A 8.25-MHz 7th-order Bessel filter built with MOSFET–C single-amplifier biquads

Hanspeter Schmid and George S. Moschytz

Signal and Information Processing Laboratory, Swiss Federal Institute of Technology,
Sternwartstrasse 7, 8092 Zürich, Switzerland. E-mail: h.p.schmid@ieee.org

Abstract — The 7th-order Bessel filter presented in this paper has an edge frequency that is continuously tunable from 4.5 MHz up to 10 MHz. It was fabricated with the 0.6-micron CMOS process by AMS, covers a chip area of 0.24 mm^2, and consumes 49 mW from a 3.3-V supply. The SNR at -40 dB of harmonic distortion is between 48 dB and 50 dB over the whole tuning range. The comparatively low power consumption and chip area could be achieved by using single-amplifier biquadratic building blocks implemented as MOSFET–C filters and generating the control voltage of the MOSFET resistors with an on-chip charge pump. This paper briefly discusses design issues of such filters, sources of harmonic distortion, clock feed-through, and noise. Finally, it is shown that the filter can also be implemented using MOSFET capacitors instead of poly-poly capacitors with only 4 dB loss of SNR.

Introduction

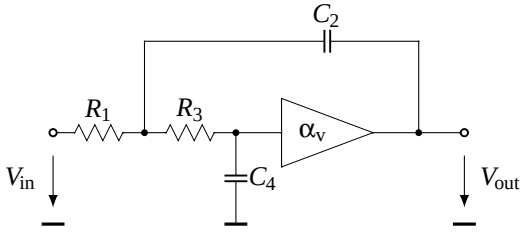
Discrete-component single-amplifier biquadratic filters (SABs) have long been used in the industry: because they require only one amplifier to generate a pair of complex poles, they are cheaper and more power-efficient than integrator-connected filters. This advantage has, however, to be paid with a higher variance of the pole-Q [1, 2].

In spite of their advantages in terms of power consumption and cost, SABs have not yet been used much on integrated circuits. The reason is that their pole frequency depends on the values of passive components only, which means that in order to tune the pole frequency, adjustable passive components have to be used. The simplest way to achieve this is to build the SAB as a MOSFET–C filter by replacing the resistors with transistors operating in the triode region.

This paper demonstrates how MOSFET–C SABs can successfully be used to build higher-order video-frequency filters. The filter described here was built to verify the theory developed in [3], where a comprehensive discussion of most theoretical and practical aspects of MOSFET–C SABs including detailed descriptions of the measurements can be found. For this reason, the present paper mainly describes how the filter was built and gives the most important measurement results.

The following sections discuss the operation principle of Sallen-and-Key filters and the effects of amplifier non-idealities on the filter transfer function. Then second-order MOSFET–C networks are introduced, and it is explained how harmonic distortion and noise are minimised. The next two sections describe the current-mode amplifier and the charge pump used to build the filter. Finally, the 7th-order filter is described, measurement results are given, and a few trade-offs are discussed.

Single-amplifier biquadratic filters (SABs)



The formulae describing a lowpass SAB become more compact if the component values are substituted as follows: $R_1 = R/n$, $R_3 = R \cdot n$, $C_2 = C/m$ and $C_4 = C \cdot m$. This substitution has also practical relevance, because R and C are the geometric means of the components, whose precision is determined by the process variations, and n and m are the *component spread factors* of the resistors and the capacitors, whose precision is determined by matching. Ideally, such a filter has the following transfer function:

$$T(s) = \frac{\alpha_\mathrm{V}\,\omega_\mathrm{p}^2}{s^2 + \frac{\omega_\mathrm{p}}{q_\mathrm{p}} s + \omega_\mathrm{p}^2}\,, \quad \text{with} \quad \omega_\mathrm{p} = \frac{1}{RC}\,, \quad \frac{1}{q_\mathrm{p}} = mn + \frac{m}{n} + \frac{1-\alpha_\mathrm{V}}{mn}\,,$$

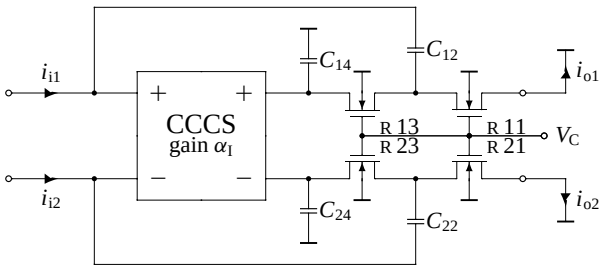
where ω_p is the pole frequency in rad/s and q_p is the pole-Q. Only R and C appear in the formula for ω_p, so in order to tune ω_p, one of these values must be adjusted. This will theoretically have no effect on q_p. If q_p should be tuned as well, which is seldom necessary for low-q_p filters, this has to be done by adjusting α_V, the gain of the amplifier.

