<table>
<thead>
<tr>
<th>Title</th>
<th>High-speed nested cascaded MASH Digital Delta-Sigma Modulator-based divider controller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Donnelly, Yann; Mo, Hongjia; Kennedy, Michael Peter</td>
</tr>
<tr>
<td>Publication date</td>
<td>2018-05-04</td>
</tr>
</tbody>
</table>
| Link to publisher's version | http://www.iscas2018.org/  
http://dx.doi.org/10.1109/ISCAS.2018.8351624  
Access to the full text of the published version may require a subscription. |
| Rights | © 2018, IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. |
| Item downloaded from | http://hdl.handle.net/10468/7236 |

Downloaded on 2019-08-03T21:23:48Z
High-Speed Nested Cascaded MASH Digital Delta-Sigma Modulator-Based Divider Controller

Yann Donnelly
School of Electrical and Electronic Engineering
University College Cork
Cork, Ireland
Email: y.donnelly@umail.ucc.ie

Hongjia Mo
Hisense
Shandong, China
Email: hongjia.mo@aliyun.com

Michael Peter Kennedy
School of Electrical and Electronic Engineering
University College Dublin
Dublin, Ireland
Email: peter.kennedy@ucd.ie

Abstract—The MASH Digital Delta-Sigma Modulator (DDSM) based divider controller represents a speed bottleneck in state of the art commercial PLL-based fractional-N frequency synthesizers. As next generation systems require higher phase detector frequencies, there is a need to make ever faster divider controllers. This paper describes a fine-grained nested cascaded MASH DDSM which is significantly faster than state of the art divider controllers, thereby eliminating the current speed bottleneck.

I. INTRODUCTION

The Digital Delta-Sigma Modulator (DDSM) is ubiquitous in a wide range of communications and consumer devices. The DDSM quantizes a discrete input signal to reduce the signal bit-width. The quantization noise is suppressed at low frequencies (shaped), while higher-frequency noise can be easily low-pass filtered. A popular class of DDSM is the MASH DDSM, which features low-frequency noise suppression [1].

A typical communications application is divide ratio generation in a fractional-N PLL [2], [3]. When the ratio between the desired output frequency and the reference frequency is fractional, a divider control signal is required with a fractional mean component. The MASH DDSM can be used to generate a modulated digital output whose mean is the constant input divided by $2^m$, where $m$ is the resolution of the quantizer. Most of the shaped quantization noise is filtered by the loop response, while the in-band noise is typically lower than other PLL noise sources, and therefore negligible [4].

There are many good reasons to maximize the MASH DDSM clock frequency. In the PLL, the DDSM is clocked at the reference clock frequency, and the increased output frequencies demanded by next generation networks require ever higher reference frequencies [5]. Additionally, a number of noise sources, such as reference clock spurs and charge pump noise, show an inverse relationship to reference frequency [6], [7]. In signal processing applications, increased MASH DDSM clock frequencies are also desirable as the noise is shaped over the frequency range $(0, f_{clk}/2)$; therefore, increasing the clock frequency increases the frequency band over which noise suppression occurs. However, the DDSM is often a speed bottleneck, as it employs a large $m$-bit adder in order to achieve quantization and feedback. The nested mixed-radix DDSM [5] employs two DDSMs, thus increasing precision without reducing the speed of the system, but noise performance can be degraded unless the parameters are carefully chosen in order to mask additional tones which might be produced by the auxiliary DDSM.

This paper describes a Nested Cascaded MASH DDSM, hereafter denoted NC-MASH, which offers identical behaviour to the conventional MASH DDSM, but with a greatly increased maximum clock speed. Specifically, the NC-MASH pipelines the DDSM blocks in order to reduce the quantization of each block, thereby dramatically reducing the latency, with only a small increase in area. The theory governing the NC-MASH 1-1-1 is presented as an example, and an efficient implementation using adders is discussed. Finally, the trade-off between maximum clock speed and area is discussed, with synthesized examples confirming that the NC-MASH offers a significant improvement over the conventional MASH.

The paper begins with an overview of the conventional MASH DDSM in Section II, demonstrating where the speed bottleneck occurs. The NC-MASH is described briefly in Section III, which is followed by the theory governing the NC-MASH 1-1-1 in Section IV. Section V addresses efficient implementation of the NC-MASH, while Section VI demonstrates the synthesized performance and discusses speed-area tradeoffs, before a final discussion is given in Section VII.

