partial \hat{f}_i / \partial \hat{x}_{pj}$. The elimination of \hat{x}_{pN} and \check{x}_{pN} is possible because they are functions of the other state variables on the Poincaré sections.

Likewise, another set of equations in *NHP*, which is the counterpart of (41), can be obtained by swapping the symbols \wedge and \vee , reversing the half-plane centers in (37) and changing into n -time functions in (38). Subsequently, defining $d\hat{\mathbf{x}}'_{p-1} = [d\hat{x}'_{p1} \ d\hat{x}'_{p2} \ \dots \ d\hat{x}'_{pN-1}]^T$, for ($\mathfrak{R} : \check{\mathbf{x}}_p \rightarrow \hat{\mathbf{x}}'_p$) we have

$$\mathbf{M}_3 d\check{\mathbf{x}}_{p-1} = \mathbf{M}_4 d\hat{\mathbf{x}}'_{p-1} \quad (45)$$

$$\hat{f}_N(\hat{\mathbf{x}}_p) = Ku + \left(\sum_{i=1}^{N_2} -(\hat{x}_{p2i-1} - \omega_i^{-1}\alpha_{2i}) \sin \omega_i \phi + (\hat{x}_{p2i} + \omega_i^{-1}\alpha_{2i-1}) \cos \omega_i \phi - \omega_i^{-1}\alpha_{2i-1} \right) \Big| + \hat{x}_{pN} - \gamma - \phi. \quad (39)$$

$$\check{f}_N(\check{\mathbf{x}}_p) = Ku + \left(\sum_{i=1}^{N_2} -(\check{x}_{p2i-1} + \omega_i^{-1}\alpha_{2i}) \sin \omega_i \eta + (\check{x}_{p2i} - \omega_i^{-1}\alpha_{2i-1}) \cos \omega_i \eta + \omega_i^{-1}\alpha_{2i-1} \right) \Big| + \check{x}_{pN} - \gamma + \eta. \quad (40)$$

$$\begin{cases} d\hat{x}_{pN} = \frac{1}{\cos \omega_{N_2} \phi} [\sin \omega_1 \phi & -\cos \omega_1 \phi & \dots & \sin \omega_{N_2-1} \phi & -\cos \omega_{N_2-1} \phi & \sin \omega_{N_2} \phi] d\hat{\mathbf{x}}_{p-1} \\ d\check{x}_{pN} = \frac{1}{\cos \omega_{N_2} \eta} [\sin \omega_1 \eta & -\cos \omega_1 \eta & \dots & \sin \omega_{N_2-1} \eta & -\cos \omega_{N_2-1} \eta & \sin \omega_{N_2} \eta] d\check{\mathbf{x}}_{p-1}. \end{cases} \quad (42)$$

$$\begin{cases} d\hat{x}_{pN} = [\sin \omega_1 \phi & -\cos \omega_1 \phi & \dots & \sin \omega_{N_2} \phi & -\cos \omega_{N_2} \phi] d\hat{\mathbf{x}}_{p-1} \\ d\check{x}_{pN} = [\sin \omega_1 \eta & -\cos \omega_1 \eta & \dots & \sin \omega_{N_2} \eta & -\cos \omega_{N_2} \eta] d\check{\mathbf{x}}_{p-1}. \end{cases} \quad (43)$$

