Article

A Complex Model via Phase-Type Distributions to Study Random Telegraph Noise in Resistive Memories

Juan E. Ruiz-Castro [1,*], **Christian Acal** [1], **Ana M. Aguilera** [1] and **Juan B. Roldán** [2]

1 Department of Statistics and O.R. and Math Institute, University of Granada, 18071 Granada, Spain; chracal@ugr.es (C.A.); aaguiler@ugr.es (A.M.A.)
2 Department of Electronics and Computing Technology, University of Granada, 18071 Granada, Spain; jroldan@ugr.es
* Correspondence: jeloy@ugr.es

Abstract: A new stochastic process was developed by considering the internal performance of macro-states in which the sojourn time in each one is phase-type distributed depending on time. The stationary distribution was calculated through matrix-algorithmic methods and multiple interesting measures were worked out. The number of visits distribution to a determine macro-state were analyzed from the respective differential equations and the Laplace transform. The mean number of visits to a macro-state between any two times was given. The results were implemented computationally and were successfully applied to study random telegraph noise (RTN) in resistive memories. RTN is an important concern in resistive random access memory (RRAM) operation. On one hand, it could limit some of the technological applications of these devices; on the other hand, RTN can be used for the physical characterization. Therefore, an in-depth statistical analysis to model the behavior of these devices is of essential importance.

1. Introduction

In several fields, such as computing and electronics engineering, it is of great interest to analyze complex devices with several macro-states that evolve by time. It is usual to consider Markov processes for this analysis, but in multiple occasions, the spent times in each macro-state are not exponentially distributed. In this context, a new approach is to consider that the spent time in each macro-state is phase-type (PH) distributed. In this case, it is assumed that each macro-state is composed of internal performance states that have a Markovian behavior. One interesting aspect is that the modeling of the stochastic process can only occur when the macro-state process can be observed. For this new process, the Markovianity is lost.

A phase-type distribution is defined as the distribution of the absorption time in an absorbing Markov chain. These probability distributions constitute a class of distributions on the positive real axis, which seems to strike a balance between generality and tractability, thanks to its good properties. This class of distributions was introduced by Neuts [1,2] and allows the modeling of complex problems in an algorithmic and computational way. It has been widely applied in fields such as engineering to model complex systems [3–5], queueing theory [6], risk theory [7] and electronics engineering [8].

The good properties of PH distributions due to their appealing probabilistic arguments constitute their main feature of being mathematically tractable. Several well-known probability distributions, e.g., exponential, Erlang, Erlang generalized, hyper-geometric, and Coxian distributions, among others, are particular cases of PH. One of the most important PH properties is that whatever nonnegative probability distribution can be approximated as much as desired through a PH, accounting for the fact that the PH class is dense in the set of probability distributions on the nonnegative half-line [9]. It allows

the consideration of general distributions through PH approximations. Exact solutions to many complex problems in stochastic modeling can be obtained either explicitly or numerically by using matrix-analytic methods. The main features of PH distributions are described in depth in [10].

A macro-state stochastic process is built here to model the behavior of different random telegraph noise (RTN) signals. Given that the behavior of the different macro-states (levels) is not exponentially distributed, we assume each level to be composed of multiple internal phases that can be related with other levels. We show that if the internal process, by considering internal states inside the macro-states, has a Markovian behavior, the process between levels is not. For the latter, the sojourn time in each level is phase-type distributed depending on the initial time. This fact is of great interest and gives us more information about the device's internal process.

Advanced statistical techniques are key tools to model complex physical and engineering problems in many different areas of expertise. In this context, a new statistical methodology is presented; we will concentrate in the study of resistive memories [11,12], also known as resistive random access memories (RRAMs), a subgroup of a wide class of electron devices named memristors [13]. They constitute a promising technology with great potential in several applications in the integrated circuit realm.

Circuit design is a tough task because of the high number of electronic components included in integrated circuits, therefore, electronic design automation (EDA) tools are essential. These software tools need compact models that represent the physics of the devices in order to fulfil their role in aiding circuit design. Although there are many works in compact modeling in the RRAM literature [14,15], there is a lot to be done, in particular in certain areas related to variability and noise. Variability is essential in modeling due to the RRAM operation's inherent stochasticity, since the physics is linked to random physical mechanisms [16,17]. Another issue of great importance in these devices is noise. Among the different types of noise, random telegraph noise is of great concern [18]. The disturbances produced by one or several traps (physical active defects within the dielectric) inside the conductive filament or close to it alter charge transport mechanisms and consequently, current fluctuations can occur (Figure 1) that lead to RTN [14,15,19,20]. This noise can affect the correct device operation in applications linked to memory-cell arrays and neuromorphic hardware [14,16,21], posing important hurdles to the use of this technology in highly scaled integrated circuits. Nevertheless, RTN fluctuations can also be beneficial, for instance, when used as entropy sources in random number generators, an employment of most interest in cryptography [22,23].

The application in this issue is addressed in the current work, in particular with the statistical description of RTN signals in RRAMs. This study, in addition to physically characterizing the devices, can be used for compact modeling and, therefore, as explained above, for developing the software tools needed for circuit design. Accounting for the aforementioned intrinsic stochasticity of these devices, the choice of a correct statistical strategy to attack this problem is essential in the analysis. In this respect, we use the PH distributions, which have already been employed to depict some facets of RRAM variability [24]; nonetheless, as far as we know, they have not been used in RTN analysis.

In our study, we deal with devices whose structure is based on the $Ni/HfO_2/Si$ stack. The fabrication details, as well as their electric characterization, are given in [25]. RTN measurements were recorded in the time domain by means of an HP-4155B semiconductor parameter analyzer (Figure 1). The current fluctuation between levels (see the marked levels in red in Figure 1), ΔI, and the spent time in each of the levels are key parameters to analyze. The associated time with the current low level is known as emission time (τ_e), which is linked to the time the active defect keeps a "captured" electron trapped until it releases it (through this time, it is said that the defect is occupied). On the other hand, the capture time (τ_c) represents the time taken to capture an electron by an empty defect.
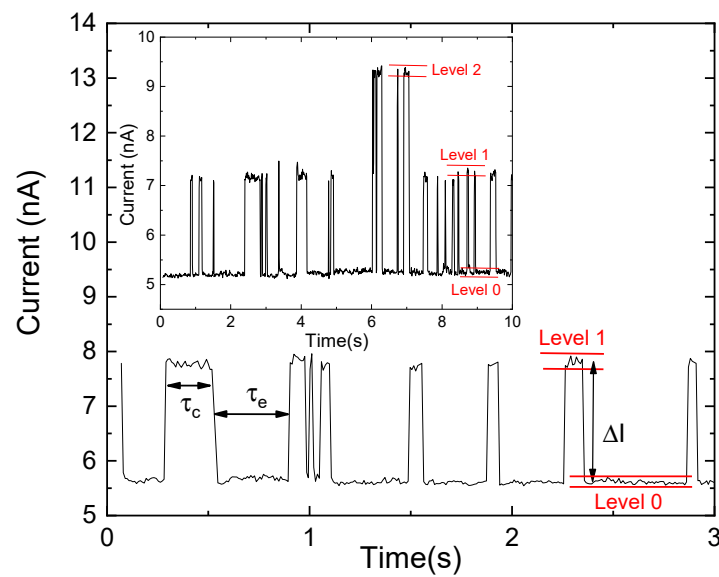
Figure 1. Current versus time trace for a Ni/HfO$_2$/Si device in the high resistance state HRS, in which the RTN noise can clearly be seen. A two-level signal is shown; the two different current levels are marked with the corresponding current thresholds. Another RTN trace is shown in the inset, measured for the same device. Three current levels are seen in this case.

The manuscript is organized as follows: in Section 2, the proposed statistical procedure is described; in Section 3, we deepen on the measures associated; and in Section 4, the number of visits to a determined macro-state is presented. The parameter estimation is unfolded in Section 5, and the application of the methodology developed here in Section 6. Finally, conclusions related to the main contributions of this work are presented in Section 7.

2. Statistical Methodology

Different RTN signals have been employed here, such as the ones shown in Figure 1. We have determined the number of current levels in order to be able to calculate the emission and capture times for a certain time interval. The current thresholds (red lines in Figure 1) were chosen to allow the extraction algorithm to calculate the time intervals corresponding to the levels defined previously.

In order to explain these issues, a new model is first built in transient and stationary regimes.

