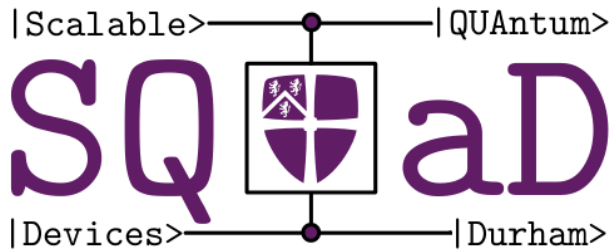
Quantum Annealing for Quantum/Classical hybrid algorithms

(partially) Based on: NJP 19, 2, 023024 (2017) and arχiv:1609.05875

Nicholas Chancellor

April 26, 2018

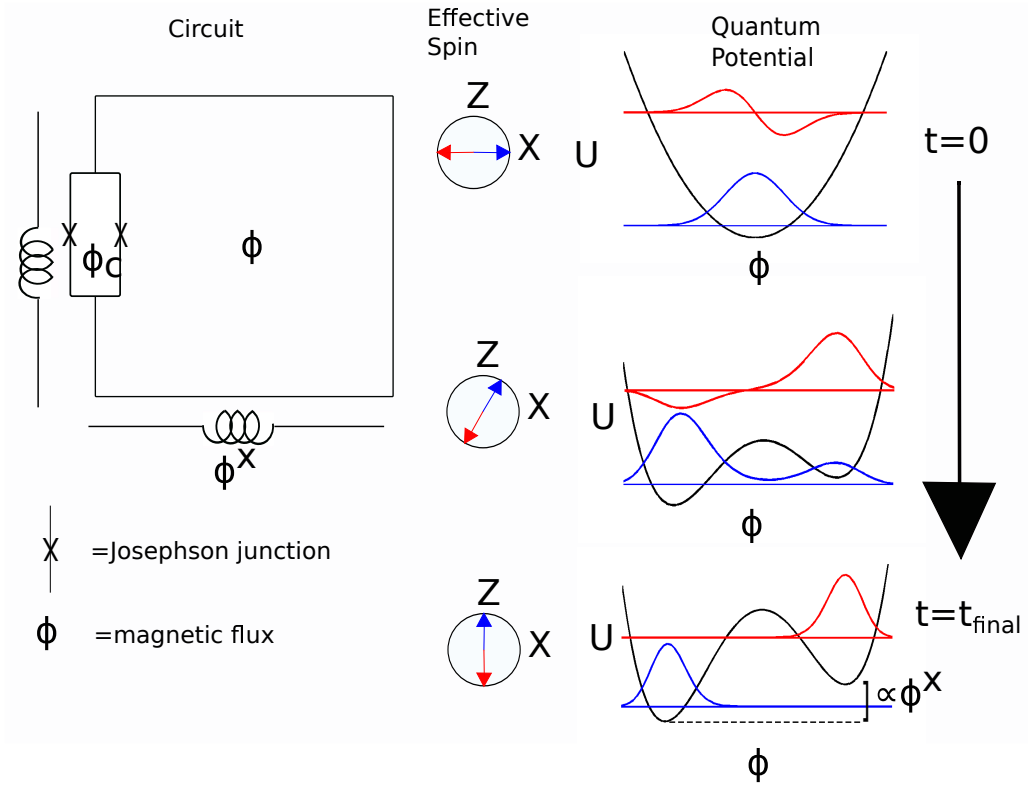


Outline

1. Introduction to quantum annealing
2. Reverse annealing and hybrid algorithms
   - Random start
   - Better algorithms inspired by classical Monte Carlo
3. Inference primitive formalism
   - Simple examples, traditional Quantum annealing and repeated local search
   - More complicated Parallel Tempering and Population Annealing algorithms
   - Genetic algorithms
4. Summary and outlook

Quantum annealing, the big picture

- Build a physical system where degrees of freedom correspond to binary variables (qubits)
- Map optimization/ machine learning problems such that lower energy $\rightarrow$ more optimal
- Use natural quantum dynamics to find highly optimal states.



