

Design of Interpolative Sigma Delta Modulators Via Semi-Infinite Programming

Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Joshua D. Reiss, Yan-Qun Liu, and Kok-Lay Teo

Abstract—This correspondence considers the optimized design of interpolative sigma delta modulators (SDMs). The first optimization problem is to determine the denominator coefficients. The objective of the optimization problem is to minimize the passband energy of the denominator of the loop filter transfer function (excluding the dc poles) subject to the continuous constraint of this function defined in the frequency domain. The second optimization problem is to determine the numerator coefficients in which the cost function is to minimize the stopband ripple energy of the loop filter subject to the stability condition of the noise transfer function (NTF) and signal transfer function (STF). These two optimization problems are actually quadratic semi-infinite programming (SIP) problems. By employing the dual-parameterization method, global optimal solutions that satisfy the corresponding continuous constraints are guaranteed if the filter length is long enough. The advantages of this formulation are the guarantee of the stability of the transfer functions, applicability to design of rational infinite-impulse-response (IIR) filters without imposing specific filter structures, and the avoidance of iterative design of numerator and denominator coefficients. Our simulation results show that this design yields a significant improvement in the signal-to-noise ratio (SNR) and have a larger stability range, compared with the existing designs.

Index Terms—Dual parameterization, interpolative sigma delta modulators (SDMs), noise shaping, semi-infinite programming (SIP), stability.

I. INTRODUCTION

Sigma delta modulation is a popular form of analog-to-digital (A/D) and digital-to-analog (D/A) conversion and is applied in most commercial A/D and D/A systems [1]–[4]. The popularity of sigma delta modulators (SDMs) is mainly due to their simple, inexpensive, and robust circuit implementation, as well as achieving very high signal-to-noise ratio (SNR) because of their ability to perform noise shaping [5].

The basic operation of SDMs is to sample the input signal at a much higher rate than the Nyquist frequency, filter the signal and perform noise shaping, and then quantize the output [5]. The block diagram of an interpolative (or feedforward) SDM is depicted in Fig. 1. It consists of a loop filter, a low-bit quantizer, and a negative feedback path. Over-sampling of the input signal and noise shaping in the loop filter is used in order to remove the quantization noise out of the passband, typically the lowpass band [5].

Optimal designs have been performed based on optimizing operational transconductance amplifier structures [6], speed, resolution, and A/D complexity [7] and the ratio of peak SNR plus distortion ratio

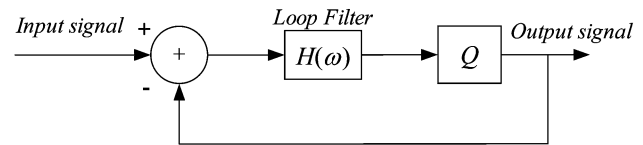


Fig. 1. Block diagram of an interpolative SDM as used for A/D conversion $|1 + (\eta_N(\omega))^T \mathbf{x}_a|^2$.

versus power consumption [8], etc. Although these designs have considered many practical issues, the solutions obtained are not global optimal because the optimization problems involved are not convex. SDMs are typically designed using Butterworth filter design rules [2], and optimal SDM designs based on comb filter [9] and Laguerre filter [10] structures were recently proposed. However, since some structures (such as all the poles of the Laguerre filters are constrained to be the same, that of the Butterworth filters are constrained on the same circle, and many zeros are in the impulse response of the comb filter) are assumed on these filters, a better solution may be obtained if these structural assumptions are relaxed. Design based on the finite horizon method [11] was also proposed. However, this method is only an approximation of an infinite horizon method. Although the approximation is improved as the length of window increases, the computational complexity increases. Genetic algorithms have also been applied to perform the optimization [12]. However, the convergence of the genetic algorithms is not guaranteed and the computational complexity of this method is very high. Other existing optimal designs, such as reported in [13]–[15], are obtained mainly based on the simulation framework and lack of the theoretical support.

