Probing tails of energy distributions using importance-sampling in the disorder with a guiding function

Mathias Körner and Helmut G. Katzgraber
Theoretische Physik, ETH Zürich, CH-8093 Zürich, Switzerland

Alexander K. Hartmann
Institut für Theoretische Physik, Universität Göttingen,
Friedrich-Hund-Platz 1, 37077 Göttingen, Germany
(Dated: March 13, 2006)

We propose a simple and general procedure based on a recently introduced approach that uses an importance-sampling Monte Carlo algorithm in the disorder to probe to high precision the tails of ground-state energy distributions of disordered systems. Our approach requires an estimate of the ground-state energy distribution as a guiding function which can be obtained from simple-sampling simulations. In order to illustrate the algorithm, we compute the ground-state energy distribution of the Sherrington-Kirkpatrick mean-field Ising spin glass to eighteen orders of magnitude. We find that the ground-state energy distribution in the thermodynamic limit is well fitted by a modified Gumbel distribution as previously predicted, but with a value of the slope parameter m which is clearly larger than 6 and of the order 11.

PACS numbers: 75.50.Lk, 75.40.Mg, 05.50.+q

I. INTRODUCTION

The ground-state energy distributions of disordered systems and in particular of spin glasses [1–4] have recently received considerable attention from different groups. Bouchaud *et al.* [5] have studied the ground-state energy distribution of short-range spin glasses in two and three dimensions and find them to become Gaussian in the thermodynamic limit, while other groups [6–9] find that the ground-state energy distribution of the mean-field Sherrington-Kirkpatrick (SK) model [10] remains skewed in the thermodynamic limit and is well described by a modified Gumbel distribution [11]. In a recent paper, we have been able to show that this transition from a Gaussian to a non-Gaussian limiting distribution coincides with the transition from long-range to infinite-range behavior for a one-dimensional spin glass with long-range power law interactions [12]. For uncorrelated random variables, when the power of the exponent in the aforementioned model is large, the central limit theorem holds and a Gaussian distribution can be expected. But when the interactions are infinite-ranged and the variables are strongly correlated, such as in the case of the SK model, deviations from the central limit theorem can be expected, and, as previously mentioned, the data follow a modified Gumbel distribution with a slope parameter expected to be [6] $m \approx 6$. This result is rather puzzling: What is the relationship between the ground-state energy distribution of the SK model and extreme value statistics? Not until Bertin and Clusel [13] analytically derived the equivalence between extreme value statistics and random sums of correlated variables it was understood that there is not necessarily an underlying extreme process in the mean-field SK model. Results by Bramwell *et al.* [14, 15] on large-scale critical fluctuations in correlated systems hinted towards such a behavior because the parameter m [see Eq. (10)], expected to be integer when studied from the context of extreme value statistics, is noninteger.

Traditionally [6, 12], simple-sampling techniques, where N_{samp} disorder realizations are computed and subsequently binned in order to obtain the ground-state energy distribution, are used. If N_{samp} samples are computed, then the maximal "resolution" of a bin is $\sim 1/N_{\mathrm{samp}}$, and thus $\sim 10^6$ samples have to be computed to resolve six orders of magnitude in the histogram. Therefore, the probing of tails in ground-state energy distributions becomes quickly intractable making it difficult to determine if a given fitting function properly describes the tails of a ground-state energy distribution, as is the case for the SK model. In this paper we outline a simple algorithm based on ideas presented in Ref. 16 that overcomes this problem by performing an importance-sampling simulation of the ground-state energy distribution *in the disorder* with a *guiding function* computed from simple-sampling simulations. Similar approaches have been used before in the studies of distributions of sequence alignment scores [16], as well as fluctuations in classical magnets [17] (albeit the latter without disorder).

By computing the tails of the disorder distribution of the SK model to up to 18 orders of magnitude we show that the slope parameter of the Gumbel distribution is considerably larger ($m \approx 11$) than found in previous studies using simple-sampling techniques [6].

We outline the method in Section II and apply it to the ground-state energy distribution of the SK model in Section III. We conclude in Sec. IV and discuss the implications of our results on studies of ground-state energy distributions.