II. CONVENTIONAL MASH DDSM

The Digital Delta-Sigma Modulator (DDSM) is a noise-shaping modulator which consists of a quantizer with a 1-bit output\(^1\) in a feedback loop, as shown in Fig. 1(a). An error signal, $e$, is fed back to the input, resulting in $+20 \text{ dB/decade}$

\(^1\)The quantizer may quantize to more than one bit, but we consider only the 1-bit case in this work. The ideas presented herein can be readily extended to the case of a multibit quantizer.
noise shaping from accumulation of the error. Due to the \( m \)-bit quantization, the output mean is attenuated by a factor of \( M = 2^m \) compared to the input. Thus:

\[
Y(z) = \frac{z^{-1}}{M} X(z) - z^{-1} \left(1 - z^{-1}\right) E(z) \tag{1}
\]

where \( Y \), \( X \) and \( E \) denote the z-transforms of the output \( y \), input \( x \) and error \( e \), respectively.

The first-order DDSM block with a 1-bit output can be easily implemented using an \( m \)-bit adder, where the sum and carry out are equivalent to the modulator error and output, respectively, as shown in Fig. 1(b).

The MASH DDSM is a development of the DDSM which employs a number of quantization stages to successively increase the noise shaping of the error term. Fig. 2 shows a MASH 1-1-1, which consists of three first-order DDSMs, and a feedback network implementing the following function:

\[
Y(z) = z^{-2}Y_1(z) + z^{-1} \left(1 - z^{-1}\right) Y_2(z) + \left(1 - z^{-1}\right)^2 Y_3(z) \tag{2}
\]

The error term is fed to a second DDSM, whose output therefore contains the error term in addition to a new error term. The second DDSM’s output is accumulated and added to the first DDSM’s output in the feedback network, cancelling the original error term and leaving behind the second error term. Likewise, the second error term is fed into a third DDSM, after which it also is cancelled, leaving behind a third error term. This error term has +60 dB/decade noise shaping, due to the effect of the three feedback loops, which results in less in-band noise after low-pass filtering.

III. NESTED CASCADED MASH

Increased precision in the output mean of the DDSM can be achieved by increasing the width of the quantizer, thus increasing both the width of the input word and the attenuation of the DDSM. However, this comes with a speed penalty as the latency of the loop is related to the width of the quantizer. Qualitatively, the greater the widths of the words, the longer it takes to add them together.

The MASH modulator achieves higher-order noise shaping by using additional DDSMs to requantize the error term. The NC-MASH takes a similar approach, and uses additional DDSMs to requantize the output term, as shown in Fig. 3.

For a given overall MASH attenuation, increasing the number of individual DDSMs in the structure reduces the required bitwidth of each DDSM, thereby reducing the delay of each DDSM, and therefore the maximum clock frequency is greatly increased with only a minimal area increase. The greater the number of levels in the cascade, the faster the clock can be.

IV. THEORY

For the purposes of illustration, we discuss the operation of the NC-MASH 1-1-1 in detail. This analysis can easily be extended to other types of NC-MASH by adding or removing terms in Eqs. (9) and (10), and modifying Eq. (11).

A. Two-level NC-MASH 1-1-1

The two-level NC-MASH replaces the chain of DDSM blocks in Fig. 2 with two levels of blocks, each with their own respective inputs \( x_1 \) and \( x_2 \), as shown in Fig. 3. The output of the DDSM blocks of the first level, shown in green, are fed into the inputs of the respective DDSM blocks of the second level, shown in red.

It can be shown that the error terms of all but the last stage of each level are cancelled, and the output has
a +60 dB/decade noise-shaped term, as in the traditional MASH 1-1-1. At level two:

\[ Y_{2,1}(z) = \frac{z^{-1}}{M_2} X_2(z) - z^{-1} (1 - z^{-1}) E_{2,1}(z) \]  
(3)

\[ Y_{2,2}(z) = z^{-1} E_{2,1}(z) - z^{-1} (1 - z^{-1}) E_{2,2}(z) \]  
(4)

\[ Y_{2,3}(z) = z^{-1} E_{2,2}(z) - z^{-1} (1 - z^{-1}) E_{2,3}(z) \]  
(5)

At level one:

\[ Y_{1,1}(z) = \frac{z^{-1}}{M_1} (Y_{2,1}(z) + X_1(z)) \]
\[ - z^{-1} (1 - z^{-1}) E_{1,1}(z) \]  
(6)

\[ Y_{1,2}(z) = z^{-1} (Y_{2,2}(z) + E_{1,1}(z)) \]
\[ - z^{-1} (1 - z^{-1}) E_{1,2}(z) \]  
(7)

\[ Y_{1,3}(z) = z^{-1} (Y_{2,3}(z) + E_{1,2}(z)) \]
\[ - z^{-1} (1 - z^{-1}) E_{1,3}(z) \]  
(8)

Substituting (3)–(5) into (6)–(8), and combining (6)–(8) according to (2), we obtain:

\[ Y(z) = \frac{z^{-3}}{M_1} X_1(z) + \frac{z^{-4}}{M_1 M_2} X_2(z) \]
\[ - \left(1 - z^{-1}\right)^3 \left( z^{-1} E_{1,1}(z) + \frac{z^{-2}}{M_1} E_{2,1}(z) \right) \]  
(9)

If the desired output mean is given by \( x/(M_1 M_2) \), then \( x_1 \) and \( x_2 \) should be chosen such that \( x = x_1 M_2 + x_2 \).