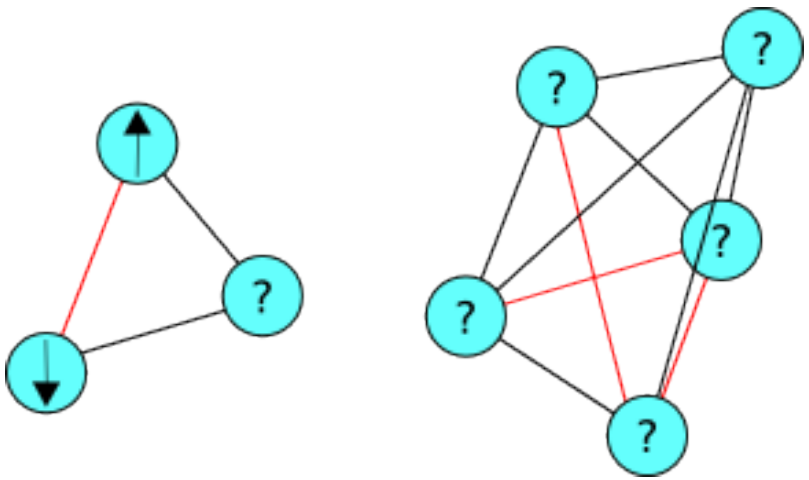
Problem Statement: Ising Spin Glass Hamiltonian

System is described by a diagonal Hamiltonian matrix:

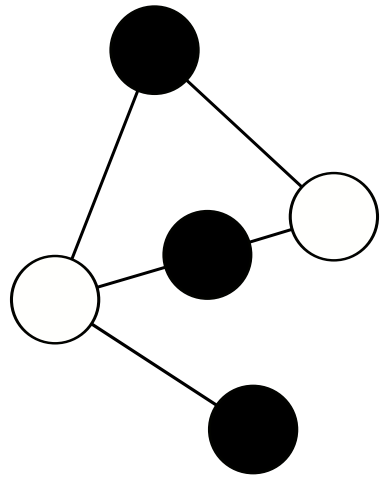
$$H_{ISG} = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z$$

$$\sigma_i^z \equiv (\bigotimes_1^{i-1} \mathbf{I_2}) \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes (\bigotimes_{i+1}^{n} \mathbf{I_2}))$$

- 'Universal' in the sense that any classical Hamiltonian can be mapped to it De las Cuevas, Cubitt Science 351 6278
- Thermal/quantum distributions also useful for inference and machine learning tasks ex. Amin et. al. arXiv:1601.02036, Chancellor et. al. Scientific Reports 6, 22318 ...



Simple example



Problem: find the maximal independent set of variables with interactions defined by a graph G
Mapping: For every edge in G set $J_{ij} = 1$, otherwise $J_{ij} = 0$, set all $h_i = -0.1$
smallest diagonal element of H_{ISG} (lowest energy state) will correspond to maximal independent set

$$H_{ISG} = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z$$

Making this quantum: adding (qu)bit flipping terms (known as transverse fields in the literature)

$$H = -A \sum_i h_i \sigma_i^x + B \left(\sum_i h_i \sigma_i^z + \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z \right)$$

$$\sigma_i^x \equiv (\bigotimes_1^{i-1} \mathbf{I_2}) \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes (\bigotimes_{i+1}^{n} \mathbf{I_2}))$$

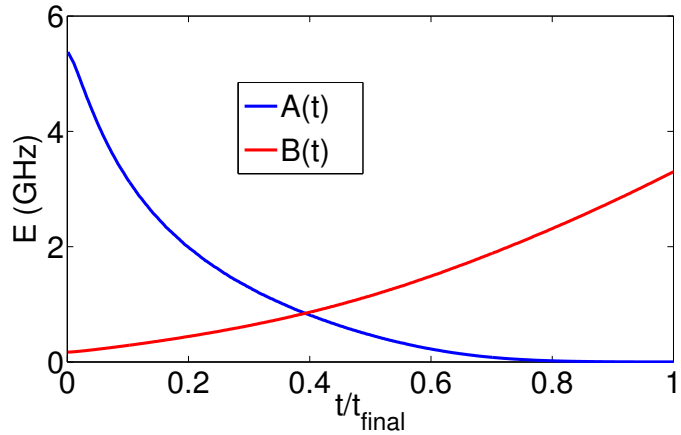
$$[\sigma_i^z, \sigma_i^x] \neq 0 \ \therefore$$