A real amplifier will have a certain input capacitance, an output resistance, and a phase lag. All of them cause pole shifts, but the non-zero output resistance also introduces a pair of complex zeros that causes the transfer to rise to a certain level for higher frequency and thus limits the achievable stopband attenuation [4, 5]. So if the stopband attenuation and a certain amplifier are given, the maximum achievable pole frequency becomes

$$\omega_\mathrm{p\,max} \approx \frac{1}{\max(m, 1/m)\, C_\mathrm{o} \cdot \max(n, 1/n)\, R_\mathrm{i} \cdot A_\mathrm{stop}}\,,$$

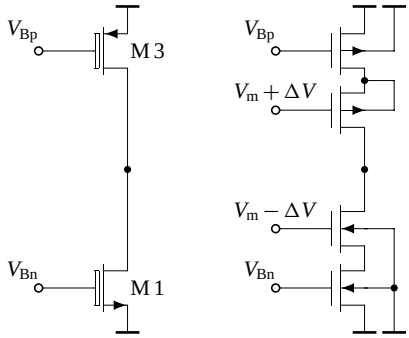
which is largest for $m = n = 1$. Note that $m = n = 1$ means that C_4 consists of the amplifier input capacitance only. This may cause an unacceptably high q_p variance, because the precision of q_p the relies on the matching of a poly-poly capaciatnce and a parasitic capacitance. Note also that in order to minimise the variance of q_p, the component spreads should generally be as large as possible, and not $m = n = 1$, as was shown in [2]. There it was also shown that a Sallen-and-Key lowpass filter with minimum q_p variance always needs an amplifier with $\alpha_\mathrm{V} < 2$.

Second-order MOSFET–C filter



The SAB from the previous section can now be converted into a MOSFET–C filter by building it in a balanced form and replacing the resistors by transistors operating in the triode region. The reason why we used a current amplifier instead of a voltage amplifier is that only a low gain is required, as mentioned above. In order to make the filter as power-efficient and small as possible, the amplifier gain should not be set using feedback. A low-gain amplifier without feedback can most easily be built using current mirrors, so we transposed the SAB into the current-mode domain [6].

The output of our current amplifier consists of composite transistors, as shown on the right. The cascode transistors are biased with a voltage $V_\mathrm{m} \pm \Delta V$, where V_m is the mid-rail voltage. Note that in the process we used, both *n*MOS and *p*MOS have the same threshold voltage $|V_{\mathrm{T0}}| = 0.85\,\mathrm{V}$. It can be shown that the ΔV that optimises the SNR for a given level of harmonic distortion (e.g. $-40\,\mathrm{dB}$) is [7]

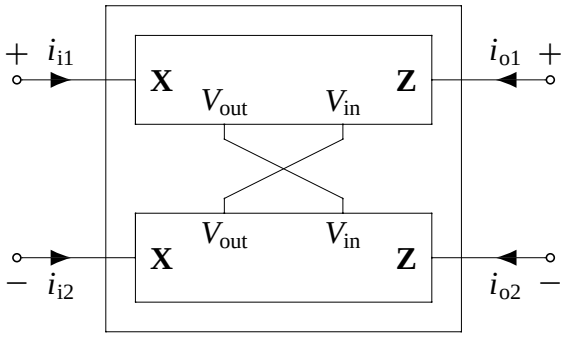


$$\Delta V = \frac{2}{7} V_\mathrm{dd} - V_{\mathrm{T0}}\,,$$

which is only $0.09\,\mathrm{V}$ for $V_{\mathrm{T0}} = 0.85\,\mathrm{V}$ and $V_\mathrm{dd} = 3.3\,\mathrm{V}$ and can safely be set to zero in the implementation.

