AFRL-RI-RS-TR-2020-017

SOLVING VERIFICATION & VALIDATION PROBLEMS WITH QUANTUM ANNEALING SAMPLERS

UNIVERSITY OF SOUTHERN CALIFORNIA

FEBRUARY 2020

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RI-RS-TR-2020-017 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /
KRISTI MEZZANO
Work Unit Manager

/ S /
LAUREN HUIE-SEVERSKY
Technical Advisor, Computing
and Communications Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
FEBRUARY 2020	FINAL TECHNICAL REPORT	MAY 2018 – SEP 2019

4. TITLE AND SUBTITLE

SOLVING VERIFICATION & VALIDATION PROBLEMS WITH QUANTUM ANNEALING SAMPLERS

5a. CONTRACT NUMBER
NA

5b. GRANT NUMBER
FA8750-18-1-0109

5c. PROGRAM ELEMENT NUMBER
62788F

6. AUTHOR(S)

Itay Hen

5d. PROJECT NUMBER
CYDT

5e. TASK NUMBER
US

5f. WORK UNIT NUMBER
CA

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Information Sciences Institute, University of Southern California
4676 Admiralty Way
Marina del Rey CA 90292

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RITQ
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER
AFRL-RI-RS-TR-2020-017

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this effort, we studied the extent to which quantum annealer computing systems can provide a powerful platform for achieving quantum enhancements in tasks of counting or listing the solutions of combinatorial optimization problems that implement verification and validation (V&V) problems. In order to generate meaningful benchmarks that can be run on experimental quantum annealers, we devised an algorithm for verifiability and certifiably test the density of states of hard optimization problems based on population annealing, which then allowed us to enumerate beforehand the degeneracies of energy levels of V&V instances. We then tested the performance of the D-Wave 2000-qubit processor on the generated instances. Our observations show that quantum annealers do indeed sample the ground state manifolds of V&V problems rather uniformly however for these devices to be effective fair groundstate samplers, their temperature and error rates must be substantially lowered. We hope that the results presented here will provide insight into the computational power of quantum devices and be useful for solving V &V problems going forward.

15. SUBJECT TERMS

Quantum Annealing, Quantum Annealing Computing (QuAnCo), Verification and Validation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			**KRISTI MEZZANO**
U	U	U	UU	25	19b. TELEPHONE NUMBER (*Include area code*) **N/A**

Table of Contents

Table of Figures

1.0 SUMMARY

In this effort, we studied the extent to which quantum annealer computing systems can provide a powerful platform for achieving quantum enhancements in tasks of counting or listing the solutions of combinatorial optimization problems that implement verification and validation (V&V) problems. In order to generate meaningful benchmarks that can be run on experimental quantum annealers, we devised an algorithm for verifiably and certifiably test the density of states of hard optimization problems based on population annealing, which then allowed us to enumerate beforehand the degeneracies of energy levels of V&V instances. We then tested the performance of the D-Wave 2000-qubit processor on the generated instances. Our observations show that quantum annealers do indeed sample the ground state manifolds of V&V problems rather uniformly however for these devices to be effective fair ground-state samplers, their temperature and error rates must be substantially lowered. We hope that the results presented here will provide insight into the computational power of quantum devices and be useful for solving V &V problems going forward.

2.0 INTRODUCTION

Many problems of relevance to DoD missions and the US Air Force in particular may be cast as a task for finding *all* the minimizing configurations, or ground states, of a given cost function. Examples are numerous — among them are SAT filtering [1], hardware fault detection and the V&V of safety-critical cyber-physical systems [2]. In the V&V of safety-critical cyber-physical systems for instance, one is concerned with testing whether a given piece of software, usually installed on an aircraft, contains a bug. This problem can naturally be cast as a constraint satisfaction problem [2] (equivalently, as an optimization problem of the Ising-type) where finding as many of the bugs as possible is critical to the success of the mission as every undetected bug could in principle result in the loss of the aircraft or, in more extreme cases, even loss of human life. Similar scenarios occur in circuit fault detection where each solution corresponds to a potential discrepancy in the implementation of a circuit and where all discrepancies must be found. Here, circuit faults could be the result of design errors but could also be the action of an adversary.

The listing of all solutions of a given cost function is, however, generally an intractable task for standard algorithms. This is not only because of the difficulty involved in finding an optimum [3], but also because of the sheer number of solutions, which may grow exponentially with input size [4]. Furthermore, the energy landscapes of certain cost functions are known to bias heuristic optimizers, as well as provable solvers, towards certain solutions and away from others [5]. Thus, the practical importance of sampling 'fairly' from the set of ground states of intricate cost functions is immense—both from the theoretical point of view and for practical reasons. In the context of V&V, one hopes that employing a suite of qualitatively dissimilar sampling algorithms will unearth nonidentical or even disjoint sets of solutions, leading eventually to the discovery of much larger sets of bugs.

Recent technological breakthroughs that have made experimental programmable quantum annealing computing (QuAnCo) optimizers containing thousands of quantum bits [6, 7] available, have rekindled the interest in quantum annealers as a revolutionary new approach to finding the minimizing assignments of discrete combinatorial cost functions [8, 9]. In this approach, gradually decreasing quantum fluctuations are used to traverse barriers in the energy landscape in search of global optima, a mechanism commonly believed to have no classical counterpart.

