Fundamental roles of extreme-value distributions in dielectric breakdown and memory applications (minimum-value versus maximum-value statistics)

Ernest Wu1*, Takashi Ando2, Baozhen Li3, Richard Southwick4, and James Stathis2

1IBM Research Division, Essex Junction, VT, 05452, United States of America
2IBM System Group, Essex Junction, VT, 05452, United States of America
3IBM Research Division, Yorktown Heights, NY, 10598, United States of America
4IBM Research Division, Albany, 12203, United States of America

E-mail: eywu@us.ibm.com

Received January 6, 2020; revised March 31, 2020; accepted May 8, 2020; published online June 1, 2020

In this paper, a thorough review of minimum- and maximum-value statistical distributions is provided. Using the Weibull model (statistics of minima) and the Gumbel model (statistics of maxima) along with the respective scaling properties of their scale-factor and distribution-percentile with device area (size), the application of these two types of extreme-value distributions to dielectric breakdown (BD) and memory operations is discussed. In the case of dielectric breakdown, the minimum-value distribution (the Weibull model) provides an indispensable tool to establish a valid voltage/field acceleration model from experimental perspectives. On the other hand, recent advances in the introduction of maximum-value distribution (the Gumbel model) overcomes the shortcomings of the conventional practice of adopting the normal distribution to characterize memory functional operations and provides much needed mathematical rigor and physical insight particularly for the rapid growing field of resistive random-access memory devices. © 2020 The Japan Society of Applied Physics

1. Introduction

Extreme-value theory refers to a mathematical framework which describes statistical phenomena involving the extreme values of random variables, also termed the statistics of extremes by Gumbel.1,2 Unlike widely used normal or log-normal distributions to characterize statistical phenomena based on the central-limit theorem, the statistics of extreme values forms a unique family of statistical distributions which can be either minimum-value or maximum-value distributions.3 While extreme-value statistics has found a wide range of applications such as in mechanical engineering, earth science, and radioactive emission as well as flood and seismic analysis, its fundamental role in microelectronics research and industry is often under-appreciated or less understood. It is a common bias that a statistical distribution merely gives a mathematical description of repetitive experimental measurements. In reality, a distribution choice fundamentally reflects the physical underpinnings of a natural statistical process versus another. Moreover, extreme-value distributions themselves alone are not adequate to capture experimental data involving different sample areas (or sizes). A spatial distribution for the occurrence of random events of interest in device areas is often required for scaling and projection purpose. In this review paper, we will examine the fundamental characteristics of both minimum-value and maximum-value distributions as well as the relevant scaling properties of scale-factor and distribution-percentile with device area. Then, we will discuss the applications of the minimum-value distribution to dielectric breakdown. Particularly, we will show how the preservation of this basic extreme-value property of dielectric breakdown can provide an unbiased tool to distinguish one voltage(field) acceleration model from another, based on either experimental interpretation or theoretical formulation of such models. Next, we will discuss the recent advances in applying the maximum-value distributions to memory operations in Sect. 4. An important distinction between dielectric breakdown and memory operation is that dielectric BD refers to failure probability statistics3 with the logarithmic random variables whereas resistive random-access memory (RRAM) conductance4,5 and static random-access memory (SRAM) minimum failure voltage ($V_{\text{min}}$) distributions6 require the functional or reliability (or survival) probability statistics with linear random variables.

2. Extreme-value statistics and scaling properties

2.1. Minimum extreme-value distribution

Weibull distribution, or the type 3 extreme-value distribution, possesses the unique minimum-value character.1,2 It has been shown to be the most appropriate description3,7-10 of the random nature of dielectric BD in the form of the cumulative density function (CDF) given by:11

$$F(x) = 1 - \exp \left[ -\left(\frac{x}{\alpha}\right)^\beta \right],$$

(1)

where $x$ is the statistical variable, which can be time-to-BD ($t_{\text{BD}}$) or voltage-to-BD ($V_{\text{BD}}$). $\alpha$ is the scale-factor at 63.2% and $\beta$ is the shape-factor (a.k.a. the Weibull slope). This CDF is commonly given in the Weibull plot as

$$W(x) = \ln(-\ln(1 - F)) = \beta \ln(x/\alpha).$$

(2)