The assumption leading to this formula is that the main source of distortion is the signal clipping that occurs when the cascode transistors at the output of the current amplifier leave the saturation region. This is indeed a good model of the reality if the gate control voltage of the MOSFET resistors is generated by a charge pump. In MOSFET–C SABs without a charge pump, clipping introduced by saturating MOSFET resistors must be taken into account [3].

Video-frequency current amplifier

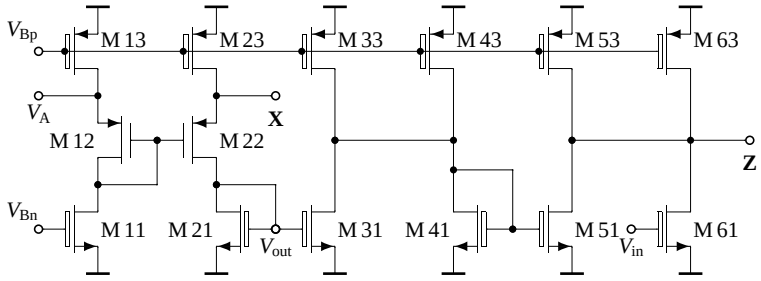


MOSFET resistors are not very linear, but since the non-linearity is mainly of second order, the difference of the two filter output signals is theoretically linear. The current amplifier used in MOSFET–C SABs must therefore amplify the difference of its input currents, and its two paths should be identical. The amplifier whose block diagram is shown on the left meets both criteria. Ideally, it is described by

$$i_{\mathrm{o1}} = \alpha_\mathrm{I}(i_{\mathrm{i1}} - i_{\mathrm{i2}}) \qquad i_{\mathrm{o2}} = -\alpha_\mathrm{I}(i_{\mathrm{i1}} - i_{\mathrm{i2}})\,.$$

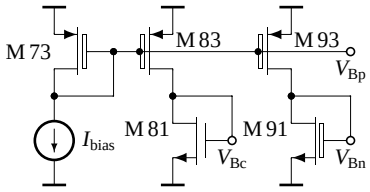
According to the discussion in the previous section, the output capacitance of the current amplifier should be as low as possible, its input resistance should be as low as possible, and it can also be shown that its phase lag should not exceed $10\ldots20^\circ$ at the pole frequency of the biquad to be built [3].

The half circuit of our current amplifier was developed from a circuit normally called class-A second-generation current conveyor in the literature [8, Chap. 11.5]. It consists of one voltage buffer and several current mirrors. M [1–6]3 and M 11 are constant current sources, while M [2–6]1 form current mirrors. M 22 is the input transistor. It provides, at its source, a current input with input resistance $R_\mathrm{in} \approx 1/g_{\mathrm{m}22}$. M 12 is a voltage level shifter that sets the operating point voltage of node X to V_A ($V_\mathrm{A} = V_\mathrm{m}$ in our implementation). Any



current flowing into X is mirrored from M 21 to M 31 and from M 41 to M 51 and flows out of Z; it is also mirrored from M 21 to M 61 of the other half-circuit, where it flows into Z. Thus the two input currents i_{i1} and i_{i2} are subtracted, and if all current mirrors have unity gain, the resulting gain is $\alpha_I = -2$. A different gain can easily be achieved by changing the width of all output transistors M [5–6][1,3].

Note that all transistors drawn with boxes as gates are composite transistors as explained in the previous section. All cascode transistors are biased by V_m, with one exception. The gate potential of the diode-connected transistor M 12 is below $V_m - V_{T0}$, so the cascode transistor of M 11 must be biased differently, namely by the voltage V_{bc} generated by transistor M 81 of the bias circuit shown on the right. The main transistor of M 11 does not necessarily need to be enlarged, since it only has to conduct the bias current, while all other *n*MOS composite transistors in the amplifier have to be able to conduct the bias current plus the maximum signal current. So in our design, M 11 has the same size as M [2–6]1.



	main transistors	cascode transistors
M [1–6]1	45 × 1.8 µm	95 × 0.6 µm
M 81	14 × 0.6 µm	—
M 91	45 × 1.8 µm	95 × 0.6 µm
M [1–2]2	120 × 0.6 µm × 2	—
M 13	87 × 1.8 µm	140 × 0.6 µm
M [2–8]3	70 × 1.8 µm	140 × 0.6 µm
M 93	87 × 1.8 µm	140 × 0.6 µm

