Confession Session: Learning from Others Mistakes

Contributors: P. Abshire, A. Bermak, R. Berner, G. Cauwenberghs, S. Chen, J. B. Christen, T. Constandinou, E. Culurciello, M. Dandin, T. Datta,T. Delbruck, P. Dudek, A. Eftekhar, R. Etienne-Cummings, G. Indiveri, M.K. Law, B. Linares-Barranco, J. Tapson, W. Tang, Y. Zhai.

Organizers and Editors: Bernabe Linares-Barranco (Sevilla Center for Microelectronics) and Tobi Delbruck (University of Zurich and ETH Zurich).

***Abstract* – People rarely put in their papers the things that didn't work, the mistakes they made, and how they found out what went wrong. Such confessions can help others learn how to avoid similar mistakes. Twenty-six confessions were collected to form the bulk of this paper. Themes that arise are errors that result from not understanding the limitations of simulation tools in modeling physical reality, chip verification errors that result from lack of clear communication between designers, and projects that are considered in their own isolated environment of technical challenges rather than the broader context of their environment or application.**

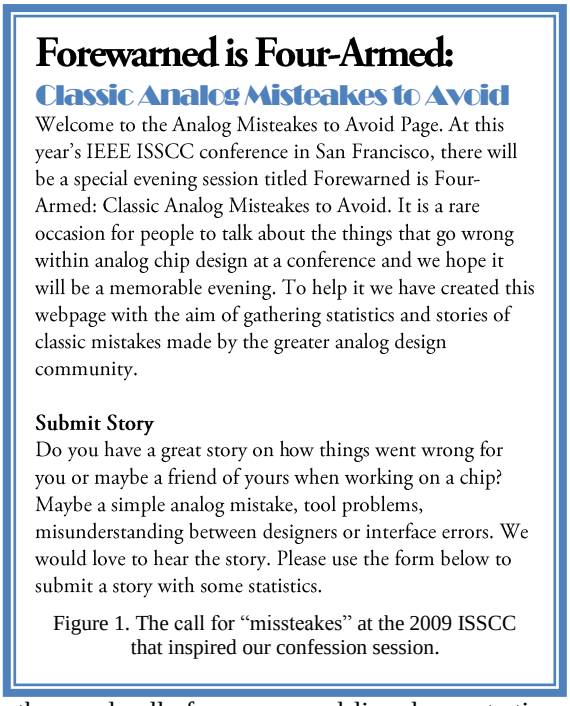
I. Introduction

In 2008, J.E.D. Hurwitz of Gigle Networks was helping to organize the IEEE International Solid State Circuits Conference. He proposed the idea of an Analog Forum dedicated to confessions of goofs in analog design. Sayed Danesh constructed a web page that people could use to submit their goofs (**Fig. 1,** Forewarned 2009). Twenty confessions were submitted to the web page and presented at the forum (but unfortunately are not published). At the session, each confessor had a few minutes and one slide to show what went wrong and why.

We thought this was a great idea and wanted to do something similar at ISCAS, but in the broader context of circuits and systems, and including not just technical blunders but also incorrect strategic decisions over longer time scales. This year's ISCAS will be held in Rio de Janeiro, which probably also qualifies as a good place to confess sins.

What do we hope will be gained? Very simply, people rarely put in their papers the things that didn't work, the mistakes they made, and how they found out what went wrong. We hope these confessions can help you learn how to avoid the same kinds of mistakes. We think this is especially appropriate at ISCAS, which has a large contingent of beginners.

We solicited confessions from the entire Circuits and Systems Society, and the session was also advertised on the main ISCAS web page along with the usual calls for papers and live demonstrations. Contributors were told they would be co-authors on the paper and that they would have between 2 and 4 minutes and one slide to present their confession in the live session. (Contributors were also told they could be anonymous but there were no takers in this category.) We made the call quite general by asking for confessions ranging over all scales, from technical blunders to scientific errors to marketing or strategic mistakes.



Forewarned is Four-Armed:

Classic Analog Misteakes to Avoid

Welcome to the Analog Misteakes to Avoid Page. At this year's IEEE ISSCC conference in San Francisco, there will be a special evening session titled Forewarned is Four-Armed: Classic Analog Misteakes to Avoid. It is a rare occasion for people to talk about the things that go wrong within analog chip design at a conference and we hope it will be a memorable evening. To help it we have created this webpage with the aim of gathering statistics and stories of classic mistakes made by the greater analog design community.

Submit Story

Do you have a great story on how things went wrong for you or maybe a friend of yours when working on a chip? Maybe a simple analog mistake, tool problems, misunderstanding between designers or interface errors. We would love to hear the story. Please use the form below to submit a story with some statistics.

Figure 1. The call for "missteakes" at the 2009 ISSCC that inspired our confession session.

After announcing the call for confessions we got quite a number of emails congratulating us on this idea and wishing us the best success. But confessions were much rarer. By cajoling our friends, we finally gathered a decent set of confessions and these form the bulk of this paper. Because this idea came from the Sensory Systems Technical Committee (SSTC), the contents are heavily (in fact entirely) formed from SSTC member contributions. And because even from the SSTC we had a hard time getting entries, about 1/3 of the confessions here come from the organizers.

We hope that, if repeated, in future years other TCs will contribute in a more balanced manner. We know we are not alone in making mistakes!

So, without further ado, here are this inaugural year's confessions.

II. MISTAKES IN PLANNING

Some confessions result from not considering a project fully in its planning stages.

Confession 1: Tiny dies are hard to handle! (Post-CMOS Processing for MEMS Integration)

Timir Datta, Marc Dandin, and Pamela Abshire, University of Maryland, College Park,

We are developing a suite of sensors for bio-sensing applications. The devices consist of the hybrid assembly of a CMOS die with micro-electro-mechanical systems (MEMS) structures (Dandin et al. 2009). The sensing functionality is achieved on the die with the aid of application-specific integrated circuits and the MEMS structures serve as support for the biological sample under study. The yield of the integration process (the post-fabrication of MEMS structures on the CMOS die) has been unsatisfactory. This results from the small size of the dice ($1.5x1.5mm^2$) that are typically available for prototyping on multi-project runs. Handling these small chips is difficult, and most importantly, impedes the accurate transfer of lithographic patterns on top of the chip; the application of a uniform photoresist layer is not easy since the tinier the sample, the more likely that photoresist will exhibit a significant edge bead after the spin-coating step. The edge bead can take up a large fraction of the available surface area of the chip. This issue is even more pronounced when thick resists are used.

Thus, we learned that when you're planning to perform post-processing on single die, it is important to allocate enough space to allow the success of subsequent lithographic processes. This is of course costly: the larger the die, the more one pays! Moreover, ways to handle the chip and perform the desired post-processing have to be thought of at the design stage in order to improve the post-processing yield.

Moral– For MEMS/CMOS integration, smaller is not always better.

Confession 2: Coordinating different instruments can be nasty work

Marc Dandin and Pamela Abshire, University of Maryland, College Park

Our `confession' is more like some practical advice. We were assembling a fancy optical test bench for measuring responsivity and quantum efficiency of semiconductor junctions and pixels. The system has a grating monochromator equipped with a broadband light source, a calibrated optical power meter, and various data acquisition hardware. The monochromator output signal is coupled to an integrating sphere whose interior is coated with a diffuse reflector, so that the optical signals at the sphere's output ports are spatially uniform and equal in magnitude. The whole system is computer controlled and can accommodate various CMOS sensor boards (Dandin et al. 2007, Sanders et al. 2008).