II. SIMULATION OF GROUND-STATE ENERGY DISTRIBUTIONS

A disordered system is defined by a Hamiltonian $\mathcal{H}_{\mathcal{J}}(\mathcal{C})$, where the disorder configuration $\mathcal{J}$ is chosen from a probability distribution $P(\mathcal{J})$ and $\mathcal{C}$ denotes the phase-space configuration of the system. The ground-state energy E of a given disorder configuration $\mathcal{J}$ is defined by

$$E(\mathcal{J}) = \min_{\mathcal{C}} \mathcal{H}_{\mathcal{J}}(\mathcal{C}). \tag{1}$$

Together with the disorder distribution $\mathcal{P}(\mathcal{J})$, this defines the ground-state energy distribution

$$P(E) = \int d\mathcal{J}\, \mathcal{P}(\mathcal{J})\, \delta\left[E - E(\mathcal{J})\right]. \tag{2}$$

To study the thermodynamic limit of the ground-state energy distribution, we use the standardized form $P_{\mathrm{s}}(x)$ defined by the equation

$$P(E) = \frac{1}{\sigma_E} P_{\mathrm{s}}\left(\frac{E - [E]_{\mathrm{av}}}{\sigma_E}\right), \tag{3}$$

where $[E]_{\mathrm{av}}$ and $\sigma_E = ([E^2]_{\mathrm{av}} - [E]_{\mathrm{av}}^2)^{1/2}$ are the average and the standard deviation of the ground-state energy, respectively [18]. Here $[\cdots]_{\mathrm{av}}$ represents an average over the disorder.

A. Simple sampling

Because $P(E)$ cannot be determined analytically for most disordered systems, approximative methods such as Monte Carlo simulations have to be used. A standard approach is to use a simple-sampling Monte Carlo algorithm to study $P(E)$. N_{samp} independent disorder configurations $\mathcal{J}_i$ are chosen from $P(\mathcal{J})$ and the ground-state energy is calculated for each disorder configuration. The calculation of the ground-state energy in itself is a difficult optimization problem and can often only be solved approximatively [19, 20]. From the ground-state energies of these disorder configurations, the ground-state energy distribution can be estimated as

$$P(E) = \frac{1}{N_{\mathrm{samp}}} \sum_{i=1}^{N_{\mathrm{samp}}} \delta\left[E - E(\mathcal{J}_i)\right], \tag{4}$$

so that the averages of functions with respect to the disorder are replaced by averages with respect to the N_{samp} random samples. The functional form of the ground-state energy distribution and its parameters can for example be estimated by a maximum likelihood fit of a distribution $F_\theta(E)$ with parameters θ to the data [21]. It corresponds to the distribution for which the observed data have the largest probability. Due to the limited range of energies sampled by the simple-sampling algorithm it is often difficult or even impossible to quantify how well the tails of the distribution are described by a maximum-likelihood fit, and improving the data by increasing the number of random samples is generally computationally very expensive. Therefore other methods have to be used to study the tails of the distribution, and we outline in the next section a simple method that can be applied if a good analytic fit to the data can be found.

B. Importance sampling with a guiding function

Assuming that we find a function $F_\theta(E)$ which accurately describes the ground-state energy distribution as calculated with a simple-sampling simulation, we now proceed by sampling the ground-state distribution with an importance-sampling Monte Carlo algorithm *in the disorder* [16, 22, 23] and use $F_\theta(E)$ as a guiding function [24]. We start from a random disorder configuration $\mathcal{J} = \mathcal{J}_0$ with ground-state energy $E(\mathcal{J}_0)$. From the i-th configuration $\mathcal{J}_i$, we generate the $i+1$-th configuration $\mathcal{J}_{i+1}$ by the following Metropolis-type [25] update:

1. Choose a candidate disorder configuration $\mathcal{J}'$ by replacing a subset of $\mathcal{J}$ chosen at random (e.g., a single bond chosen at random) with values chosen according to $P(\mathcal{J})$ [this requires that $P(\mathcal{J})$ can be written in a product form] and calculate its ground-state energy $E(\mathcal{J}')$.

2. Set $\mathcal{J}_{i+1} = \mathcal{J}'$ with probability

$$P_{\mathrm{accept}} = \min\left\{\frac{F_\theta\left[E(\mathcal{J}_i)\right]}{F_\theta\left[E(\mathcal{J}')\right]}, 1\right\} \tag{5}$$

and $\mathcal{J}_{i+1} = \mathcal{J}_i$ otherwise.