QuAnCo devices offer the exciting possibility of discovering minimizing assignments that *cannot be reached* in practice with standard algorithms. Similar to standard classical algorithms, QuAnCo systems — when tasked with solving optimization problems, will generally sample the solution space of optimization problems in a biased manner, producing certain ground states more frequently than others. Unlike the bias exhibited by thermal algorithms, the uneven sampling of quantum annealers has its origins in the *quantum nature* of their dynamics [10]. Since the distribution of minimizing configurations generated

by a quantum annealer is *intrinsically quantum*, a possibility arises that some quantum distributions cannot be efficiently generated by classical samplers. Moreover, that the choice of driver Hamiltonian determines these distributions, offers a tunable handle, or an extra knob, that potentially produces a continuum of probability distributions over the ground-state configurations. This extra 'quantum knob' may provide unique advantages in the sense that certain classically suppressed configurations, i.e., solutions that have very low probabilities of being found via classical processes, may have high probabilities of being found or sampled with suitable choices of driver Hamiltonians [11–14]. As was found by a recent study (co-authored by the PI) [10] there is an immense inherent potential for quantum annealers to complement standard solvers in thoroughly sampling the solution sets of random Ising spin glasses, giving rise to novel forms of quantum enhancements.

To test whether quantum annealers indeed provide a potentially powerful platform for achieving quantum enhancements for the counting or listing of solutions of V&V problems, we have studied in detail the potential of QuAnCo algorithms to sample ground-state configurations differently than classical algorithms. We considered prototypes of V&V problem instances in the form of Ising spin glasses, whose cost function is given as:

$$H_p = \sum_{\langle ij \rangle} J_{ij} s_i s_j + \sum_i h_i s_i \,, \tag{1}$$

where the Ising spins, $s_i = \pm 1$, are the variables to be optimized over, and the set of parameters $\{J_{ij}, h_i\}$ determines the cost function.

This report summarizes our effort, which was focused on testing the capabilities of experimental QuAnCo systems to fairly sample the ground state manifolds of problems from the above class. The first part of the effort was dedicated to generating hard test problems for which the identity and number of ground and excited states can be known in advance. The second part of the project was dedicated to testing the extent to which experimental quantum annealers [6, 15] can uniformly sample the ground state manifolds of the generated problems.

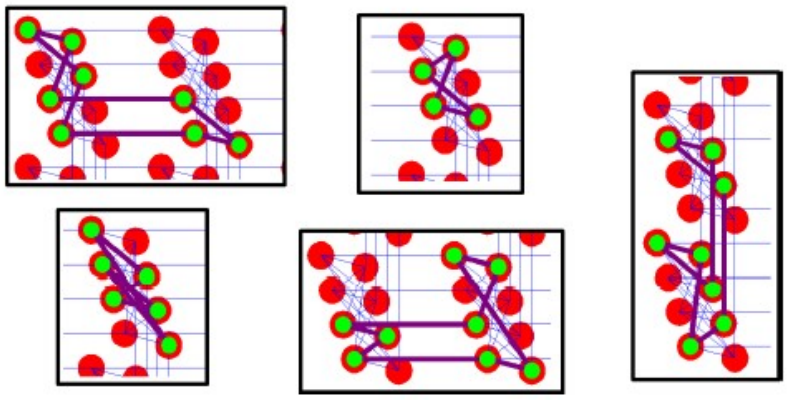

Figure 1 - A 'Planted Solution' Problem Instance

3.0 METHODS, ASSUMPTIONS AND PROCEDURES

3.1 Generating testable V&V instances

The precise enumeration of all bugs for a given V&V problem is tantamount to finding all the minimizing configurations of the instance after it has been cast in Ising form. For instances with more than a few dozen spins to do so exactly is an intractable task [16, 17]. In fact, the enumeration of all minimizing configurations of a given optimization problem is a difficult task in the general case, belonging to the complexity class #*P* as the exhaustive search for all solutions of an *n*-spin Ising problem becomes unfeasible for $n > 40$ bit problems as the search space grows exponentially with the size of the problem.

To successfully address this difficulty, for the generation of instances, we have chosen in this project to study problems constructed around "planted solutions" – an idea borrowed from constraint satisfaction (SAT) problems [18, 19]. See Fig. 1 for an illustration. In these problems, the planted solution represents a ground state configuration of Eq. (1) that minimizes the energy and is known in advance. This knowledge circumvents the need to verify the ground state energy using exact (provable) solvers, which rapidly become too expensive computationally as the number of variables grows, and which were employed in earlier benchmarking studies [20, 21]. Moreover, these problems are known to possess different degrees of "tunable hardness", achieved by adjusting the amount of frustration (see Ref. [22]). Last, studying this type of problems allowed us to devise an algorithm to find *all* minimizing configurations of the generated instances. While our approach in principle works for arbitrary graphs, we focused here on Chimera lattices, i.e., two-dimensional arrays of unit

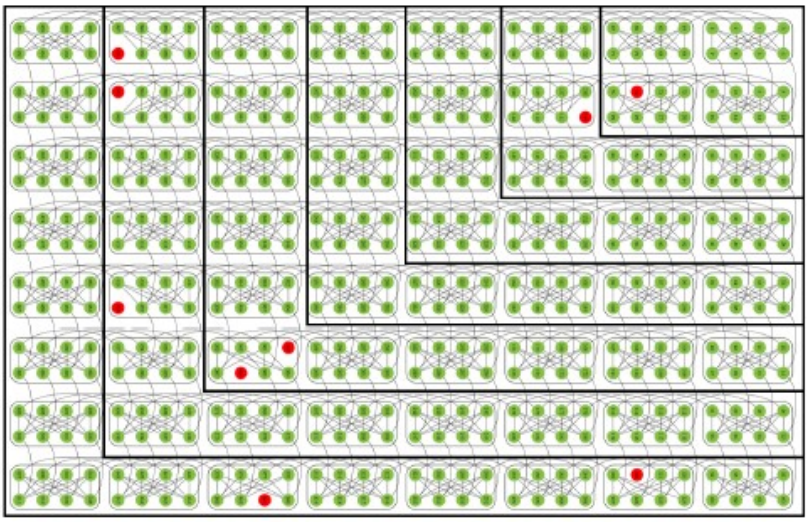

Figure 2 - The D-Wave quantum annealer chimera architecture

cells of eight spins with a $K_{4,4}$ bipartite connectivity [23, 23]. This choice further allowed us to test the capabilities of D-Wave experimental quantum annealers which feature qubits connected in a Chimera graph [7, 15, 24]. While the Chimera graph is two-dimensional in nature [25], it is also non-planar and as such gives rise to difficult spin-glass problems [26]. The Chimera architecture is shown in Fig. 2.

