Audio Engineering Society

Convention Paper 6698

Presented at the 120th Convention
2006 May 20–23 Paris, France

Estimation of Initial States of Sigma-delta Modulators

Charlotte Yuk-Fan Ho[1], Bingo Wing-Kuen Ling[2], and Joshua D. Reiss[3]

[1] Department of Electronic Engineering, Queen Mary, University of London, Mile End Road, London, E1 4NS, United Kingdom.
c.ho@qmul.ac.uk

[2] Department of Electronic Engineering, King's College London, Strand, London, WC2R 2LS, United Kingdom.
wing-kuen.ling@kcl.ac.uk

[3] Department of Electronic Engineering, Queen Mary, University of London, Mile End Road, London, E1 4NS, United Kingdom.
josh.reiss@elec.qmul.ac.uk

ABSTRACT

In this paper, an initial condition of a sigma-delta modulator is estimated based on quantizer output bit streams and an input signal. The set of initial conditions that generate a stable trajectory is characterized. It is found that this set, as well as the set of initial conditions corresponding to the quantizer output bit streams, are convex. Also, it is found that the mapping from the set of initial conditions to the stable admissible set of quantizer output bit streams is invertible if the loop filter is unstable. Hence, the initial condition corresponding to given stable admissible quantizer output streams and an input signal is uniquely defined when the loop filter is unstable, and a projection onto convex set approach is employed for approximating the initial condition.

1. INTRODUCTION

Since some sigma-delta modulators (SDM) consist of a feedback loop, an unstable or a marginally stable loop filter (Actually, a marginally stable loop filter is bounded-input bounded-output *unstable* because resonance may occur for certain bounded inputs.) and a quantizer which is characterized by a discontinuous nonlinear function, the dynamics of an SDM could be very complicated. Chaotic and fractal behaviors may occur [1], [4], [6]. As chaotic behaviors are highly dependent on initial conditions [6], the dynamics of an SDM would be very different if there is a very small change in its initial conditions. When there is a sudden change of a supply voltage or a mechanical shaking, the content in the register containing an initial condition of an SDM may be corrupted. Since signals in SDMs are reconstructed based on an initial condition and an input signal, in this case, the reconstructed signal will be very different from the original one and a serious reconstruction error would be encountered.

In order to minimize the reconstruction error, it is necessary to estimate an initial condition of an SDM based on quantizer output bit streams and a given input signal. However, some fundamental questions have not been explored yet. For example, for a certain type of SDMs, such as SDMs with unstable loop filter and bounded loop filter output, does there exist a unique initial condition that corresponding to given quantizer output bit streams and input signal? If yes, how can we find an approximate initial condition which is closed to the actual one, and what are the significance of the error between the approximate initial condition and the actual one?

One of the most common methods for estimating an initial condition is to formulate the problem as an optimization problem. In [7], constraints were imposed so that the estimated initial condition is guaranteed to generate the corresponding quantizer output bit streams. However, the obtained solution does not guaranteed to generate a bounded trajectory. In this paper, necessary and sufficient bounded conditions are characterized and constraints based on these bounded conditions are imposed so that a bounded trajectory is also guaranteed.

The outline of this paper is as follows. In Section 2, notations used throughout this paper are introduced. In Section 3, necessary and sufficient bounded conditions of state variables are derived and it is shown that the set of initial conditions generating bounded trajectories is actually convex (A convex set is the set that all the points between any two points in the set are still in the set. [2], [5], [9]). In Section 4, it is shown that the set of initial conditions corresponding to given quantizer output bit streams and an input signal is also convex. Moreover, it is shown that the mapping from the set of initial conditions to the stable admissible set of quantizer bit streams is invertible if the loop filter is unstable. Hence, by projection onto these two convex sets, an initial condition of an SDM can be estimated. In Section 5, computer simulation results are presented to illustrate the effectiveness of the proposed method. Finally, a conclusion is summarized in Section 6.