In the composite transistors, the W/L ratio of the main transistor is about six times smaller than the W/L ratio of the cascode transistor. As was shown in [9], simple cascodes are fastest when the V_{dsat} of the cascode transistor is about 40 % of the V_{dsat} of the main transistor. The factor of six results when the W/L ratios are calculated from V_{dsat} and I_d. The absolute transistor dimensions were found iteratively. First, we knew from experience that the bias current would have to be around 160 µA to achieve a pole frequency around 20 MHz. The maximum signal current to be supported by the current amplifier was designed to be 60 µA, approximately the current at which the MOSFET resistors would saturate. This determined the sizes of all current source and current mirror transistors. The input transistors M [1–2]2 were designed such that they would provide an X input resistance around 500 Ω, and then it was verified that the cascode transistor in M 11 would indeed remain in saturation by giving it a bias voltage 0.1 V below analogue ground, which determined the size of M 81.

Self-oscillating charge pump

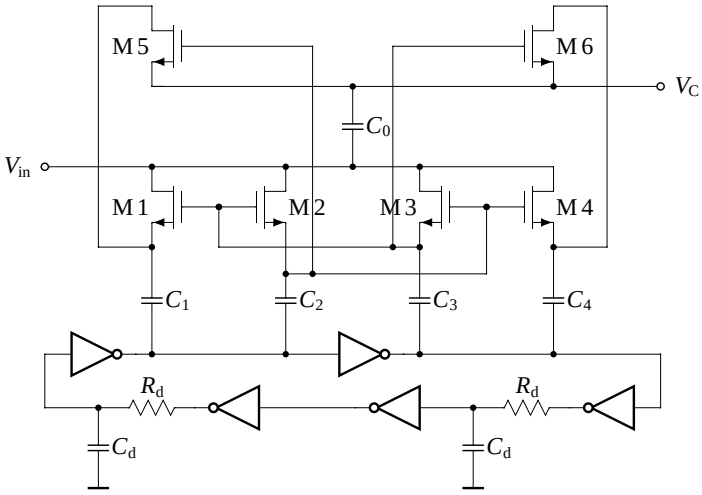
The charge pump shown on the next page combines features of a charge pump proposed by Duisters and Dijkmans in [10] with those of a five-inverter ring oscillator. It actually comprises two charge pumps. The main pump, consisting of M 1, M 4, M 5, M 6, C_1, and C_4, fills the reservoir capacitor C_0 with charge, where M 5 and M 6 alternatively conduct the charging current. A second pump driven by the same inverters, consisting of M 2, M 3, C_2, and C_3, sets the gate voltage of M 5 and M 6 to $2\,V_{in}$ while they charge C_0. Thus the output voltage becomes $V_C \leq 2\,V_{in} - V_{T5}$, where M 5's threshold voltage V_{T5} is comparatively large because of the bulk effect (we are using an *n*-well process). In our example, $V_C = 4.6\,V$ for $V_{in} = 3\,V$. The charge pump operates properly for $V_{in} = 1.3 \ldots 3.3\,V$, resulting in $V_C = 1.5 \ldots 5.0\,V$.

Charge pump components	
All *n*MOSTs	10 × 0.6 µm
All *p*MOSTs	33 × 0.6 µm
R_d	4.8 kΩ
C_d, C_2, C_3	0.5 pF
C_1, C_4	1 pF
C_0	20.5 pF

The voltage ripple of this charge pump is smaller than that of a conventional charge pump by a factor of $g_{m5}/g_{ds5} \approx 30\ldots100$. As mentioned in [10], the voltage ripple becomes

$$V_{\text{ripple}} = \frac{1}{2} \cdot \frac{I_{\text{out}}}{C_0\ f_{\text{clk}}} \cdot \frac{g_{ds5}}{g_{m5}},$$