Integration was the big headache in developing the test bench. The instrument drivers provided by the manufacturers were hard to integrate into a single software platform. We had to start by having a user be present at all times to operate the system. Needless to say, this was very inefficient in terms of time. We finally solved the issue by implementing each function of the system in MATLAB – acquisition, control, and data processing were all accomplished in one MATLAB function. For the instrument drivers written or provided in LABVIEW, we converted them to MATLAB subroutines using the Math Interface Toolkit provided by National Instruments. This allowed LABVIEW virtual instruments to be called and used in MATLAB.

In developing an automated instrumentation system, it's important to identify ahead of time which communication protocol and drivers each instrument uses. Then the next step is to figure out how to bring each piece of software together into one control and acquisition program. It is also nice if the final data processing can happen within the same program. Finally, we would advise that if a Windows-based PC is used, the experimenter turns off or reschedules the Windows Update feature. Our computer once restarted for updates in the middle of the night while acquiring data. This is especially inconvenient if one is trying to collect data for an upcoming ISCAS paper!

Moral – Software for hardware control can be a headache, and getting different components to play nicely sometimes requires Rube-Goldberg-like solutions.

Confession 3: DIP40 packages are good thermal masses

Yiming Zhai and Pamela Abshire, University of Maryland, College Park

A few years ago, we designed a small, low power temperature sensor. We included an array of these sensors on a chip with resistors to act as heaters, so that we could induce and measure temperature gradients across the chip surface (Zhai et al, 2006). We merrily sent off the chip for fabrication without thinking too much about the testing conditions. Unfortunately, once we got the chip back (wire-bonded in a typical DIP40 package) and started testing it, we realized that we were unable to induce temperature gradients across the chip of more than about one degree Celsius. It turns out the DIP40 packages are such good thermal

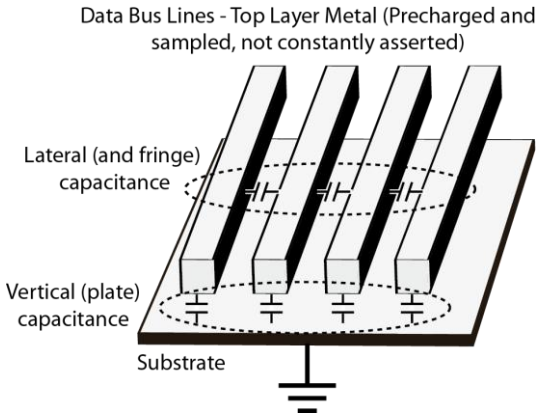


Figure 2. Lateral capacitance can be significant, especially in deep submicron processes. The hand-calculated estimated values were only 1/3 of the actual values.

conductors and thermal masses, that they do an excellent job of dissipating the heat from your chips and equalizing the temperature across the chips.

Moral – It's important to think carefully about your test plan before you bother to fab a chip, especially if you are doing anything even slightly out of the norm. A subplot to the story is that even a "failed" chip can be the topic of an ISCAS paper.

III. PARASITICS NOT CONSIDERED OR IMPROPERLY MODELED

Many confessions resulted from designers not properly considering the limitations of tools to accurately model reality. The confessions from Piotr Dudek are nicely linked to each other.

Confession 4: Bus capacitance parasitics incorrectly calculated for pre-charged output

Timothy Constandinou, Imperial College London

In 2006, we fabricated a (mainly digital) test chip using a full-custom design flow. The output bus was connected to a large number (4096) of output devices so a pre-charging scheme was implemented, to pre-charge the bus in a first phase and then assert the output in a second phase. Because we didn't have access to the parasitic extraction tools supported (Synopsys StarRC) by the target technology (ST Microelectronics 120nm), we estimated the bus parasitic capacitance (by hand calculation) and added this to circuit simulation. After fabrication, to our horror, our estimated values were only a third the actual bus capacitance. This meant that the pre-charge time we had implemented was insufficient to fully charge the bus (even with a 50% margin). What happened? Although we had calculated the capacitance between the bus lines, substrate and power supply nodes; we forgot to include the bus self-capacitance (i.e. lateral capacitance between the bus lines, **Fig. 2**). This (wrongly) wasn't considered as the bus was considered a floating node (incorrect). Also,

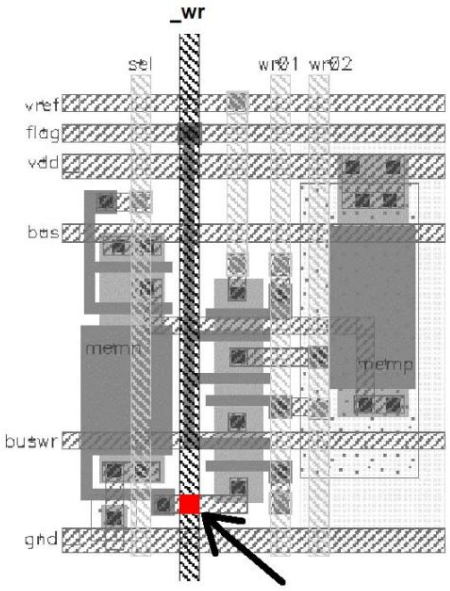


Figure 3. The wire overlap capacitance (marked by the arrow) and the accidentally produced clock phase on the vertical signal wire were just right to compensate clock feedthrough error in the SI cell.

because these bus lines were spaced at minimum spacing this was significant.

Morals –In calculating parasitic capacitances, lateral and fringe capacitances can often be larger in value than vertical (plate) capacitance. When considering parasitics associated to bus lines, always include "self-capacitance" (i.e. capacitance between adjacent lines) in addition to capacitances to substrate (or ground nodes). When parasitic extraction tools are unavailable, have a second designer review hand calculations.

Confession 5: Floating pin improves performance

Piotr Dudek, The University of Manchester

In 1998 I was designing some test circuits for compact switched-current (SI) memories to be used in an "analogue microprocessor" (Dudek and Hicks, 2000). The challenge was to design compact cells without increasing current storage error too much. With relatively little experience, I did not pay sufficient attention to screening sensitive high-impedance charge-storage nodes from digital control signals, causing a number of issues with large clock feedthrough errors. However, one day in the lab, much of the signal-independent error in a certain SI cell suddenly disappeared, turning from barely acceptable to excellent performance. The effect was intermittent, and after much debugging I traced it down to a faulty PCB track - but surprisingly the improved performance was achieved when one of the control signal pins was left floating! Connecting it back to its intended driver was making the circuit perform poorly again. What was going on? After much head-scratching I managed to establish, that when the pin was floating the capacitive coupling from a neighbor wire on a ribbon cable was causing this input pin to be

clocked. With some careful post-layout simulations I figured out that the small capacitance shown in **Fig. 3** between the wire driven from that "disconnected" pin and the charge storage node was compensating the existing clock feed-through error (in a way similar, although somewhat more complex than a regular half-size dummy switch method). With some analysis I was able to establish what perfect conditions for this compensation were – by chance the value of the coupling capacitance that I had, and the clock phase of that nearby signal, were just right! In due course, I analyzed this in much more detail, and I managed to find other, even more effective error compensation methods based on this capacitive coupling phenomenon, I wrote a chapter in my PhD thesis about it, and I have been using these methods to improve the performance of compact SI cells ever since.

Moral – If something weird is going on, always analyze in detail the causes. You never know, your original mistake may well lead to a useful discovery.

