Analytic and Simulation Results about a Compact, Reliable and Unbiased 1-bit Physically Unclonable Constant

Riccardo Bernardini and Roberto Rinaldo DIEGM – University of Udine
Via delle Scienze 208, 33100 Udine, Italy
Email: {riccardo.bernardini,rinaldo}@uniud.it

***Abstract*—Physically Unclonable Constants (PUC) are circuits used to embed unique secret bit-words in chips. We propose a simple PUC, with a complexity comparable with an SRAM cell. The proposed scheme is studied both theoretically and by means of simulations and it is shown that the proposed PUC is both unbiased and very stable. In particular, its intra-distance is predicted to be from 10 to 100 times smaller than competitor schemes. Simulations allow to conclude that the advantages of the proposed scheme are relevant enough to make it competitive even if the actual performance of a real implementation, not considered in this paper, will turn out to be an order of magnitude worse than predicted.**

***Index Terms*—security, physically unclonable functions, chip authentication**

This copy of the manuscript shows in blue the most important changes done to the manuscript. This copy is provided for reviewers' convenience and submitted as *Supporting Documentation.* An official, non-marked copy is submitted as *Main Document.*

I. Introduction

The necessity of verifying the authenticity of a chip in a simple and secure way gave rise to the introduction of Physically Unclonable Functions (PUFs) [1]–[6]. A PUF is a circuit that implements a map from bit-words to bit-words, with the actual map very sensitive to the exact values of process parameters (e.g., the exact channel length of a MOSFET or the exact doping level). As a consequence of such dependence, the map implemented by a specific chip will be unique to that chip and this can be used to verify the identity of the chip [7], [8]. Moreover, such a sensitivity makes it very difficult to replicate the PUF of a specific chip. In a sense, a PUF is like a fingerprint: as each person has a unique fingerprint whose minutiae are the result of casual variations during the fetal development, every chip has its own PUF that is the result of casual variations during chip production.

A special type of PUF is a PUF with no input arguments, that is, a constant. For this special type of PUF are called *weak PUF*s, *Physically Obfuscated Keys* (POK) or Physically Unclonable Constant (PUC). PUCs can be used to embed in chips secret bit-strings that can be used, for example, as private keys for encryption or authentication or as source of randomness in special applications [10].

The ideal PUC is a *random constant* [11], [12] in the sense that at production time a random bit value (called in the following the *preferred outcome*) is uniformly selected and every time the PUC is queried said selected value is returned. Real PUCs, however, can depart from ideality in two respects: (i) sometimes the PUC can make an *error*, not returning the preferred outcome and (ii) the selection of the preferred outcome could be not uniform. These two forms of non-idealities suggest two PUC quality indices: *stability* (a stable PUC always returns the preferred outcome with overwhelming probability) and *unbiasedness* (in an unbiased PUC the preferred outcome is uniformly selected at construction time). It is clear that reliability is very important since in many security applications a single wrong bit can render the whole system useless.

The problem of improving the reliability of a PUC stimulated research in the field of *PUC stabilizers* [7], [9], [13]–[16]. It is worth observing that every stabilizer proposed in the literature introduces some kind of "redundancy," (e.g., syndrome bits in helper-based schemes [7], [13]–[15], spare cells or repeated turn-ons in helper-less schemes [9], [16]) and that less reliable PUCs require more redundancy to be stabilized. Therefore, there is interest in designing reliable PUC schemes that can be used with low-redundancy stabilizers or, better, so reliable that no stabilizer is required.

A. Prior Work

Among the PUCs the most popular schemes are based on SRAM or similar structures [17]–[22]. In [17] the initial state of a non-initialized SRAM is used as the PUC outcome. In [18], [20] a latch-like structure is used to amplify an offset voltage. A different approach based on the measure of the data retention voltage of an SRAM is described in [21]. In [23] a PUC amplifying the difference in the threshold voltage of two NMOS is described; an approach very similar to [18], [20], [23] is described in [24] where the variations in the threshold voltage of MOSFETs are used to generate voltages that are mapped to 0 or 1 by a comparator made with two inverters; an approach based on Flash memory is proposed in [25]; [26] exploits the antenna effect in order to randomly break the gate oxide. The schemes based on uninitialized SRAM or latches [17], [18], [20] have the drawback that the underlying structure has two stable states and it can happen that the PUC ends in the "wrong" state. For example, according to [20], 4% of the latch-based cells are "unstable." According to [6], a similar result holds also for SRAM-based schemes. SRAM schemes gave rise also to some work about counteracting aging. In [27] presents some anti-aging techniques which are based on data-dependent aging effects, while [28] develops new metrics to analyze the relationship of reliability between neighboring SRAM cells. Said metrics are used to examine in detail the impact of environment.

Schemes based on comparators fed with random voltages like [23] and [24] have the drawback that, given the continuity of the transfer function of a real comparator, there is a non negligible probability of having cells whose output is sensitive to noises. In [24] a 5% of unstable bits is reported.

Although the scheme of [26] is interesting because of its stability and low power consumption, it is suggested the over-voltage used to break the oxide could cause chip degradation [23].

B. Our contribution

The very simple 1-bit PUC described in this paper originated from the analysis of the causes of the mediocre stability (i.e., an intra distance μ_{intra} [29] ranging from 3% to 12% [6], [19], [24]) that characterizes the schemes that are more similar to our proposal, namely, memory-based approaches (e.g., SRAM and latch) [17]–[20] and comparator-based approaches [23], [24]. As said above, such a stability is due to the presence of two stable states (memory-based PUCs) or to the continuity of the implemented map (comparator-based PUCs).

This observation suggested us to search for a system that (i) *has one and only one* stable state and (ii) the position of the stable state is a *discontinuous* function of the circuit "unbalance." The existence of only one equilibrium state makes the system very robust: *independently* on the initial condition or any initial transitory noise, the system will evolve, sooner or later, to the unique equilibrium point. The result is a PUC whose intra-distance μ_{intra} [29] is predicted to be *10 to 100 times smaller* than the intra-distance of similar schemes.

In this paper we describe our PUC proposal, together with a thorough analysis of its behavior, first in a qualitative way, then in a more analytic way and, finally, by means of simulations. The simulations show that the theoretical predictions still hold even when the simple models used for the theoretical analysis are replaced by the more complex and precise models used in the simulations, making the theoretical predictions quite convincing.

We choose this approach, rather than directly measuring the circuit behavior, because we wanted to achieve a good understanding of how the circuit behavior is affected by the circuit parameters, in order to determine some design guidelines. Some of the guidelines (e.g., use "long and thin" transistors) are actually not obvious and difficult to find by only experimental analysis.

Of course the final confirmation of the performance of the proposed scheme can only come from experiments that will also allow to analyze some aspects (local biasing [30], aging, temperature dependence) not considered in detail here for the lack of suitable models. The experimental assessment of the behavior of the proposed PUC will be the subject of future investigation.

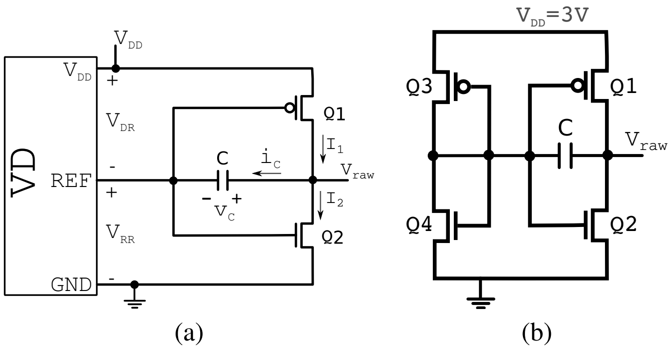


Figure 1. (a) The proposed 1-bit PUC (b) The instance used in the simulations, with a MOSFET-based voltage divider

C. Structure of this paper

In Section II we introduce some preliminary remarks about nomenclature, quality measure and security discussion. In In Section III we introduce the proposed scheme and do a first, qualitative analysis. Before moving to a more analytic study in Section V, Section IV introduces some notation and some strategies that will be used for the analysis. Section V presents an analytic study of our solution. Noise impact is studied in Section VI and the results are used in Section VII to determine the quality of our scheme. Other issues, like the dependence of the behavior from the temperature, aging and cell size, are analyzed in Section VIII. The study is completed with some simulations whose results are given in Section IX. Finally, Section X gives the conclusions and describes future research directions.

D. Summary of Results

In this section we briefly summarize the main results in this paper. In order to do this, it is convenient to anticipate briefly how our scheme works. Consider the scheme of Fig. 1a where transistors Q1 and Q2 are designed to be *nominally* matched. Call I_0 the current on C at $t = 0$. Symmetry considerations suggest that if Q1 and Q2 were exactly matched, it would be $I_0 = 0$. However, Q1 and Q2 will never be exactly matched and this will cause a current $I_0 \neq 0$ on C. It will be shown in the following that the outcome of the PUC depends only on the sign of I_0.

We will denote with I_s the saturation current of Q1 and Q2, with V_{eq} the value of v_C at equilibrium and with $\bar{I}_0 = I_0/I_s$ an adimensional version of $\bar{I}_0$. It turns our that $\bar{I}_0$ is the most important quantity in our scheme.

The main results of this paper are the following

- The key result is that the circuit of Fig. 1a has *one and only one equilibrium point* (stable) *whose position is a discontinuous function* of $\bar{I}_0$. If the channel length modulation can be neglected ($\lambda \approx 0$), it is $|V_{\text{eq}}| > V_T$ even for very small $|I_0|$. This makes our scheme robust to noise. If $\lambda > 0$ the results remain qualitatively the same, with the difference that in order to have $|V_{\text{eq}}| > V_{\text{T}}$ it is necessary to have $|\bar{I}_0| > 2\lambda V_{\text{T}}$. See Sections III, V-A and V-B for more details.

- *Current $\bar{I}_0$ is a convenient "quality measure" of a specific instance.* Not only cells with larger $|\bar{I}_0|$ are more insensitive to noise, but they are also less sensitive to temperature changes (Section VIII-A) and aging (Section IX-C). This can be exploited to disable at the first turn-on the less reliable cells, i.e., with small $|\bar{I}_0|$. See Section VIII-C3.
- We give an *analytical statistical characterization* of our scheme, including the *power of noise* affecting the cell (Section VI and (27)), the *Probability Density Function (pdf) of $\bar{I}_0$* and the *pdf of V_{eq}* (Section VII-B, equations (31) and (32)). This allows us to predict analytically that our scheme can have an *intra distance as small as* 10^{-4} and an *inter-distance* pratically *equal to* $1/2$. See Sections VII-C and IX-D.
- We provide some *design guidelines*. For example, C should be choosed as small as possible, compatibly with the load, while the transistors should be "long and thin." See Section VIII-C
- We estimate the cost in terms of *silicon area* and predict that our scheme is quite *competitive*, taking into account that the very small intra-distance allows us to use smaller error-correction codes o no correction at all. See Section IX-D2 and tables III, IV and V.
- It is possible to implement this scheme so that the *energy required* is a fraction of nJ/bit. See Section V-E.
- Our scheme is *stable* with respect *aging* (Section IX-C) and *temperature variation* (Sections VIII-A and IX-A). Moreover, property $|V_{\text{eq}}| > V_{\text{T}}$ makes our scheme *robust with respect to power variations*, since V_{T} does not depend on the supply voltage. See Section IX-B and Fig. 8.

1) Limit of the analytic/simulation approach: Although the analytic study of the circuit allows us to understand better the impact of the different variables on the final outcome, they are nevertheless based on models. It is clear that only actual experiments on prototypes can say the final word about the performance of the proposed scheme. One could wonder how much likely is that the results described in this paper will be confirmed by experiments.

- The main characteristic of our scheme (*only one stable equilibrium point*) is expected to hold in every case, since it depends only on the "saturating" and monotone behaviour of the transistors.
- Also that the *discontinuity property* $|V_{\text{eq}}| > V_{\text{T}}$ is expected to hold since it is a direct consequence of said saturating behaviour. Also the *stability* against *supply voltage variations* is expected to hold.
- The *statistical characterization* in Section VII-B are expected to hold qualitatively, since the hypothesis used to derive them are quite weak (basically, the distribution of the characteristics of the transistors is approximately Gaussian). Of course, the quantitative details can change, they could even depend on the specific foundry.
- The difference of *performance* between the proposed scheme and schemes in the literature (Section IX-D2) is so large that we expect that the proposed scheme will maintain its competitiveness.
- Because of the loss of suitable models, the predictions that are on a shaker ground and that can be verified only experimentally are those relative to the *temperature dependence* (Sections VIII-A and IX-A), *aging* (Section IX-C) and *local bias* (Section VII-E).

Summarizing, we expect that most of the characteristics described in this paper will qualitatively hold also in an actual silicon prototype, although there can be deviations on a quantitative level.

II. Preliminary Remarks

A. Nomenclature

Depending on the context, "PUC" can mean both the nominal circuit (e.g., the circuit in Fig. 1) or a specific implementation of it (e.g., a specific cell in a specific chip). This double use makes discussion difficult and introduces ambiguities. Therefore, we will say *PUC scheme* to refer to the abstract scheme and *PUC instance* to refer to a specific physical implementation [31].

Observe that a PUC can be modeled as a two-step experiments: (i) when a PUC instance is built, the instance *preferred outcome (PO)* and the corresponding probability is determined, successively (ii) when the instance is queried, an outcome is randomly drawn according to the probability selected at construction time.

Informally, a PUC *instance* is said to be *stable* if every time it is queried it gives PO with overwhelming probability. A PUC *scheme* is said to be *stable* if its instances are reliable with large probability. Finally, a PUC scheme is said to be *independent* if the POs of different instances are independent random variables (r.v.). These concepts are made more precise in the following section.

B. Quality measures

Two standard quality measures for PUF are the *inter-distance* μ_{inter} (the distance between the two responses resulting from applying the same challenge to two different PUFs) and the *intra-distance* μ_{intra} (the distance between the two responses resulting from applying the same challenge twice to the same PUF) [29]. However, since a PUC has no inputs, these two measures are not directly applicable and another approach is required [9], [31].

Let p_1 be the probability that a specific instance will return "1" when queried and let O_{pref} be the corresponding PO, that is, $O_{\text{pref}} = 1 \Leftrightarrow p_1 > 1/2$. Note that p_1 and O_{pref} are r.v. drawn at construction time. The *stability* of the *instance* is defined in [9], [31] as

$$R(p_1) \overset{\text{def}}{=} 2|p_1 - 1/2| \tag{1}$$

Note that $R(0) = R(1) = 1$ (perfect stability) and $R(1/2) = 0$ (instability). Function R measures the stability of an *instance*, in order to measure the stability of a *scheme*, one can use the *Stability Distribution Function* (SDF) [9], [31]

$$F_r(x) = \text{P}[R(p_1) \leq x] \tag{2}$$

Finally, we define the *bias* of a PUC scheme as $\varepsilon_P = |P[O_{\text{pref}} = 1] - 1/2|$. If $\varepsilon_P = 0$ the scheme is *unbiased.*

1) Relationship between the SDF, μ_{intra} and μ_{inter}: It possible to show by basic algebra (see Appendices A-A and A-B) that from the SDF one can compute μ_{intra} amd μ_{inter} as follows

$$\mu_{\text{intra}} = \int_0^1 F_r(x)\, x\, dx \lessapprox \int_0^1 F_r(x)\, dx \tag{3a}$$

$$\mu_{\text{inter}} = 2(m_p - m_p^2) = \frac{1}{2} - 2\eta^2 \tag{3b}$$

where $m_p = 1/2 + \eta = \mathbb{E}[p_1]$ is the mean of p_1 and the approximation in (3a) is good when the PUC scheme is stable.[1]

Remark II.1

Both $\varepsilon_P = 0$ and $\mu_{\text{inter}} = 1/2$ are indications of the unbiasedness of a PUC scheme. Curiously, they are not equivalent and it can happen that one condition holds, while the other does not. However, it is easy to see that they are practically equivalent when the scheme is *stable*, that is, when the pdf of p_1 is concentrated around 0 and 1.

C. Security Discussion

A detailed security analysis would require to know how the secret outcome of the PUC is used. However, on a general level, we can say that there are two possible attacks: (i) attack the cryptographic protocol employed, or (ii) use a (smart) brute-force attack to guess the outcome of the PUC. Note that the probability of success of the first kind of attack depends mostly on the protocol employed, not on the quality of the PUC. Therefore, a discussion about this type of attack is out of scope here.

The quality of the PUC determines the probability of success of the second kind of attack. Brute-force search could be done, possibly, in a smart way by trying first the most probable outcomes [36]. Clearly, the effort of the attacker is maximized when the outcomes are uniformly distributed. This happens *if and only if* the PUC is *independent* and *unbiased*. Therefore, employing a PUC with small μ_{intra} and large μ_{inter} together with a good cryptographic protocol guarantees the maximization of security.

III. Qualitative Discussion

The objective of this section is to give some intuition about our solution by means of a qualitative description. In order to keep the discussion simple, some hypothesis of ideality will be done. A more precise and analytic description, with the ideality hypothesis removed, is given in Section V.

Fig. 1 shows the proposed PUC. The block marked with VD is a *voltage divider* that, in the ideal case considered here, splits in half the supply voltage V_{DD} so that $V_{DR} = V_{RR} = V_{DD}/2$, independently from the current drawn from its terminals (a more realistic model is introduced later). Transistors Q1 and Q2 are designed to be *nominally* matched (that is, the nominal values of their threshold voltage and transconductance parameters are equal) and in saturation when the capacitor is uncharged. Of course, in a real instance Q1 and Q2 will never be exactly matched and, indeed, our scheme actually *exploits* this unavoidable mismatch.

Remark III.1

It would seem that the proposed scheme is a comparator-based scheme [23], [24] that "amplifies" V_{RR}, relying on the fact that it will always be $V_{RR} \neq V_{DD}/2$. This *would be* correct if Fig. 1 *did not include* capacitor C whose duty is to introduce a feedback that forces at equilibrium $|V_{\text{raw}} - V_{RR}| > V_{\text{T}}$ as soon as the two MOSFETs are not perfectly balanced, differently from comparator-based schemes [23], [24].

