A Fast, Modular Method for True Variation-Aware Separatrix Tracing in Nano-Scaled SRAMs

Adam Teman, Member, IEEE and Roman Visotsky

Abstract—As memory density continues to grow in modern systems, accurate analysis of SRAM stability is increasingly important to ensure high yields. Traditional static noise margin metrics fail to capture the dynamic characteristics of SRAM behavior, leading to expensive over-design and disastrous under-design. One of the central components of more accurate dynamic stability analysis is the separatrix; however, its straightforward extraction is extremely time-consuming, and efficient methods are either non-accurate or extremely difficult to implement. In this paper, we propose a novel algorithm for fast separatrix tracing of any given SRAM topology, designed with industry standard transistor models in nano-scaled technologies. The proposed algorithm is applied to both standard 6T SRAM bitcells, as well as previously proposed alternative sub-threshold bitcells, providing up to three orders-of-magnitude speedup, as compared to brute force methods. In addition, for the first time, statistical Monte Carlo separatrix distributions are plotted.

Index Terms—SRAM, Static Noise Margin, Dynamic Noise Margin, Stability Analysis, Control Theory, Separatrix, Phase Portrait, Monte Carlo Simulation

I. INTRODUCTION

As the predominant embedded memory technology in modern integrated circuits (IC), static random access memory (SRAM) has become the primary consumer of both power and silicon area in nano-scaled IC systems. In order to integrate the largest possible amount of SRAM, the traditional six-transistor (6T) SRAM bitcell has closely followed the aggressive scaling of CMOS technologies, and has employed non-standard layouts ("pushed rules") to achieve even higher densities. These trends, which according to the latest ITRS update [1] are expected to continue for at least another decade, are accompanied by an unavoidable increase in process variations, which lead to a loss of data stability [2]. Supply voltage ($V_{DD}$) scaling, targeted at lowering the power consumption of both core logic and SRAM blocks, further impedes the robustness of SRAM bitcells due to reduced noise margins and increased sensitivity to device parameter fluctuations. Furthermore, the high density and chip area consumption of SRAM blocks makes them susceptible to soft-errors (SER) caused by external radiation, as well as other single-event-upsets (SEU), such as those caused by coupled signals, further impairing data integrity [3]. Data stability has always been a central issue in SRAM design, but due to the aforementioned trends, traditional static noise margin (SNM) metrics for ensuring this stability have become impossible to maintain along with the aggressive scaling and density aspirations, and are insufficient for measuring durability in high radiation environments or due to SEUs [2]–[6].

The importance of dynamic noise modeling and stability for both logic and SRAM circuits was first noted over three decades ago [7]. However, dynamic analysis unequivocally took a back seat to SNM criteria for SRAM stability evaluation. Only recently, due to the increasing challenges in achieving high SRAM yields in modern systems, have dynamic noise margins (DNM) come into focus [2]–[5], [8]–[17]. A primary challenge to the widespread adoption of dynamic criteria is the ability to characterize dynamic stability in a simple, straightforward fashion, such as the single number SNM metric [6]. However, as opposed to static margins, which are modeled with a constant DC serial noise source, dynamic noise requires, at the minimum, a two-dimensional noise source, characterized by both amplitude and duration [3]. Therefore, despite several attempts to define a single number metric for DNM [2]–[5], [10]–[12], a toolbox of graphs and metrics is still required to evaluate stability [10]. Of these, the most important is probably the well-
known “separatrix” [9], [11], [13].

In this paper, we present a novel algorithm for fast and accurate separatrix tracing. While the majority of the previously proposed methods [2], [5], [9], [11], [13] can be applied only to a specifically and rigorously modeled 6T bitcell, rely on non-accurate transistor models, require a custom solver, cannot take into consideration statistical device variations, and/or are complex to implement, the proposed algorithm:

- Is modular and topology independent, i.e. it can be applied to any given bi-stable circuit.
- Uses technology provided device models (such as the 40 nm BSIM4 models used for demonstration), providing the highest available accuracy.
- Is operated with any industry standard SPICE solver, such as Cadence Spectre, used in this paper.
- Can take into account global and local statistical variations, as provided with the technology models.
- Is configurable to conveniently trade-off accuracy vs. runtime, as required.
- Is simple to comprehend, implement, and apply, providing the designer with an immediate solution for separatrix tracing, and therefore DNM analysis.