Since computational complexity of rational infinite-impulse-response (IIR) filters is usually lower than that of the finite-impulse-response (FIR) filters, many SDMs employ rational IIR filters [1]–[3]. However, since rational IIR filters consist of both numerator and denominator coefficients, there are some challenges for designing rational IIR filters. One way to design rational IIR filters is to first initialize a set of the denominator coefficients and then design the numerator coefficients based on this set of initialized denominator coefficients by using ripple energy as the cost function and magnitude specification as the constraints. Then, update the denominator coefficients based on the obtained numerator coefficients and iterate this procedure until the algorithm converges. However, it is not guaranteed that the iterative procedure will converge [16]. Moreover, the obtained solution depends on the initialization of the denominator coefficients, hence only a local optimal solution is obtained. Although the divergence problem can be solved via weighting the filter coefficients in each iteration, the frequency characteristics of the filter will depend on the weights and the result obtained may be degraded [17]. Furthermore, these design methods [16], [17] assume that both the desired magnitude and phase responses of the filter are known. However, sometimes it may be difficult to characterize the desired phase response. This applies to Butterworth and Laguerre filter cases because these are nonlinear phase filters. Under this circumstance, the cost function based on the error energy or the absolute error between the desired and designed energy responses will become a fourth-order function $(|H(\omega)|^2 - |H_d(\omega)|^2)^2$ or a nonsmooth function $(|H(\omega)|^2 - |H_d(\omega)|^2)$ ($H(\omega)$ and $H_d(\omega)$ denotes the designed and desired frequency response, respectively). Nevertheless, these problems are not convex.

The major issues of designing SDMs are to achieve high SNR with the guarantee of the boundedness of state variables [2], [10]. Since the SDMs consist of a quantizer, which is a nonlinear component, there is no simple relationship among the SNR [18], maximum bound of the

Manuscript received April 25, 2005; revised September 20, 2005. The associate editor coordinating the review of this paper and approving it for publication was Dr. Kenneth E. Barner.

C. Y.-F. Ho, J. D. Reiss are with the Department of Electronic Engineering, Queen Mary, University of London, London E1 4NS, U.K. (e-mail: charlotte.ho@elec.qmul.ac.uk; josh.reiss@elec.qmul.ac.uk).

B. W.-K. Ling is with the Department of Electronic Engineering, Division of Engineering, King's College London, London WC2R 2LS, U.K. (e-mail: wing-kuen.ling@kcl.ac.uk).

Y.-Q. Liu is with the Department of Mathematics and Statistics, Royal Melbourne Institute of Technology, Melbourne VIC 3001, Australia (e-mail: yanqun.liu@rmit.edu.au).

K.-L. Teo is with the Department of Mathematics and Statistics, Curtin University of Technology, Perth, 00301J, Australia (e-mail: K.L.Teo@curtin.edu.au).

Color versions of Figs. 2–4 are available online at <http://ieeexplore.ieee.org>. Digital Object Identifier 10.1109/TSP.2006.880338

input signal [19], and the filter parameters, particular when the filter order is high. Hence, it is typical to achieve high SNR by achieving good responses of both signal transfer function (STF) and noise transfer function (NTF) [10] and to achieve the boundedness of the state variables by achieving the stability conditions of STF and NTF [20]. The objective of this correspondence is to formulate an SDM design problem as optimization problems based on the characteristics of the STF and NTF, the stopband characteristics of loop filters, and the stability conditions of the STF and NTF. In order to achieve the stability conditions, the sum of the numerator and denominator polynomials of the loop filter transfer function has to be on the right-hand side of the complex plane for all frequencies [16], [17].

Since the noise-shaping characteristics, as well as the filter characteristics, are defined in the frequency domain, all the constraints are continuous. Hence, the optimization problems are actually quadratic SIP problems. Since the solution is required to satisfy the constraint for all frequencies, simple methods for solving finite number of discrete constraint problems do not apply. The most common methods for solving SIP problems are discretization methods, local reduction methods, dual-exchange methods, nondifferentiable optimization approaches, and interior point methods [21]. For the discretization methods, it is not guaranteed that the continuous constraints are satisfied among the discretized points. Although the difference between the exact upper bound of discretized constraints and that of the corresponding continuous constraints vanishes as the number of grid points increases, the computational complexity increases. For the local reduction methods, they require a good initial guess of a solution sufficiently close to the optimal solution in order to ensure its local convergence. For the dual-exchange methods, they may have numerical instabilities. For the nondifferentiable optimization approaches, they are not efficient to solve smooth problems. For the interior point methods, they are not applicable if the number of constraints tends to infinity. To solve SIP problems, the dual-parameterization method [21] is applied. It is found that this method is very efficient and effective for designing FIR filters [22], [23]. However, the problems formulated in [22] and [23] only involve linear continuous constraints, and the filters are FIR. In this correspondence, an IIR filter is designed, and the constraint is a quadratic continuous function.