In order to connect the experimental data obtained from one area ($A_1$) to another ($A_2$), a statistical spatial distribution for generated BD spots must be known. Poisson distribution is a first-order choice since it assumes uncorrelated and randomly distributed BD spots across a dielectric area. The functional probability function (or reliability function) for the formation of the first conducting filament as a result of BD is given by

$$R_{\text{Poisson}} = \exp(-DA),$$

(3)

where $D$ is the average BD spot density and $A$ is the device area. Then, the failure CDF becomes
Thus, an area-scaling relation can be readily shown as:
\[ W_{A_2} = W_{A_1} + \ln(A_2/A_1) \]  
(5)
for any two areas, \( A_1 \) and \( A_2 \). The relation given in Eq. (5) is also known as Poisson area-scaling. Consequently, the scale-factor, \( \alpha \), inversely depends on area
\[ \frac{\alpha_{A_1}}{\alpha_{A_2}} = \left( \frac{A_2}{A_1} \right)^{1/3} \]  
(6)
as this inverse relation is universally observed over all dielectric materials. The two respective relations in Eqs. (5) and (6) represent the two most fundamental characteristics of dielectric BD as they have been demonstrated for a widely range of dielectric materials as recently reviewed in.\(^3\)

### 2.2. Maximum extreme-value distribution

The type 1, or Gumbel distribution is a maximum-value distribution\(^1,2\) which also satisfies both requirements outlined above. The Gumbel CDF as the survival probability is given:
\[ R(x) = e^{-e^{-\frac{x}{\sigma}}} \]  
(7)
where \( x \) is the statistical variable, e.g. current or conductance. \( \eta \) and \( \sigma \) are the scale-factor and shape-factors of the Gumbel model, respectively. Note here we use the notation \( R \) to distinguish the functional or reliability probability function (sometimes also called as survival function) from the failure probability function, \( F \). This is an important distinction of Gumbel distribution from Weibull distribution.\(^3\) Therefore, the choice in practical applications between Gumbel distribution and Weibull distribution is not an arbitrary one but rather fundamental. It strongly depends on the nature of a physical process which can warrant such a choice as we will discuss later in the applications of dielectric BD and memory operations. The linearized form (G), also known as the Gumbel function, of the CDF Gumbel model is given:
\[ G = -\ln(-\ln R) = x - \frac{\eta}{\sigma} \]  
(8)
Note that the scale-factor, \( \eta \), corresponds to \( x \) at \( G = 0 \) or \( R = 36.8\% \). Analogous to dielectric BD, to establish a connection between device area and random variables \( x \), e.g. RRAM conductance, a spatial distribution of formed filaments is required.\(^4,5\) In this case, we also adopt the Poisson distribution to be compatible with the forming process rather than a binomial distribution model.\(^6\) Using Eqs. (3) and (7), it can be shown that the CDF’s of two devices are related to their corresponding areas, \( A_1 \) and \( A_2 \) as follows
\[ \ln(R_{A_2}) = \ln(R_{A_1}) + \ln(A_2/A_1) \]  
(9)
Alternatively, the area-dependence of Gumbel functions are given as
\[ G_{A_2} = G_{A_1} + \ln(A_2/A_1) \]  
(10)
As a result, the relation between the scale-factors, \( \eta_{A_1} \) and \( \eta_{A_2} \), of any two devices with different areas \( (A_1 \) and \( A_2) \), is given below
\[ \eta_{A_2} = \eta_{A_1} + \sigma \ln(A_2/A_1) \]  
(11)

### 3. Applications to dielectric breakdown

#### 3.1. Experimental characterization and reliability projection

Historically, the lognormal distribution was widely considered as the statistical model for the description of dielectric breakdown over almost 30–40 years prior to the advances of physics-based models\(^7\)–\(^10\) at the end of the 20th century. These ground-breaking works based on percolation theory have provided the very foundation for the use of Weibull statistical model for dielectric breakdown. In addition, the direct experimental work using a large number of samples up to \(~4000\) unequivocally demonstrates that Weibull statistical model is the appropriate model and thus rejects the lognormal distribution.\(^12\)

Dielectric breakdown (or loss of insulating property) occurs as generated defects due to applied stress form a percolating path between two electrodes. The shortest times or the lowest voltages to form such a path determines the extremes out of many nearly completed or unformed percolating paths.\(^7\)–\(^10\) In other words, those nearly formed paths would take longer times to complete their paths, even they may not be always experimentally measurable. Thus, the statistics to describe such a phenomenon represent the minimum-value distribution, a.k.a. the weakest-link statistics.

One common example of applications of the Weibull minimum-value model is to characterize the time-to-breakdown

![Fig. 1.](image)


Moreover, the introduction of low-voltage stress in BEOL interconnects and new middle of line (MOL) spacer materials such as SiBCN and SiOCN further raises reliability concerns as the BD properties of these new materials are not well known. To briefly review the \( T_{BD} \) voltage (field) acceleration models, we summarize four models commonly discussed in the literature: the exponential law of field/voltage dependence, i.e. \( E \) model, the exponential law of reciprocal field/voltage dependence, i.e. \( 1/E \) model, and the power-law voltage/frequency dependence.

\[
T_{BD} = T_{BD0}\exp(-\gamma_V V) \sim \exp(-\gamma_E E),
\]

\[
T_{BD} = T_{BD0}\exp(\frac{G_V}{V}) \sim \exp(-\frac{G_E}{E}),
\]

where \( \gamma_V, \gamma_E, \eta_V, \eta_E, G_V, G_E, \) and \( m \) represent the voltage or field acceleration factors for these four models. These models are widely used to fit experimental data with an additional parameter, the pre-factors, \( T_{BD0} \) or \( Z \) in Eqs. (12)-(15).

