

# Estimation of an Initial Condition of Sigma-delta Modulators via Projection Onto Convex Sets

Charlotte Yuk-Fan Ho, *Student Member, IEEE*, Bingo Wing-Kuen Ling and Joshua D. Reiss

**Abstract**—In this paper, an initial condition of strictly causal rational interpolative sigma-delta modulators (SDMs) is estimated based on quantizer output bit streams and an input signal. A set of initial conditions generating bounded trajectories is characterized. It is found that a set of initial conditions generating bounded trajectories but not necessarily corresponding to quantizer output bit streams is convex. Also, it is found that a set of initial conditions corresponding to quantizer output bit streams but not necessarily generating bounded trajectories is convex too. Moreover, it is found that an initial condition both corresponding to quantizer output bit streams and generating bounded trajectories is uniquely defined if the loop filter is unstable (Here, an unstable loop filter refers to that with at least one of its poles being strictly outside the unit circle). To estimate that unique initial condition, a projection onto convex set approach is employed. Numerical computer simulations show that the employed method can estimate the initial condition effectively.

**Index Terms**—Sigma-delta modulators, stability, admissibility, projection onto convex sets.

## I. INTRODUCTION

SINCE some interpolative SDMs consist of an unstable loop filter [1], [7], [8] and the input-output relationship of a quantizer is characterized by a discontinuous nonlinear function, a feedback connection of the unstable loop filter and the quantizer would cause the dynamics of these interpolative SDMs very complicated. Chaotic and fractal behaviors may occur [4], [6], [12], [13]. As chaotic behaviors are highly dependent on an initial condition, the dynamics of an interpolative SDM would be very different if there is a very small change in the initial condition. When there is a sudden change of a supply voltage or a mechanical shaking, the content in a register containing the initial condition of an interpolative SDM may be corrupted. Since signals in interpolative SDMs are constructed based on the initial condition and an input signal, in this case, the constructed signal will be very different from the actual one and a serious construction error would be encountered.

Manuscript received March 3, 2006; revised June 22, 2006. This work was supported by a research grant from Queen Mary, University of London.

Bingo Wing-Kuen Ling is with the department of Electronic Engineering, King's College London (e-mail: wing-kuen.ling@kcl.ac.uk).

Charlotte Yuk-Fan Ho and Joshua D. Reiss are with the department of Electronic Engineering, Queen Mary, University of London.

Copyright (c) 2006 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permission@ieee.org.

In order to minimize the construction error, it is necessary to estimate the initial condition of an interpolative SDM based on quantizer output bit streams and an input signal. However, some fundamental questions have not been explored yet. For examples, for a certain type of interpolative SDMs, such as interpolative SDMs with unstable loop filter, does there exist a unique initial condition both corresponding to quantizer output bit streams and generating bounded trajectories? If yes, how can we find an initial condition which is closed to that unique initial condition?

One of the most common methods for estimating the initial condition of an interpolative SDM is to formulate the problem as an optimization problem. In [7], constraints were imposed so that the estimated initial condition is guaranteed to correspond to quantizer output bit streams. However, the obtained solution does not guarantee to generate bounded trajectories because stability condition was not exploited in the optimization problem in [7]. The unbounded state responses shown in Section V of this paper illustrate this phenomenon. In this paper, necessary and sufficient bounded conditions on state variables are characterized and constraints based on these bounded conditions are imposed so that it is guaranteed to generate bounded trajectories.

The outline of this paper is as follows. In Section II, notations used throughout this paper are introduced. In Section III, analytical results are presented. This paper is to estimate an initial condition of strictly causal rational interpolative SDMs. To address this problem, projection onto convex set approach is employed. In order to apply this approach, two convex sets are needed to be characterized. The set of initial conditions generating bounded trajectories is characterized by Theorem 1 and the set of initial conditions corresponding to the given quantizer output bit streams were given in [7]. It is shown in Theorems 3 and 4 that these two sets are convex. Hence, the projection onto convex set approach can be applied for estimating the initial condition. Besides, it is shown in Theorem 2 that the initial condition both corresponding to quantizer output bit streams and generating bounded trajectories is uniquely defined if the loop filter is unstable, so it is guaranteed that the solution found by the proposed algorithm based on the projection onto convex set approach is the unique solution of the problem. In Section IV, the details of the proposed algorithm based on the projection onto convex set approach are discussed. In Section V, numerical computer simulation results are presented to illustrate the effectiveness of the method. Finally, conclusions are summarized in Section VI.

## II. NOTATIONS

The block diagram of an interpolative SDM is shown in Figure 1, in which the loop filter and the quantizer of the interpolative SDM are denoted as  $F(z)$  and  $Q(\cdot)$ , respectively. Since an interpolative SDM with a single-input single-output strictly causal rational loop filter and a single bit quantizer having the decision boundary at zero are widely employed in industries, an interpolative SDM with this type of loop filter and quantizer is considered in this paper. The state space matrices of the loop filter are denoted as  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $D$ . Due to the negative feedback configuration and the strictly causal condition,  $D=0$ . Denote an input of the interpolative SDM, an output of the loop filter, the quantizer output bit streams and the state vector of the loop filter as  $u(k)$ ,  $y(k)$ ,  $s(k)$  and  $\mathbf{x}(k)$ , respectively. Then the dynamics of the interpolative SDM can be characterized by the following standard state space equations:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}(u(k) - s(k)), \quad (1a)$$

and

$$y(k) = \mathbf{C}\mathbf{x}(k), \quad (1b)$$

where

$$s(k) = Q(y(k)), \quad (1c)$$

in which

$$Q(y(k)) \equiv \begin{cases} 1 & y(k) \geq 0 \\ -1 & y(k) < 0 \end{cases}. \quad (1d)$$

## III. ANALYTICAL RESULTS

### A. Necessary and sufficient bounded conditions of state variables

In some circuits and systems, such as audio systems [3], some eigenvalues of  $\mathbf{A}$  are strictly outside the unit circle. Hence, state variables of these interpolative SDMs may not be bounded for some bounded inputs and initial conditions. To guarantee that the state variables are bounded, define  $\Gamma$  as the set of initial conditions such that  $\mathbf{x}(k)$  is bounded  $\forall k \geq 0$ . Denote  $U(z)$  and  $S(z)$  as the  $z$ -transforms of  $u(n)$  and  $s(n)$ , respectively.

#### Theorem 1

$\mathbf{x}(0) \in \Gamma$  if and only if there exists a signal with its  $z$ -transform  $\mathbf{P}(z)$  defined by

$$\mathbf{P}(z) \equiv (1 - z^{-1}) (\mathbf{I} - \mathbf{A}z^{-1})^{-1} (\mathbf{x}(0) + z^{-1}\mathbf{B}(U(z) - S(z))) \quad (2)$$

which is analytical at  $z=1$ .

