Self-averaging of digital memcomputing machines

Daniel Primosch, Yuan-Hang Zhang ⓘ, and Massimiliano Di Ventra ⓘ[*]

Department of Physics, University of California San Diego, La Jolla, California 92093, USA

Digital memcomputing machines (DMMs) are a new class of computing machines that employ nonquantum dynamical systems with memory to solve combinatorial optimization problems. Here, we show that the time to solution (TTS) of DMMs follows an inverse Gaussian distribution, with the TTS *self-averaging* with increasing problem size, irrespective of the problem they solve. We provide both an analytical understanding of this phenomenon and numerical evidence by solving instances of the 3-SAT (satisfiability) problem. The self-averaging property of DMMs with problem size implies that they are increasingly insensitive to the detailed features of the instances they solve. This is in sharp contrast to traditional algorithms applied to the same problems, illustrating another advantage of this physics-based approach to computation.

I. INTRODUCTION

Combinatorial optimization problems emerge in a wide variety of applications in both academia and industry [1]. They require finding an assignment of variables—which typically take the values of logical 1 or logical 0—that satisfies a set of constraints. Given one of these problems, e.g., the satisfiability (SAT) with exactly three variables per constraint (known as 3-SAT), one can use different algorithms to find the solution of its instances. Since the various instances of a problem are different, even at a fixed number of variables and constraints, and one can choose different initial conditions for the algorithm, one expects the time to solution (TTS)—namely, the time it takes an algorithm to solve those instances—to vary according to some distribution.

If the problem to solve is hard, such as the 3-SAT, both the mean μ and the standard deviation σ of TTS are expected to increase exponentially with increasing problem size N. We then expect $\sigma^2(N)/\mu^2(N) \sim O(1)$ in the limit $N \to \infty$. We would then say that in this case, the TTS is *not self-averaging*: its relative variance does not go to zero in the thermodynamic limit [2]. (We will provide an explicit example of this phenomenon for traditional algorithms below.)

In recent years, a new computing paradigm, dubbed memcomputing [3–5], has been suggested which employs *time nonlocality* (memory) to tackle computational problems. In particular, its digital version (digital memcomputing machines or DMMs) [6] has been designed to solve problems in the combinatorial optimization class. Unlike traditional algorithms, DMMs map the original problem into nonlinear (nonquantum) dynamical systems whose point attractors are the solutions (if they exist) of such a problem. Since DMMs do not support chaos [7] (as well as quasiperiodic orbits), their ordinary differential equations (ODEs) can be efficiently simulated and tested on our traditional computers, even

before their hardware implementation. These simulations have already shown substantial advantages over state-of-the-art algorithms on a wide range of applications (see, e.g., [5] for a sample of problems already tackled with this paradigm).

In this paper, we show that the physics of DMMs implies that their TTS follows an inverse Gaussian distribution [8], irrespective of the problem they solve. Most importantly, the TTS *self-averages* with increasing problem size, namely, $\sigma^2(N)/\mu^2(N) \to 0$ as $N^{-\theta}$, for $N \to \infty$, with $\theta = 1$ ("strong self-averaging") for physical noise and $\theta < 1$ ("weak self-averaging") for numerical noise. Apart from an analytical understanding of this phenomenon, we corroborate this prediction by numerically solving satisfiable instances of the 3-SAT [9] in the presence of physical noise, and contrast it to the TTS of several well-known traditional algorithms [10–12]. The self-averaging property of DMMs with problem size implies that they are increasingly *insensitive* to the detailed features of the instances they solve. These results both clarify the physics behind these machines and illustrate their advantage in computing.

II. DMMs FOR 3-SAT

Without loss of generality, we will focus on the 3-SAT (with instances taken from [9]), which is a collection of clauses (OR gates) with exactly three Boolean variables, y_i ($i = 1, 2, 3$), with the clauses related by logical conjunctions (AND gates). A DMM (not necessarily unique) for such a problem can be constructed as follows [13]. The variables y_i are first transformed into continuous ones, $v_i \in [-1, 1]$ (which could be voltages in an actual circuit realization of DMMs [6]), where $v_i > 0$ corresponds to $y_i = 1$ and $v_i < 0$ corresponds to $y_i = 0$. The mth clause is represented by $(l_{m,i} \vee l_{m,j} \vee l_{m,k})$, where $l_{m,i} = \bar{y}_i$ or y_i, depending on whether or not y_i is negated, and $\vee$ is the symbol of the OR gate. Each Boolean clause is then mapped into a continuous constraint