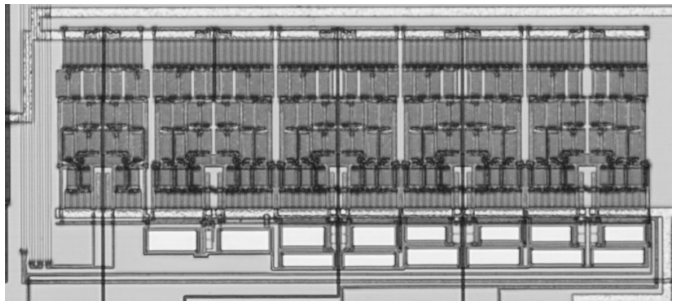
where I_{out} is the DC current flowing out of the reservoir capacitance C_0 and f_{clk} is the pump's clock frequency. This means that if V_{ripple}, I_{out} and f_{clk} are the same, the two-stage charge pump needs a reservoir capacitor which is $30\ldots100$ times smaller than the one in a conventional charge pump.



The oscillator should have an oscillation frequency that is well beyond the edge frequency of the filter. Since the inverters need to deliver small currents only, they can be built with small transistors. Using only inverters would result in an oscillation frequency of almost 1 GHz, thus two passive one-pole lowpass filters, each consisting of one poly resistor and one poly-poly capacitor, were used to slow the oscillator down to simulated 90 MHz. This has the additional advantage that it reduces the temperature dependence of f_{clk}. The measured oscillation frequency was between 62 Mhz and 71 MHz for a charge-pump supply voltage going from 2.7 V to 3.3 V.

7th-order Bessel filter

Using the biquads discussed above, the 7th-order Bessel filter shown on the right was built. From left to right, the following building blocks can be seen: the poly resistors (four vertical gray lines) and the current amplifier of the on-chip V–I converter used for making the measurements, a passive first-order low-pass filter with a normalised $f_p = 1.687$ that consists of a current amplifier and a first-order MOSFET–C lowpass filter, and three MOSFET–C SABs with $(f_p, q_p) = (2.053, 1.13)$, $(1.719, 0.53)$, and $(1.825, 0.66)$, in this order. This filter could, e.g., be used as a pulse equaliser in a $1\times$ DVD read channel [11].



In order to maximise the SNR of a biquad cascade, the gains of the individual biquads should be set to unity to make the signal levels in all biquads equal. This will, however, increase the variance of the pole-Q considerably compared to the pole-Q variance of optimum design. In our design, we decided to use a gain of two in all biquads, but measurements and simulations showed that it would have been sufficient to give the highest-q_p biquad a gain of two and the remaining stages a gain of one, which would have increased the maximum current before clipping occurs, and therefore the SNR of the filter, by a factor of three, or by 9.5 dB.

Biquadratic filter sections			
(f_p, q_p)		element dimensions	capacitance
$(1.687, -)$	C [1–2]4	68.9×28 µm	1.68 pF
	R [1–2]3	12×8 µm	
$(2.053, 1.13)$	C [1–2]2	60×26.75 µm	1.40 pF
	C [1–2]4	72.2×16.6 µm	1.05 pF
	R [1–2][1,3]	12×6 µm	
$(1.719, 0.53)$	C [1–2]2	55.5×19.1 µm	0.93 pF
	C [1–2]4	72.2×22.2 µm	1.40 pF
	R [1–2][1,3]	12×10.5 µm	
$(1.825, 0.66)$	C [1–2]2	56.5×21.25 µm	1.05 pF
	C [1–2]4	72.2×19.5 µm	1.23 pF
	R [1–2][1,3]	12×9.5 µm	

The passive component values used to build the 7th-order filter are shown on the left. Note that the signal capacitors are large compared to the output capacitance of the current amplifier, which is around 0.6 pF. It can be shown that the linearity of the filter is scarcely affected if no explicit signal capacitors are used, such that C [1–2]4 consist of parasitic capacitance only [3]. This would make it possible to increase the edge frequency of the 7th-order filter without using more power, but it would both reduce the SNR, because the noise bandwidth would increase, and it would increase the variance of all q_p, because the latter depends on the matching of C [1–2]4 and C [1–2]2, which would become worse.

Measurement results

Process parameters			
	*n*MOS	*p*MOS	
V_{T0}	0.85	−0.85	[V]
$\mu \cdot C_{ox}$	120	40	[µA/V^2]
γ	0.8	0.5	[$\sqrt{\text{V}}$]
ϕ_0	0.94	0.91	[V]

The 7th-order Bessel filter was integrated in the 0.6-µm double-poly triple-metal CMOS process by Austria Mikro Systeme, of which some parameters are shown on the left. Fifteen chips were produced, and all worked fine. The inter-chip matching measured for a single biquad with $f_p = 24$ MHz and $q_p = 3$ was sufficiently good, the standard deviations of f_p and q_p were 1.5 % and 3 %, respectively. On-chip matching can be expected to be even better, but was not measured.