**B. Four-level NC-MASH 1-1-1**

Further clock speed increases can be achieved by increasing the number of cascaded levels, thereby reducing the width, and therefore latency, of each DDSM stage.

In the four-level implementation shown in Fig. 4, expressions for the \( Y_{2,y} \) terms can be derived from (3)–(5), and expressions for the other \( Y_{x,y} \) terms can be derived from (6)–(8). Combining as before, we obtain:

\[ Y(z) = \frac{z^{-3}}{M_1} X_1(z) + \frac{z^{-4}}{M_1 M_2} X_2(z) \]
\[ + \frac{z^{-5}}{M_1 M_2 M_3} X_3(z) + \frac{z^{-6}}{M_1 M_2 M_3 M_4} X_4(z) \]
\[ - \left(1 - z^{-1}\right)^3 \left( z^{-1} E_{1,1}(z) + \frac{z^{-2}}{M_1} E_{2,1}(z) \right) \]
\[ + \frac{z^{-3}}{M_1 M_2} E_{3,3}(z) + \frac{z^{-4}}{M_1 M_2 M_3} E_{4,3}(z) \]  
(10)

**C. Fine-grained NC-MASH 1-1-1**

The above analysis can easily be extended to \( N \) levels of cascading, as follows:

\[ Y(z) = \sum_{k=1}^{N} \left( \frac{z^{-(2+k)}}{\prod_{n=1}^{k} M_n} \right) X_k(z) \]
\[ - \left(1 - z^{-1}\right)^3 \sum_{k=1}^{N} \left( \frac{z^{-k}}{\prod_{n=1}^{k-1} M_n} \right) E_{k,3} \]  
(11)

where, for a desired output mean of \( x/(\prod_{q=1}^{N} M_q) \), the inputs should be chosen such that:

\[ x = \sum_{p=1}^{N} \left( x_p \prod_{q=p+1}^{N} M_q \right) \]  
(12)

**D. Quantization noise in the NC-MASH**

A key feature of the MASH is its high-pass shaped noise profile. In the MASH 1-1-1, the only output modulation noise is due to the quantization noise of the third stage, \( e_3 \).

In the case of the NC-MASH, the quantization noise of the last stage of each level, \( e_{k,3} \), contributes to the shaped output noise term. This is not an issue as these terms are attenuated by the successive DDSM stages; hence, in the 4-level example given in Section IV-B, the amplitude of each term is given by \( e_{1,3}, e_{2,3}/M_1, e_{3,3}/(M_1 M_2) \), and \( e_{4,3}/(M_1 M_2 M_3) \). The additional contributions from the \( e_{2,3}, \ldots e_{4,3} \) terms are therefore negligible, and the noise performance of the NC-MASH is effectively identical to that of the conventional MASH DDSM.

**E. Dithering the NC-MASH**

The DDSM is a finite state machine; therefore, the conventional MASH suffers from periodic behaviour that can result in spurs appearing on the output [8]–[11]. This unwanted behaviour is also possible in the NC-MASH.

Several deterministic [11]–[13] and stochastic methods have been suggested for overcoming this problem. Of the latter, applying LFSR dither to LSBs of the inputs of the second DDSM stage of the MASH 1-1-1 has been found to be the most effective method [14], although care must be taken to ensure that the periodicity of the LFSR does not itself produce spurs [15]. More generally, Pamarti and Galton have shown [16] that dither of the order \((L-1)^{th}\) and greater will produce spurs in an \(L^{th}\)-order MASH, although a small adjustment to the DDSM structure will allow up to \(L^{th}\)-order dither [17].

Dithering can similarly be implemented in an \(N\)-level NC-MASH by dithering the relevant DDSM in the \(N^{th}\) level of the cascade. Dithering of successive levels is only necessary if a large number of levels is used, as the input from the previous level is usually sufficiently randomized to break up the periodicity. Stability is not a concern, as \(1^{st}\)-order DDSMs are unconditionally stable, and the architecture is feed-forward.

**V. IMPLEMENTATION**

DDSM latency can be minimized by implementing clocked adders as first-order modulators. In this way, the \( y \) terms can be applied to the carries in of the adders, thus minimizing the area increase of the NC-MASH compared to a conventional MASH.