The proposed algorithm is shown to provide a $1000 \times$ speedup over brute-force separatrix tracing with the same resolution. Consequently, these reduced run-times allow application of advanced and complex analysis, previously unavailable, such as statistical Monte Carlo (MC) simulation. In this paper, for the first time, statistical separatrix distributions for SRAM bitcells are displayed. Furthermore, the proposed algorithm is used to trace the separatrix of various non-standard SRAM bitcell circuits, designed for specific applications, such as subthreshold operation [17]–[19].

The rest of the paper is organized as follows: Section II provides an overview of SRAM stability analysis from both a static and dynamic perspective, including its traditional limitations; Section III presents the theory proving the existence of the separatrix and leading to the development of the proposed algorithm; the proposed algorithm is presented in Section IV and shown to correctly trace separatrices under statistical variations in Section V; Section VI presents the separatrices of several alternative SRAM topologies, as traced with the proposed algorithm; and Section VII concludes the paper.

II. SRAM Stability Analysis

Since the widespread adoption of CMOS process technology, the most common circuit for implementing single-ported SRAM cells is the 6T bitcell, shown in Fig. 1(a). The core of this cell is made up of a pair of cross-coupled CMOS inverters (M1/M3 and M4/M6), which form a bi-stable circuit with the stored data represented by the voltages of the complementary internal data nodes (Q and QB). The strong positive feedback between the two inverters results in robust bi-stability, enabling high yields and noise immunity, which have led to the popularity of this topology. Most other SRAM solutions, such
as multi-ported bitcells and bitcells targeted at sub-threshold operation, are also based on this internal core. Therefore, all discussions of SRAM stability start, and often end, with analysis of the 6T cell.

Initial discussions of digital noise margins were intended for logic gates [7], [20]–[22], primarily focusing on the DC operating points of these circuits, as extracted from their voltage transfer characteristics (VTC). With the 6T SRAM core made up of a pair of logic gates, these discussions were famously adapted for SRAM by Seevinck with his groundbreaking paper from 1987 [6]. Seevinck defined the SNM of an SRAM cell as the largest serial voltage noise that can be oppositely applied to the complementary data nodes of the 6T circuit without losing bi-stability. The conceptual setup of this definition is shown in Fig. 1(b), with \( V_n \) representing the magnitude of the DC noise sources. This definition had previously been shown to be mathematically equivalent to several other popular noise margin definitions for logic gates [21]. What caused the breakthrough was Seevinck’s description of a fast, simple, SPICE compatible method for measuring one of these equivalent definitions—the maximum square method.

The maximum square method, shown in Fig. 1(c), defines digital noise margins as the side of the largest square that can fit into the lobes of the butterfly curve plot of a logic gate. Butterfly curves are obtained by plotting the VTCs of odd and even stages of a chain of inverting gates on a single plot. As the cross-coupled inverter structure of the 6T SRAM essentially creates an infinite chain of inverting gates, this plot appropriately shows the DC characteristics of the bitcell. The logic levels of the SRAM cell are represented by the two stable operating points of this graph (the intersecting points of the VTCs, marked as \( P_0 \) and \( P_1 \) in Fig. 1(c)), with the middle intersection (marked as \( P_S \)) a metastable point. Inserting serial DC voltage sources of opposite polarities at Q and QB causes the VTCs to move towards each other, depleting the lobes of the butterfly curves. If a voltage larger than the side of the maximum square is applied, the VTCs cease to intersect at three points, resulting in a loss of bi-stability. Seevinck showed that by applying a rotation transformation to the butterfly curves and subsequently subtracting the two VTCs, the diagonal length of the squares that fit into the lobes is plotted. This enables a simple extraction of the maximum square, using a single DC sweep and a rotation circuit that can be implemented in any standard circuit simulator.