Using the importance-sampling Monte Carlo algorithm, a disorder configuration $\mathcal{J}$ is visited with probability $1/F_\theta[E(\mathcal{J})]$, such that the probability to visit a disorder configuration with ground-state energy E is $P(E)/F_\theta(E)$. If $F_\theta(E) = P(E)$, then each energy is visited with the same probability resulting in a flat-histogram sampling of the ground-state energy distribution. To be able to study a finite range of energies and to avoid the trapping of the algorithm in an extremal region of the energy space, the range of energies that the algorithm is allowed to visit can be restricted.

The main difference to the simple-sampling algorithm described in Section II A is that successive configurations visited by the algorithm are not independent. Therefore an analysis of the results requires a quantification of the correlations among configurations visited by the algorithm. This is a standard problem in Markov Chain Monte Carlo simulations, and can for example be solved with the help of the exponential autocorrelation time τ of the energy, which is the number of Monte Carlo steps after which the autocorrelation function of the ground-state energy

$$\chi(\Delta i) = \frac{\langle E_i \, E_{i+\Delta i}\rangle - \langle E_i\rangle \langle E_{i+\Delta i}\rangle}{\langle E_i^2\rangle - \langle E_i\rangle^2}, \tag{6}$$

decays to $1/e$ [23]. Here E_i is the ground state energy after the i-th Monte Carlo step and $\langle \ldots \rangle$ refers to the average over Monte Carlo time. In order to ensure that the visited ground-state configurations are not correlated, we only use every 4τ-th measurement. Once the autocorrelation effects have been quantified, the data can be analyzed with the same methods as the simple-sampling results [see Eq. (4)].

III. APPLICATION: SHERRINGTON-KIRKPATRICK MODEL

The Sherrington-Kirkpatrick [10] model is defined by the Hamiltonian

$$\mathcal{H}_{\mathcal{J}}(\{S_i\}) = \sum_{i<j} J_{ij} \, S_i \, S_j, \tag{7}$$

where the $S_i = \pm 1$ ($i = 1, \ldots, N$) are Ising spins, and the bonds $\mathcal{J} = \{J_{ij}\}$ are identically and independently distributed random variables chosen from a Normal distribution with zero mean and standard deviation $(N-1)^{-1/2}$. The sum is over all spins in the system.

The ground-state energy of a given disorder configuration $\mathcal{J}$ is defined as

$$E(\mathcal{J}) = \min_{\{S_i\}} \mathcal{H}_{\mathcal{J}}(\{S_i\}). \tag{8}$$

We are interested in the thermodynamic limit of the ground-state energy distribution and define the limiting distribution $P_\infty(x)$ in analogy to Eq. (3) by

$$P(E) = \frac{1}{\sigma_E} P_\infty \left(\frac{E - [E]_{\rm av}}{\sigma_E} \right). \tag{9}$$

For the SK model several optimization algorithms, such es extremal optimization [26], hysteretic optimization [27], as well as other algorithms such as genetic and Bayesian algorithms [19, 20], and even parallel tempering [12, 28–30] are available. In this work we compute the ground states of the system for the sake of simplicity using parallel tempering Monte Carlo [31]. Note that a combination of the proposed importance-sampling simulation in the disorder with other more efficient algorithms might yield more detailed results. We start by studying the 10^5 ground-state energies calculated for different system sizes in Ref. 12. We bin the data into 50 bins and perform a maximum-likelihood fit with a modified Gumbel distribution [6, 11]

$$G_{\mu,\nu,m}(E) \propto \exp \left[m \, \frac{E-\mu}{\nu} - m \, \exp \left(\frac{E-\mu}{\nu} \right) \right]. \tag{10}$$

TABLE I: Maximum-likelihood fit to the simple-sampling Monte Carlo data calculated for the SK model in Ref. 12. For each system size N, $N_{\rm samp} = 10^5$ random samples have been generated. The maximum-likelihood fit has been performed by binning the data into 50 bins. The error bars and the parametric estimates of $[E]_{\rm av}$ and σ_E [see Eqs. (11) and (12)] have been generated by a bootstrap method (see text).