#### Proof:

Since  $\mathbf{x}(k) = \mathbf{A}^k \mathbf{x}(0) + \sum_{n=0}^{k-1} \mathbf{A}^{k-1-n} \mathbf{B}(u(n) - s(n))$  for  $k \geq 1$ ,

$$\lim_{k \rightarrow +\infty} \mathbf{x}(k) = \lim_{z \rightarrow 1} (1 - z^{-1}) (\mathbf{I} - \mathbf{A}z^{-1})^{-1} (\mathbf{x}(0) + z^{-1}\mathbf{B}(U(z) - S(z)))$$

$\mathbf{x}(k)$  is bounded  $\forall k \geq 0$  if and only if the region of convergence of each element in  $(1 - z^{-1}) (\mathbf{I} - \mathbf{A}z^{-1})^{-1} (\mathbf{x}(0) + z^{-1}\mathbf{B}(U(z) - S(z)))$  includes the point  $z=1$ . Hence, the necessary and sufficient bounded conditions

for any bounded inputs become the existence of an initial condition such that  $\mathbf{P}(z)$  is analytical at  $z=1$ . This completes the proof. ■

It is worth noting that  $\mathbf{P}(z)$  is not the transfer function of the interpolative SDMs because  $\mathbf{P}(z)$  is nonlinear with respect to  $\mathbf{x}(0)$  and transfer functions usually refer to linear time-invariant systems only. The nonlinearity of  $\mathbf{P}(z)$  with respect to  $\mathbf{x}(0)$  is due to the term  $S(z)$  because  $S(z) = \sum_{\forall k} Q(\mathbf{C}\mathbf{x}(k))z^{-k}$ .

The set of initial conditions generating bounded trajectories depends on the nonlinearity of the interpolative SDMs. Although input, initial condition and system function of any systems are independent (that is, for any linear or nonlinear systems, we can have arbitrary input and initial condition.), based on given input, initial condition and system function, system characteristics (such as the boundedness of both state responses and output response) are uniquely defined and these characteristics depend on input, initial condition and system function.

The importance of Theorem 1 is on the characterization of the set of initial conditions generating bounded trajectories for any bounded inputs.

#### Corollary 1

If  $\mathbf{A}$  contains some unstable eigenvalues, then  $\mathbf{P}(z)$  is analytical at  $z=1$  if and only if  $\mathbf{x}(0) + z^{-1}\mathbf{B}(U(z) - S(z))$  contains unstable zeros which cancel exactly the unstable poles of  $\mathbf{A}$  and  $\mathbf{x}(0) + z^{-1}\mathbf{B}(U(z) - S(z))$  has no unstable pole.

#### Proof:

This result can be trivially derived from Theorem 1. Hence, the proof is omitted. ■

Corollary 1 is important as we have discussed before that some eigenvalues of  $\mathbf{A}$  of some circuits and systems are strictly outside the unit circle.

### B. Uniqueness of an initial condition corresponding to given stable admissible quantizer output bit streams

Denote an infinite length binary sequence with each element in the sequence being either 1 or -1 as  $\mathbf{s} \equiv (s(0), s(1), \dots)$ . Define  $\Psi \equiv \{1, -1\} \times \{1, -1\} \times \dots$  and  $n$  as the dimension of the state vectors. Denote the mapping from  $\mathfrak{R}^n$  to  $\Psi$  as  $\Lambda$  such that (1a)-(1d) are satisfied, that is  $\Lambda(\mathbf{x}(0)) = \mathbf{s}$ . The set of quantizer output bit streams is said to be admissible if  $\forall \mathbf{s}$  in the set of quantizer output bit streams,  $\exists \mathbf{x}(0) \in \mathfrak{R}^n$  such that  $\Lambda(\mathbf{x}(0)) = \mathbf{s}$ . It is worth noting that  $\Psi$  may not be admissible because it may exist  $\mathbf{s} \in \Psi$  such that  $\nexists \mathbf{x}(0) \in \mathfrak{R}^n$  satisfying  $\Lambda(\mathbf{x}(0)) = \mathbf{s}$ . Denote the admissible set of quantizer output bit streams as  $\Psi_b$ , that is  $\Psi_b \equiv \{\mathbf{s} : \forall \mathbf{x}(0) \in \mathfrak{R}^n \text{ (1a)-(1d) are satisfied}\}$ . Define a mapping from  $\mathfrak{R}^n$  to  $\Psi_b$  as  $\Lambda_b$  such that (1a)-(1d) are satisfied. Obviously,  $\Lambda_b$  is surjective, but it may not be injective. Even though it is injective, the corresponding

trajectories may not be bounded. Denote the set of stable admissible quantizer output bit streams as  $\Psi_{sb}$ , that is  $\Psi_{sb} \equiv \{s : \forall \mathbf{x}(0) \in \Gamma \text{ (1a)-(1d) are satisfied.}\}$ . Denote a mapping from  $\Gamma$  to  $\Psi_{sb}$  as  $\Lambda_{sb}$  such that (1a)-(1d) are satisfied.

*Theorem 2*

If  $\mathbf{A}$  is unstable, then  $\Lambda_{sb}$  is bijective.

*Proof:*

Suppose  $\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Gamma$  and  $\mathbf{x}^1(0) \neq \mathbf{x}^2(0)$  such that  $\Lambda_{sb}(\mathbf{x}^1(0)) = \Lambda_{sb}(\mathbf{x}^2(0)) = \mathbf{s} \in \Psi_{sb}$ , then

$$\mathbf{x}^1(k) = \mathbf{A}^k \mathbf{x}^1(0) + \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \quad \text{for } k \geq 1,$$

$$\mathbf{x}^2(k) = \mathbf{A}^k \mathbf{x}^2(0) + \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n))$$

which implies that  $\mathbf{A}^k(\mathbf{x}^1(0) - \mathbf{x}^2(0)) = \mathbf{x}^1(k) - \mathbf{x}^2(k)$  for  $k \geq 1$ . Since  $\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Gamma$ ,  $\mathbf{x}^1(k)$  and  $\mathbf{x}^2(k)$  are bounded. If  $\mathbf{A}$  is unstable, since  $\mathbf{x}^1(0) \neq \mathbf{x}^2(0)$ , then  $\mathbf{x}^1(k) - \mathbf{x}^2(k)$  will be unbounded, which is a contradiction because a subtraction of any two bounded sequences must be bounded. Hence,  $\mathbf{x}^1(0) = \mathbf{x}^2(0)$  and  $\Lambda_{sb}$  is injective. Since  $\Psi_{sb}$  is the stable admissible set of quantizer output bit streams,  $\Lambda_{sb}$  is surjective.

Hence,  $\Lambda_{sb}$  is bijective and this completes the proof. ■

The novelty of Theorem 2 is as follows. The initial conditions generating bounded trajectories but not necessarily corresponding to quantizer output bit streams are *not* uniquely defined because if  $\mathbf{x}(0) \in \Gamma$ , then  $\mathbf{x}(k) \in \Gamma \forall k \geq 0$ . Similarly, the initial conditions corresponding to quantizer output bit streams but not necessarily generating bounded trajectories are *not* uniquely defined too. However, the initial condition *both* corresponding to given quantizer output bit streams and generating bounded trajectories is uniquely defined when the loop filter is unstable.

Since Theorem 2 reveals that the initial condition both corresponding to given quantizer output bit streams and generating bounded trajectories is uniquely defined when  $\mathbf{A}$  is unstable, the importance of Theorem 2 is that if  $\mathbf{A}$  is unstable and an initial condition satisfying these two properties is found, then that initial condition is the unique solution. Moreover, Theorem 2 explains why the initial condition is sensitive to the state responses of the interpolative SDM. This result is useful for the further investigation of the occurrence of chaotic behaviors.

*Corollary 2*

If  $\mathbf{A}$  is unstable,  $\mathbf{x}(0) \in \Gamma$  and  $\mathbf{s} \in \Psi_{sb}$  such that  $\mathbf{s}$  is periodic with period  $M$ , then  $\mathbf{x}(0) = \mathbf{x}(kM) \forall k \geq 0$ .

