Poisson-process limit laws yield Gumbel max-min and min-max

Iddo Eliazar,[1,*] Ralf Metzler,[2,†] and Shlomi Reuveni[1,‡]

[1]*School of Chemistry, The Center for Physics and Chemistry of Living Systems, The Raymond and Beverly Sackler Center for Computational Molecular and Materials Science, and The Mark Ratner Institute for Single Molecule Chemistry, Tel Aviv University, Tel Aviv 6997801, Israel*

[2]*University of Potsdam, Institute of Physics & Astronomy, 14476 Potsdam, Germany*

(Received 26 August 2018; revised manuscript received 25 February 2019; published 21 August 2019)

"A chain is only as strong as its weakest link" says the proverb. But what about a collection of statistically identical chains: How long till all chains fail? The answer to this question is given by the max-min of a matrix whose (i, j) entry is the failure time of link j of chain i: take the minimum of each row, and then the maximum of the rows' minima. The corresponding min-max is obtained by taking the maximum of each column, and then the minimum of the columns' maxima. The min-max applies to the storage of critical data. Indeed, consider multiple backup copies of a set of critical data items, and consider the (i, j) matrix entry to be the time at which item j on copy i is lost; then, the min-max is the time at which the first critical data item is lost. In this paper we address random matrices whose entries are independent and identically distributed random variables. We establish Poisson-process limit laws for the row's minima and for the columns' maxima. Then, we further establish Gumbel limit laws for the max-min and for the min-max. The limit laws hold whenever the entries' distribution has a density, and yield highly applicable approximation tools and design tools for the max-min and min-max of large random matrices. A brief of the results presented herein is given in: Gumbel central limit theorem for max-min and min-max [Eliazar, Metzler, and Reuveni, Phys. Rev. E **100**, 020104 (2019)].

DOI: 10.1103/PhysRevE.100.022129

I. INTRODUCTION

Extreme value theory (EVT) is a branch of probability theory that focuses on extreme-value statistics such as maxima and minima [1–3]. EVT has major applications in science and engineering [4–6]; examples range from insurance to finance, and from hydrology to computer vision [7–9]. At the core of EVT stands its fundamental theorem, the Fisher-Tippett-Gnedenko theorem [10,11], which establishes the three extreme-value laws: *Weibull* [12,13], *Frechet* [14], and *Gumbel* [15].

The fundamental theorem of EVT applies to ensembles of independent and identically distributed (IID) real-valued random variables, and is described as follows [16]. Consider an ensemble $\{X_1, \ldots, X_n\}$ whose n components are IID copies of a general real-valued random variable X. Further consider the ensemble's maximum $M_n = \max\{X_1, \ldots, X_n\}$, and an affine scaling of this maximum:

$$\tilde{M}_n = s_n(M_n - \delta_n), \tag{1}$$

where s_n is a positive scale parameter, and where δ_n is a real location parameter. The fundamental theorem of EVT explores the convergence in law (as $n \to \infty$) of the scaled maximum $\tilde{M}_n$ to a nontrivial limiting random variable $\mathcal{L}$.

First, the fundamental theorem determines its admissible "inputs": the classes of random variables X that yield nontrivial limits $\mathcal{L}$. Second, given an admissible input X, the fundamental theorem specifies the adequate scale parameter s_n and location parameter δ_n. Third, as noted above, the fundamental theorem establishes that its "outputs" are the three extreme-value laws: the statistics of the nontrivial limits $\mathcal{L}$ are either Weibull, Frechet, or Gumbel. The *domain of attraction* of each extreme-value law is the class of inputs X yielding, respectively, each output law.

The fundamental theorem of EVT yields asymptotic approximations for the maxima of large ensembles of IID real-valued random variables. Indeed, consider the scaled maximum $\tilde{M}_n$ to converge in law (as $n \to \infty$) to a nontrivial limit $\mathcal{L}$. Then, for a given large ensemble ($n \gg 1$), the ensemble's maximum M_n admits the following *extreme-value asymptotic approximation* in law:

$$M_n \simeq \mathcal{L}_* := \delta_n + \frac{1}{s_n}\mathcal{L}. \tag{2}$$

The extreme-value asymptotic approximation of Eq. (2) has the following meaning: the deterministic asymptotic approximation of the ensemble's maximum M_n is the location parameter δ_n; the magnitude of the random fluctuations about the deterministic asymptotic approximation is $1/s_n$, the inverse of the scale parameter s_n; and the statistics of the random fluctuations about the deterministic asymptotic approximation are that of the limit $\mathcal{L}$, which is governed by one of the three extreme-value laws.

The three extreme-value laws are *universal* in the sense that they are the *only* nontrivial limiting statistics obtainable (as $n \to \infty$) from the scaled maximum $\tilde{M}_n$. However, universality holds neither for the corresponding domains of attraction nor for the corresponding scale parameter s_n and location parameter δ_n. Indeed, each extreme-value law has a very specific and rather narrow domain of attraction [16]. Also,

*eliazar@post.tau.ac.il
†rmetzler@uni-potsdam.de
‡shlomire@tauex.tau.ac.il

for any given admissible input X, the scale parameter s_n and location parameter δ_n are "custom tailored" in a very precise manner [16].

In essence, the fundamental theorem of EVT considers a random-vector setting: the maxima of what can be perceived as vector-structured ensembles of IID real-valued random variables. This paper elevates from the random-vector setting to the following *random-matrix* setting: the *max-min* and the *min-max* of matrix-structured ensembles of IID real-valued random variables. The max-min is obtained by taking the minimum of each matrix row, and then taking the maximum of the rows' minima. The min-max is obtained by taking the maximum of each matrix column, and then taking the minimum of the columns' maxima.

The max-min and the min-max values of matrices emerge naturally in science and engineering. Perhaps the best known example of the max-min and the min-max comes from *game theory* [17,18]. Indeed, consider a player that has a set of admissible strategies, and that faces a set of viable scenarios. A payoff matrix determines the player's gains—or, alternatively, losses—for each strategy it applies and for each scenario it encounters. The player's goal is to optimize with respect to the *worst-case scenario*. Hence, in the case of gains, the player goes *max-min*: calculate the minimal gain per each scenario, and then pick the strategy that yields the largest minimal gain. And, in the case of losses, the player goes *min-max*: calculate the maximal loss per each scenario, and then pick the strategy that yields the smallest maximal loss. In the field of game theory the max-min and the min-max values appear also in the context of game-searching procedures on trees [19,20].

Architectural illustrations of the max-min and the min-max values come from *reliability engineering* [21,22], where one is interested in calculating the failure time (or the failure load) of a given system. Two important system-architectures are "series-parallel" and "parallel-series" [23–25]. In the series-parallel architecture a system is a parallel array of subsystems, and each subsystem is a serial array of components. In the parallel-series architecture a system is a serial array of subsystems, and each subsystem is a parallel array of components. The max-min and the min-max values correspond, respectively, to the failure times (or the failure loads) of systems with series-parallel and with parallel-series architectures [23–25].

There are several limit-law results—counterparts of the fundamental theorem of EVT—for the max-min and the min-max of random matrices (with IID entries). The pioneering mathematical results were presented by Chernoff and Teicher [26], reliability-engineering results were presented by Kolowrocki [23–25], and relatively recent reliability-engineering results were presented by Reis and Castro [27]. All these limit-law results use affine scalings—similar to that of Eq. (1)—for the max-min and the min-max. Also, all these limit-law results employ asymptotic couplings of the dimensions of the random matrices (as these dimensions are taken to infinity).

Chernoff and Teicher established that the limit laws for the max-min and the min-max are the three extreme-value laws [26]: Weibull, Frechet, and Gumbel. Kolowrocki investigated limit laws for the max-min and the min-max in the context of systems with the aforementioned series-parallel and parallel-series architectures [23–25]. Considering the max-min, and applying the fundamental theorem of EVT iteratively—first to the minimum of each matrix-row, and then to the maximum of the rows' minima—Reis and Castro established a Gumbel limit law [27]; this limit law applies to matrix entries that belong to subsets of the domains of attraction of the three extreme-value laws.

For the results of [23–27]—as in the case of the fundamental theorem of EVT—universality holds neither with regard to the domains of attraction nor with regard to the affine scalings. Also, for these results, universality does not hold with regard to the asymptotic couplings of the dimensions of the random matrices. Moreover, as the results of [23–27] involve very intricate mathematical conditions and schemes, their practical implementation is extremely challenging.

The limit-law results of [23–27] are derived via an "EVT machinery," i.e., methods similar to the Fisher-Tippett-Gnedenko theorem, together with other EVT results (e.g., [28]). In this paper we take an altogether different approach: a "bedrock" Poisson-process method. Specifically, we dive down to the bedrock level of the rows' minima and the columns' maxima (of random matrices with IID entries), and establish *Poisson-process limit laws* for these minima and maxima. Then, elevating back from the bedrock level to the max-min and the min-max, we establish *Gumbel limit laws* for these values. For a brief of the results presented herein we refer the reader to Ref. [29].