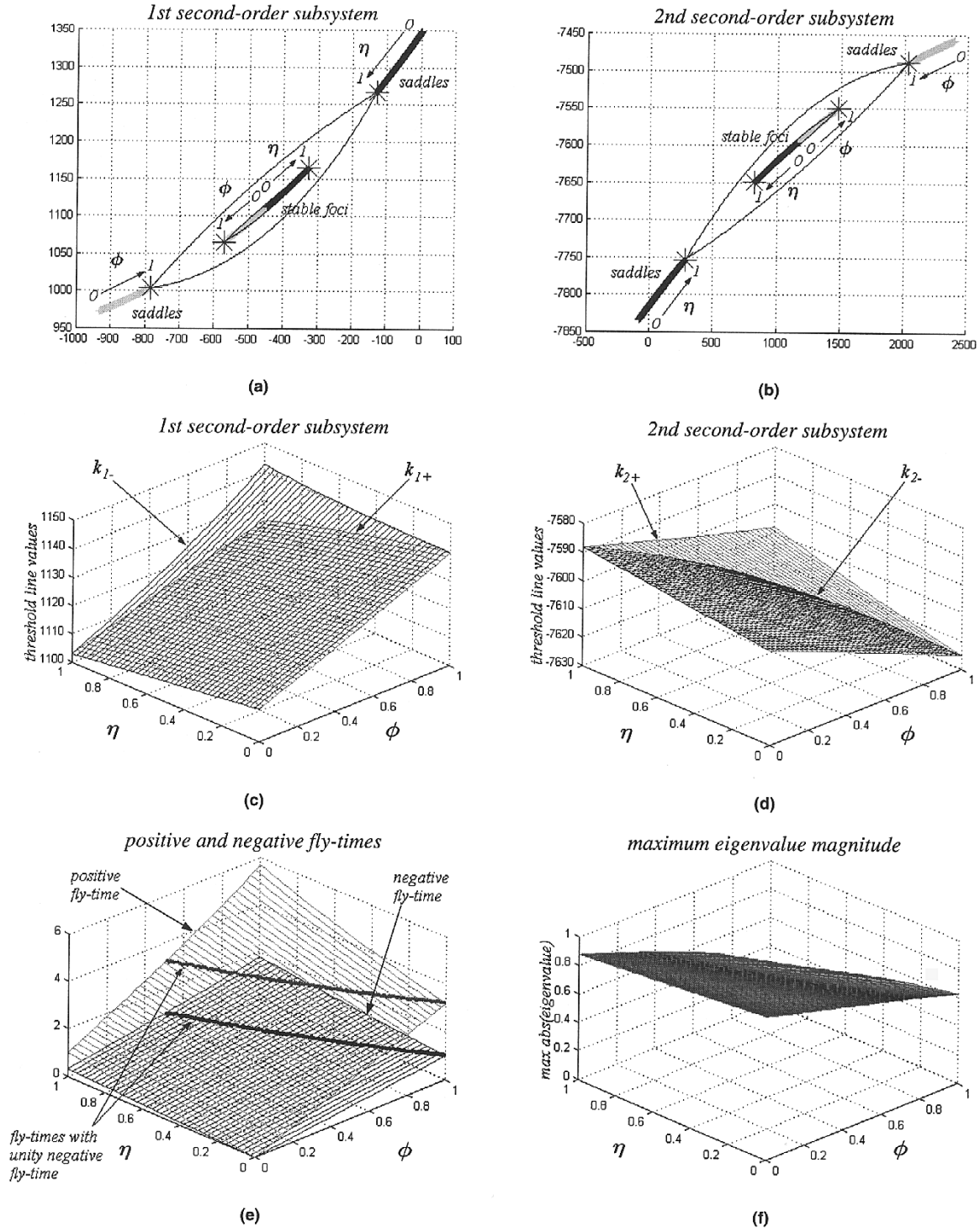


Fig. 6. Example embedded fourth-order $\Sigma\Delta$ modulator ($u = 0.523$, OSR = 64, maximum NTF gain = 1.5, optimized NTF zeros): (a),(b) fixed-points of different types of transition flow; (c),(d) threshold line values corresponding to stable fixed-points; (e) Positive and negative fly-times for stable limit cycles; (f) maximum eigenvalue magnitude for the tangent linear manifolds of stable fixed-points.

where the (i, j) th element of the $(N - 1) \times (N - 1)$ matrix \mathbf{M}_3 is $\partial \hat{f}_i / \partial \hat{x}_{pj}$ and similarly that of \mathbf{M}_4 is $-\partial \hat{f}_i / \partial \hat{x}_{pj}$. Now combining (44) and (45) and evaluating the matrices at the fixed-points

$$d\hat{\mathbf{x}}'_{p-1} = \mathbf{M}_4^{-1} \mathbf{M}_3 \mathbf{M}_2^{-1} \mathbf{M}_1 \Big|_{(\hat{x}_p^*, \hat{x}_p^*)} d\hat{\mathbf{x}}_{p-1}. \quad (46)$$

For compactness, we denote $\mathbf{R} = \mathbf{M}_4^{-1} \mathbf{M}_3 \mathbf{M}_2^{-1} \mathbf{M}_1 \Big|_{(\hat{x}_p^*, \hat{x}_p^*)}$. Stability of the fixed-points can then be assured if all the eigen-

values of \mathbf{R} have magnitude less than unity ([12], [32]). Clearly the formulations introduced in this section are scalable to systems of any order. Fig. 6(f) shows the maximum eigenvalue magnitude of \mathbf{R} against different tilt factors for the set of stable fixed-points in Fig. 6(a) and (b). Also, it can be easily deduced that for $n = 0$, n_time is equal to ϕ while that for $\phi = 0$, p_time is equal to η [see Fig. 6(e)]. Accordingly, for simple rational inputs such as $1/3$, $1/2$, $3/5$, $2/3$, etc., stable limit cycles (even in