2.1. The Model

We assume a stochastic process $\{X(t); t \geq 0\}$ with macro-state space $E = \{1, 2, \ldots, \mathbf{r}\}$. Each macro-state $\mathbf{k}$, level $\mathbf{k}$, is composed of n_k internal phases or states denoted as i_h^k for $h = 1, \ldots, n_k$. We assume that the device internal performance is governed by a Markov process, $\{J(t); t \geq 0\}$, with state space $E = \{i_1^1, \ldots, i_{n_1}^1, i_1^2, \ldots, i_{n_2}^2, \ldots, i_1^r, \ldots, i_{n_r}^r\}$, and with an initial distribution $\mathbf{\theta}$, and the following generator matrix expressed by blocks:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \cdots & \mathbf{Q}_{1k} & \cdots & \mathbf{Q}_{1r} \\ & \ddots & & & \\ \vdots & & \mathbf{Q}_{kk} & & \vdots \\ & & & \ddots & \\ \mathbf{Q}_{r1} & \cdots & \mathbf{Q}_{rk} & \cdots & \mathbf{Q}_{rr} \end{pmatrix}$$

where the matrix block $\mathbf{Q}_{ij}$ contains the transition intensities between the states of the macro-states from $\mathbf{i}$ to $\mathbf{j}$.

Throughout this work, we will denote $\mathbf{Q}_k$ as the matrix with zeros except the column matrix block k of $\mathbf{Q}$, and $\mathbf{Q}_{-k}$ as the matrix $\mathbf{Q}$ with zeros in the *k-th* block column. These matrices give the transition intensities of macro-state **k** and of any macro-state different to **k,** respectively.

Thus:

$$
\mathbf{Q}_k = \begin{pmatrix} 0 & \dots & 0 & \mathbf{Q}_{1k} & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & \mathbf{Q}_{rk} & 0 & \dots & 0 \end{pmatrix}
$$

$$
\mathbf{Q}_{-k} = \begin{pmatrix} \mathbf{Q}_{11} & \dots & \mathbf{Q}_{1,k-1} & 0 & \mathbf{Q}_{1,k+1} & \dots & \mathbf{Q}_{1r} \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{Q}_{r1} & \dots & \mathbf{Q}_{r,k-1} & 0 & \mathbf{Q}_{r,k+1} & \dots & \mathbf{Q}_{rr} \end{pmatrix}
$$

It can be seen that the following equality holds: $\mathbf{Q} = \mathbf{Q}_k + \mathbf{Q}_{-k}$.

Given the Q-matrix of the Markov process, the transition probability matrix for the process $\{J(t); t \geq 0\}$ is given by $\mathbf{P}(t) = \exp\{\mathbf{Q}t\}$. From this and from the initial distribution θ, the transient distribution at time t is given by $(P\{X(t) = i_1^1\}, P\{X(t) = i_2^1\}, \dots, P\{X(t) = i_{n_r}^r\}) = \mathbf{a}(t) = \theta \cdot \mathbf{P}(t)$. The order of this vector is $1 \times \sum_{i=1}^{r} n_i$.

The transition probabilities for the process $\{X(t); t \geq 0\}$ can be obtained from the transition probabilities of the Markov process $\{J(t); t \geq 0\}$. If it is denoted as $\mathbf{H}(\cdot, \cdot)$ where the transition between macro-states $\mathbf{i} \rightarrow \mathbf{j}$ is given by:

$$
h_{\mathbf{ij}}(s, t) = P\{X(t) = \mathbf{j} | X(s) = \mathbf{i}\} = \frac{\sum_{h \in \mathbf{j}} \sum_{k \in \mathbf{i}} P\{J(t) = h | J(s) = k\} P\{J(s) = k\}}{\sum_{k \in \mathbf{i}} P\{J(s) = k\}}
$$

$$
= \frac{\sum_{k \in \mathbf{i}} \left[a_k(s) \sum_{h \in \mathbf{j}} p_{kh}(t - s) \right]}{\sum_{k \in \mathbf{i}} a_k(s)}.
$$

This transition probability can be expressed in a matrix-algorithmic form as follows. Denoting by $\mathbf{A_k}$ a matrix of zeros with order $\sum_{i=1}^{r} n_i \times \sum_{i=1}^{r} n_i$ except the block corresponding to the macro-state **k**, which is the identity matrix, then:

$$
h_{\mathbf{ij}}(s, t) = P\{X(t) = \mathbf{j} | X(s) = \mathbf{i}\} = \frac{\mathbf{a}(s)\mathbf{A_i}\mathbf{P}(t - s)\mathbf{A_j} \cdot \mathbf{e}}{\mathbf{a}(s)\mathbf{A_i} \cdot \mathbf{e}}
$$

where $\mathbf{a}(t)\mathbf{A_i}\mathbf{e}$ is the probability of being in the macro-state **i** at time t, with **e** being a column vector of ones with an appropriate order. A similar reasoning can be done with the embedded jumping probability.

Note that the matrix **H** is a stochastic matrix, and therefore it is a transition matrix of a Markov process. However, an important remark is that the Markov process related to the matrix **H** does not correspond to the process defined in this section. In fact, this process is non-homogeneous and not Markovian.

2.2. Stationary Distribution

The stationary distribution for the process $\{X(t); t \geq 0\}$ can be calculated by blocks from the stationary distribution for the process $\{J(t); t \geq 0\}$. It is well known that this distribution verifies the balance equations $\pi\mathbf{Q} = 0$ and the normalization equation $\pi\mathbf{e} = 1$. We denote π_k as the block of π corresponding to the macro-state **k**. It is denoted as $\pi_k\mathbf{e}$ (**k** = 1,...,**r**) to the stationary distribution for the macro-state **k** of the process $\{X(t); t \geq 0\}$.

The stationary distribution has been worked out by blocks in order to reduce the computational cost, according to the macro-states, and applying matrix-analytic methods.

As stated above, the stationary distribution verifies the following system:

$$
\begin{aligned}
\boldsymbol{\pi}_1\mathbf{Q}_{11} + \boldsymbol{\pi}_2\mathbf{Q}_{21} + \boldsymbol{\pi}_3\mathbf{Q}_{31} + \cdots + \boldsymbol{\pi}_r\mathbf{Q}_{r1} &= 0 \\
\boldsymbol{\pi}_1\mathbf{Q}_{12} + \boldsymbol{\pi}_2\mathbf{Q}_{22} + \boldsymbol{\pi}_3\mathbf{Q}_{32} + \cdots + \boldsymbol{\pi}_r\mathbf{Q}_{r2} &= 0 \\
\boldsymbol{\pi}_1\mathbf{Q}_{13} + \boldsymbol{\pi}_2\mathbf{Q}_{23} + \boldsymbol{\pi}_3\mathbf{Q}_{33} + \cdots + \boldsymbol{\pi}_r\mathbf{Q}_{r3} &= 0 \\
&\cdots \\
\boldsymbol{\pi}_1\mathbf{Q}_{1r} + \boldsymbol{\pi}_2\mathbf{Q}_{2r} + \boldsymbol{\pi}_3\mathbf{Q}_{3r} + \cdots + \boldsymbol{\pi}_r\mathbf{Q}_{rr} &= 0 \\
\boldsymbol{\pi}\cdot\mathbf{e} &= 1.
\end{aligned}
$$

It has been solved by using matrix-analytics methods, and the solution is given by

$$
\boldsymbol{\pi}_j = -\sum_{i=1}^{j-1}\boldsymbol{\pi}_i\mathbf{R}_{ij}^{r-j+1}; j = 2,\ldots,r
$$

where $\mathbf{R}_{ij}^1 = \mathbf{Q}_{ij}\mathbf{Q}_{jj}^{-1}$ for any $i,j = 1,\ldots,r$ with $i \neq j$

$$
\mathbf{R}_{ij}^{r-j+1} = \left(\mathbf{R}_{ij}^1 + \mathbf{H}_{ij}^{r-j+1}\right)\left(\mathbf{I} + \mathbf{H}_{jj}^{r-j+1}\right)^{-1} \text{ for } 1 \le i < j < r \text{ being}
$$

$$
\begin{aligned}
\mathbf{H}_{ij}^{r-j+1} = \ &- \sum_{j<u_1\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 j}^1 + \sum_{j<u_1<u_2\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{R}_{u_2 j}^1 \\
&- \sum_{j<u_1<u_2<u_3\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{R}_{u_2 u_3}^{r-u_3+1}\mathbf{R}_{u_3 j}^1 \\
&+ \sum_{j<u_1<u_2<u_3<u_4\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{R}_{u_2 u_3}^{r-u_3+1}\mathbf{R}_{u_3 u_4}^{r-u_4+1}\mathbf{R}_{u_4 j}^1 - \ldots \quad \text{for } 1 \le i < j < r \\
&\pm \sum_{j<u_1<u_2<\cdots<u_{r-j}\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\cdots\mathbf{R}_{u_{r-3}u_{r-2}}^{r-u_{r-2}+1}\mathbf{R}_{u_{r-2} j}^1 \\
&\mp I_{\{j<r-1 \text{ or } i\neq j\}}\mathbf{R}_{1,j+1}^{r-j}\mathbf{R}_{j+1,j+2}^{r-j}\cdots\mathbf{R}_{r-1,r}^1\mathbf{R}_{rj}^1.
\end{aligned}
$$

