

Simulation eines rekonfigurierbaren G_m -C filter arrays

F. Henrici, J. Becker, and Y. Manoli

Universität Freiburg, Institut für Mikrosystemtechnik – IMTEK, Lehrstuhl für Mikroelektronik, Freiburg, Germany

Zusammenfassung. Es wird ein G_m -C Filter für den Einsatz in rekonfigurierbaren Analogfiltern (FPAA) präsentiert. Das Filter ist auf den Einsatz in FPAA mit hexagonalem Grid und auf den Verzicht auf Transmissiongates optimiert. Trotzdem können nicht nur die Parameter der instanziierten Filter geändert werden, sondern auch ihre Struktur. Beim Design des digital programmierbaren Transkonduktors musste auf die hohe Anzahl parallel geschalteter, gleichartiger Transkonduktoren und ihre parasitären Kapazitäten Rücksicht genommen werden. Ein FPAA mit 49 G_m -Zellen erreicht in Simulationen in einem 130 nm 1.2 V CMOS Prozess eine maximale Bandbreite von 164 MHz. Die Verzerrung beträgt weniger als -70 dB bei einem 50 mV @ 1 MHz Signal.

1 Einleitung

Analoge Signalquellen, wie z.B. Sensoren oder Empfänger in der Kommunikationstechnik, müssen gefiltert werden, bevor die Daten weiterverarbeitet werden können. Es besteht einerseits ein Bedarf an Filtern hoher Bandbreite für die drahtlose Datenübertragung, andererseits sollen diese Filter immer flexibler an neue Standards angepasst werden. Klassische Ansätze sind fast ausschließlich zeitdiskrete (DT) Lösungen. Beispiele hierfür sind DSPs oder FPGAs in der digitalen Welt, und switched-capacitor in der analogen. Da es sich hierbei um abtastende Systeme handelt, muss die Abtastfrequenz mehr als doppelt so hoch wie die Signalfrequenz sein, was einen hohen Stromverbrauch nach sich zieht.

Zeitkontinuierliche Systeme hingegen vermeiden die Abtastung, daher muss die Bandbreite des Systems gerade so groß wie die des Signals sein. Auf diese Weise kann entweder Energie gespart werden, die Bandbreite erhöht werden, oder sogar beides. Zumeist sind in zeitkontinuierlichen rekonfigurierbaren Analogfiltern (FPAA) aber nur die passiven Komponenten, z.B. Widerstände und Kapazitäten, einstellbar, die aktiven Elemente hingegen wie z.B. Operations-

verstärker bleiben konstant. Diese Arbeit präsentiert einen FPAA bestehend aus zeitkontinuierlichen G_m -C Filtern. Diese bauen auf einer Weiterentwicklung der programmierbaren G_m -Zelle von (Pavan and Tsividis, 2000) auf. Die neuartige Architektur des FPAA vermeidet den Einsatz von Transmissiongates (Becker and Manoli, 2004), da diese die maximale Bandbreite des FPAA beschneiden (Gray et al., 2005).

2 Architektur des FPAA

Der FPAA basiert auf einer hexagonalen Topologie wie in Abb. 1 gezeigt. Die dicken, roten Kreise symbolisieren die Knoten des Chips. Diese sind durch programmierbare G_m -Zellen mit den Nachbarknoten verbunden, dargestellt durch die schwarzen Pfeile. Jeder Knoten kann also über eine programmierbare G_m -Zelle von den sechs umgebenden Knoten erreicht werden. Zusätzlich gibt es eine programmierbare G_m -Zelle, die den Knoten mit sich selbst verknüpft, um die Selbstrückkoppelung zu ermöglichen. Das Routen eines Signals durch den Chip ist durch passende An- oder Abschaltung der programmierbaren G_m -Zellen erreichbar, dadurch sind keine Transmissiongates notwendig. Jede programmierbare G_m -Zelle zusammen mit der Integrationskapazität an den Knoten stellt einen G_m -C Filter dar. Diese Kapazität ist die Summe der parasitären Aus- und Eingangskapazitäten der G_m -Zellen. Eine zusätzliche Kapazität ist nicht notwendig. Die hexagonale Topologie zusammen mit der Selbstrückkoppelung erlaubt den Aufbau von Filtern beliebiger Ordnung.