The limit laws presented here have the following key features. First, their domain of attraction is vast: the limit laws hold whenever the entries' distribution has a density. Second, they use affine scalings similar to that of Eq. (1) with a location parameter that is tunable (it can be set as we wish within the interior of the support of the IID entries), and with a scale parameter that depends on the entries' distribution only up to a coefficient. Third, their asymptotic couplings (of the dimensions of the random matrices) are geometric. Due to these features the practical implementation of the limit laws presented here is easy and straightforward, and hence these results are highly applicable.

Figure 1 demonstrates the potency of the Gumbel limit law for the max-min (see Sec. III for the details). This figure depicts numerical simulations of the max-min of random matrices whose IID entries are drawn from an assortment of distributions: exponential, gamma, log-normal, inverse Gauss, uniform, Weibull, beta, Pareto, and normal. For all these distributions, the convergence of the simulations to the theoretical prediction of the max-min result is evident. The MATLAB code that was used in order to generate the simulations is detailed in the Appendix; this short code shows just how easy it is to apply, in practice, the novel Gumbel limit laws presented here.

The reminder of this paper is organized as follows. Section II presents the random-matrix setting, and the bedrock Poisson-process limit law for the rows' minima. Then, Sec. III establishes the Gumbel limit law for the max-min, which is motivated by the following question: Within a collection of IID chains, how long will the strongest chain hold? Section IV further establishes the counterpart Gumbel limit law for the min-max, which is based on a counterpart bedrock Poisson-process limit law for the columns' maxima, and which is motivated by the following question: Using a collection of IID

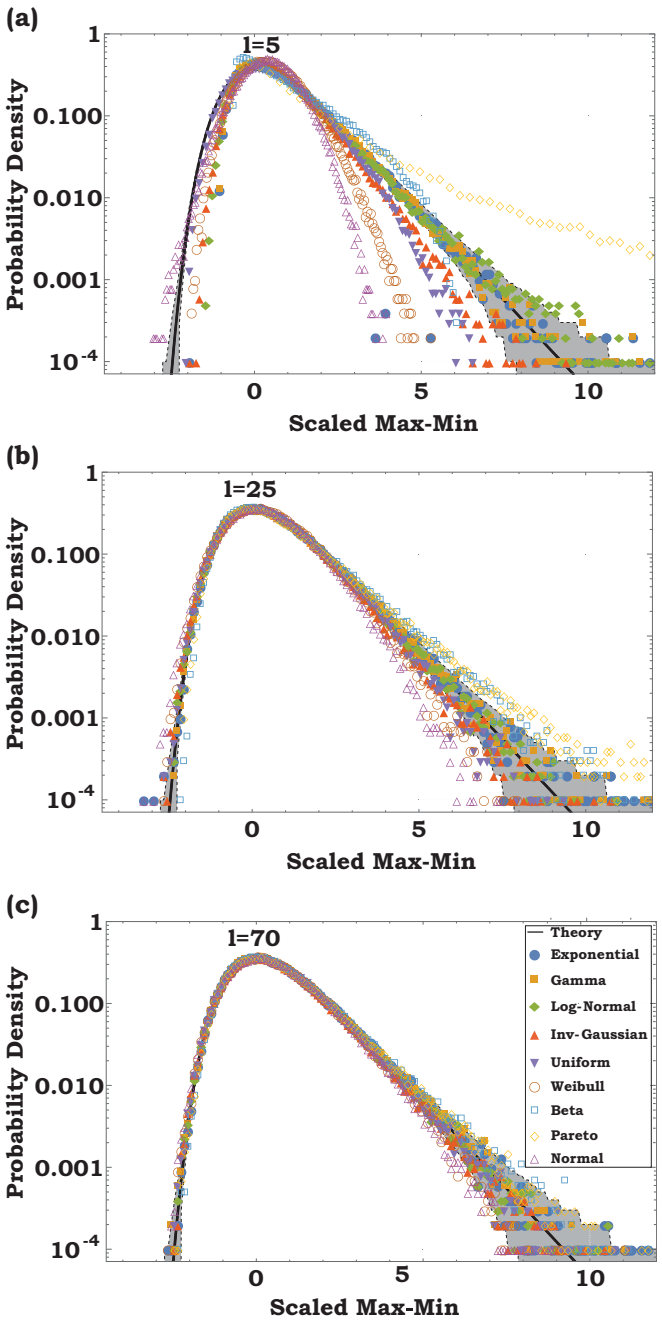

FIG. 1. Numerical simulations demonstrating Proposition 2: the convergence of the scaled max-min $\tilde{\wedge}_{\max}/\bar{\eta}$, in law, to the standard Gumbel random variable $\mathcal{G}$ (see Sec. III for details). 10^5 random matrices were simulated, with $l = 5, 25, 70$ links and $c \simeq 1.25^l$ chains [see Eq. 3]. The colored symbols depict the simulated data points of the scaled max-min. The solid black line depicts the density function of the standard Gumbel random variable $\mathcal{G}$ (with its 95% confidence interval shaded in grey). Nine different distributions of the generic failure time T are considered: exponential, gamma, log-normal, inverse-Gauss, uniform, Weibull, beta, Pareto, and normal. As the number of links grows from $l = 5$ (top) to $l = 25$ (middle) and to $l = 70$ (bottom), the convergence of the simulated data to the standard Gumbel density function is evident.

data-storage backup copies, how long can the data be stored reliably by the backup copies? Section V describes the application of the Gumbel limit laws as approximation tools and as design tools. An in-depth discussion of the limit laws is held in Sec. VI. Finally, Sec. VII concludes, and the proofs of the key results stated along the paper are detailed in the Appendix.

II. BEDROCK

Consider a collection of c *chains*, labeled by the index $i = 1, \ldots, c$. Each chain comprises l *links*, and all the $c \times l$ links are IID copies of a generic link. In this paper we take a temporal perspective and associate the *failure time* of the generic link with a real-valued random variable T. Namely, T is the random time at which the generic link fails mechanically.

As the analysis to follow is probabilistic, we introduce relevant statistical notation. Denote by $F(t) = \Pr(T \leqslant t)\ (-\infty < t < \infty)$ the *distribution function* of the generic failure time T, and by $\bar{F}(t) = \Pr(T > t)\ (-\infty < t < \infty)$ the corresponding *survival function*. These functions are coupled by $F(t) + \bar{F}(t) = 1\ (-\infty < t < \infty)$. The *density function* of the generic failure time T is given by $f(t) = F'(t) = -\bar{F}'(t)\ (-\infty < t < \infty)$. In particular, this notation covers the case of a positive-valued generic failure time T. We note that, alternatively to the temporal perspective taken here, the random variable T can manifest any other real-valued quantity of interest of the generic link, e.g., its mechanical strength (in which case T is positive valued).

The following *random matrix* underlies the collection of chains:

$$\mathbf{T} = \begin{pmatrix} T_{1,1} & \cdots & T_{1,l} \\ \vdots & \ddots & \vdots \\ T_{c,1} & \cdots & T_{c,l} \end{pmatrix}. \tag{3}$$

The dimensions of the random matrix $\mathbf{T}$ are $c \times l$, and its entries are IID copies of the generic failure time T. The ith row of the random matrix $\mathbf{T}$ represents the l links of chain i, and the entries of this row manifest the respective failure times of the links of chain i. Specifically, the entry $T_{i,j}$ is the failure time of link j of chain i.

"A chain is only as strong as its weakest link" says the proverb. So, chain i fails as soon as one of its links fails. Hence the chain's failure time is given by the *minimum* of the failure times of its links:

$$\wedge_i = \min\{T_{i,1}, \ldots, T_{i,l}\} \tag{4}$$

$(i = 1, \ldots, c)$. Namely, the random variable $\wedge_i$ is the minimum over the entries of the ith row of the random matrix $\mathbf{T}$.

Now, consider an arbitrary reference time t_* of the generic failure time T, e.g., its median, its mean (in case the mean is finite), or its mode [in case the density function $f(t)$ is unimodal]. In general, the reference time t_* can be any real number that satisfies two basic requirements: (i) $0 < F(t_*) < 1$, which is equivalent to $0 < \bar{F}(t_*) < 1$; and (ii) $0 < f(t_*) < \infty$. These requirements are met by all the interior points in the support of the input T.

With respect to the reference time t_*, we apply the following affine scaling to the failure time of the ith chain:

$$\tilde{\wedge}_i = l(\wedge_i - t_*) \tag{5}$$

TABLE I. Summary of Propositions 1 and 3. Rows 1–3 summarize the underlying settings: the quantities under consideration, their affine scalings, and the ensembles under consideration. Rows 4–6 summarize the Poisson-process limit-law results: the required asymptotic geometric couplings of c and l, the intensity functions of the limiting Poisson processes (to which the ensembles converge in law), and their exponents.