2. NOTATIONS

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 the input of the interpolative SDM, the output of the loop filter, the output of the quantizer 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 state space equations:

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

$$y(k)\equiv\mathbf{C}\mathbf{x}(k), \tag{1b}$$

and

$$s(k)\equiv Q(y(k)), \tag{1c}$$

where

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

3. BOUNDED CONDITIONS OF STATE VARIABLES AND CONVEXITY OF THE CORRESPONDING SET OF INITIAL CONDITIONS

3.1. Bounded conditions of state variables

In some circuits and systems, such as audio systems [3], some eigenvalues of $\mathbf{A}$ are outside the unit circle. Hence, state variables of the interpolative SDM may be unbounded for any bounded inputs and initial conditions. To guarantee the state variables being bounded, define Γ as the set of initial conditions such that $\mathbf{x}(k)$ is bounded. Since

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

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

where $U(z)$ and $S(z)$ are denoted as z-transform of $u(n)$ and $s(n)$, respectively. $\mathbf{x}(k)$ is bounded 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 and initial conditions become the existence of a stable transfer function* $\mathbf{P}(z)$, *where*

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

The importance of this result is on the characterization of the set of initial conditions generating bounded trajectories for any bounded inputs and quantizer output bit streams. This result will be employed in our algorithm for estimating an initial condition of the interpolative SDM.

If $\mathbf{A}$ contains some unstable eigenvalues, then $\mathbf{P}(z)$ is stable 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. To illustrate this result, we consider the loop filter with the following state space matrices because this type of interpolative SDMs is employed in audio systems [3]:

$$\mathbf{A}\equiv\begin{bmatrix}1&0&0&0&0\\1&1&-f_1&0&0\\0&1&1&0&0\\0&0&1&1&-f_2\\0&0&0&1&1\end{bmatrix},\quad(3a)$$

$$\mathbf{B}\equiv\begin{bmatrix}1&0&0&0&0\end{bmatrix}^T,\quad(3b)$$

and

$$\mathbf{C}\equiv\begin{bmatrix}c_1&c_2&c_3&c_4&c_5\end{bmatrix},\quad(3c)$$

where $f_1,f_2\in\mathbb{R}^+$ and $c_i\in\mathbb{R}$ for $i=1,2,\cdots,5$ are filter coefficients, in which $f_1\neq f_2$, $\mathbb{R}$ and $\mathbb{R}^+$ denote the sets of real numbers and positive real numbers, respectively. $\forall t_{1,5},t_{2,3},t_{2,4},t_{4,1},t_{4,2}\in\mathbb{C}\backslash\{0\}$, denote

$$\mathbf{T}\equiv\begin{bmatrix}0&0&0&0&t_{1,5}\\0&0&t_{2,3}&t_{2,4}&0\\0&0&-\frac{jt_{2,3}}{\sqrt{f_1}}&\frac{jt_{2,4}}{\sqrt{f_1}}&\frac{t_{1,5}}{f_1}\\t_{4,1}&t_{4,2}&\frac{t_{2,3}}{f_2-f_1}&\frac{t_{2,4}}{f_2-f_1}&0\\-\frac{jt_{4,1}}{\sqrt{f_2}}&\frac{jt_{4,2}}{\sqrt{f_2}}&-\frac{jt_{2,3}}{(f_2-f_1)\sqrt{f_1}}&\frac{jt_{2,4}}{(f_2-f_1)\sqrt{f_1}}&\frac{t_{1,5}}{f_1f_2}\end{bmatrix},(4a)$$

and

$$\mathbf{D}\equiv\left(1+j\sqrt{f_2},1-j\sqrt{f_2},1+j\sqrt{f_1},1-j\sqrt{f_1},1\right),\quad(4b)$$

where $j\equiv\sqrt{-1}$ and $\mathbb{C}$ denotes the set of complex numbers. Then, $\mathbf{A}=\mathbf{TDT}^{-1}$, in which

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

By expanding (2) based on the interpolative SDMs described by (3a)-(3c), we have:

$x_1(k)$ is bounded if and only if $\exists B\in\mathbb{R}$ such that

$$\lim_{z\to1}(U(z)-S(z))=B,\quad(5a)$$

$x_2(k)$ is bounded if and only if there exists a stable transfer function $P(z)$ such that

$$\begin{aligned}U(z)-S(z)=&\frac{\sqrt{f_1}z^2}{r\sin\theta}C_1(z)P(z)\\&-\frac{\sqrt{f_1}z^2}{r\sin\theta}2R(\cos\phi-rz^{-1}\cos(\theta-\phi))\end{aligned},\quad(5b)$$

$x_3(k)$ is bounded if and only if there exists a stable transfer function $P'(z)$ such that

$$U(z)-S(z)=\frac{f_1P'(z)C_1(z)}{C_1(z)-z^{-1}(1-z^{-1})(1-r\cos\theta z^{-1})}-\frac{f_1R'(1-z^{-1})(\sin\phi'+rz^{-1}\sin(\theta-\phi'))}{C_1(z)-z^{-1}(1-z^{-1})(1-r\cos\theta z^{-1})}, \quad (5c)$$