The dual-parameterization method is to parameterize the measure in the dual problems so that it transforms the SIP problems into equivalent finite-dimensional nonlinear programming problems via sequences of regular convex programs. The basic working principle of the dual-parameterization method is as follows: Since the constraint functions are continuous with respect to their index parameters and the index set is compact Hausdorff [21], the constraint functions can be redefined as an operator whose range is the Banach space consisting of continuous functions defined on the index set and equipped with the uniform norm. The order in the range space is given by a cone consisting of all non-negative functions on the index set. The assumption of the dual-parameterization method is the existence of a solution that strictly satisfies the continuous constraints. This condition is also known as the Slater's condition. It is worth noting that the Slater's condition can be satisfied if the filter length is long enough. Once the Slater's condition is satisfied, the Karush–Kuhn–Tucker (KKT) conditions [21] would be satisfied, which guarantees a necessary optimality condition for such a cone-constrained nonlinear programming problem, where the Lagrange multiplier is defined as a regular Borel measure [21] on the index set. As a result, the set of multipliers satisfying the KKT condition necessarily includes a measure with finite support unless it is empty. Hence, any constraint qualification ensures the existence of such a discrete measure, which is also called the Haar measure. On the other hand, strong duality holds for convex programming under Slater's constraint qualification. Hence, the corresponding dual problem for SIP can then be formulated in the space of finite signed regular Borel measures on the

index set. The local KKT theory and the global duality theory are naturally related through the fact that the set of multipliers satisfying the KKT condition coincides with the set of solutions to the dual problem, which leads to the consequence that the set of dual solutions always includes a measure with finite support under the Slater condition. Hence, the dual-parameterization method is guaranteed to obtain a global optimal solution that satisfies the continuous constraint if the filter length is long enough.

For the implementation of the dual-parameterization method, first initialize a sequence of index set, then compute a local optimal solution by solving a finite dimensional nonlinear programming problem. Finally, compute the global optimal solution via a local search for the finite dual problem. For the details of the theory and the implementation of the dual-parameterization method, it can be found in [21].

According to the simulations, the SDM produces a higher SNR and has a higher stability range compared with the existing designs. The outline of this correspondence is as follows. The problem formulation is presented in Section II. The simulation results are shown in Section III. Finally, a conclusion is summarized in Section IV.

II. PROBLEM FORMULATION

For practical reasons, it is easier to realize the SDMs if all the filter coefficients are real and the transfer function of loop filters is rational, causal, and with a unit sample delay in the numerator [1]. Moreover, since we only consider the lowpass SDMs [1], there is usually at least one dc pole in the transfer function of the loop filters. The frequency response of the loop filters is assumed to be as follows:

$$H(\omega) = \frac{e^{-j\omega} \sum_{m=0}^M b_m e^{-jm\omega}}{(1 - e^{-j\omega})^r \left(1 + \sum_{n=1}^N a_n e^{-jn\omega}\right)} \quad (1)$$

where M and N are the numbers of roots of the polynomial of $e^{-j\omega}$ in the numerator and denominator of the transfer function of the loop filter (excluding the dc poles and pure delay elements), respectively, r is the number of dc poles, a_n, b_m for $n = 1, 2, \dots, N$, and $m = 0, 1, \dots, M$ are the filter coefficients. In our consideration, $a_n, b_m \in \mathbb{R}$, $r \geq 1$ and $N + r \geq M + 1$. The design problem is equivalent to finding an appropriate set of filter coefficients a_n and b_m . However, our design method can still be applied to the cases when the IIR filter is not causal or there is no dc pole in the transfer function.

By grouping the filter coefficients in the numerator and denominator as

$$\mathbf{x}_b \equiv [b_0, \dots, b_M]^T \quad \text{and} \quad \mathbf{x}_a \equiv [a_1, \dots, a_N]^T \quad (2)$$

respectively, where the superscript T denotes the transpose, and defining

$$\boldsymbol{\eta}_M(\omega) \equiv [1, e^{-j\omega}, \dots, e^{-jM\omega}]^T \quad (3)$$

and

$$\boldsymbol{\eta}_N(\omega) \equiv [e^{-j\omega}, e^{-j2\omega}, \dots, e^{-jN\omega}]^T \quad (4)$$

then

$$H(\omega) = \frac{e^{-j\omega} (\boldsymbol{\eta}_M(\omega))^T \mathbf{x}_b}{(1 - e^{-j\omega})^r \left(1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a\right)}. \quad (5)$$

The STF and NTF of the SDM can be expressed as

$$\text{STF}(\omega) = \frac{H(\omega)}{1 + H(\omega)} \quad \text{and} \quad \text{NTF}(\omega) = \frac{1}{1 + H(\omega)} \quad (6)$$

respectively.