A major limitation of dielectric stress for BD study is that the experimental time window usually spans from 1 to \( \sim 10^5 \) s in wafer-level testing. To expand beyond this conventional time window, several different institutions have carried out long-term stress of up to three years at relatively lower stress-voltages. Nevertheless, though long-term stress over many years are certainly beneficial, its throughput is rather limited because of the long turn-around time. This exercise is clearly impractical for technology qualification and development including evaluation of multiple process recipes and material choices. Alternatively, without relying on direct long-term stress, the weakest-link principle (a.k.a. Poisson area-scaling) as the universal dielectric BD property can be used as a criterion to judge a most appropriate acceleration model. This is referred to as the self-consistent acceleration Poisson statistics (SCAPS) method.

This method was originally developed to demonstrate the power-law voltage dependence as the most reliable description of the acceleration model in SiO\(_2\) gate dielectrics in the direct tunneling (DT) regime and in the Fowler–Nordheim regime. The application of this methodology for high-k gate dielectrics and MOL spacer materials is shown in Fig. 3. This figure plots \( T_{BD} \) versus field/voltage with several different areas in the form of four different models of Eqs. (12)-(15). First, the parallel lines from fitting experimental \( T_{BD} \) data indicate that the area-independent acceleration factors are consistent with the weakest-link BD principle. Secondly, a systematic deviation from the parallel lines means that the extrapolation lines of the \( T_{BD} \) data from different areas would encounter a cross-over at low-voltages. This would mean that \( T_{BD} \) of smaller areas would be shorter than that of larger areas beyond the cross-over point, an unphysical result in a violation of the weakest-link property for a given model.

Figure 4 shows a comparison of \( T_{BD} \) voltage acceleration power-law model for both thin and thick SiO\(_2\) stress in DT and FN regimes using both long-term (up to 3 years) and the SCAPS methodology. It is evident that the normalized \( T_{BD} \) data from different areas yield the comparable power-law exponents in agreement with the results of long-term stress, demonstrating the validity of SCAPS methodology. Moreover, the normalized \( T_{BD} \) data of SCAPS methodology covers a similar BD time span \( \sim 10^5 \) s without actually doing long-term stress. The results of these long-term stress for the demonstration of different voltage/field acceleration models are summarized in Table I. In addition, some researchers also employed a method of expanding time window to much shorter time from milli-seconds to nano-seconds to verify the power-law model in dielectric BD and in RRAM set operation. These results are also included in Table I.

It is interesting to note that while new spacer materials of SiBCN and SiOCN follow the power-law model, the acceleration of Si\(_3\)N\(_4\) dielectrics is best described by a \( 1/E \) model,
an observation consistent with a previous publication using pulsed voltage tests and covering ten-orders of time.\textsuperscript{41} We have also shown previously that BD data for SiO$_2$ stressed in the FN regime follows a 1/$E$ or 1/$V$ model,\textsuperscript{23} but the power-law model can provide a good approximate description for the 1/$E$ model due to the fact that 1/$E$ dependence arises from current field-dependence.\textsuperscript{23} Based on these results, we can rule out the $E$- and $\sqrt{E}$-models as the appropriate acceleration models. Coupling the long-term stress data with the BD data obtained from the SCAPS method, we can clearly see in

![Figure 3](image1.png)

**Fig. 3.** (Color online) $T_{BD}$ at 63.2% versus electric fields for several areas are plotted in the four acceleration models.\textsuperscript{16–23} (Column 1: $E$-model; 2: $\sqrt{E}$ model; 3: 1/$E$ model; and 4: Power-law model) to examine their compatibility with the weakest-link property of dielectric BD for several dielectric materials. The blue stars indicate the correct acceleration models which preserve the weakest-link property as discussed in the text.