* diventra@physics.ucsd.edu

function,

$$C_m(v_i, v_j, v_k) = \tfrac{1}{2} \min(1 - q_{m,i}v_i, 1 - q_{m,j}v_j, 1 - q_{m,k}v_k),$$
(1)

where $q_{m,i} = 1$ if $l_{m_i} = y_i$ and $q_{m,i} = -1$ if $l_{m_i} = \bar{y}_i$. It is clear that the mth clause evaluates to true if and only if $C_m < 1/2$.

Time nonlocality is introduced in the form of additional "memory degrees of freedom," which guarantee that the only fixed points of the dynamics correspond to the solution of the problem, and no other critical points, other than saddle points, are present in the phase space [5]. Following [13], we introduce two additional memory variables for each OR gate: a "short-term" memory $x_{s,m}$ and a "long-term" memory $x_{l,m}$. The dynamics of DMMs for a 3-SAT instance with N variables and M clauses are then ($i = 1, \ldots, N; m = 1, \ldots, M$)

$$\dot{v}_i = \sum_m x_{l,m} x_{s,m} G_{m,i} + (1 + \zeta x_{l,m})(1 - x_{s,m}) R_{m,i}, \quad (2)$$

$$\dot{x}_{s,m} = \beta(x_{s,m} + \epsilon)(C_m - \gamma), \quad x_{s,m} \in [0, 1],$$

$$\dot{x}_{l,m} = \alpha(C_m - \delta), \quad x_{l,m} \in [1, 10^4 M], \quad (3)$$

where the "gradientlike" term $G_{m,i} = \tfrac{1}{2} q_{m,i} \min(1 - q_{m,j}v_j, 1 - q_{m,k}v_k)$, while the "rigidity" term $R_{m,i} = \tfrac{1}{2}(q_{m,i} - v_i)$ if $C_m = \tfrac{1}{2}(1 - q_{m,i}v_i)$, and $R_{m,i} = 0$ otherwise. As in [13], we have chosen the parameters $\alpha = 5, \beta = 20, \gamma = 0.25, \delta = 0.05, \epsilon = 10^{-3}, \zeta = 0.1$.

We refer to [5,13] for a thorough discussion of how Eqs. (2) and (3) have been obtained. Here, we take them at face value as representations of DMMs and point out that they can be compactly written as $\dot{\mathbf{x}}(t) = F(\mathbf{x}(t))$, with $\mathbf{x}$ the collection of continuous variables v_i and memory variables $x_{s,m}, x_{l,m}$, and F the flow vector field that can be read from the right-hand side of Eqs. (2) and (3).

To make direct contact with experiments, as done in [14], we add Gaussian white noise $l(t)$ to the memory variables of Eqs. (3). $l(t)$ satisfies $\langle l(t) \rangle = 0$ and $\langle l(t)l(t') \rangle = \Gamma \delta(t - t')$, where Γ is the noise strength. Of course, the errors introduced by the numerical integration of these ODEs also play the role of "noise," and we refer to the Supplemental Material (SM) [15] for the results in the absence of additive noise.