A. Determination of Denominator Coefficients

Denote the passband of the loop filter as B_P , which is also the band of interest. For SDMs having a good SNR, the *magnitude* of the STF should be approximately equal to 1, and that of the NTF should be approximately equal to 0 for all frequencies in the passband of the loop filter. This holds if $|1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a| \rightarrow 0, \forall \omega \in B_P$. Hence, we can define the cost function as follows:

$$J_a(\mathbf{x}_a) \equiv \int_{B_P} \left| 1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right|^2 d\omega. \quad (7)$$

Since

$$\begin{aligned} \mathbf{x}_a^T & \left(\text{Re}(\boldsymbol{\eta}_N(\omega)) \text{Re}(\boldsymbol{\eta}_N(\omega))^T \right. \\ & \left. + \text{Im}(\boldsymbol{\eta}_N(\omega)) \text{Im}(\boldsymbol{\eta}_N(\omega))^T \right) \mathbf{x}_a \\ & = \mathbf{x}_a^T \left((\boldsymbol{\eta}_N(\omega))^* \boldsymbol{\eta}_N(\omega)^T \right) \mathbf{x}_a \end{aligned} \quad (8)$$

equation (7) may be expressed as

$$J_a(\mathbf{x}_a) = \frac{1}{2} \mathbf{x}_a^T \mathbf{Q}_a \mathbf{x}_a + \mathbf{b}_a^T \mathbf{x}_a + p_a \quad (9)$$

where

$$\begin{aligned} \mathbf{Q}_a \equiv & 2 \int_{B_P} \left(\text{Re}(\boldsymbol{\eta}_N(\omega)) \text{Re}(\boldsymbol{\eta}_N(\omega))^T \right. \\ & \left. + \text{Im}(\boldsymbol{\eta}_N(\omega)) \text{Im}(\boldsymbol{\eta}_N(\omega))^T \right) d\omega \end{aligned} \quad (10)$$

$$\mathbf{b}_a \equiv 2 \int_{B_P} \text{Re}(\boldsymbol{\eta}_N(\omega)) d\omega \quad (11)$$

$$p_a \equiv \int_{B_P} d\omega \quad (12)$$

and \mathbf{Q}_a is a positive definite matrix.

Although the cost function minimizes the energy of the function $|1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a|$ over the passband of the loop filter, which reflects the error energy of the NTF and the ripple energy of the STF over the passband, there may be a serious overshoot. If this is the case, then the SNR of the SDM will be degraded. To avoid this, a further constraint should be imposed, which bounds the function in the passband. This is given by

$$\left| 1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right|^2 \leq \delta, \quad \forall \omega \in B_P \quad (13)$$

where δ denotes the bound. Equation (13) can further be represented as

$$\frac{1}{2} \mathbf{x}_a^T \mathbf{A}_a(\omega) \mathbf{x}_a + (\mathbf{c}_a(\omega))^T \mathbf{x}_a + q_a \leq 0, \quad \forall \omega \in B_P \quad (14)$$

where

$$\begin{aligned} \mathbf{A}_a(\omega) \equiv & 2 \left(\text{Re}(\boldsymbol{\eta}_N(\omega)) \text{Re}(\boldsymbol{\eta}_N(\omega))^T \right. \\ & \left. + \text{Im}(\boldsymbol{\eta}_N(\omega)) \text{Im}(\boldsymbol{\eta}_N(\omega))^T \right), \quad \forall \omega \in B_P \end{aligned} \quad (15)$$

$$\mathbf{c}_a(\omega) \equiv 2 \text{Re}(\boldsymbol{\eta}_N(\omega)), \quad \forall \omega \in B_P \quad (16)$$

and

$$q_a \equiv 1 - \delta, \quad \forall \omega \in B_P. \quad (17)$$

Since $\mathbf{A}_a(\omega)$ is a positive definite matrix $\forall \omega \in B_P$ and the constraint is continuous, the design of the denominator coefficients can be formulated as the following SIP problem.

