Variational Neural Annealing

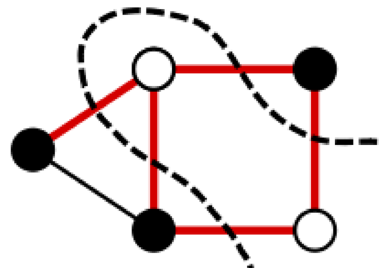
Mohamed Hibat-Allah, Estelle M. Inack, Roeland Wiersema, Roger G. Melko, Juan Carrasquilla

Vector Institute, Perimeter Institute, University of Waterloo

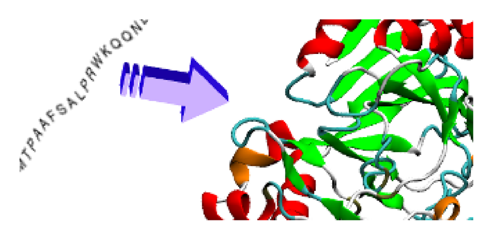
Motivation

- Optimization problems have a wide range of applications in many areas of science, as well as in real-world problems.
- Optimization problems can be formulated as the task of finding the lowest-energy state of an Ising Hamiltonian:

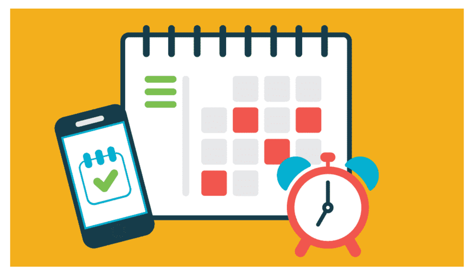
$$H_{\text{target}} = -\sum_{i<j} J_{ij}\sigma_i\sigma_j - \sum_{i=1}^{N} h_i\sigma_i,$$



Max-Cut Problem



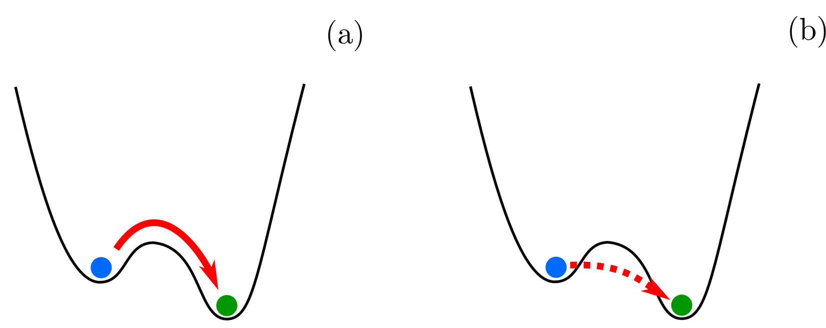
Protein Folding



Optimal Scheduling

Simulated Annealing and Quantum Annealing

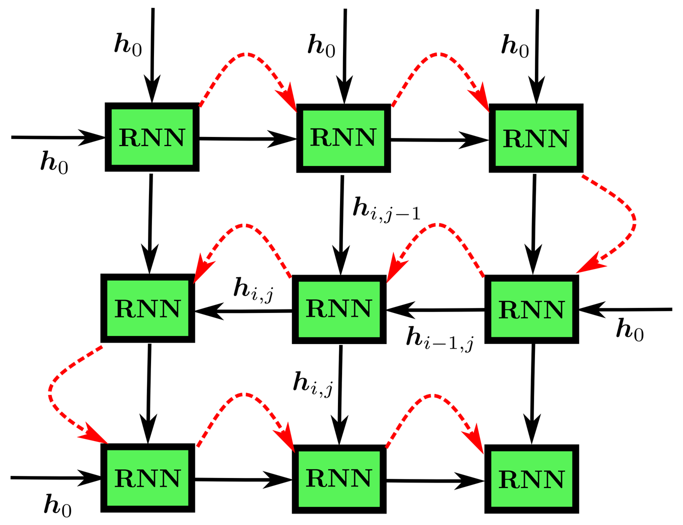
- **Simulated annealing** (SA) is a heuristic algorithm based on **thermal jumps** to find optimal minima of H_{target}.
- **Quantum annealing** (QA) is another heuristic inspired from quantum mechanics. It introduces **quantum tunneling** to overcome local minima of H_{target}.