	Proposition 1	Proposition 3
1. Quantity	$\wedge_i = \min\{T_{i,1}, \ldots, T_{i,l}\}$	$\vee_j = \max\{T_{1,j}, \ldots, T_{c,j}\}$
2. Scaling	$\tilde{\wedge}_i = l(\wedge_i - t_*)$	$\tilde{\vee}_j = c(\vee_j - t_*)$
3. Ensemble	$\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$	$\{\tilde{\vee}_1, \ldots, \tilde{\vee}_l\}$
4. Coupling	$\lim_{c\to\infty, l\to\infty} c\bar{F}(t_*)^l = 1$	$\lim_{l\to\infty, c\to\infty} lF(t_*)^c = 1$
5. Intensity	$\lambda(x) = \bar{\epsilon}\exp(-\bar{\epsilon}x)$	$\lambda(x) = \epsilon\exp(\epsilon x)$
6. Exponent	$\bar{\epsilon} = f(t_*)/\bar{F}(t_*)$	$\epsilon = f(t_*)/F(t_*)$

$(i = 1, \ldots, c)$. Namely, in the affine scaling of Eq. (5) the chains' common length l is the positive scale parameter, and the reference time t_* is the real location parameter.

Our goal is to analyze the limiting behavior of the chains' scaled failure times in the case of a multitude of long chains: $c \to \infty$ and $l \to \infty$. To that end we set our focus on the ensemble of the chains' scaled failure times: $\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$. Also, to that end we introduce the following *asymptotic geometric coupling* between the number c of the chains and the common length l of the chains: $c\bar{F}(t_*)^l \simeq 1$. Specifically, the asymptotic geometric coupling is given by the limit

$$\lim_{c\to\infty, l\to\infty} c\bar{F}(t_*)^l = 1. \tag{6}$$

With the affine scaling of Eq. (5), and the asymptotic geometric coupling of Eq. (6), we are now in position to state the following Poisson-process limit-law result.

Proposition 1. The ensemble $\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$ converges in law, in the limit of Eq. (6), to a limiting ensemble $\mathcal{P}$ that is a Poisson process over the real line with the following intensity function: $\lambda(x) = \bar{\epsilon}\exp(-\bar{\epsilon}x)$ $(-\infty < x < \infty)$, where $\bar{\epsilon} = f(t_*)/\bar{F}(t_*)$.

See the Appendix for the proof of Proposition 1. Table I summarizes Proposition 1 and its underlying setting. We now elaborate on the meaning of this Proposition.

A *Poisson process* is a countable collection of points that are scattered randomly over its domain, according to certain Poisson-process statistics that are determined by its *intensity function* [30–32]. Poisson processes are of key importance in probability theory, and their applications range from insurance and finance [8] to queueing systems [33], and from fractals [34] to power laws [35].

In the case of the Poisson process $\mathcal{P}$ of Proposition 1 the domain is the real line $(-\infty < x < \infty)$, and the intensity function is $\lambda(x) = \bar{\epsilon}\exp(-\bar{\epsilon}x)$. The points of the Poisson process $\mathcal{P}$ of Proposition 1 manifest, in the limit of Eq. (6), the chains' scaled failure times. The informal meaning of the intensity function $\lambda(x)$ is the following: the probability that the infinitesimal interval $(x, x + dx)$ contains a point of the Poisson process $\mathcal{P}$ is $\lambda(x)dx$, and this probability is independent of the scattering of points outside the interval $(x, x + dx)$.

The exponent $\bar{\epsilon} = f(t_*)/\bar{F}(t_*)$ of the intensity function $\lambda(x)$ manifests the *hazard rate* of the generic failure time T at time t_* [21,22]: $\bar{\epsilon}$ is the likelihood that the generic link will fail right after time t_*, conditioned on the information that the generic link did not fail up to time t_*. Specifically, this hazard rate is given by the following limit:

$$\bar{\epsilon} = \lim_{\Delta\to 0} \frac{1}{\Delta} \Pr(T \leqslant t_* + \Delta | T > t_*). \tag{7}$$

The hazard rate is a widely applied tool in reliability engineering and in risk management [21,22].

III. MAX-MIN

With Proposition 1 at our disposal, we now set the focus on the *strongest chain*, i.e., the last chain standing. The failure time of the strongest chain is given by the *maximum* of the chains' failure times:

$$\wedge_{\max} = \max\{\wedge_1, \ldots, \wedge_c\}. \tag{8}$$

Namely, the random variable $\wedge_{\max}$ is the *max-min* over the entries of the random matrix $\mathbf{T}$: for each and every row of the matrix pick the minimal entry, and then pick the rows' largest minimal entry.

As with the chains' failure times, we apply the affine scaling of Eq. (5) to the failure time of the strongest chain:

$$\tilde{\wedge}_{\max} = l(\wedge_{\max} - t_*), \tag{9}$$

where t_* is the above reference time. Also, as with the ensemble $\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$, we analyze the limiting behavior of the random variable $\tilde{\wedge}_{\max}$ in the case of a multitude of long chains: $c \to \infty$ and $l \to \infty$.

Here and hereinafter $\mathcal{G}$ denotes a "standard" *Gumbel* random variable. Namely, $\mathcal{G}$ is a real-valued random variable whose statistics are governed by the following "standard" Gumbel distribution function:

$$\Pr(\mathcal{G} \leqslant t) = \exp[-\exp(-t)] \tag{10}$$

$(-\infty < t < \infty)$. We note that within the three extreme-value laws, Gumbel is the only law whose support is the entire real line.

The three extreme-value laws have one-to-one correspondences with the maximal points of specific Poisson processes [36]. In particular, the Gumbel extreme-value law has a one-to-one correspondence with the maximal point of the Poisson process $\mathcal{P}$ of Proposition 1. This connection leads to the following Gumbel limit-law result.

Proposition 2. The random variable $\tilde{\wedge}_{\max}$ converges in law, in the limit of Eq. (6), to a limiting random variable $\bar{\eta}\mathcal{G}$, where $\bar{\eta} = \bar{F}(t_*)/f(t_*)$, and where $\mathcal{G}$ is the "standard" Gumbel random variable of Eq. (10).

See the Appendix for the proof of Proposition 2. Table II summarizes Proposition 2 and its underlying setting. In Fig. 1 we use numerical simulations to demonstrate Proposition 2.

TABLE II. Summary of Propositions 2 and 4. Rows 1–2 summarize the quantities under consideration and their affine scalings. Rows 3–5 summarize the Gumbel limit-law results: the required asymptotic geometric couplings of c and l, the limiting Gumbel random variables (the convergences being in law), and the coefficients of the limiting Gumbel random variables. The term $\mathcal{G}$ appearing in row 4 is the standard Gumbel random variable of Eq. (10).

	Proposition 2	Proposition 4
1. Quantity	$\wedge_{\max} = \max\{\wedge_1, \ldots, \wedge_c\}$	$\vee_{\min} = \min\{\vee_1, \ldots, \vee_l\}$
2. Scaling	$\tilde{\wedge}_{\max} = l(\wedge_{\max} - t_*)$	$\tilde{\vee}_{\min} = c(\vee_{\min} - t_*)$
3. Coupling	$\lim_{c\to\infty,\, l\to\infty} c\bar{F}(t_*)^l = 1$	$\lim_{l\to\infty,\, c\to\infty} lF(t_*)^c = 1$
4. Limit	$\lim_{c\to\infty,\, l\to\infty} \tilde{\wedge}_{\max} = \bar{\eta}\mathcal{G}$	$\lim_{l\to\infty,\, c\to\infty} \tilde{\vee}_{\min} = -\eta\mathcal{G}$
5. Coefficient	$\bar{\eta} = \bar{F}(t_*)/f(t_*)$	$\eta = F(t_*)/f(t_*)$

To that end nine different distributions of the generic failure time T are considered: exponential, gamma, log-normal, inverse-Gauss, uniform, Weibull, beta, Pareto, and normal. In all nine cases, the convergence of the simulations to the theoretical prediction of Proposition 2 is evident. See the Appendix for the MATLAB code that was used in order to generate the numerical simulations.

Proposition 2 yields an asymptotic approximation for the max-min of large random matrices with dimensions $c>l\gg 1$. Indeed, consider the matrix dimensions (c and l) and the reference time (t_*) to satisfy the relation $c\bar{F}(t_*)^l \simeq 1$. Then, the max-min random variable $\wedge_{\max}$ admits the following *Gumbel asymptotic approximation* in law:

$$\wedge_{\max} \simeq \mathcal{G}_* := t_* + \frac{\bar{\eta}}{l}\mathcal{G}, \tag{11}$$

where $\bar{\eta}$ and $\mathcal{G}$ are as in Proposition 2.

The Gumbel asymptotic approximation of Eq. (11) has the following meaning: the deterministic asymptotic approximation of the max-min $\wedge_{\max}$ is the reference time t_*, the magnitude of the random fluctuations about the deterministic asymptotic approximation is $\bar{\eta}/l$, and the statistics of the random fluctuations about the deterministic asymptotic approximation are Gumbel. Table III summarizes the Gumbel asymptotic approximation of Eq. (11), and details the key statistical features of this approximation.