N	μ	ν	m	$[E]_{\rm av}$ [a]	σ_E [a]
16	−10.373(6)	2.48(4)	4.92(16)	−10.634(4)	1.180(3)
24	−16.189(6)	2.81(5)	4.99(17)	−16.481(4)	1.326(3)
32	−22.076(6)	3.23(6)	5.59(19)	−22.375(5)	1.432(3)
48	−33.935(8)	3.81(8)	6.24(26)	−34.249(5)	1.590(4)
64	−45.848(10)	3.97(9)	5.87(26)	−46.196(5)	1.712(4)
96	−69.840(9)	4.53(9)	6.37(25)	−70.205(6)	1.867(5)
128	−93.918(13)	4.89(14)	6.49(39)	−94.305(6)	1.997(5)

[a] Parametric estimate calculated from μ, ν, and m.

The modified Gumbel distribution is parameterized by the "location" parameter μ, the "width" parameter ν, and the "slope" parameter m. To obtain error bars we generate 200 bootstrap replicates [32] of the data and perform the maximum-likelihood fit for each bootstrap replicate. Because the average and standard deviation of the modified Gumbel distribution are related to the parameters μ, ν, and m by

$$[E]_{\rm av} \;=\; \mu + E_m \, \nu \tag{11}$$

$$\sigma_E \;=\; \sigma_m \, \nu, \tag{12}$$

where E_m and σ_m are the average and standard deviation of $G_{0,1,m}$, a parametric estimate of the average energy and its standard deviation can be calculated for the data (and each bootstrap replicate). The results of the analysis are summarized in Table I. Note that for the simple-sampling data we find $m \approx 6$ as suggested in Ref. 6 and that the parametric estimates of $[E]_{\rm av}$ and σ_E agree within error bars with the direct estimates calculated in Ref. 12 using Eqs. (4) and (5) therein.

Importance-sampling simulation

Using the maximum-likelihood estimates shown in Table I, we now perform an importance-sampling simulation in the disorder as described in Section II B. To perform a step in the Monte Carlo algorithm, we choose a site at random, replace all bonds connected to this site (the expected change in the ground-state energy is then of the order $\sim 1/N$), calculate the ground-state energy of the new configuration, and accept the new configuration with the probability given in Eq. (5), which in this case is a modified Gumbel distribution, Eq. (10), with the parameters listed in Table II. To avoid a trapping of the simulation in the double-exponential forward tail of the distribution, we limit the maximum energy allowed in the simulation by $[E]_{\rm av} + 3\sigma_E$. Initial runs indicated that the

TABLE II: Input parameters for the guiding function used in the simulation. Initial runs using the parameters shown in Table I indicated that $m > 6$, so that we have chosen $m = 8$ for the production runs. The parameters μ and ν have been calculated from $[E]_{\rm av}$, σ_e, and m with the help of Eqs. (11) and (12).

N	$[E]_{\rm av}$	σ_e	m	μ	ν
16	-10.635	1.180	8	-10.429	3.233
24	-16.481	1.325	8	-16.249	3.631
32	-22.321	1.469	8	-22.064	4.025
48	-34.248	1.589	8	-33.970	4.355
64	-70.205	1.710	8	-45.961	4.685
96	-94.305	1.867	8	-69.789	5.116
128	-142.627	2.186	8	-93.956	5.472

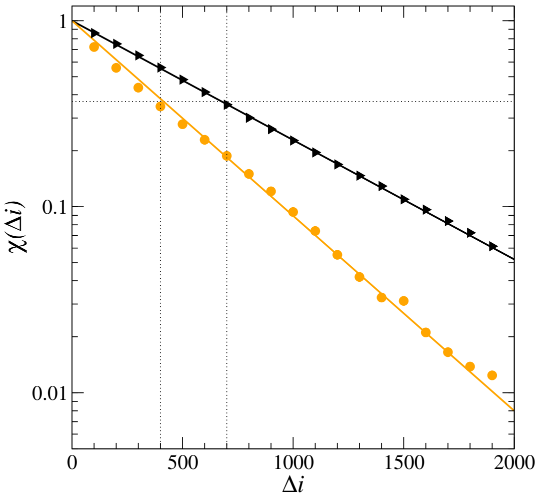

FIG. 1: Autocorrelation function as defined in Eq. (6) for system sizes $N = 16$ (circles) and $N = 128$ (triangles) for the simulation with parameters shown in Table II. The value $1/e$ is marked by the horizontal dotted line.