*Proof:*

If  $\mathbf{x}(0) \in \Gamma$ , then  $\mathbf{x}(kM) \in \Gamma \forall k \geq 0$ . Since  $\mathbf{A}$  is unstable,  $\Lambda_{sb}$  is bijective. As  $\mathbf{s}$  is periodic with period  $M$ , this implies that  $\mathbf{x}(0) = \mathbf{x}(kM) \forall k \geq 0$ . This completes the proof. ■

Although it was reported in [10] that periodicity of quantizer

output bit streams implies periodicity of state vectors, the analysis in [10] is based on the study of interpolative SDMs with DC poles and without DC poles. For interpolative SDMs without DC poles, it assumes that the null space of  $\mathbf{I} - \mathbf{A}^M$  is  $\{\mathbf{0}\}$ . In fact, this assumption is not true if  $\mathbf{A}$  is marginally stable or strictly stable [1].

When  $\mathbf{A}$  is unstable, although Theorem 2 provides information on the uniqueness of an initial condition both corresponding to given quantizer output bit streams and generating bounded trajectories, the method for finding that initial condition is not addressed yet. In order to find that initial condition, these two properties are considered separately. That is, the set of initial conditions corresponding to given quantizer output bit streams but not necessarily generating bounded trajectories, and the set of initial conditions generating bounded trajectories but not necessarily corresponding to given quantizer output bit streams, are considered separately.

*C. Convexity of admissible set of initial conditions*

The admissible condition for quantizer output bit streams were characterized in [7]. Here is a summary. For a given  $\mathbf{s} \in \Psi_b$ , since  $s(k) = 1$  if  $y(k) \geq 0$  and  $s(k) = -1$  if  $y(k) < 0 \forall k \geq 0$ , this implies that  $s(k)y(k) \geq 0 \forall k \geq 0$ , that is,  $s(0)\mathbf{C}\mathbf{x}(0) \geq 0$  and

$$s(k) \left( \mathbf{C}\mathbf{A}^k \mathbf{x}(0) + \mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \right) \geq 0 \quad (3a)$$

for  $k \geq 1$ . For given quantizer output bit streams, denote the set of initial conditions that satisfies (3a) as  $\Phi$ , that is  $\Phi \equiv \{\mathbf{x}(0) \text{ such that}$

$$\begin{bmatrix} s(0)\mathbf{C} \\ s(1)\mathbf{C}\mathbf{A} \\ \vdots \\ s(k)\mathbf{C}\mathbf{A}^k \end{bmatrix} \mathbf{x}(0) + \begin{bmatrix} 0 \\ s(1)\mathbf{C}\mathbf{B}(u(0) - s(0)) \\ \vdots \\ s(k)\mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \end{bmatrix} \geq \mathbf{0} \quad (3b)$$

for  $k \geq 1$ .}

*Theorem 3*

$\Phi$  is convex.

*Proof:*

If  $\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Phi$ , then  $\forall \lambda \in [0, 1]$ ,  $\lambda s(0)\mathbf{C}\mathbf{x}^1(0) \geq 0$ ,  $(1-\lambda)s(0)\mathbf{C}\mathbf{x}^2(0) \geq 0$ ,

$\lambda s(k) \left( \mathbf{C}\mathbf{A}^k \mathbf{x}^1(0) + \mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \right) \geq 0$  and

$(1-\lambda)s(k) \left( \mathbf{C}\mathbf{A}^k \mathbf{x}^2(0) + \mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \right) \geq 0$  for

$k \geq 1$ . Hence,  $s(0)\mathbf{C}(\lambda \mathbf{x}^1(0) + (1-\lambda)\mathbf{x}^2(0)) \geq 0$  and

$s(k) \left( \mathbf{C}\mathbf{A}^k (\lambda \mathbf{x}^1(0) + (1-\lambda)\mathbf{x}^2(0)) + \mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n) - s(n)) \right) \geq 0$

for  $k \geq 1$ . This implies that  $\lambda \mathbf{x}^1(0) + (1-\lambda)\mathbf{x}^2(0) \in \Phi$  and it completes the proof. ■

It is worth noting that although  $\mathbf{s} \in \Psi_b$  and  $\Lambda_{sb}$  is bijective

when  $\mathbf{A}$  is unstable, there may exist  $\mathbf{x}(0) \in \mathfrak{R}^n \setminus \Gamma$  such that  $\Lambda_b(\mathbf{x}(0)) = \mathbf{s} \in \Psi_b \setminus \Psi_{sb}$  because  $\mathfrak{R}^n \neq \Gamma$ .

The importance of Theorem 3 is to allow the estimation of an initial condition based on projection onto convex set approach, which will be discussed in Section IV.

#### D. Convexity of the bounded set of initial conditions

For any given  $\mathbf{s} \in \Psi$ , define  $\Theta \equiv \{\mathbf{x}(0)$  such that  $(1-z^{-1})(\mathbf{I}-\mathbf{A}z^{-1})^{-1}(\mathbf{x}(0)+z^{-1}\mathbf{B}(U(z)-S(z)))$  is analytical at  $z=1\}$ . It is worth noting that there may exist  $\mathbf{x}(0) \in \Theta \setminus \Gamma$  because  $\mathbf{s} \in \Psi \setminus \Psi_b$ . In other words,  $\mathbf{x}(0)$  may not satisfy (1a)-(1d).

#### Theorem 4

$\Theta$  is convex.

*Proof:*

For any given  $\mathbf{s} \in \Psi$ , suppose  $\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Theta$ , then there exist two signals with z-transform  $\mathbf{P}^1(z)$  and  $\mathbf{P}^2(z)$  denoted as

$$\mathbf{P}^1(z) = (1-z^{-1})(\mathbf{I}-\mathbf{A}z^{-1})^{-1}(\mathbf{x}^1(0)+z^{-1}\mathbf{B}(U(z)-S(z)))$$

and

$$\mathbf{P}^2(z) = (1-z^{-1})(\mathbf{I}-\mathbf{A}z^{-1})^{-1}(\mathbf{x}^2(0)+z^{-1}\mathbf{B}(U(z)-S(z)))$$

such that they are analytical at  $z=1$ . Since  $\forall \lambda \in [0,1]$ ,

$$\begin{aligned} & \lambda \mathbf{P}^1(z) + (1-\lambda) \mathbf{P}^2(z) = \\ & (1-z^{-1})(\mathbf{I}-\mathbf{A}z^{-1})^{-1}(\lambda \mathbf{x}^1(0) + (1-\lambda) \mathbf{x}^2(0) + z^{-1} \mathbf{B}(U(z)-S(z))) \end{aligned}$$

and  $\lambda \mathbf{P}^1(z) + (1-\lambda) \mathbf{P}^2(z)$  is analytical at  $z=1$ , this implies that  $\lambda \mathbf{x}^1(0) + (1-\lambda) \mathbf{x}^2(0) \in \Theta$ . This completes the proof. ■

Theorem 4 is useful because we can estimate  $\mathbf{x}(0)$  via a projection onto convex set approach, which will be discussed in Section IV.