The circuit is turned on at $t = 0$ with C uncharged (it will be clear in the following that any initial charge on C has no effect on the equilibrium). After a time $t_{\max}$ the value of V_{raw} is acquired and mapped to "0" or "1". We are interested in understanding qualitatively how V_{raw} is related to the asymmetries of a specific instance of Fig. 1.

Consider first the impossible case where all the components have their nominal values. Since the MOSFETs are matched, $I_{D,1} = I_{D,2}$, no current flows in the capacitor branch, the capacitor remains uncharged and $V_{\text{raw}} = V_{RR}$ forever.

Suppose now that the components have not their nominal value, so that one of the MOSFETs, say Q1, conducts more, that is, $I_{D,1} > I_{D,2}$. It follows that $i_C > 0$ which causes v_C to increase. As long as v_C is smaller than the threshold voltage V_{T} of Q1, both Q1 and Q2 remain in saturation, currents $I_{D,1}$, $I_{D,2}$ and i_C remain constant and C charges linearly with time. However, when $v_C > V_{\text{T}}$, Q1 enters the triode region, I_1 decreases and i_C decreases, too. The system reaches an equilibrium when the v_C is equal to a value $V_{\text{eq}} > V_{\text{T}}$ such that $I_{D,1} = I_{D,2}$. A similar reasoning would show that if Q2 conducts more at the equilibrium, $v_C < -V_{\text{T}}$. Note that in both cases we are *granted* that $|V_{\text{eq}}| > V_{\text{T}}$ since in order to have *equilibrium* one of the two MOSFETs must be in the *triode* region.

We can summarize the results of this qualitative analysis in the following observation.

Observation 1. *The absolute value $|V_{eq}|$ of equilibrium voltage is larger than V_T and the sign of V_{eq} is equal to the* sign *of the difference of the saturation currents $I_{D,i}$, $i = 1, 2$.*

It follows that the region $[-V_{\text{T}}, V_{\text{T}}]$ is *forbidden* for V_{eq} since map $(I_{D,1} - I_{D,2}) \mapsto V_{\text{eq}}$ is *discontinuous*, jumping from $-V_{\text{T}}$ to V_{T} at $I_{D,1} - I_{D,2} = 0$. This is very different from comparator-based schemes [23], [24] that are continuous functions that map small deviations into small output differences.

IV. Preliminary Remarks

Before moving to a more quantitative analysis of Fig. 1, it is useful to summarize few conventions used in the paper.

A. A model for the voltage divider

As said before, the block marked with VD in Fig. 1 represents a generic "voltage divider." In the ideal case, the block VD is characterized by having always $V_{DR} = V_{RR}$ for every value of the currents drawn from its terminals. In practice we expect two deviations from ideality: (i) V_{DR} and V_{RR} will not be equal (but their sum will be always equal to V_{DD}) and (ii) the value of V_{RR} will depend on i_C. In this paper we will use the following very general model for VD (see also Fig. 2):

[1] A stable scheme has $F_r(x) \approx 0$ as soon as x is slightly less than 1.

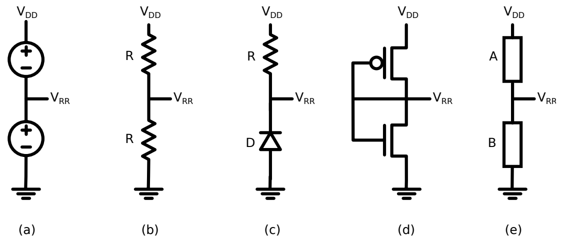


Figure 2. Examples of dividers that agree with the model considered in this paper. (a) An ideal and unbiased divider (b) A resistor-based divider. (c) A zener-based divider. (d) The MOSFET based divider used in the simulations. (e) A generic divider using two dipoles with positive differential resistance.

Hypothesis 1. *With the conventions used in Fig. 1, V_{RR} is a* monotone non-decreasing *function of i_C. The value of V_{RR} when $i_C = 0$ will be written V_{RR}°. If $V_{RR} = V_{RR}^\circ$ for every value of i_C, the divider is* ideal*; if $V_{RR}^\circ = V_{DD}/2$, the divider is* unbiased.

B. Notation

In the following we will use the common convention of using lowercase letters for time-varying values and uppercase letters for constant values. About the MOSFETs, we will use the convention that makes the voltages and currents associated with the MOSFETs always positive [37]. For example, we will consider the gate-source voltage of `Q2` (denoted as $V_{GS,2}$), but the source-gate voltage of `Q1` (denoted as $V_{GS,1}$). Note that with this convention both transistors have a positive threshold voltage and their I-V characteristic can be written as [37]

$$I_{D,i} = \begin{cases} \frac{\beta_i}{2}\left[2(V_{GS,i} - V_{\mathrm{T}i}) - V_{DS,i}\right]V_{DS,i} & V_{DS,i} < V_{DSS} \\ \frac{\beta_i}{2}V_{DSS}^2[1+\lambda(V_{DS,i} - V_{DSS})] & V_{DS,i} \geq V_{DSS} \end{cases} \tag{4}$$

where $V_{DSS} = V_{GS,i} - V_{\mathrm{T}i}$ and λ is the channel length modulation parameter. Finally, W_i, L_i, μ_i and $N_{d,i}$ denote, as usual, the width, the length, the mobility and the nominal doping of `Qi`. Oxide capacity and nominal thickness will be denoted as C_{ox} and t_{ox}. When discussing deviations from nominal values, we will use $\widetilde{N}_{d,i}$ and $\widetilde{t}_{\mathrm{ox}}$ to denote the actual doping levels and oxide thickness.

We will denote the *nominal* values of transconductance parameter and threshold voltage as β and V_{T}, while the corresponding *actual* values for transistor `Qi`, $i = 1,2$, will be denoted as $V_{\mathrm{T}i}$ and β_i. It is convenient to have a notation for $V_\Delta = V_{DD}/2 - V_{\mathrm{T}}$ and for the saturation currents (actual and nominal)

$$I_{s,i} = \frac{\beta_i}{2}(V_{GS,i} - V_{\mathrm{T}i})^2 \qquad i = 1,2 \tag{5a}$$

$$I_s = (\beta/2)(V_{DD}/2 - V_{\mathrm{T}})^2 = \beta V_\Delta^2/2 \tag{5b}$$

Finally, a special role will be played by the difference of the saturation currents

$$I_0 \overset{\text{def}}{=} I_{s,1} - I_{s,2} = \frac{\beta_1}{2}(V_{DR} - V_{\mathrm{T}1})^2 - \frac{\beta_2}{2}(V_{RR} - V_{\mathrm{T}2})^2 \tag{6}$$

that will be shown to represent the "unbalance" of the circuit. Note that V_{T}, β and I_s are design parameters, while $V_{\mathrm{T}i}$, β_i, $I_{s,i}$ and I_0 are r.v., since they depend on construction time variations.

C. Adimensional Equations

It will be convenient to write most of the equations in an adimensional form, obtained by dividing tensions, currents, ... by suitable reference values. The reference values for voltages and currents will be, respectively, the nominal threshold voltage V_{T} and the nominal saturation current $I_s = \beta V_\Delta^2/2$, while the reference value for times will be $\tau = CV_{\mathrm{T}}/I_s$ that can be interpreted as a "time constant" of the circuit. The adimensional version of a variable will be denoted with a line above. For example, the adimensional version of v_C is $\bar{v}_{\mathrm{C}} = v_C/V_{\mathrm{T}}$. We will also use $\theta = 2\lambda V_{\mathrm{T}}$.

Remark IV.1
It is convenient to select some typical parameter values (e.g., W_i, L_i, t_{ox}, ...) in order to have an idea of the order of magnitude of the values involved. Table I shows the values used in this paper. Parameters S_V, S_β and S_X are defined in Section VII.

V. Quantitative Analysis

In this section we do a more quantitative analysis of the proposed circuit. Our first step will be to find an analytic expression for map $v_C \mapsto i_C$. This will be instrumental to find the equilibrium voltage V_{eq} and a suitable value of $t_{\max}$. For the sake of simplicity we will first suppose `VD` ideal, but not necessarily unbiased (see Hypothesis 1). The case of a non-ideal `VD` is analyzed in Section V-C1.

A. Map $v_C \mapsto i_C$

By writing $I_{D,1}$ and $I_{D,2}$ as functions of v_C and taking their difference, it is possible to show by basic algebra that (see Appendix B-A)

$$\bar{i}_{\mathrm{C}}(\bar{v}_{\mathrm{C}}) = \begin{cases} \bar{I}_0 - \operatorname{sgn}\bar{v}_{\mathrm{C}}\left[\theta - \frac{\theta}{2}(1 - |\bar{v}_{\mathrm{C}}|) + \frac{(1-|\bar{v}_{\mathrm{C}}|)^2}{\bar{V}_\Delta^2}\right] & |\bar{v}_{\mathrm{C}}| > 1 \\ \bar{I}_0 - \theta\bar{v}_{\mathrm{C}} & |\bar{v}_{\mathrm{C}}| < 1 \end{cases} \tag{7}$$

Fig. 3 shows some examples of function $\bar{i}_{\mathrm{C}}(\bar{v}_{\mathrm{C}})$. Function $\bar{i}_{\mathrm{C}}(\bar{v}_{\mathrm{C}})$ can be described as an odd function with an offset $\bar{I}_0$. Function $\bar{i}_{\mathrm{C}}$ has three segments: two quadratic ones (for $|\bar{v}_{\mathrm{C}}| > 1$) and a linear one with slope $\theta = 2\lambda V_{\mathrm{T}}$. In the ideal case $\lambda = 0$, the segment is horizontal. Note that all the asymmetries of the circuit are collected inside $\bar{I}_0$ that can be interpreted as a measure of the asymmetry of the circuit.

B. Equilibrium analysis

It is obvious that v_C evolves according to

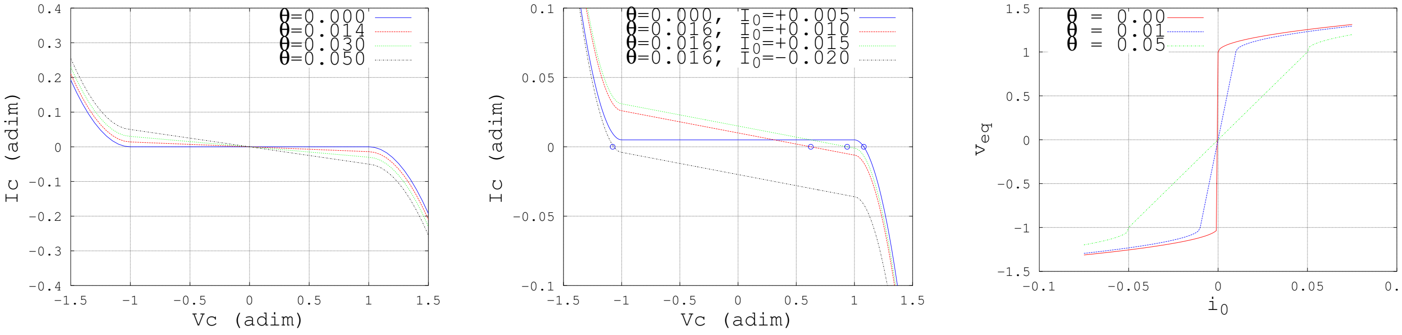
$$C\,\dot{v}_C(t) = i_C(v_C(t)) \qquad t \geq 0 \tag{8}$$

with $v_C(0) = 0$. In (8), as usual, $\dot{v}_C$ denotes the time derivative of v_C. The circuit is at equilibrium if and only if $i_C = 0$. Therefore, in order to find V_{eq} we need to find the zeros of $i_C(v_C)$. A key property of our scheme is that $i_C(v_C)$ has only one zero, so the cell has only one equilibrium point.

Property 1. *Map $i_C(v_C)$ is monotone non-increasing and it is strictly decreasing if and only if $\lambda > 0$. If $\lambda = 0$, $i_C(v_C)$ is constant for $v_C \in [-V_{T2}, V_{T1}]$ and strictly decreasing otherwise.*

Table I
REFERENCE VALUES USED IN THE PAPER AND IN THE SIMULATIONS. MOBILITIES μ_p AND μ_n ARE IN CM^2/V S. PARAMETERS $\overline{V}_\Delta$, θ, ..., S_X ARE ADIMENSIONAL.

V_T	W_1	W_2	$L_1 = L_2$	μ_n	μ_p	t_ox	C	A_V	A_β	V_{DD}	λ
0.7 V	1.15 μm	0.35 μm	0.35 μm	660	220	6.5 nm	5 pF	9 mVμm	2%μm	3 V	0.01 V^{-1}
β	I_s	τ	$\overline{V}_\Delta$	θ	S_V	S_β	S_X				
350 μA/V^2	112 μA	31 ns	1.14	0.014	3.7%	6.6%	9.9%				



(a)

(b)

(c)

Figure 3. (a) Curves $\overline{i}_\mathrm{C}$ vs $\overline{v}_\mathrm{C}$ for several values of θ, $\overline{V}_\Delta = 1.14$ and $\overline{I}_0 = 0$. (b) Curves $\overline{i}_\mathrm{C}$ vs $\overline{v}_\mathrm{C}$ for several values of θ, $\overline{V}_\Delta = 1.14$ and different values of $\overline{I}_0$. The little circles mark the equilibrium values of $\overline{V}_\mathrm{eq}$. Note the large value of $\overline{V}_\mathrm{eq}$ for the case $\theta = 0$, despite it has the smallest $|\overline{I}_0|$. (c) Plots of $\overline{V}_\mathrm{eq}$ vs $\overline{I}_0$ for different values of θ.

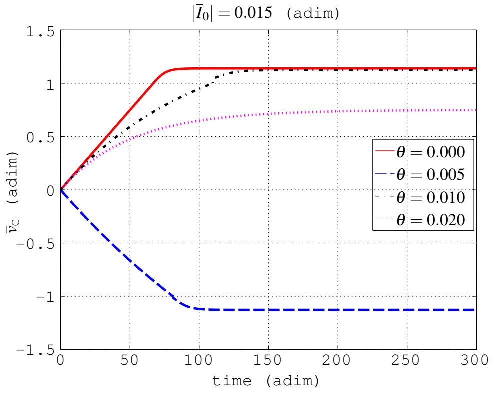


Figure 4. Examples of time evolution for several values of θ, $\overline{V}_\Delta = 1.14$ and $\overline{I}_0 = \pm 0.015$

Proof. It suffices to observe that $i_C = I_{D,1} - I_{D,2}$, and that $I_{D,1}$ ($I_{D,2}$) is monotone decreasing (increasing) with v_C. □

An immediate consequence of monotonicity is the following corollary that, albeit simple, is a cornerstone result.

Corollary 1. *If $I_0 \neq 0$ there is one and only one V_{eq} such that $i_C(V_{eq}) = 0$. Moreover, $i_C{}'(V_{eq}) < 0$, so V_{eq} is a stable equilibrium point.*

Indeed, the fact that our cell has only one equilibrium point grants that every time the cell is turned on *it will always evolve toward the same value* of V_eq, independently on any temporary disturbance such as a turn-on noise at $t = 0$ or some residual charge on C at $t = 0$. This is to be compared with the behavior of the SRAM which, having *two* equilibrium points, can sometimes end in the "wrong" one.

It is possible to write explicitly V_eq as a function of I_0

$$\overline{V}_\mathrm{eq}(\overline{I}_0) = \begin{cases} \overline{I}_0/\theta & |\overline{I}_0| < \theta \\ \operatorname{sgn}\overline{I}_0 \left[(R+1) + \sqrt{R^2 + \overline{V}_\Delta^2 (|\overline{I}_0| - \theta)} \right] & |\overline{I}_0| \geq \theta \end{cases} \tag{9}$$

where $R = \theta \overline{V}_\Delta^2 / 4$. As for (7), the proof is simple but long and given in detail in Appendix B-B. Equation (9) in the special case $\lambda = 0$ becomes

$$\overline{V}_\mathrm{eq}(\overline{I}_0) = \operatorname{sgn}(\overline{I}_0) \left(1 + \overline{V}_\Delta \sqrt{|\overline{I}_0|} \right) \tag{10}$$

A graph of (9) can be found in Fig. 3c for some values of θ and $\overline{V}_\Delta = 0.8/0.7 \approx 1.14$. Note that for $\lambda = 0$ function (9) is discontinuous in $\overline{I}_0 = 0$, as anticipated in Section III. If $\lambda > 0$, function (9) is continuous with a central part which is linear and with a slope proportional to $1/\lambda$. Note that if $\overline{I}_0 > \theta$, $|\overline{V}_\mathrm{eq}| > 1$, that is $|V_\mathrm{eq}| > V_\mathrm{T}$. This suggests the following definition

Definition 1. *An* almost balanced instance *has* $|\overline{I}_0| < \theta$.

C. Time evolution

Now we solve (8) in order to find an expression for $v_C(t)$ that will be used in Section V-D to find t_max. First, rewrite (8) in adimensional form.

Lemma 1. *Let $\tau = CV_T/I_s$ and let*

$$u(\overline{t}) = \overline{v}_C(\tau \overline{t}) \tag{11}$$

be the version of $\overline{v}_C$ with adimensional time. Function u satisfies the following adimensional form of (8) with $u(0) = 0$.

$$\dot{u} = \overline{i}_C(u; \overline{I}_0, \theta, \overline{V}_\Delta) \tag{12}$$

Proof. Bring (8) in adimensional form by dividing by I_s and V_{T} to obtain

$$\begin{cases} \tau\,\dot{\bar{v}}_{\mathrm{C}}(t) = \bar{i}_{\mathrm{C}}(\bar{v}_{\mathrm{C}}(t)) \\ \bar{v}_{\mathrm{C}}(0) = 0 \end{cases} \tag{13}$$

By rewriting (11) for $t = \tau\bar{t}$ and observing that $\dot{u}(\bar{t}) = \tau\,\dot{\bar{v}}_{\mathrm{C}}(\tau\bar{t})$, (13) becomes (12). □

Remark V.1

According to (11), $\bar{v}_{\mathrm{C}}(t)$ can be obtained by time-stretching by a factor τ the solution of (12). In a sense, τ can be interpreted as a "time constant" of the proposed scheme.