One of the most important characteristics of the Seevinck method for measuring SRAM stability is its compatibility with statistic parametric analysis. The SNM can be measured under MC statistical sampling to provide stability distributions of SRAM cells in light of global process variations and local mismatch. An example distribution is shown in Fig. 2(a) for a 40 nm 6T bitcell, biased at its nominal 1 V supply voltage. This distribution shows that not only is the mean (\( \mu \)) SNM relatively high (in this case 298 mV), but it is non-negative over many standard variations (\( \sigma \)). This analysis is used to ensure high yield for millions of products, each containing millions of bitcells. In addition, by simply changing the node biases in the testbench, read and write operations can be simulated, providing similar distributions for read SNM (RSNM) and write SNM (WSNM), as shown in Fig. 2(b) and 2(c), respectively, for the same circuit. These metrics show

![Fig. 2. SNM Distributions for a 40 nm 6T bitcell with \( V_{DD} = 1 \) V (10 K MC samples of global and local variations at 25°C)](image)
depleted margins in terms of both mean and standard deviation as compared to hold SNM, appropriately tracking the increased sensitivity of the 6T circuit during access operations. In addition, several alternative metrics for WSNM and RSNM have been considered with various advantages and drawbacks as compared to the Seevinck method [12], [23]–[25].

Despite being the de-facto standard for SRAM stability, the SNM metric has many flaws. As a static metric, it fails to capture the fundamental non-linear dynamics of the 6T bitcell, as well as the time-dependent characteristics of read, write, and SEU events. It assumes infinite DC biases in a standalone environment, rendering it unable to capture the effects of the array architecture, peripheral circuitry, and power supplies. Finally, and maybe most significantly, the SNM metric models a non-physical noise source – a serial voltage source with opposite polarities at the data nodes, which on the one hand represents a much more extreme situation that can actually occur, and on the other hand fails to provide a metric that can be captured in terms of an actual noise specification. As a result of these flaws, designing SRAMs exclusively based on the SNM metric leads to overdesign in some cases, and underdesign in others. For hold analysis, the extreme noise modeling leads to overdesign, requiring unnecessary margins that are increasingly hard to achieve in state-of-the-art technologies and under aggressive layouts. Read operations assume infinite access time, causing artificial failures when a state-flip would only occur after the de-assertion of the word line (WL), while not taking into account the time-dependent discharge of the bitlines (BL) or any coupling effects during signal transistions. Write SNM analysis, while also disregarding the dynamic behavior of the system, can further lead to missed write failures due to the assumed infinite write pulse, as the actual access time may be too short to cause a state flip.

Accordingly, several groups have started to focus on developing methods to evaluate the dynamic stability of SRAMs and define dynamic noise margins [2]–[5], [9]–[12], [16]. Dynamic noise analysis most often introduces noise as a time-variant current pulse applied to one or more circuit nodes. This enables modeling physical noise sources, such as particle strikes and coupling. In addition, actual access operations can be simulated by applying read and write operations to the full array architecture. Finally, alternative SRAM implementations, which rely on various dynamic operations or behavior during access or state variation, can be correctly modeled for an impartial comparison.

Whereas dynamic stability analysis provides many levels of improvement and accuracy over static metrics, it also adds inherent complexity. Evaluating the stability of an SRAM topology with a static metric is straightforward, based on the $\mu$ and $\sigma$ of its SNM distribution. However, dynamic stability lacks a universally accepted definition, and furthermore depends on many varying factors for analysis, such as the duration of access pulses and the exact attributes of a noise event. The most well known definitions of SRAM dynamic stability are those described in [26] and [27], providing a failure criterion for an SRAM array under noise or an access operation. While these definitions provide a means to estimate the yield of a given design, operated at a specific voltage, speed, etc., they do little to provide insight into the actual causes of failure, and therefore are ineffective as a tool for improving a design. The additional information provided by DNM measurements can fill this gap.

Several groups have provided definitions for DNM; however, none of these has yet to emerge as a clear cut winner. Lostroh introduced the noise immunity curve (NIC) in his groundbreaking paper from 1979 [7], presenting a shmoo plot showing regions of failure and success for a logic gate in the presence of a trapezoidal noise pulse of a given amplitude and duration. Zurada [28] displayed dynamic (AC) VTCs for logic gates, which led to Ding’s definition of maximum square based dynamic noise margins [4]. However, these definitions do not sufficiently capture the dynamic requirements of SRAM stability, leading to Dong’s definition of DNM [5], as the margin of time between the applied pulse and the minimum time to cause a state flip (in write) or maintain data retention (in read). Wang [12] proposes a similar model for write DNM, modeling the injected current waveforms during a write access and measuring the critical time for a successful write operation. Sharifkhani [10] invokes small signal loop-gain analysis to show the existence of stable states and rigorous state space analysis to test dynamic stability criteria. Wieckowski, et al. [11], define separatrix affinity as a metric for measuring the dynamic stability of SRAM cells, including non-standard topologies, such as their
As a result of the wide array of, often contradicting, definitions for DNM, we have reached the following conclusions:

- In order to carry out dynamic stability analysis, no matter what the chosen DNM definition is, a toolbox of graphs and metrics is required.
- The most important of these tools is the stability boundary between the regions-of-attraction of the stable states. This boundary, better known as the separatrix, is defined, presented, and discussed in the next section.