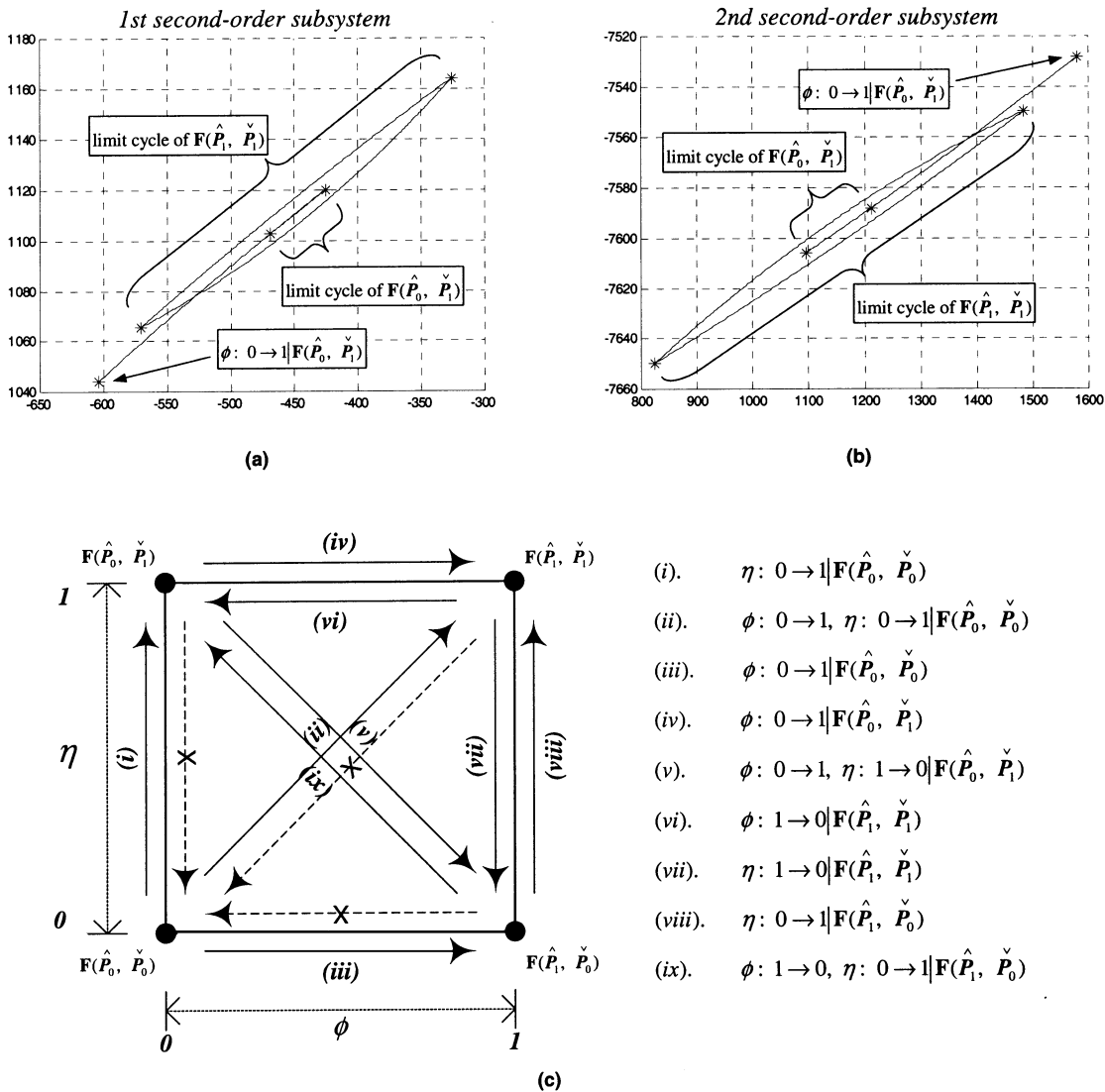


Fig. 7. (a),(b) Illustration of the trajectory change of $\phi: 0 \rightarrow 1 | F(\hat{P}_0, \check{P}_1)$ for the fourth-order modulator in Fig. 6, (c) the nine tests for the limit cycle switching.

some cases when the boundary transition flow is unstable) with integer fly-times can be found at $(\phi, \eta) = (1, 0)$ where n_{time} is one. These limit cycles represent theoretically stable limit cycles for the real, discrete-time $\Sigma\Delta$ modulators, but due to noise and component tolerance they are impossible to maintain in reality and quickly diverge to other trajectories.

In summary, for an N th-order system, we have N_2 equations describing the circular trajectories of second-order subsystem(s) within a half-plane. There are $N_2 - 1$ equations equating the fly-times of these circular trajectories and in case of odd-order modulators there is one additional equation equating the fly-times of the last second-order subsystem and the first-order subsystem. Finally, there is one pair of equations that can be combined to one for the Poincaré sections. Therefore, there are totally N independent equations for generating the tangent linear manifold.

VI. DC STABILITY FOR DISCRETE-TIME MODULATORS

This section presents the dc stability criteria and derivation of dc input bounds for practical, discrete-time $\Sigma\Delta$ modulators. Because the $\Sigma\Delta$ dynamics is symmetric about the modulator

dc input u , a positive u is assumed. The investigation begins with the concept of overlapped DOAs of various stable, embedded limit cycles featured in Sections II–V. Specifically, when u falls below a certain value, denoted by ξ_1 , it is found that every fixed-point on the Poincaré section (and their associated limit cycles) is within the DOAs of all other fixed-points. However, the existence of such a “common” DOA only guarantees a stable one-time switching from a continuous limit cycle to another. In practice, the discrete nature of the state trajectory and its intractable transition within the transition wedges prevent it from falling exactly onto any continuous limit cycles. Instead, the trajectory switches constantly between the DOAs of different stable limit cycles that may eventually lead to instability. To tackle the difficulty arising from the discrete constraint, it is first revealed that only certain fixed-points (and their corresponding limit cycles) are of interest whereupon all discrete trajectories tend to converge. With the help of a special computer algorithm, it is shown that for a u below ξ_2 ($0 < \xi_2 < \xi_1$), a domain on the hyperplane P (where all trajectories must cross despite their transition flow) can be found around these fixed-points such that discrete trajectories cannot escape. Such domain constitutes a

TABLE I
 NUMERICAL RESULTS FOR 15 DIFFERENT LP $\Sigma\Delta$ MODULATORS WITH OPTIMIZED NTF ZEROS AND OSR = 64. EACH SIMULATED BOUND IS ESTIMATED BY THE MODULATOR BEING STABLE WITHIN 2^{20} SIMULATION STEPS WITH ZERO-STATE INITIAL CONDITIONS IN $\mathbf{x}^{(0)}$

| Max NTF gain = 1.48 | order | | 3 | 4 | 5 | 6 | 7 |
|----------------------------------------|-----------------------|---------------|-------|--------|--------|--------|--------|
| | u | $ eig _{max}$ | | | | | |
| $F(\hat{P}_1, \check{P}_1)$ assumed | 0 | | 0.285 | 0.540 | 0.716 | 0.821 | 0.880 |
| | 0.1 | | 0.280 | 0.535 | 0.712 | 0.818 | 0.878 |
| | 0.2 | | 0.262 | 0.519 | 0.699 | 0.809 | 0.871 |
| | 0.3 | | 0.228 | 0.488 | 0.675 | 0.790 | 0.856 |
| | 0.4 | | 0.171 | 0.433 | 0.630 | 0.755 | 0.828 |
| | 0.5 | | 0.171 | 0.318 | 0.536 | 0.682 | 0.774 |
| | 0.6 | | 0.460 | 0.670 | 0.570 | 0.557 | 0.704 |
| | 0.7 | | 0.725 | 1.404 | 1.804 | 1.926 | 1.935 |
| | 0.8 | | 0.609 | | | | |
| bounds | $u @ eig _{max} = 1$ | | N/A | 0.647 | 0.639 | 0.644 | 0.648 |
| | simulated | | 0.776 | 0.650 | 0.576 | 0.545 | 0.535 |
| | ξ_1 | | 0.762 | 0.639 | 0.545 | 0.534 | 0.514 |
| | ξ_2 | | 0.722 | 0.547 | 0.474 | 0.425 | 0.436 |
| SNR (dB) | | | 92.12 | 105.80 | 116.21 | 126.42 | 135.87 |
| Max NTF gain = 1.5 | order | | 3 | 4 | 5 | 6 | 7 |
| | u | $ eig _{max}$ | | | | | |
| $F(\hat{P}_1, \check{P}_1)$ assumed | 0 | | 0.262 | 0.518 | 0.699 | 0.809 | 0.871 |
| | 0.1 | | 0.256 | 0.512 | 0.694 | 0.805 | 0.869 |
| | 0.2 | | 0.237 | 0.495 | 0.680 | 0.795 | 0.860 |
| | 0.3 | | 0.202 | 0.461 | 0.653 | 0.773 | 0.843 |
| | 0.4 | | 0.139 | 0.399 | 0.602 | 0.732 | 0.811 |
| | 0.5 | | 0.248 | 0.243 | 0.484 | 0.646 | 0.752 |
| | 0.6 | | 0.493 | 0.799 | 0.796 | 0.699 | 0.693 |
| | 0.7 | | 0.708 | 1.442 | 1.979 | 2.201 | 2.263 |
| | 0.8 | | 0.253 | | | | |
| bounds | $u @ eig _{max} = 1$ | | N/A | 0.631 | 0.620 | 0.625 | 0.629 |
| | simulated | | 0.766 | 0.630 | 0.561 | 0.530 | 0.509 |
| | ξ_1 | | 0.743 | 0.599 | 0.522 | 0.510 | 0.505 |
| | ξ_2 | | 0.711 | 0.523 | 0.451 | 0.394 | 0.406 |
| SNR (dB) | | | 92.23 | 105.63 | 117.12 | 127.26 | 135.52 |
| Max NTF gain = 1.52 | order | | 3 | 4 | 5 | 6 | 7 |
| | u | $ eig _{max}$ | | | | | |
| $F(\hat{P}_1, \check{P}_1)$ assumed | 0 | | 0.239 | 0.495 | 0.681 | 0.795 | 0.861 |
| | 0.1 | | 0.233 | 0.489 | 0.676 | 0.792 | 0.858 |
| | 0.2 | | 0.213 | 0.470 | 0.661 | 0.780 | 0.849 |
| | 0.3 | | 0.175 | 0.433 | 0.630 | 0.755 | 0.829 |
| | 0.4 | | 0.103 | 0.361 | 0.570 | 0.708 | 0.793 |
| | 0.5 | | 0.300 | 0.283 | 0.415 | 0.607 | 0.731 |
| | 0.6 | | 0.514 | 0.904 | 0.994 | 0.937 | 0.878 |
| | 0.65 | | 0.609 | 1.193 | 1.519 | 1.598 | 1.588 |
| | 0.75 | | 0.562 | | | | |
| bounds | $u @ eig _{max} = 1$ | | N/A | 0.618 | 0.601 | 0.606 | 0.610 |
| | simulated | | 0.756 | 0.605 | 0.519 | 0.494 | 0.485 |
| | ξ_1 | | 0.723 | 0.592 | 0.508 | 0.484 | 0.480 |
| | ξ_2 | | 0.699 | 0.511 | 0.432 | 0.363 | 0.376 |
| SNR (dB) | | | 92.97 | 106.39 | 117.82 | 127.03 | 136.79 |