$x_4(k)$ is bounded if and only if there exists a stable transfer function $P'''(z)$ such that

$$\begin{aligned}U(z)-S(z)=&\\ &\frac{C_1(z)C_2(z)P'''(z)}{\frac{z^{-1}(1-z^{-1})}{f_2-f_1}\left(\frac{C_2(z)r\sin\theta}{\sqrt{f_1}}-\frac{C_1(z)r'\sin\theta'}{\sqrt{f_1}}\right)}\\ &-\frac{(1-z^{-1})C_1(z)R''(\sin\phi''+r'z^{-1}\sin(\theta'-\phi''))}{\frac{z^{-1}(1-z^{-1})}{f_2-f_1}\left(\frac{C_2(z)r\sin\theta}{\sqrt{f_1}}-\frac{C_1(z)r'\sin\theta'}{\sqrt{f_1}}\right)}, \quad (5d)\\ &-\frac{(1-z^{-1})C_2(z)R'''(\sin\phi'''+rz^{-1}\sin(\theta-\phi'''))}{\frac{z^{-1}(1-z^{-1})}{f_2-f_1}\left(\frac{C_2(z)r\sin\theta}{\sqrt{f_1}}-\frac{C_1(z)r'\sin\theta'}{\sqrt{f_1}}\right)}\end{aligned}$$

$x_5(k)$ is bounded if and only if there exists a stable transfer function $P^{'''''}(z)$ such that

$$\begin{aligned}U(z)-S(z)=&\frac{C_1(z)C_2(z)P^{'''''}(z)}{\frac{C_3(z)(1-z^{-1})z^{-1}}{f_2-f_1}+\frac{C_1(z)C_2(z)}{f_1f_2}}\\ &-\frac{(1-z^{-1})C_1(z)R^{''''}(\sin\phi^{''''}+r'z^{-1}\sin(\theta'-\phi^{''''}))}{\frac{C_3(z)(1-z^{-1})z^{-1}}{f_2-f_1}+\frac{C_1(z)C_2(z)}{f_1f_2}}, \quad (5e)\\ &-\frac{(1-z^{-1})C_2(z)R^{'''''}(\sin\phi^{'''''}+rz^{-1}\sin(\theta-\phi^{'''''}))}{\frac{C_3(z)(1-z^{-1})z^{-1}}{f_2-f_1}+\frac{C_1(z)C_2(z)}{f_1f_2}}\end{aligned}$$

where

$$\mathbf{x}(k)\equiv[x_1(k)\quad x_2(k)\quad x_3(k)\quad x_4(k)\quad x_5(k)]^T, \quad (5f)$$

$$r\equiv\sqrt{1+f_1}\,, \quad (5g)$$

$$r'\equiv\sqrt{1+f_2}\,, \quad (5h)$$

$$\theta\equiv\tan^{-1}\left(\sqrt{f_1}\right), \quad (5i)$$

$$\theta'\equiv\tan^{-1}\left(\sqrt{f_2}\right), \quad (5j)$$

$$R\equiv\sqrt{\left(\frac{x_2(0)}{2}\right)^2+\left(\frac{x_3(0)\sqrt{f_1}-\frac{x_1(0)}{\sqrt{f_1}}}{2}\right)^2}\,, \quad (5k)$$

$$R'\equiv\sqrt{\left(\frac{x_2(0)}{\sqrt{f_1}}\right)^2+\left(x_3(0)-\frac{x_1(0)}{f_1}\right)^2}\,, \quad (5l)$$

$$R''\equiv\sqrt{\left(\frac{x_2(0)}{f_1-f_2}+x_4(0)\right)^2+\left(\frac{\sqrt{f_2}x_3(0)-\frac{x_1(0)}{\sqrt{f_2}}}{f_2-f_1}-\sqrt{f_2}x_5(0)\right)^2}, \quad (5m)$$

$$R'''\equiv\sqrt{\left(\frac{x_2(0)}{f_2-f_1}\right)^2+\left(\frac{\frac{x_1(0)}{\sqrt{f_1}}-\sqrt{f_1}x_3(0)}{f_2-f_1}\right)^2}\,, \quad (5n)$$