The above properties make our generated instances optimally suited to function as stand-ins for genuine V&V problem instances. We generated multiple planted-solution instances of varying sizes, following a technique described in detail in Ref. [27] wherein the clauses H_j are chosen to be 'frustrated loops' along the Chimera graph. We further used the fact that our generated instances are frustration free, namely, consist of a sum of terms each of which has all minimizing configurations as their ground state. To enumerate all minimizing configurations, we implemented a form of the 'bucket' algorithm [28] designed to eliminate variables one at a time to perform an exhaustive search efficiently. A more detailed discussion of planted Ising problems can be found in Ref. [27].

3.2 Estimating the density of states of V&V problems

In classical computing, to uniformly sample the ground-state manifolds of optimization problems, or more generally to 'Boltzmann sample' a given problem, one usually employs via 'thermal sampling' algorithms the most notable of which is the parallel tempering (PT) scheme [29, 30] wherein multiple copies of the problem are equilibrated in parallel at different temperatures and spin configurations at adjacent temperatures are regularly swapped [31].

In PT, each copy performs a temperature random-walk. At high temperatures, free-energy barriers are easily overcome, allowing for a global exploration of configuration space. At lower temperatures on the other hand, the local minima are explored in more detail. PT was run on the instance until equilibration is reached, at which point all minimizing configurations appear with equal probabilities and can be recorded and stored accordingly. PT equilibration was tested using the concept of 'mixing time' [32–35] using it as the figure of merit for thermalization. The mixing time of a PT simulation may be thought of as the average time it takes a PT replica to fully traverse the temperature mesh, indicating equilibration of the simulation. A 'healthy' PT simulation requires an unimpeded temperature flow and so the total length of the simulation should be longer than the temperature mixing time [32, 33].

On the other hand, the main mechanisms that determine the distributions from which quantum annealing output configurations are drawn are thus far unclear. Further insights into the role of temperature, and the ability of experimental quantum annealing optimizers to quickly thermalize, are challenging to obtain due to the limited ability to probe the inner workings of these machines, as well as the lack of control over most operating parameters [16, 17, 36]. The question of whether such devices indeed provide superior performance as compared to classical algorithms for some problem classes has been the subject of much recent debate [14, 16, 17, 27, 36–41].

Furthermore, it is unclear whether the open-system dynamics of a system naturally has the thermal state as its steady state, and if so, whether the dissipative dynamics can be used in lieu of a true quantum algorithm to prepare such a state. Having the steady state of the dissipative dynamics be the Boltzmann distribution of the problem Hamiltonian is a non-trivial assumption; it is known to be the case of Markovian weak-coupling limit master equations satisfying the Kubo-Martin-Schwinger (KMS) condition. An open-system adiabatic theorem provides a guarantee that in the long-time limit the state of the system will be close to the desired thermal state.

Here, we tested the extent to which currently available commercial realizations of quantum annealing do equilibrate to the Boltzmann distributions of the problems they are to solve. The results of these tests provide insight into the whether these devices sample uniformly the low-lying energy configurations of V&V problems coded into the couplers connecting an array of superconducting flux qubits.

The first hurdle we faced when testing the performance of experimental quantum annealers in this context is the need to estimate the density of states of systems with rugged free energy landscapes, which is a notoriously difficult task of the utmost importance in many areas of physics ranging from spin glasses to biopolymers.

For a thermalized system (a system in equilibrium) with inverse-temperature β, the probability of obtaining a spin configuration $\mathbf{s}$ is

$$P(\mathbf{s}) = \frac{e^{-\beta E(\mathbf{s})}}{\sum_{\mathbf{s}'} e^{-\beta E(\mathbf{s}')}} \,. \tag{2}$$

It follows then that in equilibrium i) the probability of obtaining a ground-state configuration is always higher than the probability of obtaining an excited state. ii) the probability of obtaining the various ground state is the same for all ground states, which in turn means that in equilibrium experimental quantum annealers are fair samplers. Moreover, for a model with discrete energy values, one may also ask about the probability of obtaining a given energy values. This is given by

$$P(E) = \frac{\Omega(E) e^{-\beta E}}{\sum_{E'} \Omega(E') e^{-\beta E'}} \,. \tag{3}$$

Here, $\Omega(E)$ is the degeneracy of the energy value E, or the number of configurations $\mathbf{s}$ that have energy E.

Some of the standard approaches taken so far to estimate the density of states, such as the widely known Wang-Landau algorithm [42], suffer from a spurious convergence of the estimates to metastable minima, and these cases are particularly hard to detect. Such difficulties apply in particular to systems with complex energy landscapes that are typically accompanied by frustration in the interactions such as in the V&V and in the spin-glass systems that result from a combination of frustration and quenched disorder [43]. The latter may be viewed as prototypical classically-hard optimization problems, and they are so challenging that specialized hardware has been built to simulate them [44]. Here, we introduce a sampling technique based on population annealing enhanced with a multi-histogram analysis and report on its performance for spin glasses [45]. We demonstrate its ability to overcome the pitfalls of other entropic samplers, resulting in some cases in orders of magnitude scaling advantages that can result in the uncovering of new physics.