Confession 6: Layout extraction tool does not understand Maxwell's equations

Piotr Dudek, The University of Manchester

In 1999 I was designing a vision chip with switch-current memory cells (Dudek and Hicks, 2001). I figured out a very nifty error compensation scheme (that itself was an accidental discovery - see my previous confession) and all I had to do for a perfect result was to have a capacitor of 0.723fF between a polysilicon gate and a digital control line. Of course I am no fool, getting exact 0.723fF was never going to happen, but the error compensation method was clever enough to allow tuning against process variation using a bias voltage, but the 0.723fF was right in the middle of the tuning range for typical process conditions, so I aimed to get exactly there. With some calculations (inter layer area capacitance, fringing capacitance etc.) I determined what overlap I needed between the metal wire and gate poly, and carefully crafted the layout to match, confirmed all with extraction of parasitics (using a top EDA software tool), post layout simulations, etc. I had the value of capacitance I wanted, and was sure that I can tune the bias for best compensation, across all process corners. Everything was perfect. Until chips came back from the foundry, and the error compensation did not work. All results were pointing to the fact that the 0.7fF capacitor that I so carefully designed was not there. It took some head-scratching before I figured out what had gone wrong. The capacitance of about 0.7fF would indeed be there between the poly plate and the metal2 plate - if there was no metal1! Of course, I did not have a metal1 plane in-between; even if I had overlooked this the extraction tool would have caught it. But because the control signal that I was coupling to the gate was routed on metal2, I had simply cut out a notch in metal1 ground plane around the intended capacitor, so the parasitic capacitance extraction based on metal2-poly area and perimeter overlap values was giving me the right number. But looking at the 3D structure (**Fig. 4**) it became obvious that I had a pretty solid shielding in between the plates of my capacitor! Indeed, 3D electromagnetic field simulations of the poly, metal1 and metal2 structure were telling me that I had only about 0.035fF of capacitance there, 20 times less that I designed.

Moral – Don't assume that the layout extraction tools based on 2D geometry will produce a perfect circuit model of your parasitic capacitances. They may be way off the mark.

Confession 7: One of the most complex 3-bit DAC and ADCs ever realized

Gert Cauwenberghs, University of California San Diego

My first VLSI chip design was a memorable and rewarding experience as part of taking Carver Mead's analog VLSI and neural systems course at Caltech in 1989. Inspired by Hopfield and Tank's analog neural model of an optimization network for analog-to-digital conversion, I decided to give it a try and implement this neural model in silicon. Having learned about the problems of mismatch with transconductance-based circuits from the labs, I further decided to do the whole thing with switched capacitors, and figured out an architecture and clocking scheme to get rid of capacitance mismatch altogether. Tobi Delbruck, my TA, thought I was nuts, but was supportive and encouraging nevertheless. The Mathematica symbolic calculations and the Analog transistor-level simulations proved that I was right: the circuit converged to arbitrary precision limited only by the number of binary stages, and by switch injection noise. Not deterred by any potential source of imprecision, I further implemented one of the fancy schemes of switch injection noise cancellation that I

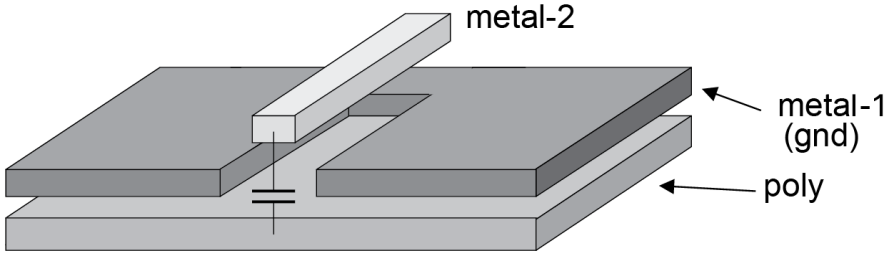


Figure 4. The 3D structure of the poly to metal2 capacitor. There is no metal 1 directly in the designed capacitor area, but it is all around, and it makes all the difference!

found in the literature. With some polygon "pushing" effort I was able to cram the entire circuit of a 16-bit DAC and ADC onto a single Tiny (really tiny!) chip in 3um 2P2M CMOS technology. The layout looked beautiful, certainly when staring at the Wolcomp screen after a 48-hour shift on the Chipmunks! When the chip came back from MOSIS after summer, I was elated to give it the carefully designed test sequence to measure its performance. My enthusiasm was contained by the three effective bits of INL and DNL observed on the oscilloscope. Even though the switch injection cancellation circuits worked as advertised, the mismatch was severe because of one critical oversight: the capacitors were not floating (as modeled in Analog), but had significant backplate capacitance to the substrate.

Moral – As Carver told us all along, listen to the technology and find out what it is telling you. No device in silicon, no matter how elegantly modeled in Mathematica and Analog, is far away from the electrons and holes in the substrate.

Afterthought: The circuit may work a lot better in Silicon-on-Sapphire! If anyone is interested in giving it a try, contact me and you're in for another great experience.

Confession 8: Ground and power connections to core are too slim

Bernabe Linares-Barranco, Sevilla Microelectronics Institute

In 2006 a new graduate student started to design a new retina chip. The student mainly designed a new pixel, as most of the peripheral circuitry was already available or needed minor readjustments. We paid a lot of attention in guiding his pixel design and layout. However, in the final layout, he connected the chip ground to the pixel array ground through minimum metal width connections crossing several metal layers with just one ohmic contact, and connecting at a central point of the array on one side of the array only. As a consequence, there was a relatively high resistance between the chip ground and pixel array ground, producing a significant gradient between the ground of the pixels near the connection and those further apart. The retina was fairly low power, so the connection did not blow out. However, in-pixel bias currents showed a strong local mismatch near the connection. We could feed one of the pixel bias currents to in-pixel integrate-and-fire circuits. The resulting frequency distribution is shown in **Fig. 5**. We can see a background random mismatch with 6 x sigma of about 600Hz, a background gradient plane of about Delta(freq)=600Hz, but at the array ground connection (at column x=17) there is a strong deviation of up to Delta(freq)=2000Hz. The result

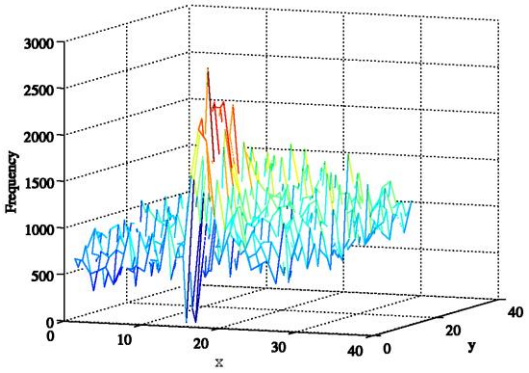


Figure 5. High-resistance power and ground connections from the pad ring to the core caused this distribution of background spike activity across x,y space of the chip.

was that in the retina we were always seeing a thick bar at x=17 running half way through the array.

Moral – Bias power/ground of large arrays uniformly around the periphery, and always have someone experienced fully inspect the first layout of a new student.

Confession 9: Beware of parasitic photodiodes in CMOS image sensor design

M. K. Law and A. Bermak, Hong Kong University of Science and Technology

We designed an image sensor array with a fixed-pattern-noise (FPN) reduction scheme that required no calibration current source. We modeled the expected photocurrents in each pixel photodiode during simulation and everything worked great. We expected the FPN after correlated double sampling (CDS) should be improved by a factor of 15 to 20. However, measurement results showed that there was only an approximately 2 or 3 times improvement. After a long and tedious debugging process, we finally realized the problem was caused by the fact that the pixel output was also light dependent even during calibration. What we overlooked is that there are also PN-junctions in other pixel transistors (**Fig. 6**). In that case, they are also photodiodes when illuminated with light, but we did not include this effect in the simulation!! We should have noticed this effect and put dummy metal over the transistors to shield incoming light. Fortunately we can use some post processing techniques at the sensor output to improve the overall FPN.

Moral – You have to fully understand your circuit before running simulations. Most importantly, never blindly believe in simulation results.

