MSDM 5001 Week 10

11 November 2023

Quantum Annealing

Outline:

10.1 Combinatorial Optimization

- Combinatorial optimization problems consist of assigning discrete variables in order to minimize cost functions.
- There are many interesting examples:
  - Graph partitioning
  - Traveling salesman problem
  - Max-cut
  - Traffic routing
  - Image restoration
  - K-SAT (K-satisfiability)
  - Machine learning

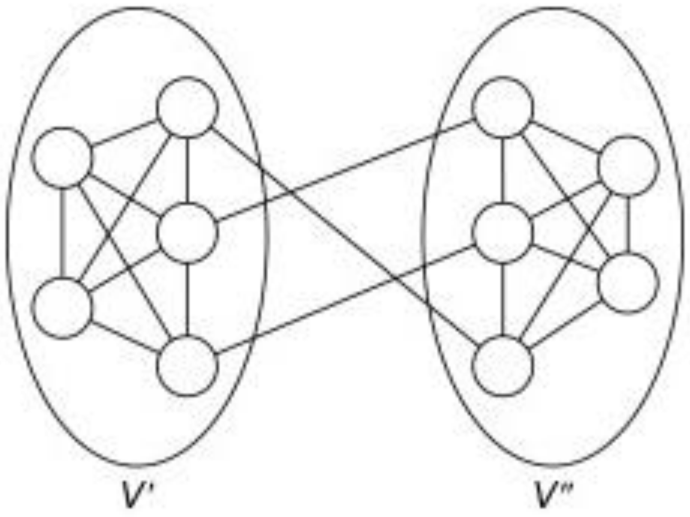
Applications

According to Wikipedia, applications for combinatorial optimization include, but are not limited to:

- Logistics
- Supply chain optimization
- Developing the best airline network of spokes and destinations
- Deciding which taxis in a fleet to route to pick up fares
- Determining the optimal way to deliver packages
- Working out the best allocation of jobs to people
- Designing water distribution networks
- Earth Science problems (e.g. reservoir flow-rates)

Graph Partitioning

- How to divide the nodes into two sets such that the links between the two sets is minimized?
- Constraint: The two sets have the same size.
- Useful in chip design.
- Formulation: Let $S_i = +1$ for $i \in V'$ and $S_i = -1$ for $i \in V"$.
- Minimize $E = -\sum_{i,j} A_{ij} S_i \, S_j$.
- Subject to the constraint $\sum_i S_i = 0$.



Traveling Salesman Problem

- How to design a trip around N cities such that the total traveling distance is minimized?
- Constraint: Every city is visited once only, and at every instant, the agent can only stay in one city.
- Formulation: Let $S_{ij} = 1$ if city i is visited at the j^{th} turn, and $S_{ij} = 0$ otherwise.
- Minimize $E = -\sum_{i,k,j} D_{i,k} S_{ij}\, S_{k,j+1}$.
- Subject to the constraints $\sum_j S_{ij} = 1$ and $\sum_i S_{ij} = 1$.

