Design Approach for Ring Amplifiers

Joschua Conrad, *Student Member, IEEE*, Patrick Vogelmann, *Student Member, IEEE*,
Mohamed Aly Mokhtar, *Student Member, IEEE*, and Maurits Ortmanns, *Senior Member, IEEE*

Abstract— Ring amplifiers are an advantageous new amplifier architecture, but designing them needs to be done in a time consuming process using manual transient simulations. AC based design criteria are of limit usability, because of their large signal operation. This paper presents an alternative design approach for ring amplifiers. A set of cost functions and parameters is developed and the parameters' effect on the amplifier's stability, accuracy and power consumption is explained in detail. The optimization process is derived from a simple ring-amplifier model and a stability criterion. This criterion is analyzed in the context of circuit parameters, design constraints and other ring-amplifier designs. A new battery-based self-biased ring-amplifier architecture is introduced and the first stage integrator of a switched-capacitor Delta-Sigma modulator is realized using the described ring-amplifier architecture. The proposed optimization process is successfully performed using 180 nm and 40 nm technology nodes (second node with two supply voltages) and the three designs are analyzed and compared. The optimization process is derived from a basic ring-amplifier model and can be applied to other ring-amplifier structures.

Index Terms— RAMP, ring amplifier, ring amplification, incremental delta sigma modulator, IDSM, discrete time, optimization process, monticelli, battery, floating current source, switched capacitor, CMOS analogue integrated circuits, inverter based amplifier, ADC, data converter.

I. INTRODUCTION

RING amplifiers (RAMPs) are an emerging amplifier structure for switched-capacitor (SC) circuits. The amplifier's basic building element is an inverter. The inverter chain is forked, where different offset voltages are applied in each branch such that the inverters do not form an oscillator.

RAMPs scale well with scaled CMOS technologies, provide a rail-to-rail output and utilize slew-based charging of the load capacitor C_l and the feedback network [1]. There is an increasing interest in this type of amplifier as can be seen by the number of recent publications.

In [1], the theory behind RAMPs is explained, constraints for a stable amplifier are derived and it is explained how to exploit advantages of split correlated level shifting (CLS) [2], [3]. The work of [4] explains the design of fully differential RAMPs and [5] proposes a technique to replace the offset-voltage capacitors by a resistor. Finally, [6] and [7] identified the size of the offset voltage as a key parameter and provide techniques to tune this voltage dynamically. However, to the best of the author's knowledge, there is no intuitive and straightforward way or an automated method for designing RAMPs. Time-consuming transient simulations and manual tuning on circuit level has to be done to obtain a stable and efficient amplifier. This work proposes a design process based on optimizations and cost functions, which allows to do the design-space exploration using transient simulations automatically.

The paper is structured as follows: Section II reviews the settling mechanism and the condition for a RAMP to be stable and explains in detail the relations between stability issues, RAMP parameters and design constraints, which is needed for the reader to be able to follow the subsequent optimization trade-offs. Section III introduces the circuit architecture of the battery-biased RAMP, which is then used to evaluate the design approach. Section IV derives the circuit-optimizer-based design approach from the simple RAMP model described in Section II, which sizes all circuit elements depending on given constraints. The optimized parameters are again linked to the stability criterion from Section II. Three RAMPs in 180 nm at 3 V, 40 nm at 2.5 V and 40 nm at 1.1 V are designed and tested in an incremental Delta-Sigma (I-$\Delta\Sigma$) modulator. The results of the optimization and the performance of the I-$\Delta\Sigma$ modulators are shown in Section V. Section VI summarizes this work.

II. STABILITY IN RING AMPLIFIERS

A basic RAMP with capacitive feedback network is shown in Fig. 1 [1]. The amplification phase is visible and the amplifier itself is shown inside the dashed box. The amplifier consists of three inverter stages A_1, A_2 and M_p/M_n. Without any modifications, the chain of inverters would start to oscillate when applying feedback. Therefore, a dead-zone voltage V_{os} is stored on the two capacitors C_{dz1} and C_{dz2} acting as constant-voltage sources during the amplification phase. The potential of the node V_{a2p} is lifted and the potential of V_{a2n} is lowered. The output inverter is pushed towards a state, where both transistors are non-conducting. This state is called the dead zone.

A. Settling Mechanism

The settling is explained using an exemplary pseudo-differential version of the circuit in Fig. 1. It is investigated in a transient simulation. A single-ended amplifier would not reveal
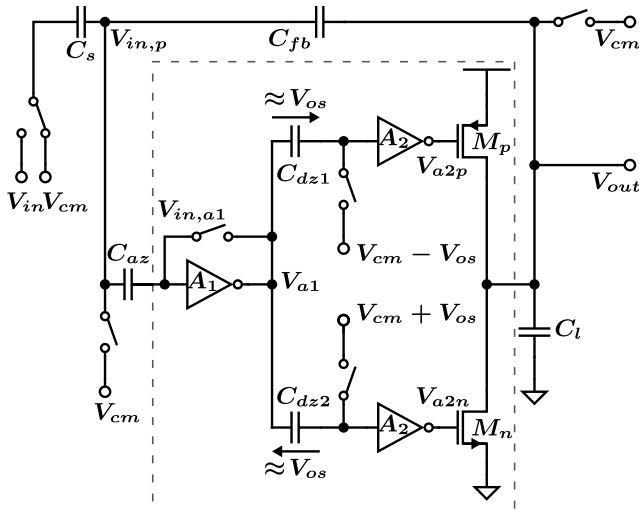
Fig. 1. Basic RAMP structure as introduced in [1]. Switch positions during amplification phase.

Fig. 2. Exemplary transient settling process showing the differential output of different stages.

all effects due to asymmetry. Fig. 2 shows the differential output potentials of the three stages. This illustration example uses 200 fF capacitors. The gain of A_1 and A_2 is 20 dB each. A delay time $t_d = 100\,ps$ is added to A_1 and marked by the vertical red bar. t_d is also sufficient for modeling the limited slew rate of a digitally operated inverter stage. The delay time of A_2 is neglected for the sake of simplicity, even though it contributes to t_d in the later analyses. The voltage V_{os} is 500 mV and the input-referred dead zone is labeled using the horizontal red bar. All capacitors and switches are ideal elements in this teaching example. A_1 and A_2 are implemented using voltage-controlled voltage source (VCVS) elements and delay primitives. This exemplary model is only used in Section II. The exemplary amplifier is designed to operate at the edge of stability in order to visualize the settling mechanism.

During the initial auto-zeroing (AZ) phase, the offset error of the first stage is counteracted using the voltage stored on capacitor C_{az}. A_1 is driven to its trip point and the dead zone is stored on C_{dz1} and C_{dz2} with respect to this trip point. The input signal is sampled on C_s and C_{fb} is reset. This prepares the RAMP for the amplification phase, which is divided in three sub-phases:

a) Initial ramping: The first phase starts after applying a step to the node $V_{in,a1}$ by switching from the AZ phase to the amplification phase. As the output of the amplifier is a high-impedance node, a portion of the step feeds through to the output. It takes $t_d = 100\,ps$, until the step is visible at the output of A_1, which drives the first stage to one rail, from where it is passed to the output of A_2. V_{a2p} switches M_p on, such that the transistor drives a current to the output, which is only limited by the large-signal channel resistance of the transistor, whereas M_n is switched off. The slew current creates a ramp shaped output voltage V_{out}, which feeds back to the node $V_{in,a1}$. Now, the first stage is driven towards the trip point of A_1. Without the delay, the output would be switched off when entering the dead zone and the amplifier would be
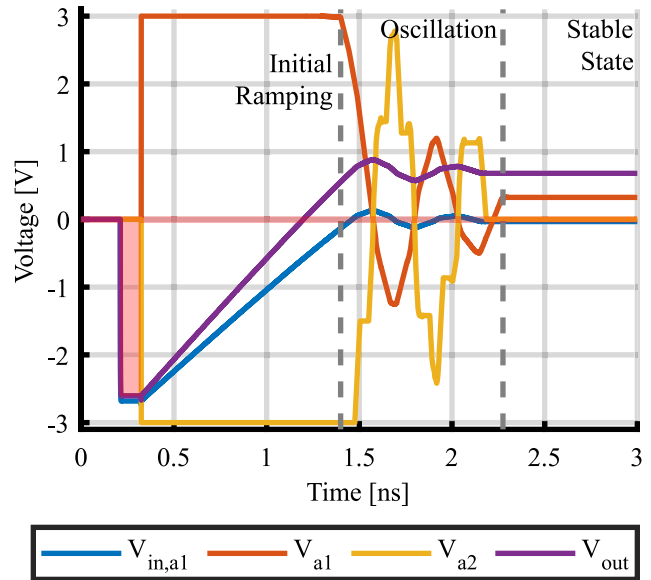
settled. With the delay, the output continues to slew for t_d after entering the dead zone. $V_{in,a1}$ leaves the dead zone, where the output stage is switched off, which results in an output current with opposite sign. The larger V_{os}, the larger the dead zone and the larger the maximum overshoots that still allow for a stable operation.

b) Oscillation: A stable RAMP starts to oscillate with decreasing amplitude around the dead zone. V_{a2}, which describes the differential output voltage as

$$V_{a2} = V_{a2p} + V_{a2n} - V_{cm} \qquad (1)$$

shows steps. This is due to different slew currents, when one pseudo-differential branch slews using the NMOS in one direction, the other branch using the PMOS into the other direction. This causes a different time needed for the initial slewing to end in each branch. The oscillations in the two branches start at different time instants, which causes common-mode fluctuations. The effect is explained in more detail in Section V-D.

c) Stable state: The oscillation amplitudes become steadily smaller, until the following oscillation amplitude of the internal signals stays inside the dead zone. The upcoming oscillation period is not able to switch the output stage to the conducting state due to the input-referred undershoot being too small. The amplifier has settled. No output current alters the voltages in the capacitive feedback network anymore.