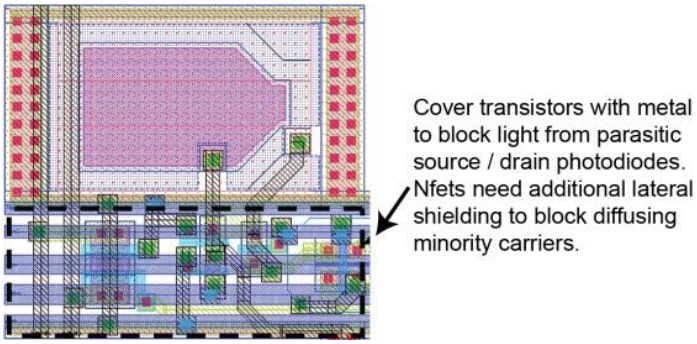


Figure 6. Because the transistors were not shielded from light, their parasitic photodiodes reduced performance of the FPN correction mechanism. Color figure.

IV. Process Design Errors

Some confessions resulted from not understanding or properly designing chip fabrication process flows.

Confession 10: A bipolar imager with one giant pixel

Tobi Delbruck, University of Zurich and ETH Zurich

In 1996 at National Semiconductor and Synaptics in an enterprise that was later to become Foveon we were trying to build a new type of image sensor which Carver Mead invented. It was based on pulsed bipolar phototransistors (**Fig. 7**, Delbruck et. al, 1997). These pixels required developing our own "poly emitter" bipolar process with vertical NPN bipolar transistors and poly emitters, where the base region was self-aligned by the thin oxide regions. After we got the sensor back from fab, we tried to make it work for many weeks, but could never see any image! The pressure was on. All the circuits seemed to be working but all we could see was that the "picture" changed brightness depending on the light intensity. Dick Merrill finally figured it out. He examined the hundreds of lines of detailed process specifications and noticed that the base implant had been set to 400keV rather than 40keV: The base implant was penetrating right through the field oxide so that we had built one single giant photodiode! I still remember the meeting in a conference room in the National fab development center. Dick said something like "Any idiot would know that 400keV penetrates through FOX!" Well, at that point I certainly didn't know it, but nodded my head wisely as if I did. Anyhow, instead of an image sensor with 4 million pixels, we had one giant pixel measuring 4mm by 4mm! Eventually we got it all to work, but after many rounds of silicon we concluded that the FPN in the bipolar gain and the image lag (because the base is never fully reset) was a killer and Foveon went in their storied direction of vertical color separation (Gilder 2005).

Moral? Even an experienced team can be tripped up by a typo.

Confession 11: Metal density rules are there for a reason

Tobi Delbruck, University of Zurich and ETH Zurich

During the early days of imager company Foveon we were working closely with National Semiconductor on process development and were turning a new wafer lot at least once a month (Gilder, 2005). This was in the days of 250nm fab development and the fab guys were using us as testers for their process development. We kept having problems with reliable metal in the pixel arrays. It seemed as though the wires were just not making it across the array, or shorting to each other. Finally, a meeting with the group doing the fab development cleared up the mystery: Our pixel

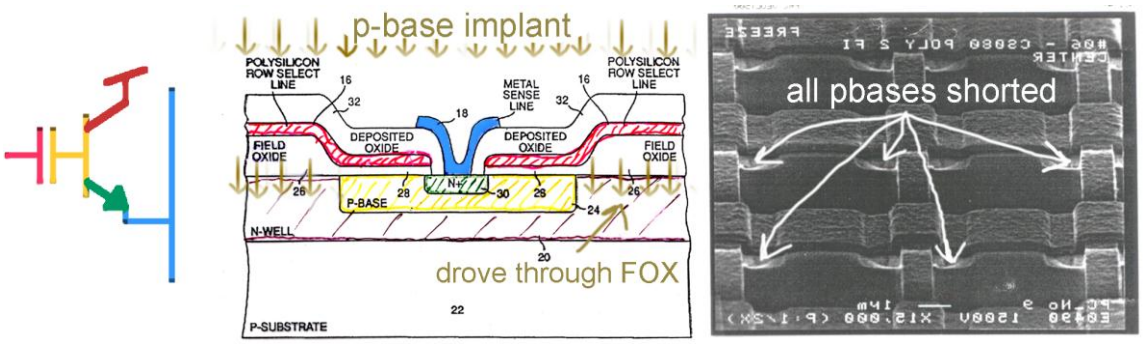


Figure 7. Because the pulsed biopolar pixel (left) p-base implant (center) was set to 400keV instead of 40keV, it penetrated through the FOX, creating a single giant pixel (right). Color figure.

arrays just didn't have enough metal in them. As a result, the chemical mechanical polishing (CMP) was leaving the surface of the array at a different height than the periphery, so that the lithography equipment, which focused using the alignment markers at a corner of the chip, was putting the array out of focus. This defocus blurred out the resist exposure, leading to bad metal. Even the very experienced crew of professionals didn't consider the reason for this metal density rule. CMP was pretty new then, but it didn't help that (as usual) the design rule documentation gave no reasons for any of the rules.

Moral – Design rules have a physical basis. DR documents should provide a bit of motivation to the designers for following the rules.

V. LOOKING AT THE BIGGER PICTURE

Some confessions were inspired by "not seeing the forest for the trees".

Confession 12: I was a PID controller

Jennifer Blain Christen, Arizona State University

As a graduate student, I was attempting to grow cultured cells on a silicon die. Fabricated on the die were a resistive heater and a temperature sensor. Initially, I characterized the temperature sensor to figure out how much heater current was needed to keep the temperature sensor at 37C (body temperature). Using this current I started running experiments but, my data had a strange oscillating pattern high and low, high and low that didn't correlate with any structure on the chip. Hmm...the air conditioning turns on and off, on and off just like the measurements of the temperature on the surface of the chip, could that be a coincidence? The day several of my lab mates all stopped by after their class to hover over my setup to see what I was doing and the temperature spiked like never before, I had to admit that the chip didn't live in its own little world.

How about the nice controlled environment in the cleanroom? That brilliant idea wasn't so brilliant; the temperature didn't stay very steady even though I was in a supposedly stable environment. To make matters worse, every time I turned the microscope light on to take a look at how the cells were doing the temperature jumped by almost 20C. Being a stubborn graduate student, I attempted to grow cells anyway. I could just adjust the power to the heater whenever the temperature sensor didn't indicate 37C...for the 60 hours it would take for the cells to grow and proliferate. Did I mention I was dedicated?

After reporting back to my advisor, the infamous Andreas Andreou, expecting praise for my ability to stay awake for 60 hours straight even though all the cells were dead, I was instead met with a "Jennifer, come on, what are you doing?". Boy did I hate hearing that phrase (for the millionth time) with the requisite head shaking, but that lead my learning how to implement a PID controller...with electronics resulting in some nice papers and a PhD Thesis.

Moral – Hey silly, you are an engineer; start acting like one!

Note added by editors: Christen and Andreou (2007) won the 2010 Transactions on Biomedical Circuits and Systems Best Paper award for their paper on this work.

Confession 13: Optical Order Sorting Filters Matter!

Marc Dandin and Pamela Abshire, University of Maryland, College Park

Order sorting filters are optical filters that are used to block undesired harmonics of an optical signal obtained from a diffraction grating. That is at least what we found out when our measurements indicated that p+/nwell silicon junctions from a 0.5 µm CMOS process were highly responsive at 1100 nm! We knew from experience and from the laws of physics, that this type of junction should not have such a high response in that wavelength band. After investigating the issue, we found that the discrepancy originated from not using order sorting filters at the input slit of the monochromator. When the grating turret was positioned to yield a signal with wavelength = 1100 nm at the output slit, the monochromator was indeed sourcing a signal with the selected wavelength, but its higher orders were sourced as well. Namely, optical signals with wavelength = 550 nm, = 275 nm, …, etc…, in addition to = 1100 nm, were also incident on the detector. The junction is highly responsive at = 550 nm, and we concluded that this caused the unusually high response. The issue was rectified by purchasing a set of order sorting filters from the vendor (Sander et al., 2007).