The vector $\boldsymbol{\pi}_1$ is worked out as follows:

$$
\boldsymbol{\pi}_1 = (1,0)\left[\mathbf{B}\big|(\mathbf{I} + \mathbf{H}_{11}^r)^*\right]^{-1}
$$

Where, given a matrix $\mathbf{A}$, $\mathbf{A}^*$ is the matrix $\mathbf{A}$ without the last column, and:

$$
\begin{aligned}
\mathbf{B} = \ &\mathbf{e} - \sum_{j<u_1\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{e} + \sum_{j<u_1<u_2\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{e} - \sum_{j<u_1<u_2<u_3\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{R}_{u_2 u_3}^{r-u_3+1}\mathbf{e} \\
&+ \sum_{j<u_1<u_2<u_3<u_4\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\mathbf{R}_{u_2 u_3}^{r-u_3+1}\mathbf{R}_{u_3 u_4}^{r-u_4+1}\mathbf{e} - \ldots \\
&\pm \sum_{j<u_1<u_2<\cdots<u_{r-j}\le r}\mathbf{R}_{iu_1}^{r-u_1+1}\mathbf{R}_{u_1 u_2}^{r-u_2+1}\cdots\mathbf{R}_{u_{r-3}u_{r-2}}^{r-u_{r-2}+1}\mathbf{e}.
\end{aligned}
$$

Finally, the stationary distribution of the process $\{X(t); \ t \ge 0\}$ is given by $(\boldsymbol{\pi}_1\cdot\mathbf{e}, \boldsymbol{\pi}_2\cdot\mathbf{e}, \cdots, \boldsymbol{\pi}_r\cdot\mathbf{e})$.

Algorithm to calculate the stationary distribution

Step 1. For $i,j = 1,\ldots,r$ and $i \neq j$
   Compute $\mathbf{R}_{ij}^1$
Step 2. For $j = r-1,\ldots,2$ {
   For $i = 1,\ldots,j$ {
   Compute $\mathbf{H}_{ij}^{r-j+1}$}
   For $i = 1,\ldots,j$ {
   Compute $\mathbf{R}_{ij}^{r-j+1}$ ($i \neq j$)}}
Step 3. Compute $\mathbf{H}_{11}^r$ and B
Step 4. Compute $\boldsymbol{\pi}_1$
Step 5. Compute $\boldsymbol{\pi}_2,\ldots,\boldsymbol{\pi}_r$ and $\boldsymbol{\pi}_1\cdot\mathbf{e},\ldots,\boldsymbol{\pi}_r\cdot\mathbf{e}$

Example. Case r = 4

Step 1.

$$\mathbf{R}_{12}^1 = \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}; \mathbf{R}_{13}^1 = \mathbf{Q}_{13}\mathbf{Q}_{33}^{-1}; \mathbf{R}_{14}^1 = \mathbf{Q}_{14}\mathbf{Q}_{44}^{-1}$$
$$\mathbf{R}_{21}^1 = \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}; \mathbf{R}_{23}^1 = \mathbf{Q}_{23}\mathbf{Q}_{33}^{-1}; \mathbf{R}_{24}^1 = \mathbf{Q}_{24}\mathbf{Q}_{44}^{-1}$$
$$\mathbf{R}_{31}^1 = \mathbf{Q}_{31}\mathbf{Q}_{11}^{-1}; \mathbf{R}_{32}^1 = \mathbf{Q}_{32}\mathbf{Q}_{22}^{-1}; \mathbf{R}_{34}^1 = \mathbf{Q}_{34}\mathbf{Q}_{44}^{-1}$$

Step 2.

$$\mathbf{H}_{13}^2 = -\mathbf{R}_{14}^1\mathbf{R}_{43}^1; \mathbf{H}_{23}^2 = -\mathbf{R}_{24}^1\mathbf{R}_{43}^1; \mathbf{H}_{33}^2 = -\mathbf{R}_{34}^1\mathbf{R}_{43}^1$$
$$\mathbf{R}_{13}^2 = \left(\mathbf{R}_{13}^1 + \mathbf{H}_{13}^2\right)\left(\mathbf{I} + \mathbf{H}_{33}^2\right)^{-1}; \mathbf{R}_{23}^2 = \left(\mathbf{R}_{23}^1 + \mathbf{H}_{23}^2\right)\left(\mathbf{I} + \mathbf{H}_{33}^2\right)^{-1}$$
$$\mathbf{H}_{12}^3 = -\mathbf{R}_{13}^2\mathbf{R}_{32}^1 - \mathbf{R}_{14}^1\mathbf{R}_{42}^1 + \mathbf{R}_{13}^2\mathbf{R}_{34}^1\mathbf{R}_{42}^1$$
$$\mathbf{H}_{22}^3 = -\mathbf{R}_{23}^2\mathbf{R}_{32}^1 - \mathbf{R}_{24}^1\mathbf{R}_{42}^1 + \mathbf{R}_{23}^2\mathbf{R}_{34}^1\mathbf{R}_{42}^1$$
$$\mathbf{R}_{12}^3 = \left(\mathbf{R}_{12}^1 + \mathbf{H}_{12}^3\right)\left(\mathbf{I} + \mathbf{H}_{22}^3\right)^{-1}$$

Step 3.

$$\mathbf{H}_{11}^4 = -\mathbf{R}_{12}^3\mathbf{R}_{21}^1 - \mathbf{R}_{13}^2\mathbf{R}_{32}^1 - \mathbf{R}_{14}^1\mathbf{R}_{41}^1 + \mathbf{R}_{12}^3\mathbf{R}_{23}^2\mathbf{R}_{31}^1 + \mathbf{R}_{12}^3\mathbf{R}_{24}^1\mathbf{R}_{31}^1 + \mathbf{R}_{13}^2\mathbf{R}_{34}^1\mathbf{R}_{41}^1 - \mathbf{R}_{12}^3\mathbf{R}_{23}^2\mathbf{R}_{34}^1\mathbf{R}_{41}^1$$
$$\mathbf{B} = \mathbf{e} - \mathbf{R}_{12}^3 \cdot \mathbf{e} - \mathbf{R}_{13}^2 \cdot \mathbf{e} - \mathbf{R}_{14}^1 \cdot \mathbf{e} + \mathbf{R}_{12}^3\mathbf{R}_{23}^2 \cdot \mathbf{e} + \mathbf{R}_{12}^3\mathbf{R}_{24}^1 \cdot \mathbf{e} + \mathbf{R}_{13}^2\mathbf{R}_{34}^1 \cdot \mathbf{e} - \mathbf{R}_{12}^3\mathbf{R}_{23}^2\mathbf{R}_{34}^1 \cdot \mathbf{e}$$

Step 4.

$$\boldsymbol{\pi}_1 = (1,0)\left[\mathbf{B}\middle|\left(\mathbf{I} + \mathbf{H}_{11}^4\right)^*\right]^{-1}$$

Step 5.

$$\boldsymbol{\pi}_2 = -\boldsymbol{\pi}_1\mathbf{R}_{12}^3; \boldsymbol{\pi}_3 = -\boldsymbol{\pi}_1\mathbf{R}_{13}^2 - \boldsymbol{\pi}_2\mathbf{R}_{23}^2; \boldsymbol{\pi}_4 = -\boldsymbol{\pi}_1\mathbf{R}_{14}^1 - \boldsymbol{\pi}_2\mathbf{R}_{24}^1 - \boldsymbol{\pi}_3\mathbf{R}_{34}^1$$

3. Associated Measures

Several associated measures, such as the sojourn time in each level (macro-state) and the number of visits to each macro-state by time, have been worked out.