![Figure 4](image2.png)

**Fig. 4.** (Color online) Comparison of experimental $T_{BD}$ data from long-term stress up to 3 years with the experimental $T_{BD}$ data using different areas based on SCAPS methodology as discussed in text. In the case of the SCAPS methodology, the measured $T_{BD}$ data from different areas is normalized to a reference area of 0.5 $\mu$m$^2$ using Eq. (5). $m$ represents the power-law exponent defined in Eq. (15). The data of the long-term stress of 2.2 nm and 6.2 nm are from\textsuperscript{13} and\textsuperscript{24} with the device areas of 2.475 $\times$ 10$^7$ $\mu$m$^2$ and 10$^5$ $\mu$m$^2$, respectively.
The number of publications supports the respective acceleration models. The number of data sets is given in the parenthesis. The data and the references considered here satisfy the criterion of $T_{BD}$ span in seconds $\geq 7$ decades.

| Testing method      | $E$ mode          | Power-law          | $1/E$ model       | $\sqrt{E}$ model |
|---------------------|-------------------|--------------------|-------------------|------------------|
| Long-term up to 3 years | $1(2) + 1^{10}$   | $5(8)$             | $1(3) + 1^{10}$   | None             |
| SCAPS               | $b$) 3(5)         | 2(2)               | None              |
| Extend to high voltages | 8(12)         | 2(3)               | None              |
| Total No.           | 2(3)              | 16(25)             | 5(8)              | None             |
| Materials           | SiO$_2$, SiO$_2$, HfO$_2$ stacks, SiBCN, SiOCH, SiOCH, ZrO$_2$ etc. | SiO$_2$, Si$_3$N$_4$ |
| Applications        | FEOL              | FEOL, BEOL, MOL, MIMCAPS | FEOL, MOL         |

a) Refers to Ref. 33 that shows the coexistence of $E$-model and $1/E$-model. b) Refers to Ref. 34 which adopts the SCAPS methodology incorrectly.

Fig. 3 that the power-law acceleration model or $1/E$ model is the correct acceleration models for a wide range of dielectric materials. It is worthwhile to point out that in the case of applying the SCAPS methodology to support a voltage or field acceleration model, the verification of Poisson area-scaling is also required as properly done in the case of FEOL gate dielectrics and MOL spacer dielectrics.

### 3.3. Common misunderstandings and mistakes regarding the use of the Weibull model

Although the Weibull model has been widely accepted over the last twenty years, the incorrect use of this model to extract its parameters ($\alpha$, $\beta$) unfortunately spreads widely, especially in the reliability community. First of all, many researchers and engineers fail to recognize that given a statistical model such as Weibull in Eq. (1), $T_{BD}$ is defined to be the sole random variable without any additional sources of statistical variables. In reality, as in the case of BD, random changes in dielectrics thickness ($t_{dilc}$) from sample to sample can introduce additional random variable, $t_{dilc}$. This variation in $t_{dilc}$ can become larger enough to cause the BD distributions to violate Poisson area-scaling relation in Eq. (5), particularly in back end of line (BEOL) and middle of line (MOL) applications. Figure 5 displays the $t_{BD}$ distributions from different sizes of BEOL line-via structures with low-$k$ dielectrics. As compared with Fig. 1(a), Fig. 5(a) shows that the translated distributions fail to merge together according to Eq. (5) with the $\beta$ values of $1.47 \pm 0.27$ by direct fit of these three distributions. In contrast, the $T_{BD}$ versus-area relation in Fig. 5(b) yields a fictitious “$\beta$’ value of $3.19 \pm 0.89$. It is evident that these two $\beta$ values vastly disagree with each other. This disagreement stems from the fact that these distributions are actually non-Weibull with an extra random variable due to thickness variation even though they may appear to fit well with the straight lines. The variability issue and its impact on dielectric BD has been

---

**Fig. 5.** (Color online) (a) The original $T_{BD}$ distributions of 533 and 53.3 K (Grey symbols) and 5.3 K (blue symbols) for different size of the line-via structures in BEOL low-$k$ dielectrics. The black and red symbols represent the $T_{BD}$ distributions of 533 and 53.3 K translated from the original distributions to 5.3 K structure as the reference area by applying Poisson area-scaling, Eq. (5). (b) $T_{BD}$ at 63% versus area with a fictitious “$\beta$” value of 3.19 using Eq. (6) in disagreement with the $\beta$ value of 1.47 from (a).

**Fig. 6.** (Color online) The current–voltage characteristics during the initial forming and the switching operation (reset/set) of a typical RRAM device.
extensively reviewed in.43) As discussed above, thousand samples are required to distinguish a Weibull distribution from other distributions. To avoid the use of thousand samples, the application of Poisson area-scaling in Eq. (5) is required to verify whether the experimental data obey the weakest-link property such as the Weibull model. Because of violation of Poisson area-scaling as seen in Fig.5(a), the use of Eq. (6) to extract the $\beta$ value is completely invalid because the $BD$ distributions are non-Weibull even if they may appear to be Weibull-distributed. Unfortunately, this kind of practice without carefully verifying whether the BD data follow Poisson area-scaling widely spreads in reliability community and lead to much confusion and misunderstandings in literature.