$$R^{''''}\equiv\sqrt{\left(\frac{x_1(0)}{f_2(f_2-f_1)}+\frac{x_3(0)}{f_1-f_2}+x_5(0)\right)^2+\left(\frac{\frac{x_2(0)}{f_1-f_2}+x_4(0)}{\sqrt{f_2}}\right)^2}, \quad (5o)$$

$$R^{'''''}\equiv\sqrt{\left(\frac{x_3(0)-\frac{x_1(0)}{f_1}}{f_2-f_1}\right)^2+\left(\frac{x_2(0)}{\sqrt{f_1}(f_2-f_1)}\right)^2}\,, \quad (5p)$$

$$\phi\equiv\tan^{-1}\left(\frac{x_3(0)\sqrt{f_1}-\frac{x_1(0)}{\sqrt{f_1}}}{x_2(0)}\right), \quad (5q)$$

$$\phi'\equiv\tan^{-1}\left(\frac{x_3(0)-\frac{x_1(0)}{f_1}}{\frac{x_2(0)}{\sqrt{f_1}}}\right), \quad (5r)$$

$$\phi'' \equiv \tan^{-1}\left(\frac{\frac{x_2(0)}{f_1-f_2}+x_4(0)}{\frac{\sqrt{f_2}x_3(0)-\frac{x_1(0)}{\sqrt{f_2}}}{f_2-f_1}-\sqrt{f_2}x_5(0)}\right), \tag{5s}$$

$$\phi''' \equiv \tan^{-1}\left(\frac{x_2(0)}{\frac{x_1(0)}{\sqrt{f_1}}-\sqrt{f_1}x_3(0)}\right), \tag{5t}$$

$$\phi'''' \equiv \tan^{-1}\left(\frac{\frac{x_1(0)}{f_2(f_2-f_1)}+\frac{x_3(0)}{f_1-f_2}+x_5(0)}{\frac{\frac{x_2(0)}{f_1-f_2}+x_4(0)}{\sqrt{f_2}}}\right), \tag{5u}$$

$$\phi''''' \equiv \tan^{-1}\left(\frac{x_3(0)-\frac{x_1(0)}{f_1}}{\frac{x_2(0)}{\sqrt{f_1}}}\right), \tag{5v}$$

$$C_1(z) \equiv 1-2r\cos\theta z^{-1}+r^2z^{-2}, \tag{5w}$$

$$C_2(z) \equiv 1-2r'\cos\theta' z^{-1}+r'^2z^{-2}, \tag{5x}$$

and

$$C_3(z) \equiv \frac{C_1(z)\left(1-r'\cos\theta' z^{-1}\right)}{f_2}-\frac{C_2(z)\left(1-r\cos\theta z^{-1}\right)}{f_1}. \tag{5y}$$

If $x_1(0)=x_2(0)=x_3(0)=0$, then $x_2(k)$ is bounded if and only if there exist two zeros of $U(z)-S(z)$ located at $re^{j\theta}$ and $re^{-j\theta}$, respectively, and $x_3(k)$ is bounded if and only if both $x_1(k)$ and $x_2(k)$ are bounded. If $\mathbf{x}(0)=\mathbf{0}$, then $x_4(k)$ is bounded 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 $r'e^{-j\theta''}$, respectively, and $x_5(k)$ is bounded if and only both $x_1(k)$ and $x_4(k)$ are bounded. It is worth noting that the zeros of $U(z)-S(z)$ for a bounded trajectory 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 only when $\mathbf{x}(0)=\mathbf{0}$. However, $x_1(k)$ is bounded if and only if the average value of quantizer output bit streams is equal to that of the input signal, that is, $\lim_{k\to+\infty}\frac{1}{k}\sum_{n=0}^{k-1}u(n)=\lim_{k\to+\infty}\frac{1}{k}\sum_{n=0}^{k-1}s(n)$. This result does not directly depend on $x_1(0)$. Figure 1a and Figure 1b plot $|U(z)-S(z)|$ 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 a random input signal with the amplitude bounded by 0.1, that is, $|u(k)|<0.1$ for $k\geq 0$. We choose this interpolative SDM with this random input for an illustration because this set of coefficients is employed in audio systems [3] and a random input with small amplitude can guarantee a bounded trajectory. Also, random inputs have wide frequency spectra, so it is a more general input signal compared to step or sinusoidal inputs. According to the simulation, it can be seen from the figure that the values of $\left|U\left(re^{j\theta}\right)-S\left(re^{j\theta}\right)\right|$ and $\left|U\left(r'e^{j\theta'}\right)-S\left(r'e^{j\theta'}\right)\right|$ are closed to zero, which implies that there are two zeros located at $re^{j\theta}$ and $r'e^{j\theta'}$, respectively.