III. THE DYNAMICS OF DMMs

It was shown both analytically (using supersymmetric topological field theory) and numerically that DMMs, with or without additive noise, find a solution to the problem instance by following specific trajectories ("instantons") in the phase space [16,17]. In particular, DMMs "transition" from a critical point [a saddle point for which $F(\mathbf{x}(t)) = 0$ in Eqs. (2) and (3)] with a certain number of unstable directions to a more stable critical point, until a solution is found. An example of such dynamics for a 3-SAT instance with 3000 variables and 21 000 constraints is provided in Fig. 1, where the sudden jumps of the continuous variables v_i correspond to the instantons connecting critical points. These results have been obtained by integrating the DMMs' Eqs. (2) and (3) using a forward Euler method [18] with fixed integration step ($\Delta t = 0.2$ in arbitrary time units throughout the main text), and no additional noise.
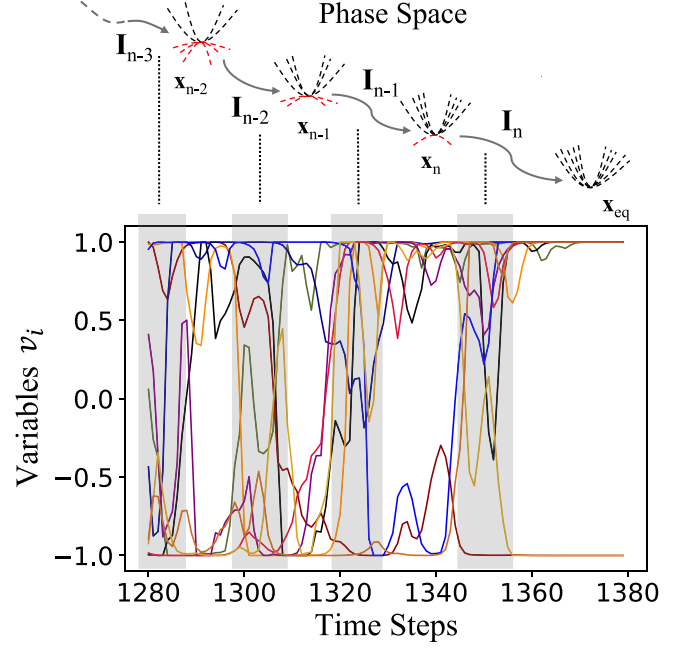

FIG. 1. Top panel: Schematic of the state dynamics of a DMM which transitions from a critical point (e.g., $\mathbf{x}_{n-1}$) to a more stable critical point ($\mathbf{x}_n$) in the phase space in such a way as to reduce the number of unstable directions after each instantonic jump ($\mathbf{I}_{n-1}$). Bottom panel: Numerical example of the DMM represented by Eqs. (2) and (3) solving a 3-SAT instance with 3000 variables and 21 000 constraints. The instantonic jumps are the shaded regions where negative (logical 0) variable values suddenly cross to positive (logical 1) ones, and vice versa. The area in between corresponds to a critical point.

A. Distribution of TTS for DMMs

As anticipated, in the presence of noise (whether physical and/or numerical), we expect the TTS to follow some distribution. In order to determine what type of distribution we expect—irrespective of the problem to solve—we can draw on a physical analogy from the dynamics of DMMs as depicted in Fig. 1. The state vector $\mathbf{x}(t)$ in phase space can be interpreted as "position of a particle" which is *directed* towards the solution and subject to some noise. The particle then has a "drift velocity," $\delta > 0$, which is analogous to how strongly the DMM is driven towards the solution of the problem instance. In this physical system, the noise, modeled by a diffusion constant ν, is analogous to the noise a DMM experiences.

We can then ask how long it takes such a particle to reach an arbitrary distance L (the solution of the problem) from its starting point (the initial condition). We may call this the time to barrier (TTB), which is the physical analog of the TTS of DMMs. This TTB is well known in the literature and is given by an inverse Gaussian probability density function of the form [8]

$$P_{\delta, L, \nu}(t) = \sqrt{\frac{L^2}{2\pi \nu t^3}} e^{-\delta^2 (t - \frac{L}{\delta})^2 / 2\nu t}, \quad (4)$$

where $t > 0$ corresponds to the time for the particle to reach the distance L. We can rewrite Eq. (4) with just two independent parameters by making the substitution $\mu = L/\delta$,
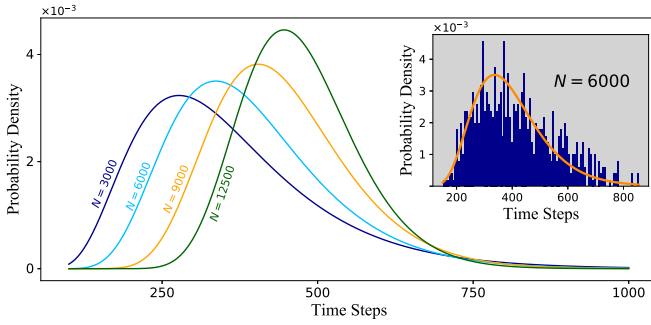
FIG. 2. Various fits (inverse Gaussian distributions) of the TTS (in number of integration steps) of DMMs out of 1000 3-SAT instances (taken from [9]) of varying number of variables, N, at a fixed clause-to-variable ratio of 7. The noise strength is $\Gamma = 0.12$ (in inverse time units). Inset: the histogram of TTS (in number of integration steps) of the DMM with $N = 6000$ variables and 42 000 clauses. The fit is an inverse Gaussian distribution, given by Eq. (5), with $\lambda = 3720 \pm 420$ and $\mu = 390 \pm 10$.