For example, in the case of use SCAPS methodology to support the $E$-model in BEOL low-$k$ dielectrics,34) the authors fail to check whether their $BD$ data actually fail to follow Poisson area-scaling.34) However, they used the incorrectly extracted $\beta$ value from area-scaling, Eq. (6) in the application of the SCAPS methodology for their $BD$ data. Therefore, their conclusion to claim the supportive evidence of $E$-model cannot be warranted.34)

4. Applications to memory operations

RRAM technology has received a great deal of attention in its research and development for neuromorphic and in-memory computing. In this section, after reviewing the RRAM forming process, we will summarize the recent advances in applications of maximum extreme-value distribution to RRAM conductance4,5) and SRAM $V_{\text{min}}$ statistical characterization.6) Figure 6 shows the forming IV curve of a typical RRAM device as well as its switching IV characteristics. The switching operation between a low-resistance state (LRS) and high-resistance state (HRS) can occur using a procedure of reset and set operations as shown in Fig. 6.

4.1. Two-stage forming process in RRAM devices

Analogous to dielectric breakdown, the forming process for the initial creation of a percolating filament by performing either ramp-voltage stress (RVS) in Fig. 6 or a constant-voltage stress (CVS) as commonly performed in BD tests.4) These two methods are shown to be equivalent if the results of CVS tests are properly converted to voltage domain or vice versa using the appropriate conversion procedure.44) Figure 7(a) compares the converted forming voltage distributions from CVS tests in good agreement with the directly measured $V_{\text{forming}}$ whereas Fig. 7(b) shows that the translated $V_{\text{forming}}$ distributions by applying Eq. (5) merge into a universal line.

One of the important questions is whether forming process follows a two-stage process in RRAM devices4,45,46) similar to dielectric BD as extensively reviewed in Ref. 43. Unlike RVS method with compressed voltage window, the advantage of CVS technique is that it reveals the filament evolution process involving two stages commonly known as progressive BD phenomenon. Figure 8 displays the current time-evolutions using a CVS test,4) showing the initial current jumps defined as nano-path nucleation, the 1st soft BD ($SBD$) in the 1st stage. Then the currents of these samples continue to grow with substantial fluctuations and eventually reach to the current compliance in the 2nd stage. More multiple paths or filaments are generated in the 2nd stage and continue to grow in competition and eventually culminate into the completion of a switching filament as the current

---

**Fig. 7.** (Color online) (a) Comparison of the $V_{\text{forming}}$ distributions with the converted $V_{\text{forming}}$ distributions from CVS test. (b). The normalized $V_{\text{forming}}$ distributions of different areas according to Poisson area-scaling using Eq. (5). The line is from the progressive BD model involving multiple competing paths (or) for two-stage forming process.43,45) In both (a) and (b), the Weibull function in Eq. (2) is used in the vertical axis.

---

**Fig. 8.** (Color online) The time-evolution of the measured currents for three devices during the forming process using CVS method.4)
compliance is researched. Note many incomplete paths/ filaments remain as a result of stress as we will discuss their impact on variability in the last section. The schematics in Fig. 9 illustrates the forming process: (a) the initial generation of the nano-conducting path (1st SBD), (b) the generation of multiple paths and their growth and conglomeration. (c) the formation of the final conducting filament with many incomplete or nearly formed filaments.

The time-to-the nano-path nucleation (1st SBD) and the time-to-filament formation (HBD) are plotted in Fig. 10 for three different areas. These data are normalized to the smallest area according to Poisson area-scaling, Eq. (5). It is worthwhile to point out that Eq. (5) can directly be derived from Eq. (4) without knowledge of the specific CDF forms either Weibull model or Gumbel model. The merged universal distributions in Fig. 6(b) and in Fig. 10 clearly reveals some curvatures at high percentiles, i.e. non-Weibull distribution. This proves that Poisson area-scaling is not only limited to the Weibull distribution but also applicable to the case of non-Weibull distribution, consistent with Eq. (5).

4.2. RRAM switching conductance statistics
As discussed above in the case of dielectric BD or forming process, the first spot presents the complete conducting filament. In this case, one of the paths is completed while all other paths are incomplete or nearly formed.

Fig. 9. (Color online) The evolution of two-stage forming process: (a) the nucleation of the 1st nano-conducting path (1st soft BD). (b) During the 2nd stage, the generation of multiple paths and enlargement of these paths. (c) the forming completes with the final percolating filament and many unformed filaments. In this case, one of the paths is completed while all other paths are incomplete or nearly formed.

Fig. 10. (Color online) The distributions of time-to-the first nano-path (1st SBD) nucleation ($T_C$) and time-to-filament formation ($T_F$). These distributions are translated to the reference area of $10^2 \mu m^2$ according to Poisson area-scaling, Eq. (5). The lines are from MC simulation of two-stage forming process based on the multiple competing spot progressive BD methodology.

---

**Fig. 9.** (Color online) The evolution of two-stage forming process: (a) the nucleation of the 1st nano-conducting path (1st soft BD). (b) During the 2nd stage, the generation of multiple paths and enlargement of these paths. (c) the forming completes with the final percolating filament and many unformed filaments. In this case, one of the paths is completed while all other paths are incomplete or nearly formed.