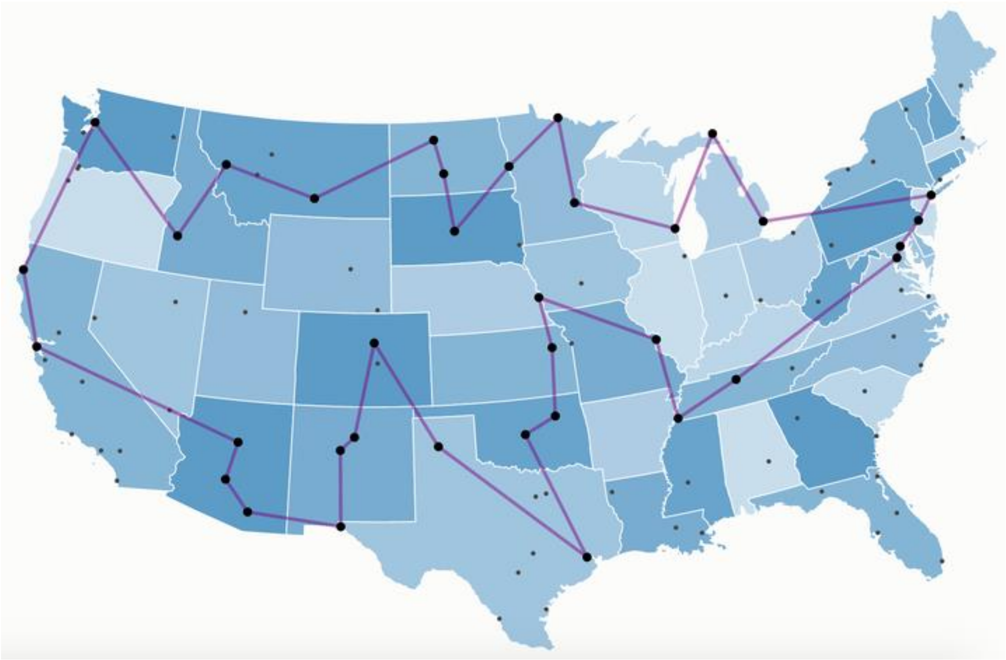


Image Restoration

- Consider an original image of binary pixels $S_i = \pm 1$. Suppose the corrupted image is T_i. The likelihood function is given by $P(T_i|S_i) \propto e^{bT_iS_i}$.
- Image restorations are performed with the help of prior distributions, e.g. smoothness prior $P(S_i, S_j) \propto e^{cA_{ij}S_iS_j}$ for neighboring pixels.
- Using Bayes' rule, $P(\{S_i\}|\{T_i\}) \propto e^{\sum_i bT_iS_i + \frac{1}{2}\sum_{i,j} cA_{ij}S_iS_j}$.
- Maximum a posteriori restoration: minimize $E = -\sum_i bT_iS_i - \frac{1}{2}\sum_{i,j} cA_{ij}S_iS_j$

Lena

General Formulation

Therefore, combinatorial optimization problems can be generally formulated as the minimization of the energy function

$$E = -\sum_i h_i S_i - \frac{1}{2} \sum_{i,j} J_{ij} S_i S_j \, .$$

For example, the graph partitioning problem can be formulated as

$$E = -\sum_{i,j} A_{ij} S_i S_j + P \left(\sum_i S_i \right)^2 = -\sum_{i,j} B_{ij} S_i S_j \, .$$

Spin Glasses

The cost function of combinatorial optimization problems is very similar to the energy function of spin glasses.

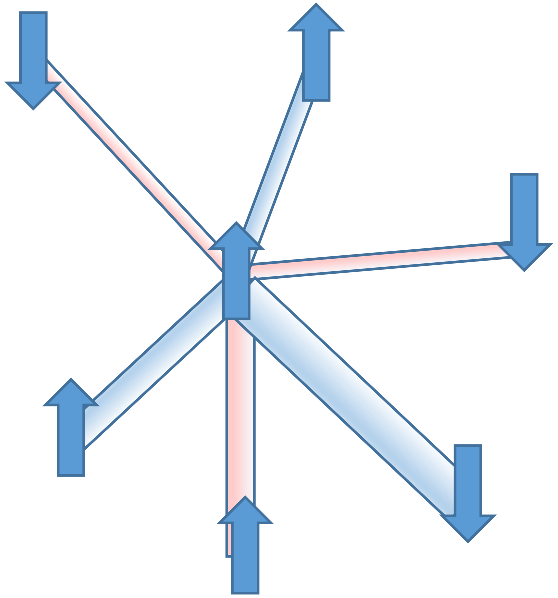
$$E = -\sum_i h_i S_i - \frac{1}{2}\sum_{i,j} J_{ij} S_i S_j \, .$$

Spin glasses are formed when a non-magnetic metal (e.g. copper) is randomly doped with magnetic atoms (e.g. iron).

The magnetic coupling between the spins are mediated by electrons and are non-trivial functions of distance.

At some distances, the coupling is ferromagnetic (i.e. tend to align the spins in parallel direction).

At other distances, the coupling is antiferromagnetic (i.e. tend to align the spins in antiparallel direction).