B. Stability Criterion and Considerations

The initial ramping happens with the maximum possible overdrive for the output stage. If the first undershoot does not drive V_{a2} to the rail, the overdrive for the second ramping in the opposite direction is smaller than for the initial ramping. Because of the smaller slew current, the overshoot voltage at $V_{in,a1}$ caused by this slewing is smaller and the overdrive of the output stage at the next reversal point is even smaller.

The oscillation amplitude decreases. In [1] a stability criterion is derived from that explanation:

$$\frac{t_d\, I_{slew}}{\psi\, C_l} \leq \left(\frac{V_{dd} - V_{ss}}{A_2} - 4\, V_{os}\right) \frac{1}{2\, A_1} \tag{2}$$

Here, t_d is the overall time delay from $V_{in,a1}$ to V_{a2} and ψ is the feedback factor of the amplifier. C_l is the load capacitor and I_{slew} the maximum slew current of the output stage. V_{dd} and V_{ss} are the respective supply rails. A_1 and A_2 is the gain of the first and the second inverter stage, respectively. V_{os} is the dead-zone voltage. All parameters influence the amplifier's stability, accuracy and power consumption:

d) Offset voltage: The voltage V_{os} determines the size of the dead zone. The size of V_{os} needs to be referred to the input of the RAMP with

$$V_{dz} = 2\, V_{os} \tag{3a}$$
$$V_{dz,in} = \frac{V_{dz}}{A_1} = \frac{2\, V_{os}}{A_1}. \tag{3b}$$

After $V_{in,a1}$ enters the dead zone, the output continues to slew for the time t_d. If the input is still in the dead zone after that time, the amplifier is settled. This means that if the dead zone is large, it is easier to obtain a stable amplifier. On the other side, $V_{dz,in}$ also describes the uncertainty of the settled voltage $V_{in,a1}$, which is similar to an input referred error as described in Section II-D. It is thus easier to design a stable RAMP, if the required accuracy is small. The parameter V_{os} has an indirect influence on the power consumption: if the amplifier is designed for high accuracy, t_d has to be reduced to avoid large overshoots that make the amplifier unstable.

e) Delay time: t_d describes the large-signal delay of the stages A_1 and A_2. A_1 usually dominates this delay time, because A_2 needs to provide the dead-zone offsets at its outputs, which moves each inverter in A_2 towards the triode operating region, where the inverter has smaller gain and a larger bandwidth [4]. After $V_{in,a1}$ enters the dead zone during the initial ramping, the output continues to slew for t_d, until the output stage is switched to a current with the opposite sign. Therefore, t_d determines the severeness of the first over- or undershoot and needs to be small for a stable amplifier. The parameter can be used to tune the amplifier towards stability by choosing a smaller technology node or by spending more power in the inverter [1].

f) Slew rate: The factor $\frac{I_{slew}}{C_l}$ in Section II-B refers to the slew rate during the initial ramping. If this parameter is too small, the settling is too slow. The slew rate is constrained by the clock speed of the SC design. Around the stability determining settling point, the output voltage V_{out} overshoots with this slew rate $\frac{I_{slew}}{C_l}$ for the amplifier's delay time t_d. If the amplifier is intended to be fast, its slew rate is increased and the amplifier tends to be unstable. This needs to be compensated by reducing t_d, which costs power. The output stage can be considered to be off in the settled state and its slew rate is not determined by any biasing current [1].

g) Feedback factor: ψ is determined by the feedback network around the RAMP and describes the ratio of V_{out} to $V_{in,a1}$. If the factor is large, the SC amplifier has a large gain, the first overshoot of V_{out} translates to a small overshoot
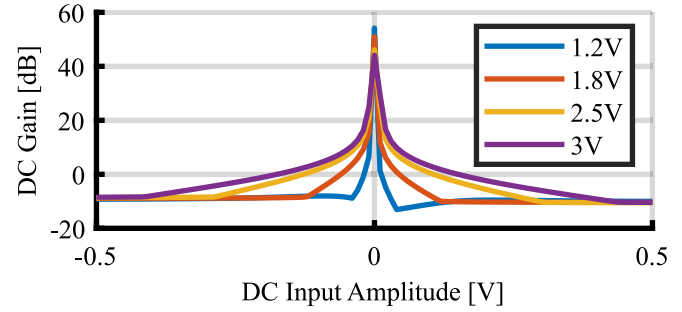

Fig. 3.　Inverter small-signal gain for multiple supply voltages.

of the input voltage $V_{in,a1}$, which tends to drive V_{a2p} and V_{a2n} less towards V_{dd} or V_{ss}. Thus, a large feedback factor creates stable RAMPs according to Section II-B. Reference [8] further analyzes the impact of the feedback factor.

h) Gain: A_1 and A_2 describe the AC gain of the first two inverter stages at the first reversal point, which is called "effective gain" [1]. If the two gains are large, an overshoot in $V_{in,a1}$ is translated into a large overshoot in V_{a2p} and V_{a2n} and a large slewing in the upcoming oscillation period, which makes the amplifier unstable. The two gains should thus be small for stability, but are increased when designing the first stage for small noise or when employing linear settling (see Section II-D). Disadvantageously, designing the two stages for a larger gain while maintaining the speed makes them consume more power. Also, the increased gains make the amplifier unstable, which needs to be compensated with a smaller t_d, which again costs power. A_2 is usually smaller than A_1 due to its triode operating region [4].

i) Supply: The supply rails $V_{dd} - V_{ss}$ define the potentials, which need to be reached by V_{a2p} and V_{a2n} at the first reversal point for making the amplifier unstable. If the span is wide, this is unlikely and the amplifier tends to be more stable. The supply voltage is defined by the used technology. Small technology nodes provide a smaller t_d (good for stability) and also a smaller supply voltage (bad for stability). The positive influence of t_d is though larger, which makes RAMPs scale well with recent technologies [1].

Fig. 3 shows the small-signal DC gain of a single inverter designed in 180 nm for different supply voltages and different large-scale inputs. All inverters have the same NMOS size and a PMOS designed for a mid-supply trip point. The x-axis describes an inverter's large-scale input signal, as it fluctuates during the settling. It does not describe an offset visible in a sampled output of an SC circuit. The plot shows that inverters with a smaller supply voltage provide a larger maximum gain because of smaller current, larger output resistance and more weak inversion operation. For small supply voltages, the devices also operate closer to the triode region and already a small input offset pushes one transistor into the triode region, where the gain is reduced. This behavior is desired: the gain at the trip point is large supporting accurate settling; when leaving the trip point, e.g. at the first reversal point during the settling, the effective small signal gain is reduced, which determines the gain in the stability criterion Section II-B.
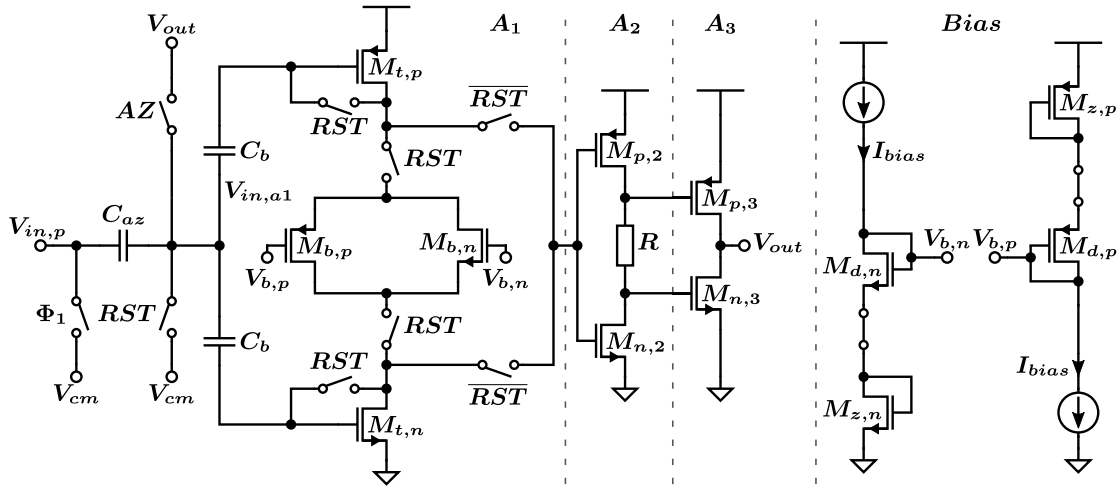
Fig. 4. The redesigned battery-self-biased RAMP. One pseudo-differential half branch with the biasing network is shown.

The effect is the most important in the first stage, because A_1 is larger than A_2. A biasing current, which defines a gate-source voltage close to the threshold voltage in the first stage shows the same stabilizing effect. Such a biasing structure was described in [4], [9].

Unfortunately, Section II-B is not really practical for designing a RAMP. Not only exist many counteracting arguments, as outlined above, also the amplifier AC gains A_1 and A_2 noticeably change during the large-signal excitation. Moreover, the voltage V_{os}, which determines the dead-zone size, is also signal dependent in advanced RAMP architectures [4], [5], [7] and t_d can also depend on the propagated signal. This makes a design-by-equation cumbersome and not realizable.

C. ADC Design Constraints

This section analyzes the relationships between symbols in Section II-B and design constraints in the context of recent RAMP designs [1]–[7].

The speed of an ADC sets a requirement for $\frac{I_{slew}}{C_l}$. Fast ADCs require a large slew rate, which requires a larger t_d compensation. Accurate ADCs require a small dead-zone voltage V_{os} and large AC gains A_1 and A_2 (see Section II-D), which again requires a t_d compensation. "Split-CLS" decouples the slew-rate requirement and the accuracy requirement by splitting the RAMP into a fast, but inaccurate RAMP with a large dead zone and a second RAMP in parallel with a small slew rate and a small dead zone. Both amplifiers achieve stability without making excessive use of t_d based compensation. This technique is used in all RAMP based ADCs with an effective number of bits (ENOB) larger than 13 [1]–[3], except [6]. Choosing a small technology node and the inherently smaller t_d also helps making the amplifier stable for a fast [7] or accurate ADC.

Recent works use RAMPs as different types of residue amplifiers in pipelined ADCs. The residue amplifiers have a gain of a factor of two [1]–[3], [5], [7], 32 [4] or 64 [6]. A large gain results in a large feedback factor ψ, which enhances stability.