3.2. Interesting behaviors of SDMs

If the input is a rational step signal, then we can denote $u(k)\equiv\bar{u}$ for $k\geq 0$, where $\bar{u}\in\mathbb{Q}$, in which $\mathbb{Q}$ denotes the set of rational numbers. For the interpolative SDMs defined by (3a)-(3c), 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. As a result, $x_1(k)-x_1(0)$ *can only be an integer multiple of the reciprocal of the denominator of the input step size.*

3.3. Convexity of the set of initial conditions corresponding to bound trajectories

If $\mathbf{x}(0)\in\Gamma$, then $\mathbf{x}(k)\in\Gamma$ for $k\geq 0$. Hence, the trajectory is confined within Γ. Suppose

$\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Gamma$, then there exist two stable transfer functions $\mathbf{P}^1(z)$ and $\mathbf{P}^2(z)$ where

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

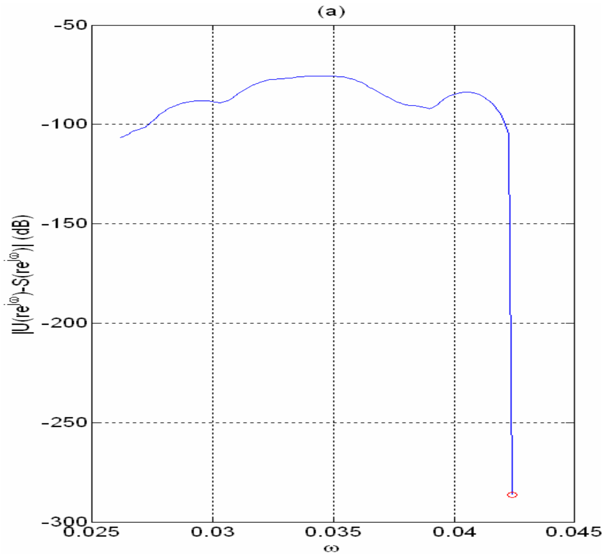
and

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

Since $\forall \lambda \in [0,1]$,

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

and $\lambda\mathbf{P}^1(z) + (1-\lambda)\mathbf{P}^2(z)$ is a stable transfer function, this implies that $\lambda\mathbf{x}^1(0) + (1-\lambda)\mathbf{x}^2(0) \in \Gamma$. Hence, Γ *is a convex set.* This result is useful because we can estimate $\mathbf{x}(0)$ via a projection onto convex set approach.



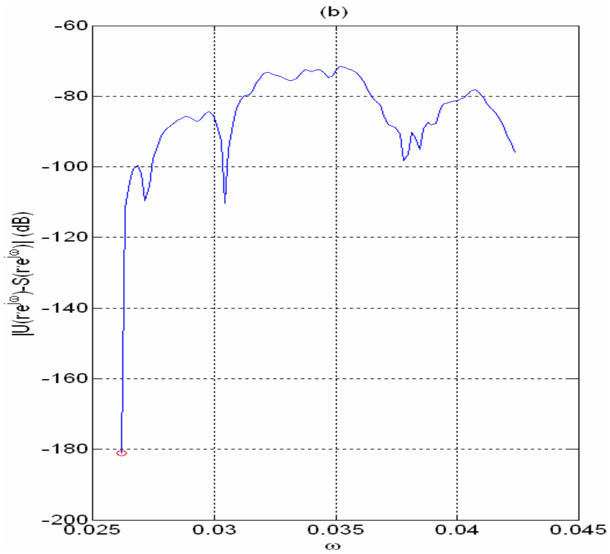


Figure 1. Plot of $|U(z) - S(z)|$ against (a) $z = re^{j\omega}$ and (b) $z = r'e^{j\omega}$, where $\omega \in [\theta', \theta]$.