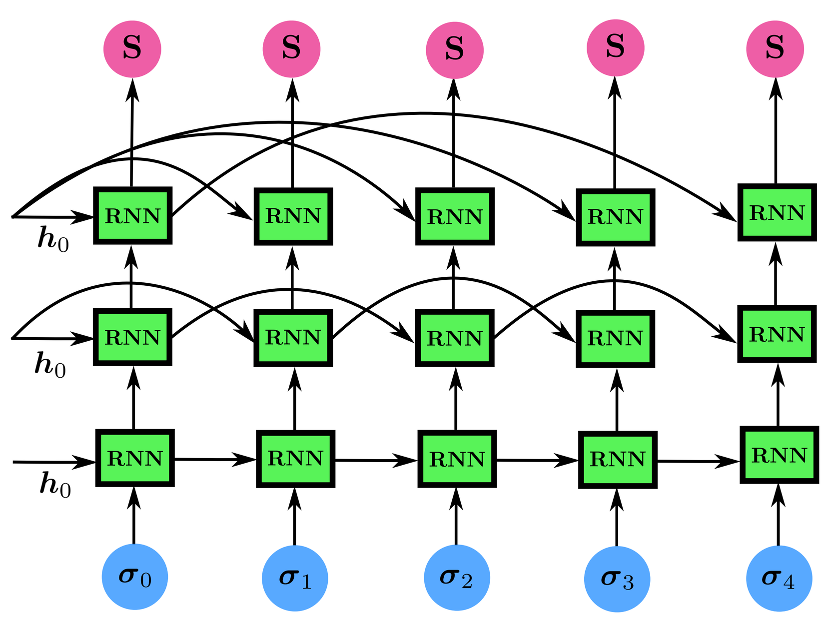
(a) Classical annealing, (b) Quantum annealing.

Recurrent Neural Networks (RNNs)

- The inputs of the RNN are spins. The latter are fed sequencially, and can be also generated autoregressively.
- We use **two-dimensional RNNs** to model the equilibrium states of two-dimensional spin glasses.



- We use **Dilated RNNs** [1] to model the equilibrium states of fully-connected spin glasses.
- In this study, we use RNNs to emulate simulated annealing and quantum annealing, for the purpose of finding better solutions to optimization problems.



Variational Classical Annealing (VCA)

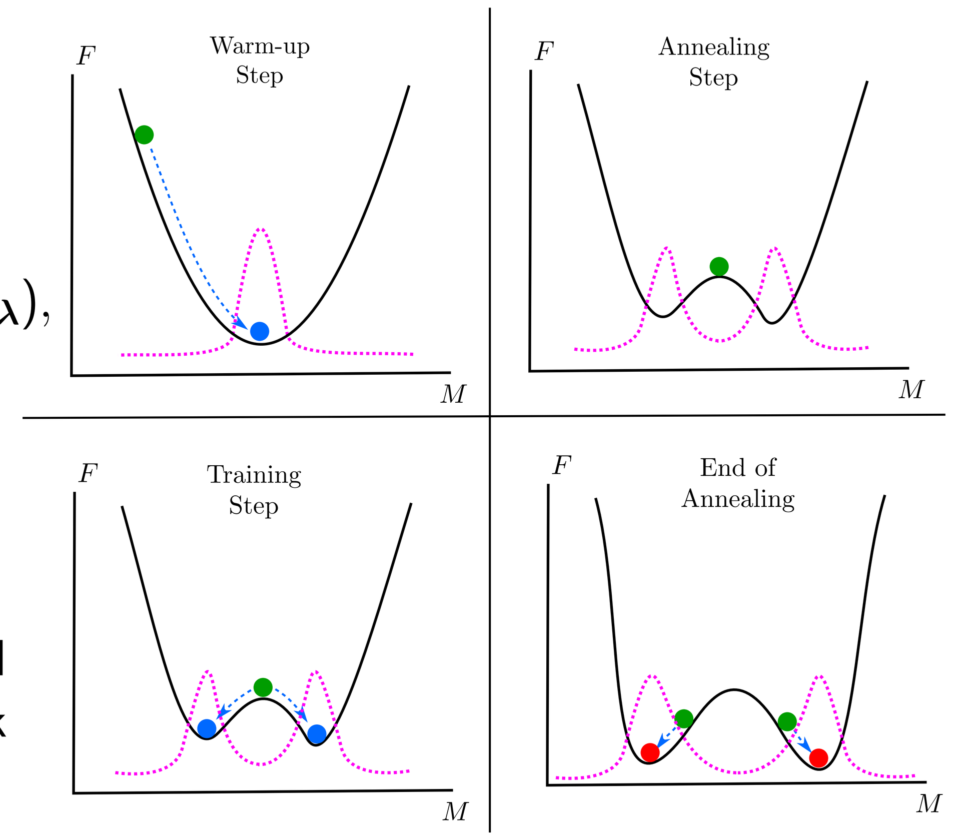
- **We use the variational free energy as a cost function:**