IV. MIN-MAX

So far we addressed the *max-min* of the random matrix $\mathbf{T}$: pick the minimum of each row $\wedge_i = \min\{T_{i,1}, \ldots, T_{i,l}\}$ ($i = 1, \ldots, c$), and then pick the maximum of these minima $\wedge_{\max} = \max\{\wedge_1, \ldots, \wedge_c\}$. Analogously, we can address the *min-max* of the random matrix $\mathbf{T}$: pick the maximum of each column

$$\vee_j = \max\{T_{1,j}, \ldots, T_{c,j}\} \tag{12}$$

($j = 1, \ldots, l$), and then pick the minimum of these maxima

$$\vee_{\min} = \min\{\vee_1, \ldots, \vee_l\}. \tag{13}$$

To illustrate the min-max $\vee_{\min}$ consider the collection of the aforementioned c chains to be copies of a given DNA strand. The chains' l links represent l sites along the DNA strand, where each of these sites codes a critical information item. The links' generic failure time T manifests the time at which the information coded by a specific DNA site is damaged; namely, the matrix entry $T_{i,j}$ is the time at which the jth information item on the ith DNA copy is damaged. The jth information item is lost once all its c copies are damaged, and hence the failure time of the jth information item is given by Eq. (12). As all the l information items are critical, a system failure occurs once any of the l information items is lost. Hence, the time of the system failure is given by the *min-max* of Eq. (13).

TABLE III. Summary of the max-min and the min-max Gumbel asymptotic approximations of Eqs. (11) and (18). Rows 1–3 summarize the approximations: the required asymptotics, the resulting approximations, and the coefficients of the magnitudes of the approximations' stochastic parts. The term $\mathcal{G}$ appearing in row 2 is the standard Gumbel random variable of Eq. (10). Rows 4–7 summarize the approximations' key statistical features: modes, medians, means, and standard deviations (SD). The term γ appearing in row 6 is the Euler-Mascheroni constant: $\gamma = 0.577\ldots$.

	Max-Min	Min-Max
1. Asymptotics	$l \gg 1$ and $c\bar{F}(t_*)^l \simeq 1$	$c \gg 1$ and $lF(t_*)^c \simeq 1$
2. Approximation	$\mathcal{G}_* := t_* + \frac{\bar{\eta}}{l}\mathcal{G}$	$\mathcal{G}_* := t_* - \frac{\eta}{c}\mathcal{G}$
3. Coefficient	$\bar{\eta} = \bar{F}(t_*)/f(t_*)$	$\eta = F(t_*)/f(t_*)$
4. Mode	t_*	t_*
5. Median	$t_* - \ln[\ln(2)]\bar{\eta}\frac{1}{l}$	$t_* + \ln[\ln(2)]\eta\frac{1}{c}$
6. Mean	$t_* + \gamma\bar{\eta}\frac{1}{l}$	$t_* - \gamma\eta\frac{1}{c}$
7. SD	$\frac{\pi\bar{\eta}}{\sqrt{6}}\frac{1}{l}$	$\frac{\pi\eta}{\sqrt{6}}\frac{1}{c}$

More generally, the min-max $\vee_{\min}$ applies to a setting in which l critical information items are stored on c different backup copies, where $j = 1, \ldots, l$ is the index of the information items, $i = 1, \ldots, c$ is the index of the copies, and $T_{i,j}$ is the time at which the jth information item on the ith backup copy is damaged. The above "DNA model" was for the sake of illustration, following the "chains model" of Sec. II, which we used in order to illustrate the max-min.

The analysis presented above was with regard to the max-min. Analogous analysis holds with regard to the min-max. Indeed, consider the above reference time t_*, and apply the following affine scaling to the failure time of the jth information item:

$$\tilde{\vee}_j = c(\vee_j - t_*) \tag{14}$$

($j = 1, \ldots, l$). Namely, in the affine scaling of Eq. (14) the number c of the copies is the positive scale parameter, and the reference time t_* is the real location parameter.

Also, introduce an *asymptotic geometric coupling* between the number l of the information items and the number c of the copies: $lF(t_*)^c \simeq 1$. Specifically, the asymptotic geometric coupling is given by the limit

$$\lim_{l\to\infty,\, c\to\infty} lF(t_*)^c = 1. \tag{15}$$

With the affine scaling of Eq. (14) and the asymptotic geometric coupling of Eq. (15), we are now in position to state the following counterpart of Proposition 1.

Proposition 3. The ensemble $\{\tilde{\vee}_1, \ldots, \tilde{\vee}_l\}$ converges in law, in the limit of Eq. (15), to a limiting ensemble $\mathcal{P}$ that is a Poisson process over the real line with the following intensity function: $\lambda(x) = \epsilon \exp(\epsilon x)$ $(-\infty < x < \infty)$, where $\epsilon = f(t_*)/F(t_*)$.

See the Appendix for the proof of Proposition 3. Table I summarizes Proposition 3 and its underlying setting. The notion of Poisson processes was described right after Proposition 1. The exponential intensity function $\lambda(x) = \epsilon \exp(\epsilon x)$ of Proposition 3, and the Poisson process $\mathcal{P}$ that this intensity characterizes, are most intimately related to the notion of *accelerating change* [37]; readers interested in a detailed analysis of the (rich) statistical structure of this Poisson process are referred to [37]. The exponent $\epsilon = f(t_*)/F(t_*)$ has the following limit interpretation:

$$\epsilon = \lim_{\Delta \to 0} \frac{1}{\Delta} \Pr(T > t_* - \Delta | T \leqslant t_*), \tag{16}$$

which is a time-reversal analog of the hazard rate of Eq. (7).

Continuing on from Proposition 3, and considering the above reference time t_*, we apply the affine scaling of Eq. (14) to the time of the system failure:

$$\tilde{\vee}_{\min} = c(\vee_{\min} - t_*). \tag{17}$$

Then, as Proposition 1 led to Proposition 2, Proposition 3 leads to the following Gumbel limit-law result, which is the min-max counterpart of Proposition 2.

Proposition 4. The random variable $\tilde{\vee}_{\min}$ converges in law, in the limit of Eq. (15), to a limiting random variable $-\eta\mathcal{G}$, where $\eta = F(t_*)/f(t_*)$, and where $\mathcal{G}$ is the "standard" Gumbel random variable of Eq. (10).

See the Appendix for the proof of Proposition 4. Table II summarizes Proposition 4 and its underlying setting. Proposition 4 yields an asymptotic approximation for the min-max of large random matrices with dimensions $l > c \gg 1$. Indeed, consider the matrix dimensions (l and c) and the reference time (t_*) to satisfy the relation $lF(t_*)^c \simeq 1$. Then, the min-max random variable $\vee_{\min}$ admits the following *Gumbel asymptotic approximation* in law:

$$\vee_{\min} \simeq \mathcal{G}_* := t_* - \frac{\eta}{c}\mathcal{G}, \tag{18}$$

where η and $\mathcal{G}$ are as in Proposition 4.

The Gumbel asymptotic approximation of Eq. (18) is the min-max counterpart of the max-min Gumbel asymptotic approximation of Eq. (11). Specifically, the deterministic asymptotic approximation of the min-max $\vee_{\min}$ is the reference time t_*, the magnitude of the random fluctuations about the deterministic asymptotic approximation is η/c, and the statistics of the random fluctuations about the deterministic asymptotic approximation are Gumbel. Table III summarizes the Gumbel asymptotic approximation of Eq. (18) and details the key statistical features of this approximation.

V. APPLICATION

The Gumbel asymptotic approximations of Eqs. (11), (18) can be applied in two modalities: as *approximation tools* and as *design tools* for the max-min and the min-max, respectively. Both applications are based on the fact that—for Eqs. (11) and (18) to hold—it is required that the matrix dimensions (c and l) and the reference time (t_*) be properly coupled. In this section we describe and demonstrate these applications.

We start with the max-min and its Gumbel asymptotic approximation of Eq. (11). This approximation requires the following coupling between the matrix-dimensions and the reference time: $c\bar{F}(t_*)^l \simeq 1$, where $c > l \gg 1$. Consequently, if the matrix dimensions are given ($c > l \gg 1$) then the approximation of Eq. (11) holds with the following *implied reference time*:

$$t_* = \bar{F}^{-1}\left[\left(\frac{1}{c}\right)^{1/l}\right]. \tag{19}$$

For example, if $c = 2^l$ then the implied reference time is the *median* of the generic failure time T. This application is an *approximation tool*: given the random matrix $\mathbf{T}$, Eq. (11) with the implied reference time of Eq. (19) approximates the max-min of the matrix.

To demonstrate the *design-tool* application of the Gumbel asymptotic approximation of Eq. (11), consider a system with a "series-parallel" architecture: the system is a parallel array of c subsystems (labeled $i = 1, \ldots, c$), and each subsystem is a serial array of l components (labeled $j = 1, \ldots, l$). In terms of the random matrix $\mathbf{T}$ of Eq. (3), the failure time of component j in subsystem i is $T_{i,j}$. The series-parallel architecture implies that the system's failure time is the max-min $\wedge_{\max}$. Now, assume that our goal is to design a system whose failure time has the following properties: its deterministic approximation is t_* and the magnitude of its random fluctuations about its deterministic approximation is $\bar{m}$, where t_* and $\bar{m}$ are specified target values. Then, to meet the goal, the dimensions of the system should be designed as follows:

$$l \simeq \frac{1}{\bar{m}} \frac{\bar{F}(t_*)}{f(t_*)} \quad \text{and} \quad c \simeq \frac{1}{\bar{F}(t_*)^l}. \tag{20}$$

Let us turn now to the min-max and its Gumbel asymptotic approximation of Eq. (18). This approximation requires the following coupling between the matrix-dimensions and the reference time: $lF(t_*)^c \simeq 1$, where $l > c \gg 1$. Consequently, if the matrix dimensions are given ($l > c \gg 1$) then the approximation of Eq. (18) holds with the following *implied reference time*:

$$t_* = F^{-1}\left[\left(\frac{1}{l}\right)^{1/c}\right]. \tag{21}$$

For example, if $l = 2^c$ then the implied reference time is the *median* of the generic failure time T. This application is an *approximation tool*: given the random matrix $\mathbf{T}$, Eq. (18) with the implied reference time of Eq. (21) approximates the min-max of the matrix.