D. Signal Error Mechanisms

The error of the output voltage of the RAMP can have multiple causes. For a RAMP with medium accuracy, the dead zone is large and switches off the output stage when being settled. In this case, the size of the dead zone dominates the error, because the amplifier chain is settled somewhere in the dead zone, not at an accurate voltage. This deterministic behavior creates harmonics [10]. If the RAMP needs to be accurate, its dead zone needs to be much smaller and becomes a weak zone. The output stage is no more completely switched off at the settling spot and the output error is now influenced by an occurring linear settling [1], [11], circuit noise and the time needed to enter the stable state, where the linear settling happens. Traditional small-signal-based noise design can be done, because for most of the practical designs, noise has such small amplitudes that even the RAMP, which does not necessarily uses biasing techniques, can be modeled as a linear amplifier [1]. Distortion might still occur and can be reduced by increasing the small-signal gain bandwidth or by making the oscillation end earlier, which leaves more time for the linear settling. If the time constant formed by the last stage's output resistance and the load capacitor is much larger than the entire settling time, the linear effect is negligible.

The stability compensation using t_d reduces all error mechanisms: it allows a smaller dead zone, which can reduce distortion, or makes the oscillation end earlier, which leaves more time for the linear settling. It spends more power in the first inverter stage and therefore improves circuit noise and the gain bandwidth for the linear settling mechanism. Therefore, an over-compensated RAMP becomes not only stable, but also accurate.

III. PROPOSED RAMP CIRCUIT ARCHITECTURE

The architecture used to evaluate the design process proposed in Section IV uses the self-biased RAMP circuit from [4], but replaces the first stage. Instead of using a current-starved fully-differential inverter for the first stage, two pseudo-differential inverters are used. A battery biasing
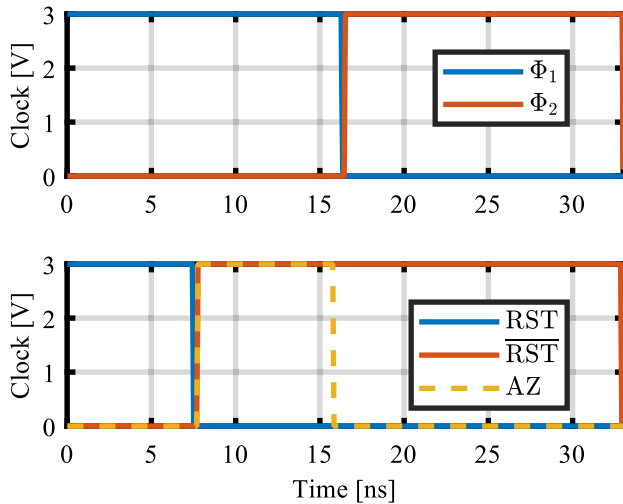
Fig. 5.   Timing diagram showing the clock signals used in the proposed circuit.

stores offset voltages on capacitors in the gate path of each transistor of A_1. During the first half of the sampling phase of the amplifier, A_1 is switched to biasing mode and the biasing capacitor voltages are set. Otherwise, the circuit operates as inverter with additional gate voltage offsets [12], [13]. Self biasing adds the dead-zone voltage offsets at the output of A_2 using a resistor instead of using two capacitors at the input of A_2. The input-referred dead zone is smaller, because it is divided by the gain of A_2 and A_1 [5].

Fig. 4 shows one branch of the proposed pseudo-differential RAMP and the biasing network. The corresponding operation phases are shown in Fig. 5. The self-biasing resistor R adds the dead-zone offset. When the amplifier is settled, $M_{p,2}$ and $M_{n,2}$ conduct a current, which creates a voltage drop across R. Both output transistors are switched off more due to its voltage drop. During the initial settling, only one transistor $M_{p,2}$ or $M_{n,2}$ conducts. The current through R is then switched off and so is the dead-zone offset. This provides larger overdrive voltages for the output transistors during the initial slewing. Additionally, the self-biased dead zone is more robust against process-voltage-temperature (PVT) variations, allows AZ over all inverter stages and reduces the number of switches [5].

$M_{t,p}$ and $M_{t,n}$ are the inverter devices of the first stage. The two biasing branches generate the voltages $V_{b,p}$ and $V_{b,n}$. All PMOS and NMOS devices of the first stage and the biasing network are sized the same, but $M_{t,p}$ and $M_{t,n}$ have a multiplier of two. The circuit operates in three different phases:

First, the switches controlled by RST and Φ_1 are closed. In this phase, a translinear loop [14] is formed. The first stage's biasing current is defined as $2 \cdot I_{bias}$ using the battery devices $M_{b,p}$ and $M_{b,n}$, which bias $M_{t,p}$ and $M_{t,n}$ as class AB stage [13]. The two inverter devices act as diodes and their gate potentials are stored on the offset-storage capacitors C_p. The inverter operates at its trip-point, where both devices conduct the same current with a mid-supply output voltage. This current at the amplifiers settling point is now defined by the C_b capacitors, which act as constant-voltage sources. The two RST switches connecting the battery and the two $\overline{RST}$ switches connecting the first stage's output need to have the

same multiplier as the inverter devices. The two RST switches connecting the battery also need to be placed in the biasing branches for maintaining the translinear loops.

Second, the AZ phase starts. All switches except the switches labeled with RST close. The devices $M_{t,p}$ and $M_{t,n}$ operate as inverter with the two offset-storage capacitors C_b in their gate paths. The AZ is done across all three inverter stages. The battery biasing operates similar to an AZ, but the circuit architecture of the first stage is switched when leaving the battery-biasing phase, which is not the case at the end of the AZ phase. The AZ is therefore used for suppressing low-frequency noise and for defining an accurate input-referred trip point.

Third, during the following amplification phase Φ_2, only the switches $\overline{RST}$ are closed. The initial slewing starts here. The first stage operates as an inverter stage with three offset storage capacitors: the two C_b set the trip point of the first inverter to V_{cm} and define the current in the first inverter's devices at that point. The offset on V_{az} is around 10 mV and suppresses low-frequency noise and sets the trip point of the entire inverter chain to V_{cm}. When being settled, the inverting and non-inverting inputs are at their trip points, which makes it important to set the exact input-referred trip point using the AZ.

During reset and AZ, the two pseudo differential branches operate exactly the same and the switches cannot cause differential-mode disturbances. Still, the added parasitic capacitors and the on-resistance of the switches are a concern especially for the wide switches bridging or connecting $M_{b,p}$ and $M_{b,n}$. This is a major drawback of this architecture.

The circuit proposed in [4] is fully differential and makes use of less switches, but needs tail current sources. Its first stage inverter operates as current-starved inverter. The proposed circuit in Fig. 4 has a larger transistor-width-limited slew current of the first inverter stage during the initial settling, which reduces the delay time t_d. Still, the current at the final settling point through both inverter devices is defined as in [4], which makes it possible to design the first stage for small thermal noise.

IV. OPTIMIZATION PROCESS

The general idea of the optimization process is independent of the exact application of the RAMP and its circuit topology. It is derived from Section II-B, which is derived from a simple RAMP model. Though, a part of the optimization is simulated with the RAMP in its application. The optimization process analyzes each design iteration purely in time domain and does not require any manual tuning, which otherwise might be necessary [15]. This section introduces this optimization process and explains how it is applied to an exemplary design using device-level models. The intention of this design is the implementation of a RAMP based first stage integrator in an I-$\Delta\Sigma$ modulator. Here, the integrator gain is only 0.24, which poses tough requirements on the RAMP due to the correspondingly worse feedback factor and thus stability, which is in contrast to most recent RAMP applications in pipelined ADC's residue amplifiers [1]–[7]. The first stage
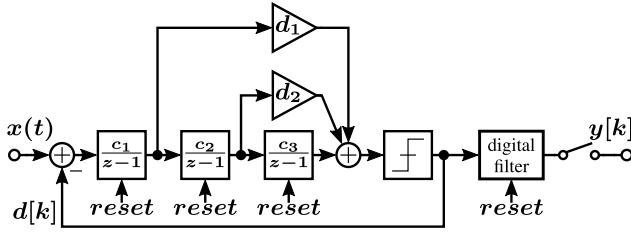
Fig. 6.   The I-$\Delta\Sigma$ modulator as RAMP test scenario [16].

integrator produces samples with a sample rate of 30 MHz. Due to the filtering and downsampling process of the I-$\Delta\Sigma$ modulator, the output sample rate of the ADC is smaller by a factor of 150 resulting in 200 kHz.

Fig. 6 shows the I-$\Delta\Sigma$ modulator in which the RAMP is used inside the first integrator. All other circuit elements are ideal components, such that the performance of the RAMP can be evaluated. In this chain of integrators with feed-forward-summation architecture, the first integrator is crucial for the overall performance [16].

The aim is to achieve an ENOB of 15 bit. For the later optimization steps, this requirement needs to be translated into an RAMP-output-referred rms (root-mean-square) error. The ENOB requirement relates to an input-referred SNDR require-ment. With an assumed single-ended input-signal amplitude $A = 1\,V$, an input-referred rms error voltage can be calculated. The first integrator's gain is $c_1 = 0.24$ [16]. The I-$\Delta\Sigma$ modulator records $M = 150$ samples before averaging them in the digital filter to one Nyquist sample and resetting all integrators and filters. When assuming a DC signal for $x(t)$, the integrator will sum up M samples, which equals a gain of $M \cdot c_1$ [16]. This gain is used for referring the input-referred error to the output, which results in

$$\varepsilon_{rms} = 2\frac{A}{\sqrt{2}} \cdot 10^{-(ENOB \cdot 6.02 + 1.76)/20} \cdot c_1 \cdot M = 1.27\,mV.$$

(4)

The output-referred error is typically larger than the input-referred error. Thus it is numerically easier to handle in the following optimization process.