## IV. ALGORITHM FOR ESTIMATING AN INITIAL CONDITION

Projection onto convex set approach is widely used in the construction of signals [2], [5], [9]. To estimate an initial condition, a projection onto convex set approach is employed. The algorithm is as follows:

#### Algorithm

Step 1: Initialize  $\hat{\mathbf{x}}^0(0) \in \Phi$  and  $k=0$ .

Step 2: Solve the following optimization problem:

$$\min_{\bar{\mathbf{x}}^k(0) \in \Theta} \|\bar{\mathbf{x}}^k(0) - \hat{\mathbf{x}}^k(0)\|_2. \quad (4a)$$

This optimization problem is equivalent to the following optimization problem:

$$\min_{\bar{\mathbf{x}}^k(0)} \|\bar{\mathbf{x}}^k(0) - \hat{\mathbf{x}}^k(0)\|_2, \quad (4b)$$

subject to

$$(1-z^{-1})(\mathbf{I}-\mathbf{A}z^{-1})^{-1}(\bar{\mathbf{x}}^{(k)}(0)+z^{-1}\mathbf{B}(U(z)-S(z))) \text{ is analytical at } z=1. \quad (4c)$$

By Corollary 1, if  $\mathbf{A}$  has  $r$  unstable modes, denoted as  $\lambda_i$  for  $i=1,2,\dots,r$ , then  $\bar{\mathbf{x}}^{(k)}(0)+z^{-1}\mathbf{B}(U(\lambda_i)-S(\lambda_i))=\mathbf{0}$ . Hence, the constraint of this optimization problem becomes:

$$\bar{\mathbf{x}}^{(k)}(0)+z^{-1}\mathbf{B}(U(\lambda_i)-S(\lambda_i))=\mathbf{0} \text{ for } i=1,2,\dots,r \quad (4d)$$

and

$$\bar{\mathbf{x}}^{(k)}(0)+z^{-1}\mathbf{B}(U(z)-S(z)) \text{ is stable.} \quad (4e)$$

This problem is a standard quadratic programming problem with LMI constraints and a linear continuous constraint. This problem can be solved via the dual parameterization method [8] and it has a unique solution. Denote the solution as  $\bar{\mathbf{x}}^k(0)$ .

Step 3: Solve the following optimization problem:

$$\min_{\hat{\mathbf{x}}^{k+1}(0) \in \Phi} \|\hat{\mathbf{x}}^{k+1}(0) - \bar{\mathbf{x}}^k(0)\|_2. \quad (4f)$$

This optimization problem is equivalent to the following optimization problem:

$$\min_{\hat{\mathbf{x}}^{k+1}(0)} \|\hat{\mathbf{x}}^{k+1}(0) - \bar{\mathbf{x}}^k(0)\|_2, \quad (4g)$$

subject to

$$\begin{bmatrix} s(0)\mathbf{C} \\ s(1)\mathbf{C}\mathbf{A} \\ \vdots \\ s(k)\mathbf{C}\mathbf{A}^k \end{bmatrix} \hat{\mathbf{x}}^{(k+1)}(0) + \begin{bmatrix} 0 \\ s(1)\mathbf{C}\mathbf{B}(u(0)-s(0)) \\ \vdots \\ s(k)\mathbf{C} \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1} \mathbf{B}(u(n)-s(n)) \end{bmatrix} \geq \mathbf{0} \quad (4h)$$

for  $k \geq 1$ . This problem is also a standard quadratic programming problem with LMI constraints. There are many existing solvers for solving this problem and it has a unique solution. Denote the solution as  $\hat{\mathbf{x}}^{k+1}(0)$ .

Step 4: Iterative Step 2 and Step 3 until  $\|\hat{\mathbf{x}}^{k+1}(0) - \hat{\mathbf{x}}^k(0)\|_2 \leq \varepsilon$ ,

where  $\varepsilon$  is a prescribed acceptable error.

#### Corollary 3

When  $\mathbf{A}$  is unstable, if  $\Theta \cap \Phi \neq \emptyset$ , where  $\emptyset$  denotes the empty set, then the proposed Algorithm converges to the actual initial condition.

*Proof:*

Since both  $\Theta$  and  $\Phi$  are fixed for given quantizer output bit streams, they are convex sets, and the initial condition both generating bounded trajectories and corresponding to given quantizer output bit streams is uniquely defined when  $\mathbf{A}$  is unstable, the result follows directly and it completes the proof. ■

The importance of Corollary 3 is to provide a condition for the existence of an initial condition both generating bounded trajectories and corresponding to given quantizer output bit streams. When  $\mathbf{A}$  is unstable, if  $\Theta \cap \Phi \neq \emptyset$ , then the projection onto convex set approach can guarantee that the obtained solution is closed to the actual one.

The computational complexity of the proposed algorithm depends on that of solving the corresponding LMI problem and the linear continuous constraint optimization problem. In general, the computational complexity of solving LMI problem is much lower than that of the linear continuous constraint optimization problem. Hence, the analysis of the computational complexity of the proposed algorithm can be simplified by just considering that of the linear continuous constraint optimization problem. The detail analysis of the computational complexity of solving the linear continuous constraint optimization problem can be found in [14]. In [14], the index

set is constructed by adding only one of the most violated points in a refined set of grid points. Hence, the computational complexity is much reduced.

## V. NUMERICAL COMPUTER SIMULATION RESULTS

Since the interpolative SDM is nonlinear, the actual system is quite complicated. In particular, it is hard to understand why the state trajectories are bounded when the loop filter is unstable. To understand this phenomenon, the following example is used to illustrate and account for this phenomenon. Consider the following state space matrices, in which they are employed in audio systems [3]:

$$\mathbf{A} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -f_1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -f_2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad (5a)$$

$$\mathbf{B} \equiv [1 \ 0 \ 0 \ 0 \ 0]^T, \quad (5b)$$

and

$$\mathbf{C} \equiv [c_1 \ c_2 \ c_3 \ c_4 \ c_5], \quad (5c)$$

where  $f_1, f_2 \in \mathfrak{R}^+$  and  $c_i \in \mathfrak{R}$  for  $i=1,2,\dots,5$  are loop filter coefficients, in which  $f_1 \neq f_2 \cdot \forall t_{1,5}, t_{2,3}, t_{2,4}, t_{4,1}, t_{4,2} \in \mathbb{C} \setminus \{0\}$ , denote

$$\mathbf{T} \equiv \begin{bmatrix} 0 & 0 & 0 & 0 & t_{1,5} \\ 0 & 0 & t_{2,3} & t_{2,4} & 0 \\ 0 & 0 & -\frac{jt_{2,3}}{\sqrt{f_1}} & \frac{jt_{2,4}}{\sqrt{f_1}} & t_{1,5} \\ t_{4,1} & t_{4,2} & \frac{t_{2,3}}{f_2 - f_1} & \frac{t_{2,4}}{f_2 - f_1} & 0 \\ -\frac{jt_{4,1}}{\sqrt{f_2}} & \frac{jt_{4,2}}{\sqrt{f_2}} & -\frac{jt_{2,3}}{(f_2 - f_1)\sqrt{f_1}} & \frac{jt_{2,4}}{(f_2 - f_1)\sqrt{f_1}} & t_{1,5} \\ & & & & f_1 f_2 \end{bmatrix}. \quad (6a)$$

and

$$\mathbf{D} \equiv \text{diag}(1 + j\sqrt{f_2}, 1 - j\sqrt{f_2}, 1 + j\sqrt{f_1}, 1 - j\sqrt{f_1}, 1), \quad (6b)$$

where  $j \equiv \sqrt{-1}$  and  $\mathbb{C}$  denotes as the set of complex numbers.