To demonstrate the *design-tool* application of the Gumbel asymptotic approximation of Eq. (18), consider a system with a "parallel-series" architecture: the system is a serial array of l subsystems (labeled $j = 1, \ldots, l$), and each subsystem is a parallel array of c components (labeled $i = 1, \ldots, c$). In terms of the random matrix $\mathbf{T}$ of Eq. (3), the failure time of component i in subsystem j is $T_{i,j}$. The parallel-series architecture implies that the system's failure time is

the min-max $\vee_{\min}$. Now, assume that our goal is to design a system whose failure time has the following properties: its deterministic approximation is t_* and the magnitude of its random fluctuations about its deterministic approximation is m, where t_* and m are specified target values. Then, to meet the goal, the dimensions of the system should be designed as follows:

$$c \simeq \frac{1}{m}\frac{F(t_*)}{f(t_*)} \quad \text{and} \quad l \simeq \frac{1}{F(t_*)^c}. \tag{22}$$

Equations (19) and (21) are explicit formulas facilitating the approximation of the max-min and min-max of large random matrices. Equations (20) and (22) are explicit formulas facilitating the design of systems with, respectively, series-parallel and parallel-series architectures. The practical implementation of these formulas is easy and straightforward.

VI. DISCUSSION

We opened this paper with the fundamental theorem of EVT, and with a short discussion of the extreme-value asymptotic approximations emerging from this theorem. We now continue with this discussion, and expand it to include the Gumbel asymptotic approximations of Eqs. (11) and (18), as well as the asymptotic approximation emanating from the central limit theorem (CLT) of probability theory [38,39]. To that end we begin with a succinct review of the CLT.

As in the case of the fundamental theorem of EVT, the CLT applies to ensembles of IID real-valued random variables, $\{X_1, \ldots, X_n\}$, where the ensemble's n components are IID copies of a general real-valued random variable X. The input X is assumed to have a finite (positive) standard deviation σ, and hence also a finite (real) mean μ. We consider the ensemble's average $A_n = (X_1 + \cdots + X_n)/n$, and further consider the following affine scaling of this average:

$$\tilde{A}_n = \frac{1}{\sigma}\sqrt{n}(A_n - \mu)\,. \tag{23}$$

Equation (23) is the CLT counterpart of Eq. (1), with the term $\sqrt{n}/\sigma$ assuming the role of the positive scale parameter [s_n in Eq. (1)] and with the mean μ assuming the role of the real location parameter [δ_n in Eq. (1)].

The CLT asserts that the scaled average $\tilde{A}_n$ converges in law (as $n \to \infty$) to a limiting random variable $\mathcal{N}$ that is "standard" normal; i.e., the statistics of the limit $\mathcal{N}$ are normal (Gauss) with zero mean and with unit variance. Consequently, for a given large ensemble ($n \gg 1$), the ensemble's average A_n admits the following *normal asymptotic approximation* in law:

$$A_n \simeq \mathcal{N}_* := \mu + \frac{\sigma}{\sqrt{n}}\mathcal{N}. \tag{24}$$

The normal asymptotic approximation of Eq. (24) has the following meaning: the deterministic asymptotic approximation of the ensemble's average A_n is the mean μ, the magnitude of the random fluctuations about the deterministic asymptotic approximation is $\sigma/\sqrt{n}$, and the statistics of the random fluctuations about the deterministic asymptotic approximation are normal.

It is illuminating to compare the extreme-value asymptotic approximation of Eq. (2), the normal asymptotic approximation of Eq. (24), and the Gumbel asymptotic approximations of Eqs. (11) and (18). Such a comparison will highlight the analogies and the differences between these asymptotic approximations, as we shall now see.

The extreme-value asymptotic approximation of Eq. (2) has the following key features. (I) The domains of attraction are characterized by *narrow tail conditions*: regular-variation conditions for the Weibull and Frechet extreme-value laws, and a complicated condition for the Gumbel extreme-value law (see Theorems 8.13.2–8.13.4 in [16], and [28]). (II) The deterministic asymptotic approximation δ_n is highly dependent on the input X. (III) The fluctuations' magnitude $1/s_n$ is highly dependent on the input X. (IV) The limit $\mathcal{L}$ is either Weibull, Frechet, or Gumbel. (V) The information required in order to apply this asymptotic approximation is infinite-dimensional: the input's distribution function.

The normal asymptotic approximation of Eq. (24) has the following key features. (I) The domain of attraction is characterized by a *wide moment condition*: inputs X with a finite variance. (II) The deterministic asymptotic approximation μ is the input's mean. (III) The fluctuations' magnitude $\sigma/\sqrt{n}$ depends on the input X only via the coefficient σ (which is the input's standard deviation); hence the asymptotic order $O(1/\sqrt{n})$ of the fluctuations' magnitude is independent of the input X. (IV) The limit $\mathcal{N}$ is standard normal. (V) The information required in order to apply this asymptotic approximation is two-dimensional: the input's mean and standard deviation.

The Gumbel asymptotic approximations of Eqs. (11) and (18)—for a preset reference time t_*—have the following key features. (I) The domain of attraction is characterized by a *wide smoothness condition*: inputs T with a density function. (II) The deterministic asymptotic approximation t_* is the preset reference time. (III) The fluctuations' magnitudes $\bar{\eta}/l$ and η/c depend on the input T only via the coefficients $\bar{\eta}$ and η, respectively; hence the asymptotic orders $O(1/l)$ and $O(1/c)$ of the fluctuations magnitudes are independent of the input T. (IV) The limit $\mathcal{G}$ is standard Gumbel. (V) The information required in order to apply these asymptotic approximations is two-dimensional: the values of the input's distribution function and density function at the reference time t_*.

On the one hand, the key features of the Gumbel asymptotic approximations of Eqs. (11) and (18) are quite different from those of the extreme-value asymptotic approximation of Eq. (2). On the other, the key features of these Gumbel asymptotic approximations are markedly similar to those of the normal asymptotic approximation of Eq. (24). Thus, the Gumbel asymptotic approximations presented here are "as universal" as the normal asymptotic approximation; the similarities between these approximations are summarized in Table IV.

As its name suggests, a cornerstone of the central limit theorem (CLT) is its *centrality*. In terms of the normal asymptotic approximation of Eq. (24), centrality is manifested as follows: the ensemble's average A_n is approximated about the "center point" of the input X, its mean μ. In effect, the CLT "magnifies" the statistical behavior of the ensemble's average A_n about the center point μ.

The fundamental theorem of EVT is diametric to the CLT. Indeed, denote by x^* the upper bound of the support of the input X; this upper bound can be either finite ($x^* < \infty$) or

TABLE IV. Summary of the similarities between the max-min and the min-max Gumbel asymptotic approximations of Eqs. (11) and (18) (for a preset reference time t_*) and the normal asymptotic approximation of Eq. (24). In row 1, the term $\mathcal{G}$ is the standard Gumbel random variable of Eq. (10), and the term $\mathcal{N}$ is a standard normal random variable (i.e., with zero mean and with unit variance). Rows 2–4 summarize the approximations' structures: their deterministic parts, the coefficients of the magnitudes of their stochastic parts, and the orders of the magnitudes of their stochastic parts. For the normal column, $f(x)$ $(-\infty < x < \infty)$ denotes the density function of the input X, and n is the size of the ensemble $\{X_1, \ldots, X_n\}$ (see Sec. VI for the details).

	Max-Min	Min-Max	Normal
1. Approximation	$\mathcal{G}_* := t_* + \frac{\bar{\eta}}{l}\mathcal{G}$	$\mathcal{G}_* := t_* - \frac{\eta}{c}\mathcal{G}$	$\mathcal{N}_* := \mu + \frac{\sigma}{\sqrt{n}}\mathcal{N}$
2. Deterministic	t_*	t_*	$\mu = \int_{-\infty}^{\infty} x f(x) dx$
3. Coefficient	$\bar{\eta} = \bar{F}(t_*)/f(t_*)$	$\eta = F(t_*)/f(t_*)$	$\sigma = \sqrt{\int_{-\infty}^{\infty}(x-\mu)^2 f(x) dx}$
4. Order	$O(1/l)$	$O(1/c)$	$O(1/\sqrt{n})$

infinite ($x^* = \infty$). Specifically, in the Weibull case x^* is finite, in the Frechet case x^* is infinite, and in the Gumbel case x^* is either (see Theorems 8.13.2–8.13.4 in [16] and [28]). In effect, the fundamental theorem of EVT magnifies the statistical behavior of the ensemble's maximum M_n about the upper bound x^*.