positively invariant set (PIS, Schreier *et al.* [21]–[23]) on P , eventually guaranteeing dc stability.

As seen from Section V, a stable tangent linear manifold (i.e., all eigenvalues of magnitude less than unity) is accompanied with each stable fixed-point or limit cycle of a particular transition flow. This implies that a DOA can be found or estimated (by

the Lyapunov's methods, Ch. 5, [32]). The DOAs for different stable fixed-points are lying on their corresponding Poincaré sections. To unify the description all fixed-points and their associated DOAs, they are mapped back accordingly from the Poincaré sections onto the hyperplane P where all trajectories, irrespective of the transition flow, must cross [cf. Fig. 3(d)].

The description of all fixed-points and DOAs thereafter will always refer to their back-mapped versions on \mathbf{P} unless otherwise stated. Also, all fixed-points/limit cycles refer to first-order ones as obtained from the (first-return) Poincaré map, instead of higher order return maps.

A. Convergence of a One-Time Change in Transition Flow

The simplified case of the stability of a one-time switching from one stable limit cycle of a particular transition flow to another is first investigated, i.e., whether a stable limit cycle will converge asymptotically to another when the tilt factors ϕ , $\eta \in [0, 1]$ are changed. Numerical analysis as in Fig. 6(f) shows that for all properly designed $\Sigma\Delta$ modulators, a stable boundary transition flow (i.e., less-than-unity eigenvalues at $\phi = \eta = 1$) implies stable limit cycles corresponding to all other tilt factor tuples (ϕ, η) . Stability of the one-time switching or change of transition flow can then be assured by testing if all the back-mapped DOAs form an overlapped region on \mathbf{P} that contains all the stable, back-mapped fixed-points. In fact, the set of back-mapped fixed-points in respective subsystems reveals that extreme trajectory switching must occur when ϕ and/or η make their farthest jumps from 0 to 1 or otherwise. And that their DOAs guarantee the stability of intermediate switching (i.e., change of ϕ and/or η being less than one) provided extreme switching is stable. In other words, the stability of a one-time limit cycle switching can be assured by testing if the trajectory at one extremity converges to those stable limit cycles at other extremities, i.e., within the DOAs of them. This is done by first back-mapping the stable fixed-point (on $\hat{\mathbf{P}}_\phi$ or $\check{\mathbf{P}}_\eta$) onto \mathbf{P} , assigning it as the initial condition and subjecting it to another transition flow to see if it converges to the new limit cycle. The notation to denote a trajectory switching from $\mathbf{F}(\hat{\mathbf{P}}_{\phi_1}, \check{\mathbf{P}}_{\eta_1})$ to $\mathbf{F}(\hat{\mathbf{P}}_{\phi_2}, \check{\mathbf{P}}_{\eta_2})$ is

$$\phi : \phi_1 \rightarrow \phi_2, \quad \eta : \eta_1 \rightarrow \eta_2 | \mathbf{F}(\hat{\mathbf{P}}_{\phi_1}, \check{\mathbf{P}}_{\eta_1}). \quad (47)$$

Fig. 7(a) and (b) show the example of $\phi : 0 \rightarrow 1 | \mathbf{F}(\hat{\mathbf{P}}_0, \check{\mathbf{P}}_1)$ for the fourth-order modulator in Fig. 6. The nine cases of extreme switching to be tested are summarized in Fig. 7(c). The test of switching to $\mathbf{F}(\hat{\mathbf{P}}_0, \check{\mathbf{P}}_0)$ is not needed (denoted by broken arrows) because by starting a trajectory on \mathbf{P} with $\mathbf{F}(\hat{\mathbf{P}}_0, \check{\mathbf{P}}_0)$ assumption, it simply means that the trajectory will stay there forever and is, therefore, stable. The bound on u , denoted by ξ_1 , below which all fixed-points are within the DOA of one another, is tabulated in Table I.