Then,  $\mathbf{A}^k = \mathbf{T}\mathbf{D}^k\mathbf{T}^{-1} \ \forall k \geq 0$ , where

$$\mathbf{T}^{-1} = \begin{bmatrix} \frac{j}{2t_{4,1}(f_2 - f_1)\sqrt{f_2}} & \frac{1}{2t_{4,1}(f_1 - f_2)} & \frac{j\sqrt{f_2}}{2t_{4,1}(f_1 - f_2)} & \frac{1}{2t_{4,1}} & \frac{j\sqrt{f_2}}{2t_{4,1}} \\ -\frac{j}{2t_{4,2}(f_2 - f_1)\sqrt{f_2}} & \frac{1}{2t_{4,2}(f_1 - f_2)} & -\frac{j\sqrt{f_2}}{2t_{4,2}(f_1 - f_2)} & \frac{1}{2t_{4,2}} & -\frac{j\sqrt{f_2}}{2t_{4,2}} \\ -\frac{j}{2t_{2,3}\sqrt{f_1}} & \frac{1}{2t_{2,3}} & \frac{j\sqrt{f_1}}{2t_{2,3}} & 0 & 0 \\ \frac{j}{2t_{2,4}\sqrt{f_1}} & \frac{1}{2t_{2,4}} & -\frac{j\sqrt{f_1}}{2t_{2,4}} & 0 & 0 \\ \frac{1}{t_{1,5}} & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (6c)$$

Since the eigenvalues of  $\mathbf{A}$  are the diagonal elements of the matrix  $\mathbf{D}$ , the unstable modes being cancelled are the unstable diagonal elements of the matrix  $\mathbf{D}$ . By expanding (2), we have the following necessary and sufficient bounded conditions on state variables relating  $U(z) - S(z)$  to the initial condition:

$x_1(k)$  is bounded  $\forall k \geq 0$  if and only if  $\exists B \in \mathfrak{R}$  such that

$$\lim_{z \rightarrow 1} (U(z) - S(z)) = B, \quad (7a)$$

$x_2(k)$  is bounded  $\forall k \geq 0$  if and only if there exists a signal with z-transform  $P(z)$  analytically defined at  $z=1$  such that

$$U(z) - S(z) = \frac{\sqrt{f_1} z^2}{r \sin \theta} C_1(z) P(z) - \frac{\sqrt{f_1} z^2}{r \sin \theta} 2R(\cos \phi - rz^{-1} \cos(\theta - \phi)), \quad (7b)$$

$x_3(k)$  is bounded  $\forall k \geq 0$  if and only if there exists a signal with z-transform  $P'(z)$  analytically defined at  $z=1$  such that

$$U(z) - S(z) = \frac{f_1 P'(z) C_1(z)}{C_1(z) - z^{-1}(1 - z^{-1})(1 - r \cos \theta z^{-1})} - \frac{f_1 R'(1 - z^{-1})(\sin \phi' + rz^{-1} \sin(\theta - \phi'))}{C_1(z) - z^{-1}(1 - z^{-1})(1 - r \cos \theta z^{-1})}, \quad (7c)$$

$x_4(k)$  is bounded  $\forall k \geq 0$  if and only if there exists a signal with z-transform  $P''(z)$  analytically defined at  $z=1$  such that

$$U(z) - S(z) = \frac{C_1(z) C_2(z) P''(z)}{\frac{z^{-1}(1 - z^{-1}) \left( \frac{C_2(z) r \sin \theta}{\sqrt{f_1}} - \frac{C_1(z) r' \sin \theta'}{\sqrt{f_1}} \right)}{f_2 - f_1} - \frac{(1 - z^{-1}) C_1(z) R''(\sin \phi'' + r' z^{-1} \sin(\theta' - \phi''))}{f_2 - f_1} - \frac{z^{-1}(1 - z^{-1}) \left( \frac{C_2(z) r \sin \theta}{\sqrt{f_1}} - \frac{C_1(z) r' \sin \theta'}{\sqrt{f_1}} \right)}{f_2 - f_1} - \frac{(1 - z^{-1}) C_2(z) R'''(\sin \phi''' + rz^{-1} \sin(\theta - \phi'''))}{f_2 - f_1} - \frac{z^{-1}(1 - z^{-1}) \left( \frac{C_2(z) r \sin \theta}{\sqrt{f_1}} - \frac{C_1(z) r' \sin \theta'}{\sqrt{f_1}} \right)}{f_2 - f_1}}, \quad (7d)$$

$x_5(k)$  is bounded  $\forall k \geq 0$  if and only if there exists a signal with z-transform  $P'''(z)$  analytically defined at  $z=1$  such that

$$U(z) - S(z) = \frac{C_1(z) C_2(z) P'''(z)}{\frac{C_3(z)(1 - z^{-1})z^{-1}}{f_2 - f_1} + \frac{C_1(z) C_2(z)}{f_1 f_2} - \frac{(1 - z^{-1}) C_1(z) R''''(\sin \phi'''' + r' z^{-1} \sin(\theta' - \phi''''))}{f_2 - f_1} - \frac{C_3(z)(1 - z^{-1})z^{-1}}{f_2 - f_1} + \frac{C_1(z) C_2(z)}{f_1 f_2} - \frac{(1 - z^{-1}) C_2(z) R''''(\sin \phi'''' + rz^{-1} \sin(\theta - \phi''''))}{f_2 - f_1} + \frac{C_3(z)(1 - z^{-1})z^{-1}}{f_2 - f_1} + \frac{C_1(z) C_2(z)}{f_1 f_2}}, \quad (7e)$$

where

$$\mathbf{x}(k) \equiv [x_1(k) \ x_2(k) \ x_3(k) \ x_4(k) \ x_5(k)]^T, \quad (7f)$$

$$r \equiv \sqrt{1 + f_1}, \quad (7g)$$

$$r' \equiv \sqrt{1 + f_2}, \quad (7h)$$

$$\theta \equiv \tan^{-1}(\sqrt{f_1}), \quad (7i)$$

$$\theta' \equiv \tan^{-1}(\sqrt{f_2}), \quad (7j)$$

$$R \equiv \sqrt{\left(\frac{x_2(0)}{2}\right)^2 + \left(\frac{x_3(0)\sqrt{f_1} - \frac{x_1(0)}{\sqrt{f_1}}}{2}\right)^2}, \quad (7k)$$

$$R' \equiv \sqrt{\left(\frac{x_2(0)}{\sqrt{f_1}}\right)^2 + \left(x_3(0) - \frac{x_1(0)}{f_1}\right)^2}, \quad (7l)$$

$$R'' \equiv \sqrt{\left(\frac{x_2(0)}{f_1 - f_2} + x_4(0)\right)^2 + \left(\frac{\sqrt{f_2}x_3(0) - \frac{x_1(0)}{\sqrt{f_2}}}{f_2 - f_1} - \sqrt{f_2}x_5(0)\right)^2}, \quad (7m)$$

$$R''' \equiv \sqrt{\left(\frac{x_2(0)}{f_2 - f_1}\right)^2 + \left(\frac{\frac{x_1(0)}{\sqrt{f_1}} - \sqrt{f_1}x_3(0)}{f_2 - f_1}\right)^2}, \quad (7n)$$