Problem (P₁):

$$\min_{\mathbf{x}_a} J_a(\mathbf{x}_a) = \frac{1}{2} \mathbf{x}_a^T \mathbf{Q}_a \mathbf{x}_a + \mathbf{b}_a^T \mathbf{x}_a + p_a \quad (18a)$$

$$\text{subject to } \frac{1}{2} \mathbf{x}_a^T \mathbf{A}_a(\omega) \mathbf{x}_a + (\mathbf{c}_a(\omega))^T \mathbf{x}_a + q_a \leq 0, \quad \forall \omega \in B_P. \quad (18b)$$

Since the constraint function is convex in \mathbf{x}_a and continuously differentiable with respect to both \mathbf{x}_a and ω , the SIP problem can be solved by the dual-parameterization method [21], which guarantees the global optimal solution and satisfies the continuous quadratic constraint if the filter length is long enough.

B. Determination of Numerator Coefficients

Though the characteristics of the NTF and STF are captured in the design, the stability of these two transfer functions and the frequency characteristics of the loop filter should also be considered. Our objective is to minimize the ripple energy of the loop filter in the stopband subject to the stability condition of the transfer functions. Let the desired *magnitude* response of the loop filter be $\tilde{H}(\omega)$. In order to have good frequency characteristics of the loop filter, we want to achieve $|H(\omega)| \approx \tilde{H}(\omega)$, which implies that

$$\left| (\boldsymbol{\eta}_M(\omega))^T \mathbf{x}_b \right|^2 \approx \left| \frac{(1 - e^{-j\omega})^r \tilde{H}(\omega)}{e^{-j\omega}} \right|^2 \left| 1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right|^2. \quad (19)$$

Since \mathbf{x}_a is obtained from solving the problem P₁, r is known from the design specifications, and $\tilde{H}(\omega)$ is zero in the stopband, the cost function can be formulated as

$$J_b(\mathbf{x}_b) \equiv \int_{B_S} \left| (\boldsymbol{\eta}_M(\omega))^T \mathbf{x}_b \right|^2 d\omega. \quad (20)$$

The stability condition of the NTF and STF is [16], [17]

$$\begin{aligned} \text{Re} \left(e^{-j\omega} (\boldsymbol{\eta}_M(\omega))^T \mathbf{x}_b + (1 - e^{-j\omega})^r \right. \\ \left. \times \left(1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right) \right) \geq 0, \quad \forall \omega \in [-\pi, \pi], \end{aligned} \quad (21)$$

which is equivalent to

$$\begin{aligned} (\boldsymbol{\eta}'_M(\omega))^T \mathbf{x}_b + \text{Re} \left((1 - e^{-j\omega})^r \right. \\ \left. \times \left(1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right) \right) \geq 0, \quad \forall \omega \in [-\pi, \pi] \end{aligned} \quad (22)$$

where

$$\boldsymbol{\eta}'_M(\omega) \equiv [\cos \omega, \cos 2\omega, \dots, \cos(M+1)\omega]^T. \quad (23)$$

Hence, the optimization problem can be represented as the following SIP problem.

Problem (P₂):

$$\min_{\mathbf{x}_b} J_b(\mathbf{x}_b) = \frac{1}{2} \mathbf{x}_b^T \mathbf{Q}_b \mathbf{x}_b, \quad (24a)$$

$$\text{subject to } \mathbf{A}_b(\omega) \mathbf{x}_b + \mathbf{c}_b(\omega) \leq 0, \quad \forall \omega \in [-\pi, \pi] \quad (24b)$$

where

$$\begin{aligned} \mathbf{Q}_b \equiv & 2 \int_{B_S} \left(\text{Re}(\boldsymbol{\eta}_M(\omega)) \text{Re}(\boldsymbol{\eta}_M(\omega))^T \right. \\ & \left. + \text{Im}(\boldsymbol{\eta}_M(\omega)) \text{Im}(\boldsymbol{\eta}_M(\omega))^T \right) d\omega \end{aligned} \quad (25)$$

$$\mathbf{A}_b(\omega) \equiv -(\boldsymbol{\eta}'_M(\omega))^T, \quad \forall \omega \in [-\pi, \pi] \quad (26)$$

and

$$\begin{aligned} \mathbf{c}_b(\omega) \equiv & -\text{Re} \left((1 - e^{-j\omega})^r \left(1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a \right) \right), \\ & \forall \omega \in [-\pi, \pi]. \end{aligned} \quad (27)$$