Fig. 8 illustrates why this convergence analysis of one-time change in transition flow is not sufficient to ensure dc stability. In terms of trajectory intersections being mapped back from $\hat{\mathbf{P}}_\phi$ onto \mathbf{P} , Fig. 8(a) shows that for a one-time switching, when two fixed-points are within the DOA (mapped onto \mathbf{P}) of each other, they will converge asymptotically to one another if the tilt factor tuple is changed. But in the case of constantly changing transition flow as in the case of real, discrete-time $\Sigma\Delta$ modulators requiring integer fly-times, a trajectory may eventually become unstable. For example, in the numbered sequence in Fig. 8(a), the transition flow changes from the right (triangular) fixed-point toward the left (square) fixed-point. But before the convergence is complete, another transition flow change, say,

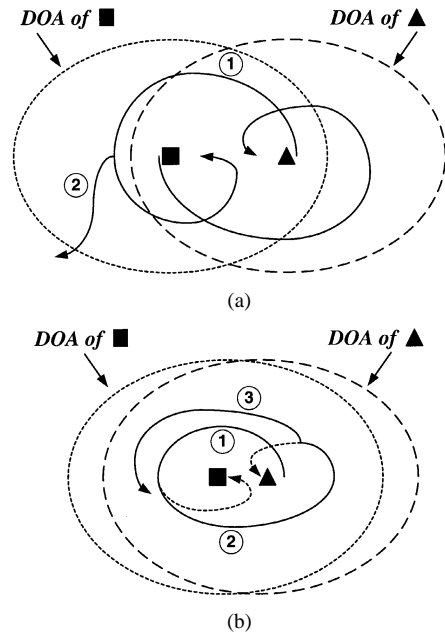


Fig. 8. Illustration of the (back-mapped) trajectory switching as on the hyperplane \mathbf{P} : (a) one-time switching and possible instability due to a second switching and (b) stable multiple switching.

back to the right fixed-point will render the trajectory unstable because it is now outside the DOA of the right fixed-point. Subsequently, to ensure dc stability, it is required that the DOAs of different fixed-points to be at least close enough such that trajectory changes are always encompassed by all DOAs, as in Fig. 8(b).

B. Convergence of Practically Changing Transition Flow

In practical $\Sigma\Delta$ modulators, the integer fly-time constraint lends itself to certain modifications from the continuous-time analysis. For example, the fly-times in Fig. 6(e) now need to be quantized and it is easy to see that near the dc input bound the negative fly-time is always quantized to one. In other words, practical discrete trajectories will tend to converge to those limit cycles of unity negative fly-time, as highlighted by the bold lines in the same plot. Fig. 9 shows some examples of tilt factor tuples recorded in real, discrete trajectories on the fly, demonstrating this inversely proportional relationship between ϕ and η . A lower u may have more than one level of quantization as in Fig. 9(b) and (d). For a positive input u , ϕ is dependent on η (and vice versa for a negative u , for example, when $\eta = 0$ ϕ must be equal to 1 and ϕ decreases when η increases). During stable dc operation near the bound, while η varies randomly between 0 and 1 to satisfy the integer positive fly-time constraint, ϕ will take on values near those of the fixed-points with an integer negative fly-time.

Fig. 8(b) suggests that the worst-case trajectory switching occurs when the trajectory is constantly switching between extreme fixed-points. This is justified by simulation showing that instability occurrence is preceded by tilt factor tuples alternating near the two ends of the series of fixed-points with unity negative fly-time. In view of this, a special computer algorithm for generating state trajectories is used to investigate discrete-time dc stability. It works in a way that for a positive input u , η is