The optimization is split into two steps: First, the transistor sizes of the three amplifying unit inverters are found. This step is not yet directly related to any design constraint. Second, the multipliers for these unit inverters and three other parame-ters are found using the design constraints in an optimization of the overall RAMP within its application circuit, i.e. the I-$\Delta\Sigma$ modulator. Multiple iterations of the process might be necessary, if the result of the top-level optimization indicates bad parameters used in the unit-inverter optimization. The process is shown in Fig. 7 and explained in detail in the following sections.

A. First Optimization Step: Sizing of the Unit Inverters

This first optimization is used to determine the transistor sizes of the three inverters, without constraining their mul-tipliers. This step needs to be done once when using a new technology and is independent of the application of the RAMP.
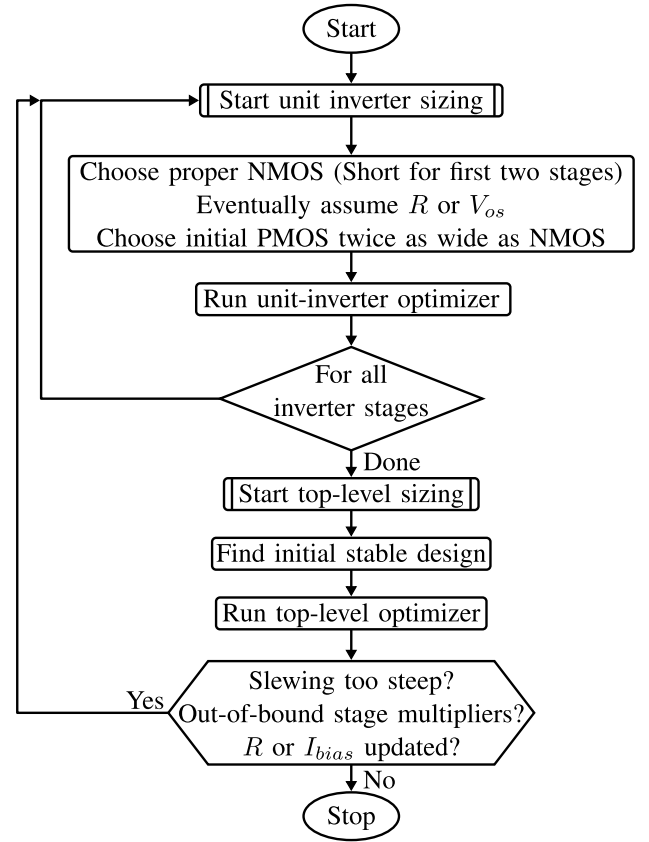
Fig. 7.   Flowchart describing the proposed optimization process.
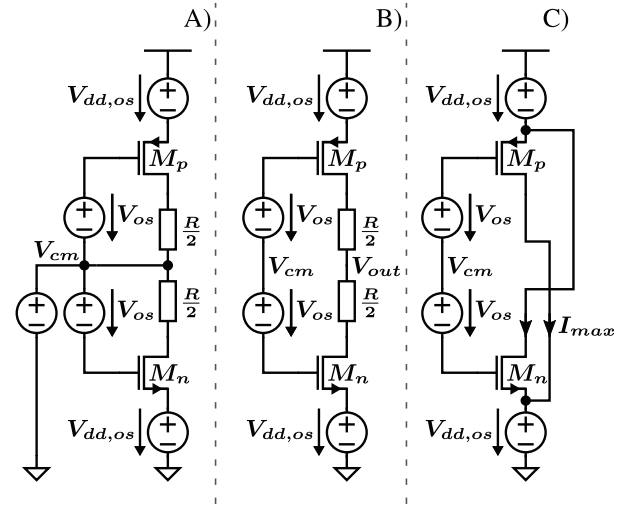
Fig. 8.   Testbench for designing the unit inverter cells.

Fig. 8 shows the testbench, which is simulated for evaluating the cost function, which rates the quality of the design during one optimization step. It simulates three instances A), B) and C) of one inverter at its trip point and extracts operating points and small signal parameters. The size of the transistors M_p and M_n are the parameters found by the optimization. $V_{dd,os}$ can be used when designing the inner inverter transistors of a current-starved inverter. The resistors $\frac{R}{2}$ are used for taking the self biasing resistor of the second stage into account, otherwise

it is set to zero. It can also be used for finding the size of the tail-current-source transistors, if a current-starved inverter is used. The V_{os} sources are used for applying the gate offset, which is otherwise applied by the offset-storage capacitors C_b or the dead zone. This offset voltage and the self-biasing resistor are not known when starting the first optimization step, thus their values need to be assumed. The actual values are found by the second optimization step. The two optimization steps are then iteratively executed. For the two designs in this paper, it was sufficient to iterate this optimization twice. In Fig. 8, the testbench in

A) is used to simulate the circuit at the desired trip-point. In- and output are forced to the trip-point and it can be analyzed, if the small-signal parameters of the two transistors are as required at the desired trip point.

B) is used to force the inverter's input to the desired trip point and to measure if the output reaches V_{cm}. Thus, it is analyzed if the true trip point is the desired one.

C) is used to analyze the maximum slew current of the two devices and to investigate the effects caused by the slew-current asymmetry as described in Section II-A and Section V-D. V_{os} might be chosen larger in this branch, if the initial slewing is investigated.

Having these testbenches for the optimization, next the cost function and parameters for the optimization have to be defined.

1) Cost Functions: The testbenches A)-C) optimize for two goals and the compliance of each goal is rated with a cost function. First, the output-referred trip point obtained in B) should be close to V_{cm}. Second, the relative drain-source resistance r_{ds} mismatch between the two devices should be small. These two constraints should be chosen tight, but relaxed enough to allow the optimizer to converge. With the same overdrive voltage V_{ov} and the same drain-source current I_{bias}, both devices have the same transconductance g_m, when considering the square law model. Thus, if both devices also have the same r_{ds}, both contribute the same to the overall small signal gain, which then is one intrinsic transistor gain [13]. This makes the design efficient for small-signal gain.

An alternative for the r_{ds} cost function in smaller technology nodes is the design for symmetric slewing in C). For large technology nodes, the devices sized for a matching intrinsic gain show also symmetric slew behavior. Devices in smaller technology nodes do not follow the square law model and their currents will mismatch. Thus, the r_{ds} cost function can be replaced by a cost function weighting the similarity of the two slew currents.

2) Parameters: The described optimization finds the PMOS width and an additional PMOS area multiplier. An NMOS size needs to be given before running the optimization: its length should be chosen close to minimum length in the first two stages in order to reduce their delay time. The transistors of the output stage should be chosen long enough to feature a very large AC output resistance in the settled state. The width of the NMOS can be chosen narrow, but not minimum size. The devices should not suffer narrow-channel effects, while very large widths are not necessary, because the width

of the inverter can be scaled up by the second optimization step. If the initial slewing after the top-level optimization is very fast or if the found stage multipliers are extremely large or small, another optimization iteration with adjusted NMOS sizes needs to be done.

All PMOS devices get the same length L as the NMOS devices and a larger width W for a starting point of the optimization. The width is optimized together with a multiplier m_{area}, which is applied to length and width. The width controls the trip point as observed in B). This might change the current flowing through NMOS and PMOS, but according to

$$r_{ds} \propto \frac{L}{I_{bias}}, \qquad (5)$$

this affects r_{ds} in both transistors equally and does not influence their matching [13].

The area multiplier changes width and length by the same factor. According to the square-law model, this does not change the current in the transistor and thus the trip point is not affected. According to Section IV-A.2, the area multiplier proportionally affects the output resistance of the PMOS. It is used for matching the output resistance of the PMOS to the one of the NMOS. Due to process variation, the design cannot rely on this matching during operation. It only provides a raw estimation of e.g. the trip point, but the exact trip-point needs to be set using AZ to protect it against PVT variations.

Two optimized parameters are used, where each parameter influences only one of the two cost functions (when assuming ideal square-law behavior): the one for the output-referred trip point and the one for the NMOS vs. PMOS r_{ds} mismatch. The optimizer algorithm is able to find gradients, which do not change their direction during the optimization iterations, which makes this optimization converge fast. In the exemplary performed optimizations, the 40 nm design needs more iterations to converge, because the assumption of the square-law behavior does not match as good as for the 180 nm node.

B. Second Step: Top-Level Optimization

After finding the unit-inverter parameters, they are transferred to a schematic of the entire RAMP and simulated within the exemplary application of an I-$\Delta\Sigma$ modulator. The RAMP is evaluated using transient simulations, since it operates with large signal fluctuations during most of the settling process. For evaluating the cost function, the entire I-$\Delta\Sigma$ modulator including the RAMP-based first-stage integrator is simulated for $2 \cdot M$ samples with one reset of all integrators in between and a ramp shaped input signal. This simulation yields a similar RAMP performance compared to a full spectral simulation and is run for evaluating the cost functions. The integrator is compared to a reference circuit using ideal amplifiers, which is run in parallel to the RAMP based I-$\Delta\Sigma$ under test.

The transistor sizes in the inverters, the biasing network and the battery are already fixed as described in Section III and have been found using the first optimization step. The next section introduces the parameters that are optimized during the top-level optimization.

1) Parameters: A parameter m_{a1} is the multiplier for the first inverter stage. It is applied to C_{az}, C_b, the *RST* switches connecting the battery, the $\overline{RST}$ switches connecting the output of the first stage and all transistors except the devices in the biasing network. This parameter is used for performing the delay-time compensation. If m_{a1} is increased, transistors in the first inverter become wider and its delay time is reduced. The drawback of this operation is that one pseudo-differential half-side of the first stage inverter drives the current $2 \cdot m_{a1} \cdot I_{bias}$ near to the final settling point. Increasing m_{a1} enhances stability, but increases the power consumption. It is optimized like all multipliers on a linearly distributed value space. The current density is not influenced by m_{a1}, such that $\frac{g_m}{I_d}$ is fixed. g_m scales proportionally with m_{a1} and influences the thermal noise performance of the first stage. Thus, a large multiplier m_{a1} also results in better noise performance and more power consumption.

For m_{a2}, the same rules apply as for m_{a1} but it refers to the second inverter stage. It can again be used for decreasing t_d at the cost of increasing the power consumption, but the contribution of the second stage to t_d is smaller than the contribution of the first stage (see Section II-B) even when using self biasing [4]. Excessively increasing m_{a2} would load the first stage more and would increase its delay time.