$\lambda = L^2/\nu$. This gives the simplified equation

$$P_{\lambda,\mu}(t) = \sqrt{\frac{\lambda}{2\pi t^3}} e^{-\lambda(t-\mu)^2/2\mu^2 t}. \qquad (5)$$

The mean of this distribution is μ, while its variance is $\sigma^2 = \mu^3/\lambda$ [8]. If the analogy we have just made holds, we should find a similar distribution for the TTS of DMMs.

We verify this analytical prediction by numerically integrating Eqs. (2) and (3), using again a forward Euler method, where we added Gaussian white noise to the memory variables in Eqs. (3) with a strength $\Gamma = 0.12$ (in inverse time units) [19]. The TTS can be simply defined as the amount of integration steps it took the DMM to find a solution. In the inset of Fig. 2, we report the histogram of TTS for a DMM solving 1000 3-SAT instances [9] with $N = 6000$ variables and 42 000 clauses. The histogram is well fitted by the inverse Gaussian distribution (5) with parameters $\lambda = 3720 \pm 420$ and $\mu = 390 \pm 10$ [20]. The main panel of Fig. 2 shows the fits for different numbers of variables with constant clause-to-variable ratio. It indicates a sharpening of the inverse Gaussian distribution with increasing number of variables. (See the SM [15] for more statistics using different integration methods and for different clause-to-variable ratios.)

B. Self-averaging of TTS

To understand this better, we need to determine how this distribution varies with problem size (the number of variables, N, if we fix the clause-to-variable ratio). In particular, we want to know how the ratio σ^2/μ^2 varies with increasing N. To do this, let us again use the physical analogy we have just developed.

We can easily relate the distance L from the initial point of the dynamics to the number of variables, N, by noting that L would correspond to the average radius of an N sphere, $S^N(L)$, in phase space. In the limit of large N, this is simply $L \sim \sqrt{N}$ [21].

The diffusion constant ν is related to the amount of noise in the system. For *physical* noise (such as the additive one we consider here), this is only determined by some external
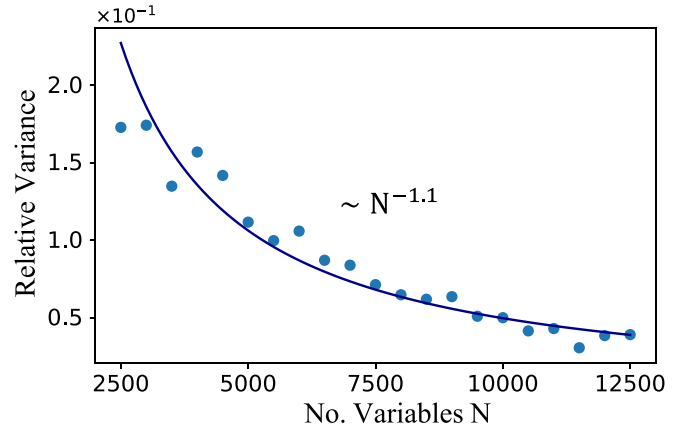

FIG. 3. Relative variance σ^2/μ^2 as a function of number of variables, N, of the TTS distributions of the inset of Fig. 2. The relative variance is well fitted by a power law $\sigma^2/\mu^2 \sim N^{-1.10\pm0.07}$, consistent with self-averaging behavior.

parameter, such as the temperature [2]. In other words, for physical noise, ν is *independent* of N.

Finally, the dependence of the drift velocity δ on N can be obtained as follows. If we have N continuous variables v_i, the initial energy injected into the system is $E = N \times E_0$, where E_0 is the energy associated with a single variable (and at the initial time $t = 0$, the variables are independent of each other; they are coupled immediately by the memory variables for $t > 0$). Assuming the N variables have all unit mass (the mass value is irrelevant as long as it is a constant), the N dependence of the drift velocity is then $\delta \sim \sqrt{E} = \sqrt{N \times E_0}$. This means that the variance $\sigma^2(N) = (L\nu)/\delta^3 \sim \nu/N$.