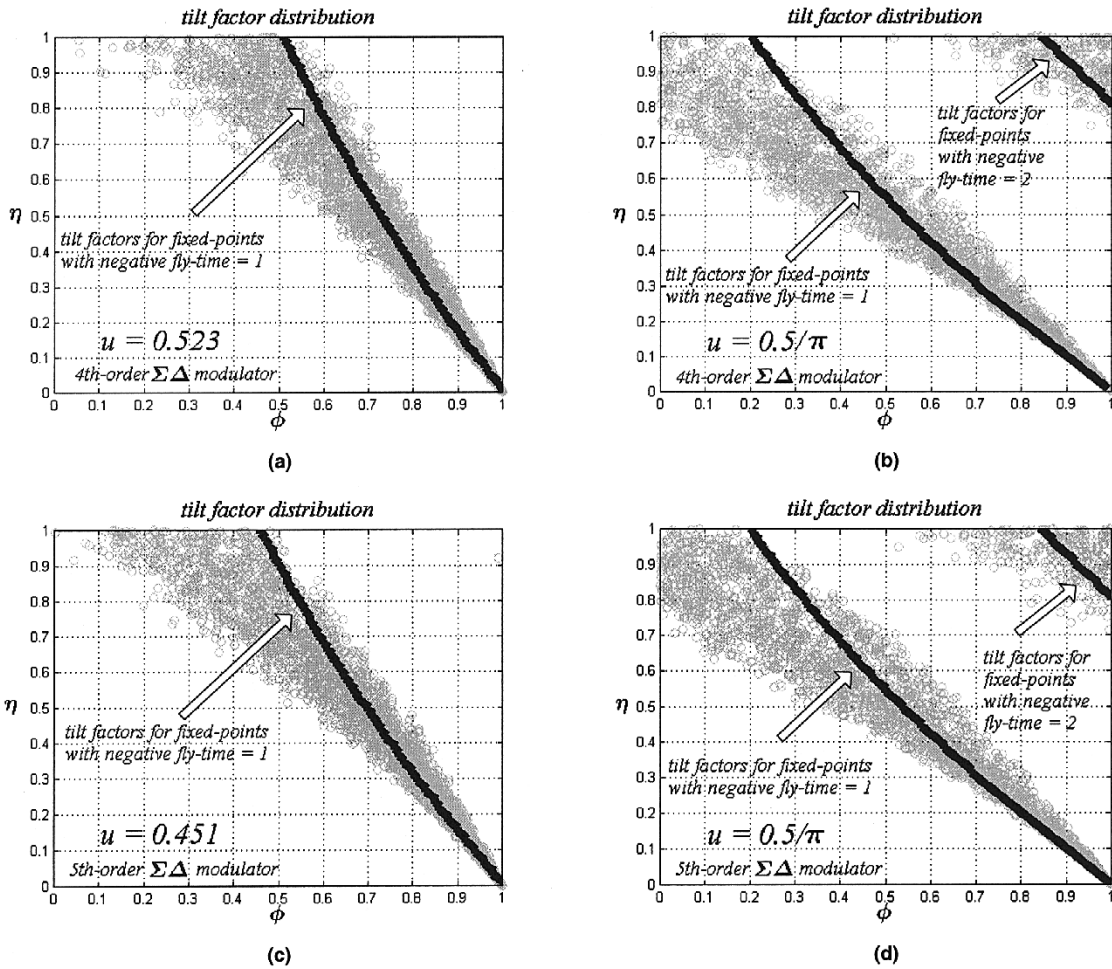


Fig. 9. Tilt factor tuples recorded for real $\Sigma\Delta$ modulators (OSR = 64, maximum NTF gain = 1.5, optimized NTF zeros, 2500 tuples for each figure) under different dc inputs: (a),(b) fourth-order and (c),(d) fifth-order.

an input parameter (which varies randomly in practice), while ϕ is determined from the integer negative fly-time constraint. It, therefore, represents a hybrid algorithm that is partially constrained by the integer negative fly-time, and partially dependent on the user input for the positive fly-time (which is not necessarily an integer but at least one). The worst-case switching sequence, studied in terms of the trajectory intersections on \hat{P}_ϕ mapped back to P , is found by starting a trajectory on one end of the series of fixed-point of unity negative fly-time with η set for the other end. While the trajectory is evolving, η is constantly alternated between the two extremes (0 and 1) such that the trajectory, if it is ever convergent, is farthest stretched. Experiments show that such η switching sequence is easy to locate and is in fact a periodic sequence. The artificial manipulation, therefore, simulates the worst-case η variation which will rarely occur, if not impossible, in practice. Subsequently, it is found that there exists a modulator input ξ_2 ($0 < \xi_2 < \xi_1$) below which the trajectory behavior on P is always bounded even under the contrived, worst-case changes in η . Practical trajectory intersections on P , with initial conditions near the fixed-points of unity negative fly-time, will then collapse into this bounded, worst-case region (effectively a PIS) due to the attraction of intermediate fixed-points corresponding to η between 0 and 1. In an actual modulator, the intractability in the transition flow, however, pre-

vents the process from settling on any fixed-points. For a lower u ($0 < u < \xi_2$) where negative fly-time can be quantized to integers larger than one, the corresponding fixed-points are situated in between the fixed-points (and associated DOAs) of unity negative fly-time and the above worst-case analysis still applies. Finally, considering the circular or oscillating nature of trajectories in each decomposed subsystem, it is easy to verify that a global trajectory is bounded in the whole state space provided it is bounded on the Poincaré sections. In other words, when u falls below ξ_2 , a PIS is formed. dc stability can then be assured for any discrete trajectories starting within this PIS, or more specifically, near those fixed-points of unity negative fly-time. For dc inputs above ξ_2 , trajectories as given by the algorithm may not be bounded due to possible couplings from second or higher order return maps. Trajectories taking on specific paths may travel beyond the attractive region and become unstable.

C. Numerical Examples and Results

Example fourth- and fifth-order modulators are used for exposition in Fig. 10 where only the first second-order decomposed subsystem is shown. Visualization is provided for the trajectory intersections on P back-mapped from \hat{P}_ϕ . The behavior in other subsystems as well as for the intersections back-mapped from \hat{P}_η onto P all look alike and are omitted