3.1. Sojourn Time Phase-Type Distribution

One interesting aspect to address at this point is the probability distribution of the sojourn time in a macro-state. It is well known that for the Markov process *J(t)*, the sojourn time in any state is exponentially distributed. For the stochastic process *X(t)*, it is different.

We denote $T(X(s))$ as the random sojourn time in macro-state $X(s)$ from time *s*. If the macro-state is known at time *s*, then:

$$P\{T(X(s)) > t | X(s) = \mathbf{i}\} = \frac{\mathbf{a}(s)\mathbf{A_i} \cdot \exp\{\mathbf{A_i}\mathbf{Q}\mathbf{A_i}t\} \cdot \mathbf{e}}{\mathbf{a}(s)\mathbf{A_i} \cdot \mathbf{e}}$$

Therefore, the probability distribution of $T(X(s))|X(s) = \mathbf{i}$, for any *s* is phase-type distributed for any *s* with representation $\left(\frac{\mathbf{a}(s)\mathbf{A_i}}{\mathbf{a}(s)\mathbf{A_i}\cdot\mathbf{e}}, \mathbf{A_i}\mathbf{Q}\mathbf{A_i}\right)$. Obviously, this distribution is the same as $\left(\frac{\mathbf{b}(s)}{\mathbf{b}(s)\cdot\mathbf{e}}, \mathbf{Q_{ii}}\right)$, with $\mathbf{b}(s)$ being the vector $\boldsymbol{\theta}\mathbf{P}(s)$ restricted to the states of the macro-state $\mathbf{i}$.

If the process $\{X(t); t \geq 0\}$ has reached the stationary regime, then, if the process is in macro-state $\mathbf{i}$, the probability distribution of the sojourn time is PH with representation $\left(\frac{\boldsymbol{\pi}\mathbf{A_i}}{\boldsymbol{\pi}\mathbf{A_i}\cdot\mathbf{e}}, \mathbf{A_i}\mathbf{Q}\mathbf{A_i}\right)$.

3.2. First Step Time

It is well-known that the first step-time distribution for the process $\{J(t); t \geq 0\}$ from state *l* (out of macro-state $\mathbf{k}$) to a macro-state $\mathbf{k}$ is phase-type distributed with representation $(\boldsymbol{\theta}_l, \mathbf{Q}_{-2\mathbf{k}})$ with $\boldsymbol{\theta}_l = (0,\ldots,0,1,0,\ldots,0)$, where the value 1 corresponds to state *l* and $\mathbf{Q}_{-2\mathbf{k}}$ is the matrix $\mathbf{Q}$ with row and column blocks of zeros corresponding to the macro-state $\mathbf{k}$. If we denote $T_{\mathbf{h}}(s, \mathbf{k})$ as the first step time from macro state $\mathbf{h}$, at time *s*, to macro-state $\mathbf{k}$, then:

$$P(T_{\mathbf{h}}(s, \mathbf{k}) > t) = \frac{\sum\limits_{j \in \mathbf{k}}\sum\limits_{i \in \mathbf{h}} P(T(j) > t | J(s) = i)P(J(s) = i)}{\sum\limits_{i \in \mathbf{h}} P(J(s) = i)}$$
$$= \frac{\mathbf{a}(s)\mathbf{A_h} \cdot \exp\{\mathbf{Q}_{-2\mathbf{k}}t\} \cdot \mathbf{e}}{\mathbf{a}(s)\mathbf{A_h}\mathbf{e}}.$$

4. Number of Visits to a Macro-State

One interesting measure is the number of visits to a determined state between any two times s and t. We denote $N_k(s,t)$ as the number of visits to the macro-state **k** from time s up to time t. We denote $\mathbf{p}_k(n,s,t)$ as the matrix whose (i,j) element is:

$$
\begin{aligned}
[\mathbf{p}_k(n,s,t)]_{ij} &= P\{N_k(s,t) = n, J(t) = j | J(s) = i\} \\
&= P\{N_k(t-s) = n, J(t-s) = j | J(0) = i\} \\
&= [\mathbf{p}_k(n,t-s)]_{ij}.
\end{aligned}
$$

The probability matrix verifies the following differential equations:

$$
\mathbf{p}'_k(n,t) = \mathbf{p}_k(n,t)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right) + \mathbf{p}_k(n-1,t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right); n \geq 1,
$$

with initial condition $\mathbf{p}_k(n,0) = 0$, with $\widetilde{\mathbf{Q}}_{kk}$ a matrix of zeros, with the same order than $\mathbf{Q}$, except for the matrix block $\mathbf{Q}_{kk}$.

For $n = 0$:

$$
\mathbf{p}'_k(0,t) = \mathbf{p}_k(0,t)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)
$$

$$
\mathbf{p}_k(0,0) = \mathbf{I}
$$

where, from the last two expressions we have that:

$$
\mathbf{p}_k(0,t) = \exp\left\{\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)t\right\}
$$

To obtain the probability matrix, we use Laplace transforms. It is well known that given a locally integrable function $f(t)$, its Laplace transform is defined as $f^*(u) = \int_0^\infty e^{-ut}f(t)dt$. Therefore:

$$
u\mathbf{p}_k^*(0,u) - \mathbf{I} = \mathbf{p}_k^*(0,u)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)
$$

$$
u\mathbf{p}_k^*(n,u) = \mathbf{p}_k^*(n,u)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right) + \mathbf{p}_k^*(n-1,u)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right); n \geq 1
$$

Then:

$$
\mathbf{p}_k^*(0,u) = \left[u\mathbf{I} - \left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)\right]^{-1}
$$

$$
\mathbf{p}_k^*(n,u) = \mathbf{p}_k^*(n-1,u)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)\left[u\mathbf{I} - \left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)\right]^{-1}; n \geq 1
$$

Thus, it can be proved that:

$$
\mathbf{p}_k^*(n,u) = \left[u\mathbf{I} - \left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)\right]^{-1}\mathbf{A}^n(u); n \geq 0
$$

with A(u) being:

$$
\mathbf{A}(u) = \left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)\left[u\mathbf{I} - \left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right)\right]^{-1}
$$

Taking the inverse Laplace transform, the function $\mathbf{p}_k(n,t)$ is achieved, and for the non-homogeneous Markov process $\{X(t); t \geq 0\}$, we have:

$$
P\{N_{\mathbf{k}}(n,s,t) = n, X(t) = \mathbf{j} | X(s) = \mathbf{i}\} = \frac{\theta\mathbf{P}(s)\mathbf{A_i}\mathbf{p_k}(n,t-s)\mathbf{A_j} \cdot \mathbf{e}}{\theta\mathbf{P}(s)\mathbf{A_i} \cdot \mathbf{e}}
$$

Therefore:

$$
P\{N_k(s,t) = n\} = \theta \cdot \mathbf{P}(s) \cdot \mathbf{p}_k(n,t-s) \cdot \mathbf{e}, \text{ for } n = 0, 1, 2, \ldots
$$

4.1. Expected Number of Visits to Determine a Macro-State

The mean number of visits to the macro-state $\mathbf{k}$ can be obtained as follows. This measure is given by:

$$E\{N_k(t)\} = \boldsymbol{\theta} \cdot \sum_{n=0}^{\infty} n \cdot \mathbf{p}_k(n,t) \cdot \mathbf{e} = \boldsymbol{\theta} \cdot \mathbf{M}_k(t) \cdot \mathbf{e}$$

From the differential equations above, we have:

$$\sum_{n=1}^{\infty} n \cdot \mathbf{p}_k'(n,t) = \sum_{n=1}^{\infty} n \cdot \mathbf{p}_k(n,t)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right) + \sum_{n=1}^{\infty} n \cdot \mathbf{p}_k(n-1,t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)$$

$$\sum_{n=1}^{\infty} n \cdot \mathbf{p}_k'(n,t) = \sum_{n=1}^{\infty} n \cdot \mathbf{p}_k(n,t)\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right) + \sum_{n=1}^{\infty} n \cdot \mathbf{p}_k(n,t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right) + \mathbf{P}(t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)$$

$$\mathbf{M}_k'(t) = \mathbf{M}_k(t)\left[\left(\mathbf{Q}_{-k} + \widetilde{\mathbf{Q}}_{kk}\right) + \left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)\right] + \mathbf{P}(t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)$$

$$\mathbf{M}_k'(t) = \mathbf{M}_k(t)\mathbf{Q} + \mathbf{P}(t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right)$$

Given that $\mathbf{Q}$ is a conservative matrix, then:

$$\mathbf{M}_k'(t) \cdot \mathbf{e} = \mathbf{P}(t)\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right) \cdot \mathbf{e}$$

with initial condition $\mathbf{M}_k(0) = 0$ if the initial state is not considered, or $\mathbf{M}_k(0) = \mathbf{A}_k$ if the initial state is counted.