$$R'''' \equiv \sqrt{\left(\frac{x_1(0)}{f_2(f_2 - f_1)} + \frac{x_3(0)}{f_1 - f_2} + x_5(0)\right)^2 + \left(\frac{x_2(0)}{f_1 - f_2} + x_4(0)\right)^2}, \quad (7o)$$

$$R'''' \equiv \sqrt{\left(\frac{x_3(0) - \frac{x_1(0)}{f_1}}{f_2 - f_1}\right)^2 + \left(\frac{x_2(0)}{\sqrt{f_1}(f_2 - f_1)}\right)^2}, \quad (7p)$$

$$\phi \equiv \tan^{-1} \left( \frac{x_3(0)\sqrt{f_1} - \frac{x_1(0)}{\sqrt{f_1}}}{x_2(0)} \right), \quad (7q)$$

$$\phi' \equiv \tan^{-1} \left( \frac{x_3(0) - \frac{x_1(0)}{f_1}}{\frac{x_2(0)}{\sqrt{f_1}}} \right), \quad (7r)$$

$$\phi'' \equiv \tan^{-1} \left( \frac{\frac{x_2(0)}{f_1 - f_2} + x_4(0)}{\frac{\sqrt{f_2}x_3(0) - \frac{x_1(0)}{\sqrt{f_2}}}{f_2 - f_1} - \sqrt{f_2}x_5(0)} \right), \quad (7s)$$

$$\phi''' \equiv \tan^{-1} \left( \frac{x_2(0)}{\frac{x_1(0)}{\sqrt{f_1}} - \sqrt{f_1}x_3(0)} \right), \quad (7t)$$

$$\phi'''' \equiv \tan^{-1} \left( \frac{\frac{x_1(0)}{f_2(f_2 - f_1)} + \frac{x_3(0)}{f_1 - f_2} + x_5(0)}{\frac{\frac{x_2(0)}{f_1 - f_2} + x_4(0)}{\sqrt{f_2}}} \right), \quad (7u)$$

$$\phi'''' \equiv \tan^{-1} \left( \frac{x_3(0) - \frac{x_1(0)}{f_1}}{\frac{x_2(0)}{\sqrt{f_1}}} \right), \quad (7v)$$

$$C_1(z) \equiv 1 - 2r \cos \theta z^{-1} + r^2 z^{-2}, \quad (7w)$$

$$C_2(z) \equiv 1 - 2r' \cos \theta' z^{-1} + r'^2 z^{-2}, \quad (7x)$$

and

$$C_3(z) \equiv \frac{C_1(z)(1 - r' \cos \theta' z^{-1})}{f_2} - \frac{C_2(z)(1 - r \cos \theta z^{-1})}{f_1}. \quad (7y)$$

If  $x_1(0) = x_2(0) = x_3(0) = 0$ , then  $x_2(k)$  is bounded  $\forall k \geq 0$  if and only if there exist two zeros of  $U(z) - S(z)$  located at  $re^{j\theta}$  and  $re^{-j\theta}$ , and  $x_3(k)$  is bounded  $\forall k \geq 0$  if and only if both  $x_1(k)$  and  $x_2(k)$  are bounded  $\forall k \geq 0$ . If  $\mathbf{x}(0) = \mathbf{0}$ , then  $x_4(k)$  is bounded  $\forall k \geq 0$  if and only if there exist four zeros of  $U(z) - S(z)$  located at  $re^{j\theta}$ ,  $re^{-j\theta}$ ,  $r'e^{j\theta'}$  and  $r'e^{-j\theta'}$ , and  $x_5(k)$  is bounded  $\forall k \geq 0$  if and only if both  $x_1(k)$  and  $x_4(k)$  are bounded  $\forall k \geq 0$ . It is worth noting that the zeros of  $U(z) - S(z)$  for bounded trajectories are in general not located at  $re^{j\theta}$ ,  $re^{-j\theta}$ ,  $r'e^{j\theta'}$  and  $r'e^{-j\theta'}$ , and this is true when  $\mathbf{x}(0) = \mathbf{0}$ . However,  $x_1(k)$  is bounded  $\forall k \geq 0$  if and only if the average value of quantizer output bit streams is equal to that of the input signal, that is,  $\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^{k-1} u(n) = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{n=0}^{k-1} s(n)$ . This

result does not directly depend on  $x_1(0)$ . Figure 2a and Figure 2b plot  $20 \log_{10} |U(z) - S(z)|$  against  $z = re^{j\omega}$  and  $z = r'e^{j\omega}$ , respectively, where  $\omega \in [\theta', \theta]$ ,  $\mathbf{x}(0) = \mathbf{0}$ ,  $f_1 = 0.0018$ ,  $f_2 = 0.000685$ ,  $c_1 = 0.8637566182$ ,  $c_2 = 0.3613814738$ ,  $c_3 = 0.090003709$ ,  $c_4 = 0.0132091570$ ,  $c_5 = 0.0009083750$  and  $u(k)$  being a normalized sum of sinusoidal signals within the signal band of the interpolative SDM. Here, we select

$$u(k) = \frac{\sum_{n=1}^{100} \sin\left(\frac{kn\pi}{100R}\right)}{\max_{\forall k \geq 0} \left( \sum_{n=1}^{100} \sin\left(\frac{kn\pi}{100R}\right) \right)}, \quad \text{in which } R = 64 \text{ for this}$$

interpolative SDM. A normalized sum of sinusoidal signals within the signal band is used for an illustration because it covers the whole signal band. 100 sinusoidal signals, instead of a single sinusoidal signal, are employed for an illustration because it can avoid the occurrence of limit cycles. According to the numerical computer simulations, it can be seen from Figure 2 that there are two zeros located at  $re^{j\theta}$  and  $r'e^{j\theta'}$ .

If the input is a rational step signal, then  $u(k)$  can be denoted as  $u(k) \equiv \bar{u}$  for  $k \geq 0$ , where  $\bar{u} \in \mathbb{Q}$ , in which  $\mathbb{Q}$  denotes the set of rational numbers.

*Corollary 4*

For a rational step input signal and the interpolative SDM defined by (5a)-(5c),  $x_1(k) - x_1(0)$  can only be an integer multiple of the reciprocal of the denominator of the input step size.

*Proof:*

As  $x_1(k) = x_1(0) + k\bar{u} - \sum_{n=0}^{k-1} Q(y(n))$  for  $k \geq 1$ ,  $\exists q_1 \in \mathbb{Z}$  and

$\exists q_2 \in \mathbb{Z}^+$  such that  $x_1(k) - x_1(0) = \frac{q_1}{q_2}$  for  $k \geq 1$ , where  $\mathbb{Z}$  and

$\mathbb{Z}^+$  denote the sets of integers and positive integers, respectively. This completes the proof. ■

Although it was reported in [11] that when  $\bar{u} \in \mathbb{Q}$ , then  $s(k)$  is periodic with the period being a multiple of the denominator of the input step size, this result is different from that stated in Corollary 4. This is because periodicity was studied in [11], while magnitude is studied in Corollary 4. In fact,  $x_1(k) - x_1(0)$  may be aperiodic. Also, the quantizer output bit streams were studied in [11], while the first state variable of the interpolative SDM is studied in Corollary 4.