By putting all this together, we conclude that the TTS of DMMs subject to physical noise (strongly) *self-averages*: $\sigma^2(N)/\mu^2(N) \to 0$ as N^{-1}, for $N \to \infty$. This is the same result one would obtain for a self-averaging observable in the central limit theorem [22]. It shows that DMMs become increasingly insensitive to the microscopic details of the instances they solve (how the different variables are distributed in the various clauses). Note also that the result obtained by this physical analogy is valid *irrespective* of the problem the DMMs solve. It is only related to their general dynamical behavior.

In Fig. 3, we show numerical results corroborating the previous analysis, where we solve Eqs. (2) and (3) (the latter ones with added noise of strength $\Gamma = 0.12$) for various numbers of variables, N, and at a fixed clause-to-variable ratio. The relative variance is well fitted by the curve $\sigma^2(N)/\mu^2(N) \sim N^{-1.10\pm0.07}$, showing also that in this case, the numerical noise does not much affect the stochastic dynamics.

If we had only numerical noise, or the latter strongly couples to the additive noise, the assumption that the diffusion coefficient ν is independent of N would no longer be valid because of the accumulation of errors. In this case, we expect $\nu(N)$ to increase with N, with the form of this function dependent on the numerical method that is employed. This is shown in Figs. S7 and S8 of the SM [15], where we still find the TTS to be self-averaging but weakly $[\sigma^2(N)/\mu^2(N) \sim N^{0.39\pm0.02}$ for a forward Euler and $N^{0.29\pm0.02}$ for Runge-Kutta fourth order].
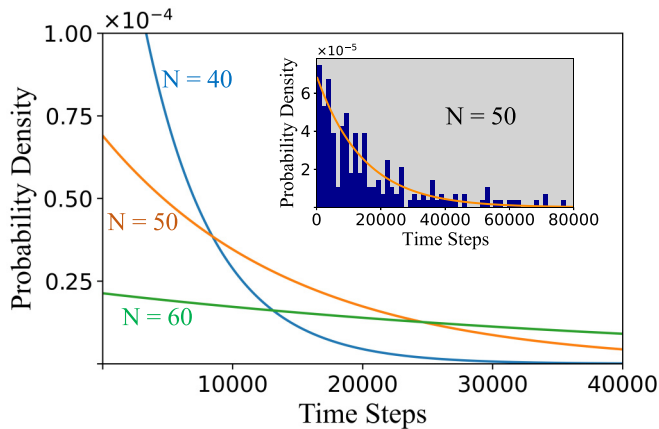
FIG. 4. The distributions of TTS for the local search algorithm WalkSAT plotted for $N = 40, 50, 60$ variables (over 200 instances) at a fixed clause-to-variable ratio of 7. The TTS is well fitted by an exponential probability distribution for all cases considered. Inset: an example for $N = 50$ variables of the histogram of TTS used to obtain the fits in the main panel.

C. Non-self-averaging of traditional algorithms

We finally conclude by showing that the self-averaging property of the TTS of DMMs is not shared by traditional algorithms applied to the same problems. We show this explicitly in Fig. 4 for a widely used (and representative) local search algorithm (WalkSAT) [10] applied to the same instances used previously [9]. (In the SM [15], we show additional numerical results using two different conflict-driven

clause learning algorithms, i.e., MiniSAT [11] and Kissat [12], and we find, as expected, that they are also not self-averaging.)

This algorithm simply flips one or a few variables at each iteration according to some prescribed rule. Unlike a DMM, the probability distribution of the TTS of this local search algorithm follows well an *exponential* fit of the type $P_\lambda(t) = \lambda e^{-\lambda t}$. (See the SM [15] for the analytical justification of why such a distribution has to be an exponential.) Here, λ is the inverse of the average ($\mu = 1/\lambda$) and the variance of this distribution is $\sigma^2 = 1/\lambda^2$. We then see that the relative variance is $\sigma^2/\mu^2 = 1$, *irrespective* of the size of the problem. As anticipated, this traditional algorithm is *not* self-averaging.

IV. CONCLUSIONS

We have shown that the TTS of DMMs follows an inverse Gaussian distribution with a relative variance that goes to zero as the size of the problem increases. This is true both in the presence of physical noise (showcasing strong self-averaging) as well as numerical noise (weak self-averaging). This means that DMMs are increasingly insensitive to the microscopic details of the instances they solve with increasing problem size. This property is not shared by traditional algorithms, illustrating a substantial advantage of DMMs in the solution of combinatorial optimization problems.