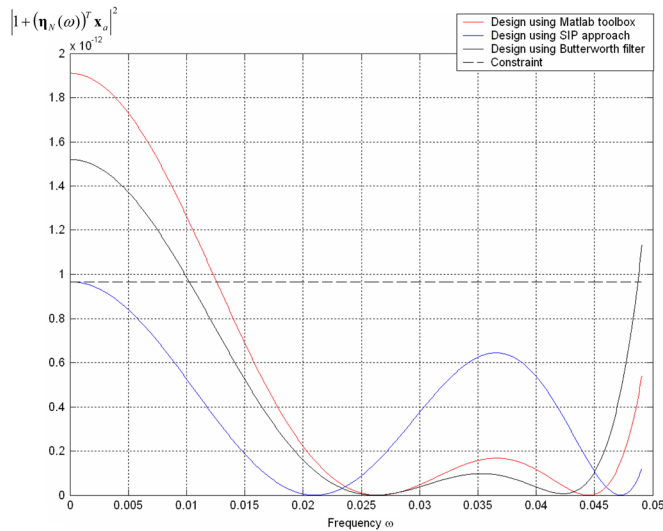


Fig. 2. Magnitude responses of $|1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a|^2$. (Color version available online at <http://ieeexplore.ieee.org>.)

Problem P_1 does not depend on the numerator coefficients, so the global optimal solution of problem P_1 can be obtained via the dual-parameterization method [21]. Since the denominator coefficients are obtained from solving problem P_1 , the global optimal solution of problem P_2 can then be obtained similarly. In this formulation, iterative design of the numerator and denominator coefficients is avoided. This is advantageous because convergence of the iterative design is not guaranteed [16].

III. SIMULATION RESULTS

To compare our design with the existing optimal designs, similar cost function and constraints should be used. However, few of them have exactly the same cost function and constraints. The most related existing design approach is the one based on the Butterworth filter structure [10] and the one via the Matlab sigma-delta toolbox [24], [25] because these two design methods employ SNR as criterion.

Consider a fifth order SDM with a dc pole, a pure delay in the numerator of the loop filter transfer function, and an oversampling ratio of 64, that is

$$M = 5, N = 4, r = 1, B_P = \left[-\frac{\pi}{64}, \frac{\pi}{64}\right]$$

and

$$B_S = [-\pi, \pi] \setminus \left[-\frac{\pi}{64}, \frac{\pi}{64}\right].$$

We choose this configuration because the order of SDM, and its oversampling ratio is typical of many audio applications [1], [2]. It can be seen from Fig. 2 that the maximum bound on $|1 + (\boldsymbol{\eta}_N(\omega))^T \mathbf{x}_a|^2$ for the design using the Matlab sigma-delta toolbox [24], [25] is 1.9101×10^{-12} , while that of based on the Butterworth structure [10] is 1.5203×10^{-12} . Hence, we would expect that our result should achieve the error bounded by 10^{-12} for $\forall \omega \in B_P$. By selecting $\delta = 10^{-12}$, the optimal SDM design problem can now be formulated as SIP problems as discussed in Section II, and these problems can be solved via the dual-parameterization method [21]. According to the simulation, it is found that our design can achieve the error bounded by 9.6658×10^{-13} , as shown in Fig. 2. It is worth noting that the designs based on the Matlab sigma-delta toolbox or the Butterworth

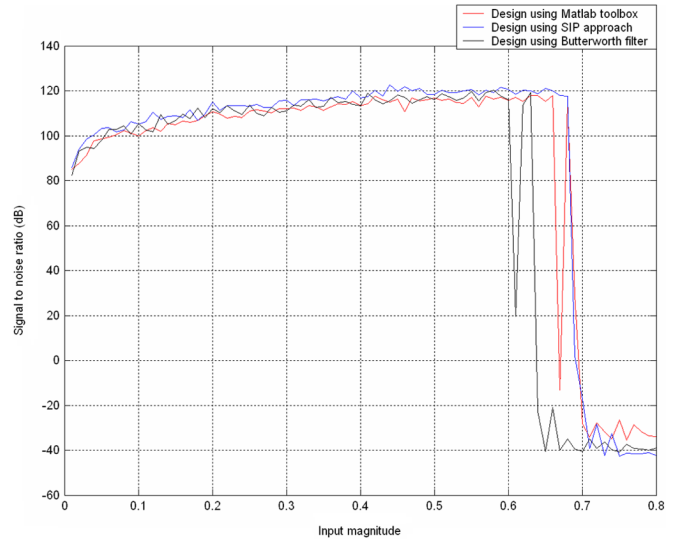


Fig. 3. Signal-to-noise ratios. (Color version available online at <http://ieeexplore.ieee.org>.)