The third inverter stage rather acts like a current source and therefore other scaling rules apply. The multiplier m_{a3} for the output stage scales the slew-current. A faster slewing causes larger overshoots and results in more oscillation periods, which compromises the advantage of the shorter initial slewing [1]. If the multiplier m_{a3} is too small, the time needed for the initial slewing might exceed the settling time. If the slew rate becomes excessively large, the RAMP becomes unstable according to Section II-B. The optimum for m_{a3} is not defined by a t_d vs. power relation like for m_{a1} and m_{a2}: the parameter does not have a noticeable influence on the power consumption. The output stage is off after entering the dead zone and does not draw a noticeable current. Before that state, it delivered the current that is needed to recharge the capacitive feedback network, which is not influenced by m_{a3}.

The biasing current for the first stage I_{bias} is the second option to realize t_d compensation. The current is referred to a single unit inverter (not to all m_{a1} unit inverters in the first stage). Each one of the m_{a1} parallel inverters conduct I_{bias} in the settled state. A large I_{bias} results in smaller gate offsets, which reduces t_d and thus enhances stability. A small I_{bias} reduces the overdrive voltage of the two input devices and reduces the effective gain A_1 as shown in Fig. 3, which also increases stability. There is an optimum spot between the two criteria. An increased I_{bias} increases the power consumption and g_m of the first stage. In contrast to m_{a1}, I_{bias} sets the current density in A_1. $\frac{g_m}{I_d}$ can be reduced, because the stage operates closer to strong inversion with a larger current density. I_{bias} and the following two parameters are optimized on a logarithmic value space. Small values are thus distributed more densely.

The self biasing resistor R is also referred to a single unit inverter element of the second stage. The real resistor in the RAMP is one R instance with the multiplier m_{a2}.

When increasing the multiplier, the inverter current at the final settling point increases, but flows through more parallel resistors. The dead-zone voltage across the complete self biasing resistor R/m_{a2} remains constant. The size of the dead-zone can only be influenced by the value of R. A large value creates a large dead zone and boosts stability, but also increases the input referred error by increasing the uncertainty of the final settling point. The power consumption is not influenced directly. An inaccurate amplifier is corrected with a smaller R, which decreases the stability requiring again t_d compensation, which costs power.

The last optimized parameter is C_{az}. In the exemplary used design, C_b equals $C_{az}/2$, such that the battery biasing and the AZ capacitors share the same voltage drop in a capacitive divider. The capacitors need to be large enough to reduce the impacts of noise folding, drifts and charge injection. Small capacitors divide the input signal of the first inverter stage and increase the input-referred noise, but also decrease the effective gain of the first stage A_1, which enhances stability. The capacitors C_{az} and C_b are scaled with the first stage's multiplier m_{a1}, such that m_{a1} does not influence the capacitive divider.

Moving one of the six parameters into its extreme maximum or minimum value influences the overall RAMP performance (power, accuracy, stability) in opposed directions. Thus the necessity of an optimizer.

2) Initial Parameter Values: The starting point of the optimization and thus the starting values for all parameters, should result in a stable RAMP. The power consumption can though be large since it is minimized by the optimizer. In the unstable region, the output might oscillate and the optimizer might e.g. phase shift this oscillation to a minimum error instead of making the RAMP itself stable. This would make the optimization converge slowly or not at all.

If these rules are followed, the optimization process yields the same parameters for different initial parameters.

Parameters can be added, as long as they influence at least one of the cost functions and as long as there is an optimum value in the mid of the parameter space.

3) Cost Functions: The performance of the overall RAMP in the top-level simulation is evaluated using two cost functions: the power consumption and the output referred error are extracted from the transient I-$\Delta\Sigma$ simulation including the RAMP. The ideal reference circuit is thereby used for measuring the output referred rms error of the first integrator at the end of each amplification phase. A relation between ENOB and the rms error is given in (4). This makes it possible to estimate the performance of the I-$\Delta\Sigma$ modulator by means of a very short simulation. The offset error is ignored for calculating the rms error, because it influences only the DC bin in the spectrum. A gain error can also be referred to the input and does not influence the spectral performance. The gain error is thus ignored, too. Both error types are important when designing residue amplifier stages, but not important in I-$\Delta\Sigma$ modulators, and moreover the switching scheme from [5] could be used for compensation.

The output-referred errors explained in Section II-D are reduced by the optimizer with the two mentioned actions:

adjusting the dead zone or reducing t_d. The best action is found by the optimizer, which looks for the steepest gradient of the error measure on the parameter space.

Both cost functions, namely mean error and power consumption, are normalized to a target value to transfer them into the same order of magnitude. For the rms error, the normalization is the value depicted in (4). For the power consumption, the normalization is 1.2 mW as this is the state-of-the-art performance using a non-RAMP I-$\Delta\Sigma$ modulator [16].

The two cost functions are minimized during the optimization process. Reducing either function below 1 theoretically indicates that the constrained target value of power or accuracy are fulfilled. Though, the accuracy measure is obtained using the described short-time transient simulation (c.f. Section IV-B). This forecast of the SNDR assumes e.g. a uniform distribution of all input voltage values by simulating with a ramp shaped signal. This assumption is actually not correct when driving the I-$\Delta\Sigma$ modulator with a sinusoidal signal and the actual feedback signal pattern to obtain an output spectrum and the real SNDR. This made it necessary in the considered example to aim for an over-designed accuracy measure when using the simplified simulation setup in order to forecast the target SNDR after full transient simulation at the end. Therefore, the optimizer goal for the accuracy was readjusted x4 smaller, i.e. 0.25 for the cost function, which turned out to obtain the 15-bit SNDR performance.

Obviously, while compromising accuracy vs. power consumption, the optimizer will try to get the accuracy cost function to the target value and the power cost function to a value as small as possible.

V. RAMP OPTIMIZATION EXAMPLE

This section presents three different designs carried out in 180 nm and in 40 nm for two supply voltages. The RAMP from Section III is used in the first integrator of the I-$\Delta\Sigma$ modulator shown in Fig. 6. All other parts of the I-$\Delta\Sigma$ modulator as well as all switches are ideal. The three designs are not preliminary meant to be comparable with implemented and manufactured RAMP-based circuits or other ADCs, because the influence of e.g. the second and third integrators on the system's performance is neglected for the sake of simplicity and for being able to focus on the performance of the RAMP. Furthermore, the reference design [16] showed, that the first stage integrator dominates the overall SNDR and that this integrator is responsible for 77 % of the overall system's power consumption, whereas the other two stages account for only 12 %. The reason for that is, that all input-referred DC and low-frequency errors of the second integrator in Fig. 6 are divided by $M \cdot c_1$ when referring them to the input of the ADC, which relaxes the constraints of the second integrator. Similar rules apply for the third integrator and the inherently linear 1-bit quantizer. But the input of the first integrator is the input of the entire ADC and all input-referred errors of the first integrator are directly visible at the ADC's output. Therefore, the first integrator is the building block, which dominates the performance of the entire ADC. The designs are synthesized and simulated with ideal switch models, in order to solely

TABLE I
FOUND UNIT INVERTER PARAMETERS

	180 nm @ 3 V				
	NMOS		**PMOS**		
	Length	**Width**	**Length**	**Width**	**Area Mult**
A_1	350 nm	1 um	350 nm	2.92 um	2.9
A_2	350 nm	1 um	350 nm	3.48 um	1.42
A_3	1 um	1 um	1 um	2.07 um	1.63

	40 nm @ 2.5 V				
	NMOS		**PMOS**		
	Length	**Width**	**Length**	**Width**	**Area Mult**
A_1	300 nm	1 um	300 nm	9.97 um	1.15
A_2	300 nm	350 nm	300 nm	1.28 um	1.35
A_3	1 um	350 nm	1 um	5.06 um	0.85

	40 nm @ 1.1 V				
	NMOS		**PMOS**		
	Length	**Width**	**Length**	**Width**	**Area Mult**
A_1	40 nm	1 um	40 nm	3.49 um	1.17
A_2	40 nm	1 um	40 nm	2.69 um	1.11
A_3	500 nm	200 nm	500 nm	845 nm	2

depend on the performance of the active transistors and not on the technology's switch performance. Resimulation with transmission gates in the 40 nm designs reveals similar performance, whereas the same switch design in the 180 nm node yields a severe performance drop. The idea of this example is to show how the proposed optimization-based design process, which was derived from a simple RAMP model, can be applied to a more complex circuit architecture; also, the influence of design choices onto the performance is thereby shown and it can be shown how the optimization process is embedded in the system design and that it is applicable to actual design cases. The three designs are carried out in the *Cadence Virtuoso Analog Design Environment* using the *Spectre* simulator and the optimization framework provided by *ADE XL*.

The used RAMP schematics are identical in all three cases, only the transistor instances are obviously from the respective technology design kit. As the I-$\Delta\Sigma$ architecture, scaling, input amplitude (dBFS) and input-referred noise (SNDR) are identical, also the feedback networks are the same in the three cases and only the RAMP transistor level design would differ. The first two supply voltages of 3 V and 2.5 V for the 180 nm and 40 nm designs (using I/O devices for the higher possible supply) are similar, but a third design makes use of the 1.1 V core devices of the 40 nm technology, where the minimum length of the devices is significantly smaller and the intrinsic t_d compensation can be exploited.

A. RAMP Parameter Optimization

This section describes the parameters, which are found during the optimization process.

1) Sizing of the Unit Inverters: Table I shows the transistor dimensions found by the unit-inverter optimization outlined in Section IV-A. The NMOS dimensions and the PMOS length have even values, because their values are chosen manually before starting the first optimization step (see Section IV-A.2).