estimates of m shown in Table I are too small, and we therefore choose an estimate of $m = 8$ for the production runs. μ and ν are determined from the simple-sampling results for the average and standard deviation with the help of Eqs. (11) and (12), and a summary of the input parameters used is shown in Table II. It is important to note that this change in the parameters does not lead to a systematic error or bias in the results and merely constitutes a change of the guiding function, which can either improve or degrade the range of energies visited by the algorithm. Figure 1 shows the energy-energy autocorrelation function for system sizes $N = 16$ and 128. Autocorrelation times are of the order of 400 to 700 Monte Carlo steps resulting in 92 to 686 independent measure-

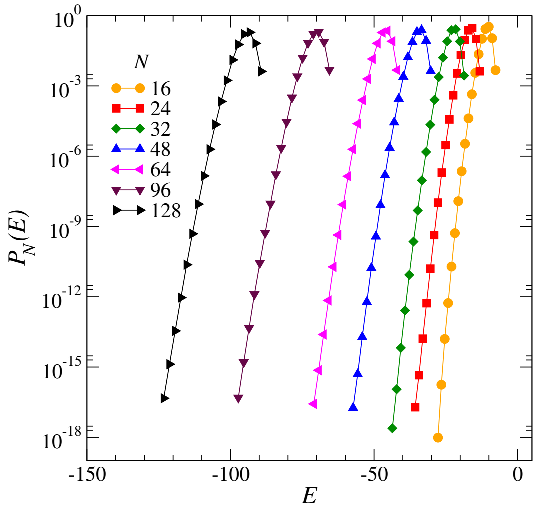

FIG. 2: Unscaled ground-state energy distributions of the Sherrington-Kirkpatrick model for different system sizes, obtained by the guiding-function simulation with the parameters given in Table II. For each system size, between 92 and 686 independent samples are simulated.

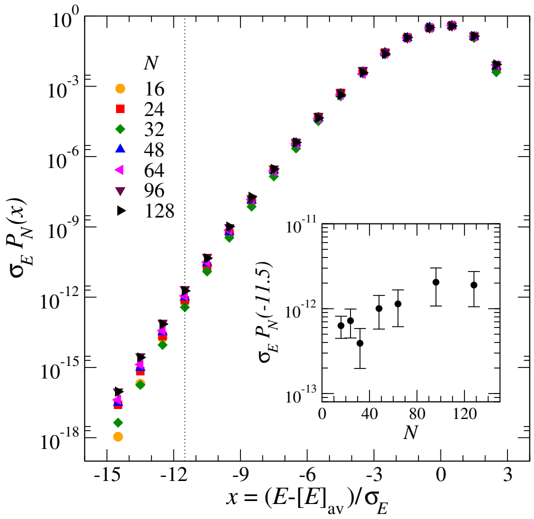

FIG. 3: Scaled ground-state energy distributions of the Sherrington-Kirkpatrick model for different system sizes. The curves are scaled by the average energy and standard deviation obtained from simple-sampling simulations shown in Table I. The scaled distributions obtained for different system sizes collapse on a common curve, although finite size effects are clearly visible in the tail. The inset shows the behavior of the ground-state energy distribution at $x = -11.5$ (vertical dotted line in the main panel) as a function of the system size N in order to illustrate the convergence to a limiting value with increasing system size.

TABLE III: Three-parameter fit in the parameters μ, ν, and m to the data [rescaled to $x = (E - [E]_{\rm av})/\sigma_E$] for the SK model. For each system size N, between $N_{\rm samp} = 92$ and 686 independent samples have been generated. $z = \nu/m$ describes the asymptotic behavior of the single-exponential tail (error bar obtained from independent fits). $\chi^2/{\rm dof}$ represents the χ^2 per degree of freedom of the fit [18].