Property 2. *Suppose $\lambda > 0$ and define*

$$R = \frac{\theta\bar{V}_\Delta^2}{4} \approx 0 \tag{14a}$$

$$A = \sqrt{(|\bar{I}_0| - \theta)\bar{V}_\Delta^2 + R^2} \approx \bar{V}_\Delta\sqrt{|\bar{I}_0|} \tag{14b}$$

$$B = \frac{A}{\bar{V}_\Delta^2} = \sqrt{\frac{|\bar{I}_0| - \theta}{\bar{V}_\Delta^2} + \frac{\theta^2}{16}} \approx \sqrt{|\bar{I}_0|/\bar{V}_\Delta} \tag{14c}$$

$$C = \tanh^{-1}\left(\frac{R}{A}\right) = \tanh^{-1}\left(\frac{\theta}{4B}\right) \approx 0 \tag{14d}$$

$$U = \frac{1}{\theta}\left(\bar{I}_0 + \frac{\theta(\bar{I}_0 - \bar{\Delta}_v)}{2}\right) \approx \frac{\bar{I}_0}{\theta} \tag{14e}$$

$$\bar{t}_{sw} = \begin{cases} -\frac{1}{\theta}\ln\left(1 - \frac{\theta}{|\bar{I}_0|}\right) \approx \frac{1}{|\bar{I}_0|} & |\bar{I}_0| > \theta \\ \infty & |\bar{I}_0| \leq \theta \end{cases} \tag{14f}$$

where the approximations are valid when $\lambda \to 0$. Let also $w : \mathbb{R}_+ \to \mathbb{R}$ be defined as

$$w(t;\bar{I}_0,\theta,\bar{V}_\Delta) \overset{def}{=} A\tanh(Bt + C) + R \qquad t \geq 0 \tag{15}$$

With the notation above, the solution of (12) can be written as

$$u(\bar{t}) = \begin{cases} U[1 - \exp(-\theta\bar{t})] & 0 \leq \bar{t} < \bar{t}_{sw} \\ \operatorname{sgn}\bar{I}_0\left[w(\bar{t} - \bar{t}_{sw};|\bar{I}_0|,\theta,\bar{V}_\Delta) + 1\right] & \bar{t}_{sw} \leq \bar{t} \end{cases} \tag{16}$$

The proof is just basic algebra, but very long and it is given in detail in Appendix B-C. By taking the limit $\lambda \to 0$ one obtains the evolution in the ideal case

$$u(\bar{t}) = \begin{cases} \bar{I}_0\bar{t} & 0 \leq \bar{t} < \bar{t}_{\mathrm{sw}} = 1/|\bar{I}_0| \\ \bar{V}_\Delta\sqrt{\bar{I}_0}\tanh\left(\frac{\sqrt{\bar{I}_0}}{g}(\bar{t} - \bar{t}_{\mathrm{sw}})\right) & 1/|\bar{I}_0| \leq \bar{t} \end{cases}$$

Fig. 4 shows few examples of time evolution for several values of θ, $\bar{V}_\Delta = 1.14$ and $\bar{I}_0 = \pm 0.015$.

1) The case of a non-ideal divider: The main results of the non-ideal voltage divider is contained in this Property.

Property 3. *Let $i_C(v_C)$ be the current on C when an ideal divider with open voltage V_{RR}° is employed, let V_{eq} be the corresponding equilibrium voltage and let $\hat{\imath}_C(v_C)$ be the current on C when a non-ideal divider, with the same V_{RR}°, is employed. Moreover, define $R_{max} = \sup dV_{RR}/dI \geq 0$ and $\Gamma = \sup(-\partial i_C/\partial V_{RR})$.*

Thesis: *Function $\hat{\imath}_C(v_C)$ is monotone non-increasing it has the same zero as i_C, that is,*

$$\hat{\imath}_C(V_{eq}) = i_C(V_{eq}) = 0 \tag{17}$$

and, moreover, the following inequalities hold

$$\frac{1}{1+\Gamma}i_C(x) \leq \hat{\imath}_C(x) \leq i_C(x) \qquad x \leq V_{eq} \tag{18a}$$

$$\frac{1}{1+\Gamma}i_C(x) \geq \hat{\imath}_C(x) \geq i_C(x) \qquad x > V_{eq} \tag{18b}$$

Property 3 can be informally summarized by saying that a non-ideal divider does not change the equilibrium point (because of (17)), but it increases the time to reach it (because of (18b) that shows that i_C is smaller in the non-ideal case). The proof is elementary and involves some standard inequality arguments. See Appendix B-D for the details.

D. Transient length

The criterion for choosing $\bar{t}_{\max}$ is that a large fraction of the instances is very close to the equilibrium value at $\bar{t}_{\max}$. More precisely, denote with $t_{\varepsilon,\bar{I}_0}$ the time required for a circuit to reach the equilibrium value within ε, that is

$$\bar{v}_{\mathrm{C}}(t_{\varepsilon,\bar{I}_0}) = (1-\varepsilon)\bar{V}_{\mathrm{eq}}(\bar{I}_0) \tag{19}$$

Fix $\varepsilon, \eta \in (0,1)$ and search for $\bar{t}_{\max}$ such that

$$\mathrm{P}[t_{\varepsilon,|\bar{I}_0|} > \bar{t}_{\max}] \leq \eta \tag{20}$$

It can be easily shown (see details in Appendix B-E) that

$$\bar{t}_{\max} \geq t_\varepsilon\left(\bar{\sigma}_I\Phi^{-1}\left(\frac{1+\eta}{2}\right)\right) \approx t_\varepsilon\left(\sqrt{\frac{\pi}{2}}\,\bar{\sigma}_I\,\eta\right) \tag{21}$$

where

$$t_\varepsilon(\bar{I}_0) = \begin{cases} -\frac{\ln\varepsilon}{\theta} & \text{if } |\bar{I}_0| < \theta \\ \bar{V}_\Delta\frac{\tanh^{-1}\left(-\varepsilon + \frac{1-\varepsilon}{\bar{V}_\Delta\sqrt{\bar{I}_0}}\right)}{\sqrt{\bar{I}_0}} + \frac{1}{\bar{I}_0} & \text{if } |\bar{I}_0| \geq \theta \end{cases} \tag{22}$$

It turns out that typically $\bar{t}_{\max}$ is approximately few hundreds (see also Fig. 4) that with the values in Table I corresponds to $t_{\max} \approx 10\ \mu$s.

E. Energy consumption

The proposed circuit at steady state consumes a current approximately equal to I_s. In order to minimize power consumption, the cell is powered only for $t_{\max}$ seconds, successively the outcome is copied to an SRAM cell[2] (for example) and the cell turned off. The energy required is

$$E = t_{\max}V_{DD}\,I_s = \bar{t}_{\max}\tau V_{DD}\,I_s = \bar{t}_{\max}CV_{\mathrm{T}}V_{DD} \tag{23}$$

Since I_s is of the order of μA and $t_{\max}$ is of the order of tens of μs, E is a fraction of nJ/bit.

[2] Of course, it is advisable that the SRAM cell is on the same chip of the PUC, in order to avoid the obvious security issues related to the transfer of the outcome to an external SRAM.

VI. Noise analysis

So far we supposed the system noiseless. In practice, however, the currents of the two MOSFETs will be affected by noise that will be integrated by the capacitance `C` and this will affect the value of V_{raw} acquired at $t_{\max}$. Note that any noise of temporary nature (such as a turn-on noise at $t = 0$) has no effect on the outcome, due to the fact that the circuit evolves toward the unique equilibrium point. Therefore, the only noise that is necessary to consider is the combination of pink and white noise affecting the MOSFET currents.

At room temperature the Power Spectrum Density (PSD) of the noise of `Qi` can be modeled as [38]

$$R_i(f) = K_1 \frac{I_i}{C_{\text{ox}} L_i^2} \frac{1}{f} + K_0\, g_m (1 + g_{mbs}/g_m) = \frac{a_{i,1}^2}{f} + a_{i,0} \tag{24}$$

where $K_0 = 0.0094$ eV, $K_1 = 10^{-28}$ F·A. Since the noises of `Q1` and `Q2` are independent, the PSD of the noise on i_C is

$$R_C(f) = \frac{a_{1,1}^2 + a_{2,1}^2}{f} + (a_{1,0} + a_{2,0}) = \frac{a_1^2}{f} + a_0 \tag{25}$$

A useful parameter is the frequency $f_K = a_1^2/a_0$ where the two components of (25) (pink and white) are equal. With the values in Table I, $a_0 \approx 4.6 \cdot 10^{-25}\,\text{A}^2/\text{Hz}$, $a_1^2 \approx 1.7 \cdot 10^{-17}\,\text{A}^2$ and $f_K \approx 36.5$ MHz.

Since the analysis is complicated by the fact that the differential equation (8) is non linear, we consider the approximate problem of determining the variance σ_ξ^2 of the voltage ξ across a capacitor with capacity C charged, during a time $t_{\max}$, by a current with PSD (24). In Appendix C-A it is shown by elementary means (taking into account the peculiarities of pink noise) that

$$\sigma_\xi^2 \approx 0.045 \cdot a_1^2 \frac{t_{\max}^2}{C^2} + a_0 \frac{t_{\max}}{C^2} \approx 0.045 \cdot a_1^2 \frac{t_{\max}^2}{C^2} \tag{26}$$

where the last approximation is valid when $t_{\max} \gg 1/f_K = a_0/a_1^2$, that is, as soon as $t_{\max}$ is more than few microseconds. From $t_{\max} = \tau \bar{t}_{\max}$ one deduces

$$\overline{\sigma}_\xi = 0.212 \cdot \bar{t}_{\max} \overline{a}_1 \tag{27}$$

where $\overline{a}_1 = a_1/I_s$. Note that $\overline{\sigma}_\xi$ does not depend on C.

VII. Quality figures

The objective of this section is to predict the quality indexes of the proposed cell, namely, its *unbiasedness*, its *stability* and its *independence* and derive from them a prediction for μ_{intra} and μ_{inter}. We will achieve this by first determining the statistical distributions of $\overline{I}_0$ and $\overline{V}_{\text{eq}}$. Symmetry properties of the distribution of $\overline{I}_0$ will allow us to show unbiasedness, while stability will be obtained by using the distribution of $\overline{V}_{\text{eq}}$ together with result (27) of Section VI. Finally, at the end of this section we will discuss briefly the *local biasing* effect that can be induced by process gradients.

We will write $X \sim \mathcal{N}(m, \sigma^2)$ when X is normally distributed with mean m and variance σ^2 and $\phi_{m,\sigma}$ will denote the corresponding density. For notational convenience we will write $\phi(x)$ in place of $\phi_{0,1}(x)$.

A. Statistical model

The parameters that can be modeled as r.v. are $V_{\text{T}i}$, β_i and V_{RR}. According to the literature [41]

$$V_{\text{T}i} \sim \mathcal{N}(V_{\text{T}}, \sigma_V^2) \qquad ; \qquad \beta_i \sim \mathcal{N}(\beta, \sigma_\beta^2) \tag{28}$$

where $\sigma_V^2 = A_V^2/(L_i W_i)$ and $\sigma_\beta^2 = \beta^2 A_\beta^2/(L_i W_i)$. Although $V_{\text{T}i}$ and β_i depend on the same physical parameters, it has been seen in practice that they can be considered independent [41]. We will need the zero-mean versions of β_i, $V_{\text{T}i}$ and V_{RR} as

$$\delta_{\beta,i} \overset{\text{def}}{=} \beta_i - \beta \;;\; \delta_{V,i} \overset{\text{def}}{=} V_{\text{T}i} - V_{\text{T}} \;;\; \delta_R \overset{\text{def}}{=} V_{RR} - V_{DD}/2 \tag{29}$$

the adimensional values

$$\begin{aligned} S_\beta^2 &= \frac{\sigma_{\beta,p}^2 + \sigma_{\beta,n}^2}{\beta^2} = A_\beta^2 \left(\frac{1}{L_1 W_1} + \frac{1}{L_2 W_2} \right) \\ S_V^2 &= \frac{\sigma_{V,p}^2 + \sigma_{V,n}^2}{V_\Delta^2} = \frac{A_V^2}{V_\Delta^2} \left(\frac{1}{L_1 W_1} + \frac{1}{L_2 W_2} \right) \end{aligned} \tag{30}$$

and their combination $S_X^2 = S_\beta^2 + 4 S_V^2$. With the values given in Table I, S_β, S_V and S_X are approximately 6%, 3% and 10%. Finally, it is reasonable to assume V_{RR} independent from both $V_{\text{T}i}$ and β_i and $V_{RR} \sim \mathcal{N}(V_{DD}/2, \sigma_R^2)$, for some σ_R.

B. Probability density of $\overline{I}_0$ and $\overline{V}_{eq}$

The key result is the following property.

Property 4. *Let f_{δ_R} and f_{I_0} be the probability density functions of, respectively, δ_R and I_0. The following claims hold (1) If f_{δ_R} is even, then f_{I_0} is even as well and (2) if $\delta_R \sim \mathcal{N}(0, \sigma_R^2)$, then*

$$\overline{I}_0 \sim \mathcal{N}\left(0,\, S_X^2 + 16(\overline{\sigma}_R/\overline{V}_\Delta)^2\right) \tag{31}$$

with $\sigma_I^2 = {I_s}^2 S_X^2 + 4\beta^2 V_\Delta^2 \sigma_R^2 = {I_s}^2 (S_X^2 + 16\sigma_R^2/V_\Delta^2)$.

The proof involve some standard approximation and conditioning. The details can be found in Apendix D, Proof D.1.

From the pdf (31) of I_0 one obtains the pdf of $\overline{V}_{\text{eq}}$ as

$$f_{eq}(v) = h'(v)\, \phi_{0,\overline{\sigma}_I^2}(h(v)) = \frac{1}{\overline{\sigma}_I} h'(v)\, \phi(h(v/\overline{\sigma}_I)) \tag{32}$$

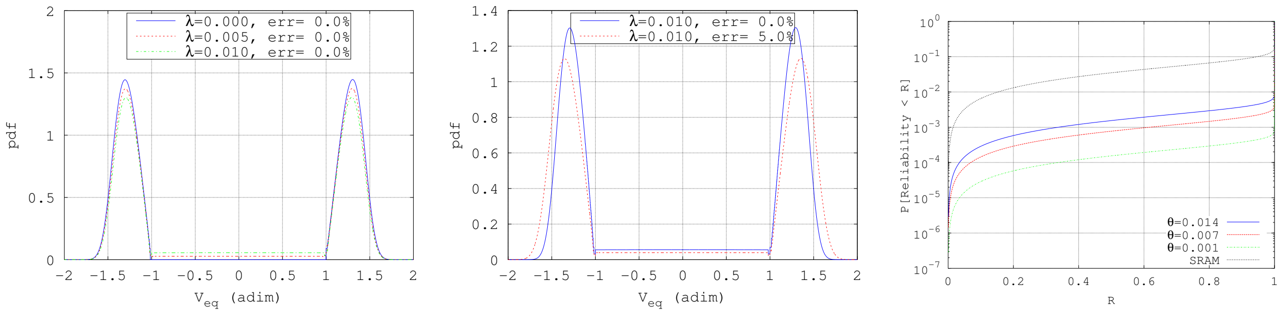
where h is the inverse of (9) and h' its derivative, namely

$$h(\overline{V}_{\text{eq}}) = \begin{cases} \theta \overline{V}_{\text{eq}} & |\overline{V}_{\text{eq}}| < 1 \\ \operatorname{sgn} \overline{V}_{\text{eq}} \frac{(|\overline{V}_{\text{eq}}| - R - 1)^2 - R^2 + 4R}{\overline{V}_\Delta^2} & |\overline{V}_{\text{eq}}| \geq 1 \end{cases} \tag{33a}$$

$$h'(\overline{V}_{\text{eq}}) = \begin{cases} \theta & |\overline{V}_{\text{eq}}| < 1 \\ 2 \frac{|\overline{V}_{\text{eq}}| - 1 - R}{\overline{V}_\Delta^2} & |\overline{V}_{\text{eq}}| \geq 1 \end{cases} \tag{33b}$$

Note that if $\lambda = 0$, $h'(v) = 0$ if $|v| < 1$ and this implies, via (32), $f_{eq}(v) = 0$ for $|v| < 1$, coherently with the fact that if $\lambda = 0$ map $\overline{I}_0 \mapsto \overline{V}_{\text{eq}}$ is discontinuous and the anti-image of $(-1, 1)$ is empty. Fig. 5 shows some examples of f_{eq}, for different values of λ and σ_R. The following result is obvious

Property 5. *The probability of having an* almost balanced *instance (see Definition 1) is equal to* $\operatorname{erf}(\theta/(\sqrt{2}\overline{\sigma}_I))$.



(a) (b) (c)

Figure 5. Some examples of probability density f_{eq} of $\overline{V}_{\text{eq}}$ and RDF. (a) $\sigma_R = 0$ and different values of λ. (b) $\lambda = 0.010$ and different values of σ_R (c) Comparison of the RDF for several values of θ and the RDF of the SRAM

C. Performance measures

From the knowledge of f_{eq} is possible to predict the stability of the proposed scheme. Let $\overline{\xi}$ be the noise affecting the output $\overline{V}_{\text{raw}} \approx \overline{V}_{\text{eq}}$ of the cell at $t = t_{\max}$, let F_ξ be the distribution of $\overline{\xi}$ and suppose that the density of $\overline{\xi}$ is even. (From Section VI we know that $\overline{\xi} \sim \mathcal{N}(0, \sigma_\xi^2)$, so that $F_\xi(x) = \Phi(x/\sigma_\xi)$, but we will not need this). Since the PUC outcome is “1” when $\overline{V}_{\text{eq}} + \overline{\xi} > 0$ one can write

$$p_1 = F_\xi(\overline{V}_{\text{eq}}) = F_\xi(\overline{V}_{\text{eq}}(\overline{I}_0)) \tag{34}$$

By using (34) we can derive the quality measures for our cell.