4. CONVEXITY OF ADMISSIBLE SET OF INITAIL CONDITONS AND INVERTIBILITY OF A MAPPING

4.1. Convexity of the set of initial conditions corresponding to quantizer output bit streams

Denote an infinite length binary sequence with each element in the sequence being either 1 or -1 and their corresponding set as $\mathbf{s} = (s(0), s(1), \cdots)$ and Ψ, respectively. Define the mapping from Γ to Ψ as Λ such that (1a)-(1d) are satisfied. The set of quantizer output bit streams is said to be stable and admissible if $\forall \mathbf{s} \in \Psi$, $\exists \mathbf{x}(0) \in \Gamma$ such that $\Lambda(\mathbf{x}(0)) = \mathbf{s}$. It is worth noting that Ψ is not necessary a stable admissible set because there may not exist $\mathbf{x}(0) \in \Gamma$ such that (1a)-(1d) are satisfied. To characterize the admissible condition, the approach in [7] is employed and summarized below. Since $s(k) = 1$ if $y(k) \geq 0$ and $s(k) = -1$ if $y(k) < 0$ for $k \geq 0$, we have $s(k)y(k) \geq 0$ for $k \geq 0$, that is, $s(0)\mathbf{C}\mathbf{x}(0) \geq 0$ and

$$s(k)\left(\mathbf{C}\mathbf{A}^k\mathbf{x}(0) + \mathbf{C}\sum_{n=0}^{k-1}\mathbf{A}^{k-n-1}\mathbf{B}\left(u(n) - s(n)\right)\right) \geq 0 \qquad (6a)$$

for $k \geq 1$. Denote the set of initial conditions that satisfies (6a) as Φ. Then the stable admissible set of quantizer output bit streams is

$$\Psi_b = \left\{ \begin{array}{l} \mathbf{s} : s(0) = Q(\mathbf{C}\mathbf{x}(0)), \\ s(k) = \\ Q\left(\mathbf{C}\mathbf{A}^k \mathbf{x}(0) + \mathbf{C}\sum_{n=0}^{k-1} \mathbf{A}^{k-n-1}\mathbf{B}(u(n) - s(n)) \right) \\ \text{for } k \geq 1 \text{ and } \mathbf{x}(0) \in \Gamma \end{array} \right\}. \quad (6b)$$

Hence,

$\mathbf{x}(0) \in \Phi$ if and only if

$$\begin{bmatrix} s(0)\mathbf{C} \\ s(1)\mathbf{C}\mathbf{A} \\ \vdots \\ s(k)\mathbf{C}\mathbf{A}^k \end{bmatrix} \mathbf{x}(0) + \begin{bmatrix} 0 \\ s(1)\mathbf{C}\mathbf{B}(u(0) - s(0)) \\ \vdots \\ s(k)\mathbf{C}\sum_{n=0}^{k-1} \mathbf{A}^{k-n-1}\mathbf{B}(u(n) - s(n)) \end{bmatrix} \geq \mathbf{0} \quad (6c)$$

for $k \geq 1$. The importance of this result is that the set of initial conditions generating a given quantizer output bit streams for a given input can be characterized. This result will be employed in our proposed algorithm.

Besides, 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 Φ *is a convex set.*

This result is useful because we can estimate an initial condition based on a projection onto convex set approach. However, it is worth noting that $\mathbf{x}(0) \in \Phi$ does not imply that $\mathbf{x}(0) \in \Gamma$, that means initial conditions corresponding to quantizer output bit streams may cause the output of the loop filter unbounded.

4.2. Invertibility of the mapping from Γ∩Φ to Ψ_b

It is worth to know whether there is a unique initial condition corresponding to a bounded loop filter output, given quantizer output bit streams and input signal. To address this problem, define a mapping Λ_b from the set of initial conditions to the stable admissible set of quantizer output bit streams, that is $\Lambda_b : \Gamma \cap \Phi \to \Psi_b$. Suppose $\mathbf{x}^1(0), \mathbf{x}^2(0) \in \Gamma \cap \Phi$ and $\mathbf{x}^1(0) \neq \mathbf{x}^2(0)$ such that $\Lambda_b(\mathbf{x}^1(0)) = \Lambda_b(\mathbf{x}^2(0)) = \mathbf{s}$, then

$$\begin{array}{l} \mathbf{x}^1(k) = \mathbf{A}^k \mathbf{x}^1(0) + \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1}\mathbf{B}(u(n) - s(n)) \\ \mathbf{x}^2(k) = \mathbf{A}^k \mathbf{x}^2(0) + \sum_{n=0}^{k-1} \mathbf{A}^{k-n-1}\mathbf{B}(u(n) - s(n)) \end{array} \quad \text{for } k \geq 1,$$

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, $\mathbf{x}^1(0) = \mathbf{x}^2(0)$, which implies that *if* $\mathbf{A}$ *is unstable, then* Λ_b *is invertible.*

The importance of this result is to guarantee that the initial condition corresponding to a bounded loop filter output, given quantizer output bit streams and input signal is uniquely defined if $\mathbf{A}$ is unstable.