Moral – When using diffraction gratings, make sure that higher order harmonics of the optical signals are blocked using an appropriate order sorting filter.

Confession 14: Pay Attention to the Statistics of Extreme Events

Jonathan Tapson, Department of Electrical Engineering, University of Cape Town

In his books *The Black Swan* and *Fooled By Randomness*, Nassim Nicholas Taleb has highlighted that people are bad at making common sense judgments about extremely unlikely events, and particularly bad at allowing for the consequences of these events; we always underestimate both the likelihood and the consequences. As a grad student or lab researcher, we tend to work with small numbers of components, and so we seldom see the full range of weirdness which can be exhibited by otherwise sensible designs, when the extremes of the range of

component parameters is explored. One gets in the habit of reading only the "Typ." column of the datasheet, without considering the "Min." and "Max." columns. After all, the consequences of getting an extreme component aren't high. On the other hand, when you produce a system for industrial use and it gets deployed in large numbers, you *will* see every possible value in the entire parameter space, as well as some values which are well outside the stated limits. You had better be prepared for the consequences of industrial-scale malfunction.

This happened to me in 1999 when I was using the first generation of MEMS accelerometers to measure vibration on factory machinery. We built thousands of MEMS devices into robust little packages and stuck them on machinery from Seoul to Sydney. One of the issues we had to contend with, was that these MEMS devices had an undesirable mechanical resonance frequency at about 10x the measurement bandwidth (resonating at about 40kHz with a measurement bandwidth DC-4kHz). We figured we didn't need to worry about it because in the first place, we had a low-pass anti-aliasing filter which cut off more than a decade lower than the resonance, and secondly the Q of the resonance was so high it was unlikely there would be an vibration source at exactly the right frequency, and thirdly big industrial machines don't vibrate with any real energy at 40 kHz, do they? Well, if you put enough devices on enough machines, you find that some of them do indeed vibrate at exactly the wrong frequency around 40kHz, and if the resonance Q is really high, a two-pole low-pass filter at 1kHz doesn't help much at all. I had had a big argument with a colleague as to whether we should go for a higher order cut-off filter, and every time a report came in from the field of another rogue accelerometer, he would shout across the lab "Yeah, we don't need no stinkin' filter!" We wound up having to replace a great many of the units. It would be nice to say that since then I have never again been caught by this kind of mistake, but it would be a lie...

VI. STRATEGIC ERRORS

Although we actively solicited confessions of strategic errors in planning or judgment ranging over a longer time scale, it could be that some of these are still too "raw". The following confessions will hopefully inspire discussion and future confessions.

Confession 15: Don't cough up your core technology

Tobi Delbruck, University of Zurich and ETH Zurich

In forming a development agreement with an industry partner, we had the bright idea that they might be able to help us pay for chip fabrication. That would have been fine except that we also let them manage the actual submission and direct payment of the run. As a result, they got the chips and the full layout of our sensor and were in a position that they could re-fabricate the design, even without asking us - which they did. We were left in a position of having to buy our own design from another party instead of having them buy it from us.

Moral – Consider how your technology will be used if you let someone else have it.

Confession 16: You should start with the big picture

Giacomo Indiveri, University of Zurich and ETH Zurich

An error that I have made over and over again throughout the years, in projects that involve the design of full custom VLSI devices is the following: given a specific research objective (e.g. vision sensors, motion chips, visual tracking sensors, simplified models of selective attention, etc.) I made the error of focusing straight away on the VLSI design aspect and working only on the circuit design, simulation, and layout parts of the project, without thinking about the more general aspects of the projects such as:

1. possible application areas and use cases
2. system-level constraints (power, printed circuit board, size, packaging, etc.)
3. system integration aspects (interfacing to micro-controllers, to input sensors, robotic actuators, etc.)

An important aspect that I have ignored is that of high-level simulations. We have done low-level SPICE circuit simulations, as well as basic C and Matlab behavioral simulations. But we always neglected to do even higher-level, more abstract simulations that take into account possible use-cases. For example, in the case of vision sensors we could have done simulations of sensors in automotive applications (using scenes of cars driving in freeways, traffic monitoring, etc.), or simulations of vehicles landing, including how the scenes change with the sensor in the loop, controlling the vehicle. These types of high-level simulations (e.g. at the level of Java or Simulink) would have made a significant impact on the analog circuit properties (e.g. tuning of dynamic range of speed selectivity), on the types of output interfacing circuits (e.g. on-chip ADC or PWM circuits) or power management (on-chip voltage regulator, bias generators, etc.)

In my specific case, we have been designing event-based multi-neuron chips with spike-based learning capabilities. The learning synapses on each neuron can be stimulated by sending them spike-trains using the AER protocol. Each synapse is trained by stimulating it individually. The neurons on the chip can be used as perceptrons trained to recognize specific input patterns. After several

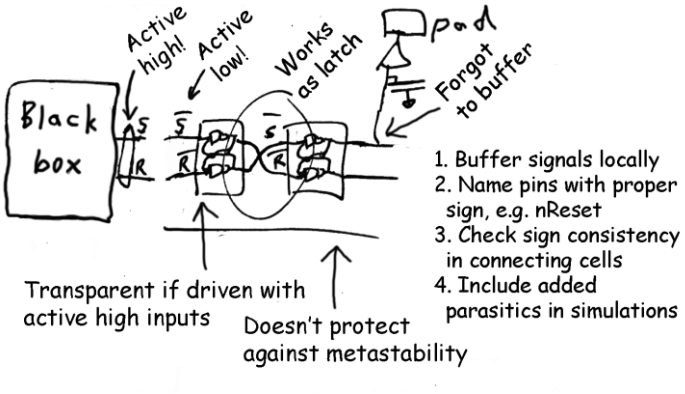


Figure 8. What went wrong in in this asynchronous logic

generations of chip design, testing, debugging, and refining, we are now at a stage where we can start to use these chips in practical applications. The neurons on the chip can be trained to recognize specific patterns, and after training the chip can be used to report if a specific input pattern is recognized or not.

To test if an input pattern is recognized by any of the trained neurons, this should be transmitted to all the neurons in parallel, and the neuron with the highest output firing rate (above some set threshold) would represent the class of the recognized pattern. But using the standard AER circuits, input patterns are sent to the synapses of individual neurons, one synapse at a time. Had we planned on larger scales, we would have designed in the chip a "broadcast" circuit, so that each input pattern is sent in parallel to all neurons. This would have been a simple option to include. But in the current design we have to waste bandwidth and send the same input pattern <n> times in series (where <n> is the number of neurons used), and compare the output rates of each neuron in series, rather than in parallel.

Moral – High-level simulations, together with planning of long-term usage of the device designed at the system level are an important aspect that should be considered from the first steps of a project definition.

VII. LOGIC GOOFS

Logic design errors are still made, particularly for designs using delay-insensitive asynchronous styles.