Phase Transitions in Spin Glasses

Sherrington and Kirkpatrick proposed a model of spin glasses (SK model). Result:

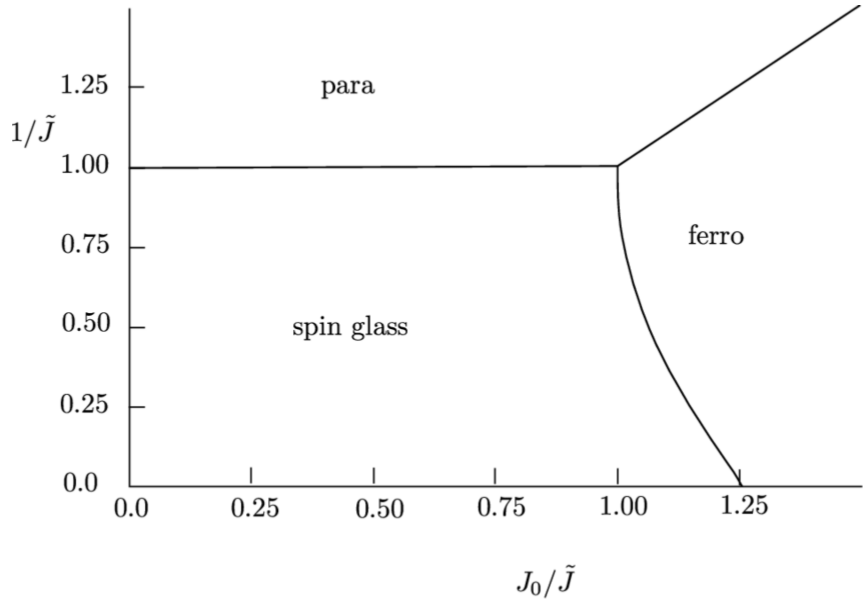
Horizontal axis: ferromagnetic bias (increases with the fraction of ferromagnetic couplings)

Vertical axis: temperature

High temperature: paramagnetic phase (the spins have no definite directions).

High ferromagnetic bias: ferromagnetic phase (the spins are aligned with the same direction).

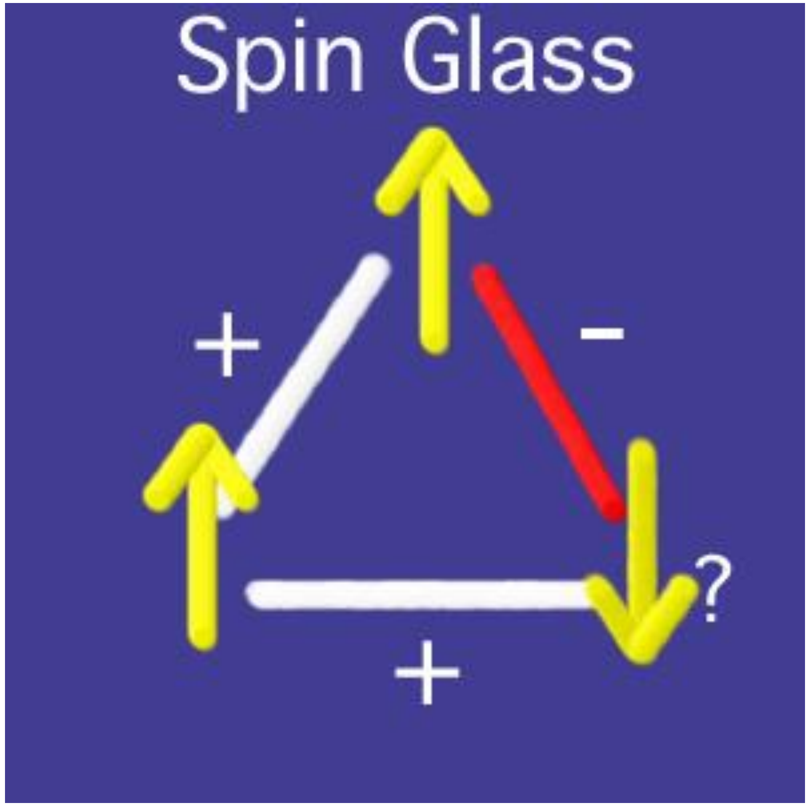
Low temperature and weak ferromagnetic bias: spin glass phase (each spin is frozen along its own direction, but their aligned directions are not uniform).