Therefore, for the first and second cases, we have, respectively:

$$E[N_k(t)] = \boldsymbol{\theta} \cdot \mathbf{M}_k(t) \cdot \mathbf{e} = \boldsymbol{\theta} \cdot \int_0^t \mathbf{P}(u)du\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right) \cdot \mathbf{e}$$

or

$$E[N_k(t)] = \boldsymbol{\theta} \cdot \mathbf{A}_k \cdot \mathbf{e} + \boldsymbol{\theta} \cdot \int_0^t \mathbf{P}(u)du\left(\mathbf{Q}_k - \widetilde{\mathbf{Q}}_{kk}\right) \cdot \mathbf{e}$$

5. Parameter Estimation

The likelihood function when the exact change time between macro-states is known is achieved. For a device l we observe m_l transition times, denoted as:

$$0 = t_0^l, t_1^l, t_2^l, \ldots, t_{m_l-1}^l, t_{m_l}^l$$

where the last time is a complete or a censoring time.

The time t_a^l corresponds to the transition from x_a^l to x_{a+1}^l. These macro-states are: $x_0^l, x_1^l, x_2^l, \ldots, x_{m_l-1}^l, x_{m_l}^l$.

This device contributes to the likelihood function: likelihood function:

$$L_l = \boldsymbol{\alpha}_{x_0^l}\left[\prod_{a=0}^{m_l-1} e^{\mathbf{Q}_{x_a^l x_a^l}\left(t_{a+1}^l - t_a^l\right)}\left(\mathbf{Q}_{x_a^l x_{a+1}^l}\right)^{\tau_l}\right]\mathbf{e}$$

where τ_l is 0 if the last time is a censoring time, and is 1 otherwise.

When m independent devices are considered:

$$L = \prod_{l=1}^{m} L_l = \prod_{l=1}^{m} \boldsymbol{\alpha}_{x_0^l}\left[\prod_{a=0}^{m_l-1} e^{\mathbf{Q}_{x_a^l x_a^l}\left(t_{a+1}^l - t_a^l\right)}\left(\mathbf{Q}_{x_a^l x_{a+1}^l}\right)^{\tau_l}\right]\mathbf{e}$$

This function is maximized by considering that $\mathbf{Q}_{aa}$ is a square sub-stochastic matrix whose main diagonal is negative and the rest are positive values, and $\mathbf{Q}_{ab}$ a matrix with positive elements with $\sum_{b=1}^{r} \mathbf{Q}_{ab}\mathbf{e} = 0$ for any a and b.

Other log-likelihood function is given from the transition probabilities of the process

$$\log L = \sum_{l=1}^{m} \sum_{a=0}^{m_l - 1} \log\left(h_{x_a^l x_{a+1}^l}\left(t_a^l, t_{a+1}^l\right)\right).$$

6. Application of the Developed Methodology

We have made use of measurements of current-time traces in a unipolar RRAM. The RTN signals had an easy pattern, so different features of the device could be studied. Current traces showed, for instance, the number of current levels in the signal, and also the frequency at which each of these levels was active. The developed methodology allows the modeling of the internal states under an approach based on hidden Markov processes that produces the observed output (RTN signal measured). By analyzing the signals shown in Figure 1, we are able to determine these states and characterize them, probabilistically speaking, in order to understand their nature. It is possible to reproduce similar signals in case they are needed for a circuit simulation or physical characterization of the traps that help to generate the signals.

Specifically, we are going to analyze the behavior of several different current-time traces (RTN25, RTN26, RTN27) that show RTN for the described devices in the introduction. In addition, a long (more than three hours with millions of measured data) RTN current-time trace was measured and used here for the same device. Due to the fact that all signals come from devices with the same characteristics of fabrication, the difference between them lies in the applied voltage that produces the variations of electric current. Naturally, the measurement time for the long RTN was different. These measurements have previously been characterized in different ways [19,20]; however, in this new approach, the internal Markov chain that leads to the observed data set was identified, and the model, whose methodology is developed in this work, was estimated. In particular, in this application, it was shown that the short signals (RTN25, RTN26, and RTN27) had a Markovian internal behavior, whereas the developed new methodology was applied to the long RTN trace for its modeling.
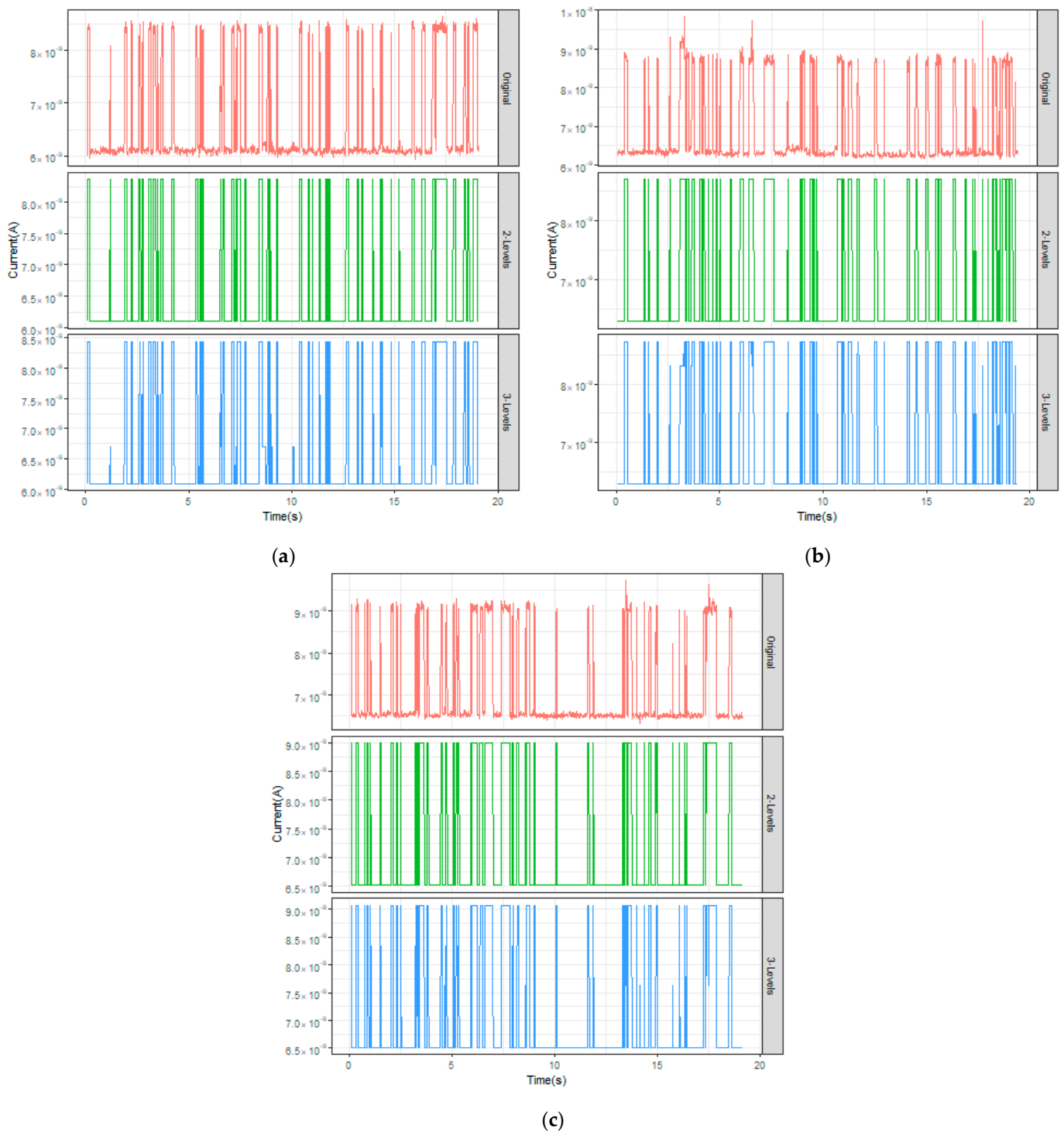
6.1. Series RTN25-26-27

As stated above, the signals RTN25, RTN26, and RTN27 were produced by devices whose structure is based on the Ni/HfO$_2$/Si stack. Nevertheless, different voltages were applied: 0.34 volts, 0.35 volts, and 0.36 volts, respectively. On the other hand, the measurement time was similar for each one of these series.