1) SDF and μ_{intra}: In order to derive the SDF, observe that

$$R(\overline{V}_{\text{eq}}) = 2F_\xi(|\overline{V}_{\text{eq}}|) - 1 \tag{35}$$

Define, for notational convenience, $W_x = R^{-1}(x) = F_\xi^{-1}((1+x)/2)$. By using (32) and the fact that h is odd it is easy to show that the SDF is (see Appendix D-A for details)

$$F_r(x) = 2F_I\left(\frac{h(W_x)}{\overline{\sigma}_I}\right) - 1 = 2\Phi\left(\frac{h\left(F_\xi^{-1}\left(\frac{1+x}{2}\right)\right)}{\overline{\sigma}_I}\right) - 1 \tag{36}$$

where F_I is the distribution of $\overline{I}_0$ and where we used the fact that $\overline{I}_0 \sim \mathcal{N}(0, \overline{\sigma}_I^2)$.

Example VII.1

Fig. 5c shows few examples of SDF (36) for several values of θ, together with the SDF of the SRAM, according to [6]. Note that the SDFs for the proposed scheme are always below the SDF of the SRAM and this means that unreliable cells are less probable and that reliability improves as θ decreases.

By using (35) in (3a) one can obtain μ_{intra} of the proposed PUC. An observation that helps in estimating μ_{intra} is that in Fig. 5c the SDF curves for the proposed PUC can be obtained by lowering the SRAM curve. Since Fig. 5c is in logarithmic scale, this implies that the SDF for our solution can be approximately[3] obtained by multiplying the SDF of the SRAM by a constant $\alpha < 1$. This, together with (3a), implies $\mu_{\text{intra}}^{\text{ours}} = \alpha\mu_{\text{intra}}^{\text{SRAM}}$. According to Fig. 5c, our solution is 10 to 100 times better than the SRAM (result confirmed by simulations in Section IX).

2) Inter distance μ_{inter}: The following property (proved in Appendix D, Proof D.2) is instrumental to determine μ_{inter}.

Property 6. *If the pdfs of noise $\overline{\xi}$ and $\overline{I}_0$ are even, then the pdf f_p of p_1 is symmetric around $1/2$.*

From the symmetry of f_p one predicts $P[O_{\text{pref}} = 1] = 1/2$, $m_p = 1/2$ and (via (3b)) $\mu_{\text{inter}} = 1/2$. Therefore, our PUC is predicted to be perfectly unbiased in both senses of Remark II.1.

D. Independence

Remember that a scheme is independent if the POs of different instances are independent. In the proposed scheme, the PO is a function of I_0 which depends on $V_{\text{T}i}$ and β_i, $i = 1, 2$, that in turn depend on $N_{d,i}$ and t_{ox}. It is commonly accepted that variations in $N_{d,i}$ can be modeled as a consequence of the 2D Poisson process associated with doping [41]. Since there is no overlap between the areas of different transistors, the r.v. $N_{d,i}$ are independent. About oxide thickness t_{ox}, according to [42], it can be considered uncorrelated after few nanometers of distance. Therefore we can predict that the PO associated with different cells will be independent.

E. Process Gradients and Local Biasing

The results above show that the proposed scheme is unbiased in the sense that if one selects at random a cell from a pool of cells, the probability of selecting a cell with PO “1” is $1/2$. However, some mechanisms (e.g., gradients in dopant density) together with the intrinsic asymmetry of the cell could induce some *local biasing*, i.e., it may happen that the PO of the cells implemented in a specific area are biased toward “1” or “0” [30].[4] Local biasing is quite a general issue and every PUC scheme can be expected to be subject to it. Its impact can be studied in a general setting [30]. In the specific case of our scheme, a local biasing will cause $\overline{I}_0$ to have a non-null average $\overline{m}_I \overset{\text{def}}{=} \mathbb{E}\left[\overline{I}_0\right] \neq 0$. It is trivial to show that the corresponding bias is $p_1 - 1/2 = \Phi(\overline{m}_I/\overline{\sigma}_I) - 1/2$. It follows that a large $\overline{\sigma}_I$ counteracts the impact of local bias. The actual impact of local biasing is very dependent on process details and it can be measured only by experimental means.

[3] It cannot be exact since any SDF is equal to 1 in $R = 1$.

[4] *Local biasing* should not be confused with *statistical dependence*.

VIII. Other Considerations

A. Temperature Dependence: Crossover Temperature

A key characteristic that a PUC must have is the stability of its behavior with respect to changes in the temperature. On a qualitative level, since both MOSFETs are affected in the same way by changes in the temperature, we expect that if, say, $I_0 > 0$ at room temperature T_0, then I_0 will maintain the same sign also at other temperatures. Against this qualitative reasoning one could object that even if the effect of temperature change on both MOSFETs is *qualitatively* the same (e.g., both I_s increase), it could be that one transistor changes more and it "catches up" with the other. A more quantitative reasoning is, therefore, required.

Observe that in our case we are not interested in the actual value of I_0, but only in its sign. Therefore, the "critical" case that we would like to avoid is that I_0 at some temperature T has a sign different from the sign that it has at T_0. If this happens, we know that there will be a *crossover temperature* T_X between T_0 and T where $I_0(T_X) = 0$. Note that in the neighborhood of T_X, I_0 will be close to zero, so that the cell becomes unstable when the temperature is near T_X.

It turns out that the validity range of analytic results obtained by using simple temperature dependency models [37] is not sufficient to cover a wide range of temperatures. Therefore, we decided to study the behavior of T_X by means of simulations. Results and details are given in Section IX-A with other simulation results. Here we can anticipate that it turns out that T_X is a function almost deterministic of $\overline{I}_0$ and that most temperature-sensitive cells are those that are *almost balanced* (Definition 1). This allows us to recognize and disable the temperature-sensitive instances. See Section VIII-C3 for a detailed discussion about this.

B. Power supply variations

An advantage of our scheme is that non-*almost balanced* cells are insensitive to power supply variations. Indeed, if a cell is not almost balanced, its equilibrium value V_{eq} is larger in absolute value than V_{T} and V_{T} is a characteristic of the MOSFET, independent on the power supply. This reasoning is confirmed by the simulations described in Section IX-B.

C. Design Guidelines

1) Transistor size: A key requirement in the design is keeping λ small, in order to have a small probability of almost balanced instances (Property 5). In order to have λ small one must use large values of L. This not only increases the size of the cell, but also reduces the variability of V_{T} and β.

In order to reduce the area one can reduce W. This reduces I_0, making the charging of C slower. This can be compensated with a smaller C and/or a larger $t_{\max}$. Reducing I_0 also increases the noise, according to (27), since $\overline{a}_1 = a_1/I_s$, but the impact of the increased noise turns out to be negligible. Simulations show that one can reduce the area of the cell down to[5] 1 μm^2 by using a "long and thin" transistor without reducing the performance, taking into account the noise too.

Table II
Dimensions (in μm) of the MOSFET used in the simulations.

Label	L	W_p	W_n	Label	L	W_p	W_n
0.5/0.5	0.5	0.3	0.1	0.5/20	20.0	1.1	0.5
0.5/1	1.0	0.7	0.3	1/20	20.0	2.2	1.0
1.5/5	5.0	3.3	1.5	3/20	20.0	6.2	2.8
1.5/10	10.0	3.3	1.5	7/50	50.0	16.5	6.9

Table III
Predicted μ_{intra} and σ_{intra} for the proposed PUC (128-bit PUCs). $\mu_{\text{inter}} \approx 50\%$ for all the schemes.

PUC	μ_{intra}	PUC	μ_{intra}	PUC	μ_{intra}
SRAM [6]	12%	Butterfly [19]	6%	Latch [20]	3%
0.5/0.5	1.2%	0.5/1	0.4%	1.5/5	0.05%

Note that although the proposed cell could be larger than other PUC cells, the stability of our proposal is such that one does not need costly (in terms of area) error-correction circuits.

2) Load effect and the choice of C: According to the analysis above and the results of Section VI, the value of C impacts only τ and the energy consumption. This suggests to choose C as small as possible, but large enough to make the effect of any load negligible.

3) Handling almost balanced cells: It turns out that $\overline{I}_0$ can be consider a "quality measure" of a specific instance. Almost balanced cells (that is, with $\overline{I}_0$ smaller than a threshold) are "bad" in many senses: their smaller $\overline{V}_{\text{eq}}$ (see (9)) makes them more sensitive to noise, they are more sensitive to temperature (Section IX-A) and more sensitive to aging (Section IX-C). This suggests to measure $|\overline{I}_0|$ (or, equivalently, $\overline{V}_{\text{eq}}$) at enrollment phase and "disable" those cells with small $|\overline{I}_0|$. This procedure requires, of course, a surplus of cells that grows when the probability of having an almost balanced cell grows.

IX. Simulation results

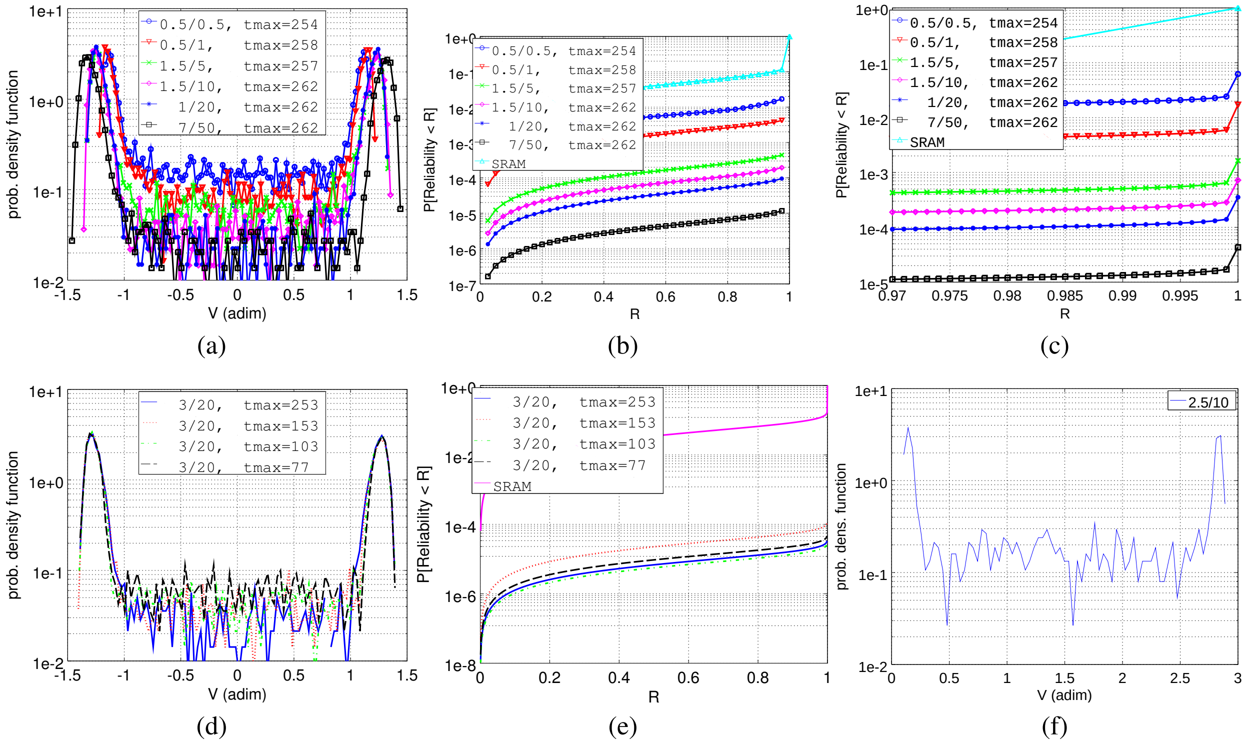
We run several simulations in order to verify that the results predicted by the analysis above – done using simple models suitable for theoretical analysis – still hold when the circuit is simulated using more complex and realistic models. More precisely, we verified the results about stability (SDF and μ_{intra}) and unbiasedness (μ_{inter}), and what happens when temperature and power supply change. We also take into account the effects of aging.

A. Temperature Dependence

We simulated the proposed circuit by using transistors of different sizes varying, for every size, doping $\widetilde{N}_d$ and $\widetilde{t}_{\text{ox}}$ around their nominal values. More precisely, for t_{ox} we scanned the interval $\pm\sigma_{\text{ox}} = \pm 2$Å, while for $\widetilde{N}_d$ we scanned the interval $\pm 3\sigma_{\text{N}}$, where $\sigma_{\text{N}} = \sqrt{N_d/(WL\,d)}$ where d is the doping depth and N_d is the nominal doping (see Remark IX.1 in Section IX-D for an explanation of σ_{ox} and σ_{N}). For every choice of t_{ox} and N_d, the circuit was simulated at temperatures ranging from -150°C to 100°C in order to find T_X. We repeated the simulations with supply voltages 3.5 V and 5 V in order to increase $\overline{V}_\Delta$ without changing the transistors.

The results can be seen in Fig. 7 that shows the scatter plot of T_X vs $\overline{I}_0$ (shown as percentage of I_s) at $T_0 = 300$ K.

[5] Maybe the area can be further reduced, but with smaller transistor the BSIM3 model used in the simulations becomes unreliable.



(a) (b) (c)

(d) (e) (f)

Figure 6. Results for different cell sizes and $\bar{t}_{\max} \approx 260$. (a) Density probability function of $\overline{V}_{\rm eq}$ (b) RDFs compared with the RDF of the SRAM, (c) is (b) with the x-axis expanded. (d) and (e) are like (a) and (b), but for different $\bar{t}_{\max}$ and fixed cell size. (f) Example of behavior of a cell scaled down at $V_{DD} = 1.5V$.

Table IV

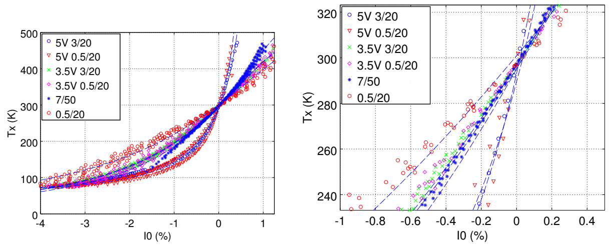
Total number of bit required and relative complexity of FEC implementation for an N-bit PUC and error probability $\leq \eta$. The area required for power control is $\approx 10\%$ the area required by the PUC.

		SRAM		Butterfly		Latch		0.5/0.5		0.5/1		1.5/5		1.5/10		1.5/20		7/50	
η	N	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost	Bits	Cost
10^{-2}	64	184	30.0	116	13.0	92	7.0	80	4.0	72	2.0	68	1.0	**64**	**0**	**64**	**0**	**64**	**0**
	128	324	49.0	208	20.0	172	11.0	148	5.0	140	3.0	132	1.0	132	1.0	**128**	**0**	**128**	**0**
	256	588	83.0	388	33.0	324	17.0	288	8.0	276	5.0	260	1.0	260	1.0	260	1.0	**256**	**0**
10^{-4}	64	216	38.0	128	16.0	100	9.0	84	5.0	76	3.0	68	1.0	68	1.0	68	1.0	**64**	**0**
	128	364	29.5	224	12.0	180	6.5	156	3.5	144	2.0	136	1.0	132	0.5	132	0.5	**128**	**0**
	256	636	95.0	412	39.0	336	20.0	296	10.0	280	6.0	264	2.0	260	1.0	260	1.0	260	1.0
10^{-6}	64	304	60.0	168	26.0	124	15.0	100	9.0	88	6.0	76	3.0	72	2.0	72	2.0	68	1.0
	128	464	84.0	272	36.0	208	20.0	176	12.0	156	7.0	140	3.0	136	2.0	136	2.0	132	1.0
	256	760	126.0	468	53.0	372	29.0	320	16.0	296	10.0	272	4.0	268	3.0	264	2.0	260	1.0

Table V

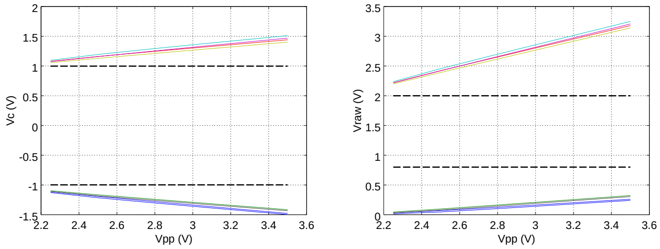
Number of iterations required for a PUC (η, δ)-stable [9]

$1-\eta$	δ	SRAM	Butterfly	Latch	0.5/0.5	0.5/1	1.5/5	1.5/10	1.5/20	7/50
10^{-2}	$5 \cdot 10^{-3}$	536	130	29	**1**	**1**	**1**	**1**	**1**	**1**
10^{-4}	$5 \cdot 10^{-4}$	138 298	34 564	8 631	1 370	208	**1**	**1**	**1**	**1**
10^{-4}	10^{-4}	3 457 758	864 429	216 097	34 564	5 519	42	**1**	**1**	**1**



(a) (b)

Figure 7. (a) Scatter plots of T_X vs $\overline{I}_0$. (b) Zoom of (a).



(a) (b)

Figure 8. (a) V_{eq} vs. supply voltage for differently unbalanced PUCs. (b) V_{raw} vs. supply voltage.

The curves are labeled with the *plot labels* in Table II. It is interesting to observe that T_X is almost a function of $\overline{I}_0$ in the sense that for every $\overline{I}_0$ there is a limited range of possible T_X, especially for small $\overline{I}_0$. Observe that $|\overline{I}_0|$ can be used as a measure of quality of the cell also from the viewpoint of insensitivity to temperature variations (see Section VIII-C3).

Of special interest is the slope of the curves in Fig. 7 at $\overline{I}_0 = 0$, since a large slope implies that T_X is far from T_0 even for small $|\overline{I}_0|$. It is clear that the slope increases with $\overline{V}_\Delta$.