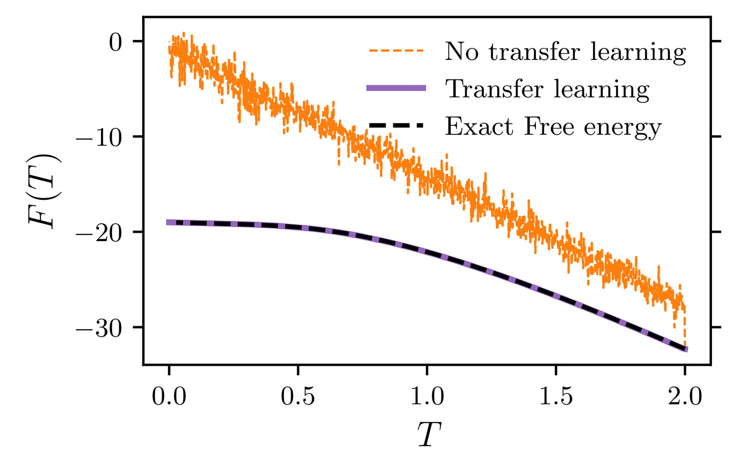
$$F_\lambda(t) = \langle \hat{H}_{\text{target}}\rangle_\lambda - T(t)S(p_\lambda),$$

- We gradually decrease the temperature using a linear schedule:

$$T(t) = T_0(1-t)$$

- The probability p_λ is modeled by a recurrent neural network (RNN).
- At $T=0$, we expect the RNN to output the ground state(s) of $\hat{H}_{\text{target}}$.
- **Proof of principle of adiabaticity** for the uniform Ising chain with $N=20$ spins. $\hat{H}_{\text{target}} = -\sum_{i=1}^{N-1}\hat{\sigma}_i^z\hat{\sigma}_{i+1}^z$.





Variational Quantum Annealing (VQA)

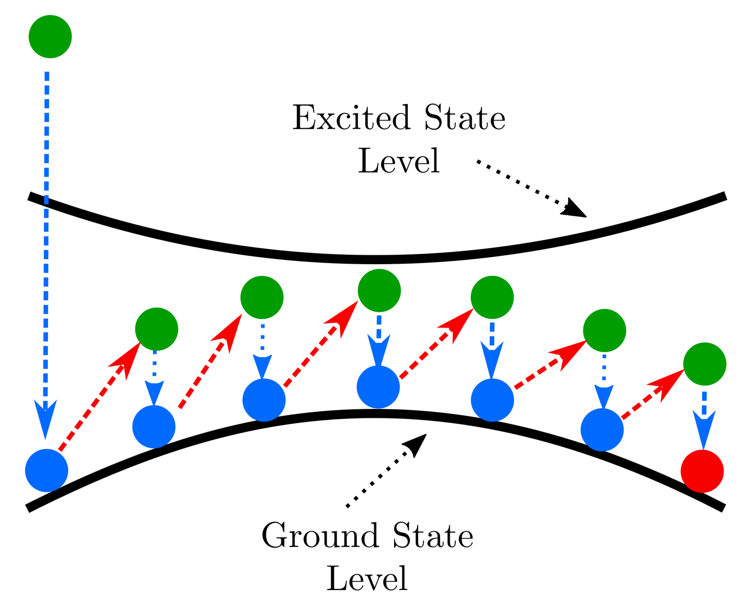
- **We use the variational energy as a cost function:**

$$\langle \hat{H}(t)\rangle_\lambda = \langle \hat{H}_{\text{target}}\rangle_\lambda + f(t)\langle \hat{H}_D\rangle_\lambda,$$