**Fig. 10.** (Color online) The distributions of time-to-the first nano-path (1st SBD) nucleation ($T_C$) and time-to-filament formation ($T_F$). These distributions are translated to the reference area of $10^2 \mu m^2$ according to Poisson area-scaling, Eq. (5). The lines are from MC simulation of two-stage forming process based on the multiple competing spot progressive BD methodology.
filament among many nearly complete filaments. Consequently, this conductance or current through this filament is the highest in comparison with other filaments. In this regard, the distribution of the measured conductance or currents represents the statistics of the maximum-values out of potentially completed filaments. This important aspect is sharply in contrast to dielectric BD in which a minimum extreme-value statistical process controls the outcome of the competition among many potentially formed filaments. This consideration renders two important requirements: first the cumulative probability function corresponds to the functional or reliability (or survival) probability distribution rather than the failure distribution as in the case of BD.

Figure 11 shows some typical reset IV characteristics of a large number of RRAM devices in the LRS state after the forming process. To demonstrate the validity or applicability of the Gumbel statistical model with the maxima-value distribution, we plot the experimental distributions of RRAM conductance at $V = -0.1V$ in the Gumbel plot in Fig. 12(a). The distributions are also translated using Eq. (10) to verify the Poisson area-scaling in the Gumbel plot as shown in Fig. 12(b). These results clearly reveal the translated distributions overlay well as expected from Eq. (10).

To further verify the validity of the Gumbel model in its application of RRAM conductance, we plot the scale-factor, $(q)$ as a function of area in Fig. 13. The derived shape-factor $(\sigma)$ using Eq. (12) is used to compare with the merged distributions in Fig. 12(b). The results clearly demonstrate the maximum-value distribution model can well capture the experimental data with reasonable agreement. These results support the validity of the Gumbel model as the maxima-value distribution to characterize the RRAM conductance or current statistics. Figure 13 also shows the forming-voltages as a function of area in comparison with the area-dependence of the LRS conductance. It exhibits the remarkably distinctive characteristics of the area-scaling of their respective scale-factors for the minimum-value distribution [Eq. (6)] and the maximum-value distribution [Eq. (11)]. This suggests any arbitrary choices to replace extreme-valued statistical process can be misguided in research and technology development with unwanted consequences.

### 4.3. Applications to SRAM $V_{\text{min}}$ statistics

Recently, the Gumbel model has successfully been applied to investigate the $V_{\text{min}}$ voltage distribution in the SRAM circuits. The $V_{\text{min}}$ value is generally defined as the smallest supply voltage for which a SRAM array remains functional. At any voltages lower than this $V_{\text{min}}$ value, one cell of this SRAM array would fail by definition. Therefore, the $V_{\text{min}}$ values measured from a population of SRAM arrays actually represent the maximum-operational voltages from the viewpoint of SRAM functionality. This consideration leads to the realization that $V_{\text{min}}$ issue is consistent with the property of maximum extreme-value statistics. The Gumbel model has been successfully applied to characterize SRAM $V_{\text{min}}$ distribution. Figure 14 shows the cumulative distributions of SRAM $V_{\text{min}}$ measurements for two different array sizes as well as the normalized distributions according to Eq. (10). These results clearly demonstrate the applicability of the Gumbel model to characterize the SRAM $V_{\text{min}}$ distributions.

![Fig. 11.](image1) The IV characteristics of the first reset for 900 $\mu m^2$. The red dots are defined as $I_{\text{Reset}}$ and $V_{\text{Reset}}$.

![Fig. 12.](image2) (Color online) The LRS and HRS conductance distributions at ±0.1 V for three areas (a) and the merged (or normalized) distributions (b) according to Eq. (4) in Table I. The reference area is the largest area of 104 $\mu m^2$. $G_0$ is the quantum conductance of $7.75 \times 10^{-5}$ S.
The slight deviation of the $V_{\text{min}}$ distribution for 20 Mb SRAM arrays at lower percentiles from that of 1 Mb arrays is believed to be caused by the silicon process variation from wafer- to wafer-edge.\cite{6}

It is interesting to note Poisson area-scaling of Eqs. (5) and (10) remains a common feature as shown in Figs. 1, 10, 12(b) and 14, even though arising from the two sets of opposite extreme-value considerations: minimum-value versus maximum-value distribution or drastically different CDF forms of the Weibull model versus the Gumbel model. This is because the linearization of these CDF's removes the dependence on the specific functional forms as given in Eqs. (1) and (7). Since Poisson random statistics of Eq. (3) is adopted in either the Weibull model and the Gumbel model, the translation of the Weibull function or the Gumbel function remains unchanged as in Eqs. (5) and (10).