III. Separatrix Theory

The 6T SRAM bitcell is a dynamic, non-linear, bi-stable system requiring rigorous non-linear control theory for mathematical analysis. The behavior of the circuit can be described by choosing the internal data node voltages (\(V_Q\) and \(V_{QB}\) from Fig. 1(a)) as the state variables and writing the state equations [9]:

\[
\begin{align*}
\frac{\partial V_Q(t)}{\partial t} &= f_f(V_Q, V_{QB}) \\
\frac{\partial V_{QB}(t)}{\partial t} &= f_b(V_Q, V_{QB})
\end{align*}
\]

(1)

where \(f_f(\cdot)\) is the feed-forward non-linear function from Q to QB, and \(f_b(\cdot)\) is the non-linear feedback function from QB to Q. Bi-stability of this state space can easily be demonstrated with the butterfly curve plot of Fig. 1(c). As previously mentioned, these curves plot the DC characteristics of the cross-coupled inverters that comprise the circuit. A corollary to this is that these curves also trace the points at which the derivatives of \(f_f\) and \(f_b\) are zero, better known as the nullcline vectors [11]. Therefore, the intersections of the nullclines are points at which:

\[
\begin{align*}
\frac{\partial V_{QB}(t)}{\partial t} &= f_f(v_q, v_{qb}) = 0 \\
\frac{\partial V_Q(t)}{\partial t} &= f_b(v_q, v_{qb}) = 0
\end{align*}
\]

(2)

such that the points \((v_q, v_{qb})\) are equilibria. Therefore, the SRAM circuit has three equilibria, denoted as \(P_0\), \(P_1\), and \(P_S\).

Evaluation of the stability of a particular equilibrium point is achieved by analyzing the eigenvalues of the Jacobian matrix around the equilibrium point [29]. For the two extreme points, \(P_0\) and \(P_1\), the real part of the Jacobian eigenvalues are both negative, resulting in stable states; namely the bitcell’s data ‘0’ and data ‘1’ states, respectively (designated according to the logic state of node Q). The eigenvalues of the third equilibrium, \(P_S\), on the other hand, have one positive and one negative real parts, and therefore, this is an unstable equilibrium, better known as a saddle point.

Plotting the phase portrait of the 6T circuit provides more insight into these conclusions. The phase portrait of a balanced 6T bitcell is shown in Fig. 3(a), with the three equilibria pointed out. Following the phase vectors in the top left of the plot will bring the state space to \(P_0\), while following the bottom right vectors will proceed to \(P_1\). The boundary between the regions of convergence (RoC) or basins of attraction of \(P_0\) and \(P_1\) is an unstable manifold made up of phase vectors that lead to \(P_S\). Therefore, a theoretical initial condition upon this manifold would result in DC convergence at the unstable equilibrium; however, any small perturbation from this manifold will result in convergence at one of the stable points. This stability boundary is better known as the separatrix of the system [29].

The importance of the separatrix in dynamic stability analysis is straightforward. Given an initial state, for example \(P_1\), following a write ‘1’ operation, an SEU can be simulated. Depending on the amplitude, duration, shape, and attack point of the SEU, the state space of the system will be pushed to a different point on the graph. Following the passing of this noise event, the circuit will either converge back to \(P_1\) or “flip” and converge to \(P_0\), as illustrated in Fig. 3(b). The condition for state retention or loss of data is if the state space following the event crossed the separatrix or not. In a similar fashion, both read and write operations can be modeled as a dynamic noise event, and knowledge of the separatrix enables evaluation of the success of the operation (i.e. retaining or toggling the stored data). Therefore, the majority of the DNM metrics, mentioned in Section II, rely on knowledge of the separatrix for calculation.