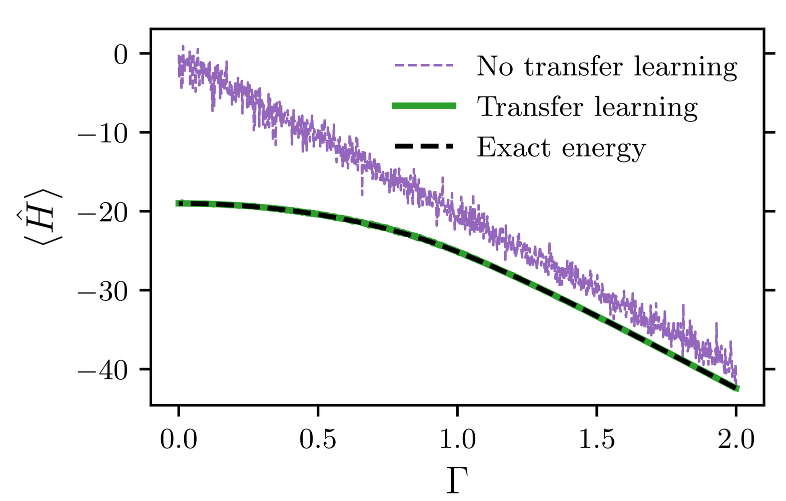
- The driving term $\hat{H}_D$ is initially the dominant term.
- We anneal $\hat{H}_D$ using a linear schedule:

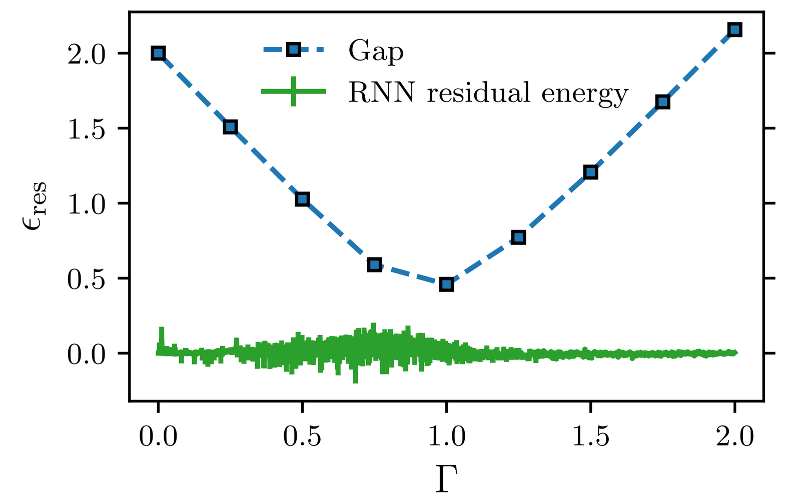
$$f(t) = (1-t)$$

- The wave function ansatz $|\Psi_\lambda\rangle$ is modeled by a recurrent neural network (RNN).



Proof of principle of adiabaticity on the 1D uniform Ising chain with $N=20$ spins and a driving term $\hat{H}_D = -\Gamma_0\sum_x \hat{\sigma}_i^x$. **Residual energy:** $\epsilon_{\text{res}} = \langle \hat{H}\rangle_\lambda - E_G$.