N	$N_{\rm samp}$	μ	ν	m	ν/m	$\chi^2/{\rm dof}$
16	686	−0.059(63)	3.87(32)	13.8(16)	0.279(9)	6.29
24	274	−0.024(41)	3.55(19)	11.9(9)	0.298(6)	1.04
32	311	−0.017(43)	3.31(18)	11.2(9)	0.295(6)	1.47
48	221	0.010(40)	3.50(17)	11.5(8)	0.303(5)	0.77
64	168	0.063(44)	3.42(18)	11.1(8)	0.309(5)	0.72
96	92	0.026(36)	3.44(14)	10.9(6)	0.314(4)	0.23
128	112	0.066(42)	3.39(16)	10.7(7)	0.317(5)	0.42

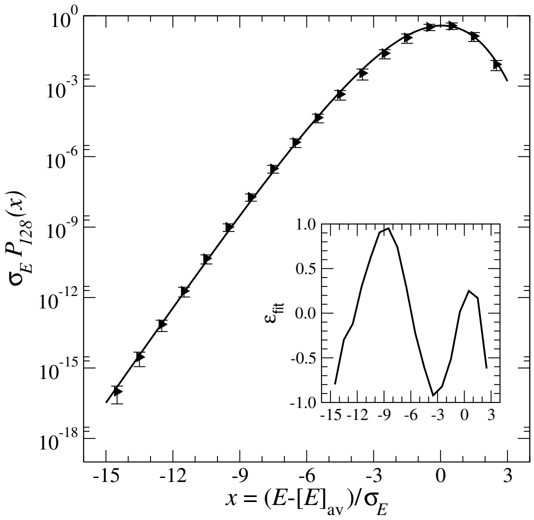

FIG. 4: Scaled ground-state energy distribution for $N = 128$ and the fit to the modified Gumbel distribution as described in the text. The inset shows $\epsilon_{\rm fit} = (P_{\rm fit} - P_{128})/\Delta P_{128}$, the fit deviation normalized by the error of the calculated ground-state energy distribution.

ments for the different system sizes. While the number of samples used is small, the method is able to probe the tails in this particular case down to 18 orders of magnitude, a result *impossible* to obtain with simple-sampling techniques. Figure 2 shows the unscaled ground-state energy distributions obtained from the simulation. Figure 3 shows the same distributions standardized with the average energy and the standard deviation as measured from simple-sampling simulations. We choose to standardize with $[E]_{\rm av}$ and σ_E from the simple-sampling simulations because of the greater number of statistically – for these quantities – relevant measurements. The standardized distributions obtained for different system sizes collapse on a single curve, although finite size effects are clearly visible in the tail of the distribution (see also inset of Figure 3).

Next we try to determine the functional form of the distribution by fitting a modified Gumbel distribution to the data. We have used the standard fitting function of the `gnuplot` package, which implements the nonlinear least-squares (NLLS) Marquardt-Levenberg algorithm. The results of the fit are shown as μ, ν, and m in Table III. In particular, the values for m are approximately two times larger than previous results [6].

Figure 4 shows the ground-state energy distribution for $N = 128$ together with the resulting fit. The data have been standardized by the simple-sampling results. The distribution is well fitted by a modified Gumbel distribution. To examine the systematic deviations more closely, the inset shows the deviation $\epsilon_{\rm fit}$ of the fit from the observed data scaled by the measurement error ΔP obtained from a bootstrap analysis. If no systematic deviations were present, one would expect the deviations to be of order one with an irregularly changing sign. Indeed $\epsilon_{\rm fit}$ varies between -1 and 1, showing the good quality of the fit. Nevertheless, the differences between fit and data have small systematic deviations, which we ascribe to the finite system sizes studied. Note that we also excluded by a visual comparison that the distribution has the form of a Tracy-Widom distribution [7, 33].

The fact that the limiting distribution appears to be a modified Gumbel distribution is confirmed by the fact that the deviations between fitted function and data decrease systematically with system size (not shown). We extrapolate to infinite system size by fitting a power law $m(N) = m_\infty + aN^{-b}$ to the data for the values of m shown in Table III, resulting in an estimated limiting value

$$m_\infty = 10.9(5), \tag{13}$$

where for the fit $\chi^2/{\rm dof} = 0.331$. The data points together with the resulting fit are shown in Fig. 5. The inset shows a similar analysis for $z = \nu/m$ which describes the behavior in the exponential tail of the Gumbel distribution, resulting in $z_\infty = 0.34(4)$. Thus our data suggest hat the limiting distribution is a modified Gumbel distribution with a slope parameter m clearly larger than $m = 6$.