Jeder Knoten, zusammen mit den von ihm ausgehenden programmierbaren G_m -Zellen, formt einen so genannten configurable analog block (CAB), zu sehen in Abb. 2. Die grauen Bauteile gehören zu den Nachbarknoten. Die Ausgänge der von den Nachbarknoten kommenden programmierbaren G_m -Zellen sind alle mit dem zentralen Knoten verbunden (der schwarze Ring in Abb. 2). An diesen Ring sind auch die Eingänge der wegführenden programmierbaren G_m -Zellen, und der Ein- und Ausgang der Selbstrückkoppelung angeschlossen.

Correspondence to: F. Henrici
(henrici@imtek.de)

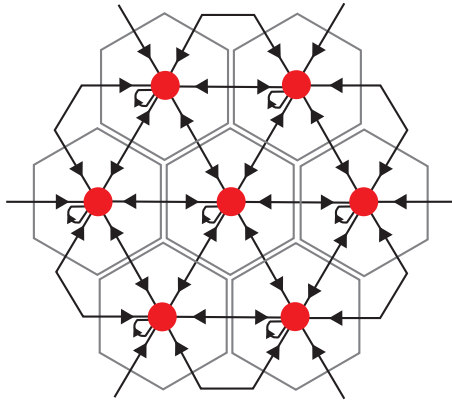


Abb. 1. Schaltplan der Chiparchitektur.

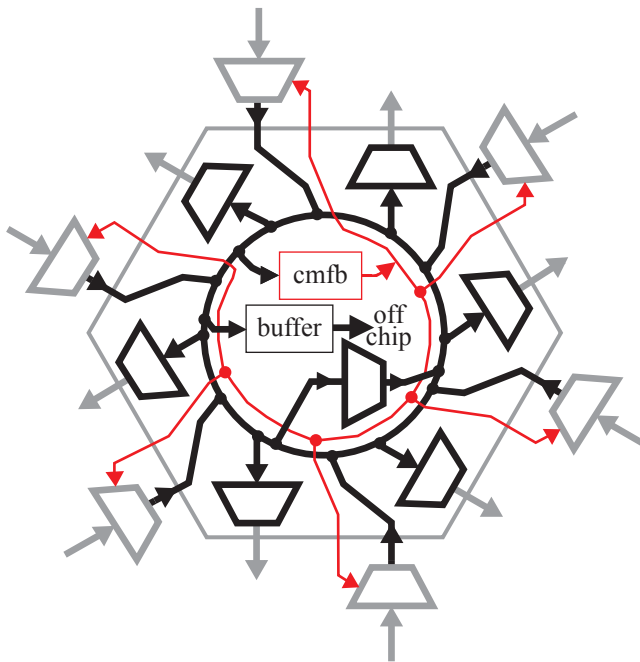


Abb. 2. Schaltplan eines configurable analog blocks (CAB).

Um den Dynamikbereich, die total harmonic distortion und das Signal/Rauschverhältnis (SNR) zu verbessern, ist der komplette FPAA voll differentiell aufgebaut. Dies macht ein common mode feedback (CMFB) notwendig, gezeigt in Abb. 2 in dünn und rot. Die zum Knoten führenden programmierbaren G_m -Zellen sind alle parallel geschaltet, daher haben sie alle den gleichen common mode level (CM) und können von einer einzigen CMFB gesteuert werden. Diese CMFB misst den CM am zentralen Knoten ihres CABs und steuert die G_m -Zellen aller umliegenden CABs, die zu ihrem Knoten führen. In jedem CAB wird ein Buffer eingesetzt, der das Signal an diesem Knoten an die Ausgänge (Pads) des Chips führt. Der Buffer ist notwendig um den Knoten vor resistiver Belastung von ausserhalb des Chips zu schützen.

Eine programmierbare G_m -Zelle ist aus 12 so genann-

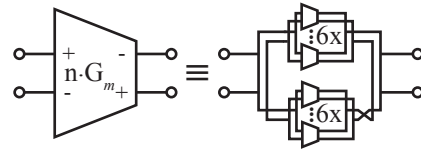


Abb. 3. Schaltplan einer programmierbaren G_m -Zelle.