Confession 17: Confusing active high and active low signals

Raphael Berner, University of Zurich and ETH Zurich

We use cross coupled NAND- or NOR-gates as state-holding elements in asynchronous state machines. Usually we put two of those in series to make sure that the output only starts switching when the new state is already latched, to avoid meta-stability when the state bits are part of the equation for the calculation of the next state, and these state bits are switching slowly due to parasitic capacitive load. For one of our chips, we used the cross coupled NAND-gates in **Fig. 8**, which have active low set/reset inputs. However, if you put two of those in series in the right (or actually the wrong) way, you get latching behavior also with active high inputs, but without safety against meta-stability described above. In this chip, we brought the state bits to pads for diagnostics, but forgot to buffer them with local inverters. This added parasitic capacitance created meta-stability, so the state machines did not work properly. How did we verify this was the problem? On one of the chips, we spent some expensive FIB money to cut the wires from the state machine to the pad. The reduction of the capacitive load was enough to make the chip work. It could have been worse. If we had used local buffers, we may never have seen the bug, leaving it lurking to pop up years from now!

Moral – Double-check whether your circuit uses active low or active high signals (a working simulation is not good enough!!) and always correctly buffer diagnostic outputs. Name your signals according to their active polarity. Simulate your circuit with (at least) the large parasitics that you add for debugging.

Confession 18: Don't just improve asynchronous circuits by electrical simulation

T. Serrano-Gotarredona and B. Linares-Barranco, Sevilla Microelectronics Institute

We had designed several array based circuits with asynchronous event based read out. We used reported event read out circuitry because we were not experts in digital asynchronous circuit design. Reported cells were for unsigned events, but our pixels had signed events, sent through independent lines. So, for each line we used one of the reported circuitry, and just used each pixel as if it was made out of two unsigned pixels. Everything worked fine. However, we felt that we could possible merge the circuitry as the two signs would never be active simultaneously, so we could share and save some circuitry and make pixels smaller. So, we came up with a merged circuitry. Everything simulated very well, so we fabbed a chip with this improvement. When it came back, it would only send out events

for 1-10ms and then freeze. Fortunately, since we were not experts in asynchronous design, from the very first prototype we fabbed, we had the habit of including a reset pin for the full asynchronous read out circuitry. Using this reset pin, we could always revive the chip for another 1-10ms. So, at the end, we fed a 0.1ms period reset signal to the pin, and could test the chip. We found out the error in the asynchronous circuitry (too complex to explain here), but it was due to designing the circuits based on delays, and not on strict handshaking-based delay-insensitive principles. Correcting the pixel using orthodox design principles resulted in a pixel as complex as the originally duplicated one. So, we decided to stick with the original design.

Moral – If you want to do asynchronous design, go by the well-established design principles. Never fully trust an electrical simulation.

Confession 19: A 7 Amp SRAM memory - smaller is not always better

Ralph Etienne-Cummings, Johns Hopkins University

I have a technical one from the Corticon days (Van der Spiegel et al. 1992): a system that was supposed to be low power drew 7Amps per board due to a slight SRAM design error. I designed an SRAM cell that was optimized for size rather than DC current draw. A few by themselves did not have a problem, but when it was scaled up to chip level and then to board level and then multi-board level, this minor omission became hugely significant and the final system statically consumed 85 amps! I usually teach it in my class because smaller is not always better – even for memory cells – and you need to see the forest and not just the tree in front of you. The root cause was using pass-transistors rather than T-gates and not properly sizing the transistors (**Fig. 9**).

Moral – When designing VLSI circuits, we often focus on optimizing area, performance or power at the expense of each other. But system level considerations typically receive short shrift when optimizing at the cell level. Don't forget the context!

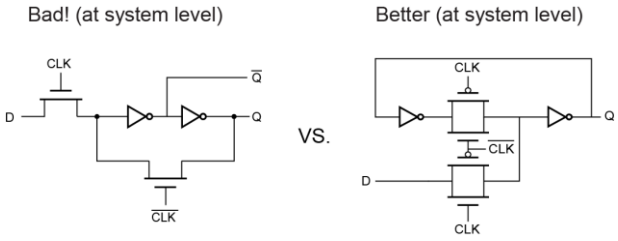


Figure 9. Incorrectly sizing the FETs and using pass gates in the left SRAM cell led to a huge static power consumption at the system level. A better way is shown on the right.

Confession 20: Address Decoding Glitches Reset Pixel

Shoushun Chen, Nanyang Technological University

In 2008 we designed and fabricated a motion detection image sensor using an address counter and decoder. The motivation to use an address counter instead of a scanner is to obtain flexibility in reading out regions of interest. However, we found that in the captured image there were a few rows of pixels having abnormal brightness, or row based mismatch. We also found that the error increased with integration time. At the beginning we thought the problem was due to power supply noise. Finally we realized that the error came from glitches of the decoder, which resets the pixel in the middle of integration (**Fig. 10**). We finally confirmed the mistake in post-layout simulation.

Moral – Firstly, without special delay balancing techniques, the decoder always produces glitches. Secondly, the reset node of the pixel is highly sensitive. Any glitch applied to this node will destroy the integration signal. The decoded row/column reset signal should be resynchronized using a register to filter the glitch.

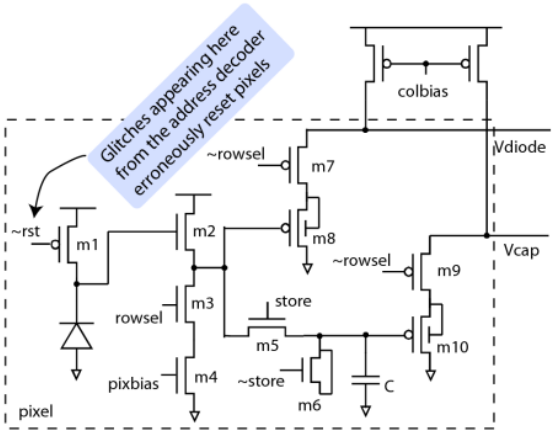


Figure 10. Glitching from decoder logic ruined this design for a motion detection image sensor

VIII. LVS CONFESSIONS

Layout versus Schematic chip design verification errors are presented here.

Confession 21: Wrong pads used on the pad-ring

Amir Eftekhar, Imperial College London

In 2006 we designed and fabricated a chip with two parts of a system that would be linked in a later iteration. The input of the part I designed was an Address Event Register (AER) signal that drove one of these systems, i.e. defining its state and the type of output. We didn't leave ourselves a lot of time for layout and so had several people on board to help. This coupled with insufficient planning and communication meant that certain elements were rushed through including the pad-ring. The result was that the AER input pads were assigned as an output rather than digital inputs. As such, after fabrication, we found our system tests to be showing quite random outputs (although correct, just not controllable from the AER input) - due to whatever floating voltage was on the input. The time we spent debugging and testing was comparable to the design-time. It was only when looking back at the design schematics and layout did we find our mistake. The problem was that the core system was designed and simulated by one person but as mentioned the layout was assigned to several people. Having a few sleepless nights of layout it was inevitable that a simple miscommunication would result in such a mistake.

Moral – Designing a chip without planning and check-lists is not advisable (especially with a few days to go). With lots to do and not always a lot of time you tend to not double check specifications and progress especially if you are relying on a lot of different people – always use planning documents and check-lists!

Confession 22: Chip lacks a global reset

Tobi Delbruck, University of Zurich and ETH Zurich

We recently fabricated a test chip and forgot to bring the global reset node to a pin. How was this possible? It happened because one person was designing the chip core and the other the pad-frame and periphery. The core designer brought all the core outputs out, but not to one place all gathered together and clearly labeled as being intended for pad-frame connections. The periphery designer wired up all the core outputs, but not the hard-to-see global reset. The chip passed LVS because the reset pad was missing from the layout and schematic. As a result, our chip may either not be resettable or stuck in reset, depending on the floating voltage of the reset node.

Moral – Bring all your core signals clearly out to a few clearly labeled peripheral connection buses! As soon as you know a signal needs to go to a pad, put it in your schematic!