	180 nm @ 3 V	40 nm @ 2.5 V	40 nm @ 1.1 V
	Optimization Parameters		
m_{a1}	19	9	15
m_{a2}	6	2	8
m_{a3}	1	1	4
I_q	51.3 μA	2.4 μA	1.7 μA
R	14.5 kΩ	24.4 kΩ	30.5 kΩ
C_{az}	40 fF	128 fF	76 fF
	Cost Functions		
Error	0.243	0.184	0.306
Power	12.99	0.94	0.334

The PMOS is obviously wider due to the mobility difference of positive charge carriers [13]. For the 40 nm designs, the factor by which the PMOS is wider than the NMOS largely varies between the different stages. For a simple square-law transistor, the ratio is rather constant, as seen for the 180nm design. The PMOS area needs to be increased to increase its r_{ds}, because it is wider than the NMOS. The output PMOS of the 40 nm at 2.5 V design has to be downsized by the optimizer. This is a second hint for the output stage in 40 nm not behaving according to the square-law model. Still, the optimizer could find a sizing, which fulfills both cost functions for the unit inverter optimization. The 40 nm design at 1.1 V uses core devices and can therefore utilize the minimum length of the technology. The output stage of this design is not matched for a mid-supply output voltage, but for symmetric slew currents as described in Section IV-A.1.

2) Top-Level Optimization: Section IV-B.1 described the tradeoffs of all top-level-optimization parameters on power consumption, accuracy and stability. Table II shows the found top-level parameters and the value of the cost functions for the parameters after transient simulation based optimization. The top-level error cost function reaches its target specification, when it approaches or falls below 0.25, the power cost function aims for a value of 1 (see Section IV-B.3). For the 40 nm designs, the input amplitude is scaled with the smaller supply voltage and so the targeted absolute rms error and the normalization value for the error cost function also scale down. The optimization for the 180 nm design could not achieve to satisfy both cost functions. The accuracy constraint is weighted larger, such that the 180 nm design fulfills the accuracy constraint, but fails the power consumption constraint. More power is needed for the t_d compensation to achieve stable operation. This strong influence of technology scaling on RAMP performance is also investigated in [1]. This effect is also visible when comparing the two supply voltages used for the 40 nm designs: The power consumption of the 1.1 V devices is even smaller than for the 2.5 V devices, because the core devices can exploit the minimum length and the intrinsic t_d compensation. An exact error-cost-function value of 0.25 cannot be reached, because for that the optimizer would need to do very fine-grained steps with e.g. 1 fF.

The multipliers m_{a1} and m_{a2} are used for t_d compensation. The 40 nm at 2.5 V devices have intrinsically smaller delay times in the inverters and therefore use smaller multipliers than

the 180 nm devices. The current I_{bias} is used in the same way and here the effect is much more visible. In 40 nm at 2.5 V, the optimizer uses a large multiplier and a small current for the stability compensation, which results in a smaller current density in the first stage, weak inversion and superior noise performance. On the contrary, the 180 nm design needs a large m_{a1} and a large I_{bias} to reduce t_d. The larger I_{bias} is the main reason why the 180 nm design consumes much more power as noted by the power measure. In both designs, m_{a1} is larger than m_{a2}, because the second stage has a smaller contribution to the overall delay time t_d and therefore needs less compensation. When comparing the two 40 nm designs, the lower supply design needs large multipliers when taking the unit inverters with widths in the same region and much smaller lengths into account. The widths and multipliers always need to be analyzed in the context of the smaller supply voltage, which causes smaller currents. Beside that, short-channel effects are much more likely to occur for the shorter devices and can reduce the conducted current.

The multiplier m_{a3} is one for the first two designs, which indicates that a larger length could have been initially chosen for the unit inverters of the output stages. The optimizer does not choose a larger multiplier, because that increases the slew rate more than necessary and makes the amplifier unstable. For the 40 nm at 1.1 V design, the multiplier of the output stage seems to be relatively large, even when taking the smaller supply and overdrive voltages into account. This is discussed in Section V-C and Section V-D.

The size of R influences the size of the dead zone, which is addressed more concerning stability in Section V-C. The optimized C_{az} in Table II is smaller for the 180 nm design, such that the capacitive divider acts stronger. The first stage uses a larger multiplier and a much larger biasing current in 180 nm. C_{az} is used to suppress the resulting larger DC gain as described in Section IV-B.1. This increases the input-referred noise, which would otherwise be smaller due to the larger power consumption in A_1. In 40 nm at 1.1 V the capacitors are smaller, because the shorter devices also provide smaller gate capacities.

B. Small Signal Performance After Optimization

An AC analysis cannot be applied for designing the RAMP, because the three inverter stages leave their operating points during most of the settling process. Still, the amplifier must be stable at its settling point, otherwise it might oscillate around the final settling point. This is investigated in this section. The small-signal analysis also shows parameters, which are important for analyzing the circuit noise, which is caused by the RAMP in the dead zone.

Table III shows the AC model parameters for the three inverter stages in the three designs at their settling points, where each inverter's input is at its trip point. The $\frac{g_m}{I_d}$ efficiency is related to the inversion region of the inverters. The first stage in 180 nm has a large current density with a large I_{bias} and therefore operates in strong inversion. The 40 nm designs operate with a much smaller I_{bias} and therefore in weak inversion. The t_d compensation with m_{a1} or I_{bias} creates

TABLE III
COMPARISON OF SMALL-SIGNAL PARAMETERS

	$\frac{g_m}{I_d}$ $[\frac{1}{V}]$	Max gain [dB]	Pole frequency [kHz]
180 nm @ 3 V			
A_1	3.68	17.7	181 MHz
A_2	1.86	17.7	984 MHz
A_3	18.03	61.2	9.1 kHz
40 nm @ 2.5 V			
A_1	15.1	16.7	198 MHz
A_2	1.5	9.5	1.433 GHz
A_3	13.3	55	94.9 kHz
40 nm @ 1.1 V			
A_1	20	15.8	370 MHz
A_2	12.6	16.3	1.869 GHz
A_3	23.9	31.7	1.513 kHz

a large maximum gain in all three designs with a medium speed. C_{az} attenuates the maximum gain, which would be much larger otherwise. The second stage has gate-source voltages of $|V_{cm}|$ and lower drain-source voltages due to the voltage over R. The stage operates in triode region, where it has a smaller $\frac{g_m}{I_d}$ efficiency. The small output resistance results in a small DC gain and a large bandwidth of this stage [4]. The output stage operates close to/in weak inversion in all three designs. The dead-zone related gate offsets of the output stage push its gate-source voltages close to the threshold voltage. The drain-source voltage is V_{cm}. The dead zone in the 180 nm design is larger than in the 40 nm at 2.5 V design (see Section V-C), which results in the stage operating even deeper in weak inversion. The 40 nm at 1.1 V design uses a wide output stage, which drives this stage the most into weak inversion. The large output resistance results in the output stages having a large DC gain and the most dominant pole for all three designs. The output stage of the 40 nm at 1.1 V design is optimized for symmetric slew currents, not for a large small-signal gain, which can also be seen in Table III. The phase margin for the 180 nm and 40 nm (both supplies) designs are 66°, 79° and 63°, respectively.

C. Stability Plot After Optimization

The stability plot, as introduced in [1], is obtained by simulating the RAMP in an open loop transient simulation. For multiple differential input voltages, the output current of the RAMP with outputs shorted to V_{cm} is plotted. This shows the maximum slew currents, output current symmetry between PMOS and NMOS and the size of the input referred dead zone.

Fig. 9 shows the stability plot for the three designed RAMPs. The output devices are optimized for a matching r_{ds}, except for 40 nm at 1.1 V, which is designed for symmetric slewing. This third design shows symmetry around the x axis in the stability plot. The width of the formed valley in Fig. 9 determines the input-referred dead zone of the RAMP [1]. The 180 nm design shows the largest dead zone, because it is needed for stabilization there due to the slower inverters.
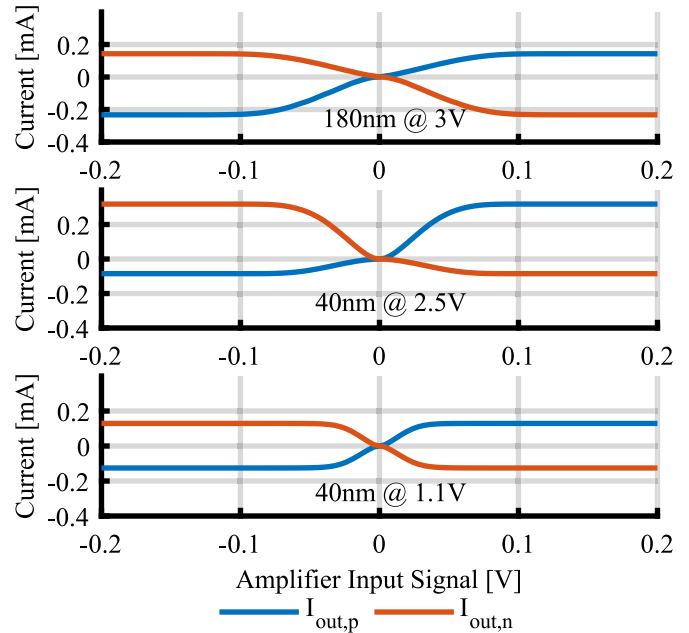

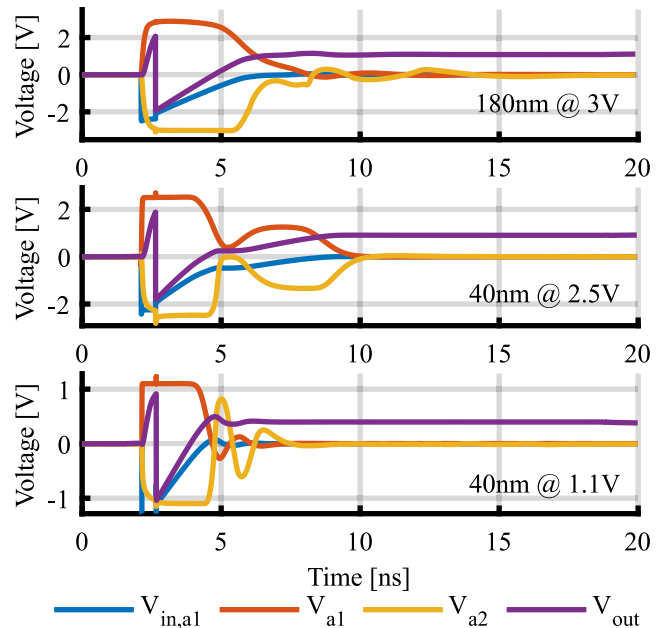Fig. 9.   Comparison of the stability plots [1].