ten Einheits- G_m -Zellen zusammengesetzt, siehe Abb. 3. Jede dieser Einheits- G_m -Zellen kann mit einem einzelnen Bit ein- oder ausgeschaltet werden. Eine kleine Logik zieht dann die Bias Potentiale der PMOS und NMOS Transistoren gegen VDD, bzw. GND. Die Einheits- G_m -Zellen sehen alle die gleiche Eingangsspannung und produzieren – falls eingeschaltet – den gleichen Ausgangsstrom. Da ihre Ausgänge ebenfalls parallelgeschaltet sind, addieren sich ihre Ausgangsströme. Dadurch wird das effektive G_m des Verbundes linear abhängig von der Anzahl der eingeschalteten Zellen. Jede Einheits- G_m -Zelle hat einen bestimmten, endlichen Ausgangswiderstand falls eingeschaltet, und einen näherungsweise unendlichen falls ausgeschaltet. Das G_m steigt also linear mit der Anzahl eingeschalteter Einheits- G_m -Zellen, während der Ausgangswiderstand um den gleichen Faktor sinkt. Da die Leerlaufverstärkung das Produkt beider ist bleibt sie konstant.

Sechs der Einheits- G_m -Zellen sind direkt mit dem Ausgang verbunden, die anderen sechs mit invertierten Ausgängen. Dies erlaubt die Inversion des Eingangssignales mit normalen G_m -Zellen und ohne Einsatz von Transmissiongates (Becker and Manoli, 2004).

3 Die parasitären Kapazitäten

In den meisten Designs ist die Integrationskapazität eines G_m -C Filter aus einer festen Kapazität C_{intcap} und den verschiedenen parasitären Kapazitäten der MOS Transistoren $C_{parasitic}(W, L)$ zusammengesetzt:

$$C_{int} = C_{intcap} + C_{parasitic}(W, L) \tag{1}$$

Eine genauere Betrachtung dieser Kapazitäten ist wichtig. Ohne Beschränkung der Allgemeinheit wird im folgenden eine gefaltete Kaskode angenommen, der Miller Effekt ignoriert, und angenommen, dass PMOS und NMOS Transistoren die gleiche Breite haben. Die für den FPAA signifikanten parasitären Kapazitäten $C_{parasitic}$ sind die Eingangs- und Ausgangskapazitäten C_{in} und C_{out} der Einheits- G_m -Zellen (Razavi, 2001):

$$C_{in} = C_{GS} + C_{GD} = W \cdot \left[\frac{2}{3} \cdot L \cdot C_{OX} + 2 \cdot C_{OV} \right]$$

$$C_{out} = C_{outp} + C_{outn} = 2 \cdot (C_{DG} + C_{DB}) \tag{2}$$

$$= W \cdot 2 \cdot [C_{OV} + E \cdot C_J + 2 \cdot C_{JSW}] + 4 \cdot E \cdot C_{JSW}, \quad \text{mit } E: \text{ Drain Länge}$$

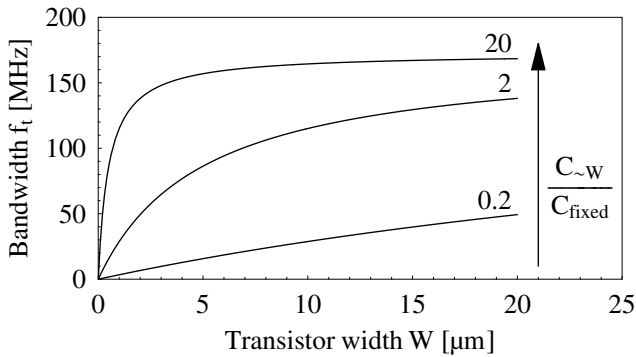


Abb. 4. Abhängigkeit der Bandbreite von den Transistordimensionen. Alle Biasströme und Transistorbreiten werden linear mit der Breite W der Eingangstransistoren skaliert.

Es zeigt sich, dass, zumindest in unserem Fall, die parasitären MOS Kapazitäten fast komplett proportional zur Breite W der Transistoren sind. Die Länge der Transistoren wird so klein wie möglich gewählt, aber groß genug um einen akzeptablen Ausgangswiderstand und SNR zu erreichen. (1) kann umgeformt werden um die Abhängigkeit von einem konstanten Anteil C_{fixed} und einem von W abhängigen Anteil $C_{\alpha W}$ zu zeigen:

$$C_{int}(W) = C_{fixed} + C_{\alpha W} \cdot W \quad (3)$$

Um die Bandbreite des FPAA's zu erhöhen wird auf C_{intcap} verzichtet, stattdessen werden nur die parasitären Kapazitäten genutzt. So lange alle Transistoren in Sättigung bleiben, hat die Spannungsabhängigkeit dieser Kapazitäten weniger Einfluss auf die Linearität als die des G_m s der Zellen.