Confession 23: Pull-up transistors are missing

Rafael Serrano-Gotarredona, T. Serrano-Gotarredona and B. Linares-Barranco, Sevilla Microelectronics Institute

In 2003 we designed our first AER Convolution test chip prototype with 16x16 pixels. The design included a VHDL designed synchronous controller, a mixed mode pixel array designed in full custom, and asynchronous read out circuitry to send out the events produced by the pixel array. The complete schematic was fairly complex and could not be fully simulated at the electrical level. Therefore, electrical simulations were performed by parts only. To simulate the full schematic, we described the different parts by AHDL. AHDL descriptions were very simplified versions of the actual operation. The pixel array sends out digital signals (wired-or by rows) to a peripheral arbiter. These lines had pull-ups (or pull-downs) at the periphery. It turned out that at the end, the peripheral pull-up transistors were missing, since they played no role for the AHDL description we made. In the layout, we also forgot to include them, and LVS checks were OK. However, when testing the chip, it was dead: no events ever came out, although the chip would successfully consume input events. We quickly found out the mistake. Fortunately, we included some spare pads in the pad ring and could FIB some samples to route 2 of the lines to spare pads and provide the pull-down transistors outside of the chip. This way, we could verify that the rest of circuits operated correctly, redesigned quickly the chip to include pull-ups, and had a corrected chip back in about 5 months.

Moral – Always include spare components for FIBing, learn more about how to do mixed signal simulations and verification, and do your behavioral descriptions as faithfully as possible.

IX. ANALOG GOOFS

The art of analog design has become heavily reliant on simulation tools and theory. As can be seen, these often both fall short of describing reality.

Confession 24: Problems of asynchronous delta modulator used in bio-potential signal recording

Wei Tang, Eugenio Culurciello, Department of Electrical Engineering, Yale University

In 2008 we designed a bio-potential sensor using an asynchronous delta modulator to perform A/D conversion (**Fig. 11**). The advantage of the design is the low complexity of the A/D which can even be merged seamlessly into the front end amplifier (Tang and Culurciello 2009). However, we did not put enough consideration into the amplifier reset. In the measurement we found if we try to increase the resolution, we need to reset the delta modulator more frequently, which increases the fraction of time in reset. The problem is that during the resetting time, the input analog signal is lost. So

finally we had worse distortion in the output when higher resolution is desired. Also we saw that non-ideal effects such as digital noise from the output pulse to the input signal is higher than simulation and that noise and the mismatch between the positive delta and the negative delta cause an unpredictable DC shift in the final recovered output signal. How did we fix these problems? We redesigned the system by putting a high speed asynchronous switched-capacitor buffer between the analog amplifier and the delta modulator. We also used a digital filter to remove the DC shift (Tang et al. 2010).

Moral – In mixed signal system design we need to put more attention on the interface between analog blocks and digital blocks, especially when new techniques are used. Such design relies on experience and working designs more than simulation models.

Confession 25: Don't give up your freedoms before you know you don't need them

Raphael Berner and Tobi Delbruck, University of Zurich and ETH Zurich

We have our own confession about switch injection and can confess it now that we finally understand it. It's a result of trying to be too tricky and (just like in software) doing the bad thing of optimizing before getting your thing working in the first place. It's also a case of giving up your degrees of freedom without really needing to. We were building the two stage self-timed switched capacitor amplifier with two amplifiers, *A1* and *A2* in **Fig. 12**. Resetting (balancing) these amplifiers requires two clock signals, $\phi 1$ and $\phi 2$. We generated these using a source follower to generate $\phi 1$ from $\phi 2$, which was supposed to guarantee that *A2* would be held in reset until after *A1* was done being reset. Why? So that the switch injection from resetting *A1* would not appear (in a greatly amplified form) at the output of *A2*. The problem is that our circuit doesn't do that properly. Instead, *A1*'s output o1 is still changing when the $\phi 2$ switch is opened, so that we get a huge offset appearing at o2, the output of *A2*.

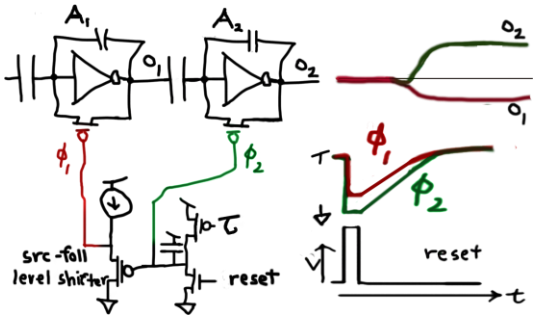


Figure 12. Because of our fancy scheme for generating ϕ_1 from ϕ_2, o_1 is still changing when o_2 is leaving reset, causing a huge offset in o_2. Color figure.

We could have avoided this problem if we had separated the timing of $\phi 1$ from $\phi 2$. Instead, by not being able to control the $\phi 1/\phi 2$ timing separately we cannot avoid this problem. It is especially bad because the more we turn on the source follower to split $\phi 1$ from $\phi 2$,the higher is the $\phi 1$ reset voltage level, and the more slowly $\phi 1$ can reset *A1*. We could have seen this if we had done a proper simulation of the circuit that used the bias currents necessary in real operation!

Moral – Get something working before you optimize it. Ensure your simulation makes sense. If you measure something that doesn't make sense, make sense of it.

Confession 26: Blinded by theory to reality for 9 long years

Tobi Delbruck, University of Zurich and ETH Zurich

As a young and inexperienced PhD student in Carver Mead's lab I invented a circuit that John Harris coined as the "bump circuit". It is useful for nonlinear similarity and dissimilarity measurements (Delbruck, 1991). It takes a differential input voltage and produces output currents that tell you if the voltages are the same or different (**Fig. 13a&b**). I was so proud of my very first baby and so happy

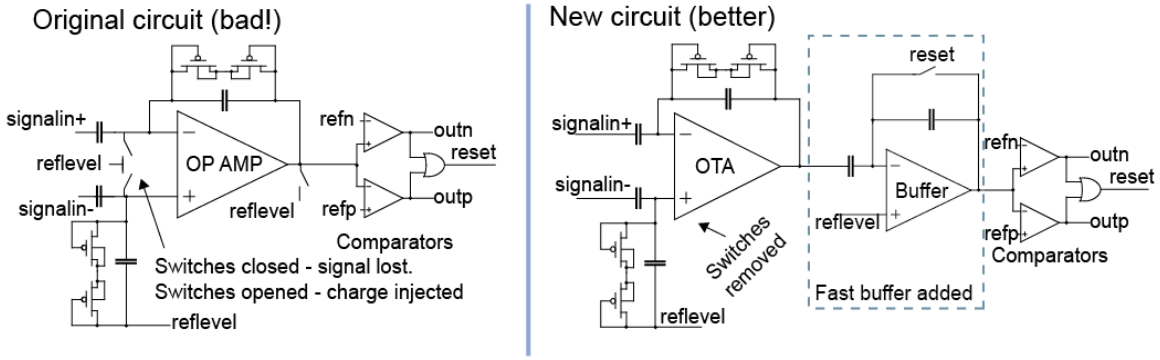


Figure 11. Switches in the original circuit at the sensitive analog input were a bad idea. The new circuit runs the front end in continuous time and only switches after impedance is low and signal is amplified.