Hidden Markov Models and the Latent Markov chain for RTN25/26/27 traces

To study the number of possible latent levels hidden into the signals, we considered hidden Markov models (HMMs) [26,27]. Whereas in simple Markov models, the state of the device is visible to the observer at each time, in HMMs, only the output of the device is visible, while the state that leads to a determined output is hidden. Each state is associated with a set of transition probabilities (one per each state) defining how likely it is for the system, being in a given state at a given instant of time, to switch to another possible state (including a transition to the same state) at the successive instants of time.

We analyzed different data sets for RTN25, RTN26, RTN27, and the hidden states from the corresponding signals were discriminated. The best fit was achieved for latent levels 2 and 3 for RTN25, RTN26, RTN27. Figure 2 shows the original observed RTN signals and the corresponding latent levels given by the model for both cases. This analysis was performed using the depmixS4 package of R-cran [28].

Figure 2. Fit obtained with the HMM for multiple latent states for the different RTN signals under study: (**a**) RTN25; (**b**) RTN26; (**c**) RTN27.

The RTN25/26/27 series was analyzed. The proportional number of times that the chain is in the latent states for the different proposed models is shown in Table 1.

Table 1. Proportional number of times that each chain was in the latent states.

Signal	Model	Latent State 1	Latent State 2	Latent State 3
RTN25	2 latent states	0.784	—	0.216
	3 latent states	0.762	0.037	0.201
RTN26	2 latent states	0.756	—	0.244
	3 latent states	0.753	0.0305	0.2165
RTN27	2 latent states	0.7665	—	0.2335
	3 latent states	0.7575	0.0255	0.2170

For each model, the continuous Markov chain associated to the latent states was estimated, and the corresponding stationary distribution was achieved. This is shown in Table 2.

Table 2. Stationary distribution for the different traces.

Signal	Model	Latent State 1	Latent State 2	Latent State 3
RTN25	2 latent states	0.7974	—	0.2026
	3 latent states	0.7847	0.0185	0.1968
RTN26	2 latent states	0.7708	—	0.2292
	3 latent states	0.7746	0.0165	0.2089
RTN27	2 latent states	0.7736	—	0.2264
	3 latent states	0.7785	0.0026	0.219

These values can be interpreted as the proportional time that the device is in each latent state in the stationary regime for the embedded continuous Markov chain. We observed that these values were very close to zero for the RTN25/26/27 devices. We tested this fact, and do not think it can be rejected.

Therefore, we assumed that the internal performance of the devices behaved as a Markov model with two latent states for the RTN25/26/27 traces. The Markov chains for two latent states for the different cases were estimated. The exponentiality of the sojourn time in each state cannot be rejected according to the p-values obtained by means of a Kolmogorov–Smirnov test, and the expected number of visits up to a certain time as explained in Section 4.1 also was estimated. These estimations are shown in Table 3.

Table 3. Expected number of visits to each level for the different traces.

Signal	Level	$t = 5$	$t = 10$	$t = 15$	$t = 20$
RTN25	Level 1	12.3497	23.8609	35.3721	46.8833
	Level 2	11.5523	23.0635	34.5747	46.0859
RTN26	Level 1	12.8198	24.8162	36.8127	48.8092
	Level 2	12.0490	24.0455	36.0419	48.0384
RTN27	Level 1	11.5645	22.5305	33.4965	44.4626
	Level 2	11.7909	22.7569	33.7229	44.6890

6.2. Long RTN Trace

A similar exploratory analysis for this trace was carried out. This signal was supplied with 0.5 V, and its behavior was measured for more than three hours. To study the number of possible latent levels hidden in the signal, we again considered hidden Markov chains. The best fit was achieved for latent levels 3 and 4 for this long RTN trace. Figure 3 shows the original observed signal (a time interval of the whole signal was considered for the sake of computing feasibility) and the corresponding latent levels given by the model for both cases. This analysis also was performed using the depmixS4 package of R-cran.
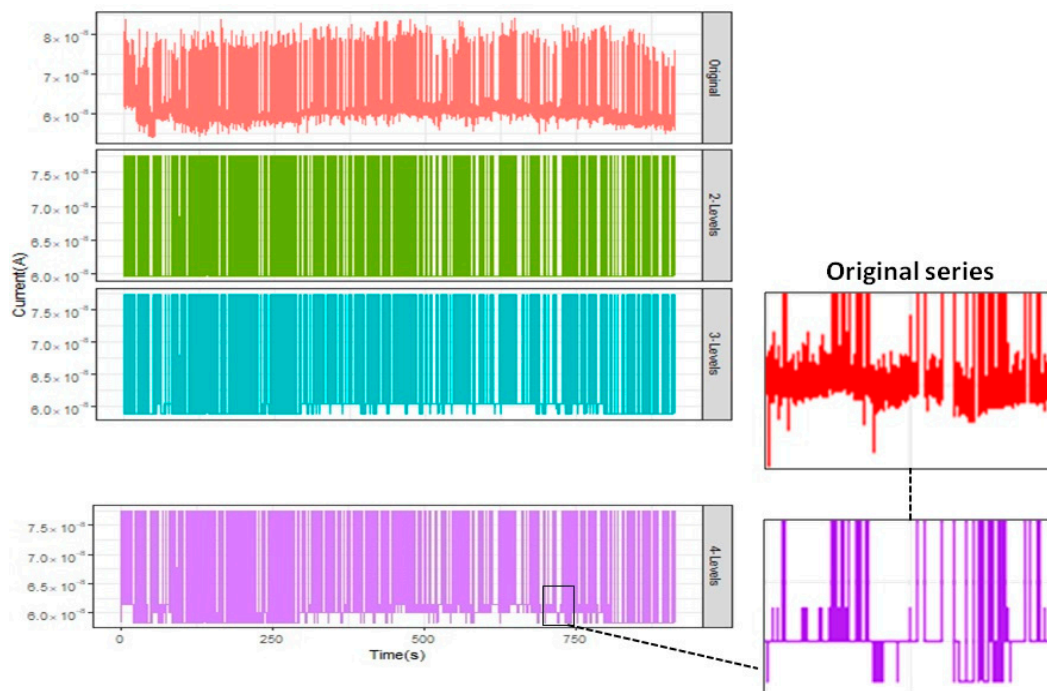
Figure 3. Fit obtained with the HMM for multiple latent states for the RTN different signals under study.

We have focused on the four latent states. If the HMM is considered, the proportional time that the process was in each latent state was 0.3243048, 0.1219429, 0.1635429, and 0.3902095. If more latent states were assumed, negligible proportions would appear. Therefore, four different latent states were assumed. Before applying the new model to the data set, a statistical analysis was performed that considered classical techniques. For these latent states, the exponentiality of the sojourn time was tested and rejected through a Kolmogorov–Smirnov test for any latent state by obtaining the following p-values: 0.00027, 0.0045, 0.0000, and 0.0001, respectively. Next, we studied whether these times could be described as PH distributions. After multiple analyses, the best PH distributions for the latent states had 2, 2, 4, and 3 internal states, respectively. The structures of these PH distributions were generalized Coxian/Erlang distributions. Then, the internal behavior for each latent state passed across multiple internal states in a sequential way, one-by-one. The Anderson–Darling test was applied to test the goodness of fit, and obtained *p*-values of 0.8972, 0.4405, 0.0752, and 0.9876, respectively, for each latent state.