The basic method for separatrix tracing is through brute-force state space sampling. In this method, illustrated in Fig. 3(c), the state space is divided into a dense grid, and a transient simulation is run with initial conditions for each grid coordinate. The convergence state of each simulation is evaluated, and accordingly, the RoC of each stable point is mapped, revealing the separatrix as the boundary between the two RoCs. This method has often been shown to be computationally inefficient, with a quadratic relationship between run-time and accuracy [9]. This
all but rules out statistical analysis, a mandatory requirement for high-yield SRAM design.

Accordingly, several methods for the derivation of the separatrix have been proposed in the recent past. Zhang [3] developed a closed-form expression for dynamic stability based on non-linear state space analysis. However, this analysis was based on piece-wise linear device models and assumed the separatrix was a straight line, diagonally separating the state space along the $V_Q = V_{QB}$ vector. Huang’s group proposed an algorithm for separatrix tracing, based on only two transient simulations [5], [9]. They choose an initial point on the separatrix (such as the non-stable equilibrium) and integrate backwards to reveal the entire manifold. This analysis introduces variability into the SRAM model, displaying the effect of device mismatch on the separatrix. However, their algorithm requires the extraction of a modified dynamic system that allows the backward integration. This is a non-trivial operation and may not be adaptable to high-complexity device models, such as BSIM4. In [2], the separatrix is approximated as two line segments based on dynamic analysis and alpha-power law device models. Finally, in [11], a black box method for stability analysis is developed with a method for separatrix tracing without the need for any transient analyses or DC convergence by representing each initial condition in the state-space grid with a circuit model.

While these works propose very efficient methods for separatrix tracing, each with its specific advantages (summarized in Table I for the equivalent of $N \times N$ brute-force accuracy), none of them provide a method as simple and modular as the brute force sampling method, but with short enough run times to enable statistical stability analysis. The following subsection presents our algorithm for separatrix tracing, providing a uniform solution for all of these requirements.

### IV. PROPOSED ALGORITHM

The proposed algorithm is presented below in two parts. Algorithm 1 provides the frame procedure that controls the simulation flow, while Algorithm 2 describes the subroutine for finding the separatrix in proximity to a given point. As inputs, the procedure accepts the netlist of the bitcell, comprised of device models from any technology, in addition to the target supply voltage, and two configuration parameters – the resolution ($x_{step}$) and the error tolerance ($d_{tol}$) of the output.

An initial guess is taken for the first point on the separatrix. This point can have any value, but for symmetry, we show the initial $V_Q$ value to be

<table>
<thead>
<tr>
<th>Method</th>
<th># Sims</th>
<th>Accuracy</th>
<th>Solver</th>
<th>Model</th>
<th>Topology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brute Force</td>
<td>$N^2$ transients</td>
<td>High</td>
<td>Standard</td>
<td>Any</td>
<td>Any</td>
</tr>
<tr>
<td>Zhang [3]</td>
<td>Analytical solution</td>
<td>Low</td>
<td>None</td>
<td>Simple</td>
<td>Only 6T</td>
</tr>
<tr>
<td>Huang [5], [9], [13]</td>
<td>2 DC Sweeps, 2 transients</td>
<td>Low</td>
<td>Custom</td>
<td>Custom</td>
<td>Only specific</td>
</tr>
<tr>
<td>This work</td>
<td>$O(N \log N)$ transients</td>
<td>High</td>
<td>Standard</td>
<td>Any</td>
<td>Any</td>
</tr>
</tbody>
</table>
An additional simulation is run with this new point. ∆y converges to the ‘0’ state (above that are modified according to the convergence of iterative transient simulations with initial conditions on separatrix. This is done by running a series of

Algorithm 1 Modular Separatrix Tracing

Require: Define process technology, bitcell netlist, VDD, resolution (xstep), and tolerance (dtol)
1: Guess initial point, \( p_0^g = (V_{DD}/2, y_0^g) \)
2: Find point \( p_0^u = (V_{DD}/2, y_0^u) \) on separatrix according to Algorithm 2.
3: for \( i = 1; i \leq V_{DD}/2 \times x_{step}; i = i + 1 \) do
4: Set \( x_− = V_{DD}/2 - i \times x_{step}; x_+ = V_{DD}/2 + i \times x_{step} \)
5: Guess: \( p_0^g = (x_0^g, y_0^g); p_0^u = (x_0^u, y_0^u) \), where \( y_0^g \) and \( y_0^u \) are set according to a linear approximation of previous points on the separatrix.
6: Find points \( p_i^u = (x_−, y_i^u) \) and \( p_i^u = (x_+, y_i^u) \) on separatrix according to Algorithm 2.
7: Add points \( p_i^u \) and \( p_i^u \) to Separatrix trace.
8: end for