Thus, on the one hand, the normal asymptotic approximation of Eq. (24) "anchors" at the mean μ, which is an *interior point* of the range spanned by the lower bound and upper bound of the input's supports. And, on the other hand, the extreme-value asymptotic approximation of Eq. (2) anchors at the upper bound x^*, which is a *boundary point* of the support of the input X. So, also from an anchoring perspective, the Gumbel asymptotic approximations of Eqs. (11) and (18) are different from the extreme-value asymptotic approximation of Eq. (2), and are similar to the normal asymptotic approximation of Eq. (24). Indeed, these Gumbel asymptotic approximations anchor at the reference time t_*, which is an *interior point* of the support of the input T.

Notably, in the design-tool modality, the Gumbel asymptotic approximations of Eqs. (11) and (18) offer a feature that even the CLT does not offer: *tunability*. The center point at which the normal asymptotic approximation of Eq. (24) anchors is the mean μ, and this anchoring point is *fixed*. The center point at which the Gumbel asymptotic approximations of Eqs. (11) and (18) anchor is the reference time t_*, and this anchoring point is *tunable*. Namely, Propositions 1–4 allow us to set the reference time t_* as we wish within the support of the input T.

Perhaps the most straightforward approach to tackle the max-min and the min-max of random matrices is to apply the fundamental theorem of EVT iteratively. Reis and Castro did precisely so for the max-min [27]: they applied the fundamental theorem first to the minimum of each and every row of the random matrix $\mathbf{T}$ [of Eq. (3)] and then to the maximum of the rows' minima. Interestingly, the results of Reis and Castro and our results both yield Gumbel limit laws. Nonetheless, these seemingly identical limit-law results are profoundly different. The details and features of the approaches are important, as we shall now elucidate.

Consider the iterative EVT approach. The first iteration of the fundamental theorem implicitly confines the input T to one of the theorem's narrow domains of attraction (Weibull, Frechet, and Gumbel); moreover, as noted above, this iteration anchors at the the upper bound of the support of the input T. To apply the second iteration one has to impose further conditions, as well as to introduce an asymptotic coupling between the dimensions of the random matrix $\mathbf{T}$. Consequently, the iterative EVT approach comes with an expensive "intricacy price tag." Specifically, for the limit law of [27] the following are highly dependent on the input T, and are also highly elaborate: the max-min domain of attraction, scaling scheme, and asymptotic coupling. Matters are as intricate also in the max-min and min-max results of [25,26] (which are derived via "EVT machineries").

Here, rather than mimicking the fundamental theorem of EVT, we mimicked the CLT. First, we set a vast domain of attraction: inputs T with a density function. Second, we devised particular asymptotic couplings and affine scalings: Eqs. (6) and (9) for the max-min and Eqs. (15) and (17) for the min-max. Third, we showed that these particular asymptotic couplings and affine scalings always yield the Gumbel limit laws of Propositions 2 and 4; i.e., they do so for all inputs T that belong to the vast domain of attraction. These novel Gumbel limit laws were achieved via a Poisson-process approach: the bedrock Poisson-process limit laws of Propositions 1 and 3. This approach enabled us to circumvent the use of the fundamental theorem of EVT.

The Gumbel limit laws of Propositions 2 and 4 are truly workable tools for the max-min and the min-max of random matrices with IID entries. In turn, so are the Gumbel asymptotic approximations of Eqs. (11) and (18). A short MATLAB code given in the Appendix shows just how easy it is to apply these tools in practice.

VII. CONCLUSION

This paper explored the max-min value $\wedge_{\max}$ and the min-max value $\vee_{\min}$ of a random matrix $\mathbf{T}$ with c rows, l columns, and entries that are IID real-valued random variables. This IID setting is common to random-matrix theory, to the fundamental theorem of extreme value theory, and to the central limit theorem. The max-min and the min-max values of matrices emerge naturally in science and engineering, e.g., in game theory and in reliability engineering. We motivated the max-min value $\wedge_{\max}$ by the following question: Within a collection of c IID chains, each with l links, how long will the strongest chain hold? And, we motivated the min-max value $\vee_{\min}$ by the following question: How long can l critical information items be stored reliably on c IID backup copies?

We showed that if the number of rows c and the number of columns l are large, and are coupled geometrically, then the max-min value $\wedge_{\max}$ and the min-max value $\vee_{\min}$ admit, respectively, the Gumbel asymptotic approximations of Eq. (11) and of Eq. (18) (in law). These Gumbel asymptotic approximations are similar, in form, to the normal asymptotic approximation that follows from the central limit theorem. Moreover, in their design-tool modality, the Gumbel asymptotic approximations display a special feature: their deterministic part—the reference time t_*—is tunable. Hence, these Gumbel asymptotic approximations can be used, via Eqs. (20) and (22), to design the max-min and min-max values.

The Gumbel asymptotic approximations are founded on the Gumbel limit laws of Propositions 2 and 4. In turn, the Gumbel limit laws are founded on the bedrock Poisson-process limit laws of Propositions 1 and 3. These four novel limit laws have a vast domain of attraction, have simple affine scalings, and use geometric asymptotic couplings (of c and l). With their generality, their CLT-like structure, their straightforward practical implementation, and their many potential applications, the results established and presented in this paper are expected to serve diverse audiences in science and engineering.

ACKNOWLEDGMENTS

R.M. acknowledges Deutsche Forschungsgemeinschaft for funding (ME 1535/7-1) and support from the Foundation for Polish Science within an Alexander von Humboldt Polish Honorary Research Scholarship. S.R. gratefully acknowledges support from the Azrieli Foundation and the Sackler Center for Computational Molecular and Materials Science.

APPENDIX

1. A general Poisson-process limit-law result

In this subsection we establish a general Poisson-process limit-law result. The setting of the general result is as follows. Consider $X_1, \ldots, X_n$ to be n IID copies of a generic random variable X. The random variable X is real-valued, and its density function is given by

$$f_\theta(x) = \kappa_\theta g_\theta(x) \tag{A1}$$

$(-\infty < x < \infty)$, where θ is a positive parameter κ_θ is a positive constant, and $g_\theta(x)$ is a non-negative function.

Consider the joint limits $n \to \infty$ and $\theta \to \infty$. We assume that the parameter n and the constant κ_θ admit the following asymptotic coupling:

$$\lim_{n\to\infty,\,\theta\to\infty} n\kappa_\theta = \kappa, \tag{A2}$$

where κ is a positive limit value. Also, we assume that

$$\lim_{\theta\to\infty} g_\theta(x) = g(x) \tag{A3}$$

$(-\infty < x < \infty)$, where $g(x)$ is a non-negative limit function.

Now, let us analyze the asymptotic statistical behavior of the ensemble $\{X_1, \ldots, X_n\}$ in the joint limits $n \to \infty$ and $\theta \to \infty$. To that end we take a real-valued "test function" $\phi(x)$ $(-\infty < x < \infty)$ and compute the *characteristic functional* of the ensemble $\{X_1, \ldots, X_n\}$ with respect to this test function:

$$\begin{aligned}
&\mathbf{E}[\phi(X_1)\cdots\phi(X_n)] \\
&= \mathbf{E}[\phi(x)]^n = \left\{\int_{-\infty}^{\infty} \phi(x) f_\theta(x) dx\right\}^n \\
&= \left\{1 - \int_{-\infty}^{\infty} [1-\phi(x)] f_\theta(x) dx\right\}^n \\
&= \left\{1 - \frac{1}{n}\int_{-\infty}^{\infty} [1-\phi(x)][(n\kappa_\theta) g_\theta(x)] dx\right\}^n
\end{aligned} \tag{A4}$$

[in Eq. (A4) we used the IID structure of the ensemble $\{X_1, \ldots, X_n\}$, and Eq. (A1)]. Applying the limits of Eqs. (A2) and (A3), Eq. (A4) implies that

$$\begin{aligned}
&\lim_{n\to\infty,\,\theta\to\infty} \mathbf{E}[\phi(X_1)\cdots\phi(X_n)] \\
&\quad = \exp\left\{-\int_{-\infty}^{\infty} [1-\phi(x)][\kappa g(x)] dx\right\}.
\end{aligned} \tag{A5}$$

The *characteristic functional* of a Poisson process $\mathcal{P}$ over the real line, with intensity function $\lambda(x)$ $(-\infty < x < \infty)$, is given by [30]:

$$\mathbf{E}\left[\prod_{x\in\mathcal{P}} \phi(x)\right] = \exp\left\{-\int_{-\infty}^{\infty} [1-\phi(x)]\lambda(x) dx\right\}, \tag{A6}$$

where $\phi(x)$ $(-\infty < x < \infty)$ is a real-valued test function. We emphasize that the characteristic functional of Eq. (A6) is indeed characteristic [30]: if $\mathcal{P}$ is collection of real points that satisfies Eq. (A6), then $\mathcal{P}$ is a Poisson process over the real line, with intensity function $\lambda(x)$ $(-\infty < x < \infty)$. Hence, combined together, Eqs. (A5) and (A6) yield the following general result:

Proposition 5. The ensemble $\{X_1, \ldots, X_n\}$ converges in law, in the joint limits $n \to \infty$ and $\theta \to \infty$, to a Poisson process $\mathcal{P}$ over the real line with intensity function $\lambda(x) = \kappa g(x)$ $(-\infty < x < \infty)$.