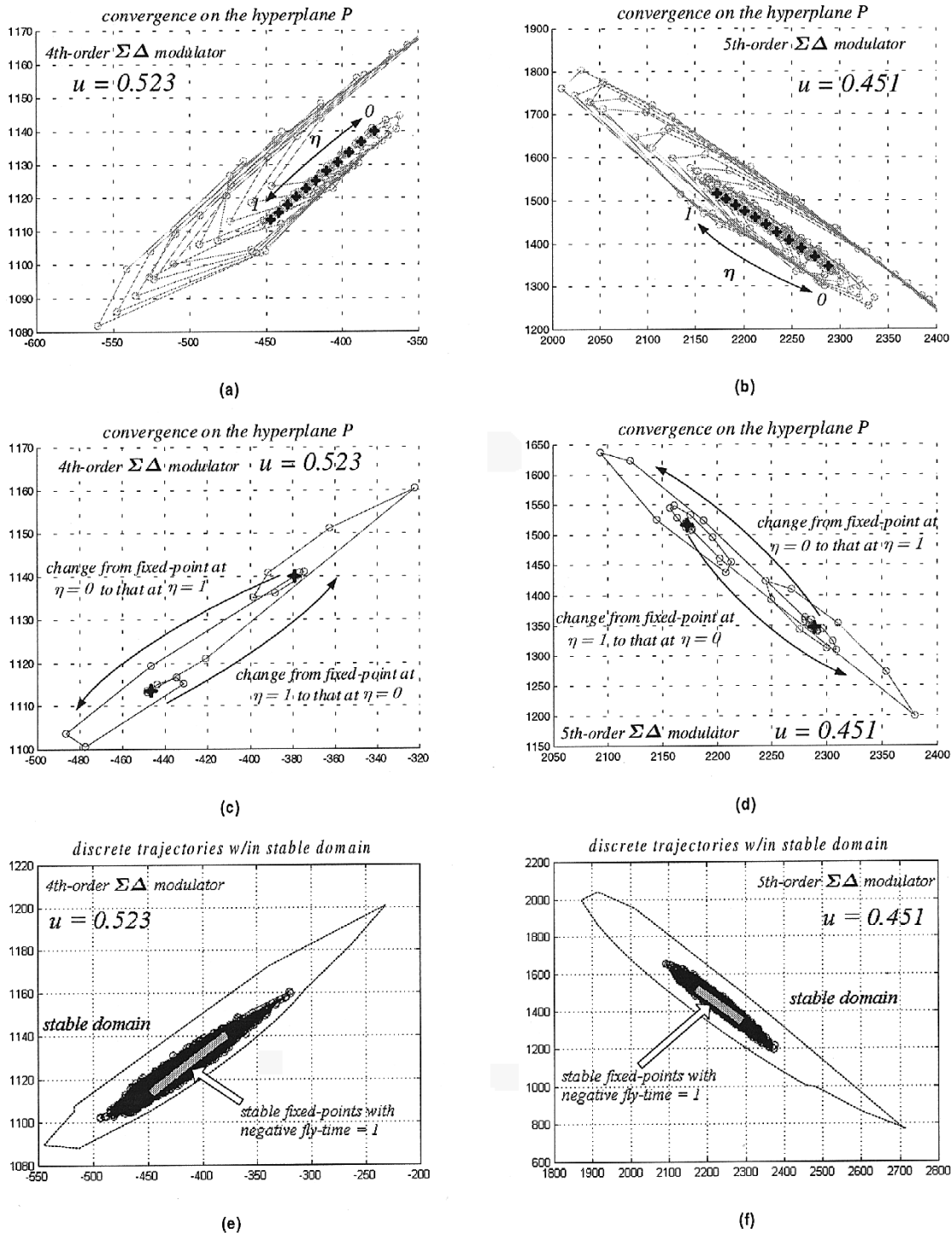


Fig. 10. Converging trajectory intersections on P (shown for the first second-order subsystems of the fourth- and fifth-order modulators in Figs. 5 and 6) by the computer algorithm: (a),(b) for different values of η with initial conditions from $\phi : 0 \rightarrow 1$ $\mathbf{F}(\hat{P}_0, \hat{P}_1)$; (c),(d) one-time switching between the two extreme fixed-points; (e),(f) the corresponding worst-case bounds or stable domains (dotted lines) together with 5000 simulated intersections.

for simplicity. Fig. 10(a) and (b) show that by fixing η at different values, the state trajectories are all convergent to the corresponding fixed-points of unity negative fly-time for $u = \xi_2$. Fig. 10(c) and (d) illustrate trajectory convergence of a one-time switching from one extreme fixed-point to another. Fig. 10(e) and (f) show the worst-case bounds (stable domains) found by varying η as described and the actual trajectory intersections obtained from simulation, verifying the enclosure by the bounds (on P). Finally, 15 modulators

of different orders and maximum NTF gains are designed with optimized NTF zeros using the $\Delta\Sigma$ design toolbox [3] at OSR = 64. The maximum eigenvalue magnitudes of their tangent linear manifolds, assuming boundary transition flow $\mathbf{F}(\hat{P}_1, \hat{P}_1)$, against different dc inputs are tabulated in Table I. The simulated dc bounds and the corresponding theoretical settings as discussed before, are given. Also shown are the estimated SNRs for a sine wave input of amplitude ξ_2 situated

in the middle of the baseband, confirming the advantage of using higher order modulators for higher resolution. It can be observed from the table that relying on the stability of the farthest stretched boundary transition flow $\mathbf{F}(\hat{\mathbf{P}}_1, \check{\mathbf{P}}_1)$ is unreliable since the dc input bounds, thus, estimated (where the maximum eigenvalue magnitude just reaches unity) show an increasing trend when modulator order is greater than four, contradicting the decreasing trend borne out by simulation and experience. Moreover, eigenvalues of unit magnitude cannot even be found for $\mathbf{F}(\hat{\mathbf{P}}_1, \check{\mathbf{P}}_1)$ in third-order modulators. The application of the proposed stability analysis eliminates such contradiction and produces a theoretical dc bound ξ_2 consistent with and, due to its worst-case assumption, more conservative than the simulated dc bound. The discrete trajectories in practice, depending on their initial conditions, may never reach this worst-case trajectory and a bound between ξ_2 and ξ_1 is expected. It should be stressed that no amount of simulation can guarantee dc stability since a modulator can go unstable even after a million simulation steps as we have observed. In contrast, derivation of the dc bound ξ_2 is a one-off process in finite time. Interestingly, all the simulated bounds are slightly larger than ξ_1 . This demonstrates both the possible pitfalls of simulation and the safety margin provided by the proposed approach in dc bound estimation.

VII. CONCLUSION

This paper has presented an analytical investigation of the state trajectory behavior and dc stability of high-order (order > 2), LP $\Sigma\Delta$ modulators with distinct unit circle NTF zeros. Algebraic difficulties arising from the highly nonlinear quantizer function have been tackled through combining and generalizing the techniques of state-space decomposition, continuous-time embedding, and Poincaré map analysis. The study of both even- and odd-order modulators has been presented in a unified and scalable manner. It has been shown that a high-order modulator system can be decomposed into second- and first-order subsystems, thus facilitating graphical interpretation and analysis. Systems of nonlinear equations have been established for the location of embedded state-space limit cycles using efficient numerical methods. It has been revealed that a necessary condition for dc stability requires the overlapping of DOAs associated with all types of transition flow and the inclusion of all stable limit cycles within this overlapped region. A special computer algorithm has been written to solve for the sufficient condition for dc stability under the regime of constantly changing transition flow present in real, discrete-time $\Sigma\Delta$ modulators. Numerical examples featuring modulators of different orders and maximum NTF gains have been given to illustrate the results.

REFERENCES

- [1] J. C. Cy and G. C. Temes, Eds., *Oversampling Delta-Sigma Data Converters*. New York: IEEE Press, 1992.
- [2] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters*. New York: IEEE Press, 1997.
- [3] R. Schreier, "An empirical study of high-order single-bit delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 461–466, Aug. 1993.