The new algorithm introduced here, which we call entropic population annealing (EPA), is not based on Markov chains but on the sequential Monte Carlo method. Population annealing (PA) was first studied in Refs. [46, 47] and more recently developed further in Refs. [48–54]. It is based on the initialization of a population of replicas drawn from the equilibrium distribution at high temperatures, which is then subsequently cooled to lower and lower temperatures. During this process, a combination of population control and spin flips is used to ensure that the ensemble remains in equilibrium. The simulation entails the following steps [48, 51]:

1. Set up an equilibrium ensemble of R0 = R independent copies (replicas) of the system at

inverse temperature β0 = 1/kBT0.

2. Take a step to inverse temperature $\beta_i > \beta_{i-1}$ by resampling the configurations $j = 1, \ldots, R_{i-1}$ with their relative Boltzmann weight $\hat{\tau}_i(E_j)$, leading to $R_i \neq R_{i-1}$ replicas in general.

3. Update each replica by θ rounds of an MCMC algorithm at inverse temperature β_i.

4. Go to step 2 unless the inverse target temperature β_f has been reached.

During resampling, the expected number of copies is

$$\hat{\tau}_i(E_j) = \frac{R}{R_{i-1}} \frac{e^{-(\beta_i - \beta_{i-1})E_j}}{Q(\beta_{i-1}, \beta_i)}, \tag{4}$$

with a normalizing factor

$$Q(\beta_{i-1}, \beta_i) = \frac{1}{R_{i-1}} \sum_{j=1}^{R_{i-1}} e^{-(\beta_i - \beta_{i-1})E_j}. \tag{5}$$

The actual number of copies is taken to be the integer part $\lfloor \hat{\tau}_i(E_j) \rfloor$ plus an additional copy added with a probability corresponding to the fractional part, $\hat{\tau}_i(E_j) - \lfloor \hat{\tau}_i(E_j) \rfloor$. While initially, constant (inverse) temperature steps were used on increasing $\beta_i > \beta_{i-1}$ [48], it turns out that a better, parameter-free method consists of choosing β_i to ensure a certain overlap of the energy distributions between the two temperatures [51]. This overlap can be computed from the resampling factors,

$$\alpha(\beta_{i-1}, \beta_i) = \frac{1}{R_{i-1}} \sum_{j=1}^{R_{i-1}} \min\left(1, \frac{R \exp[-(\beta_i - \beta_{i-1})E_j]}{R_{i-1} Q(\beta_{i-1}, \beta_i)}\right) \tag{6}$$

and β_i is adapted using a bisection search such as to ensure an overlap α^* of energy histograms. The method is not very sensitive to the precise value of α^*, and we choose $\alpha^* = 0.86$ in the runs below.

While the algorithm described above is just population annealing [48] improved by adaptive temperature steps [51, 53], the possibility of sampling the entropy arises from a combination of the method with multi-histogram techniques [55]. An estimator of the free energy follows directly from the resampling factors,

$$-\beta_i \hat{F}(\beta_i) = \ln Z_{\beta_0} + \sum_{k=1}^{i} \ln Q_k, \tag{7}$$

where Z_{β_0} is the partition function at the initial temperature β_0. In the following, we always choose $\beta_0 = 0$, such that simply $Z_{\beta_0} = 2^N$, where N is the number of spins. We can then estimate the density of states by combining the histograms at all temperature steps, and a variance-optimized estimator is given by [51, 55].

$$\hat{\Omega}(E) = \frac{\sum_{i=1}^{N_\beta} \hat{H}_{\beta_i}(E)}{\sum_{i=1}^{N_\beta} R_i \exp[\beta_i \hat{F}(\beta_i) - \beta_i E]}. \tag{8}$$

Here, N_β is the total number of temperatures, and the energy histogram $\hat{H}_{\beta_i}(E)$ at inverse temperature β_i is normalized such that $\sum_E \hat{H}_{\beta_i}(E) = R_i$.

The approach is naturally suited for (moderately or massively) parallel calculations as the R replicas are simulated independently of each other and the only interaction occurs during resampling. An efficient GPU implementation was discussed in Ref. [51]. Importantly for our application, EPA does not require any prior knowledge of the range of realized energies. A detailed analysis of systematic and statistical errors of PA can be found in Ref. [56].

One of the advantages of the approach based on population annealing is that it does not require any prior knowledge about the energy spectrum, which in contrast needs to be acquired in an additional pre-run for the Wang-Landau method. The notorious problem of premature and false convergence that plagues the latter approach is not so much of an issue for the newly introduced technique, where a re-distribution of weights can occur at all stages of the algorithm. The main advantage of the approach, however, lies in the ideal suitability for massively parallel calculations, where given sufficient parallel resources the accuracy of the approximation can be arbitrarily improved at a constant wall-clock time by increasing the size of the population.

As we have found [45], EPA performs significantly better at estimating $\Omega(E)$ for hard spin-

glass samples than the Wang-Landau entropic sampler [42], which is the most widely used entropic sampler to date.