The measured performance of the 7th-order Bessel filter is summarised in the table on the right. The values for power consumption and chip area include the charge pump, the latter also includes the wiring around the filter block and a margin of a few µm in every direction. The clock feed-through to the filter outputs was measured as well (actually, this was how we determined the oscillation frequency of the filter), but it was scarcely visible over the noise floor and changes the SNR of the filter by much less than 1 dB.

Measured performance	
Charge pump supply	2.7…3.3 V
V_C	4.4…5.0 V
Edge frequency	4.5…10 MHz
SNR for 1 % THD	48…50 dB
Supply voltage	3.3 V
Power consumption	53 mW
Chip area	0.28 mm^2

As discussed in the previous section, changing the gain of all but the highest-q_p stage to one would increase the SNR by 9.5 dB, and at the same time reduce the power consumption a little, since the bias current of the current amplifiers are scaled by the gain as well. Then the filter would have properties that are very similar to the 4-MHz, fifth-order Gm–C ladder filter simulation from [12], but would be smaller by a factor of six: The filter in [12] uses 0.25 mm^2 per pole, ours uses only 0.04 mm^2 per pole. Other comparisons made in [3] show that, generally, the performance of MOSFET–C SABs is similar to the performance of conventional Gm–C filters, but results in circuits that use only 15…30 % of the chip area.

Discussion

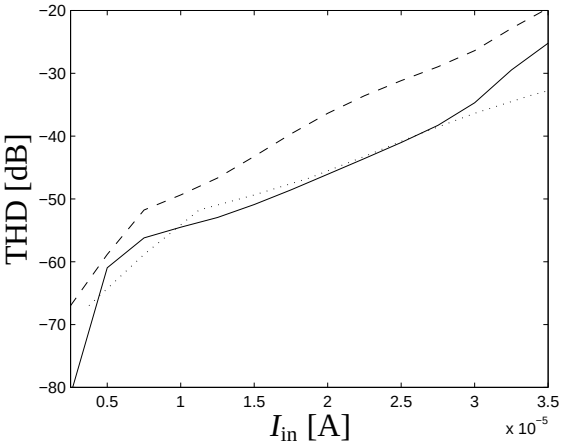
Several trade-offs that are important during the design of our filters were already discussed above. This final section covers a few important trade-offs from a wider perspective; the aim is to give the reader an impression of what can or be done with MOSFET–C SABs.

Amplifier input resistance and output capacitance. As described above, the maximum achievable pole frequency of an SAB is determined by the stopband attenuation, the input resistance, and the output capacitance. Since the output capacitance cannot be decreased much without reducing the voltage swing (and with it the signal swing), the only viable alternative is to reduce the input resistance. One way to do this is to simply increase the supply current of the current amplifier input stage. However, since this current is mirrored to all other stages, this makes it necessary to enlarge the current mirrors, which again increases the output capacitance and limits the use of this method.

Another idea is to reduce the input resistance by using local feedback with a very high unity-gain bandwidth. Then the local feedback amplifier would consume the major part of the total power, which is possibly the only viable alternative to considerably increase the maximum possible pole frequency by trading power off for speed.

Signal swing, THD, and SNR. It was explained above how the signal swing in charge-pumped MOSFET–C SABs should be set in order to maximise the SNR of the filter at a certain level of THD. However, there is little correlation between the level of THD and the SNR in a certain filter. Since the THD is mostly caused by clipping, it rises very quickly when a certain input current is exceeded. Thus the SNR for -40 dB THD would normally be only 2 dB larger than for -60 dB THD.

Double-poly or single-poly process. One may also use gate-oxide capacitors instead of poly-poly capacitors to implement the signal capacitances. The resulting filter is then compatible with standard digital CMOS processes. The figure on the right shows the simulated THD of a conventional MOSFET–C SAB (solid), of a MOSFET-only SAB (dashed), and the latter with I_{in} scaled by 1.5 (dotted). The capacitor block becomes only slightly larger (by about 25 %). The maximum possible input current is reduced to 65 % by the conversion to a MOSFET-only filter, which amounts to a loss of only 4 dB of SNR. This shows that single-amplifier biquadratic MOSFET–C filters can also be used on standard digital processes.