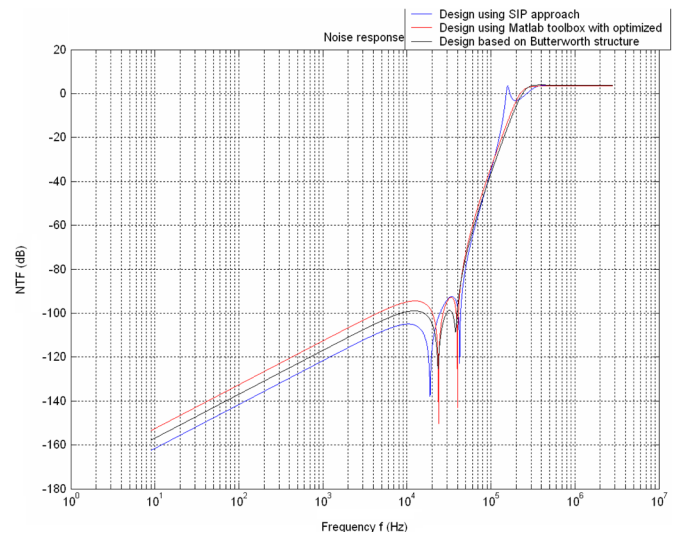


Fig. 4. Noise transfer function. (Color version available online at <http://ieeexplore.ieee.org>.)

filter structure have larger response values on the first lobe, while our design has larger value on the second lobe. This implies that our design has a higher ability to push the noise to the higher frequency band compared with the previous design.

Fig. 3 shows the SNRs of our design as well as the optimal design via the Butterworth filter structure [10] and Matlab sigma-delta toolbox [24], [25] based on the sinusoidal input with input frequency equal to $(2/3)$ of the passband bandwidth [24], [25]. It can be seen from Fig. 3 that our design achieves an average 3.7483-dB improvement compared with that of [24] and [25] and 3.0380-dB improvement compared with that of [10] when these SDMs operate normally. In addition, it is worth noting that the design via the Matlab sigma-delta toolbox [24], [25] diverges when the input sinusoidal magnitude is at 0.67, and the design via the Butterworth structure diverges at 0.61, while our design operates normally before 0.69. Hence, our design provides a high SNR and has a higher stability range. It is found that the magnitude of the poles of STF and NTF of our design are 0.9928, 0.9928, 0.8556, 0.8556, and 0.6143,

respectively, in which all are strictly inside the unit circle. Hence, the transfer functions are strictly stable.

Fig. 4 shows NTFs of our proposed design, as well as the design via the Matlab sigma-delta toolbox [24], [25] and Butterworth filter structure [10] based on oversampling ratio at 64 and Nyquist rate at 44.1 kHz, which is widely adopted in the audio applications [1], [2]. According to the simulation results, our design produces at least 9.4750-dB improvement on the passband of the loop filter compared to the design in [24] and [25] and 5.1113-dB improvement compared with the design in [10], which is a significant improvement on the suppression of the noise on the frequency band in which we are interested.

IV. CONCLUSION

In this correspondence, we have formulated SDM design problems as SIP problems and solved the problems via the dual-parameterization method. The advantages of this formulation are the guarantee of the stability of NTF and STF if the filter length is long enough, applicability to design of rational IIR filters without imposing specific filter structures such as Laguerre filter and Butterworth filter structures, and the avoidance of the nonconvergent iterative design of numerator and the denominator coefficients. Our simulation results show that the proposed design yields a significant improvement in the SNR and achieving a higher stability range compared to the existing designs.

ACKNOWLEDGMENT

The work obtained in this correspondence was supported by a research grant from Queen Mary, University of London, and a research grant from Australian Research Council