Our EPA algorithm as well as Wang-Landau (WL) intrinsically estimate entropy *differences*, i.e., ratios of degeneracies for neighboring energy values, and the absolute scale is only achieved through an additional normalization such as that given by Z_{β_0} in Eq. (7). It is therefore reasonable to study their performance in estimating

$$r_{10} = \frac{\Omega(E_1)}{\Omega(E_0)}. \tag{9}$$

the ratio of degeneracies of first excited and ground states. In Fig. 3 we show the relative deviations of the ratios r_{10} from the exact values known through the planting as estimated from WL and EPA. WL found the correct ground-state energy for 622 of the 625 samples. For some samples the relative deviations are so large that they exceed the scale of the plot of Fig. 3, some by many orders of magnitude. These samples are shown at the boundary of the box and in a different color. It is clear that for most samples the deviations are substantially smaller for EPA than for WL. In total, EPA outperforms WL in 89% of the instances. The

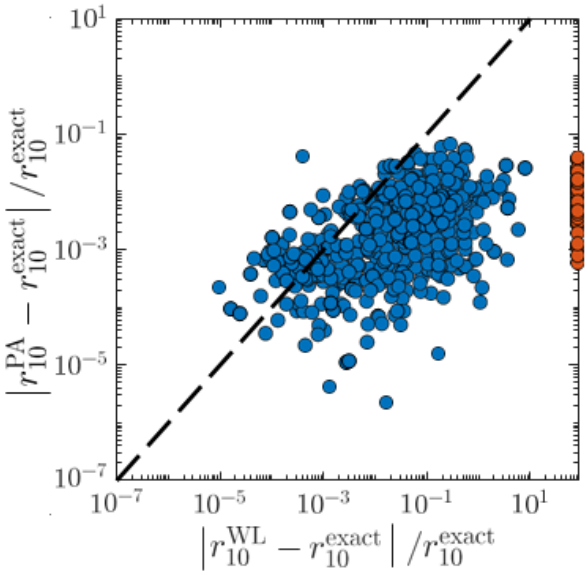

Figure 3 - Density of States Estimation Algorithm

error of WL is larger than 10% for 39% of samples and it is difficult to distinguish between the accurate and inaccurate WL results. In contrast, the EPA results are accurate to within 7% for all samples.

4.0 RESULTS AND DISCUSSION: THE D-WAVE 2000Q AS A V&V SAMPLER

A handful of recent studies suggest that quantum annealers too may be well suited to function as fast thermal samplers [16, 17, 39, 57]. By taking advantage of their finite temperature nature [16, 17, 34, 35], quantum annealing devices may be used as samplers from Boltzmann distributions of input cost functions. Such a capability opens up the exciting possibility of applications of quantum annealing to so-far-uncharted avenues of research, with immediate applications to domains such as deep learning networks and restricted Boltzmann machines [16, 36, 39].

To test this hypothesis, we devised an experiment to ascertain the performance of an experimental quantum annealer functioning as a Boltzmann sampler, using the sets of V&V-type problem instances discussed above, for which we calculated the degeneracies $\Omega(E)$.

Using those instances, we tested the performance of the experimental D-Wave 2000Q processor (DW2000Q) to sample the ground state manifolds and additional low-lying states according to the expected Boltzmann distribution. The DW2000Q quantum annealing optimizer consists of an array of superconducting flux qubits and is designed to solve Ising model instances defined on the graph hardware via a gradually decreasing transverse field Hamiltonian. This device is designed to solve optimization problems by evolving a known initial configuration—the ground-state of a transverse field towards the ground-state of the classical Ising-model Hamiltonian of Eq. (1). The device's Hamiltonian is given by

$$H(s) = A(s)H_d + B(s)H_p\,, \tag{10}$$

where $H_p = \sum_{\langle i,j\rangle} J_{ij}\sigma_i^z\sigma_j^z + \sum_i h_i\sigma_i^z$ is the programmable Ising spin-glass problem (the final Hamiltonian) to be sampled defined by the parameters $\{J_{ii}, h_i\}$, and $H_d = -\sum_i \sigma_i^x$ is a transverse-field Hamiltonian, which proved the quantum fluctuation (the initial Hamiltonian). Figure 4 shows the schedule for A(s) and B(s) as well as the device temperature (the hardware architecture of the device is given in Fig. 2). While the device implements only a stoquastic transverse-field Hamiltonian, it allows us access to the performance on problems of much larger sizes than can be computed numerically, namely, problems of up to > 2000 qubits.

We ran each V&V instance multiple (> 10^6) times on the experimental device with each anneal ending up with a measurement in the computational basis yielding a spin configuration with a given energy which was then stored. To overcome the inhomogeneity of the processor as well as other systematic errors, each anneal is carried out with a randomly generated gauge (see Ref.[20] for more details).

By calculating an estimator for the probability that the device finds a configuration with

energy E, namely $\tilde{P}(E)$, we estimate the effective temperature β^{eff}. To do so, we use the energy level degeneracies computed with our EPA algorithm. We estimate the inverse-