Charge pump or no charge pump. The advantages of having a charge-pump to drive the MOSFET resistor gates are so great that it should be done if possible. Also, the clock feed-through to the output of our filters is small enough for most applications. There are two things that could prevent the use of a charge pump.

First, although our filters reject the substrate noise generated by the charge pump quite well, it must be made sure that the same is true for all other signal processing circuits on the chip. This may be a problem on purely analogue ICs, but is not really an issue on mixed-signal ICs, since there the substrate noise of the digital part dominates anyway.

Second, the charge pump described above is constructed that although its output voltage can reach 5 V, no terminal voltage difference on any elements will exceed 3.3 V, so theoretically

no break-down will occur even if the process used does not support 5 V (the 0.6-μm CMOS process is actually a 5-V process). The same is true for the MOSFET–C SABs. However, over-peaking during the transients might change this, and it must be made sure by careful simulations that the charge pump is compatible with the process at hand, lest the yield becomes unacceptably low.

References

[1] George S. Moschytz, "Single-amplifier active filters: A review," *Scientia Electrica*, vol. 26, no. 1, pp. 1–46, 1980.

[2] Hanspeter Schmid and George S. Moschytz, "Minimum-sensitivity single-amplifier biquadratic filters," in *Proc. ECCTD*, Stresa, Italy, Aug. 1999, vol. 2, pp. 1027–1030.

[3] Hanspeter Schmid, *Single-Amplifier Biquadratic MOSFET–C Filters for Video Frequencies*, Ph.D. thesis, Swiss Federal Institute of Technology, Zürich, 2000, (Will be published in November 2000, copies can be obtained from h.p.schmid@ieee.org).

[4] Hanspeter Schmid and George S. Moschytz, "Active-MOSFET-C single-amplifier biquadratic filters for video frequencies," *IEE Proc. Circuits Devices Systems*, vol. 147, no. 1, pp. 35–41, Feb. 2000.

[5] Hanspeter Schmid and George S. Moschytz, "Fundamental frequency limitations in current-mode Sallen-Key filters," in *Proc. ISCAS*, Monterey, California, June 1998, vol. 1, pp. 57–60.

[6] George S. Moschytz and A. Carlosena, "A classification of current-mode single-amplifier biquads based on a voltage-to-current transformation," *IEEE Trans. CAS–II*, vol. 41, no. 2, pp. 151–156, Feb. 1994.

[7] Hanspeter Schmid and George S. Moschytz, "A charge-pump-controlled mosfet–c single-amplifier biquad," in *Proc. ISCAS*, Geneva, Switzerland, May 2000, vol. 2, pp. 677–680.

[8] John Lidgey, Chris Toumazou, Alison Payne, Doug C. Wadsworth, Sittichai Pookaiyaudom, and Erik Bruun, "Tutorial 10: Current-mode analog signal processing; Part 2: current conveyors," in *Proc. ISCAS*, London, 1994, pp. 569–641.

[9] Thomas Burger and Qiuting Huang, "A 100 dB, 480 MHz OTA in 0.7 μm CMOS for sampled-data applications," in *Proc. CICC*, San Diego, California, May 1996, pp. 101–104.

[10] Tonny A. F. Duisters and Eise Carel Dijkmans, "A -90-dB THD rail-to-rail input opamp using a new local charge pump in CMOS," *IEEE J. Solid-State Circuits*, vol. 33, no. 7, pp. 947–955, July 1998.

[11] Chun-Sup Kim, Gea-Ok Cho, Yong-Hwan Kim, and Bang-Sup Song, "A CMOS $4\times$ speed DVD read channel IC," *IEEE J. Solid-State Circuits*, vol. 33, no. 8, pp. 1168–1178, Aug. 1998.

[12] Changsik Yoo, Seung-Wook Lee, and Wonchan Kim, "A $\pm$1.5-V, 4-MHz CMOS continuous-time filter with a single-integrator based tuning," *IEEE J. Solid-State Circuits*, vol. 33, no. 1, pp. 18–27, Jan. 1998.