In order to verify the effectiveness of the algorithm, the same filter ( $f_1 = 0.0018$ ,  $f_2 = 0.000685$ ,  $c_1 = 0.8637566182$ ,  $c_2 = 0.3613814738$ ,  $c_3 = 0.090003709$ ,  $c_4 = 0.0132091570$ ,  $c_5 = 0.0009083750$ ) and the input signal

$$u(k) = \frac{\sum_{n=1}^{100} \sin\left(\frac{kn\pi}{100R}\right)}{\max_{\forall k \geq 0} \left( \sum_{n=1}^{100} \sin\left(\frac{kn\pi}{100R}\right) \right)}, \text{ in which } R = 64 \text{ ) employed}$$

above are used for an illustration here. An initial condition is generated randomly with the first state variable being uniformly distributed between  $-0.1$  and  $0.1$  and the other state variables being uniformly distributed between  $-0.0001$  and  $0.0001$ . The first state variable has a larger variance than the others because it has larger stability margin. In the algorithm, we choose  $\varepsilon = 10^{-12}$  because it is small enough for most circuits and systems. Also, a random vector with the same distribution as the initial condition is generated and employed as the initialized vector for the algorithm. First, it is tested to see if it satisfied (3b) or not. If it is not satisfied, a new random vector is re-generated until (3b) is satisfied. Second, run Steps 2 to 4 of the proposed algorithm. Figures 3a-3e plot the actual state responses and Figure 3f plots the actual quantizer output bit streams. Figures 4a-4e plot the differences between the actual and new state responses using the estimated initial condition, while Figure 4f plots the difference between the actual and new quantizer output bit streams. It can be seen from Figure 4b and Figure 4c that the differences diverge transiently. This is because as  $\mathbf{A}$  is unstable, the interpolative SDM is chaotic. Although the 2-norm error between the actual and the

estimated initial condition is guaranteed to be bounded by  $\varepsilon$ ,  $\varepsilon \neq 0$  and a small deviation from the actual initial condition would cause very different state responses. However, the differences in the steady state responses are bounded because the constructed trajectories are guaranteed to be bounded. Also, there is no difference between the actual and the constructed quantizer output bit streams as shown in Figure 4f because the quantizer output bit streams is admissible. Figure 5a-5e plot the difference between the actual and new state responses using a random initial condition with zero mean and variance 0.0001. It can be seen from Figure 5a-5e that the transient differences are much more than that of our estimated initial condition. Also, there is a great difference between the actual and the new quantizer output bit streams, as shown in Figure 5f.

The run-time complexity of the algorithm depends on the solvers employed for solving the corresponding LMI problem and the linear continuous constraint optimization problem. In our numerical computer simulations, Matlab optimization toolbox is employed for solving these problems. Based on a PC with Pentium 1.2GHz CPU and 256M bytes DDRAM, the numerical computer simulation time is about 13 seconds.

## VI. CONCLUSIONS

In this paper, an initial condition of an interpolative SDM is estimated based on projection onto convex set approach. The set of initial conditions generating bounded trajectories is characterized and it is shown that a set of initial conditions generating bounded trajectories but not necessarily corresponding to quantizer output bit streams is convex. Also, it is shown that a set of initial conditions corresponding to quantizer output bit streams but not necessarily generating bounded trajectories is convex too. Moreover, an initial condition both generating bounded trajectories and corresponding to quantizer output bit streams is uniquely defined if the loop filter of an interpolative SDM is unstable. Hence, by using a projection onto convex set approach, the initial condition can be estimated. One of the advantages of the projection onto convex set approach is the guarantee of the convergence to the actual solution if the intersection of the above two convex sets is nonempty.

## ACKNOWLEDGEMENT

The work obtained in this paper was supported by a research grant from Queen Mary, University of London.

## REFERENCES

- [1] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling and Joshua D. Reiss, "Fuzzy impulsive control of high order interpolative lowpass sigma delta modulators," to appear in *IEEE Transactions on Circuits and Systems—I: Regular Papers*.
- [2] Søren Hein, "A fast block-based nonlinear decoding algorithm for  $\Sigma\Delta$  modulators," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1360-1367, June, 1995.
- [3] Derk Reefman and Erwin Janssen, "Signal processing for direct stream digital: a tutorial for digital sigma delta modulation and 1-bit digital audio processing," *Philips Research, Eindhoven, White Paper*, 2002. (<http://www.elec.qmul.ac.uk/people/josh/documents/ReefmanReissJanssenSandler-2003-AES115.pdf>)
- [4] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Joshua D. Reiss and

Xinghuo Yu, "Occurrence of elliptical fractal patterns in multi-bit bandpass sigma delta modulators," *International Journal of Bifurcation and Chaos*, vol. 15, no. 10, pp. 3377-3380, October, 2005.

[5] Nguyen T. Thao and Martin Vetterli, "Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates," *IEEE Transactions on Communications*, vol. 42, no. 3, pp. 519-531, March, 1994.

[6] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Joshua D. Reiss and Xinghuo Yu, "Nonlinear behaviors of bandpass sigma delta modulators with stable matrices," to appear in *IEEE Transactions on Circuits and Systems—II: Express Briefs*.

[7] Derk Reefman and Peter Nuijten, "Editing and switching in 1-bit audio streams," *Convention Paper of Audio Engineering Society, AES*, Paper Number 5399, Amsterdam, Netherlands, 12-15, May, 2001.

[8] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Joshua D. Reiss, Yan-Qun Liu and Kok-Lay Teo, "Design of interpolative sigma delta modulators via semi-infinite programming," to appear in *IEEE Transactions on Signal Processing*.

[9] Soren Hein and Avideh Zakhor, "Reconstruction of oversampled band-limited signals from  $\Sigma\Delta$  encoded binary sequences," *IEEE Transactions on Signal Processing*, vol. 42, no. 4, pp. 799-811, April, 1994.

[10] Derk Reefman, Joshua D. Reiss, Erwin Janssen and Mark Sandler, "Description of limit cycles in sigma-delta modulators," *IEEE Transactions on Circuits and Systems—I: Regular Papers*, vol. 52, no. 6, pp. 1211-1223, June, 2005.

[11] Vladimiiir Friedman, "The structure of the limit cycles in sigma-delta modulation," *IEEE Transactions on Communications*, vol. 36, no. 8, pp. 972-979, August, 1988.

[12] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling and Joshua D. Reiss, "Difference between irregular chaotic patterns of second-order double-loop  $\Sigma\Delta$  modulators and second-order interpolative bandpass  $\Sigma\Delta$  modulators," to appear in *Chaos, Solitons and Fractals*.

[13] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling and Joshua D. Reiss, "Stability of sinusoidal responses of marginally stable bandpass sigma delta modulators," to appear in *International Journal of Circuit Theory and Applications*.

[14] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Yan-Qun Liu, Peter Kwong-Shun Tam and Kok-Lay Teo, "Efficient algorithm for solving semi-infinite programming problems and their applications to nonuniform filter bank designs," to appear in *IEEE Transactions on Signal Processing*.

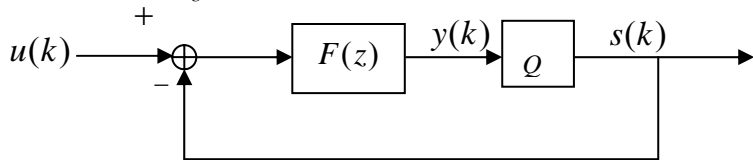


Figure 1. The block diagram of an interpolative SDM.