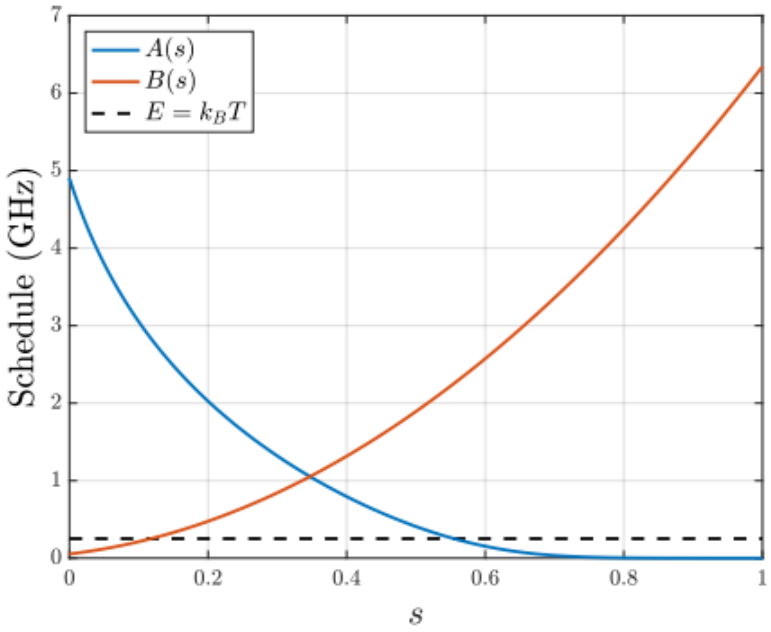

Figure 4 - The D-Wave 2000Q Schedule

temperature β^{eff} for each instance by minimizing the total distances between the exact $P(E)$ as given by Eq. (3) to the estimated probability.

Our results show that typically, the fits of the observed data points to the predictions are rather accurate, implying that the D-Wave quantum annealing sampler does indeed equilibrate and reach a thermal Boltzmann distribution. Two examples for such fits are given in Fig. 5. Our results further imply that the configurations of equal energy are sampled with approximately similar probabilities, or in other words, are sampled uniformly.

Nonetheless, by inspecting the values of effective inverse temperatures, we find that these are five to ten times higher than the real temperature device. This high effective temperature translates to a substantial decrease in success probability of the device. This is illustrated in Fig. 6.