$A \gg B \rightarrow$ non-interacting qubits classical for technical reasons

$A \ll B \rightarrow$classical Ising spin glass, low energy states are good solutions to problem

$A \approx B \rightarrow$complicated quantum spin glass, quantum tunneling and thermal dissipation can team up to solve the problem

Quantum annealing (traditional)

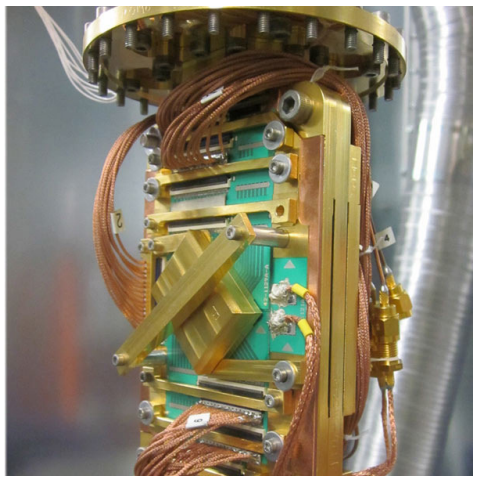


$$H = -A(t) \sum_i h_i \sigma_i^x + B(t) \left(\sum_i h_i \sigma_i^z + \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z \right)$$

- $\sigma_i^x \equiv (\bigotimes_1^{i-1} \mathbf{I_2}) \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes (\bigotimes_{i+1}^{n} \mathbf{I_2}))$
- $\sigma_i^z \equiv (\bigotimes_1^{i-1} \mathbf{I_2}) \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes (\bigotimes_{i+1}^{n} \mathbf{I_2}))$
- $A \approx B \rightarrow$complicated quantum spin glass, quantum tunneling and thermal dissipation can team up to solve the problem
- Slowly change Hamiltonian $\rightarrow$ nature solves problem for you

Is this just a nice idea, can you build one?

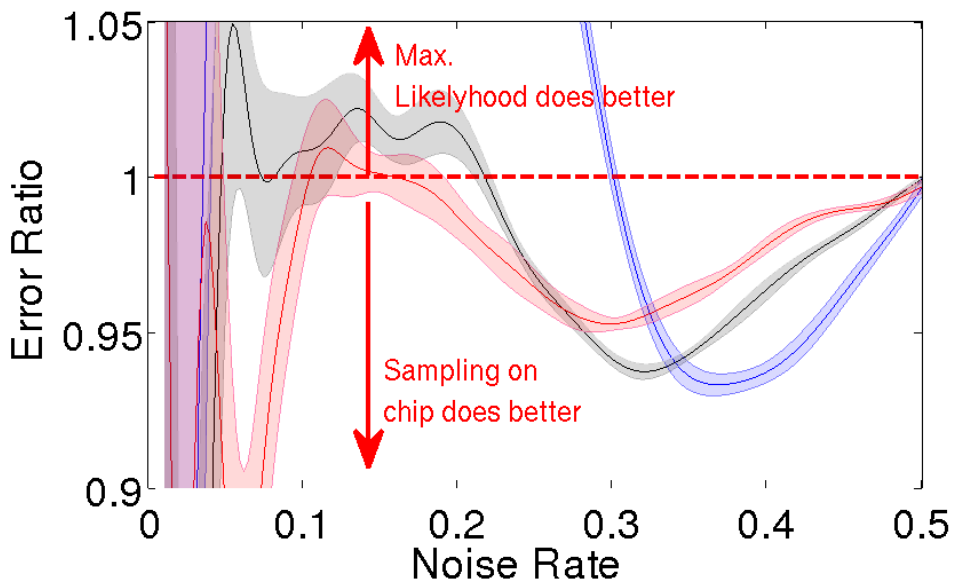
- Superconducting circuit quantum annealers commercially available from D-Wave systems Inc.
- Devices have over $2,000$ qubits
- You can own one for $\approx$£10M (some have been sold)
- Interest from big names like Google, Lockheed Martin, and NASA
- No conclusive 'killer app' yet, but people are working on this



(image courtesy of D-Wave systems Inc.)