10.2 Simulated Annealing

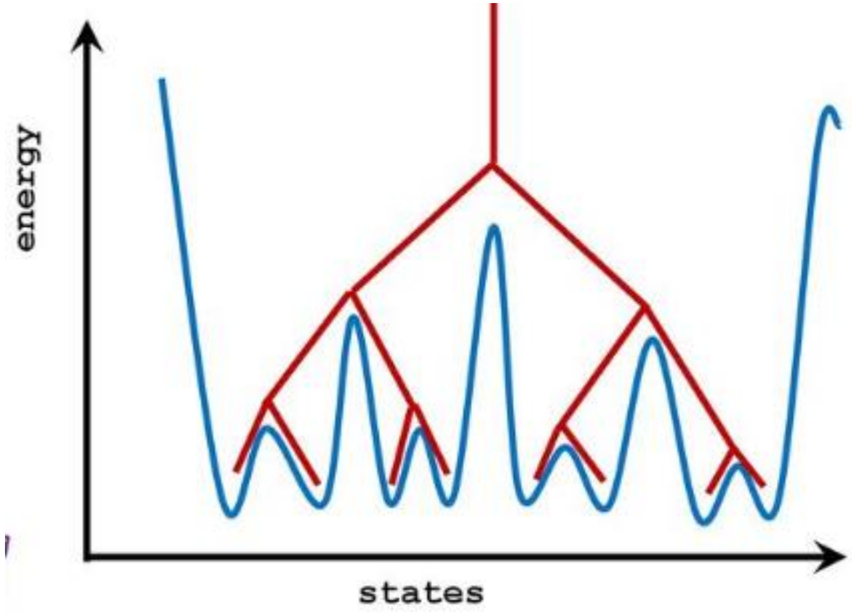
Frustration

- However, the SK model still cannot explain many strange properties of spin glasses.
- For example, their behaviors depend on history, and their approach to equilibrium has very long time-scale (similar to glasses, which are supercooled liquids).
- Then it was realized that their properties are caused by frustration.
- Due to the mixture of ferromagnetic and antiferromagnetic couplings, frustrations exist in spin glasses, that is, not all couplings are "happy".

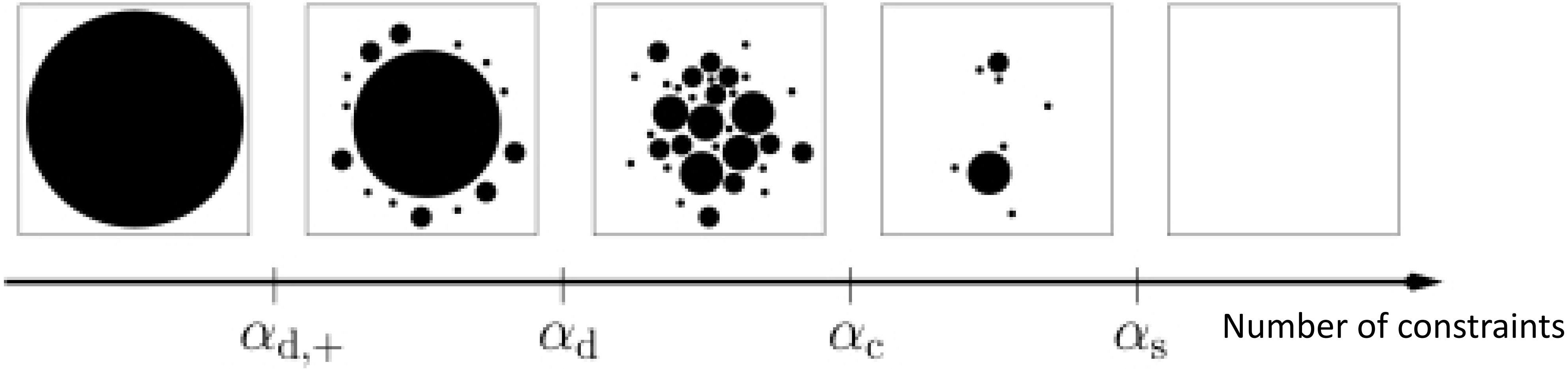


Rugged Energy Landscape

- A consequence of frustration is that the energy landscape becomes very rough.
- Besides the ground state (global minimum), there are many metastable states (local minima).
- In fact, the energy landscape of the SK model consists of hierarchies of valleys.