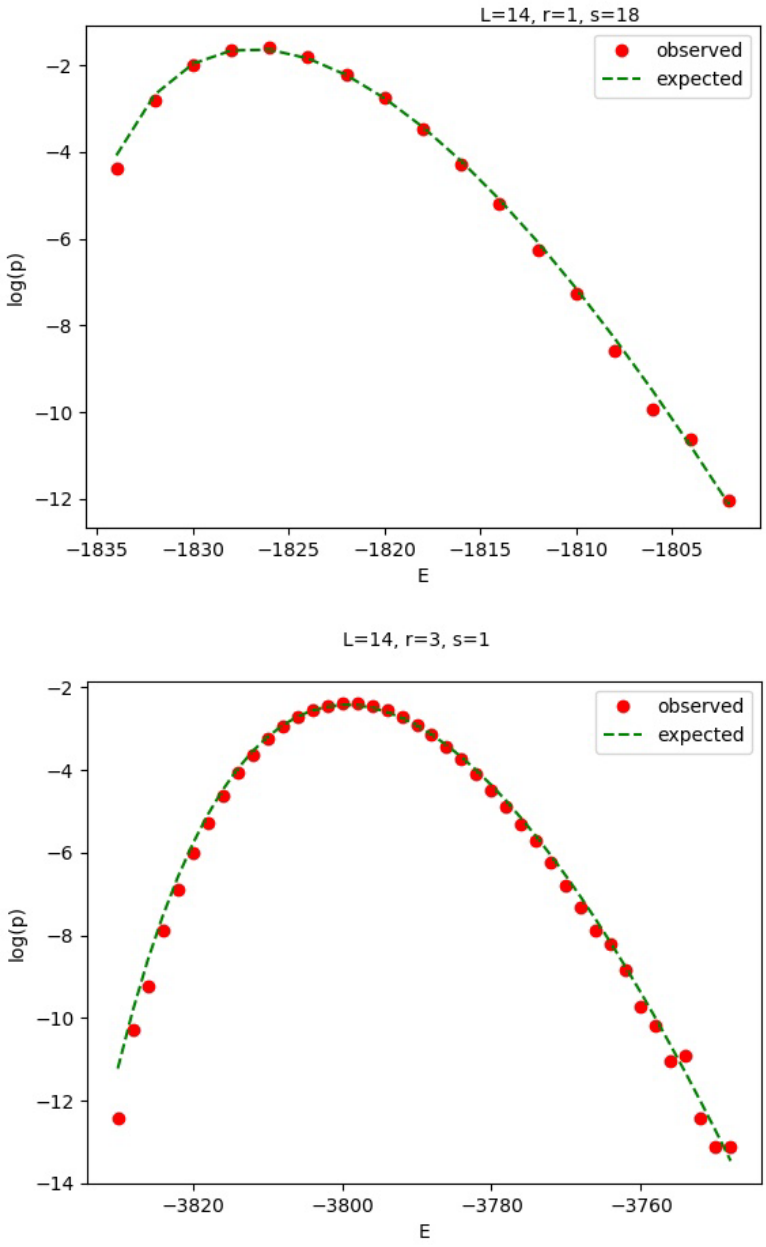

Figure 5 - Observed vs Expected Degeneracies

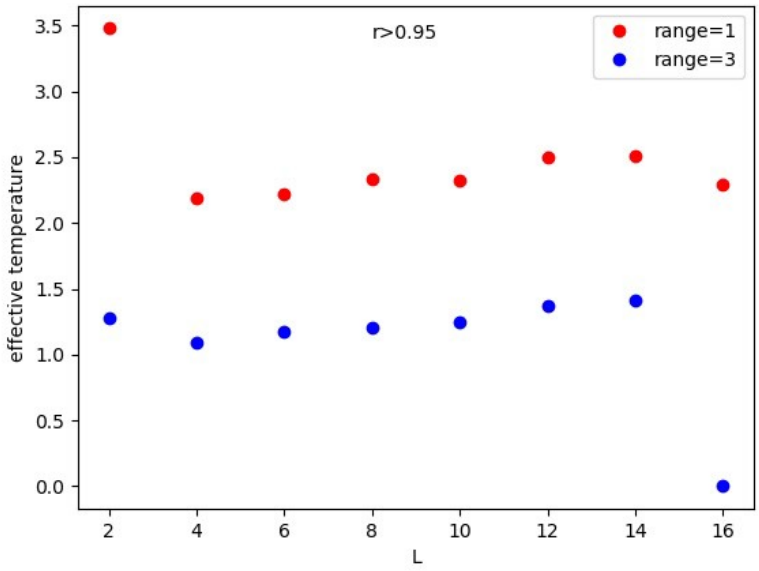

Figure 6 - Effective Quantum Annealer Temperature

5.0 CONCLUSIONS

In this effort, we studied the extent to which quantum annealer computing (QuAnCo) systems can provide a powerful platform for achieving quantum enhancements in tasks of counting or listing the solutions of verification and validation (V&V) problems. To that aim, we tested the potential of experimental quantum annealing systems, specifically the D-Wave 2000 qubit processor, to sample uniformly the ground-state manifolds of hard optimization problems which correspond to 'bugs' in V&V instances.

In order to generate meaningful benchmarks, we devised an algorithm for verifiably and certifiably test the density of states of hard optimization problems based on population annealing. We have investigated the performance of sampling methods for estimating the density of states for systems with complex free-energy landscapes, focusing on samples of spin glasses. We proposed a novel sampling technique based on sequential Monte Carlo on a large population of copies and demonstrated that it outperforms the most widely used entropic sampler, the Wang-Landau algorithm, in the vast majority of cases — in some situations providing improvements by several orders of magnitude thereby allowing for the reliable studying of spin-glass, and hence also V&V, degeneracies for larger problems than ever before.

A notorious problem with benchmarking algorithms for estimating the density of states lies in safely assessing convergence. Here we solved this issue by studying samples with planted solutions that can be run on experimental quantum annealers and for which the exact degeneracies of the ground and first excited states can be calculated using a 'bucket' algorithm. The main advantage of the approach, however, lies in the ideal suitability for massively parallel calculations, where given sufficient parallel resources the accuracy of the approximation can be arbitrarily improved at a constant wall-clock time by increasing the size of the population.

Tasking the D-Wave 2000Q experimental annealer with finding the ground states of these instances, we observed the following. The ground states of the tested instances were found to be indeed sampled uniformly in general, which further means that these quantum devices are potentially a useful tool for solving V&V problems which require finding all sets of solutions.

However, the effective temperatures were found to be up to five times higher than one would have expected. This high effective temperature translates to a substantial decrease in success probability of the device. Our observations show that for future quantum annealers to be effective fair ground-state samplers, the temperature and error rates of experimental quantum annealing devices must be substantially lowered.

Our results have been (or will be) published in peer-reviewed journals:

- L. Barash, J. Marshall, M. Weigel and I. Hen, “Estimating the Density of States of Frustrated Spin Systems”, New Journal of Physics 21 073065 (2019). arXiv:1808.04340. (Enclosed with this report)
- X. Wang, L. Barash and I. Hen, “Performance of Experimental Quantum Annealers Functioning as Boltzmann Samplers”, in preparation.
- Z. Gonzalez-Izquierdo, T. Albash and I. Hen, “Testing a quantum annealer as a quantum thermal sampler”, in preparation.

We hope that the results presented above will provide useful insight into the computational power of quantum devices and be useful for solving V &V problems going forward.

6.0 REFERENCES

1. A. Douglass, A. D. King, and J. Raymond, "Theory and applications of satisfiability testing - sat 2015: 18th international conference, austin, tx, usa, september 24-27, 2015, proceedings," (Springer International Publishing, Cham, 2015) Chap. Constructing SAT Filters with a Quantum Annealer, pp. 104-120.
2. K. L. Pudenz and D. A. Lidar, Quantum Information Processing 12, 2027 (2013).
3. C. Papadimitriou and K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Dover Books on Computer Science (Dover Publications, 2013).
4. L. Pauling, Journal of the American Chemical Society 57, 2680 (1935).
5. S. Bastea and P. M. Duxbury, Phys. Rev. E 58, 4261 (1998).
6. M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose, Nature 473, 194 (2011).
7. A. J. Berkley et al., Phys. Rev. B 87, 020502(R) (2013).
8. T. Kadowaki and H. Nishimori, Phys. Rev. E 58, 5355 (1998).
9. E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science 292, 472 (2001).
10. B. H. Zhang, G. Wagenbreth, V. Martin-Mayor, and I. Hen, Scientific Reports 7, 1044 (2017).
11. S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Nat. Commun. 4, 2067 (2013).
12. T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, Phys. Rev. A 91, 042314 (2015).
13. Y. Matsuda, H. Nishimori, and H. G. Katzgraber, Journal of Physics: Conference Series 143, 012003 (2009).
14. S. Mandra`, Z. Zhu, and H. G. Katzgraber, Phys. Rev. Lett. 118, 070502 (2017).
15. P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. Przybysz, and J. Whittaker, IEEE Transactions on Applied Superconductivity 24, 1 (Aug 2014).
16. M. Benedetti, J. Realpe-Go´mez, R. Biswas, and A. Perdomo-Ortiz, Phys. Rev. A 94, 022308 (2016).
17. B. H. Zhang, G. Wagenbreth, V. Martin-Mayor, and I. Hen, Scientific Reports 7, 1044 (2017).
18. W. Barthel, A. K. Hartmann, M. Leone, F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Phys. Rev. Lett. 88, 188701 (2002).
19. F. Krzakala and L. Zdeborov´a, Phys. Rev. Lett. 102, 238701 (2009).