Thermal sampling

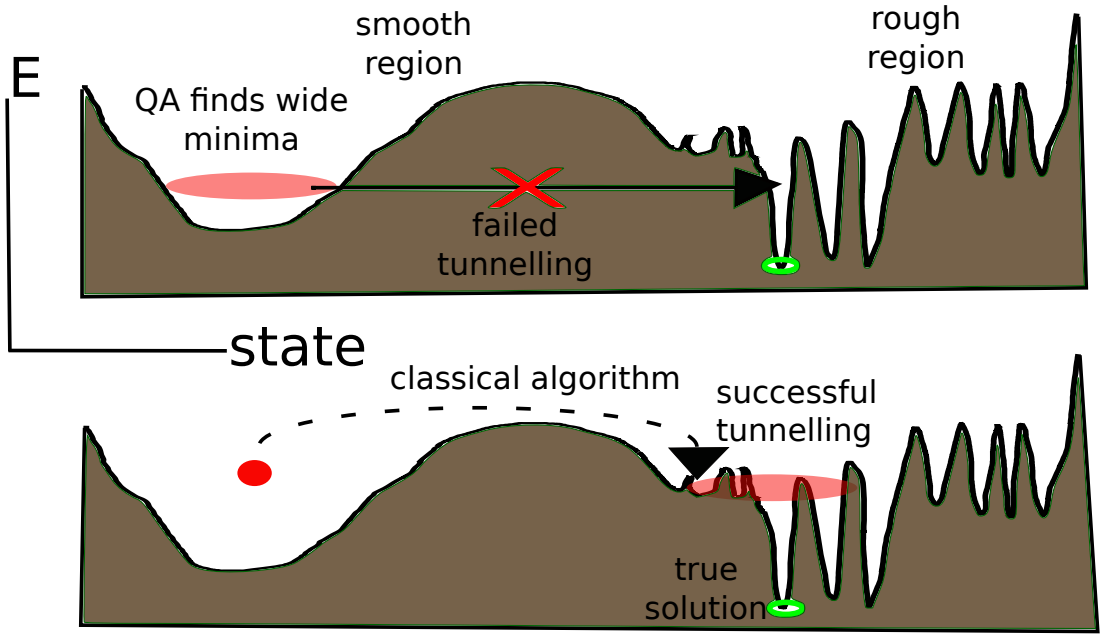
- Because of the role thermal fluctuations play in these devices, they not only act as *optimizers*, but also thermal *samplers*
- Can actually be used for a maximum entropy decoder $\rightarrow$ can do better than maximum likelihood for some error rates and codes[1]: Scientific Reports 6, Article number: 22318 (2016)
- Work has also been done on Boltzmann machines &c. but I will not talk about this here