IV. CONCLUSIONS

In this paper we have explained an importance-sampling algorithm with a guiding function to simulate the ground-state energy distribution of a disordered system to high order using a *considerably smaller* numerical effort than with simple-sampling techniques. When compared to full multicanonical simulation schemes to sample distributions, such as used in Ref. 16 and Ref. 17, our algorithm has several advantages due to its simplicity: Instead of iterating towards a good guiding function, which may be quite expensive computationally, we use a maximum likelihood fit as a guiding function. Therefore,

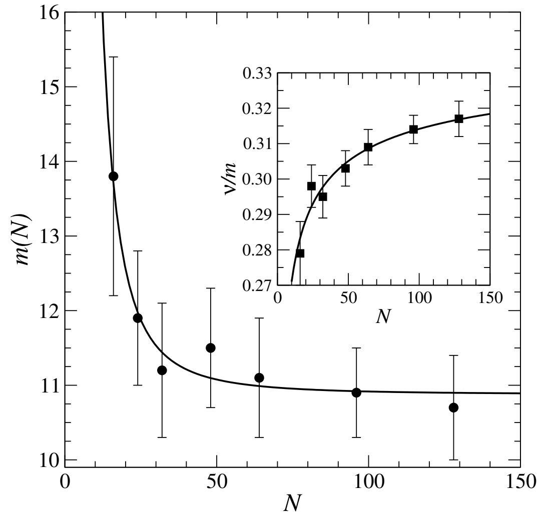

FIG. 5: Finite-size dependence of the parameter m [see Eq. (10)] from the fits to the modified Gumbel distribution, together with a power-law fit (see text). The inset shows a similar analysis for the tail parameter $z = \nu/m$. The data converge to $m_\infty = 10.9(5)$ and $z_\infty = 0.34(4)$, respectively.

the algorithm we propose is straightforward to implement and considerably more efficient than traditional approaches, provided a good guiding function, i.e., a good maximum-likelihood fit to the simple-sampling results, can be found.

We have illustrated the algorithm with the computation of the ground-state energy distribution of the Sherrington-Kirkpatrick model and find that the ground-state energy distribution is well-described for $N \to \infty$ by a modified Gumbel distribution with a slope parameter $m \approx 11$, a value significantly larger than estimated before. We do not expect the estimate of m to change considerably for larger system sizes as the data seem to have converged to a limiting function (see inset to Fig. 3). Note that our results for the mean energy $E_{\rm av}$ as well as the fluctuations σ_E are not influenced by the importance sampling technique because the method probes the tails where probabilities are small and thus contributions to the moments are negligible.

Since the method can be applied more generally to systems where simple-sampling results exist, revisiting the one-dimensional Ising chain with random power-law interactions [12] together with a better ground-state search algorithm would be desirable in order to probe the crossover from mean-field to non-mean-field behavior in more detail.

Acknowledgments

We would like to thank A. P. Young for comments on the manuscript. We also thank M. Troyer for his input and C. Tracy and M. Dieng for providing the Mathematica and Matlab code [33] to calculate the Tracy-Widom distribution. The simulations have been performed on the Hreidar cluster at ETH Zürich. A. K. Hartmann has been funded by the *VolkswagenStiftung* (Germany) within the program "Nachwuchsgruppen an Universitäten."

[1] K. Binder and A. P. Young, *Spin glasses: Experimental facts, theoretical concepts and open questions*, Rev. Mod. Phys. **58**, 801 (1986).

[2] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[3] A. P. Young, ed., *Spin Glasses and Random Fields* (World Scientific, Singapore, 1998).

[4] H. T. Diep, *Frustrated Spin Systems* (World Scientific, Singapore, 2005).

[5] J.-P. Bouchaud, F. Krzakala, and O. C. Martin, *Energy exponents and corrections to scaling in Ising spin glasses*, Phys. Rev. B **68**, 224404 (2003).

[6] M. Palassini, *Ground-state energy fluctuations in the Sherrington-Kirkpatrick model* (2003), (cond-mat/0307713).

[7] A. Andreanov, F. Barbieri, and O. C. Martin, *Large deviations in spin-glass ground-state energies*, European Physical Journal B **41**, 365 (2004).

[8] S. Böttcher, *Extremal Optimization for Sherrington-Kirkpatrick Spin Glasses*, E. Phys. J. B **44**, 317 (2005).