at the middle of the voltage swing (VDD/2), and lacking additional knowledge about the cell, the arbitrary point \( p_0^g = (x_0, y_0^g) = (V_{DD}/2, V_{DD}/2) \), could be used, where \( p \) denotes a point in the state-space with coordinates \((x, y) = (V_Q, V_{QB})\) and the superscript \( g \) denotes a guess. Using this initial guess, a point with \( p_0^g = (x_0, y_0^g) \) within \( dtol \) of the separatrix is found according to Algorithm 2, with the superscript \( u \) denoting a point on the unstable manifold. Subsequently, \( x_0 \) is both incremented and decremented by \( x_{step} \) and \( y_{step}^g \) are approximated at 45°/225° angles from \( p_0^g \). Algorithm 2 is again invoked for these guesses, and additional points within \( dtol \) of the separatrix are added to the solution vector. This procedure is repeated until the full voltage swing is covered, using linear approximation from the previous points on the separatrix to provide initial guesses at the next increments of \( x_i \).

The subroutine, presented here as Algorithm 2, receives an initial guess, \( p^g = (x, y^g) \), and returns a point \( p^u = (x, y^u) \) that is within \( dtol \) of the separatrix. This is done by running a series of iterative transient simulations with initial conditions that are modified according to the convergence of the previous runs. For example, if the initial guess converges to the ‘0’ state (above the separatrix), the next guess is set with \( y^u = y^g - \Delta Y \), where \( \Delta Y \) is a configurable voltage step (e.g., 0.1 \times V_{DD}). An additional simulation is run with this new point. If the system converges to the opposite state (‘1’), the separatrix is on the interval connecting the two guesses. If, however, the system converges to the same state (‘0’), the separatrix is still below the guess, and therefore, another guess is taken with a further decreased \( y \) value. Once it is known that the separatrix is on a finite interval between two \( y \) values, a simple binary search is applied until the voltage step (\( \Delta Y \)) between subsequent guesses is smaller than \( dtol \). This procedure is illustrated in Fig. 4.

It is important to note that this algorithm is both highly configurable and enables parallelism for distributed calculation. Following an initial run on a given topology, parameters, such as \( y_0^g \), the initial value of \( \Delta Y \), the angle for the guess of the first step, \( x_{step} \), and \( dtol \) can be updated to provide a sufficient result with even lower runtime. The increasing and decreasing searches (\( x_+ \) and \( x_− \), respectively) are independent of each other, and the algorithm can be further modified by providing \( n \) initial guesses and advancing outward in \( 2n \) independent processes. Therefore, the overall procedure can be linearly ac-
Fig. 4. Illustrative example of the operation of Algorithm 2.

Fig. 5. Comparison of separatrix extraction with proposed algorithm to brute force method with the same accuracy.

The proposed algorithm was implemented as a Matlab script running Cadence Spectre simulations on a commercial 40 nm CMOS technology based on BSIM4 models. The sampling density of our algorithm is based on two parameters: resolution and accuracy. The resolution is the distance between samples on one of the axes (e.g. distance between points on the Q axis), and the accuracy is the maximum proximity to the separatrix in direction of the other axis (QB) for each sample. For example, when searching for a point on the separatrix with \( V_Q = 400 \text{ mV} \) with an accuracy of 1% and a 1 V supply, the resulting point (e.g. \((V_Q, V_{QB}) = (0.400, 0.523)\)) is less than 1% \( \cdot V_{DD} = 10 \text{ mV} \) away from the actual separatrix in the vertical (QB) direction. Therefore, for comparison with a 10 mV brute-force sampling grid at \( V_{DD} = 1 \text{ V} \), our algorithm is run with a 10 mV resolution and 1% accuracy.