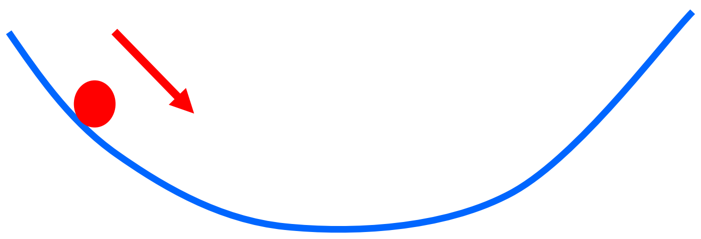
Phase transitions in random constraint satisfaction problems

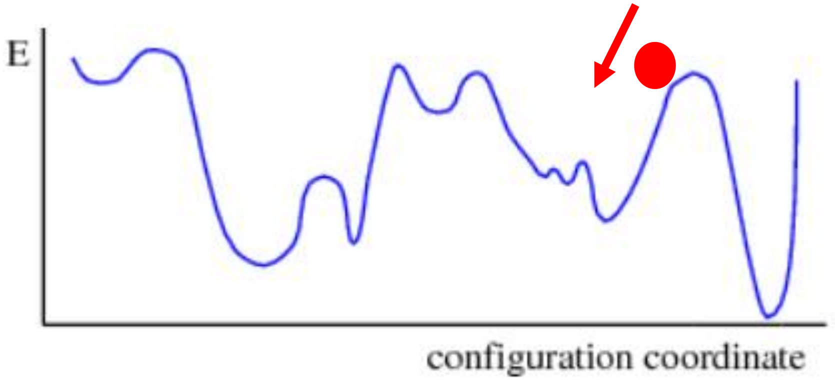


Florent Krzakala, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian, Lenka Zdeborova. Gibbs States and the Set of Solutions of Random Constraint Satisfaction Problems. Proc. Natl. Acad. Sci. 104, 10318 (2007).

Implications to Optimization Problems

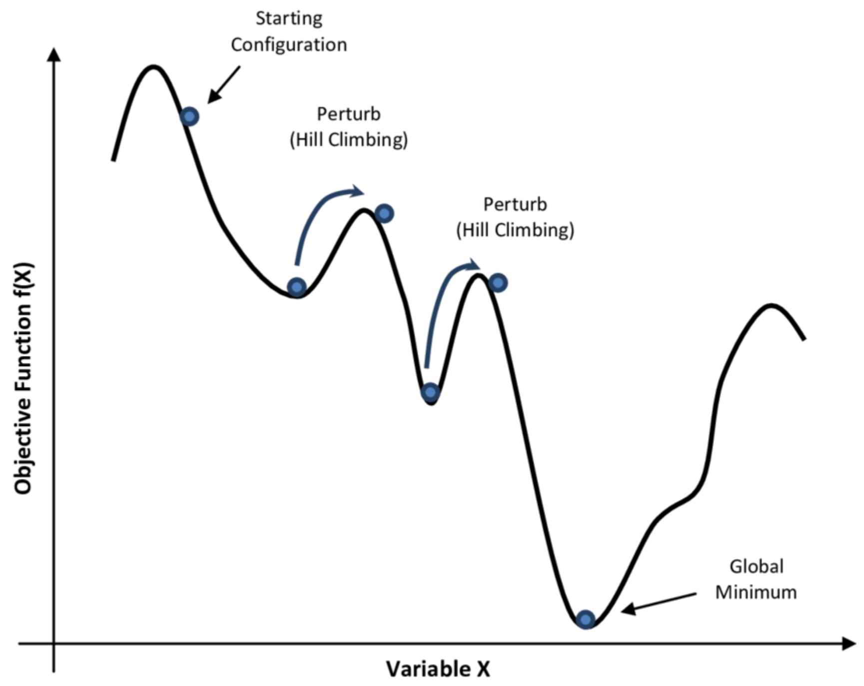
- The most direct approach to optimization is to start with an initial condition and flip the spins repeatedly until it reaches a minimum (gradient descent).
- However, in a complex energy landscape, the final solution becomes heavily dependent on the initial condition.
- There is a high probability that one ends up at a local minimum.