2. Proof of Proposition 1

Equation (4) implies that

$$\begin{aligned}
\Pr(\wedge_i > t) &= \Pr[\min\{T_{i,1}, \ldots, T_{i,l}\} > t] \\
&= \Pr(T_{i,1} > t)\cdots\Pr(T_{i,l} > t) \\
&= \Pr(T > t)^l = \bar{F}(t)^l
\end{aligned} \tag{A7}$$

$(-\infty < t < \infty)$. Equations (5) and (A7) imply that

$$\begin{aligned}
\Pr(\tilde{\wedge}_i > t) &= \Pr[l(\wedge_i - t_*) > t] \\
&= \Pr\left(\wedge_i > t_* + \frac{t}{l}\right) = \bar{F}\left(t_* + \frac{t}{l}\right)^l
\end{aligned} \tag{A8}$$

$(-\infty < t < \infty)$. Differentiating Eq. (A8) with respect to the variable t implies that the density function of the scaled random variable $\tilde{\wedge}_i$ is given by

$$-\frac{d}{dt}\Pr(\tilde{\wedge}_i > t) = \bar{F}\left(t_* + \frac{t}{l}\right)^l \bar{h}\left(t_* + \frac{t}{l}\right) \tag{A9}$$

$(-\infty < t < \infty)$, where $\bar{h}(t) = f(t)/\bar{F}(t)$. In what follows we use the shorthand notation $\bar{\epsilon} = \bar{h}(t_*)$. Note that the two basic

requirements $0 < F(t_*) < 1$ and $0 < f(t_*) < \infty$ imply that $0 < \bar{\epsilon} < \infty$.

Now, apply Proposition 5 to the following setting: $n = c$, $\theta = l$, and $X_i = \tilde{\wedge}_i$ $(i = 1, \ldots, c)$. Equation (A9) implies that

$$f_\theta(x) = \underbrace{\bar{F}(t_*)^\theta}_{\kappa_\theta} \underbrace{\left[\frac{\bar{F}\left(t_* + \frac{x}{\theta}\right)}{\bar{F}(t_*)}\right]^\theta \bar{h}\left(t_* + \frac{x}{\theta}\right)}_{g_\theta(x)} \tag{A10}$$

$(-\infty < x < \infty)$. Note that

$$\left[\frac{\bar{F}\left(t_* + \frac{x}{\theta}\right)}{\bar{F}(t_*)}\right]^\theta = \left[\frac{\bar{F}(t_*) - f(t_*)\frac{x}{\theta} + o\left(\frac{1}{\theta}\right)}{\bar{F}(t_*)}\right]^\theta = \left[1 - \frac{\bar{\epsilon}x}{\theta} + o\left(\frac{1}{\theta}\right)\right]^\theta \underset{\theta\to\infty}{\longrightarrow} \exp(-\bar{\epsilon}x) \tag{A11}$$

$(-\infty < x < \infty)$. Equations (A10) and (A11) imply that

$$\lim_{\theta\to\infty} g_\theta(x) = g(x) := \bar{\epsilon}\exp(-\bar{\epsilon}x) \tag{A12}$$

$(-\infty < x < \infty)$. Also, the asymptotic geometric coupling of Eq. (6) implies that the asymptotic coupling of Eq. (A2) holds with $\kappa = 1$. Hence, the result of Proposition 5 holds with the intensity function

$$\lambda(x) = \bar{\epsilon}\exp(-\bar{\epsilon}x) \tag{A13}$$

$(-\infty < x < \infty)$. This proves Proposition 1.

3. Proof of Proposition 2

Set $\mathcal{P}$ to be a Poisson process, over the real line, with intensity function $\lambda(x) = \bar{\epsilon}\exp(-\bar{\epsilon}x)$ $(-\infty < x < \infty)$ and exponent $\bar{\epsilon} = f(t_*)/\bar{F}(t_*)$. Consider the number of points $N(t)$ of the Poisson process $\mathcal{P}$ that reside above a real threshold t. The Poisson-process statistics imply that the number $N(t)$ is a Poisson-distributed random variable with mean

$$\mathbf{E}[N(t)] = \int_t^\infty \lambda(x)dx = \int_t^\infty \bar{\epsilon}\exp(-\bar{\epsilon}x)dx = \exp(-\bar{\epsilon}t). \tag{A14}$$

Now, consider the *maximal point* M of the Poisson process $\mathcal{P}$. This maximal point is no larger than the threshold t if and only if no points of the Poisson process $\mathcal{P}$ reside above this threshold: $\{M \leqslant t\} \Leftrightarrow \{N(t) = 0\}$. Hence, as $N(t)$ is a Poisson-distributed random variable with mean $\mathbf{E}[N(t)]$, Eq. (A14) implies that the distribution function of the maximal point M is given by

$$\Pr(M \leqslant t) = \exp[-\exp(-\bar{\epsilon}t)] \tag{A15}$$

$(-\infty < t < \infty)$. The distribution function of Eq. (A15) characterizes the Gumbel law. A standard Gumbel-distributed random variable $\mathcal{G}$ is governed by the distribution function of Eq. (10): $\Pr(\mathcal{G} \leqslant t) = \exp[-\exp(-t)]$ $(-\infty < t < \infty)$. Equations (A15) and (10) imply that the maximal point M admits the following Gumbel representation in law:

$$M = \bar{\eta}\mathcal{G}, \tag{A16}$$

where

$$\bar{\eta} = \frac{1}{\bar{\epsilon}} = \frac{\bar{F}(t_*)}{f(t_*)}. \tag{A17}$$

Proposition 1 established that the ensemble $\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$ converges in law—in the limit of Eq. (6)—to the Poisson process $\mathcal{P}$. Consequently, the maximum $\tilde{\wedge}_{\max}$ of the ensemble $\{\tilde{\wedge}_1, \ldots, \tilde{\wedge}_c\}$ converges in law—in the limit of Eq. (6)—to the maximal point M of the Poisson process $\mathcal{P}$. Hence, Eq. (A16) proves Proposition 2.

4. Proof of Proposition 3

For the random variable $\vee_j = \max\{T_{1,j}, \ldots, T_{c,j}\}$ we have

$$\begin{aligned}\Pr(\vee_j \leqslant t) &= \Pr[\max\{T_{1,j}, \ldots, T_{c,j}\} \leqslant t] \\ &= \Pr(T_{1,j} \leqslant t)\cdots\Pr(T_{c,j} \leqslant t) \\ &= \Pr(T \leqslant t)^c = F(t)^c\end{aligned} \tag{A18}$$

$(-\infty < t < \infty)$. In turn, for the scaled random variable $\tilde{\vee}_j = c(\vee_j - t_*)$ Eq. (A18) implies that

$$\Pr(\tilde{\vee}_j \leqslant t) = \Pr[c(\vee_j - t_*) \leqslant t] = \Pr\left(\vee_j \leqslant t_* + \frac{t}{c}\right) = F\left(t_* + \frac{t}{c}\right)^c \tag{A19}$$

$(-\infty < t < \infty)$. Differentiating Eq. (A19) with respect to the variable t implies that the density function of the scaled random variable $\tilde{\vee}_j$ is given by

$$\frac{d}{dt}\Pr(\tilde{\vee}_j \leqslant t) = F\left(t_* + \frac{t}{c}\right)^c h\left(t_* + \frac{t}{c}\right) \tag{A20}$$

$(-\infty < t < \infty)$, where $h(t) = f(t)/F(t)$. In what follows we use the shorthand notation $\epsilon = h(t_*)$. Note that the two basic requirements $0 < F(t_*) < 1$ and $0 < f(t_*) < \infty$ imply that $0 < \epsilon < \infty$.

Now, apply Proposition 5 to the following setting: $n = l$, $\theta = c$, and $X_i = \tilde{\vee}_j$ $(j = 1, \ldots, l)$. Equation (A20) implies that

$$f_\theta(x) = \underbrace{F(t_*)^\theta}_{\kappa_\theta} \underbrace{\left[\frac{F\left(t_* + \frac{x}{\theta}\right)}{F(t_*)}\right]^\theta h\left(t_* + \frac{x}{\theta}\right)}_{g_\theta(x)} \tag{A21}$$

$(-\infty < x < \infty)$. Note that

$$\left[\frac{F\left(t_* + \frac{x}{\theta}\right)}{F(t_*)}\right]^\theta = \left[\frac{F(t_*) + f(t_*)\frac{x}{\theta} + o\left(\frac{1}{\theta}\right)}{F(t_*)}\right]^\theta = \left[1 + \frac{\epsilon x}{\theta} + o\left(\frac{1}{\theta}\right)\right]^\theta \underset{\theta\to\infty}{\longrightarrow} \exp(\epsilon x) \tag{A22}$$

$(-\infty < x < \infty)$. Equations (A21) and (A22) imply that

$$\lim_{\theta\to\infty} g_\theta(x) = g(x) := \epsilon\exp(\epsilon x) \tag{A23}$$

$(-\infty < x < \infty)$. Also, the asymptotic geometric coupling of Eq. (15) implies that the asymptotic coupling of Eq. (A2) holds with $\kappa = 1$. Hence, the result of Proposition 5 holds with the intensity function

$$\lambda(x) = \epsilon\exp(\epsilon x) \tag{A24}$$

$(-\infty < x < \infty)$. This proves Proposition 3.