[9] S. Böttcher and T. M. Kott, *Exact Enumeration of Ground States in the Sherrington-Kirkpatrick Spin Glass*, Phys. Rev. B **72**, 212408 (2005).

[10] D. Sherrington and S. Kirkpatrick, *Solvable model of a spin glass*, Phys. Rev. Lett. **35**, 1792 (1975).

[11] E. J. Gumbel, *Multivariate Extremal Distributions*, Bull. Inst. Internat. de Statistique **37**, 471 (1960).

[12] H. G. Katzgraber, M. Körner, F. Liers, M. Jünger, and A. K. Hartmann, *Universality-class dependence of energy distributions in spin glasses*, Phys. Rev. B **72**, 094421 (2005).

[13] E. Bertin and M. Clusel, *Generalised extreme value statistics and sum of correlated variables* (2006), (cond-mat/0601189).

[14] S. T. Bramwell, K. Christensen, J.-Y. Fortin, P. C. W. Holdsworth, H. J. Jensen, S. Lise, J. M. López, M. Nicodemi, J.-F. Pinton, and M. Sellitto, *Universal Fluctuations in Correlated Systems*, Phys. Rev. Lett. **84**, 3744 (2000).

[15] S. T. Bramwell, J. Y. Fortin, P. C. W. Holdsworth, S. Peysson, J. F. Pinton, B. Portelli, and M. Sellitto, *Magnetic fluctuations in the classical XY model: The origin of an exponential tail in a complex system*, Phys. Rev. E **63**, 041106 (2001).

[16] A. K. Hartmann, *Sampling rare events: Statistics of local sequence alignments*, Phys. Rev. E **65**, 056102 (2002).

[17] R. Hilfer, B. Biswal, H. G. Mattutis, and W. Janke, *Multicanonical Monte Carlo study and analysis of tails for the order-parameter distribution of the two-dimensional Ising model*, Phys. Rev. E **68**, 046123 (2003).

[18] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1995).

[19] A. K. Hartmann and H. Rieger, *Optimization Algorithms in Physics* (Wiley-VCH, Berlin, 2001).

[20] A. K. Hartmann and H. Rieger, *New Optimization Algorithms in Physics* (Wiley-VCH, Berlin, 2004).

[21] G. Cowan, *Statistical Data Analysis* (Oxford Science Publications, 1998).

[22] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press Inc., New York, USA, 1999).

[23] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).

[24] Note that the method outlined here has advantages over the method using parallel tempering presented in Ref. 16: Parallel tempering fails when the logarithm of the energy distribution, $\log[P(E)]$, is not concave. Since $d\log[P(E)]/dE$ corresponds to an inverse effective temperature at which the energy E is sampled, different points in the distribution would correspond to the same temperature if the logarithm of the distribution is not concave. Thus parallel tempering would be difficult to tune. A prime example of a system where $\log[P(E)]$ is not concave is when a first-order phase transition is present. In that case $P(E)$ has a two-peak structure.

[25] N. Metropolis and S. Ulam, Journal of the American Statistical Association **44**, 335 (1949).

[26] S. Boettcher and A. G. Percus, *Optimization with Extremal Dynamics*, Phys. Rev. Lett. **86**, 5211 (2001).

[27] K. F. Pal, *Hysteretic optimization for the Sherrington-Kirkpatrick spin glass* (2006), (cond-mat/0601027).

[28] E. Marinari and G. Parisi, *Simulated tempering: A new Monte Carlo scheme*, Europhys. Lett. **19**, 451 (1992).

[29] K. Hukushima and K. Nemoto, *Exchange Monte Carlo method and application to spin glass simulations*, J. Phys. Soc. Jpn. **65**, 1604 (1996).

[30] J. J. Moreno, H. G. Katzgraber, and A. K. Hartmann, *Finding low-temperature states with parallel tempering, simulated annealing and simple Monte Carlo*, Int. J. Mod. Phys. C **14**, 285 (2003).

[31] Details of the parallel-tempering Monte Carlo ground-state search as well as simulation parameters used to compute the ground-state energies of the Sherrington-Kirkpatrick model are presented in Ref. 12, in particular see Appendix A.

[32] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman & Hall, 1994).

[33] RMLab package for computing Tracy-Widom distributions and simulating random matrices, URL `http://math.arizona.edu/~momar/research.htm`.