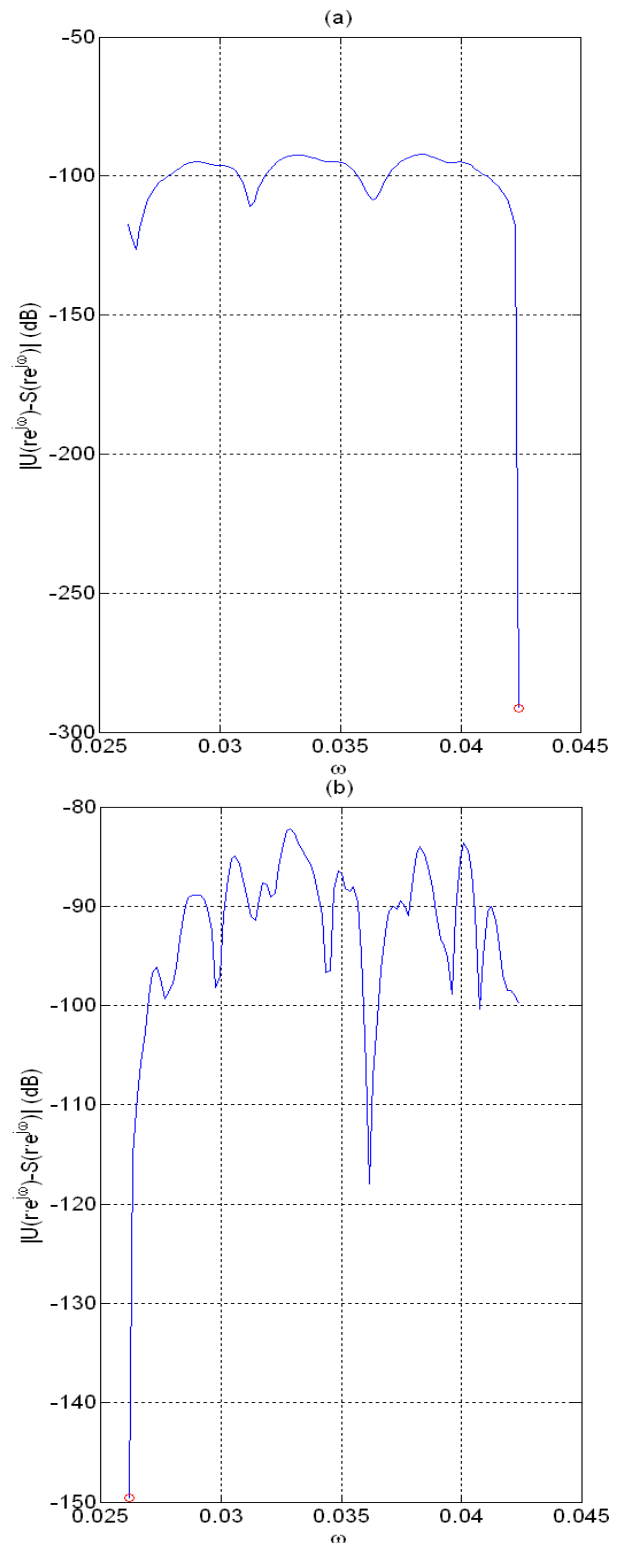


Figure 2. Plot of  $20 \log_{10} |U(z) - S(z)|$  against (a)  $z = re^{j\omega}$  and (b)  $z = r'e^{j\omega}$ , where  $\omega \in [\theta', \theta]$ .



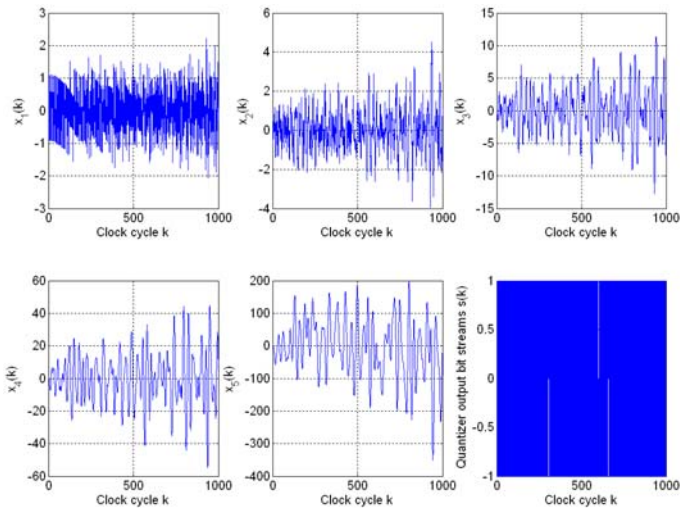


Figure 3. (a)-(e) Actual state responses. (f) Actual quantizer output bit streams.

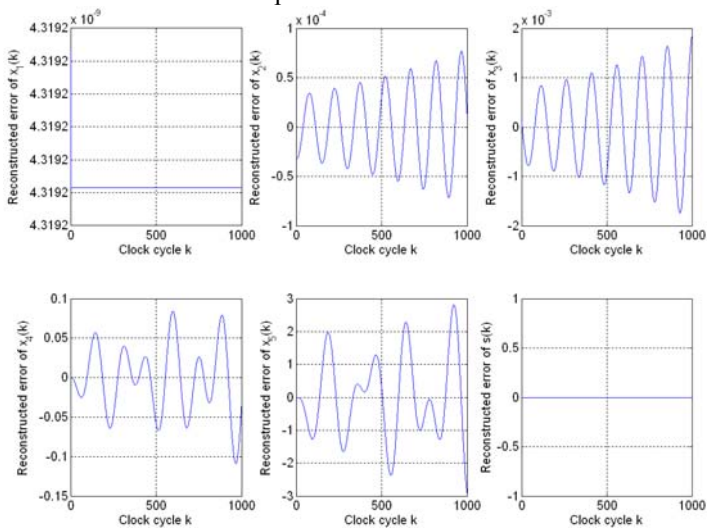


Figure 4. (a)-(e) Differences between the actual and new state responses using the estimated initial condition. (f) Difference between the actual and the new quantizer output bit streams.

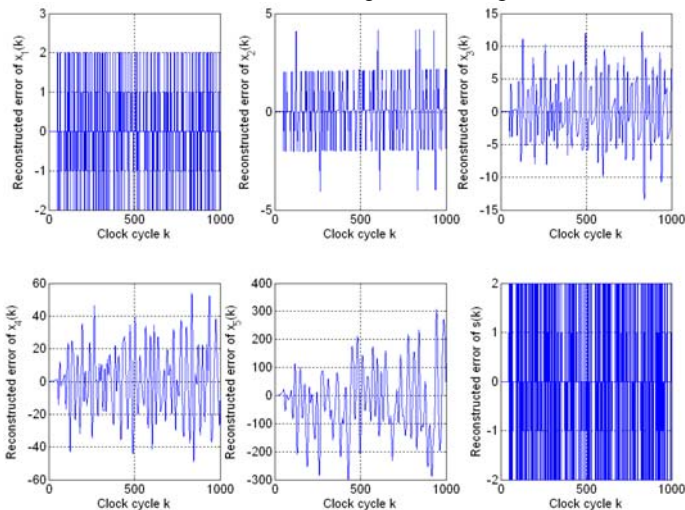


Figure 5. (a)-(e) Differences between the actual and new state responses using a random initial condition. (f) Difference between the actual and the new quantizer output bit streams.