Daraus resultiert ein großes Verhältnis von W abhängiger Kapazität $C_{\alpha W}$ zu konstanter C_{fixed} . Dies hat eine interessante Auswirkung auf die erreichbare Bandbreite. (4) zeigt die Grenzfrequenz eines G_m -C Filters und das G_m einer G_m -Zelle. Zusammen mit (2) kann man die Abhängigkeit der Grenzfrequenz von der Transistorbreite W zeigen:

$$f_t = \frac{G_m}{2 \cdot \pi \cdot C_{int}}, \quad \text{mit } G_m = \sqrt{2 \cdot \frac{W}{L} \cdot \beta_0 \cdot I_{BIAS}} \quad (4)$$

$$f_t = \frac{\sqrt{W \cdot I_{BIAS}}}{C_{fixed} + C_{\alpha W} \cdot W} \cdot \sqrt{\frac{\beta_0}{2 \cdot \pi^2 \cdot L}} \quad (5)$$

Es gibt zwei Möglichkeiten, das G_m – und damit die Bandbreite – zu erhöhen. Entweder man vergrößert das β der Eingangstransistoren indem man ihre Breite W erhöht, oder man erhöht den Biasstrom I_{BIAS} der durch sie hindurch fließt. Ein erhöhter Strom durch einen Transistor zieht allerdings auch ein erhöhtes V_{GS} nach sich, und verringert damit den Dynamikbereich. Will man diesen konstant halten, muss man die Breite der Transistoren proportional mit dem Biasstrom erhöhen. In einem System mit einer großen, konstanten Integrationskapazität hat eine Verbreiterung der Transistoren

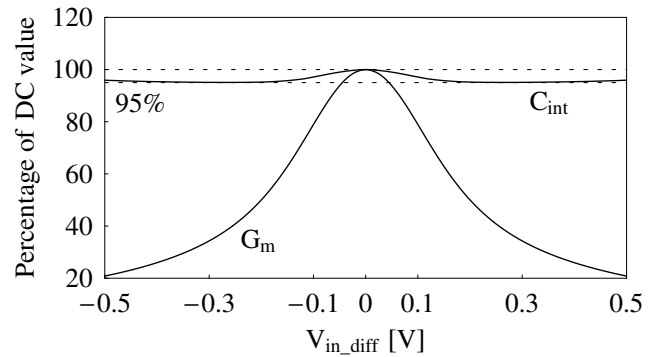


Abb. 5. Spannungsabhängigkeit der parasitären Kapazitäten und G_m in Prozent des DC Wertes. Die größte Abweichung von C_{int} beträgt minus fünf Prozent.

nur geringen Einfluss auf die Gesamtkapazität. Die Bandbreite skaliert mit \sqrt{W} falls nur β (W) erhöht wird, und mit W falls der Biasstrom proportional zu W erhöht wird. Falls aber der konstante Anteil der Kapazität sehr klein ist, würde Bandbreite sogar abnehmen, falls nur β (W) erhöht wird, und konstant bleiben, falls β und Biasstrom erhöht werden. Zur Verdeutlichung ist (5) für verschiedene Verhältnisse in Abb. 4 aufgetragen.

Durch Weglassen der Integrationskapazität und ausschließlicher Nutzung der parasitären Kapazitäten der aktiven Elemente kann man eine hohe Effizienz von Bandbreite zu Stromverbrauch erreichen. Andererseits kann man mit einem höheren Stromverbrauch keine Bandbreitenerhöhung mehr erreichen. Es lässt sich folgern, dass Transistorabmessungen so klein wie möglich gewählt werden sollten, außer man möchte ein geringeres Rauschen und höhere Linearität erreichen.