5. Proof of Proposition 4

Set $\mathcal{P}$ to be a Poisson process, over the real line, with intensity function $\lambda(x) = \epsilon \exp(\epsilon x)$ $(-\infty < x < \infty)$ and exponent $\epsilon = f(t_*)/F(t_*)$. Consider the number of points $N(t)$ of the Poisson process $\mathcal{P}$ that reside up to a real threshold t. The Poisson-process statistics imply that the number $N(t)$ is a Poisson-distributed random variable with mean

$$\begin{aligned}\mathbf{E}[N(t)] &= \int_{-\infty}^{t} \lambda(x)dx \\ &= \int_{-\infty}^{t} \epsilon \exp(\epsilon x)dx = \exp(\epsilon t). \end{aligned} \tag{A25}$$

Now, consider the *minimal point* M of the Poisson process $\mathcal{P}$. This minimal point is larger than the threshold t if and only if no points of the Poisson process $\mathcal{P}$ reside up to this threshold: $\{M > t\} \Leftrightarrow \{N(t) = 0\}$. Hence, as $N(t)$ is a Poisson-distributed random variable with mean $\mathbf{E}[N(t)]$, Eq. (A25) implies that the survival function of the minimal point M is given by

$$\Pr(M > t) = \exp[-\exp(\epsilon t)] \tag{A26}$$

$(-\infty < t < \infty)$. A standard Gumbel-distributed random variable $\mathcal{G}$ is governed by the distribution function of Eq. (10): $\Pr(\mathcal{G} \leqslant t) = \exp[-\exp(-t)]$ $(-\infty < t < \infty)$. Equations (A26) and (10) imply that the minimal point M admits the following Gumbel representation in law:

$$M = -\eta\mathcal{G}, \tag{A27}$$

where

$$\eta = \frac{1}{\epsilon} = \frac{F(t_*)}{f(t_*)}. \tag{A28}$$

Proposition 3 established that the ensemble $\{\tilde{\forall}_1, \ldots, \tilde{\forall}_l\}$ converges in law—in the limit of Eq. (15)—to the Poisson process $\mathcal{P}$. Consequently, the minimum $\tilde{\forall}_{\min} = \min\{\tilde{\forall}_1, \ldots, \tilde{\forall}_l\}$ of the ensemble $\{\tilde{\forall}_1, \ldots, \tilde{\forall}_l\}$ converges in law—in the limit of Eq. (15)—to the minimal point M of the Poisson process $\mathcal{P}$. Hence, Eq. (A27) proves Proposition 4.

6. MATLAB code for Fig. 1

```
% This function computes the scaled MaxMin/eta_bar

% N specifies the number of random matrices to be generated
N=10^5;

% MaxMin will hold the N max-min values that will be
% computed
MaxMin=zeros(1,N);

% pd specifies the distribution of the random matrix entries
pd = makedist('Exponential', 'mu', 1);

% CDF_t specifies the value of the cumulative distribution
% function at the anchor point
CDF_t=1/5;

% This computes the anchor point t by inverting cumulative
% distribution function
t=icdf(pd,CDF_t);

% l sets the number of links
l=70;

% c sets the number of chains via geometric coupling
c=floor((1-CDF_t)^(-l));

% This for-loop generates the random matrices and computes
% the MaxMin
for k=1:N
   M=random(pd,c,l);
   MaxMin(k)=max(min(M'));
end

% This computes the coefficient eta_bar
eta_bar=(1-CDF_t)/pdf(pd,t);

% This computes the scaled MaxMin/eta_bar
MaxMin=(MaxMin-t)*l/eta_bar;
```


[1] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics* (Wiley, New York, 1978).

[2] S. Coles, S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics Vol. 208 (Springer, London, 2001).

[3] L. De Haan and A. Ferreira, *Extreme Value Theory: An Introduction* (Springer, New York, 2007).

[4] E. Castillo, *Extreme Value Theory in Engineering* (Elsevier, New York, 2012).

[5] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications* (World Scientific, Singapore, 2000).

[6] J. Beirlant, Y. Goegebeur, J. Segers, and J. L. Teugels, *Statistics of Extremes: Theory and Applications* (Wiley, New York, 2006).

[7] R. D. Reiss, M. Thomas, and R. D. Reiss, *Statistical Analysis of Extreme Values* (Birkhäuser, Basel, 2007), Vol. 2.

[8] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: For insurance and Finance*, Stochastic Modelling and Applied Probability Vol. 33 (Springer, Berlin, 2013).

[9] W. J. Scheirer, Extreme Value theory-based methods for visual recognition, Synth. Lect. Comput. Visi., **7**, 1 (2017).

[10] R. A. Fisher and L. H. C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Cambr. Philos. Soc. **24**, 180 (1928).

[11] B. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math. **44**, 423 (1943); translated and reprinted in *Breakthroughs in Statistics I*, edited by S. Kotz and N. L. Johnson (Springer, New York, 1992), pp. 195–225.

[12] W. Weibull, *A Statistical Theory of Strength of Materials* (Ingeniörsvetenskapsakademiens Handlingar, Stockholm, 1939).

[13] W. Weibull, A statistical distribution function of wide applicability, J. Appl. Mech. **18**, 293 (1951).
[14] M. Fréchet, Sur la loi de probabilité de l'écart maximum, Ann. Soc. Math. Polon. **6**, 93 (1927).
[15] E. J. Gumbel, *Statistics of Extremes* (Courier, North Chelmsford, MA, 2012).
[16] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular Variation*, Encyclopedia of Mathematics and its Applications Vol. 27 (Cambridge University Press, Cambridge, UK, 1989).
[17] D. Fudenberg and J. Tirole, *Game Theory* (MIT Press, Cambridge, MA, 1991), p. 80.
[18] M. Maschler, E. Solan, and S. Zamir, *Game Theory* (Cambridge University Press, Cambridge, UK, 2013).
[19] J. Pearl, Asymptotic properties of minimax trees and game-searching procedures, Artif. Intell. **14**, 113 (1980).
[20] T. A. Khan, L. Devroye, and R. Neininger, A limit law for the root value of minimax trees, Electron. Commun. Probab. **10**, 273 (2005).
[21] R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability*, Classics in Applied Mathematics Vol. 17 (SIAM, Philadelphia, 1996).
[22] M. Finkelstein, *Failure Rate Modelling for Reliability and Risk* (Springer, London, 2008).
[23] K. Kolowrocki, Limit reliability functions of some series-parallel and parallel-series systems, Appl. Math. Comput. **62**, 129 (1994).
[24] K. Kolowrocki, On a class of limit reliability functions of some regular homogeneous series-parallel systems, Reliab. Eng. Syst. Safety **39**, 11 (1993).
[25] K. Kolowrocki, On asymptotic reliability functions of series-parallel and parallel-series systems with identical components, Reliab. Eng. Syst. Safety **41**, 251 (1993).
[26] H. Chernoff and H. Teicher, Limit distributions of the minimax of independent identically distributed random variables, Trans. Am. Math. Soc. **116**, 474 (1965).
[27] P. Reis and L. C. Castro, Limit model for the reliability of a regular and homogeneous series-parallel system, Revstat **7**, 227 (2009).
[28] A. A. Balkema and L. De Haan, On R. von Mises' condition for the domain of attraction of $\exp(-\exp(-x))$, Ann. Math. Stat. **43**, 1352 (1972).
[29] I. Eliazar, R. Metzler, and S. Reuveni, Gumbel central limit theorem for max-min and min-max, Phys. Rev. E **100**, 020104 (2019).
[30] J. F. C. Kingman, *Poisson Processes*, Oxford Studies in Probability Vol. 3 (Clarendon Press, Oxford, 1992).
[31] D. R. Cox and V. Isham, *Point Processes*, Monographs on Statistics & Applied Probability Vol. 12 (CRC Press Boca Raton, FL, 1980).
[32] R. L. Streit, *Poisson Point Processes: Imaging, Tracking, and Sensing* (Springer, New York, 2010).
[33] R. W. Wolff, *Stochastic Modeling and the Theory of Queues* (Pearson, London, 1989).
[34] S. B. Lowen and M. C. Teich, *Fractal-Based Point Processes*, Wiley Series in Probability and Statistics Vol. 366 (Wiley, New York, 2005).
[35] I. Eliazar and J. Klafter, A probabilistic walk up power laws, Phys. Rep. **511**, 143 (2012).
[36] I. Eliazar and I. M. Sokolov, Gini characterization of extreme-value statistics, Physica A **389**, 4462 (2010).
[37] I. Eliazar and M. F. Shlesinger, Universality of accelerating change, Physica A **494**, 430 (2018).
[38] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 2008), Vol. 1.
[39] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 2008), Vol. 2.