Generator of the Internal Markov Process

Given the previous analysis, we have observed that the sojourn time distribution in each macro-state (latent state) was phase-type distributed with a sequential degradation (generalized Coxian degradation). That is, the macro-states 1, 2 were composed of two phases (internal states), the macro-state 3 of four phases, and the macro-state 4 of three states. Thus, we assume the sojourn time in a macro-state i (level i) is PH distributed with representation $(\boldsymbol{\alpha}_i, \mathbf{T}_i)$ with the following structure:

Macro-state 1	Macro-state 2
$\boldsymbol{\alpha}_1 = \left(\alpha_1^1, 1 - \alpha_1^1\right)$	$\boldsymbol{\alpha}_2 = \left(\alpha_1^2, 1 - \alpha_1^2\right)$
$\mathbf{T}_1 = \begin{pmatrix} -t_{12}^1 & t_{12}^1 \\ 0 & -t_{13}^1 \end{pmatrix}$	$\mathbf{T}_2 = \begin{pmatrix} -t_{12}^2 & t_{12}^2 \\ 0 & -t_{13}^2 \end{pmatrix}$
$\mathbf{T}_1^0 = \left(0, t_{13}^1\right)'$	$\mathbf{T}_2^0 = \left(0, t_{13}^2\right)'$

Macro-state 3	Macro-state 4
$\boldsymbol{\alpha}_3 = \left(\alpha_1^3, \alpha_2^3, \alpha_3^3, 1 - \alpha_1^3 - \alpha_2^3 - \alpha_3^3\right)$	$\boldsymbol{\alpha}_4 = \left(\alpha_1^4, \alpha_2^4, 1 - \alpha_1^4 - \alpha_2^4\right)$
$\mathbf{T}_3 = \begin{pmatrix} -t_{12}^3 & t_{12}^3 & 0 & 0 \\ 0 & -t_{23}^3 & t_{23}^3 & 0 \\ 0 & 0 & -t_{34}^3 & t_{34}^3 \\ 0 & 0 & 0 & -t_{35}^3 \end{pmatrix}$	$\mathbf{T}_4 = \begin{pmatrix} -t_{12}^4 & t_{12}^4 & 0 \\ 0 & -t_{23}^4 & t_{23}^4 \\ 0 & 0 & -t_{34}^4 \end{pmatrix}$
$\mathbf{T}_3^0 = \left(0, 0, 0, t_{35}^3\right)'$	$\mathbf{T}_4^0 = \left(0, 0, 0, t_{34}^4\right)'$

We assume that when the device leaves the macro-state **i**, it goes to macro-state **j** with a probability of p_{ij}, where $p_{ii} = 0$ and $p_{14} = 1 - \sum_{i=1}^{3} p_{ij}$, and the sojourn time in this new macro-state begins with the corresponding initial distribution $\boldsymbol{\alpha}_j$.

Thus, the behavior of the device is governed by a stochastic process $\{X(t); t \geq 0\}$ with an embedded Markov process $\{J(t); t \geq 0\}$, as described in Section 2.1. The macro-state space is $\{$**1, 2, 3, 4**$\}$ and the state space $\{1 = i_1^1, 2 = i_2^1; 3 = i_1^2, 4 = i_2^2; 5 = i_1^3, 6 = i_2^3, 7 = i_3^3, 8 = i_4^3; 9 = i_1^4, 10 = i_2^4, 11 = i_3^4\}$.

In this case, the matrix blocks are $\mathbf{Q}_{ii} = \mathbf{T}_i$ and $\mathbf{Q}_{ii} = p_{ij}\mathbf{T}_i^0 \otimes \boldsymbol{\alpha}_j$ for $i, j = 1,..., r$ and $I \neq j$. Therefore, the matrix generator is given by:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T}_1 & p_{12}\mathbf{T}_1^0 \otimes \boldsymbol{\alpha}_2 & p_{13}\mathbf{T}_1^0 \otimes \boldsymbol{\alpha}_3 & p_{14}\mathbf{T}_1^0 \otimes \boldsymbol{\alpha}_4 \\ p_{21}\mathbf{T}_2^0 \otimes \boldsymbol{\alpha}_1 & \mathbf{T}_2 & p_{23}\mathbf{T}_2^0 \otimes \boldsymbol{\alpha}_3 & p_{24}\mathbf{T}_2^0 \otimes \boldsymbol{\alpha}_4 \\ p_{31}\mathbf{T}_3^0 \otimes \boldsymbol{\alpha}_1 & p_{32}\mathbf{T}_3^0 \otimes \boldsymbol{\alpha}_2 & \mathbf{T}_3 & p_{34}\mathbf{T}_3^0 \otimes \boldsymbol{\alpha}_4 \\ p_{41}\mathbf{T}_4^0 \otimes \boldsymbol{\alpha}_1 & p_{42}\mathbf{T}_4^0 \otimes \boldsymbol{\alpha}_2 & p_{43}\mathbf{T}_4^0 \otimes \boldsymbol{\alpha}_3 & \mathbf{T}_4 \end{pmatrix}$$

The parameters were estimated by considering the likelihood function built in Section 5. The estimated parameters with a value of the logL = -4066.118120 are:

$$\mathbf{P} = \begin{pmatrix} 0 & 0.6667 & 0.0407 & 0.2926 \\ 0.3870 & 0 & 0.1969 & 0.4161 \\ 0.0296 & 0.2238 & 0 & 0.7466 \\ 0.1605 & 0.3395 & 0.5 & 0 \end{pmatrix}$$

$$\boldsymbol{\alpha}_1 = (0.5730374, 0.4269626); \mathbf{T}_1 = \begin{pmatrix} -0.6790043 & 0.6790043 \\ 0 & -4.1343018 \end{pmatrix}$$

$$\boldsymbol{\alpha}_2 = (0.4699825, 0.5300175); \mathbf{T}_2 = \begin{pmatrix} -3.249849 & 3.249849 \\ 0 & -10.426533 \end{pmatrix}$$

$$\boldsymbol{\alpha}_3 = (0.0741494, 0.4258142, 0.5000364, 0)$$

$$\mathbf{T}_3 = \begin{pmatrix} -0.5533471 & 0.5533471 & 0 & 0 \\ 0 & -2.2755462 & 2.2755462 & 0 \\ 0 & 0 & -29.535695 & 29.535695 \\ 0 & 0 & 0 & -242.7465 \end{pmatrix}$$

$$\boldsymbol{\alpha}_4 = (0.3593538, 0.6406462, 0.0000); \mathbf{T}_4 = \begin{pmatrix} -0.964899 & 0.964899 & 0 \\ 0 & -4.112655 & 4.112655 \\ 0 & 0 & -28.638854 \end{pmatrix}$$

Therefore, the generator of the Markov process $\{J(t); t \geq 0\}$ is:

$$
\hat{\mathbf{Q}} = \begin{pmatrix}
-0.6790 & 0.6790 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -4.1343 & 1.2954 & 1.4609 & 0.0125 & 0.0717 & 0.0841 & 0 & 0.4347 & 0.7750 & 0 \\
0 & 0 & -3.2498 & 3.2498 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2.3122 & 1.7228 & 0 & -10.4265 & 0.1522 & 0.8742 & 1.0266 & 0 & 1.5590 & 2.7794 & 0 \\
0 & 0 & 0 & 0 & -0.5533 & 0.5533 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -2.2755 & 2.2755 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -29.5357 & 29.5357 & 0 & 0 & 0 \\
4.1174 & 3.0679 & 25.5326 & 28.7941 & 0 & 0 & 0 & -242.7465 & 65.1273 & 116.1072 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.9649 & 0.9649 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4.1127 & 4.1127 \\
2.6340 & 1.9625 & 4.5696 & 5.1533 & 1.0618 & 6.0974 & 7.1602 & 0 & 0 & 0 & -28.6389
\end{pmatrix}
$$

with initial distribution $\hat{\theta} = (0, 0, 0, 0, 0.0741494, 0.4258142, 0.5000364, 0, 0, 0, 0)$.

The stationary distribution was estimated for the process $\{X(t); t \geq 0\}$ (given in Table 4) from Section 2.2. It can be interpreted as the proportional time in each macro-state (long-run).

Table 4. Stationary distribution for the process $\{X(t); t \geq 0\}$.

	Level 1	Level 2	Level 3	Level 4
Selected interval in the long RTN trace	0.3273	0.1197	0.1612	0.3919

Finally, the mean number of visits to each macro-state up to a certain time t was obtained following Section 4.1. It is shown in Table 5.

Table 5. Expected number of visits up to a certain time for different times.