Fig. 10.   Differential mode settling comparison. All three plot axes are scaled with their respective supply voltage.

The 40 nm at 1.1 V design uses the smallest dead zone, because it utilizes the largest intrinsic t_d compensation with the shortest transistors. All three designs have maximum slew currents in the same region, which results in the 40 nm at 1.1 V design having the fastest initial slewing, because the amount of charge to be moved on the feedback capacitors scales with the supply voltage. This matches the relatively large m_{a3} for the 40 nm at 1.1 V design in Table II.

D. Settling Behavior After Optimization

Fig. 10 shows the differential output voltage of all three inverter stages for the three RAMP designs. One integration
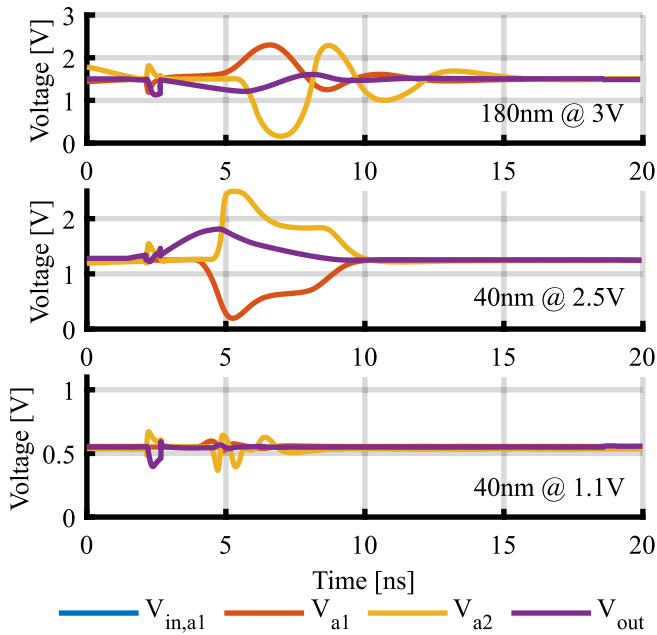
Fig. 11. Common mode settling comparison. All three plot axes are scaled with their respective supply voltage.

	180 nm @ 3 V	40 nm @ 2.5 V	40 nm @ 1.1 V
Power [mW]	15.87	1.683	0.46
SNR [dB]	91.84	93.98	94.39
THD [dB]	-100.58	-106.27	-104.81
SFDR [dB]	102.9	107.38	107.83
SNDR [dB]	91.30	93.73	94.01
ENOB	14.87	15.28	15.32
F1 [dBFS]	-3.52	-3.52	-3.52

branches end the initial ramping at the same time and the following oscillations do not alter the common modes of the three stages. The common mode fluctuations in the other two plots in Fig. 11 can be related to discontinuities in the differential-mode settling in Fig. 10: the oscillation of the 40 nm at 1.1 V design in Fig. 10 looks like an attenuated sinusoidal signal, whereas the other two oscillations show kinks, because the common mode fluctuations interrupt the oscillations.

The settled input-referred common mode of the RAMP is defined by the AZ and the settled output-referred common mode is defined relative to that by the voltage on the feedback capacitors. If the signal source feeding the sampling capacitors has a common-mode mismatch, this mismatch can be integrated on the feedback capacitors and becomes visible as a common-mode drift of the output of the integrator. Therefore, a standard common-mode feedback network feeding back the output-referred common mode via two 32 fF capacitors to the center node of the sampling capacitors during the integration phase is used. The capacitors are reset using one switch at their center node during the AZ phase. This circuitry reduces the common-mode gain of the integrator over M samples to 0.85, but does not influence the common-mode behavior during the settling of one sample, because the 32 fF common-mode feedback capacitors are much smaller than the 300 fF sampling capacitors.

phase within the I-ΔΣ is shown. All three y axes cover the respective supply range. The settlings look similar to the settling of the illustrative model as shown in Fig. 2. The initial settling can be seen in each design, but the following stabilization and oscillation phases do not show as many oscillation periods. The transistor-level amplifiers are more stable than the teaching example in Fig. 2. Even though the 40 nm at 2.5 V design uses a smaller input-referred dead zone than the 180 nm design, it looks more stable in this plot and its oscillations vanish earlier; this is because the smaller inverter delay can counteract the negative influence of a smaller dead zone on stability. The same can be said when comparing the 40 nm designs against each other. The shorter devices and their smaller intrinsic delay result in a shorter oscillation period, even though the initial slewing is steeper and can cause higher overshoots. This leaves more time to the linear settling.

Fig. 11 shows the common-mode settling of all three stages for each design. The same y axis scaling as for Fig. 10 was applied. The settling mechanism is the same as for differential signals including the stability criterion. An ideal circuit would show a straight line at V_{cm}. The time axis is the same as in Fig. 10, such that the common-mode fluctuations can directly be related to it. When being compared to the 180 nm design, the 40 nm at 2.5 V design has a larger output current asymmetry in Fig. 9, which results in larger output-referred common-mode fluctuations. The sign of the asymmetry also differs between both designs: the 180 nm design drives negative currents with a larger magnitude, the 40 nm at 2.5 V design drives positive currents stronger. Therefore, the sign of the output referred common-mode fluctuations is opposite between the two designs in Fig. 11. The output stage of the 40 nm at 1.1 V design with unit inverters matched for current symmetry counteracts the common-mode fluctuations and shows much smaller magnitudes in Fig. 11. The two pseudo-differential

E. RAMP Based ADC Performance After Optimization

This last simulation section shows the output spectra of the I-ΔΣ modulator simulated for 256 samples at Nyquist rate at an over-sampling ratio (OSR) of $M = 150$. M samples take 5 μs. The input frequency is 9.375 kHz, which makes the sampling coherent. The input data of the FFT is not windowed. The RAMP is driven with a clock frequency of 30 Mhz.

Table IV shows metrics obtained from the three spectra in Figs. 12a to 12c. In the three designs, the error caused by noise is larger than the one caused by distortion, but the difference is larger in the 40 nm designs. The smaller distortion matches the smaller oscillations as seen in Fig. 10. The smaller delay time of the 40 nm technology is able to compensate the destabilizing smaller dead zone (c.f. Fig. 9), which again is enhancing input referred noise, because the optimizer would compensate the amplifier for a stable settling and low noise with the same actions. Due to the smaller t_d, the 40 nm designs are noise limited and reduce noise with the first stage biased in weak inversion. The 180 nm design
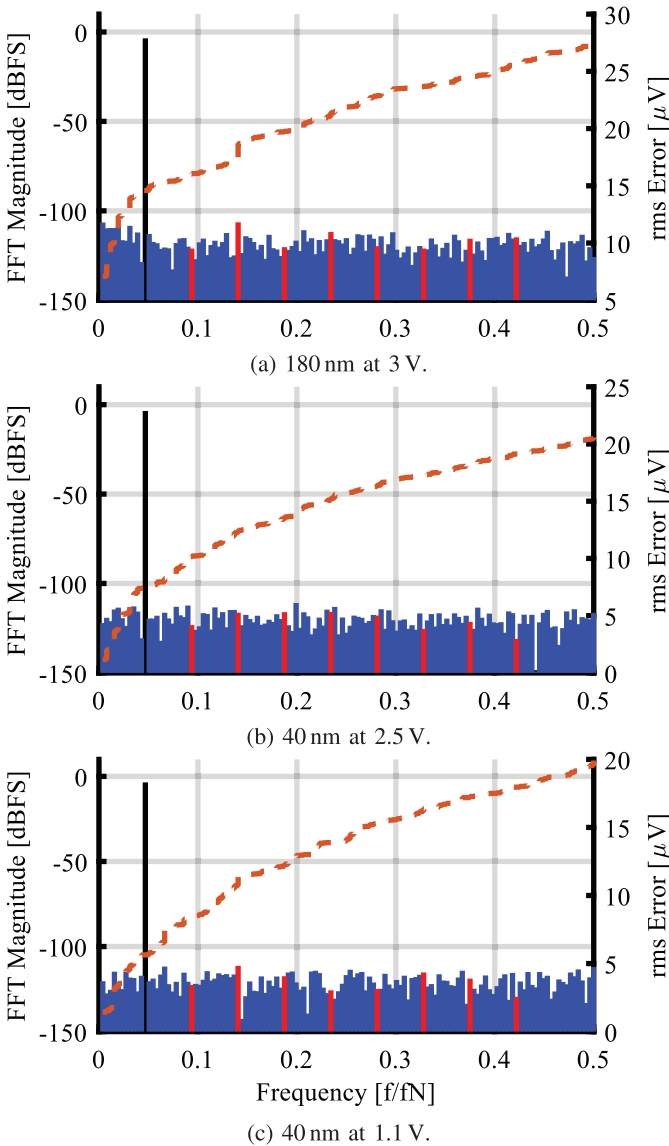
Fig. 12. Output spectra of the three designed I-$\Delta\Sigma$ modulators. The blue lines show noise bins, the black lines the F1 bins, the red lines the distortion bins and the orange dashed lines show the integrated rms noise.
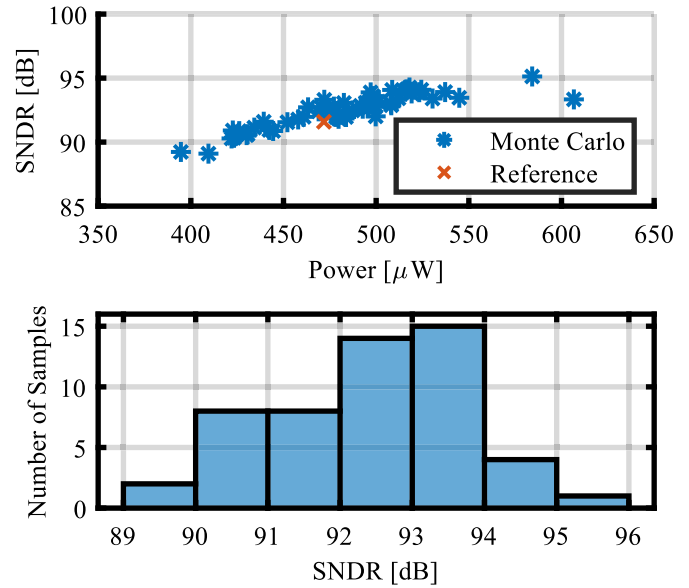

Fig. 13. Results of Monte Carlo simulation of entire ADC in 40 nm at 1.1 V with global and local mismatch. The scatter plot compares power consumption and SNDR of each design point. The histogram shows the distribution of the SNDR performance of all 52 Monte-Carlo samples.

strongly reduces the thermal noise, when achieving a stable settling, and its distortion raises more above the noise floor.