20. S. Boixo, T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, Nat. Phys. 10, 218 (2014).

21. T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, Science 345, 420 (2014).

22. M. Mezard, G. Parisi and M.A. Virasoro, Spin Glass Theory and Beyond, World Scientific Lecture Notes in Physics (World Scientific, Singapore, 1987).

23. V. Choi, Quant. Inf. Proc. 7, 193 (2008); Quant. Inf. Proc. 10, 343 (2011).

24. M. W. Johnson et al., Nature 473, 194 (2011).

25. M. Weigel, H. G. Katzgraber, J. Machta, F. Hamze, and R. S. Andrist, Phys. Rev. X5, 019901 (2015).

26. F. Barahona, J. Phys. A 15, 3241 (1982).

27. I. Hen, J. Job, T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar,Phys. Rev. A 92, 042325 (2015).

28. R. Dechter, Artificial Intelligence 113, 41 (1999).

29. K. Hukushima and K. Nemoto, Journal of the Physical Society of Japan 65, 1604 (1996), http://dx.doi.org/10.1143/JPSJ.65.1604.

30. E. Marinari, in Advances in Computer Simulation, edited by J. Kert´esz and I. Kondor (Springer-Verlag, 1998) p. 50, (arXiv:cond-mat/9612010).

31. A. Sokal, in Functional Integration: Basics and Applications, edited by C. DeWitt-Morette, P. Cartier, and A. Folacci (Plenum, 1997).

32. L. A. Fernandez, V. Martin-Mayor, S. Perez-Gaviro, A. Tarancon, and A. P. Young, Phys. Rev. B 80, 024422 (2009).

33. R. Alvarez Baños, A. Cruz, L. A. Fernandez, J. M. Gil-Narvion, A. Gordillo-Guerrero, M. Guidetti, A. Maiorano, F. Mantovani, E. Marinari, V. Martin-Mayor, J. Monforte-Garcia, Muñoz Sudupe, D. Navarro, G. Parisi, S. Perez-Gaviro, J. J. Ruiz-Lorenzo, S. F. Schifano, Seoane, A. Tarancon, R. Tripiccione, and D. Yllanes (Janus Collaboration), 2010, P06026 (2010), arXiv:1003.2569.

34. V. Martin-Mayor and I. Hen, Scientific Reports 5, 15324 EP (2015).

35. J. Marshall, V. Martin-Mayor, and I. Hen, Phys. Rev. A 94, 012320 (2016).

36. S. H. Adachi and M. P. Henderson, ArXiv e-prints (2015), arXiv:1510.06356 [quant-ph].

37. H. G. Katzgraber, F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, Phys. Rev. X5, 031026 (2015).

38. S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, J. M. Martinis, and H. Neven, Nature Physics (2018).

39. M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, R. Melko, ArXiv e-prints (2016), arXiy:1601.02036 [quant-ph].

40. M. S. K¨onz, G. Mazzola, A. J. Ochoa, H. G. Katzgraber, and M. Troyer, ArXiv e-prints (2018), arXiv:1806.06081 [quant-ph].

41. J. Marshall, E. G. Rieffel, and I. Hen, Phys. Rev. Applied 8, 064025 (2017).

42. F. Wang and D. P. Landau, Phys. Rev. Lett. 86, 2050 (2001).

43. K. Binder and A. P. Young, Rev. Mod. Phys. 58, 801 (1986).

44. F. Belletti, M. Cotallo, A. Cruz, L. A. Fernández, A. G. Guerrero, M. Guidetti, A. Maiorano, F. Mantovani, E. Marinari, V. Martín-Mayor, A. Muñoz Sudupe, D. Navarro, G. Parisi, S. P. Gaviro, M. Rossi, J. J. Ruiz-Lorenzo, S. F. Schifano, D. Sciretti, A. Tarancón, and R. L. Tripiccione, Comput. Sci. Eng. 11, 48 (2009).

45. L. Barash, J. Marshall, M. Weigel, and I. Hen, New Journal of Physics 21, 073065 (2019).

46. Y. Iba, Trans. Jpn. Soc. Artif. Intell. 16, 279 (2001).

47. K. Hukushima and Y. Iba, AIP Conf. Proc. 690, 200 (2003).

48. J. Machta, Phys. Rev. E 82, 026704 (2010).

49. W. Wang, J. Machta, and H. G. Katzgraber, Phys. Rev. E 92, 063307 (2015).

50. M. Borovsky´, M. Weigel, L. Y. Barash, and M. Zukovic, EPJ Web Conf. 108, 02016 (2016).

51. L. Y. Barash, M. Weigel, M. Borovsky´, W. Janke, and L. N. Shchur, Comput. Phys. Commun. 220, 341 (2017).

52. L. Y. Barash, M. Weigel, L. N. Shchur, and W. Janke, Eur. Phys. J. Special Topics 226, 595 (2017).

53. C. Amey and J. Machta, Phys. Rev. E 97, 033301 (2018).

54. A. Barzegar, C. Pattison, W. Wang, and H. G. Katzgraber, arXiv preprint arXiv:1710.09025 (2017).

55. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. 63, 1195 (1989).

56. M. Weigel, L. Y. Barash, W. Janke, and L. N. Shchur, "Testing and improving the population annealing algorithm," In preparation.

57. M. Amin, arXiv:1503.04216 (2015).

List of Symbols, Abbreviations and Acronyms

DW2000Q	D-Wave 2000 qubit processor
EPA	Entropy Population Annealing
QuAnCo	quantum annealing computing
SAT	satisfiability
V&V	Verification and Validation
WL	Wang-Landau