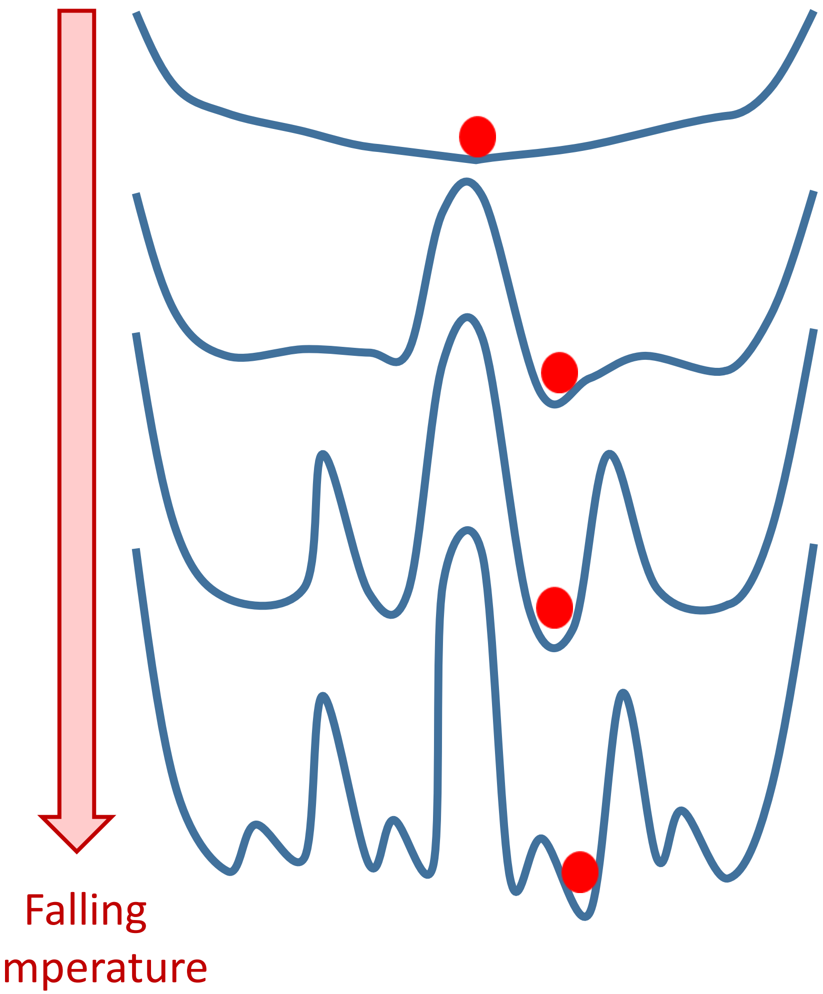
Adding Noise

- By adding noises to the algorithm, the state of the system can overcome the energy barriers and escape from the local minima.
- However, noises can also cause the state of the system to escape from the global minimum as well. So what is the correct amount of noise?
- If the probability of moving a step is proportional to the Boltzmann factor $e^{-\Delta E/T}$, then the probability of the average energy of the final solution is the same as the average energy at temperature T.