[1] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W. H. Freeman, New York, 1990).

[2] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992).

[3] M. Di Ventra and Y. V. Pershin, The parallel approach, Nat. Phys. **9**, 200 (2013).

[4] F. L. Traversa and M. Di Ventra, Universal MemComputing machines, IEEE Trans. Neural Netw. Learn. Syst. **26**, 2702 (2015).

[5] M. Di Ventra, *MemComputing: Fundamentals and Applications* (Oxford University Press, Oxford, 2022).

[6] F. L. Traversa and M. Di Ventra, Polynomial-time solution of prime factorization and NP-complete problems with digital MemComputing machines, Chaos: Interdiscipl. J. Nonlinear Sci. **27**, 023107 (2017).

[7] M. Di Ventra and F. L. Traversa, Absence of chaos in digital memcomputing machines with solutions, Phys. Lett. A **381**, 3255 (2017).

[8] G. A. Whitmore and V. Seshadri, A heuristic derivation of the inverse Gaussian distribution, Am. Stat. **41**, 280 (1987).

[9] W. Barthel, A. K. Hartmann, M. Leone, F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Hiding Solutions in Random Satisfiability Problems: A Statistical Mechanics Approach, Phys. Rev. Lett. **88**, 188701 (2002).

[10] S. Arora and B. Boaz, *Computational Complexity: A Modern Approach* (Cambridge University Press, Cambridge, 2009).

[11] N. Sorensson and N. Een, MINISAT v1. 13-a SAT solver with conflict-clause minimization, SAT **2005**, 1 (2005).

[12] A. B. K. F. M. Fleury and M. Heisinger, Cadical, kissat, paracooba, plingeling and treengeling entering the SAT competition 2020, SAT Competition **50**, 2020 (2020).

[13] S. R. B. Bearden, Y. R. Pei, and M. Di Ventra, Efficient solution of Boolean satisfiability problems with digital MemComputing, Sci. Rep. **10**, 19741 (2020).

[14] S. R. B. Bearden, H. Manukian, F. L. Traversa, and M. Di Ventra, Instantons in Self-Organizing Logic Gates, Phys. Rev. Appl. **9**, 034029 (2018).

[15] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.108.034306 for an analytical understanding of the exponential distribution of local search algorithms [24], the distributions using traditional conflict-driven clause learning algorithms [25,26], and more numerical results under different clause-to-variable ratios, different noise model, and different initial conditions.

[16] M. Di Ventra, F. L. Traversa, and I. V. Ovchinnikov, Topological field theory and computing with instantons, Ann. Phys. (Berlin) **529**, 1700123 (2017).

[17] M. Di Ventra and I. V. Ovchinnikov, Digital MemComputing: From logic to dynamics to topology, Ann. Phys. **409**, 167935 (2019).

[18] T. Sauer, *Numerical Analysis* (Pearson, New York, 2018).

[19] Following [14], we can relate the noise strength Γ (in s^{-1}) to an effective external temperature (in K). The value considered

in the present work, $\Gamma = 0.12\,\text{s}^{-1}$, would then correspond to a temperature comparable to room temperature, $\sim$300 K.

[20] Note that the TTS distribution (5) also arises for a given instance (namely, at *fixed* number of variables and clauses) over an *ensemble* of initial conditions. See Fig. S9 in the SM [15].

[21] H. Coxeter, *Regular Polytopes* (Dover, Mineola, NY, 1973).

[22] A. Aharony and A. B. Harris, Absence of Self-Averaging and Universal Fluctuations in Random Systems near Critical Points, Phys. Rev. Lett. **77**, 3700 (1996).

[23] https://github.com/yuanhangzhang98/MemComputing_3SAT_demo.

[24] H. Hoos, A mixture-model for the behaviour of SLS algorithms for SAT, AAAI-02 Proceedings (unpublished).

[25] J. P. Marques Silva and K. A. Sakallah, Grasp-A new search algorithm for satisfiability, in *The Best of ICCAD*, edited by A. Kuehlmann (Springer, New York, 2003), pp. 73–89.

[26] D. H. Wolpert and W. G. Macready, No free lunch theorems for optimization, IEEE Trans. Evol. Comput. **1**, 67 (1997).