The results of this comparison are shown in Fig. 5 for three standard 6T cell examples. As expected, the trivial, albeit non-physical, example of a perfectly balanced 6T cell displays a 45° separatrix for both methods. In fact, when only taking global variations into consideration, the resulting separatrix for any symmetric bitcell is very trivial. The more interesting cases are the actual, realistic circuits with some degree of mismatch between devices. For an initial demonstration of the resulting separatrix in such a case, two asymmetric examples are shown – one with 10% size mismatch between the pull-down NMOS devices (M1 and M4) and the other for 10% mismatch between the pull-up PMOS devices (M3 and M6). In these cases, the probability to find exact points on the separatrix is very small, such that the brute-force, reference separatrix points in Fig. 5 represent the border between initial conditions that converged to ‘0’ and those that converged to ‘1’, or in other words, up to 10 mV from the actual separatrix. The results from our algorithm are shown to very closely match the brute-force extraction, despite requiring just 4% of the run time, as summarized in Table II below.

The examples in Fig. 5 emphasize that the true importance of separatrix analysis emerges under mismatch variations, which cause a skew from the trivial 45° manifold. The efficiency of the proposed algorithm provides the ability to extract and display separatrices under statistical mismatch distributions.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Method} & \textbf{# Simulations} & \textbf{Run Time} \\
\hline
Brute Force & 10,201 & 755 min \\
Balanced Bitcell & 101 & 7.48 min. \\
Pull-Down Mismatch & 413 & 30.65 min. \\
Pull-Up Mismatch & 422 & 31.29 min. \\
\hline
\end{tabular}
\caption{Comparison with brute-force extraction}
\end{table}

\footnote{All run-times are provided for single-threaded, serial computations for an impartial comparison. Both the brute-force and the proposed algorithm benefit from a linear speedup through parallelization.}

accelerated with the use of multi-cored architectures.

V. IMPLEMENTATION RESULTS AND STATISTICAL DISTRIBUTIONS

The proposed algorithm was implemented as a Matlab script running Cadence Spectre simulations on a commercial 40 nm CMOS technology based on BSIM4 models. The sampling density of our algorithm is based on two parameters: resolution and accuracy. The resolution is the distance between samples on one of the axes (e.g. distance between points on the Q axis), and the accuracy is the maximum proximity to the separatrix in direction of the other axis (QB) for each sample. For example, when searching for a point
Fig. 6. MC statistical distributions of the separatrices of a standard 40 nm 6T bitcell with 0.5% accuracy. Plots include 1k samples taking into consideration both global variations and local mismatch at room temperature (25°C). Measured runtimes are for serial (single-thread) execution.

thereby providing both insight into circuit behavior and a basis for advanced dynamic stability analysis. Fig. 6(a) displays the separatrices of 1k MC samples of a 40 nm 6T SRAM bitcell with a nominal 1 V supply. This plot shows the significant skew that occurs to the RoCs of the stable states under mismatch variation. This skew increases as the supply voltage is scaled, as can be seen in Fig. 6(b) and Fig. 6(c) for a 650 mV and a 300 mV supply, respectively. These plots were extracted with 1% accuracy and a resolution of 0.1·V_{DD} (i.e., 10 points per separatrix), requiring 1–4 minutes (using a single thread) to extract the separatrix for each sample. For comparison, using the brute-force method would require over 165 minutes per sample or approximately 115 days to extract Fig. 6(a) at a similar resolution and accuracy, for a speedup of over 1000×.

VI. SEPARATRIX OF ALTERNATIVE SRAM TOPOLOGIES

The previous section presented our proposed algorithm and displayed the effects of device mismatch on the characteristics of the separatrix for a standard, symmetric 6T SRAM bitcell. However, as extensively described by Calhoun and Chandrakasan in [32], this traditional circuit is limited to above-threshold supply voltages of around 0.7 V or higher. Accordingly, recent years have shown many proposals for novel SRAM bitcell circuits, most of which are based on the cross-coupled 6T circuit core, but several of which are significantly different. Separatrix extraction is a central and essential component in the analysis of such circuits; however, in the majority of these works, these curves were not presented, and none of these publications included separatrix extraction under local mismatch.

In this section we revisit three of these alternative SRAM topologies, implementing them in a scaled 40 nm process, and plotting an MC distribution of their separatrices at sub-threshold operating voltages. The goal of this discussion is to display the modularity of the proposed algorithm, demonstrating its seamless application to non-standard topologies, while emphasizing the importance of separatrix distribution at scaled nodes and low supply voltages. We argue that all future analysis of dynamic stability should take these distributions into consideration, as they extremely influence the integrity of the results.