4.3. Algorithm for estimating the initial condition

To estimate the 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:

$$\min_{\overline{\mathbf{x}}^k(0) \in \Gamma} \left\| \overline{\mathbf{x}}^k(0) - \hat{\mathbf{x}}^k(0) \right\|_2 . \quad (7a)$$

This optimization problem is equivalent to the following optimization problem:

$$\min_{\bar{\mathbf{x}}^k(0)} \left\| \bar{\mathbf{x}}^k(0) - \hat{\mathbf{x}}^k(0) \right\|_2 , \tag{7b}$$

subject to

$$\left(1 - z^{-1}\right)\left(\mathbf{I} - \mathbf{A}z^{-1}\right)^{-1}\left(\bar{\mathbf{x}}^{(k)}(0) + z^{-1}\mathbf{B}\left(U(z) - S(z)\right)\right) \text{ is stable.} \tag{7c}$$

This problem is a standard convex control problem and a standard control technique [8] can be applied for solving the problem. Denote the solution as $\bar{\mathbf{x}}^k(0)$.

Step 3: Solve the following optimization problem:

$$\min_{\hat{\mathbf{x}}^{k+1}(0) \in \Phi} \left\| \hat{\mathbf{x}}^{k+1}(0) - \bar{\mathbf{x}}^k(0) \right\|_2 . \tag{7d}$$

This optimization problem is equivalent to the following optimization problem:

$$\min_{\hat{\mathbf{x}}^{k+1}(0)} \left\| \hat{\mathbf{x}}^{k+1}(0) - \bar{\mathbf{x}}^k(0) \right\|_2 , \tag{7e}$$

subject to

$$\begin{bmatrix} s(0)\mathbf{C} \\ s(1)\mathbf{CA} \\ \vdots \\ s(k)\mathbf{CA}^k \end{bmatrix} \hat{\mathbf{x}}^{(k+1)}(0) + \begin{bmatrix} 0 \\ s(1)\mathbf{CB}\left(u(0) - s(0)\right) \\ \vdots \\ s(k)\mathbf{C}\sum_{n=0}^{k-1} \mathbf{A}^{k-n-1}\mathbf{B}\left(u(n) - s(n)\right) \end{bmatrix} \geq \mathbf{0} \tag{7f}$$

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

Step 4: Iterative Steps 2 and 3 until $\left\| \hat{\mathbf{x}}^{k+1}(0) - \hat{\mathbf{x}}^k(0) \right\|_2 \leq \varepsilon$, where ε is a prescribed acceptable error.

It is worth noting that the proposed Algorithm guarantees to converge to the actual initial condition if $\Gamma \cap \Phi \neq \emptyset$, where Ø denotes the empty set, because both Γ and Φ are convex sets and the initial condition corresponding to a bounded loop filter output, given quantizer output bit streams and input signal is uniquely defined when $\mathbf{A}$ is unstable.

5. COMPUTER SIMULATION RESULTS

In order to verify the effectiveness of the proposed algorithm, the same filter and same type of random input in Section 3.1 are used for an illustration. 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 our proposed 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 our proposed algorithm. First, it is tested to see if it satisfied (6c) or not. If it is not satisfied, a new random vector is re-generated until (6c) is satisfied. Second, run Steps 2 to 4 of our proposed algorithm. Figures 2a-2e plot the original state responses and Figure 2f plots the original quantizer output bit streams. Figures 3a-3e plot the differences between the original and new state responses using the estimated initial condition, while Figure 3f plots the difference between the original and the reconstructed quantizer output bit streams. It can be seen from Figure 3c and Figure 3d that the differences diverge transiently. This is because as $\mathbf{A}$ is unstable, the SDM is chaotic. Although the 2-norm error between the original and the estimated initial condition is guaranteed to be bounded by ε, $\varepsilon \neq 0$ and the small deviation from the actual initial condition would cause very different state responses. However, there is no difference between the original and the reconstructed quantizer output bit streams as shown in Figure 3f. Figure 4a-4e plot the difference between the original and new state responses using a random initial condition with zero mean and variance 0.0001. It can be seen from Figure 4a-4e that the transient differences are much more than that of using the estimated initial condition. Also, there is a great difference between the original and the new quantizer output bit streams, as shown in Figure 4f.