Ein damit verbundenes Problem dieses schalterlosen Designs ist, dass die Anzahl von Einheits- G_m -Zellen, die eine programmierbare G_m -Zelle formen, die maximale Bandbreite nicht beeinflussen, sondern nur die Einstellmöglichkeiten verfeinern: Angenommen eine programmierbare G_m -Zelle würde nur aus zwei Einheits- G_m -Zellen bestehen, dann hätte sie zwar das doppelte G_m im Vergleich zu einer, die nur aus einer Einheits- G_m -Zelle besteht, aber auch die doppelte parasitäre Ein- und Ausgangskapazität. Die Übertragungsfunktion für parallelgeschaltete Einheits- G_m -Zellen ist in (6) gegeben, mit m der Anzahl vorhandener Einheits- G_m -Zellen, und n die Anzahl derer, die davon eingeschaltet sind.

$$T(s) = \frac{A}{1 + \frac{s}{2 \cdot \pi \cdot f_t}}, \quad \text{with } f_t = \frac{n \cdot G_m}{2 \cdot \pi \cdot m \cdot C_{int}} \quad (6)$$

Es ist zu sehen, dass, falls alle Zellen eingeschaltet sind, sich n und m gerade aufheben. Die Bandbreite kann also nicht durch zusätzliche Einheits- G_m -Zellen erhöht werden. Die einzige Möglichkeit, die Bandbreite des Schaltkreises zu erhöhen, ist daher die Architektur zu ändern oder das Zell-design zu verbessern. Es wurde daher ein Einheits- G_m -Zelle mit verkleinerten Eingangstransistoren entwickelt.

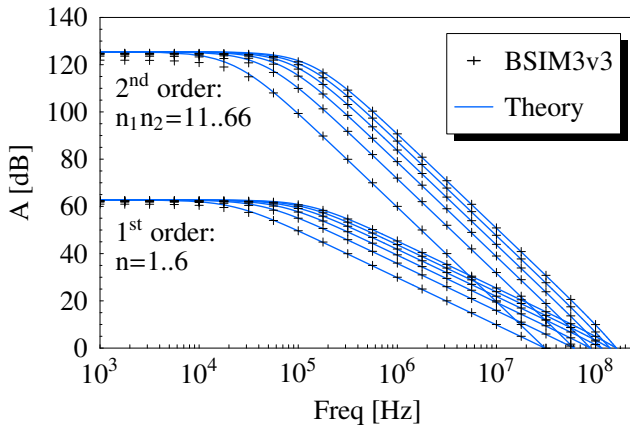


Abb. 6. Simulationsergebnisse eines Tiefpasses erster und zweiter Ordnung.

4 Linearität der parasitären Kapazitäten und des G_m

Um die Gesamtkapazität an einem Knoten C_{int} zu bestimmen, wurden AC Simulationen durchgeführt, bei denen eine Sinusquelle den Knoten über einen Serienwiderstand treibt. Die effektive Kapazität kann damit aus der RC Zeitkonstante bestimmt werden. Bei allen Simulationen wurde der ganze Chip simuliert um sämtliche parasitären Effekte zu berücksichtigen. Das Result in Prozent des DC Wertes ist in Abb. 5 zu sehen. Es zeigt sich, dass die Variation sowohl in G_m als auch in C_{int} im Arbeitsbereich zwischen -5 mV und $+50$ mV gering ist.

Die Spannungsabhängigkeit der parasitären Kapazitäten ist deutlich geringer als die von G_m . Dafür gibt es zwei Hauptgründe. Zum einen ist der Hauptbeitrag zu C_{int} die Gatekapazität C_{GS} , diese bleibt aber praktisch konstant so lange der Transistor in Sättigung ist (Razavi, 2001). Die zweite Erklärung liegt im voll differentiellen Design. C_{int} besteht eigentlich aus zwei Kapazitäten, eine für jeden differentiellen Pfad. Wenn die Spannung an einem Pfad steigt, sinkt sie am anderen. Zumindest für kleine Spannungsänderungen, kann die Spannungsabhängigkeit der Kapazitäten aber als linear angenommen werden, daher heben sich die zwei Kapazitätsänderungen mit entgegengesetzten Vorzeichen auf.

5 Simulationsergebnisse

Alle Simulationen umfassen den gesamten Chip, da nur so die parasitären Kapazitäten korrekt simuliert werden können (Becker et al., 2005). Die Simulationen in Abb. 6 zeigen die Übertragungsfunktion eines einzelnen G_m -C Filters, und die von zwei hintereinandergeschalteten G_m -C Filtern, die damit einen Tiefpass zweiter Ordnung darstellen. Die BSIM3v3 Simulationen zeigen gute Übereinstimmung mit der Theorie. Bei der Reihenschaltung sind die G_m -C Filter jeweils auf das gleiche G_m eingestellt. Die idealisierte Übertragungsfunkti-