Schematics of the three chosen topologies are shown in Fig. 7. These topologies include:

- A Schmitt Trigger based 10T bitcell (Fig. 7(a)) proposed in [30], allowing bit interleaving and differential read, was shown to operate robustly at 160 mV in 0.13 µm CMOS.
- A variation tolerant, single-ended 6T bitcell (Fig. 7(b)) proposed in [31], includes a complementary transmission gate and gated supplies to achieve full functionality at 193 mV in 0.13 µm CMOS.

Note that this speedup is resolution and accuracy dependent. Whereas the Brute-Force method requires an increased number of simulations that is quadratically proportional to increased accuracy, the number of simulations only logarithmically increases for the proposed algorithm, and can be even lower through dynamically updating the search parameters.

VI. SEPARATRIX OF ALTERNATIVE SRAM TOPOLOGIES

The previous section presented our proposed algorithm and displayed the effects of device mismatch on the characteristics of the separatrix for a standard, symmetric 6T SRAM bitcell. However, as extensively described by Calhoun and Chandrakasan in [32], this traditional circuit is limited to above-threshold supply voltages of around 0.7 V or higher. Accordingly, recent years have shown many proposals for novel SRAM bitcell circuits, most of which are based on the cross-coupled 6T circuit core, but several of which are significantly different. Separatrix extraction is a central and essential component in the analysis of such circuits; however, in the majority of these works, these curves were not presented, and none of these publications included separatrix extraction under local mismatch.

In this section we revisit three of these alternative SRAM topologies, implementing them in a scaled 40 nm process, and plotting an MC distribution of their separatrices at sub-threshold operating voltages. The goal of this discussion is to display the modularity of the proposed algorithm, demonstrating its seamless application to non-standard topologies, while emphasizing the importance of separatrix distribution at scaled nodes and low supply voltages. We argue that all future analysis of dynamic stability should take these distributions into consideration, as they extremely influence the integrity of the results.

Schematics of the three chosen topologies are shown in Fig. 7. These topologies include:

- A Schmitt Trigger based 10T bitcell (Fig. 7(a)) proposed in [30], allowing bit interleaving and differential read, was shown to operate robustly at 160 mV in 0.13 µm CMOS.
- A variation tolerant, single-ended 6T bitcell (Fig. 7(b)) proposed in [31], includes a complementary transmission gate and gated supplies to achieve full functionality at 193 mV in 0.13 µm CMOS.

2Note that this speedup is resolution and accuracy dependent. Whereas the Brute-Force method requires an increased number of simulations that is quadratically proportional to increased accuracy, the number of simulations only logarithmically increases for the proposed algorithm, and can be even lower through dynamically updating the search parameters.
A 9T Supply-Feedback (SFSRAM) bitcell (Fig. 7(c)) [18], utilizes internal feedback to gate \( V_{DD} \) for increased write margins and decreased leakage, and was shown to operate robustly at 250 mV in 40 nm CMOS.

These topologies were chosen, as opposed to other popular sub-threshold implementations (such as [32]), as their core device composition differs significantly from the standard 6T topology, already discussed extensively in the previous section. The separatrix spreads of these bitcells, implemented in 40 nm CMOS, are shown in Fig. 8 for 1k MC samples at 0.5% accuracy. The Schmitt Trigger (Fig. 8(a)) and Single-Ended (Fig. 8(b)) bitcells were operated at 200 mV, whereas the SFSRAM (Fig. 8(c)) was operated at 250 mV, as this was its previously presented lower functionality limit. All three plots show clearly separated RoCs, demonstrating bi-stability, albeit with significantly wide distributions, showing susceptibility to SEUs, as well as significantly asymmetric dynamic access behavior.

VII. CONCLUSIONS

This paper proposed a novel algorithm for SRAM separatrix tracing, a mandatory component of dynamic stability analysis. The algorithm was shown to be extremely modular, allowing simple application to any bitcell topology, while enabling simulation with accurate, industry standard technology models. For the first time, Monte Carlo statistical distributions of SRAM separatrices were displayed, emphasizing the importance of these plots, especially at scaled technologies and under low supply voltages. The proposed algorithm provides as much as a 1000\( \times \) speedup, as compared to traditional brute-force separatrix tracing, with many configuration options and parallel computing capabilities.
ACKNOWLEDGMENTS

We would like to thank Mr. O. Sharon for his intensive work in developing the algorithm and Prof. Alexander Fish and Prof. Andreas Burg for their advice and mentoring in carrying out this work.

REFERENCES