B. Power supply variations

In order to verify the analysis of Section VIII-B, we simulated the circuit behavior for several values of V_{DD} and measured V_{eq} after 1 ms of simulated evolution. The results are shown in Fig. 8a. The different slanted lines correspond to differently unbalanced instances of the circuit, while the two dashed horizontal lines show the values $\pm V_{\mathrm{T}}$ (in this case $V_{\mathrm{T}} \approx 1\mathrm{V}$). It is clear that $|V_{\mathrm{eq}}| \geq V_{\mathrm{T}}$ for every supply voltage, as predicted. Therefore, our scheme preserves its reliability also for different supply voltages.

It is also of interest to check that V_{raw} remains outside the transition region of a possible logic gate fed by the PUC. This can be seen in Fig. 8b that shows the behavior of V_{raw} when the supply voltage is changed. The horizontal dashed lines show the typical guaranteed values of $V_{IL} = 0.8$ V and $V_{IH} = 2$ V for a 3V-CMOS inverter. It is clear that V_{raw} never enters the transition region.

C. Aging

Aging is another important issue since it could happen that with time the cell will change its PO [27], [28]. We carried out some aging simulations with Relxpert. We generated several cells of different unbalance and determined $\overline{V}_{\mathrm{eq}}$ at the first turn-on and after 10 years of usage (with 50 turn-on/day). Fig. 9 shows $\overline{V}_{\mathrm{eq}}$ of the "fresh" device vs the $\overline{V}_{\mathrm{eq}}$ after aging. Note that the PO of a cell changes when the corresponding point in Fig. 9 is in the second or in the fourth quadrant (hatched). It is clear the PO changes only if the cell is very balanced, that is, its $|\overline{I}_0|$ is small. See Section VIII-C3 for a discussion about handling almost balanced cells.

D. Stability, Unbiasedness, Consumption

We simulated the scheme of Fig. 1b with the voltage divider of Fig. 2d. Both NMOS have size $W_n \times L$ and both PMOS have size $W_p \times L$. W_n, W_p and L can be found in Table II together with the labels used in the plots. We used `ngspice-26` (based on `spice3f5`) with model BSIM3 nominal values $t_{\mathrm{ox}} = 9.3$ nm and $N_d = 1.7 \cdot 10^{17}$ cm^{-3}.

For every MOSFET in the scheme, we randomly changed the $N_{d,i}$ and t_{ox} according to $\widetilde{N}_{d,i} \sim \mathcal{N}(N_{d,i}, N_{d,i}/(d_i W_i L_i))$, and $\widetilde{t}_{\mathrm{ox}} \sim \mathcal{N}(t_{\mathrm{ox}}, 0.04\ \mathrm{nm}^2)$. In this way we account also for the variations of `VD`.

Remark IX.1

Doping pdf $\mathcal{N}(N_{d,i}, N_{d,i}/(d_i W_i L_i))$ has been chosen by observing that the number M of doping atoms in a $d_i \times W_i \times L_i$ cube is a Poisson r.v. with parameter $\Lambda = N_{d,i}\ d_i\ W_i\ L_i$. By approximating the distribution of M with $\mathcal{N}(\Lambda, \Lambda)$ one deduces that the dopant concentration is a r.v. $\mathcal{N}(N_{d,i}, N_{d,i}/(d_i W_i L_i))$. About the the oxide thickness, we did the pessimistic[6] choice $\sigma_{\mathrm{ox}} = 0.2\ \mathrm{nm} = 2$ Å ($\approx$ the diameter of a silicon atom).

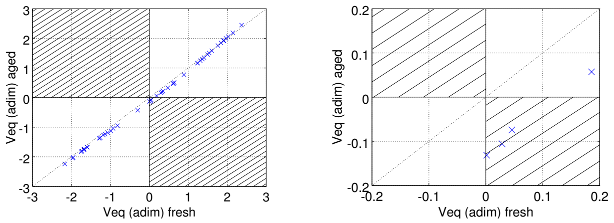
Fig. 6 shows the results of some simulations done with different cells and[7] $\overline{t}_{\max} \approx 260$. Fig. 6a shows in logarithmic scale the pdf of $\overline{V}_{\mathrm{eq}}$, while Fig. 6b and Fig. 6c show the corresponding SDFs and compares them with the SDF of the SRAM PUC. Fig. 6d and e are similar to Fig. 6a and b, but with fixed size and variable $\overline{t}_{\max}$. The energy required ranges from 1 to 2 nJ/cell (for $\overline{t}_{\max} = 100$).

1) Discussion: It is clear from the plots that the pdf of the adimensional voltage $\overline{V}_{\mathrm{eq}}$ is almost zero in $[-1, 1]$ (i.e., $[-V_{\mathrm{T}}, V_{\mathrm{T}}]$ in dimensional units), confirming the prediction that $|V_{\mathrm{eq}}| > V_{\mathrm{T}}$, even for small unbalances. Note also that the density relative to smaller transistors assumes larger values in $[-1, 1]$, probably because to a larger λ makes more probable to have almost balanced cells (Section V-B). Moreover, Fig. 6c and 6d show that increasing $\overline{t}_{\max}$ lowers the pdf around $\overline{V}_{\mathrm{eq}} = 0$, as expected, since when $\overline{t}_{\max}$ is larger more cells reach the equilibrium.

From Fig. 6 it is clear that stability improves with the cell size. However, even the smallest cell is still an order of magnitude more stable than the SRAM (see Table III and Section IX-D2). Fig. 6f shows an example of the behavior of a cell scaled for $V_{DD} = 1.5$ V. Note that the adimensional plot is almost invariant. This suggests a nice scalability of the solution.

[6] It is unlikely that the actual standard deviation will be smaller and a larger value would improve the dispersion of $\overline{I}_0$ and the SDF of our solution.

[7] The $\overline{t}_{\max}$ of the different cells are not exactly equal since $\overline{t}_{\max}$ depends on τ which in turn depends on I_s and V_{T}.



(a) (b)

Figure 9. (a) Variation of $\overline{V}_{\text{eq}}$ after 10 years. (b) Zoom of (a).

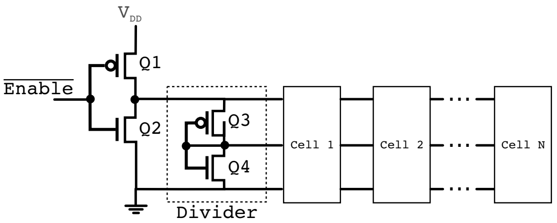


Figure 10. A possible power control solution

2) Comparison with other proposals: Table III compares the predicted performance of the proposed scheme with other schemes in the literature. The value of μ_{intra} for our solution has been computed using (3a) and the results in Fig. 6, while $\mu_{\text{inter}} = 50\%$ has been obtained from Property 6 and the fact that the pdf of V_{eq} is even (Fig. 6).

Table IV shows the number of bits (information + redundancy) used by the optimal[8] Reed-Solomon code necessary to stabilize an N-bit PUC with a failure probability smaller than η, together with the estimated *relative*[9] cost (in terms of silicon area) of the decoder. The area required by the decoder has been obtained by [43, Table III], see Appendix E-A2 for details. The cases where no correction code is necessary are marked in bold. These cases are interesting because, beyond not requiring any redundancy cell, they do not require error-correction circuits, nor a Non Volatile Memory (NVM). The predicted advantage of the proposed solution is clear. The details of the procedure used to compute Table IV are described in Appendix E-A.

Remark IX.2 (Gating cost)

As suggested in Section V-E, power consumption can be reduced by powering the cell only for the time required to reach the equilibrium, then turning them off. In order to make a farier comparison, it is necessary to take into account also the area required by the circuit required to power the cells. A possible solution for a group of N cells is shown in Fig. 10. Since the cells are powered through `Q1`, the size of `Q1` must be proportional to N. In order to estimate the required area, we determined, by means of simulations and for several different cell sizes, the minimum size of `Q1`. It turned out that an indicative planning figure for the area of the best `Q1` is approximately 10% of the area used by the cells. For example, in the case of $N = 64$, the area required for power control is equivalent to 6–7 additional cells.

Table V shows the number of iterations that would be required to implement an (η, δ)-stable helper-less stabilizer using the approach of [9]. The entries in bold are relative to the case when no iteration is required. Also in this case the advantage is clear.

Although these results need to be verified experimentally, the margin over other solutions is high enough that we expect that the proposed PUC will maintain its competitiveness.

a) Environmental Variations: The results in Table III are relative to fixed environmental conditions. Because of the results of Section IX-B, we predict that the same figures will hold even in the presence of power supply variations.

Because of the lack of suitable models, it is more difficult to predict the actual impact of temperature. If the selection procedure in Section VIII-C3 is *not* employed, we expect, from the results of Section IX-A, that over an extended range of temperatures μ_{intra} could increase by approximately 1%. The precise impact of the temperature needs to be verified in an actual circuit implementation.

X. Conclusions and Future Directions

We proposed a 1-bit PUC with a single equilibrium point that depends discontinuously on cell asymmetry. The behavior of the cell has been analyzed both analytically and by simulations. From the theoretical analysis some design guidelines were derived. We predict that $\mu_{\text{inter}} = 50\%$, μ_{intra} can be as small as 10^{-3} or 10^{-4}, and that the cell is insensitive to power supply variations. Preliminary analysis show a limited sensitivity to temperature variations, but the lack of suitable models requires that a definitive answer is obtained experimentally. Although the predicted performance need to be confirmed experimentally, the margin over other solutions is such that the we expect that the proposed PUC will maintain its advantage.

Further research will aim to verify the predictions and to investigate in more detail the effect of temperature and aging.

References

[1] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas, "Silicon physical random functions," in *Proceedings of the 9th ACM conference on Computer and communications security*, CCS '02, (New York, NY, USA), pp. 148–160, ACM, 2002.

[2] D. Lim, J. Lee, B. Gassend, G. Suh, M. van Dijk, and S. Devadas, "Extracting secret keys from integrated circuits," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 13, no. 10, pp. 1200–1205, 2005.

[3] D. Lim, "Extracting secret keys from integrated circuits," Master's thesis, MIT, May 2004.

[4] G. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Design Automation Conference, 2007. DAC '07. 44th ACM/IEEE*, pp. 9–14, 2007.

[5] D. E. Holcomb, W. P. Burleson, and K. Fu, "Initial SRAM state as a fingerprint and source of true random numbers for rfid tags," in *In Proceedings of the Conference on RFID Security*, 2007.

[6] R. Maes, P. Tuyls, and I. Verbauwhede, "A soft decision helper data algorithm for SRAM PUFs," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pp. 2101–2105, 2009.

[7] Z. Paral and S. Devadas, "Reliable and efficient PUF-based key generation using pattern matching," in *Hardware-Oriented Security and Trust (HOST), 2011 IEEE International Symposium on*, pp. 128–133, 2011.

[8] J. Delvaux and I. Verbauwhede, "Fault injection modeling attacks on 65nm arbiter and ro sum PUFs via environmental changes." Cryptology ePrint Archive, Report 2013/619, 2013. http://eprint.iacr.org/.

[8] In the sense of minimum number of total bits employed

[9] *Relative* means that all the costs have been normalized to the smallest one for a specific combination of N and η. For example, for $N = 64$ and $\eta = 0.01$, the decoder for an SRAM-based PUC will take approximately an area 30 times larger than the decoder for the 1.5/5 cell.

[9] R. Bernardini and R. Rinaldo, “Theoretical limits of helper-less stabilizers for physically unclonable constants,” *Emerging Topics in Computing, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2014. doi : 10.1109/TETC.2014.2386137.

[10] R. Bernardini and R. Rinaldo, “Physically unclonable random permutations,” in *Recent Advances in Electrical and Electronic Engineering*, (Florence, Italy), pp. 148–154, Nov. 2014.

[11] M. Bellare and P. Rogaway, “Random oracles are practical: A paradigm for designing efficient protocols,” in *Proceedings of the 1st ACM Conference on Computer and Communications Security*, CCS ’93, (New York, NY, USA), pp. 62–73, ACM, 1993.

[12] O. Goldreich, S. Goldwasser, and S. Micali, “How to construct random functions,” *J. ACM*, vol. 33, pp. 792–807, Aug. 1986.

[13] Y. Dodis, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” in *Advances in Cryptology - EUROCRYPT 2004* (C. Cachin and J. Camenisch, eds.), vol. 3027 of *Lecture Notes in Computer Science*, pp. 523–540, Springer Berlin Heidelberg, 2004.

[14] B. Skoric and N. de Vreede, “The spammed code offset method.” Cryptology ePrint Archive, Report 2013/527, 2013. http://eprint.iacr.org/.

[15] M.-D. M. Yu and S. Devadas, “Secure and robust error correction for physical unclonable functions.,” *IEEE Design & Test of Computers*, vol. 27, no. 1, pp. 48–65, 2010.

[16] E. Öztürk, G. Hammouri, and B. Sunar, “Towards robust low cost authentication for pervasive devices,” in *Sixth Annual IEEE International Conference on Pervasive Computing and Communications (PerCom 2008), 17-21 March 2008, Hong Kong*, pp. 170–178, 2008.

[17] D. E. Holcomb, W. P. Burleson, and K. Fu, “Power-up SRAM state as an identifying fingerprint and source of true random numbers,” *IEEE Transactions on Computers*, 2009.

[18] Y. Su, J. Holleman, and B. P. Otis, “A 1.6pj/bit 96% stable chip-id generating circuit using process variations,” in *2007 IEEE International Solid-State Circuits Conference, ISSCC 2007, Digest of Technical Papers, San Francisco, CA, USA, February 11-15, 2007*, pp. 406–611, 2007.

[19] E. S. Kumar, J. Guajardo, R. Maesyz, G. jan Schrijen, and P. Tuyls, “Extended abstract: The butterfly puf protecting ip on every fpga,” in *in Hardware-Oriented Security and Trust. HOST, 2008. IEEE International Workshop on*, pp. 67–70, 2008.

[20] Y. Su, J. Holleman, and B. Otis, “A digital 1.6 pJ/bit chip identification circuit using process variations,” *Solid-State Circuits, IEEE Journal of*, vol. 43, pp. 69–77, Jan 2008.

[21] X. Xu, A. Rahmati, D. Holcomb, K. Fu, and W. Burleson, “Reliable physical unclonable functions using data retention voltage of SRAM cells,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2015. doi: 10.1109/TCAD.2015.2418288.

[22] J. Lee, D. Lim, B. Gassend, G. Suh, M. van Dijk, and S. Devadas, “A technique to build a secret key in integrated circuits for identification and authentication applications,” in *VLSI Circuits, 2004. Digest of Technical Papers. 2004 Symposium on*, pp. 176–179, June 2004.

[23] K. Matsunaga, S. Oshima, T. Minotani, T. Kondo, and H. Morimura, “Automatic identification number generation circuit using NMOS pair current mismatch,” *Japanese Journal of Applied Physics*, vol. 54, no. 4S, p. 04DE12, 2015.

[24] K. Lofstrom, W. Daasch, and D. Taylor, “Ic identification circuit using device mismatch,” in *Solid-State Circuits Conference, 2000. Digest of Technical Papers. ISSCC. 2000 IEEE International*, pp. 372–373, Feb 2000.

[25] P. Prabhu, A. Akel, L. Grupp, W.-K. Yu, G. Suh, E. Kan, and S. Swanson, “Extracting device fingerprints from flash memory by exploiting physical variations,” in *Trust and Trustworthy Computing* (J. McCune, B. Balacheff, A. Perrig, A.-R. Sadeghi, A. Sasse, and Y. Beres, eds.), vol. 6740 of *Lecture Notes in Computer Science*, pp. 188–201, Springer Berlin Heidelberg, 2011.

[26] F. Tang, D. Chen, B. Wang, A. Bermak, A. Amira, and S. Mohamad, “CMOS on-chip stable true-random ID generation using antenna effect,” *Electron Device Letters, IEEE*, vol. 35, pp. 54–56, Jan 2014.

[27] R. Maes and V. van der Leest, “Countering the effects of silicon aging on sram pufs,” in *Hardware-Oriented Security and Trust (HOST), 2014 IEEE International Symposium on*, pp. 148–153, May 2014.

[28] A. Hosey, M. T. Rahman, K. Xiao, D. Forte, and M. Tehranipoor, “Advanced analysis of cell stability for reliable sram pufs,” in *2014 IEEE 23rd Asian Test Symposium*, pp. 348–353, Nov 2014.

[29] R. Maes and I. Verbauwhede, “Physically unclonable functions: A study on the state of the art and future research directions,” in *Towards Hardware-Intrinsic Security - Foundations and Practice*, pp. 3–37, 2010.

[30] R. Bernardini and R. Rinaldo, “Analysis of the security impact of local bias in physically unclonable constants.” submitted.

[31] R. Bernardini and R. Rinaldo, “A simple and reliable cell for single bit physically unclonable constants,” in *Proc. Austrochip 2014*, (Granz, Austria), Sept. 2014. doi: 10.1109/Austrochip.2014.6946317.

[32] B. Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. John Wiley and Sons, 1994.

[33] B. Diffie, M. Hellman, and R. Merkle, “Cryptographic apparatus and method.” PatentUS US 4200770 A, Apr. 1980.

[34] Y. Sheffer, G. Zorn, H. Tschofenig, and S. Fluhrer, “An EAP authentication method based on the encrypted key exchange (EKE) protocol.” RFC 6124, Feb. 2011. http://tools.ietf.org/html/rfc6124.

[35] R. Bernardini and R. Rinaldo, “Helper-less physically unclonable functions and chip authentication,” in *Proc. ICASSP 2014*, (Firenze), May 2014. doi: 10.1109/ICASSP.2014.6855198.

[36] J. L. Massey, “Guessing and entropy,” in *Proc. 1994 IEEE International Symposium on Information Theory*, p. 204, 1994.

[37] R. J. Baker, *CMOS Circuit Design, Layout and Simulation*. Wiley, 2010.