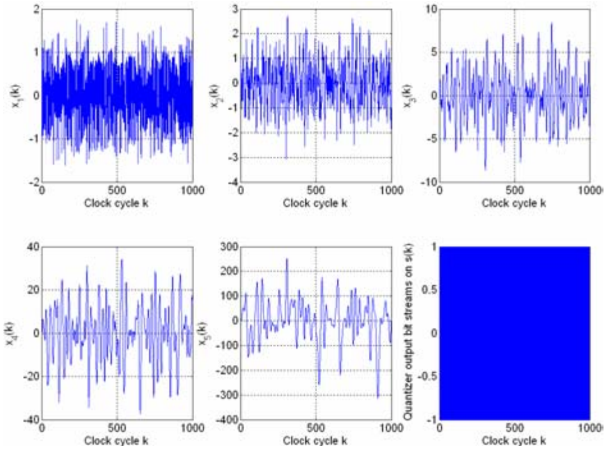


Figure 2. (a)-(e) Original state responses. (f) Original quantizer output bit streams.

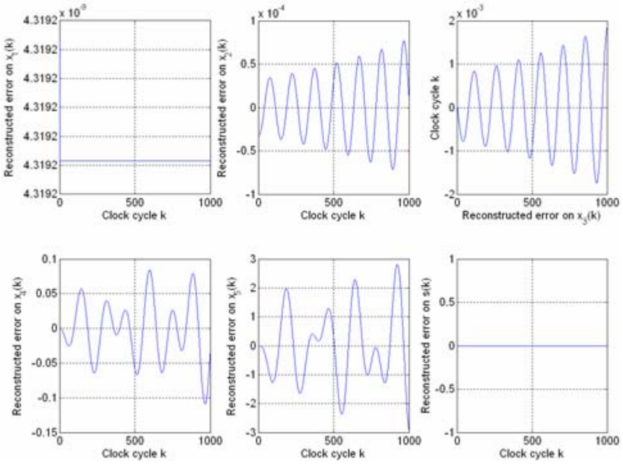


Figure 3. (a)-(e) Differences between the original and new state responses using the estimated initial condition. (f) Difference between the original and the new quantizer output bit streams.

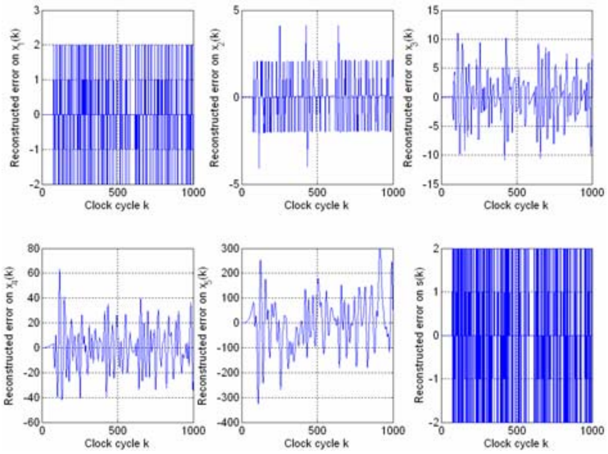


Figure 4. (a)-(e) Differences between the original and new state responses using a random initial condition. (f) Difference between the original and the new quantizer output bit streams.

6. CONCLUSION

In this paper, an initial condition of the interpolative SDM is estimated based on projection onto convex set approach. The set of initial conditions that generating a bounded trajectory is characterized and it is shown that the set is convex. Also, we show that the set of initial conditions corresponding to quantizer output bit streams is convex too. Moreover, the mapping from the set of initial conditions to the stable admissible set of quantizer output bit streams is invertible if the loop filter is unstable. Hence, by using a projection onto convex set approach, the initial condition can be estimated. One of the advantages of the proposed method is the guarantee of the convergence of the unique solution if the intersection of these two convex sets is non-empty.

7. ACKNOWLEDGEMENTS

This work was supported by a research grant from Queen Mary, University of London.

8. 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 ΣΔ modulators," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1360-1367, 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.

[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, 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, 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] Michael Rotkowitz and Sanjay Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.

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