In the design shown in Fig 3, and as seen in Equation (11), there is a mismatched latency between each of the \( x \) inputs and the output \( y \), which results in incorrect performance when the input word is not constant. This can be overcome by delaying the inputs by a sufficient amount so that the latencies from each input to the output are equal. In general, in an \(N\)-stage...
delay reduction can therefore be estimated using:

\[ D_{\text{flop}} = \sum_{k=0}^{L-1} L_k \text{clock cycles}, \text{where } L_k \text{ is the latency of the DDSM in the } k^\text{th} \text{ level}. \]

An implementation of a 4-stage NC-MASH is shown in Fig. 4, with one-bit white noise dither added to the second DDSM stage of the fourth level to produce \((L-2) = 1^\text{st}-\text{order shaped dither}.\)

**VI. PERFORMANCE**

The NC-MASH is characterized by a trade-off between area and speed, as the reduction in worst-case delay comes at the expense of the additional flip flops that are required to store the intermediate output and error signals.

In the conventional MASH, we can divide the area into four: (i) the DDSM output flip flops, (ii) the DDSM feedback flip flops, (iii) the DDSM adders and and (iv) the accumulation network: \(A_{\text{MASH}} = A_{\text{DFF(y)}} + A_{\text{DFF(e)}} + A_{\text{add}} + A_{\text{acc}}.\) In the \(N\)-level NC-MASH, the number of DDSMs is increased by a factor of \(N\); however, the width of each adder is also decreased by a factor of \(N\). As a result, the area of the adders and feedback flip flops remain roughly constant, and only the number of output flip flops increases. Hence, \(A_{\text{NC-MASH}} \approx NA_{\text{DFF(y)}} + A_{\text{DFF(e)}} + A_{\text{add}} + A_{\text{acc}}.\) The area increase can therefore be estimated using:

\[ A_{\text{NC-MASH}} \approx (A_{\text{MASH}} - A_{\text{DFF(y)}}) + NA_{\text{DFF(y)}}. \] (13)

Similarly, the worst-case delay of a DDSM block, implemented as a clocked adder, can be divided into the delay through all adder stages and the delay due to the output flip flop: \(D_{\text{MASH}} = D_{\text{DFF(y)}} + D_{\text{add}}.\) In the \(N\)-stage NC-MASH, if we assume that \(N\) is a factor of the overall modulator bit width \(m\), the width of each DDSM block is reduced by a factor of \(N\); therefore \(D_{\text{NC-MASH}} \approx D_{\text{DFF(y)}} + D_{\text{add}}/N.\) The delay reduction can therefore be estimated using:

\[ D_{\text{NC-MASH}} \approx (D_{\text{MASH}} - D_{\text{add}}) + \left(\frac{1}{N}\right)D_{\text{add}}. \] (14)

In the more general case, if \(N\) is not a factor of \(m\), then the worst-case DDSM delay can be shown to be equal to:

\[ D_{\text{NC-MASH}} \approx D_{\text{DFF}} + \frac{1}{m} \left[\frac{m}{N}\right] D_{\text{add}}, \] (15)

where \([x]\) denotes the smallest integer greater than or equal to \(x\).

The worst-case delay can therefore be more generally estimated as:

\[ D_{\text{NC-MASH}} \approx (D_{\text{MASH}} - D_{\text{add}}) + \left(\frac{1}{m} \left[\frac{m}{N}\right]\right) D_{\text{add}}. \] (16)

Note that the delay improvement relative to area is maximized when \(N\) is a factor of \(m\).

From (13) and (14) it is clear that, since the flip flop area is a small component of the overall area, and the adder latency comprises the majority of the DDSM delay, the NC-MASH offers a substantial speed improvement for a minimal area increase. Fig. 5 shows the estimated worst-path delay and area for a 16-bit conventional MASH 1-1-1 \((N = 1)\) and 16-bit NC-MASH 1-1-1 \((N = 2, 4, 8)\), normalised to the conventional MASH. Representative structures have been synthesized using Cadence Genus, using a 160 nm process and a timing-driven flow. The corresponding areas and delays are denoted by squares and circles. Appropriate curves derived from (13) and (14) are fitted. With four levels of cascading, the maximum clock speed increases *threelfold*, compared to the conventional MASH, for an area increase of only 8.6%.

**VII. CONCLUSION**

A Nested Cascaded MASH has been presented that offers identical spectral performance to the conventional MASH, but features a greatly increased maximum clock speed with a minimal area penalty. These improvements have been demonstrated
using synthesized RTL. This makes the NC-MASH a very interesting drop-in replacement for the conventional MASH DDSM.

ACKNOWLEDGMENT
This work has been funded in part by the Irish Research Council, Science Foundation Ireland and Enterprise Ireland under grants GOIPG/2014/14222, 13/IA/1979 and 13/RC/2077, and CC-2009-05, respectively.

REFERENCES