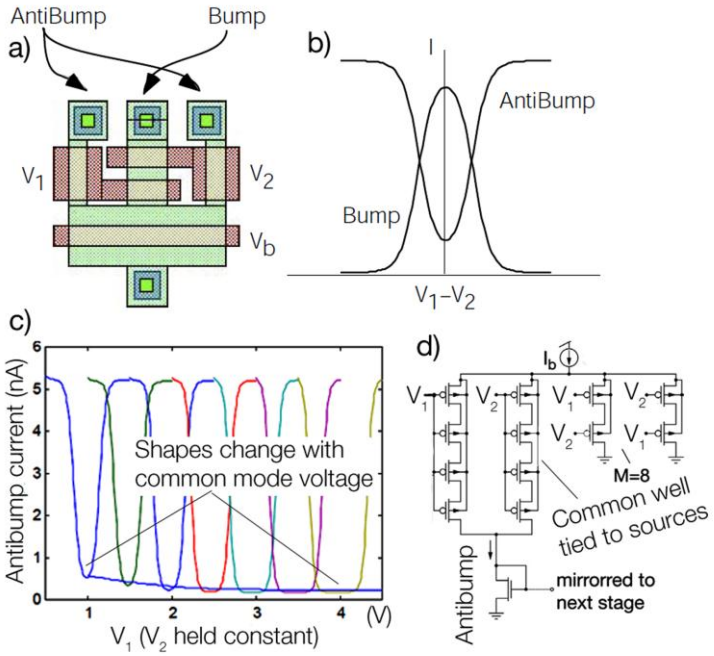


Figure 13. The bump-antibump circuit layout (a) and differential input IV curves (b). The common mode responses of the antibump output are in (c); each curve is from holding one input fixed and varying the other one. The shape of the bowl changes dramatically with common mode DC level. This change in the shape can be removed by tying the wells to the pfet sources, as shown in (d). Color figure.

and that the IV curves matched subthreshold transistor theory that I neglected to check something really basic, namely, the common mode response! It took me 9 years and a chip that didn't work very well to finally make a measurement of this very large effect (Delbruck, 2000), which greatly disturbs use of the antibump output as a form of soft rectification (**Fig. 13c**). Why does the shape of the bowl change? The bump circuit relies on transistor geometries with very different W/L aspect ratios. The middle transistors must be strong compared with the outer ones. To make the layout compact, this big ratio requires the middle transistors to be short and wide and the outer ones to be long and skinny. These constraints on compact layout really expose the mysterious "short and narrow channel" effects that are not discussed in introductory transistor physics course and also not modeled very well in SPICE. They boil down to the fact that narrow transistors act even narrower than you draw them, and short ones even shorter. (Think of the transistor channel as a mountain pass with clouds of carriers diffusing from one valley to another and you can get the picture.) And these effects are also a function of the back gate voltage and current. As a result, the effective ratio of transistor strengths between middle and outer legs is not only much larger than predicted by drawn geometry but also a function of back gate voltage. The end result is that if you try to build an antibump circuit with a nicely defined bowl-like response characteristic, then the shape of this bowl is also a strong function of the common mode voltage. You can get around this, but only by making all the FETs p-type and tying their wells to the pfet source voltages (Delbruck et al., 2009). And if you really want a shape you design, you have to use unit transistors - as usual for any precise desired geometry (**Fig. 13d**).

Moral – Theory is not all of reality. Don't be blinded by it.

X. Conclusions

Although our sample size of about 20 confessions is too small to draw reliable conclusions, we can still classify the mistakes. It should be kept in mind that there is a huge sample bias here, since the confessions all originate from the Sensory Systems Technical Committee and are dominated by simple technical blunders and not by the more interesting errors in judgment that are the basis of strategic errors.

Still, we can see that there are some trends. One theme that arises again and again are errors that result from not understanding the limitations of simulation tools. That is why it is amazing to us that papers that report simulation results on CMOS circuits are accepted to any conference. Another trend is that LVS errors often result from lack of clear communication between designers, or lazy annotation of schematics that follows after layout. Another trend is that projects are often first considered in their own isolated environment of technical challenges, and not in the broader context of their environment or application.

We hope that this year's session is lively and that it will end up encouraging future confession sessions that span more of CAS.

ACKNOWLEDGEMENTS

We thank Sayed Danesh, J.E.D. Hurwitz, the special session chairs Yehia Massoud and Eduardo da Silva, and the general organizers Barbara Wehner, Sergio Lima Netto, Paulo Diniz, and Bassen Lande for their support of this special invited confession session.

REFERENCES

Christen, J.B.; Andreou, A.G. "Design, Fabrication, and Testing of a Hybrid CMOS/PDMS Microsystem for Cell Culture and Incubation," Biomedical Circuits and Systems, IEEE Transactions on , vol.1, no.1, pp.3-18, March 2007

M. Dandin, I. D. Jung, M. Piyasena, J. Gallagher, N. Nelson, M. Urdaneta, C. Artis, P. Abshire, and E. Smela, "Post-CMOS packaging methods for integrated biosensors," in IEEE Conference on Sensors, Christchurch, New Zealand, 2009, pp. 795—798.

M. Dandin, N. M. Nelson, H. Ji, and P. Abshire, "Single photon avalanche detectors in standard CMOS," in IEEE Conference on Sensors, Atlanta, Georgia, USA, 2008, pp. 585—588.

T. Delbruck, "'Bump' circuits for computing similarity and dissimilarity of analog voltages". (1991) Proceedings of International Joint Conference on Neural Networks, July 8-12, 1991, Seattle Washington, pp I-475-479.

T. Delbrück, N. Mascarenhas, M.-H. Chi, A. Bergemont, C., "Pulsed bipolar CMOS imager". (1997) Mead. Proceedings of IEEE CCD/AIS meeting, Brugge, Belgium, June 5-7 1997.

T. Delbruck. (2000) "Silicon retina for autofocus", Invited session on Bio-inspired Microelectronics, ISCAS 2000, May 27-31, 2000, Geneva, Switzerland, pp. IV-393-396.

T. Delbruck, T. Koch, R. Berner, H. Hermansky, "Fully integrated 500uW speech detection wake-up circuit", in ISCAS 2010, Paris, pp. 2015-2018.

P.Dudek and P.J.Hicks (2000) "A CMOS General-Purpose Sampled-Data Analogue Processing Element", IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing, vol. 47, no. 5, pp. 467-473, May 2000

P.Dudek and P.J.Hicks (2001) "A General-Purpose Vision Chip with a Processor-Per-Pixel SIMD Array", European Solid State Circuits Conference, ESSCIRC 2001, pp.228-231, September 2001

Forewarned is Four-Armed: Classic Analog Mistakes to Avoid. Available: http://www.see.ed.ac.uk/~s0128661/analog/, 2009

G. Gilder, The Silicon Eye: How a Silicon Valley Company Aims to Make All Current Computers, Cameras, and Cell Phones Obsolete. New York: W.W. Norton & Company, 2005.

D. Sander, M. Dandin, H. Ji, N. Nelson, and P. Abshire, "Low-noise CMOS fluorescence sensor" in IEEE International Symposium on Circuits and Systems, New Orleans, Louisiana, USA, 2007, pp. 2007—2010.

J. Van der Spiegel, P. Mueller, D. Blackman, P. Chance, C. Donham, R. Etienne-Cummings, P. Kinget, "An analog neural computer with modular architecture for real-time dynamic computation, IEEE J. Solid State Circuits, 27:1, pp. 82-92, 1992.

Wei Tang; Culurciello, E.; , "A pulse-based amplifier and data converter for bio-potentials," IEEE International Symposium on Circuits and Systems, 2009 (ISCAS 2009). pp.337-340

Wei Tang; Chenxi Huang; Dongsoo Kim; Martini, B.; Culurciello, E.; , "4-Channel asynchronous bio-potential recording system," Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS 2010) , pp.953-956

Y. Zhai, S. B. Prakash, M. H. Cohen, and P. Abshire, "Detection of on-chip temperature gradient using a 1.5V low power CMOS temperature sensor," IEEE International Symposium on Circuits and Systems 2006 (ISCAS06), Kos, Greece, May 21-24, 2006.