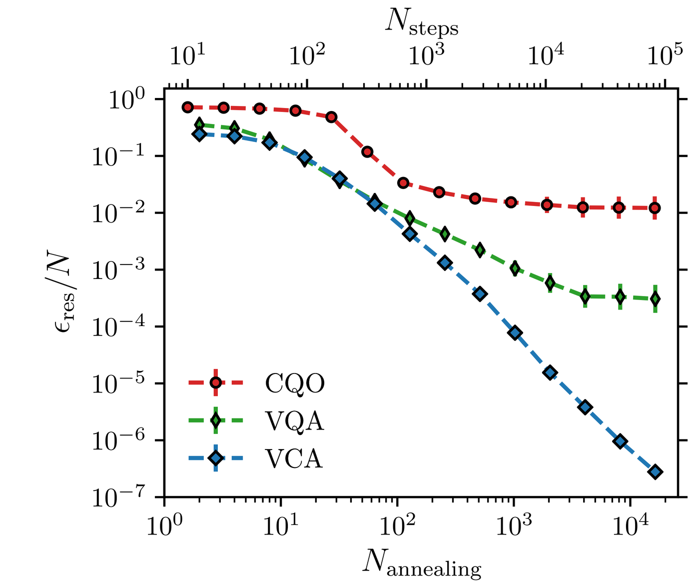




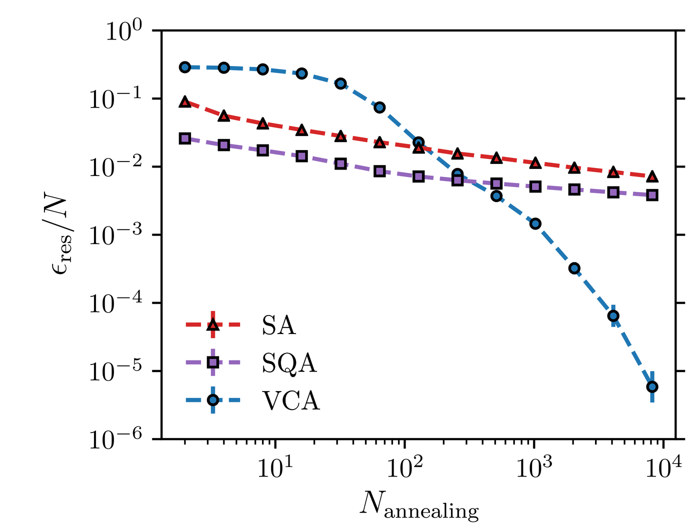
Annealing for two-dimensional spin glasses

Edwards-Anderson Hamiltonian in 2D:

$$\hat{H}_{\text{EA}} = -\sum_{\langle i,j\rangle} J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z,$$



Comparing CQO, VQA, VCA (10x10)



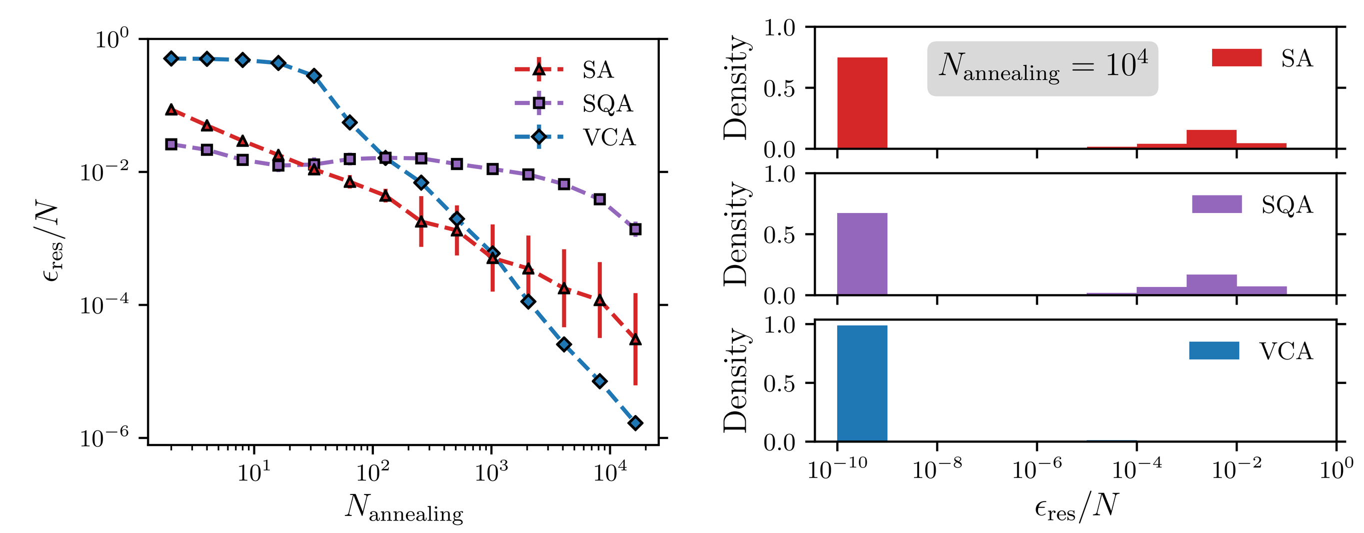
Comparing SA, SQA, VCA (40x40)

- VCA has a better scaling compared to CQO [3], VQA, SA and SQA [4].

Annealing for fully-connected spin glasses

Sherrington-Kirkpatrick Hamiltonian for $N=100$ spins:

$$\hat{H}_{\text{SK}} = -\frac{1}{2}\sum_{i\neq j}\frac{J_{ij}}{\sqrt{N}}\hat{\sigma}_i^z\hat{\sigma}_j^z,$$



- VCA is superior compared to SA and SQA [4].

Outlooks

- Understanding why VCA performs better compared to VQA.
- Taking advantage of state-of-the-art autoregressive models, besides traditional RNNs, to improve VQA and VCA.
- Applying VQA and VCA to real-world optimization problems.

References

1 Dilated Recurrent Neural Networks, NeurIPS, 2017.
2 Recurrent Neural Network Wave Functions, PRReasearch, 2020.
3 Classical Quantum Optimization with Neural Network Quantum States, 2019.
4 Theory of Quantum Annealing of an Ising Spin Glass, Science, 2002.