[1] The example here is a simple code which matches the hardware graph

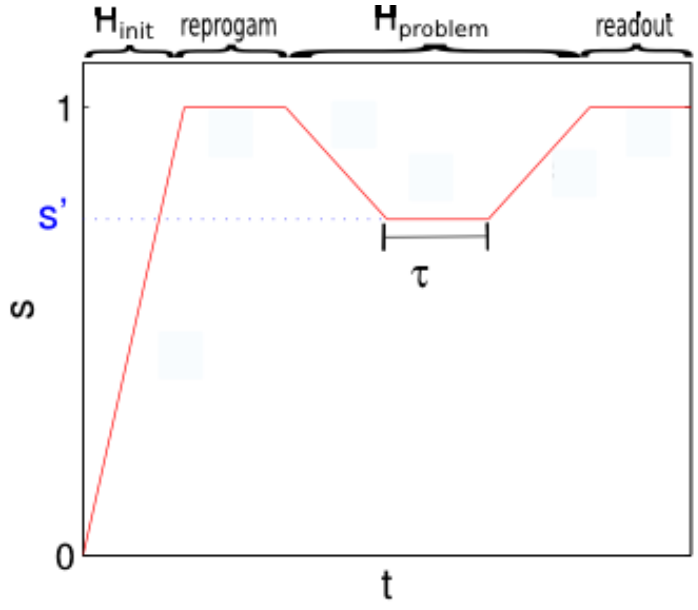
How to make this better

- First 'run' just as likely to find a good solution as the last better if we could use info from previous tries
- Already have pretty good classical algorithms for many of these problems, would be nice if these could help
- QA can be 'fooled' by broad local minima in energy landscapes, good at sharp pointy landscapes, bad at broad smooth ones



How to do this: reverse annealing

$$H = -A(s) \sum_i h_i \sigma_i^x + B(s) \left(\sum_i h_i \sigma_i^z + \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z \right)$$



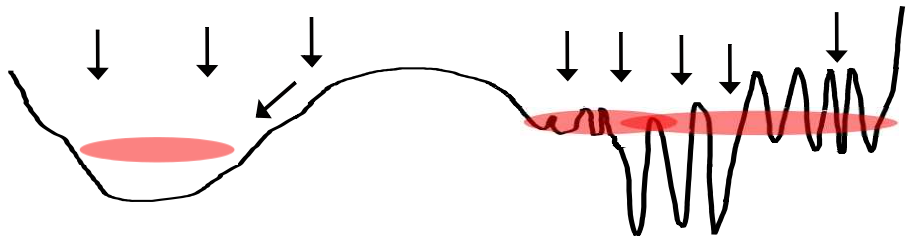
1. Start with easy problem annealer finds the (known) lowest energy state with essentially 100% probability
2. Reprogram to hard problem
3. Turn up bit flipping terms to search the space, controlled by value of s'
4. Possibly wait for some time τ and finish protocol as normal

Time to think about the algorithms this enables

(logo courtesy of D-Wave systems Inc. used with permission)

- These protocols were added to D-Wave device controls earlier this year
- Let's think about how to use them algorithmically

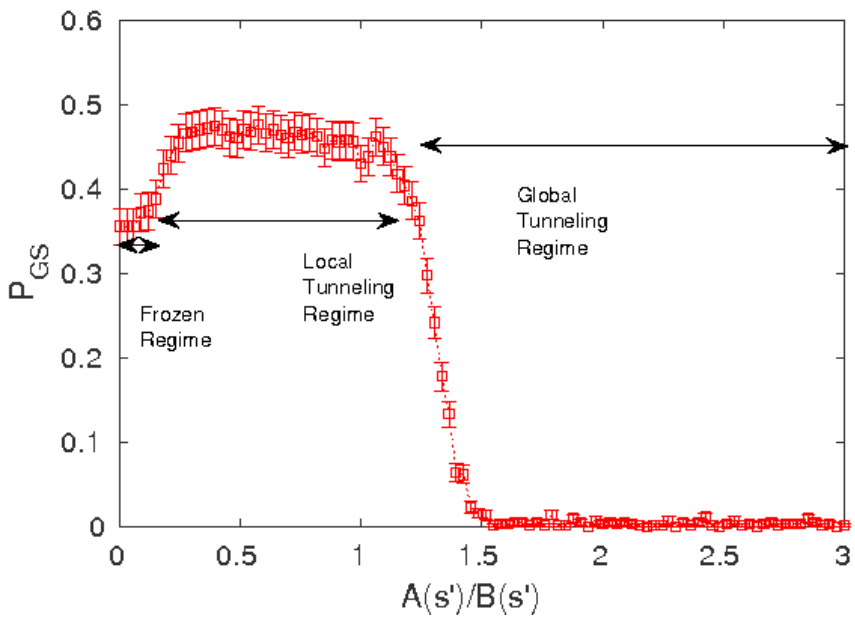
The dumbest algorithm I can think of: start in a random state