4.4. Implication on variability

It is difficult to directly compare the shape-factors ($\beta$ versus $\sigma$) of $V_{\text{forming}}$ distributions versus LRS conductance distributions because the random variables are plotted in the log-scale and the linear-scale, respectively, as discussed above. Nevertheless, it can be seen that the results in Fig. 12(b) reveal that the spread of LRS conductance data is much larger than that of forming-voltages in Fig. 7(b) and time-to-forming in Fig. 10, particularly with respect to the results of the respective model predictions. The deviations from the merged distributions are particularly evident as shown in Fig. 12(b). This large difference stems from the inherent consequences of minimum and maximum extreme-value distributions which characterize the forming-voltage and RRAM conductance, respectively. For forming-voltage, its minimum extreme-value characteristics inherently exclude any effect of nearly formed percolation paths because these incomplete paths would require longer times or higher voltages to form as illustrated in Fig. 9(c) even though they are not experimentally measurable due to the termination of the stress work. On the other hand, the maximum-value statistics merely guarantee that the minimum-valued conducting filaments are contributing the highest current in the conductance. Nevertheless, this does not necessarily prevent contributions from other nearly or unformed percolating filaments to the current measurements. Therefore, in the case of forming-voltage statistics, the variability arises solely from sampling uncertainties if external variability sources are excluded such as dielectric thickness variations. In contrast, in the case of RRAM conductance statistics, nearly or unformed filaments can all contribute to the total current as additional and intrinsic variability sources if the external sources are absent. The fact that the deviated conductance values at lower or higher percentiles are smaller than those of the model prediction may suggest strong contributions from nearly or unformed filaments without any completed filaments in these particular RRAM devices.

In the case of SRAM $V_{\text{min}}$ statistics, the inherent variation is small as shown in Fig. 14 even though the distribution is controlled by a maximum-valued stochastic process. This is because the $V_{\text{min}}$ measurements correspond to the largest functional voltages while any smaller voltages are excluded analogous to the forming process or dielectric breakdown. These physical considerations cannot be realized without a solid understanding of the role of a statistical distribution model as discussed in the introduction.

5. Conclusions

We have reviewed the fundamentals of the Weibull and Gumbel statistical model (minimum-value versus maximum-value distribution) in connection with the area-scaling
properties of their respective scale-factors and distribution-percentile. In application of the minimum-value distribution such as the Weibull model, we show the use of the weakest-link character can allow us to objectively demonstrate a voltage/field acceleration. In characterization of forming process in RRAM technology, we show that the minimum-value statistics can be either Weibull or non-Weibull depending on the nano-path nucleation to switching filament formation. On the other hand, the proper consideration of the physical random process of interest leads to a consistent and correct choice of the maximum-value distribution, Gumbel model, for RRAM conductance and SRAM $V_{\text{min}}$ voltage measurements in excellent agreement with experimental data. These examples demonstrate the fundamental roles of extreme-value distribution in dielectric breakdown and memory applications, laying the foundation for future applications and pave the paths to investigate the impact of variability encountered in realistic applications in RRAM technology development.