[38] P. Allen and D. Holberg, *CMOS Analog Circuit Design*. Oxford series in electrical and computer engineering, Oxford University Press, 2002.

[39] M. Keshner, “1/f noise,” *Proceedings of the IEEE*, vol. 70, pp. 212–218, March 1982.

[40] E. J. McDowell, X. Cui, Z. Yaqoob, and C. Yang, “A generalized noise variance analysis model and its application to the characterization of 1/f noise,” *Opt. Express*, vol. 15, pp. 3833–3848, Apr 2007.

[41] P. Kinget, “Device mismatch and tradeoffs in the design of analog circuits,” *Solid-State Circuits, IEEE Journal of*, vol. 40, pp. 1212–1224, June 2005.

[42] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, “Surface roughness at the Si(100)-SiO$_2$ interface,” *Phys. Rev. B*, vol. 32, pp. 8171–8186, Dec 1985.

[43] N. Chen and Z. Yan, “Complexity analysis of reed-solomon decoding over gf(2ˆm) without using syndromes,” *CoRR*, vol. abs/0803.0731, 2008.

Appendix A
Proofs of Section II

A. Proof of (3a)

Suppose we have a N bit PUC that we query twice. Let x_i and y_i be the outcome of the i-th PUC after the first and the second query. Collect them in $\mathbf{x}$ and $\mathbf{y}$. We are interested in computing the average Hamming distance normalized to the number of bits N, that is,

$$\begin{aligned}\mu_{\text{intra}} &= \frac{1}{N}\mathbb{E}\left[d(\mathbf{x},\mathbf{y})\right]\\ &= \frac{1}{N}\sum_i \mathbb{E}\left[d(x_i,y_i)\right]\\ &= \frac{1}{N}\sum_i P[x_i \neq y_i] = P[x \neq y]\end{aligned} \tag{37}$$

In order to compute the last probability in (37), condition with respect to p_1

$$\begin{aligned}P[x \neq y] &= \int_0^1 f_p(u) P[x \neq y | p_1 = u]\, du\\ &= \int_0^1 f_p(u) 2u(1-u)\, du\\ &= 2\mathbb{E}\left[p_1(1-p_1)\right]\end{aligned} \tag{38}$$

Now observe that

$$p_1(1-p_1) = \frac{1}{4}(1-R^2(p_1))x \tag{39}$$

Therefore, from (38)

$$\mu_{\text{intra}} = P[x \neq y] = \frac{1}{2}\left(1-\mathbb{E}\left[R^2\right]\right) \tag{40}$$

Let f_R be the density of R and observe that

$$\begin{aligned}\mathbb{E}\left[R^2\right] &= \int_0^1 f_R(u)u^2\,du \\ &= \left[F_R(u)u^2\right]_0^2 - 2\int_0^1 F_R(u)u\,du \\ &= 1-2\int_0^1 F_R(u)u\,du\end{aligned} \tag{41}$$

which used in (40) implies

$$\mu_{\text{intra}} = \int_0^1 F_R(u)u\,du \le \int_0^1 F_R(u)\,du \tag{42}$$

B. Proof of (3b)

The experiment is the following: we build two N-bit PUCs, we query them to obtain two vectors of outcomes $\mathbf{x}$ and $\mathbf{y}$. We are interested in computing

$$\mu_{\text{inter}} = \frac{1}{N}\mathbb{E}\left[d(\mathbf{x},\mathbf{y})\right] = \frac{1}{N}\sum_i \mathbb{E}\left[\chi(x_i \neq y_i)\right] = P[x_i \neq y_i] \tag{43}$$

Let f_p be the density of p_1 and observe that

$$\begin{aligned}P[x_i \neq y_i] &= \int_{[0,1]^2} f_p(u)f_p(v)P[x_i \neq y_i | p=u, q=v]dudv \\ &= \int_{[0,1]^2} f_p(u)f_p(v)[u(1-v)+v(1-u)]dudv \\ &= 2\int_0^1 f_p(a)da - 2\int_{[0,1]^2} f_p(u)f_p(v)\,uv\,dudv \\ &= 2(m_p - m_p^2)\end{aligned} \tag{44}$$

Appendix B
Proofs of Section V

A. Derivation of $i_C(v_C)$

In order to find $i_C(v_C)$, observe that the currents I_1 and I_2 can be written explicitly as functions of v_C as

$$I_{D,1}(v_C) = \begin{cases} \frac{\beta_1}{2}(V_{DR}-v_C)(V_{DR}-2V_{\mathrm{T1}}+v_C) & v_C > V_{\mathrm{T1}} \\ \frac{\beta_1}{2}(V_{DR}-V_{\mathrm{T1}})^2\left(1+\lambda(V_{\mathrm{T1}}-v_C)\right) & v_C < V_{\mathrm{T1}}\end{cases}$$

$$I_{D,2}(v_C) = \begin{cases} \frac{\beta_2}{2}(V_{RR}+v_C)(V_{RR}-2V_{\mathrm{T2}}-v_C) & v_C \le -V_{\mathrm{T2}} \\ \frac{\beta_2}{2}(V_{RR}-V_{\mathrm{T2}})^2\left(1+\lambda(V_{\mathrm{T2}}+v_C)\right) & v_C > -V_{\mathrm{T2}}\end{cases}$$

By using the saturation currents $I_{s,i}$, it is possible to rewrite (45) in a simpler form. To such an end, note that

$$\begin{aligned}I_1(v_C) &= \frac{\beta_1}{2}(V_{DR}-v_C)(V_{DR}-2V_{\mathrm{T1}}+v_C) \\ &= \frac{\beta_1}{2}[(V_{DR}-V_{\mathrm{T1}})+(V_{\mathrm{T1}}-v_C)][(V_{DR}-V_{\mathrm{T1}})-(V_{\mathrm{T1}}-v_C)] \\ &= \frac{\beta_1}{2}[(V_{DR}-V_{\mathrm{T1}})^2-(V_{\mathrm{T1}}-v_C)^2] \\ &= I_{s,1} - \frac{\beta_1}{2}(V_{\mathrm{T1}}-v_C)^2\end{aligned} \tag{46a}$$

$$\begin{aligned}I_2(v_C) &= \frac{\beta_2}{2}(V_{RR}+v_C)\left[(V_{RR}-V_{\mathrm{T2}})-(V_{\mathrm{T2}}+v_C)\right] \\ &= \frac{\beta_2}{2}\left[(V_{RR}-V_{\mathrm{T2}})+(V_{\mathrm{T2}}+v_C)\right]\left[(V_{RR}-V_{\mathrm{T2}})-(V_{\mathrm{T2}}+v_C)\right] \\ &= \frac{\beta_2}{2}[(V_{RR}-V_{\mathrm{T2}})^2-(v_C+V_{\mathrm{T2}})^2] \\ &= I_{s,2} - \frac{\beta_2}{2}(V_{\mathrm{T2}}+v_C)^2\end{aligned} \tag{46b}$$

It follows

$$I_{D,1}(v_C) = \begin{cases} I_{s,1} - \frac{\beta_1}{2}(V_{\mathrm{T1}}-v_C)^2 & v_C > V_{\mathrm{T1}} \\ I_{s,1} + \lambda I_{s,1}(V_{\mathrm{T1}}-v_C) & v_C < V_{\mathrm{T1}}\end{cases} \tag{47a}$$

$$I_{D,2}(v_C) = \begin{cases} I_{s,2} - \frac{\beta_2}{2}(V_{\mathrm{T2}}+v_C)^2 & v_C < -V_{\mathrm{T2}} \\ I_{s,2} + \lambda I_{s,2}(V_{\mathrm{T2}}+v_C) & v_C \ge -V_{\mathrm{T2}}\end{cases} \tag{47b}$$

In order to derive $i_C(v_C)$, we need to consider three cases: $v_C \ge V_{\mathrm{T1}} > -V_{\mathrm{T2}}$, $v_C \in (-V_{\mathrm{T2}}, V_{\mathrm{T1}})$ and $v_C \le -V_{\mathrm{T2}}$. Consider first the case $v_C \ge V_{\mathrm{T1}} > -V_{\mathrm{T2}}$

$$\begin{aligned}i_C(v_C) &= I_{D,1}(v_C) - I_{D,2}(v_C) \\ &= I_0 - \frac{\beta_1}{2}(V_{\mathrm{T1}}-v_C)^2 - \lambda I_{s,2}(V_{\mathrm{T2}}+v_C) \\ &= I_0 - \frac{\beta_1}{2}(V_{\mathrm{T1}}-v_C)^2 + \lambda I_{s,2}(V_{\mathrm{T1}}-V_{\mathrm{T1}}-V_{\mathrm{T2}}-v_C) \\ &= [I_0 - \lambda I_{s,2}(V_{\mathrm{T1}}+V_{\mathrm{T2}})] \\ &\quad - \frac{\beta_1}{2}(V_{\mathrm{T1}}-v_C)^2 + \lambda I_{s,2}(V_{\mathrm{T1}}-v_C)\end{aligned} \tag{48}$$

The case $v_C \in (-V_{\mathrm{T2}}, V_{\mathrm{T1}})$ is

$$\begin{aligned}i_C(v_C) &= I_{D,1}(v_C) - I_{D,2}(v_C) \\ &= I_0 + \lambda I_{s,1}(V_{\mathrm{T1}}-v_C) - \lambda I_{s,2}(V_{\mathrm{T2}}+v_C) \\ &= I_0 + \lambda(I_{s,1}V_{\mathrm{T1}} - I_{s,2}V_{\mathrm{T2}}) - \lambda(I_{s,1}+I_{s,2})v_C\end{aligned} \tag{49}$$

Finally, the case $v_C \le -V_{\mathrm{T2}}$ is

$$\begin{aligned}i_C(v_C) &= I_{D,1}(v_C) - I_{D,2}(v_C) \\ &= I_0 + \lambda I_{s,1}(V_{\mathrm{T1}}-v_C) + \frac{\beta_2}{2}(V_{\mathrm{T2}}+v_C)^2 \\ &= I_0 - \lambda I_{s,1}(V_{\mathrm{T2}}-V_{\mathrm{T2}}-V_{\mathrm{T1}}+v_C) + \frac{\beta_2}{2}(V_{\mathrm{T2}}+v_C)^2 \\ &= [I_0 + \lambda I_{s,1}(V_{\mathrm{T2}}+V_{\mathrm{T1}})] \\ &\quad - \lambda I_{s,1}(V_{\mathrm{T2}}+v_C) + \frac{\beta_2}{2}(V_{\mathrm{T2}}+v_C)^2\end{aligned} \tag{50}$$

In order to simplify (48), (49) and (50) observe that

$$\begin{aligned}\lambda I_{s,2}(V_{\mathrm{T1}}+V_{\mathrm{T2}}) &= \lambda(I_s+\delta_I)(2V_{\mathrm{T}}+\delta_{v1}) \\ &= 2V_{\mathrm{T}}\lambda I_s + 2V_{\mathrm{T}}\lambda\delta_I + \lambda I_s\delta_v\end{aligned} \tag{51}$$

Since δ_v and δ_I are expected to be at least one order of magnitude smaller than V_{T} and I_s, we can write

$$\lambda I_{s,2}(V_{\mathrm{T1}}+V_{\mathrm{T2}}) \approx 2V_{\mathrm{T}}\lambda I_s = \theta I_s \tag{52}$$

By a similar reasoning we can write

$$\lambda I_{s,1}(V_{\mathrm{T2}}+V_{\mathrm{T1}}) \approx \theta I_s \tag{53a}$$

$$\lambda I_{s,1}(V_{\mathrm{T2}}+v_C) \approx \lambda I_s(V_{\mathrm{T}}+v_C) = \frac{\theta}{2}I_s(1+\overline{v}_{\mathrm{C}}) \tag{53b}$$

$$\lambda I_{s,2}(V_{\mathrm{T1}}-v_C) \approx \lambda I_s(V_{\mathrm{T}}-v_C) = \frac{\theta}{2}I_s(1-\overline{v}_{\mathrm{C}}) \tag{53c}$$

$$\lambda(I_{s,1}V_{\mathrm{T1}} - I_{s,2}V_{\mathrm{T2}}) \approx 0 \tag{53d}$$

$$\lambda(I_{s,1}+I_{s,2})v_C \approx 2I_s\lambda v_C = I_s\theta\overline{v}_{\mathrm{C}} \tag{53e}$$

By using approximations (53) and (52) in (48), (49) and (50) and dividing the result by I_s we obtain

$$\bar{i}_{\mathrm{C}}(\overline{v}_{\mathrm{C}}) = \begin{cases} [\bar{I}_0+\theta] - \frac{\theta(1+\overline{v}_{\mathrm{C}})}{2} + \frac{(1+\overline{v}_{\mathrm{C}})^2 {V_{\mathrm{T}}}^2\beta_2/2}{I_s} & \overline{v}_{\mathrm{C}} \leq -1 \\ \bar{I}_0 - \theta\overline{v}_{\mathrm{C}} & |\overline{v}_{\mathrm{C}}| < 1 \\ [\bar{I}_0-\theta] + \frac{(1-\overline{v}_{\mathrm{C}})\theta}{2} - \frac{(1-\overline{v}_{\mathrm{C}})^2 {V_{\mathrm{T}}}^2\beta_1/2}{I_s} & \overline{v}_{\mathrm{C}} \geq 1 \end{cases} \tag{54}$$

By observing now that

$$\begin{aligned} \frac{{V_{\mathrm{T}}}^2\beta_1/2}{I_s} &\approx \frac{{V_{\mathrm{T}}}^2\beta_2/2}{I_s} \approx \frac{{V_{\mathrm{T}}}^2\beta/2}{I_s} \\ &= \frac{{V_{\mathrm{T}}}^2\beta/2}{(\beta/2)(V_{\mathrm{T}}-V_{RR})^2} = \frac{1}{\overline{V}_\Delta^2} \end{aligned} \tag{55}$$

equation (54) can be rewritten as (7).

B. Derivation of $V_{eq}(I_0)$, equation (9)

In order to find the value of V_{eq} we need to consider three different cases, corresponding to the three regions of (7).

1) Central part: Suppose that V_{eq} belongs to the central region $[-V_{\mathrm{T2}}, V_{\mathrm{T1}}]$. In this case

$$\begin{aligned} V_{\mathrm{eq}} &= \frac{I_0 + \lambda(I_{s,1}V_{\mathrm{T1}} - I_{s,2}V_{\mathrm{T2}})}{\lambda(I_{s,1}+I_{s,2})} \\ &= \frac{I_0}{\lambda(I_{s,1}+I_{s,2})} + \frac{I_{s,1}V_{\mathrm{T1}} - I_{s,2}V_{\mathrm{T2}}}{I_{s,1}+I_{s,2}} \\ &= \frac{I_0}{\lambda(I_{s,1}+I_{s,2})} + \frac{I_0V_{\mathrm{T1}} + I_{s,2}(V_{\mathrm{T1}}-V_{\mathrm{T2}})}{I_{s,1}+I_{s,2}} \\ &= \frac{I_0}{\lambda(I_{s,1}+I_{s,2})} + \frac{I_0V_{\mathrm{T1}}}{I_{s,1}+I_{s,2}} + \frac{I_{s,2}(V_{\mathrm{T1}}-V_{\mathrm{T2}})}{I_{s,1}+I_{s,2}} \\ &\approx \frac{I_0}{2\lambda I_s} + \frac{I_0V_{\mathrm{T1}}}{2I_s} + \frac{I_s(V_{\mathrm{T1}}-V_{\mathrm{T2}})}{2I_s} \\ &= \frac{I_0}{2I_s}\left(\frac{1}{\lambda} + V_{\mathrm{T1}}\right) + (V_{\mathrm{T1}}-V_{\mathrm{T2}}) \approx \frac{I_0}{2I_s\lambda} \end{aligned} \tag{56}$$

where we used the approximations $I_{s,1} \approx I_s \approx I_{s,2}$, $V_{\mathrm{T1}} \approx V_{\mathrm{T}} \approx V_{\mathrm{T2}}$ and $1/\lambda \gg V_{\mathrm{T}}$. In order for (56) being the real value of V_{eq} it must be

$$\left|\frac{I_0}{2I_s\lambda}\right| < V_{\mathrm{T}} \Leftrightarrow \left|\frac{I_0}{I_s}\right| = \bar{I}_0 < 2\lambda V_{\mathrm{T}} = \theta \tag{57}$$

In order to write (56) in adimensional form observe that

$$\overline{V}_{\mathrm{eq}} = \frac{V_{\mathrm{eq}}}{V_{\mathrm{T}}} = \frac{I_0}{2I_s\lambda V_{\mathrm{T}}} = \frac{I_0}{I_s}\frac{1}{2\lambda V_{\mathrm{T}}} = \frac{\bar{I}_0}{\theta} \tag{58}$$

2) Case $I_0 < -\theta$: In the case $v_C \leq -V_{\mathrm{T2}}$ (corresponding to $I_0 < 0$ the equation is

$$I_0 + \lambda I_{s,1}(V_{\mathrm{T1}} - v_C) + \frac{\beta_2}{2}(v_C + V_{\mathrm{T2}})^2 = 0 \tag{59}$$

Replace $v_C + V_{\mathrm{T2}}$ with u to obtain

$$\begin{aligned} 0 &= \frac{2I_0}{\beta_2} + \frac{2\lambda I_{s,1}}{\beta_2}(V_{\mathrm{T1}}+V_{\mathrm{T2}}) - \frac{2\lambda I_{s,1}}{\beta_2}u + u^2 \\ &= u^2 - \frac{2\lambda I_{s,1}}{\beta_2}u + \left[\frac{2I_0}{\beta_2} + \lambda\frac{2I_{s,1}}{\beta_2}(V_{\mathrm{T1}}+V_{\mathrm{T2}})\right] \\ &\approx u^2 - \frac{2\lambda I_s}{\beta}u + \left[\frac{2I_0}{\beta} + \lambda\frac{4I_sV_{\mathrm{T}}}{\beta_2}\right] \\ &= u^2 - \frac{2\lambda I_s}{\beta}u + \left[\frac{I_0}{I_s}\frac{2I_s}{\beta} + \lambda\frac{4I_sV_{\mathrm{T}}}{\beta}\right] \end{aligned} \tag{60}$$

Remember that $\lambda I_s/\beta = (1/2)g^2\theta V_{\mathrm{T}}$, so that (60) is equivalent to

$$\begin{aligned} 0 &= u^2 - \overline{V}_\Delta^2\theta V_{\mathrm{T}}u + \left[\bar{I}_0\overline{V}_\Delta^2{V_{\mathrm{T}}}^2 + 2\overline{V}_\Delta^2\theta{V_{\mathrm{T}}}^2\right] \\ &= u^2 - \overline{V}_\Delta^2\theta V_{\mathrm{T}}u + \overline{V}_\Delta^2{V_{\mathrm{T}}}^2\left[\bar{I}_0 + 2\theta\right] \end{aligned} \tag{61}$$

By dividing both sides of (61) by ${V_{\mathrm{T}}}^2$ in order to make it adimensional we get

$$0 = x^2 - \overline{V}_\Delta^2\theta x + \overline{V}_\Delta^2\left[\bar{I}_0 + 2\theta\right] \tag{62}$$

whose solutions are

$$\begin{aligned} &\frac{1}{2}\left(\overline{V}_\Delta^2\theta \pm \sqrt{\overline{V}_\Delta^4\theta^2 - 4\overline{V}_\Delta^2(\bar{I}_0+2\theta)}\right) \\ &= \frac{1}{2}\left(\overline{V}_\Delta^2\theta \pm \overline{V}_\Delta\sqrt{\overline{V}_\Delta^2\theta^2 - 4(\bar{I}_0+2\theta)}\right) \\ &\approx \pm\overline{V}_\Delta\sqrt{-(\bar{I}_0+2\theta)} \end{aligned} \tag{63}$$

Remember that $V_{\mathrm{eq}} = u - V_{\mathrm{T}}$ which implies $\overline{V}_{\mathrm{eq}} = x - 1$.