- Can this actually do anything useful?
- Yes! Can avoid getting stuck in a single broad local minima
- Can search different parts of the space and help get a fair thermal sampling (important for some machine learning)

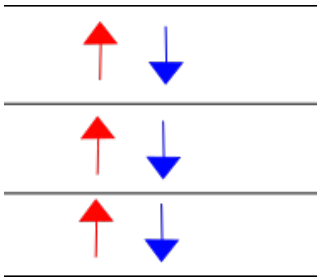
A numerical example of this 'dumb' algorithm

- Construct a small 'toy' problem with the energy landscape given on previous slide
- Details not important for the discussion here
- Simulate with Monte Carlo, look at probability to find lowest energy 'ground' state P_{GS}



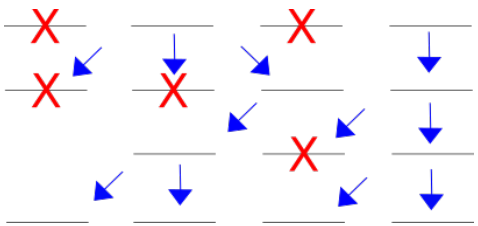
Better algorithms: inspiration from classical Monte Carlo

Parallel Tempering



- Multiple replicas at different temperatures
- Swap replicas by rules which obey detailed balance
- $P_{swap}(i,j) = \min\left[1, \exp\left(\left(\frac{1}{T(i)} - \frac{1}{T(j)}\right)(E_i - E_j)\right)\right]$

Population Annealing



- Lower T for multiple replicas
- Probabilistically remove poorly performing replicas and copy those which perform well
- Rules preserve average population and obey detailed balance
- $\bar{N}(E) = \frac{1}{Q}\exp\left(\left(\frac{1}{T_{old}} - \frac{1}{T_{new}}\right)E\right)$

Probability to transition between states in terms of temperature T:

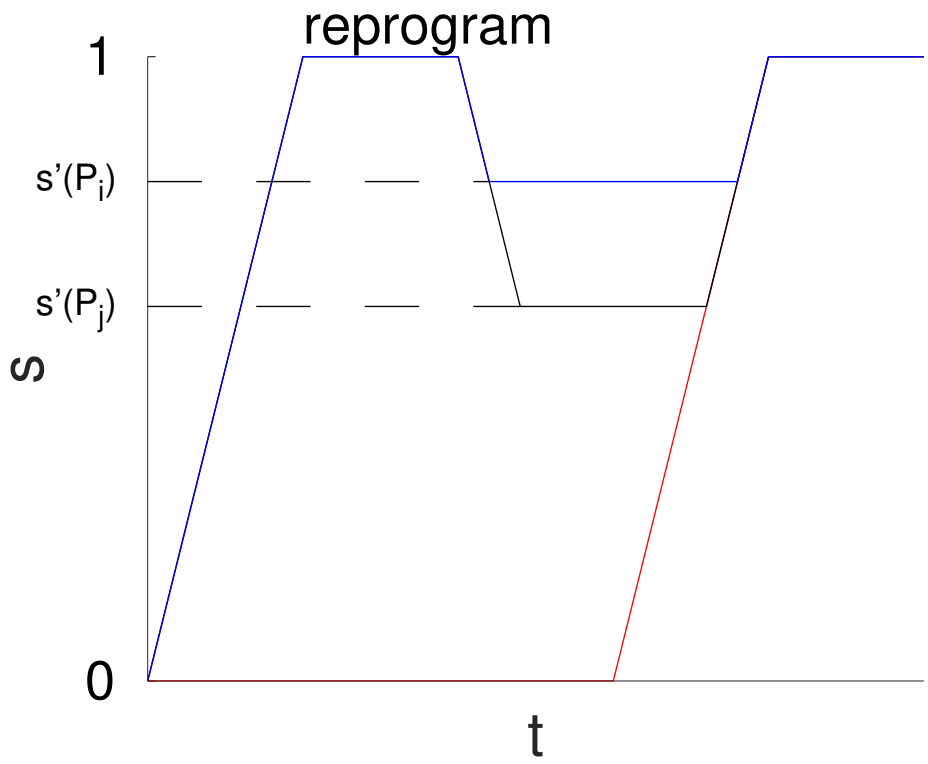
$P(S(1) \to S(2)) = \min[\exp(\frac{(E(1)-E(2))}{T})P(S(2) \to S(1)), 1]$