1) E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
2) R. Fisher and L. Tippett, Math. Proc. Cambridge Philos. Soc., 24, 180 (1928).
3) E. Wu, IEEE Trans. Electron Devices 66, 4523 (2019).
4) E. Wu, T. Ando, Y. Kim, R. Muralidhar, E. Cartier, P. Jamison, M. Wang, and V. Narayanan, Digest of Internal Electron Device Meeting, 2019, p. 709.
5) E. Wu, T. Ando, Y. Kim, R. Muralidhar, E. Cartier, P. Jamison, M. Wang, and V. Narayanan, Appl. Phys. Lett. 116, 082901 (2020).
6) R. Pompl, S. Strasser, Drex1, and M. Ostermayr, IEEE Trans. Device Mater. Reliab. 15, 289 (2015).
7) J. Suñé, I. Placencia, N. Barniol, E. Farrés, F. Martín, and X. Aymerich, Thin Solid Films 185, 347 (1990).
8) R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, IEDM Tech. Dig., 1995, p. 866.
9) J. H. Stathis, J. Appl. Phys. 86, 5757 (1999).
10) J. Suñé, IEEE Electron Device Lett. 22, 296 (2001).
11) E. Y. Wu and R.-P. Vollandt, IEEE Trans. Electron Devices 49, 2131 (2002).
12) E. Y. Wu, W. W. Abadeer, L. Han, S. Lo, and G. Huckel, Proc. Int. Reliability Physics Symp., 1999, p. 57.
13) J. Stathis, Proc. Int. Reliability Physics Symp., 2001, p. 132.
14) K. S. Toki, Microelectron. Eng. 87, 348 (2010).
15) J. H. Stathis, M. Wang, R. G. Southwick, E. Y. Wu, B. P. Linder, E. G. Liniger, G. Bobilla, and H. Kothari, Digest of Internal Electron Device Meeting, 2014, p. 522.
16) J. W. McPherson and H. C. Mogul, Proc. IRPS, 1998, p. 47.
17) J. W. McPherson, J. Appl. Phys. 95, 8101 (2004).
18) I. C. Chen, S. E. Holland, and C. Hu, IEEE Trans. Electron Devices 32, 413 (1985).
19) K. F. Schuegraf and C. Hu, J. Appl. Phys. 76, 3695 (1994).
20) K.-H. Allers, Microelectron. Reliab. 44, 411 (2004).
21) N. Suzumura, Y. Yamamoto, D. Kodama, K. Makabe, E. Marakami, S. Maegawa, and K. Kubota, Proc. IRPS, 2006, p. 484.
22) E. Wu, A. Vayshenker, E. Nowak, J. Suñé, R.-P. Vollandt, W. Lai, and D. Harmon, IEEE Trans. Electron Devices 49, 2244 (2002).
23) E. Wu and J. Suñé, IEEE Trans. Electron Devices 56, 1433 (2009).
24) R.-P. Vollandt, Microelectron. Reliab. 43, 865 (2003).
25) J. W. McPherson, V. Reddy, K. Banerjee, and H. L. Iedm, Tech. Dig., 1998, p. 171.
26) K. Croes et al., Proc. IRPS, 2013, p. 2F.4.1.
27) E. G. Liniger, S. A. Cohen, and G. Bonilla, Proc. IRPS, 2014, p. BD.4.1.
28) E. G. Liniger, S. A. Cohen, and G. Bonilla, Appl. Phys. Lett. 109, 90 (2013).
29) T.-Y. Jeong, J. Kim, Y. Jo, K. Tak, M. Lee, S. Windu, H. Choi, Y. Jo, S. Paé, and J. Park, Proc. IEEE Int. Interconnect Technology Conf., (IFTC), 2015, p. 299.
30) M. N. Chang, Y.-H. Lee, S. Y. Lee, K. Joshi, C. C. Ko, C. C. Chiu, and K. Wu, IEDM, Tech. Dig., 2015, p. 165.
31) E. Wu, J. Suñé, C. LoRow, and R. Dufrense, IEDM, Tech. Dig., 2012, p. 653.
32) R. Southwick III, E. Wu, S. Mehta1, and J. Stathis, Proc. IRPS, 2017, DG-1.1-1.5.
33) T. Pompl, K. Allers, R. Schwab, K. Hofmann, and M. Röhner, IRPS Proc., 2005, p. 388.
34) J. Kim, E. Ogawa, and J. McPherson, Proc. IRPS, 2007, p. 399.
35) M. Röhner, A. Kerber, and M. Kerber, Proc. IRPS, 2006, p. 76.
36) A. Elie, W. Stadler, A. Kerber, T. Pompl, T. Brodzbeck, K. Esmark, and A. Bravaiix, IEEE Electrical Overstress/Electrostatic Discharge Symp. (EOS/ESD), 2006, p. 5A.2.1.
37) C. LaRow, Z. Chibili, S. F. Yap, and A. Kerber, IEEE Int. Integrated Reliability Workshop (IRW), Final Report, 2017, p. 1.
38) J. Y. Seo, J. F. Seok, H. J. Kim, H. S. Park, J. E. Jeon, and W. S. Lee, Proc. 26th Int. Conf. on Microelectronics (MIEL), 2008, p. 521.
39) D. Zhou, U. Schroeder, G. Jegert, M. Kerber, S. Uppal, R. Agaby, M. Reinicke, J. Heitmann, and L. Oberbeck, J. Appl. Phys. 106, 044104 (2009).
40) W. Luo, J. Liu, H. Feng, Y. Lin, J. Huang, K. Lin, and T. Hou, IEDM, Tech. Dig., 2012, p. 215.
41) R. Kaplar, S. Habermehl, R. Apodaca, B. Havener, and E. Osmun, IEEE Trans. Electron Devices 55, 224 (2013).
42) E. Wu, B. Li, J. H. Stathis, R. Achanta, R. Filippi, and P. McLaughlin, IEDM, Tech. Dig., 2013, p. 401.
43) E. Wu, IEEE Trans. Electron Devices 66, 4535 (2019).
44) A. Kerber, L. Pantuso, A. Veloso, G. Groeseneken, and M. Kerber, Microelectron. Reliab. 47, 513 (2007).
45) N. Raghavan, A. Fantini, R. Degraeve, P. Roussel, L. Goux, B. Gouvreau, D. J. Wouters, G. Groeseneken, and M. Jurczak, Microelectron. Eng. 109, 177 (2013).
46) E. Wu and J. Suñé, Proc. IRPS, 2006, p. 54.