Time	Level 1	Level 2	Level 3	Level 4
$t = 50$	16.0207	25.0716	20.3974	30.0827
$t = 100$	31.0837	49.9364	40.9591	60.1877
$t = 200$	61.1925	99.6404	82.0612	120.3666
$t = 500$	151.4018	248.5475	205.1981	300.6553

7. Conclusions

A real problem motivates the construction of a new stochastic process that accounts for the internal performance of different macro-states by considering that it follows an internal Markovian behavior. We have shown that the homogeneity and Markovianity was lost in the developed new macro-state model. The sojourn time in each macro-state was phase-type distributed depending on the initial observed time. The stationary distribution was calculated through matrix-algorithmic methods, and the distribution of the number of visits to a determined macro-state between any two different times was calculated using a Laplace transform. The mean number of visits depending on times was worked out explicitly.

The new developed methodology enables the modeling of complex systems in an algorithmic way by solving classic calculus problems. Thanks to the proposed development, the results and measures were worked out in an easier way and could be interpreted. Matrix analysis and Laplace transform techniques were used to determine the properties of the model, and algorithms to obtain quantitative results were provided. Given that everything was carried out algorithmically, is the model was implemented computationally and was successfully applied to study different random telegraph noise signals measured for unipolar resistive memories.

Resistive memory random telegraph noise signals were analyzed in depth in order to characterize them from a probabilistic point of view. These signals are essential, since they can pose a limit in the performance of certain applications; in addition, this type of noise can be good as an entropy source in the design of random number generators for cryptography. From the proposed model, we showed that a latent state of a resistive memory RTN long signal is composed of multiple internal states. Of course, the applications of a phase-type model, as given in this work, are not restricted to the RRAM context.

Author Contributions: J.B.R. was in charge of the physical theoretical explanation and its conclusions. J.E.R.-C., C.A., and A.M.A. developed all the statistical theory, computationally implemented the methodology, and obtained the results shown. All authors contributed equally to the writing part of the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PH	Phase-type
RTN	Random Telegraph Noise
RRAM	Resistive Random Access Memory
EDA	Electronic Design Automation
HMM	Hidden Markov Model

References

1. Neuts, M.F. Probability distributions of phase type. In *Liber Amicorum Professor Emeritus H. Florin*; University of Louvain: Louvain-la-Neuve, Belgium, 1975.
2. Neuts, M.F. *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach*; John Hopkins University Press: Baltimore, MD, USA, 1981.
3. Ruiz-Castro, J.E. Complex multi-state systems modelled through Marked Markovian Arrival Processes. *Eur. J. Oper. Res.* **2016**, *252*, 852–865. [CrossRef]
4. Ruiz-Castro, J.E. A complex multi-state k-out-of-n: G system with preventive maintenance and loss of units. *Reliab. Eng. Syst. Safe.* **2020**, *197*, 106797. [CrossRef]
5. Ruiz-Castro, J.E.; Dawabsha, M. A multi-state warm standby system with preventive maintenance, loss of units and an indeterminate multiple number of repairpersons. *Comput. Ind. Eng.* **2020**, *142*, 106348. [CrossRef]
6. Artalejo, J.R.; Chakravarthy, S.R. Algorithmic Analysis of the MAP/PH/1 Retrial Queue. *Top* **2006**, *14*, 293–332. [CrossRef]
7. Asmussen, S.; Bladt, M. Phase-type distribution and risk processes with state-dependent premiums. *Scand. Actuar. J.* **1996**, *1*, 19–36. [CrossRef]
8. Ruiz-Castro, J.E.; Acal, C.; Aguilera, A.M.; Aguilera-Morillo, M.C.; Roldán, J.B. Linear-Phase-Type probability modelling of functional PCA with applications to resistive memories. *Math. Comput. Simulat.* **2020**. [CrossRef]
9. Asmussen, S. *Ruin Probabilities*; World Scientific: Singapore, 2000.
10. He, Q.M. *Fundamentals of Matrix-Analytic Methods*; Springer: New York, NY, USA, 2014.
11. Carboni, R.; Ielmini, D. Stochastic Memory Devices for Security and Computing. *Adv. Electron. Mater.* **2019**, *5*, 1900198. [CrossRef]
12. Aldana, S.; Roldán, J.B.; García-Fernández, P.; Suñe, J.; Romero-Zaliz, R.; Jiménez-Molinos, F.; Long, S.; Gómez-Campos, F.; Liu, M. An in-depth description of bipolar resistive switching in Cu/HfOx/Pt devices, a 3D Kinetic Monte Carlo simulation approach. *J. Appl. Phys.* **2018**, *123*, 154501. [CrossRef]
13. Chual, L.O. Memristor-the missing circuit element. *IEEE Trans. Circuits Syst.* **1971**, *18*, 507–519.
14. Puglisi, F.M.; Zagni, N.; Larcher, L.; Pavan, P. Random Telegraph Noise in Resistive Random Access Memories: Compact Modeling and Advanced Circuit Design. *IEEE Trans. Electron Dev.* **2018**, *65*, 2964–2972. [CrossRef]
15. Puglisi, F.M.; Larcher, L.; Padovani, A.; Pavan, P. A Complete Statistical Investigation of RTN in HfO2-Based RRAM in High Resistive State. *IEEE Trans. Electron Dev.* **2015**, *62*, 2606–2613. [CrossRef]

16. Alonso, F.J.; Maldonado, D.; Aguilera, A.M.; Roldán, J.B. Memristor variability and stochastic physical properties modeling from a multivariate time series approach. *Chaos Soliton Fract.* **2021**, *143*, 110461. [CrossRef]
17. Aguilera-Morillo, M.C.; Aguilera, A.M.; Jiménez-Molinos, F.; Roldán, J.B. Stochastic modeling of Random Access Memories reset transitions. *Math. Comput. Simulat.* **2019**, *159*, 197–209. [CrossRef]
18. Simoen, E.; Claeys, C. *Random Telegraph Signals in Semiconductor Devices*; IOP Publishing: Bristol, UK, 2017.
19. González-Cordero, G.; González, M.B.; Jiménez-Molinos, F.; Campabadal, F.; Roldán, J.B. New method to analyze random telegraph signals in resistive random access memories. *J. Vac. Sci. Technol. B* **2019**, *37*, 012203. [CrossRef]
20. González-Cordero, G.; González, M.B.; Morell, A.; Jiménez-Molinos, F.; Campabadal, F.; Roldán, J.B. Neural network based analysis of Random Telegraph Noise in Resistive Random Access Memories. *Semicond. Sci. Tech.* **2020**, *35*, 025021. [CrossRef]
21. Grasser, T. *Noise in Nanoscale Semiconductor Devices*; Springer: Cham, Switzerland, 2020.
22. Wei, Z.; Katoh, Y.; Ogasahara, S.; Yoshimoto, Y.; Kawai, K.; Ikeda, Y.; Eriguchi, K.; Ohmori, K.; Yoneda, S. True random number generator using current difference based on a fractional stochastic model in 40-nm embedded ReRAM. In Proceedings of the 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 3–7 December 2016.
23. Chen, X.; Wang, L.; Li, B.; Wang, Y.; Li, X.; Liu, Y.; Yang, H. Modeling Random Telegraph Noise as a Randomness Source and its Application in True Random Number Generation. *IEEE Trans. Comput. Aided Des.* **2016**, *35*, 1435–1448. [CrossRef]
24. Acal, C.; Ruiz-Castro, J.E.; Aguilera, A.M.; Jiménez-Molinos, F.; Roldán, J.B. Phase-type distributions for studying variability in resistive memories. *J. Comput. Appl. Math.* **2019**, *345*, 23–32. [CrossRef]
25. González, M.B.; Rafí, J.M.; Beldarrain, O.; Zabala, M.; Campabadal, F. Analysis of the Switching Variability in Ni/HfO2 -Based RRAM Devices. *IEEE Trans. Device Mat. Reliab.* **2014**, *14*, 769–771.
26. Rabiner, L.R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **1989**, *77*, 257–286. [CrossRef]
27. Puglisi, F.M.; Pavan, P. Factorial Hidden Markov Model analysis of Random Telegraph Noise in Resistive Random Access Memories. *ECTI Trans. Electr. Eng. Electron. Commun.* **2014**, *12*, 24–29.
28. Visser, I.; Speekenbrink, M. depmixS4: An R Package for Hidden Markov Models. *J. Stat. Softw.* **2010**, *36*, 1–21. [CrossRef]