Making quantum annealing versions

1. Replace updates with annealing runs consisting of calls to annealer
2. Define 'energy' and 'state' as the lowest energy solution found in an annealing run and the corresponding classical state
3. Replace $T \rightarrow T_{eff}$, effective temperature derived from s'
4. Apply replica swapping/copying/deleting rules as usual

Including uncertainty by annealing qubits differently

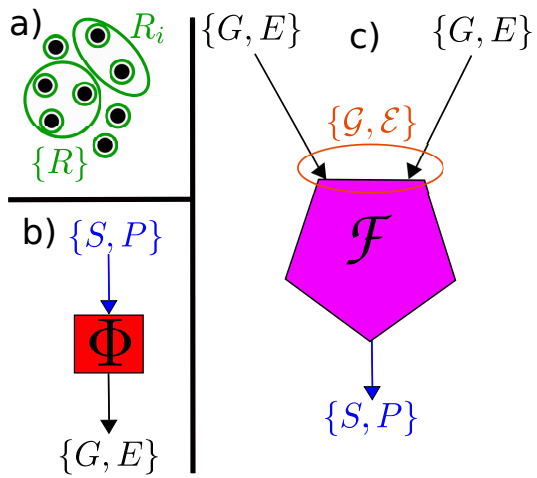
What if we are more sure about some parts of our guess then others? $\rightarrow$ anneal different qubits back to different points



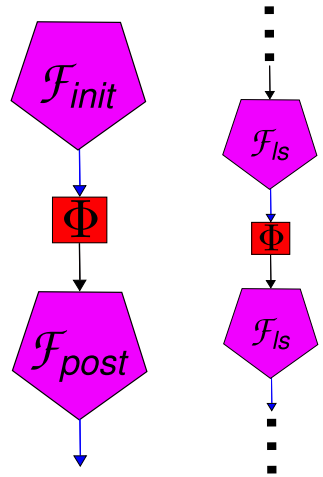
An extreme version of this, which excluded qubits where a value was expected with high certainty has already been done H. Karimi and G. Rosenberg Quantum Inf. Proc. 16(7):166 (2017) and H. Karimi and G. Rosenberg Phys. Rev. E, 96:043312

Representing this graphically: Inference Primitive Formalism

- Represent quantum annealing call as an inference primitive Φ, takes state guess $S \in \{-1, 1\}$ and uncertianty values $P \in [0, 0.5]$, outputs list of states G and energies E
- Processing function $\mathcal{F}$ represents classical processing $\rightarrow$ takes any number (including zero) of annealer outputs (found states G and energies E) and finds new guess S and uncertainty values P
- Easily generalized to multi-body drivers representing uncertainty on clusters of qubits



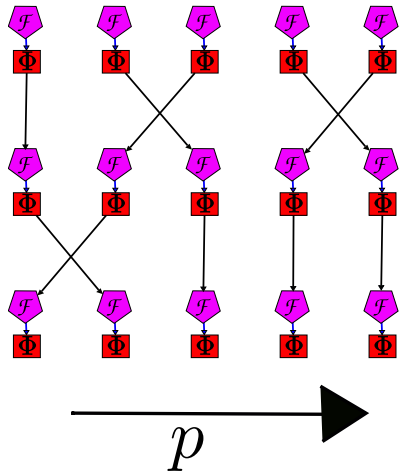
Basic Examples: traditional QA, and repeated local search in this formalism