REFERENCES

- [1] E. Janssen and D. Reefman, "Super-audio CD: An introduction," *IEEE Signal Process. Mag.*, vol. 20, no. 4, pp. 83–90, 2003.
- [2] D. Reefman and E. Janssen, "Signal processing for direct stream digital: A tutorial for digital sigma delta modulation and 1-bit digital audio processing," Philips Research Eindhoven, Eindhoven, The Netherlands, White Paper, 2002.
- [3] C. M. Zierhofer, "A multiplier-free digital sinusoid generator based on sigma-delta modulation," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 43, no. 5, pp. 387–396, 1996.
- [4] A. Gerosa, A. Maniero, and A. Neviani, "A fully integrated two-channel A/D interface for the acquisition of cardiac signals in implantable pacemakers," *IEEE J. Solid-State Circuits*, vol. 39, no. 7, pp. 1083–1093, 2004.
- [5] J. C. Candy, "A use of limit cycle oscillations to obtain robust analog-to-digital converters," *IEEE Trans. Commun.*, vol. COM-22, no. 3, pp. 298–305, 1974.
- [6] H. Lampinen and O. Vainio, "An optimization approach to designing OTAs for low-voltage sigma-delta modulators," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 6, pp. 1665–1671, 2001.
- [7] J. Márkus, J. Silva, and G. C. Temes, "Theory and applications of incremental $\Delta\Sigma$ converters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 4, pp. 678–690, 2004.
- [8] O. Bajdechi, G. E. Gielen, and J. H. Huijsing, "Systematic design exploration of delta-sigma ADCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 1, pp. 86–95, 2004.
- [9] P. C. Maulik, M. S. Chadha, W. L. Lee, and P. J. Crawley, "A 16-bit 250-kHz delta-sigma demodulator and decimation filter," *IEEE J. Solid-State Circuits*, vol. 35, no. 4, pp. 458–467, 2000.
- [10] S. S. Abeysekera, Y. Xue, and C. Charoensak, "Design of optimal and narrow-band Laguerre filters for sigma-delta demodulators," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 7, pp. 368–375, 2003.
- [11] D. E. Quevedo and G. C. Goodwin, "Multistep optimal analog-to-digital conversion," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 503–515, 2005.
- [12] K. Francken and G. G. E. Gielen, "A high-level simulation and synthesis environment for $\Delta\Sigma$ modulators," *IEEE Trans. Computer-Aided Design Integr. Circuits Syst.*, vol. 22, no. 8, pp. 1049–1061, 2003.
- [13] A. Marques, V. Peluso, M. S. Steyaert, and W. M. Sansen, "Optimal parameters for $\Delta\Sigma$ modulator topologies," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 9, pp. 1232–1241, 1998.
- [14] P. Rombouts and L. Weyten, "Systematic design of double-sampling $\Sigma\Delta$ A/D converters with a modified noise transfer function," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 51, no. 12, pp. 675–679, 2004.
- [15] T.-H. Kuo, K.-D. Chen, and J.-R. Chen, "Automatic coefficients design for high-order sigma-delta modulators," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 46, no. 1, pp. 6–15, 1999.
- [16] W. S. Lu, "Design of stable IIR digital filters with equiripple passbands and peak-constrained least squares stopbands," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 46, no. 11, pp. 1421–1426, 1999.
- [17] W.-S. Lu, S.-C. Pei, and C.-C. Tseng, "A weighted least-squares method for the design of stable 1-D and 2-D IIR digital filters," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 1–10, Jan. 1998.
- [18] N. T. Thao, "Deterministic analysis of sigma-delta modulation for linear and non-linear signal reconstruction," *Int. J. Circuit Theory Appl.*, vol. 25, pp. 369–391, 1997.
- [19] R. Schreier, M. V. Goodson, and B. Zhang, "An algorithm for computing convex positively invariant sets for delta-sigma modulators," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 44, no. 1, pp. 38–44, 1997.
- [20] E. F. Stikvoort, "Some remarks on the stability and performance of the noise shaper or sigma-delta modulator," *IEEE Trans. Commun.*, vol. 36, no. 10, pp. 1157–1162, 1988.
- [21] S. Ito, Y. Liu, and K. L. Teo, "A dual parametrization method for convex semi-infinite programming," *Ann. Oper. Res.*, vol. 98, pp. 189–213, 2000.
- [22] C. Y.-F. Ho, B. W.-K. Ling, Y.-Q. Liu, P. K.-S. Tam, and K.-L. Teo, "Optimal design of nonuniform FIR transmultiplexer using semi-infinite programming," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2598–2603, Jul. 2005.
- [23] —, "Design of nonuniform near allpass complementary FIR filters via a semi-infinite programming technique," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 376–380, Jan. 2005.
- [24] R. Schreier, Delta-Sigma Modulators Toolbox Version 6.0, Analog Devices Inc., Norwood, MA, Jan. 1, 2003.
- [25] S. R. Norsworthy, R. Schreier, and G. C. Temes, "Delta-sigma data converters: Theory, design and simulation," in *Design and Simulation*. New York: IEEE Press, 1997.