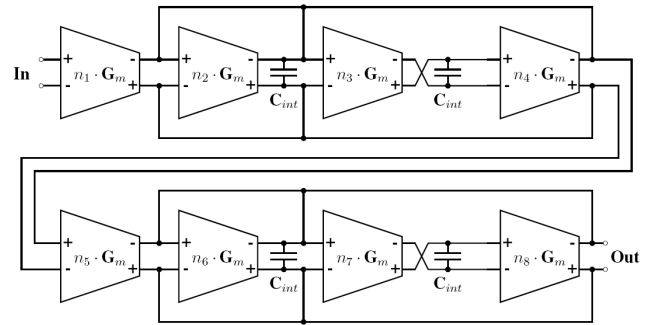


Abb. 7. Zwei kaskadierte Biquads formen einen Bandpass vierter Ordnung. Es werden vier Chipknoten benötigt.

on erster und zweiter Ordnung ist:

$$T_1(s, n) = \frac{A}{1 + \frac{s}{n \cdot \omega_0}}, \quad T_2(s, n_1, n_2) = T_1(s, n_1) \cdot T_1(s, n_2) \quad (7)$$

Die Transitfrequenz einer programmierbaren G_m -Zelle skaliert linear mit n . Eine aus 12 Einheits- G_m -Zellen mit gefalteter Kaskode aufgebaute programmierbare G_m -Zelle erreicht eine Maximalbandbreite von 164 MHz, bei einer unteren Grenze von 31.5 MHz.

Um die Fähigkeiten der Architektur zur Instanziierung komplexer Strukturen zu verdeutlichen, ist ein Bandpass vierter Ordnung, basierend auf zwei kaskadierten Biquads wie in Abb. 7 gezeigt, simuliert worden. Es werden lediglich vier Knoten benötigt. Die Selbstrückkoppelung stellt den simulierten Widerstand, die Kapazität ist C_{int} , und eine Verbindung zum nächsten Knoten und zurück liefert einen Gyrtator, also eine simulierte Induktivität. Die idealisierte Übertragungsfunktion für einen Biquad - einen Bandpass zweiter Ordnung - in Abhängigkeit von n lautet:

$$T_{BP2}(s) = \frac{\frac{A_0 \cdot s}{Q \cdot \omega_t}}{1 + \frac{1}{Q} \cdot \frac{s}{\omega_t} + \frac{s^2}{\omega_t^2}}, \quad A_0 = n_1 \cdot G_m^2, \quad (8)$$

$$\omega_t = \sqrt{\frac{n_3 \cdot n_4}{2 \cdot G_m^2}}, \quad Q = C_{int} \cdot n_2 \cdot \sqrt{2 \cdot n_3 \cdot n_4}$$

Die Verstärkung, Mittenfrequenz und Güte sind in Abhängigkeit von n angegeben. Obwohl die Kapazität nicht variabel ist, können doch alle Parameter getrennt eingestellt werden. Um eine genauere Vorhersage zu ermöglichen, müssen die endliche Ausgangsverstärkung und ein parasitärer Pol bei ungefähr der doppelten Transitfrequenz modelliert werden. Der Pol stammt vom internen Knoten der gefalteten Kaskode. Ein einfaches differentielles Paar hätte diesen Pol nicht, aber die Kaskode bietet ein ausgewogeneres Verhältnis zwischen Bandbreite und Ausgangswiderstand.

Um die Stromsummation am Knoten zu berücksichtigen, wurde das Modell in zwei Übertragungsfunktionen zerlegt. Die erste modelliert das frequenzabhängige G_m der programmierbaren G_m -Zelle (9), die zweite das Verhalten des Knotens (10). Letztere enthält den dominanten Pol von C_{int}

Tabelle 1. Parameter der verwendeten Einheits- G_m -Zelle.

| C_{int} [pF] | G_m [mS] | A_0 [dB] | p_2 [MHz] | r_{out} |
|-----------------------|------------|------------|-------------|-------------------|
| 14.8 | 2.87 | 62.7 | 293 | $\frac{A_0}{G_m}$ |

und die mehrfachen g_{out} mit denen jeder aktivierte Einheits- G_m -Zelle den Knoten belastet. Dafür müssen die n_i aller programmierbaren G_m -Zellen i , die an diesen Knoten angeschlossen sind, addiert werden und mit dem g_{out} eines Einheits- G_m -Zellen multipliziert werden.

$$T_{\text{cell}}(s) = n \cdot G_m \cdot \frac{1}{\frac{s}{2\pi \cdot p_2} + 1} \quad (9)$$

$$T_{\text{node}}(s) = \frac{1}{s \cdot C_{\text{int}} + \frac{1}{r_{\text{out}}} \cdot \sum_{i=1}^7 n_i} \quad (10)$$

Die Werte der Parameter einer Einheits- G_m -Zellen wurden aus den Simulationen extrahiert und sind in Tab. 1 angegeben.