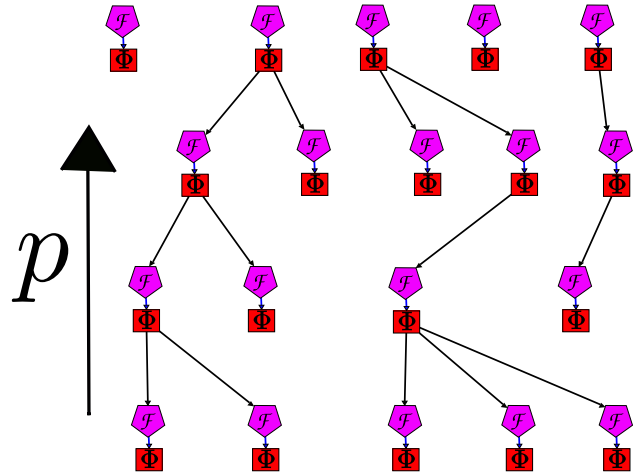


- Traditional QA (left) represented by initialization processing function which takes no inputs and gives complete uncertainty ($P_i = 0.5 \forall i$) on all qubits, followed by post processing function
- Repeated local search (right) from running annealer many times and using the output as an input to the next processing function

More advanced algorithms: Parallel tempering and Population annealing analogues

- Processing function $\mathcal{F}$ returns lowest energy state as guess and gives all qubits the same uncertainty $P_i = p \forall i$
- Assign effective temperature T to each p value and either:
  1. exchange using Parallel tempering rules (left)
     $P_{swap}(i,j) = \min\left[1, \exp\left(\left(\frac{1}{T(i)} - \frac{1}{T(j)}\right)(E_i - E_j)\right)\right]$
  2. kill or replicate states using population annealing rules (right)
     $\bar{N}(E) = \frac{1}{Q}\exp\left(\left(\frac{1}{T_{old}} - \frac{1}{T_{new}}\right)E\right)$





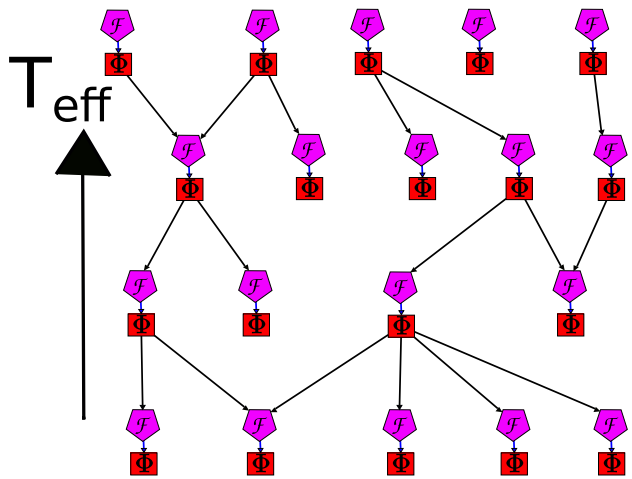
Even more advanced algorithms: Genetic algorithms

- A processing function which takes more than one input is a 'breeding' step of a genetic algorithm
- For instance could be thermally reweighted sum[2] (u indicates sum over unique states found)

$$S_i = \mathrm{sgn}(\sum_{j=1}^{N_u} G_j^{(u)} \exp(-\frac{E_j^{(u)}}{T_{eff}})),$$

$$P_i = \frac{1}{Z}(\sum_{j=1}^{N_u} \delta_{G_j^{(u)},-S_i} \exp(-\frac{E_j^{(u)}}{T_{eff}}))$$

- Could be used to add crossbreeding to Population annealing analogue, as shown below



[2] see: arχiv:1609.05875 for details

Summary and Outlook

- Quantum annealing is an alternative approach to gate based quantum computers
- 'Energy landscape' approach conducive to many important optimization and machine learning problems
- New features in annealing devices allow for (more) hybrid quantum/classical algorithms
- I have recently been given NQIT funding to test these features

- It is an exciting time to work in quantum annealing!
- This subject is multidisciplinary by nature, plenty of room for 'classical' people to contribute without having to learn quantum mechanics