All designs achieve the target performance close to 15 bit. The 40 nm at 2.5 V design is about 2.43 dB more accurate than the 180 nm design, which exactly matches its error cost function being smaller by $\frac{0.243}{0.184} = 2.42$ dB as depicted in Table II. The 180 nm design consumes 9.6 times more power than the 40 nm at 2.5 V design, as already depicted by the power cost function (see Section V-A.2). This is much more as with the original amplifier in [16] and already confirms that RAMPs benefit highly from scaled technology, whereas might not outperform conventional amplifiers in larger technology nodes. Recent work makes use of Split-CLS [1], [2] or a large feedback factor [6] when utilizing RAMPs in large technology nodes.

The performance of the 40 nm at 1.1 V design is comparable to the design in the same technology using core devices at 2.5 V, but consumes even less power, because it needs less

t_d compensation. The error measure of 0.306 (slightly worse than for the 40 nm at 2.5 V design) as depicted in Table II stands in contrast to the ENOB of 15.32 (slightly better than for the 40 nm at 2.5 V design). The test case used for the top-level optimization and the full ADC simulation seem to match better for small supply voltages. One explanation can be the missing common-mode fluctuations, which make the settling performance much more reliable.

Fig. 13 shows the results of a Monte Carlo simulation of the 40 nm at 1.1 V design. The scatter plot shows that design points with a larger power consumption perform a better SNDR and vice versa. The mean of the power consumption and the SNDR is 483 μW and 92.46 dB. These values match to the reference point at 472 μW and 91.59 dB. The reference is slightly different from the results in Table IV, which is caused by differences in the transistor models used for the spectral and the Monte Carlo simulation. The transistor models need to be changed when switching to another simulation. Running the same simulation with the two different models and no PVT results in slight differences. The standard deviation of 41.8 μW and 1.32 dB for power consumption and SNDR are small, which can be explained by the long time available for the linear settling and the used AZ.

PVT variations are not encountered during the circuit optimization, because this would require many transient simulations to evaluate one iteration of the optimizer. All test cases for analyzing PVT effects would need to be covered, where the current optimizer only needs one transient simulation. Furthermore, PVT can be reduced efficiently by using a different circuit architecture (e.g. fully instead of pseudo differential) which is not tried by the circuit optimizer. Still, optimizing a design for PVT performance requires more compute power but no major changes in the presented optimization workflow.

VI. Conclusion

Recent state-of-the-art work on RAMPs does not provide an easy to copy and robust design process. This work has provided an optimizer-based design process. A given stability criterion has been reviewed and explained qualitatively using a simple RAMP model. The dynamic behavior of the effective gain and the delay time have been discussed as key parameters to achieve a stable, accurate and power efficient RAMP.

A new fast RAMP architecture, which overcomes large-supply problems and provides noise design techniques has been presented.

The stability criterion derived from a simple RAMP model has been used to develop a two-phase optimization, which can find circuit parameters automatically depending on accuracy and power requirements. The relations between the optimized parameters and the stability criterion have been shown. The process can be used with complex RAMP architectures, even though it is derived from a simple architecture.

The abstraction of the design process onto more complex circuits has been evaluated by synthesizing three RAMP based I-$\Delta\Sigma$ modulators in two different technology nodes. The influence of the technology node, AC mode parameters, different types of errors and the cause of common-mode fluctuations were investigated. The designs in the smaller technology node showed a much better performance. Using the minimum-length devices boosts the performance even more. The optimizer was confronted with a smaller intrinsic delay time and needed much less power to compensate the effects of this delay time on stability. The optimizer automatically made the expected compromises even though the constraints are tough and is therefore an appropriate tool to reliably design RAMPs with given power/accuracy constraints in a short time.

References

[1] B. Hershberg, S. Weaver, K. Sobue, S. Takeuchi, K. Hamashita, and U.-K. Moon, "Ring amplifiers for switched capacitor circuits," *IEEE J. Solid-State Circuits*, vol. 47, no. 12, pp. 2928–2942, Dec. 2012.

[2] B. Hershberg and U.-K. Moon, "A 75.9 dB-SNDR 2.96 mW 29 fJ/conv-step ringamp-only pipelined ADC," in *Proc. Symp. VLSI Circuits*, Jun. 2013, pp. C94–C95.

[3] T.-C. Hung and T.-H. Kuo, "A 75.3-dB SNDR 24-MS/s ring amplifier-based pipelined ADC using averaging correlated level shifting and reference swapping for reducing errors from finite opamp gain and capacitor mismatch," *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1425–1435, May 2019.

[4] Y. Lim and M. P. Flynn, "A 1 mW 71.5 dB SNDR 50 MS/s 13 bit fully differential ring amplifier based SAR-assisted pipeline ADC," *IEEE J. Solid-State Circuits*, vol. 50, no. 12, pp. 2901–2911, Dec. 2015.

[5] Y. Lim and M. P. Flynn, "A 100 MS/s, 10.5 bit, 2.46 mW comparator-less pipeline ADC using self-biased ring amplifiers," *IEEE J. Solid-State Circuits*, vol. 50, no. 10, pp. 2331–2341, Oct. 2015.

[6] A. ElShater *et al.*, "3.7 A 10 mW 16b 15MS/s two-step SAR ADC with 95dB DR using dual-deadzone ring-amplifier," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019, pp. 70–72.

[7] J. Lagos, B. P. Hershberg, E. Martens, P. Wambacq, and J. Craninckx, "A 1-GS/s, 12-b, single-channel pipelined ADC with dead-zone-degenerated ring amplifiers," *IEEE J. Solid-State Circuits*, vol. 54, no. 3, pp. 646–658, Mar. 2019.

[8] S. Leuenberger *et al.*, "An empirical study of the settling performance of ring amplifiers for pipelined ADCs," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.

[9] W. Wilson, T. Chen, and R. Selby, "A current-starved inverter-based differential amplifier design for ultra-low power applications," in *Proc. IEEE 4th Latin Amer. Symp. Circuits Syst. (LASCAS)*, Feb. 2013, pp. 1–4.

[10] T. Suguro and H. Ishikuro, "Low power DT delta-sigma modulator with ring amplifier SC-integrator," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 2006–2009.

[11] F. Wang and R. Harjani, *Design of Modulators for Oversampled Converters* (The Springer International Series in Engineering and Computer Science), no. 430. New York, NY, USA: Springer, 2012.

[12] Y. Chae, K. Souri, and K. A. A. Makinwa, "A 6.3 μW 20 bit incremental zoom-ADC with 6 ppm INL and 1 μV offset," *IEEE J. Solid-State Circuits*, vol. 48, no. 12, pp. 3019–3027, Dec. 2013.

[13] R. J. Baker, "CMOS," in Series on Microelectronic Systems. S. K. Tewksbury and J. E. Brewer, Eds., 3rd ed. Piscataway, NJ, USA: IEEE Press, 2010.

[14] C. Yuhua, W. Xiaobo, and Y. Xiaolang, "Translinear loop principle and identification of the translinear loops," in *Proc. IEEE Asia–Pacific Conf. Circuits Syst. (APCCAS)*, Dec. 2006, pp. 1667–1670.

[15] K. M. Megawer, F. A. Hussien, M. M. Aboudina, and A. N. Mohieldin, "A systematic design methodology for Class-AB-Style ring amplifiers," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 9, pp. 1169–1173, Sep. 2018.

[16] P. Vogelmann, J. Wagner, M. Haas, and M. Ortmanns, "A dynamic power reduction technique for incremental $\Delta\Sigma$ modulators," *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1455–1467, May 2019.

Joschua Conrad (Student Member, IEEE) received the B.Eng. degree in electrical engineering from Baden-Wuerttemberg Cooperative State University (DHBW) Stuttgart, Stuttgart, Germany, in 2016, and the M.Sc. degree in electrical engineering from the University of Ulm, Ulm, Germany, in 2019, where he is currently pursuing the Ph.D. degree with the Institute of Microelectronics, under the supervision of Prof. M. Ortmanns, with a focus on mixed-signal neural-network accelerator hardware.

Patrick Vogelmann (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Ulm, Ulm, Germany, in 2013 and 2015, respectively, where he is currently pursuing the Ph.D. degree with the Institute of Microelectronics, under the supervision of Prof. M. Ortmanns, with a focus on the development of high-resolution and incremental sigma–delta analog-to-digital converters for biomedical applications. His research interests include CMOS analog and mixed-signal IC design.

Mohamed Aly Mokhtar (Student Member, IEEE) received the M.Sc. degree in communication technology from the University of Ulm, Ulm, Germany, in 2017, where he is currently pursuing the Ph.D. degree with the Institute of Microelectronics, under the supervision of Prof. M. Ortmanns, with a focus on power-optimized high-resolution incremental sigma-delta analog-to-digital converters.

Maurits Ortmanns (Senior Member, IEEE) is currently a Full Professor with the University of Ulm, Germany, where he is also the Head of the Institute of Microelectronics. He has authored or coauthored more than 250 IEEE journal articles and conference papers. His current research interests include mixed-signal integrated circuit design and self-correcting and reconfigurable analog circuits, with a special emphasis on data converters and biomedical applications.