Die Mittenfrequenz des Bandpasses lässt sich in einem weiten Bereich zwischen 60 und 160 MHz einstellen, siehe Abb. 8. Zur Ermittlung der Linearität der Zellen wurde ein 50 mV Signal bei 1 MHz durch einen verlustbehafteten Integrator, eingebettet in den Chip, geschickt. Dieser besteht aus einer Reihenschaltung einer programmierbaren G_m -Zelle und einem simulierten Widerstand, eingestellt war eine Bandbreite von 164 MHz und eine Verstärkung von eins. Die HD3 betrug weniger als -70 dB.

6 Zusammenfassung

Es wurde ein neuartiger FPAA vorgestellt, der ohne Transmissionsgates im Signalpfad auskommt, und nur parasitäre Kapazitäten in den G_m -C Filtern nutzt. Simulationen haben gezeigt, dass die Linearität dieser Kapazitäten ausreichend groß ist, bezogen auf die Linearität von G_m . Die Diskussion der Abhängigkeit der Bandbreite von den Transistordimensionen zeigt, dass eine höhere Bandbreite nur durch besseres Design, nicht aber durch mehr Strom erreicht werden kann. Dafür wurde eine G_m -Zelle mit verkleinerten Eingangstransistoren entwickelt. Der Testchip umfasst sieben Knoten in einem hexagonalen Gitter. Diese reichen aus, um drei Bi-quads zu instantzieren, dabei verbraucht der gesamte Chip weniger als 75 mW trotz einer erreichten Bandbreite von 164 MHz. Ausserdem wurde eine hohe Linearität mit einer HD3 geringer als -70 dB erzielt. Der Chip wurde zur Fertigung in einem 130 nm 1.2 V CMOS Prozess submittiert.

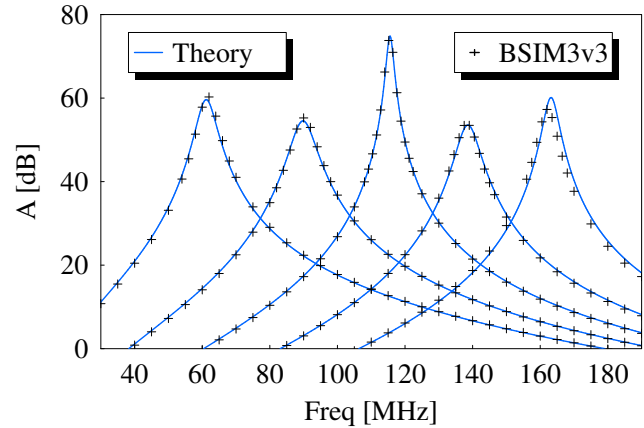


Abb. 8. Simulationsergebnisse des Bandpasses vierter Ordnung aus Abb. 7. Die Mittenfrequenz kann zwischen 60 und 160 MHz eingestellt werden.

Literatur

- Becker, J. and Manoli, Y.: A continuous-time field programmable analog array (FPAA) consisting of digitally reconfigurable Gm-cells, in: Proc. ISCAS'04, 1, 1092–1095, doi:10.1109/ISCAS.2004.1328389, 2004.
- Becker, J., Henrici, F., and Manoli, Y.: System-level analog simulation of a mixed-signal continuous-time field programmable analog array, in: Proc. IWSOC'05, 434–438, doi:10.1109/IWSOC.2005.102, 2005.
- Gray, J., Twigg, C., Abramson, D., and Hasler, P.: Characteristics and programming of floating-gate pFET switches in an FPAA crossbar network, in: Proc. ISCAS'05, 468–471, doi:10.1109/ISCAS.2005.1464626, 2005.
- Pavan, S. and Tsividis, Y.: High Frequency Continuous Time Filters in Digital CMOS Processes, Kluwer Academic Publishers, Boston, 2000.
- Razavi, B.: Design of Analog CMOS Integrated Circuits, McGraw-Hill, Boston, 2001.