3) Case $I_0 > \theta$: Similar to the case $I_0 < -\theta$

C. Proof of Property 2

First we need a lemma.

Lemma 2. *The solution of the following problem*

$$\begin{aligned} \dot{u}(t) &= a - bu^2(t) + cu(t); \qquad ab > 0, t \geq 0 && \text{(64a)} \\ u(0) &= 0 && \text{(64b)} \end{aligned}$$

is

$$u(t) = A\tanh(Bt + C) + R \qquad t \geq 0 \tag{65}$$

where

$$\begin{aligned} R &= \frac{c}{2b} && \text{(66a)} \\ A &= \sqrt{\frac{a}{b} + \frac{c^2}{4b^2}} = \sqrt{\frac{a}{b} + R^2} && \text{(66b)} \\ B &= \frac{A}{b} = \sqrt{ab + \frac{c^2}{4}} && \text{(66c)} \\ C &= \tanh^{-1}\left(-\frac{R}{A}\right) && \text{(66d)} \end{aligned}$$

Proof. We first make a change of variable in order to remove the linear term. Let $u = y + R$, so that $\dot{y} = \dot{u}$ and (64a) becomes

$$\begin{aligned} \dot{u} = \dot{y} &= a - b(y+R)^2 + c(y+R) \\ &= a - by^2 - bR^2 - 2bRy + cy + cR \\ &= (a - bR^2 + cR) + (c - 2bR)y - by^2 \end{aligned} \tag{67}$$

The linear term disappear when $R = c/(2b)$, so that (64) becomes

$$\begin{aligned} &y(0) = -R && \text{(68a)} \\ &\dot{y} = a - b\frac{c^2}{4b^2} + c\frac{c}{2b} - by^2 = \underbrace{\left(a + \frac{c^2}{4b}\right)}_{\hat{a}} - by^2 && \text{(68b)} \end{aligned}$$

where, for notational convenience, we introduced the new variable $\hat{a}$. Now search for solutions of (68b) with the form

$$y(t) = \alpha \tanh(\beta t + \gamma) \tag{69}$$

Differentiating y gives

$$\begin{aligned}\dot{y}(t) &= \alpha\beta(1 - \tanh^2(\beta t + \gamma)) \\ &= \alpha\beta - \alpha\beta \tanh^2(\beta t + \gamma)) \\ &= \alpha\beta - \frac{\beta}{\alpha} y^2(t)\end{aligned} \tag{70}$$

By comparing (64) and (70) we get

$$\hat{a} = \alpha\beta \qquad ; \qquad b = \frac{\beta}{\alpha} \tag{71}$$

which gives

$$\alpha = \sqrt{\frac{\hat{a}}{b}} = \sqrt{\frac{a}{b} + \frac{c^2}{4b^2}} = \sqrt{\frac{a}{b} + R^2} \tag{72}$$

$$\beta = b\alpha = \sqrt{\hat{a}b} = \sqrt{ab + \frac{c^2}{4}} \tag{73}$$

Note that $\alpha, \beta \in \mathbb{R}$ since $ab > 0$. In order to find the value of γ we are going to use the initial condition, searching for γ such that

$$\alpha \tanh(\gamma) = -R \tag{74}$$

whose solution is

$$\gamma = \tanh^{-1}\left(-\frac{R}{\alpha}\right) = -\tanh^{-1}\left(\frac{R}{\sqrt{a/b + R^2}}\right) \tag{75}$$

Note that the absolute value of the argument of the $\tanh^{-1}$ in (75) is smaller than one since $a/b > 0$, so that γ is real. □

Proof B.1. *Proof of Property 2* At time $\bar{t} = 0$, $\bar{v}_C(0) = 0$, so the linear region of (7) is active and the differential equation (12) can be specialized to

$$\begin{cases} \dot{u} = \underbrace{\left[\bar{I}_0\left(1 + \frac{\theta}{2}\right) + \frac{\theta\bar{\Delta}_v}{2}\right]}_{A} - \theta u \\ u(0) = 0 \end{cases} \tag{76}$$

The solution to (76) is well known to be

$$\begin{aligned} u(\bar{t}) &= \frac{A}{\theta}[1 - \exp(-\theta\bar{t})] \\ &= \underbrace{\left[\bar{I}_0\left(\frac{1}{\theta} + \frac{1}{2}\right) + \frac{\bar{\Delta}_v}{2}\right]}_{U} [1 - \exp(-\theta\bar{t})] \approx \frac{\bar{I}_0}{\bar{t}} \end{aligned} \tag{77}$$

where the approximation is valid in the limit $\lambda \to 0$. For $\bar{t} \to \infty$, $u(\bar{t})$ tends to U, the first factor in square brackets in (77). If $|U| < 1$, u never leaves the region $[-1, 1]$ and (77) is valid for any $\bar{t} > 0$. If $|U| > 1$, there will be a time t_{sw} when $|u(t_{sw})| = 1$ and one of the other two sections of (7) becomes active.

In order to compute t_{sw}, we approximate U in (77) with $\bar{I}_0/\theta$; this is possible when λ is small. It follows that $|U| > 1$ if $|\bar{I}_0| > \theta$. The value of t_{sw} can be obtained by solving

$$\frac{\bar{I}_0}{\theta}[1 - \exp(-\theta t_{sw})] = 1 \tag{78}$$

which gives

$$t_{sw} = -\frac{1}{\theta} \ln\left(1 - \frac{\theta}{\bar{I}_0}\right) \tag{79}$$

Suppose $\bar{I}_0 < -\theta < 0$. In this case $\bar{v}_C$ decreases and at time t_{sw} it will be $u(t_{sw}) = -1$, so that the section with support $(-\infty, -1)$ will become active. In this case (12) can be specialized to

$$\begin{cases} \dot{u}(\bar{t}) = \left(\bar{I}_0 + \frac{\theta}{2}\right) - \frac{\theta}{2}u(\bar{t}) + \frac{1}{\bar{V}_\Delta^2}(u(\bar{t}) + 1)^2 & \bar{t} \geq t_{sw} \\ u(t_{sw}) = -1 \end{cases} \tag{80}$$

In order to simplify the problem, define $w_-(\bar{t}) = u(\bar{t} + t_{sw}) + 1$. By rewriting (80) in $\bar{t} + t_{sw}$, $\bar{t} \geq 0$ and observing that $\dot{w}_- = \dot{u}$ one obtains

$$\begin{cases} \dot{w}_-(\bar{t}) = \bar{V}_\Delta\left(\bar{I}_0 + \frac{\theta}{2}\right) - \frac{\theta}{2}(w_-(\bar{t}) - 1) + \frac{1}{\bar{V}_\Delta^2}(w(\bar{t}))^2 & \bar{t} \geq 0 \\ \qquad = (\bar{I}_0 + \theta) - \frac{\theta}{2}w_-(\bar{t}) + \frac{1}{\bar{V}_\Delta^2}w_-^2(\bar{t}) \\ w_-(0) = 0 \end{cases} \tag{81}$$

By a similar reasoning, when $\bar{I}_0 > \theta$, after $\bar{t} = t_{sw}$, (12) can be specialized to

$$\begin{cases} \dot{u}(\bar{t}) = \left(\bar{I}_0 - \frac{\theta}{2}\right) - \frac{\theta}{2}u(\bar{t}) - \frac{1}{\bar{V}_\Delta^2}(u(\bar{t}) - 1)^2 & \bar{t} \geq t_{sw} \\ u(t_{sw}) = 1 \end{cases} \tag{82}$$

Similarly to what done before, we define $w_+(\bar{t}) = u(\bar{t} + t_{sw}) - 1$ and (82) becomes

$$\begin{cases} \dot{w}_+(\bar{t}) = \left(\bar{I}_0 - \frac{\theta}{2}\right) - \frac{\theta}{2}(w_+(\bar{t}) + 1) - \frac{1}{\bar{V}_\Delta^2}w_+^2(\bar{t})^2 & \bar{t} \geq 0 \\ \qquad = (\bar{I}_0 - \theta) - \frac{\theta}{2}w_+(\bar{t}) - \frac{1}{\bar{V}_\Delta^2}w_+^2(\bar{t})^2 \\ w_+(0) = 0 \end{cases} \tag{83}$$

Equations (81) and (83) are quite similar and we will exploit this similarity to show a relationship between w_+ and w_-. Define $r(\bar{t}) = -w_-(\bar{t})$, multiply both sides of (81) by -1 and obtain

$$\begin{cases} -\dot{w}_-(\bar{t}) = (-\bar{I}_0 - \theta) + \frac{\theta}{2}w_-(\bar{t}) - \frac{1}{\bar{V}_\Delta^2}w_-^2(\bar{t}) \\ w_-(0) = 0 \end{cases} \tag{84}$$

Now, by using the fact that $r = -w_+$ and that $|\bar{I}_0| = -\bar{I}_0$, equation (84) can be rewritten as

$$\begin{cases} \dot{r}(\bar{t}) = (|\bar{I}_0| - \theta) - \frac{\theta}{2}r(\bar{t}) - \frac{1}{\bar{V}_\Delta^2}r^2(\bar{t}) \\ r(0) = 0 \end{cases} \tag{85}$$

which is equal to (83), but with $|\bar{I}_0|$ instead of $\bar{I}_0$. It follows that

$$w_-(\bar{t}; \bar{I}_0, \theta, \bar{V}_\Delta) = -w_+(\bar{t}; |\bar{I}_0|, \theta, g) \tag{86}$$

so that it suffices to solve (83). From (86) we deduce, remembering that $u(\bar{t}) = w_(\bar{t} - t_{sw}) - -1$ that when $\bar{I}_0 < -\theta$ and $\bar{t} \ge t_{sw}$

$$u(\bar{t}) = -w_+(\bar{t} - t_{sw}; |\bar{I}_0|, \theta, \bar{V}_\Delta) - 1 \tag{87}$$

while when $\bar{I}_0 > \theta$ and $\bar{t} \ge t_{sw}$

$$u(\bar{t}) = w_+(\bar{t} - t_{sw}; |\bar{I}_0|, \theta, \bar{V}_\Delta) + 1 \tag{88}$$

Equations (88) can be unified in

$$u(\bar{t}) = \operatorname{sgn}\bar{I}_0 \left[w_+(\bar{t} - t_{sw}; |\bar{I}_0|, \theta, \bar{V}_\Delta) + 1\right] \qquad \bar{t} \ge t_{sw} \tag{89}$$

The last step is to find w_+ by solving (86). This can be done by applying Lemma 2. □

D. The case of non ideal voltage divider

In order to proof Property 3, we need first a lemma. Since in the following we will need to consider the dependence of i_C on V_{RR}, we will change, for this section only, the notation from $i_C(v_C)$ to $i_C(v_C, V_{RR})$. The same will be done for $I_{D,1}$ and $I_{D,2}$.

Lemma 3. *The following inequalities hold*

$$\frac{\partial i_C}{\partial v_C}(v_C, V_{RR}) \le 0 \tag{90a}$$

$$\frac{\partial i_C}{\partial V_{RR}}(v_C, V_{RR}) \le 0 \tag{90b}$$

In other words, i_C is a monotone decreasing function of both v_C and V_{RR}.

Proof. It has already be shown that, with fixed V_{RR}, i_C is monotone non-increasing function of v_C, that is, (90a). It remains to show (90b). Observe that

$$\begin{aligned} i_C(v_C, V_{RR}) &= I_{D,1}(v_C, V_{DR}) - I_{D,2}(v_C, V_{RR}) \\ &= I_{D,1}(v_C, V_{DD} - V_{RR}) - I_{D,2}(v_C, V_{RR}) \\ &= h(v_C, V_{RR}) - I_{D,2}(v_C, V_{RR}) \end{aligned} \tag{91}$$

Where we introduced the notation $h(v_C, V_{RR}) = I_{D,1}(v_C, V_{DD} - V_{RR})$. By differentiating (91) with respect to V_{RR} we get

$$\begin{aligned} \frac{\partial i_C}{\partial V_{RR}} &= \frac{\partial h}{\partial V_{RR}} - \frac{\partial I_{D,2}}{\partial V_{RR}} \\ &= -\frac{\partial I_{D,1}}{\partial V_{DR}} - \frac{\partial I_{D,2}}{\partial V_{RR}} < 0 \end{aligned} \tag{92}$$

where in the last step we exploited the fact that $I_{D,i}$ is a monotone increasing function of $V_{GS,i}$ when $V_{DS,i} - V_{GS,i}$ is fixed. □

Proof. Denote with $V_{RR}(i_C)$ the value of V_{RR} at the output of the divider when a current i_C enters in the middle terminal. Remember that the non-negative resistance hypothesis means that $dV_{RR}/di_C \ge 0$. If we replace the capacitance with an ideal voltage source of value v_C, the current x entering in the divider must satisfy

$$x = i_C(v_C, V_{RR}(x)) \tag{93}$$

Observe that in (93) is parameterized by v_C, therefore the solution of (93) will be function of v_C. We want to prove that the map from v_C to the solution of x is monotone non-decreasing. Define $u(v_C, x) = x - i_C(v_C, V_{RR}(x))$ and observe that the graph of the desired function is the subset of $\mathbb{R}^2$ where u is zero, that is,

$$S = \{(v, x) : u(v, x) = 0\} \tag{94}$$

Note that S is not empty since

$$u(V_{\text{eq}}, 0) = 0 - i_C(V_{\text{eq}}, V_{RR}(0)) = 0 - i_C(V_{\text{eq}}, V_{RR}^{\circ}) = 0 \tag{95}$$

so that $(V_{\text{eq}}, 0) \in S$.

We want to apply the implicit function theorem to u, so we differentiate it with respect to v and x

$$\frac{\partial u}{\partial v} = -\frac{\partial i_C}{\partial v_C} \ge 0 \tag{96a}$$

$$\frac{\partial u}{\partial x} = 1 - \frac{\partial i_C}{\partial V_{RR}} \frac{dV_{RR}}{dx} \ge 1 \tag{96b}$$

where the inequalities follow from Lemma 3 and the non-negative resistance hypothesis.

Because of (96b) the implicit function theorem grants us that there is a neighbor U of V_{eq} and a function $g : U \to \mathbb{R}$ such that $u(v, g(v)) = 0$ for every $v \in U$. Function g is the desired function. Moreover,

$$\frac{dg}{dv} = -\left(\frac{\partial u}{\partial x}\right)^{-1} \frac{\partial u}{\partial v} = \left(\frac{\partial u}{\partial x}\right)^{-1} \frac{\partial i_C}{\partial v_C} \le 0 \tag{97}$$

where the inequality follows at once from (96).