- [4] Y. Botteron and B. Nowrouzian, "An investigation of bandpass sigma-delta a/D converters," in *Proc. 40th Midwest Symp. Circuits Syst.*, vol. 1, 1998, pp. 293–296.
- [5] Y. Botteron, B. Nowrouzian, and A. T. G. Fuller, "Design and switched-capacitor implementation of a new cascade-of-resonators Σ - Δ converter configuration," in *Proc. IEEE Int. Symp. Circuits Systems*, vol. 2, 1999, pp. 45–48.
- [6] T. Ritoniemi, T. Karema, and H. Tenhunen, "The design of stable high order 1-bit sigma-delta modulators," in *Proc. IEEE Int. Symp. Circuits Systems*, vol. 4, 1990, pp. 3267–3270.
- [7] R. T. Baird and T. S. Fiez, "Stability analysis of high-order delta-sigma modulation for ADC's," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 59–62, Jan. 1994.
- [8] S. H. Ardalan and J. J. Paulos, "Analysis of nonlinear behavior in delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. CAS-34, pp. 593–603, June 1987.
- [9] N. A. Fraser and B. Nowrouzian, "Stability analysis of multiple-feedback oversampled Σ - Δ A/D converter configurations," in *Proc. IEEE Midwest Symp. Circuits Systems*, vol. 2, 2000, pp. 676–679.
- [10] H. Wang, "A geometric view of sigma delta modulations," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 402–405, June 1992.
- [11] —, " $\Sigma\Delta$ modulation from the perspective of nonlinear dynamics," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 3, 1992, pp. 1296–1299.
- [12] —, "On the stability of third-order sigma-delta modulation," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 2, 1993, pp. 1377–1380.
- [13] —, "A Study of sigma delta modulations as dynamical systems," Ph.D. dissertation, Columbia Univ., New York, 1993.
- [14] S. Hein and A. Zakhor, "On the stability of sigma delta modulators," *IEEE Trans. Signal Process.*, vol. 41, pp. 2322–2348, July 1993.
- [15] S. Pinault and P. Lopresti, "On the behavior of the double-loop sigma delta modulator," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 467–479, 1993.
- [16] O. Feely, "Nonlinear dynamics of sigma-delta modulation," in *Proc. 34th Midwest Symp. Circuits and Systems*, vol. 2, 1992, pp. 760–763.
- [17] —, "Nonlinear dynamics of sigma-delta modulation," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 6, 1994, pp. 101–104.
- [18] —, "A tutorial introduction to nonlinear dynamics and chaos and their application to sigma-delta modulators," *Int. J. Circuit Theory Applicat.*, vol. 25, pp. 347–367, 1997.
- [19] R. Farrell and O. Feely, "Application of nonlinear analysis techniques to sigma-delta modulators," *Proc. IEE Colloq. Signals, Systems and Chaos*, pp. 4/1–4/6, 1997.
- [20] —, "Bounding the integrator outputs of second-order sigma-delta modulators," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 691–702, June 1998.
- [21] R. Schreier, M. V. Goodson, and B. Zhang, "Invariant sets for general second-order low-pass delta-sigma modulators with dc inputs," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 6, 1994, pp. 1–4.
- [22] —, "Proving stability of delta-sigma modulators using invariant sets," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 1, 1995, pp. 633–636.
- [23] —, "An algorithm for computing convex positively invariant sets for delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 38–44, Jan. 1997.
- [24] P. Steiner and W. Yang, "A framework for analysis of high-order sigma-delta modulators," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 1–10, Jan. 1997.
- [25] —, "Stability analysis of the second order Σ - Δ modulator," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 5, 1994, pp. 365–368.
- [26] —, "Stability of high order sigma-delta modulators," in *Proc. Int. Symp. Circuits and Systems*, vol. 3, 1996, pp. 52–55.
- [27] V. Mladenov, H. Hegt, and A. van Roermund, "On the stability of high order sigma-delta modulators," in *Proc. 8th IEEE Conf. Electrical Circuits Systems*, vol. 3, 2001, pp. 1383–1386.
- [28] R. L. Burden and J. D. Faires, *Numerical Analysis*, 7th ed. Pacific Grove, CA: Brooks/Cole, 2001.
- [29] N. Wong and T. S. Ng, "A generalized direct-form delta-operator-based IIR filter with minimum noise gain and sensitivity," *IEEE Trans. Circuits Syst. II*, vol. 48, pp. 425–431, Apr. 2001.
- [30] R. H. Middleton and G. C. Goodwin, *Digital Control and Estimation: A Unified Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
- [31] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. New York: Springer, 1996.
- [32] P. Cook, *Nonlinear Dynamical Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.



Ngai Wong (S'98–M'02) was born in Fujian, China, in 1975. He received the B.Eng. (first class honors) in electrical and electronic engineering from The University of Hong Kong in 1999 where he is currently working toward the Ph.D. degree focusing on the areas of signal processing and telecommunications.

From 1997 to 1998, he was a Product Engineer at Motorola Semiconductor Hong Kong Ltd. under an internship program, where he worked on mixed-mode IC design and circuit testing. His

research interests include digital filtering using Δ -operators, $\Sigma\Delta$ modulator design and analysis, and the application of semidefinite programming in communication systems and VLSI synthesis.

Mr. Wong was awarded the 2nd prize of the IEEE Undergraduate Student Paper Contest, Hong Kong Section, in 1999, the P. K. Yu Memorial Scholarship, in 2000, the Sir Edward Youde Memorial Fellowship, and the Leung Wai Sun Fellowship, in 2002.



Tung-Sang Ng (S'74–M'78–SM'90–F'03) received the B.Sc.(Eng.) degree from The University of Hong Kong in 1972, and the M.Eng.Sc. and Ph.D. degrees from the University of Newcastle, Newcastle, Australia, in 1974 and 1977, respectively, all in electrical engineering.

He worked for BHP Steel International and The University of Wollongong, Australia. In 1991, he returned to The University of Hong Kong, taking up the position of Professor and Chair of the Electronic Engineering Department. Since 2000, he has been the

Head of the Department of Electrical and Electronic Engineering. He has published over 180 International journal and conference papers. His current research interests include wireless communication systems, spread spectrum techniques, CDMA and digital signal processing.

Dr. Ng is a Fellow of the Institution of Electrical Engineers (IEE), London, U.K. and the Hong Kong Institution of Engineers (HKIE). He was awarded the Honorary Doctor of Engineering Degree by the University of Newcastle, Australia in August, 1997, for his service to higher education specifically to engineering education. He received the Senior Croucher Foundation Fellowship, in 1999 and the IEEE Third Millennium medal in February, 2000. He was the General Chair of ISCAS'97 and the VP-Region 10 of IEEE Circuits and Systems (CAS) Society for 1999 and 2000. He was an Executive Committee Member and a Board Member of the IEE Informatics Divisional Board from 1999 to 2001 and was an Ordinary Member of the IEE Council from 1999 to 2001.