It is possible to say something more. Let

$$A \overset{\text{def}}{=} \sup\left(-\frac{\partial i_C}{\partial V_{RR}} \frac{dV_{RR}}{dx}\right) \le R_{\max} \sup\left(-\frac{\partial i_C}{\partial V_{RR}}\right) \tag{98}$$

and observe that $A \ge 0$. Note also that $A = 0$ if the divider is ideal. It follows that $1 \le \partial u / \partial v \le 1 + A$ and from (97) we deduce (remembering that $\partial i_C / \partial v_C \le 0$)

$$\frac{1}{1 + A} \frac{\partial i_C}{\partial v_C} \ge \frac{dg}{dv} \ge \frac{\partial i_C}{\partial v_C} \tag{99}$$

Remembering that $g(V_{\text{eq}}) = 0 = i_C(V_{\text{eq}})$, by integrating (99) from V_{eq} to x, one obtains (18). □

E. Computing $\bar{t}_{max}$

It will be shown in the following that map $|\bar{I}_0| \mapsto t_{\varepsilon,|\bar{I}_0|}$ is monotone decreasing and that $\bar{I}_0$ is approximately Gaussian with mean 0 and variance $\bar{\sigma}_I^2$. Giving for granted these two facts, one can deduce that inequality (20) is equivalent to

$$\begin{aligned} \eta &\ge \mathrm{P}[t_\varepsilon(|\bar{I}_0|) > \bar{t}_{\max}] && \text{Equation (20)} \\ &= \mathrm{P}[|\bar{I}_0| < t_\varepsilon^{-1}(\bar{t}_{\max})] && t_\varepsilon \text{ is monotone decreasing} \\ &= 1 - 2\Phi\left(-\frac{t_\varepsilon^{-1}(\bar{t}_{\max})}{\bar{\sigma}_I}\right) && \bar{I}_0 \sim \mathcal{N}(0, \bar{\sigma}_I^2) \end{aligned} \tag{100a}$$

whose solution can be found as

$$\begin{aligned} \mathrm{P}[t_{\varepsilon,|\bar{I}_0|} > \bar{t}_{\max}] &= P[|\bar{I}_0| < t_{\varepsilon,\bar{I}_0}{}^{-1}(\bar{t}_{\max})] \\ &= 2\Phi\left(\frac{t_{\varepsilon,\bar{I}_0}{}^{-1}(\bar{t}_{\max})}{\bar{\sigma}_I}\right) - 1 \end{aligned}$$

It follows

$$
\begin{aligned}
\mathrm{P}[t_{\varepsilon,|\bar{I}_0|} > \bar{t}_{\max}] \leq \eta &\Leftrightarrow 2\Phi\left(\frac{{t_{\varepsilon,\bar{I}_0}}^{-1}(\bar{t}_{\max})}{\overline{\sigma}_I}\right) - 1 \leq \eta \\
&\Leftrightarrow \Phi\left(\frac{{t_{\varepsilon,\bar{I}_0}}^{-1}(\bar{t}_{\max})}{\overline{\sigma}_I}\right) \leq \frac{1+\eta}{2} \\
&\Leftrightarrow \frac{{t_{\varepsilon,\bar{I}_0}}^{-1}(\bar{t}_{\max})}{\overline{\sigma}_I} \leq \Phi^{-1}\left(\frac{1+\eta}{2}\right) \\
&\Leftrightarrow {t_{\varepsilon,\bar{I}_0}}^{-1}(\bar{t}_{\max}) \leq \overline{\sigma}_I \Phi^{-1}\left(\frac{1+\eta}{2}\right) \\
&\Leftrightarrow \bar{t}_{\max} \geq t_{\varepsilon,\bar{I}_0}\left(\overline{\sigma}_I \Phi^{-1}\left(\frac{1+\eta}{2}\right)\right)
\end{aligned}
$$

and (21) follows.

1) Functional form of t_ε: In order to find function t_ε we need to solve (19). Three cases need to be considered

1) The first case is $|\bar{I}_0| < \theta$, so that $t_{sw} = \infty$. This implies that $t_{\varepsilon,\bar{I}_0}$ is in the first segment (i.e., $t_{\varepsilon,\bar{I}_0} < t_{sw}$) and $\bar{v}_{\mathrm{C}}(\bar{t})$ is an exponential for every $\bar{t}$. Condition (19) becomes

$$\bar{v}_{\mathrm{C}}(t_{\varepsilon,\bar{I}_0}) = \frac{\bar{I}_0}{\theta}\left(1 - \exp(-\theta t_{\varepsilon,\bar{I}_0})\right) = (1-\varepsilon)\frac{\bar{I}_0}{\theta} \tag{101}$$

which has the $\bar{I}_0$-independent solution

$$t_{\varepsilon,\bar{I}_0} = -\frac{\ln\varepsilon}{\theta} \tag{102}$$

2) In the second case $|\bar{I}_0| \geq \theta$ is large enough to have

$$\bar{V}_{\mathrm{eq}} > \frac{\bar{v}_{\mathrm{C}}(t_{sw})}{1-\varepsilon} \tag{103}$$

so that $t_{\varepsilon,\bar{I}_0} > t_{sw}$, that is, $t_{\varepsilon,\bar{I}_0}$ "falls" in the second segment. From (16), one deduces $\bar{V}_{\mathrm{eq}} = A + R + 1$, with A and R as in (14). By using this result in (19) one deduces, with the notation of Property 2

$$A\tanh(B(t_{\varepsilon,\bar{I}_0} - t_{sw})) + 1 = (1-\varepsilon)(A+1) \tag{104}$$

whose solution is

$$t_{\varepsilon,\bar{I}_0} = \bar{V}_\Delta \frac{\tanh^{-1}\left(-\varepsilon + \frac{1-\varepsilon}{\bar{V}_\Delta\sqrt{\bar{I}_0}}\right)}{\sqrt{\bar{I}_0}} + \frac{1}{\bar{I}_0} \tag{105}$$

Note that (105) is a decreasing function of $\bar{I}_0$, as one would expect, since a larger $\bar{I}_0$ causes a faster charging of $\mathtt{C}$.

3) The third case happens when $|\bar{I}_0| \geq \theta$, but $\bar{I}_0$ is not large enough to guarantee (103). In this case $t_{\varepsilon,\bar{I}_0} \leq t_{sw}$, that is, $t_{\varepsilon,\bar{I}_0}$ still "falls" in the first segment. Since there is a small interval of values of $\bar{I}_0$ that causes this behavior, the probability of this case is negligible and we do not consider it for the sake of simplicity.

Appendix C
Proofs of Section VI

A. Derivation of (26)

A well-known problem with integrating pink noise is the $1/f$ divergence around $f = 0$. Because of this, following [39], [40], we "cut" the spectrum for frequencies smaller than $1/t_{\max}$, by forcing to zero $R(f)$ for $|f| < 1/t_{\max}$.

Let ν be the noise current on $\mathtt{C}$. The result of integrating ν on $\mathtt{C}$ for $t_{\max}$ seconds can be written as

$$\xi = \int_0^{t_{\max}} \frac{\nu(t)}{C}dt = \frac{1}{C}\nu \star \mathrm{rect}_{t_{\max}}(0) \tag{106}$$

where $\mathrm{rect}_{t_{\max}}(t) = 1$ if $t \in [0, t_{\max}]$ and 0 otherwise. It follows that the variance of ξ can be computed as

$$\sigma_\xi^2 = \frac{1}{C^2}\int_{-\infty}^{+\infty} R(f)\, t_{\max}^2 \operatorname{sinc}^2(f t_{\max})\, df \tag{107}$$

where we used the fact that $t_{\max}|\operatorname{sinc}(f t_{\max})|$ is the modulus of the Fourier transform of $\mathrm{rect}_{t_{\max}}$.

The contribution of the constant a_0 (the white component) to σ_ξ^2 is

$$
\begin{aligned}
\int_{-\infty}^{+\infty} a_0 \frac{t_{\max}^2}{C^2}\operatorname{sinc}^2(t_{\max} f)df &= \frac{a_0 t_{\max}}{C^2}\int_{-\infty}^{+\infty}\operatorname{sinc}^2(w)dw \\
&= \frac{a_0 t_{\max}}{C^2}
\end{aligned}
\tag{108}
$$

Therefore, it is proportional to $t_{\max}$, as expected. In order to compute the contribution of $S(f)$ (the pink component), observe that

$$
\begin{aligned}
&2\int_{1/t_{\max}}^{\infty} a_1 \frac{t_{\max}^2}{C^2}\frac{\operatorname{sinc}^2(t_{\max} f)}{f}df \\
&\qquad = 2a_1\frac{t_{\max}^2}{C^2}\int_1^{\infty}\frac{\operatorname{sinc}^2(w)}{w}dw \approx 0.045 \cdot a_1 \frac{t_{\max}^2}{C^2}
\end{aligned}
\tag{109}
$$

By summing (108) and (109) we get (26).

Appendix D
Proofs of Section VII

The key result is the following property. Equation (32) follows at once from it.

Proof D.1. *Proof of Property 4* Note that $I_0 = I_{s,1} - I_{s,2}$ where $I_{s,1}$ and $I_{s,2}$ are not independent, since they both depend on δ_R. Because of this, we will first find the density of I_0 conditioned by δ_R and successively we will average the result with respect to δ_R.

The saturation currents $I_{s,i}$ in (5) can be rewritten by using the new variables (29) as

$$2I_{s,1|\delta_R} = (\beta + \delta_{\beta,1})(-\delta_{V,1} - \delta_R + V_\Delta)^2 \tag{110a}$$

$$2I_{s,2|\delta_R} = (\beta + \delta_{\beta,2})(-\delta_{V,2} + \delta_R + V_\Delta)^2 \tag{110b}$$

Subscript "$|\delta_R$" reminds us that we are conditioning with respect to δ_R. Equations (110) can be unified in

$$2I_{s|\delta_R} = (\beta + \delta_\beta)(U - \delta_V)^2 \tag{111}$$

where $U = V_\Delta \pm \delta_R$, with the sign depending on the considered MOSFET. By expanding (111) one obtains

$$2I_{s|\delta_R} = \beta U^2 + \delta_\beta U^2 - 2\beta U\delta_V \underbrace{-2U\delta_\beta\delta_V + \beta{\delta_V}^2 + \delta_\beta{\delta_V}^2}_{\approx 0}$$

It is not difficult to verify that, with typical values of β, V_{T}, ..., the nonlinear terms underlined by the under-brace are at least one order of magnitude smaller than the others, so they can be ignored. By accepting this approximation, we see that $I_{s|\delta_R}$ is a Gaussian variable with mean $\beta U^2/2$ and variance $(\sigma_\beta^2 U^4 + 4\beta^2 U^2 \sigma_{V_T}^2)/4$. By remembering the definition of U we deduce

$$2I_{s,1|\delta_R} \sim \mathscr{N}\left(\beta(V_\Delta - \delta_R)^2, \sigma_{\beta,p}^2 (V_\Delta - \delta_R)^4 + 4\beta^2 \sigma_{V,p}^2 (V_\Delta - \delta_R)^2\right)$$
$$2I_{s,2|\delta_R} \sim \mathscr{N}\left(\beta(V_\Delta - \delta_R)^2, \sigma_{\beta,n}^2 (V_\Delta + \delta_R)^4 + 4\beta^2 \sigma_{V,n}^2 (V_\Delta + \delta_R)^2\right)$$

By neglecting δ_R with respect to V_Δ, we can obtain, with some easy algebra, the density of I_0 conditioned by δ_R

$$I_{0|\delta_R} \sim= \mathscr{N}\left(-2\beta V_\Delta \delta_R,\, {I_s}^2 S_X^2\right) \tag{113}$$

By averaging (113) one obtains the following result about the density of $\bar{I}_0$ (see Appendix D).

Let $\phi_{m,\sigma} : \mathbb{R} \to \mathbb{R}$ denote the density of $\mathscr{N}(m, \sigma^2)$ and let, as usual, $\phi(x) = \phi_{0,1}(x)$. Let, for notational convenience, $b = -2\beta V_\Delta$ and $s = S_X I_s$. By conditioning with respect to δ_R we can write

$$f_{I_0}(x) = \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x - bu}{s}\right) f_{\delta_R}(u) du \tag{114}$$

In order to prove the first claim (f_{δ_R} is even $\Rightarrow$ f_{I_0} is even) observe that

$$\begin{aligned}
f_{I_0}(-x) &= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{-x - bu}{s}\right) f_{\delta_R}(u) du \\
&= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x + bu}{s}\right) f_{\delta_R}(u) du && \phi \text{ is even} \\
&= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x - bv}{s}\right) f_{\delta_R}(-v) dv && \text{Change of variable } u = -v \\
&= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x - bv}{s}\right) f_{\delta_R}(v) dv && f_{\delta_R} \text{ is even} \\
&= f_{I_0}(x)
\end{aligned} \tag{115a}$$

In order to prove the second claim (f_{δ_R} is Gaussian $\Rightarrow$ f_{I_0} is Gaussian) observe that

$$\begin{aligned}
f_{I_0}(x) &= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x - bu}{s}\right) \frac{1}{\sigma_R} \phi\left(\frac{u}{\sigma_R}\right) du \\
&= \int_{\mathbb{R}} \frac{1}{s} \phi\left(\frac{x - v}{s}\right) \frac{1}{b\sigma_R} \phi\left(\frac{v}{b\sigma_R}\right) dv \\
&= \phi_{0,s} * \phi_{0,b\sigma_R}(x)
\end{aligned} \tag{116}$$

According to (116), I_0 can be written as the sum of two independent Gaussian variables with zero mean and variances s^2 and $b^2\sigma_R^2$, therefore I_0 is Gaussian. Equations (31) follow by remembering the definitions of s and b and $2I_s = \beta V_\Delta^2$. □

Proof D.2. *Proof of Property 6* Let F_p be the distribution of p_1 and observe that the density f_p of p_1 is symmetric around $1/2$ if and only if

$$F_p(x) = 1 - F_p(1 - x) \tag{117}$$

while the symmetry of f_I implies $F_I(x) = 1 - F_I(-x)$, similarly for F_ξ. Now observe that

$$\begin{aligned}
F_p(x) &= P[p_1 < x] && && (118)\\
&= P[F_\xi(\overline{V}_{\mathrm{eq}}(\bar{I}_0)) < x] && && (119)\\
&= P[\overline{V}_{\mathrm{eq}}(\bar{I}_0) < F_\xi^{-1}(x)] && F_\xi \text{ monotone} && (120)\\
&= P[\bar{I}_0 < h(F_\xi^{-1}(x))] && \overline{V}_{\mathrm{eq}} \text{ monotone} && (121)\\
&= F_I(h(F_\xi^{-1}(x))) && && (122)
\end{aligned}$$

where we used h defined in (33a). Now compute

$$\begin{aligned}
F_p(1 - x) &= F_I(h(F_\xi^{-1}(1 - x))) && && (123)\\
&= F_I(h(-F_\xi^{-1}(x))) && \text{Symmetry of } f_\xi && (124)\\
&= F_I(-h(F_\xi^{-1}(x))) && h \text{ is odd} && (125)\\
&= 1 - F_I(h(F_\xi^{-1}(x))) && \text{Symmetry of } f_\xi && (126)\\
&= 1 - F_p(x) && && (127)
\end{aligned}$$

which proves the thesis. □

A. *Proof of (36)*

By definition of SDF,

$$\begin{aligned}
\mathrm{P}[R(V_{\mathrm{eq}}) < x] &= \mathrm{P}[|\overline{V}_{\mathrm{eq}}| < R^{-1}(x)] \\
&= \int_{-W_x}^{W_x} f_{eq}(y) dy \\
&= \int_{-W_x}^{W_x} \frac{1}{\overline{\sigma}_I} h'(y) \phi(h(y/\overline{\sigma}_I)) dy \\
&= \int_{-h(W_x)/\overline{\sigma}_I}^{h(W_x)/\overline{\sigma}_I} \phi(u) du
\end{aligned} \tag{128}$$

where we exploited the fact that h is odd. Equation (36) follows at once from (128).

Appendix E
Proofs of Section IX

A. *Construction of Table IV*

In constructing Table IV we supposed to protect the PUC outcome using a Reed-Solomon (RS) (n, k, t) code over the Galois field of 2^b elements $\mathbb{F}_{2^b}$. The field size 2^b was chosen in order to minimize the total number of bits, namely, nb under the constraint that the probability of failure, i.e., the probability of having more than t errors, is smaller than η. In order to construct Table IV we first found the best RS code, successively we estimated its relative complexity (in silicon area).

1) Finding the best RS code: Let $P_{\mathrm{bit}} = \mu_{\mathrm{intra}}$ the probability of a PUC bit error. If we collects the PUC bits into b-bit symbols, the probability of error on the symbols becomes

$$P_{\mathrm{sym}} = 1 - (1 - P_{\mathrm{bit}})^b \tag{129}$$

since we have a symbol error as soon as at least one bit is wrong. Note that b must divide N and that the number of information symbols is $k = N/b$.

In order to have a failure we must have more than $t = (n-k)/2$ errors over n symbols. It follows that

$$\begin{aligned} P[\text{failure}] &\leq P[\text{N. errors} > t] \\ &= 1 - P[\text{N. errors} \leq t] \\ &= 1 - F_B(t;n,P_{\text{sym}}) \\ &\approx 1 - \Phi\left(\sqrt{\frac{n}{q}}\left(\frac{t}{n} - P_{\text{sym}}\right)\right) \\ &= \Phi\left(\sqrt{\frac{n}{q}}\left(P_{\text{sym}} - \frac{1}{2}\right) + \frac{k}{2\sqrt{qn}}\right) \end{aligned} \tag{130}$$

where $q = P_{\text{sym}}(1-P_{\text{sym}})$ and $F_B(t;n,P_{\text{sym}})$ is the distribution of a binomial r.v. with n trials and success probability P_{sym}. If $P_{\text{sym}} < 1/2$ the last term of (130) goes to zero when $n \to \infty$, therefore it must exist at least one n that makes (130) not larger than η. We select the smallest among those n. Call it $n_{\text{opt}}(b)$

Remark E.1

If $P_{\text{sym}} > 1/2$ it is not granted that there is a n good enough. This because $t = (n-k)/2 < n/2$, so that any code cannot correct more than $n/2$ errors. If $P_{\text{sym}} > 1/2$ the number of errors for large n is greater (with good probability) than $n/2$. Indeed, if $P_{\text{sym}} > 1/2$, the last term of (130) goes to 1 when $n \to \infty$.

Of course, if one changes b, probability (129) changes and optimal $n_{\text{opt}}(b)$ changes as well. In order to find the best combination, we numerically tried all the possible b and choose the value that minimizes $b\,n_{\text{opt}}(b)$.

2) Estimation of Decoder Relative Cost: In order to get an estimate of the silicon area required by the RS decoder, we refer to [43, Table III] that shows the cost of a syndrome-based RS decoder in terms of number of multipliers, adders, registers, ... required. It is easy to see that the cost of the decoder, in terms of number of components, is linear in t, the maximum number of recoverable errors. Moreover, every component (but the multipliers) require an area that is proportional to b. Therefore, the area required by a RS decoder for a (n,k,t) code over $\mathbb{F}_{2^b}$ is proportional to

$$tb = \frac{nb-kb}{2} \propto nb - kb \tag{131}$$

and the last term is proportional to the number of redundancy bits required. Therefore, the relative cost can be computed by taking the ratio of the number of redundancy bits required for a given combination of PUC, N and η.