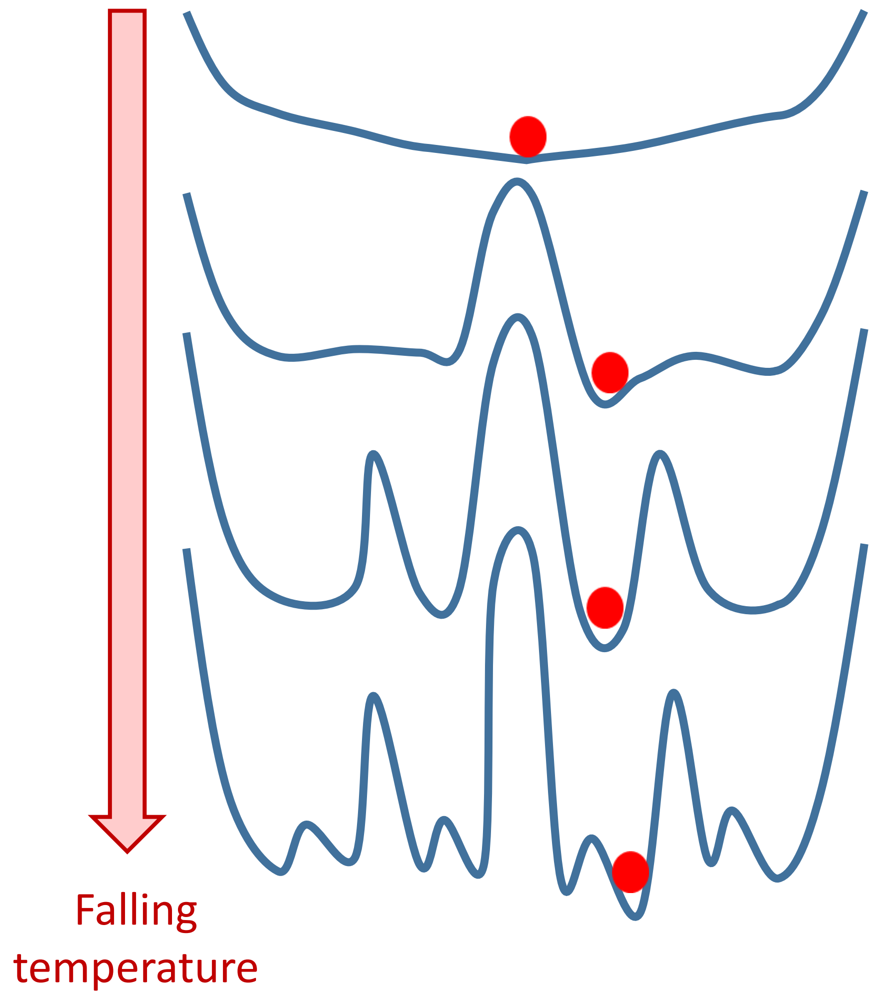
Free Energy

- The average energy at temperature T is not the same as the minimum energy at $T = 0$ (with no noise). At temperature T,
$$P(E) \propto \Omega(E) e^{-\frac{E}{T}}.$$
- $\Omega(E)$ is the volume of states with energy E. It is related to the entropy by
$$S(E) = \ln \Omega(E).$$
- Therefore, $P(E) \propto e^{-\frac{E - TS(E)}{T}}$.
- $F = E - TS$ is called the free energy.
- That is, the system is minimizing the free energy at temperate T, not minimizing the energy.



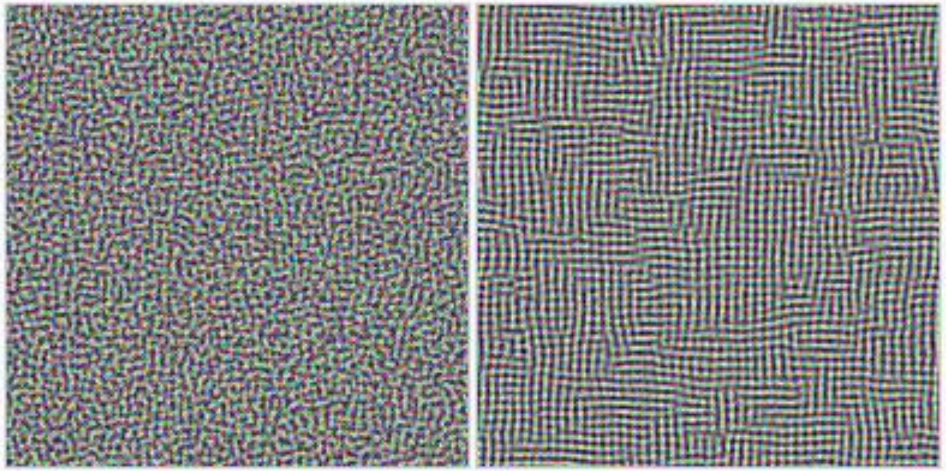
Simulated Annealing

- At high temperature, the free energy landscape is smooth. (Recall the paramagnetic phase at high temperature of spin glass.)
- To overcome the problem of rugged energy landscape at low temperature (recall the low temperature spin glass phase), simulated annealing is proposed.
- The system is simulated at high temperature, and the temperature is gradually reduced, hoping that the system is gradually led to the state of minimum energy when the temperature approaches 0.



Cooling Schedule

- The quality of simulated annealing depends on the cooling schedule.
- The temperature change should be slow enough to ensure that the system is not far from equilibrium during the cooling process.
- This is particularly important near the phase transition temperature. Critical slow down exists near the transition temperature.
- For image restoration, the theoretical cooling rate is $T \propto 1/\ln(1+t)$.



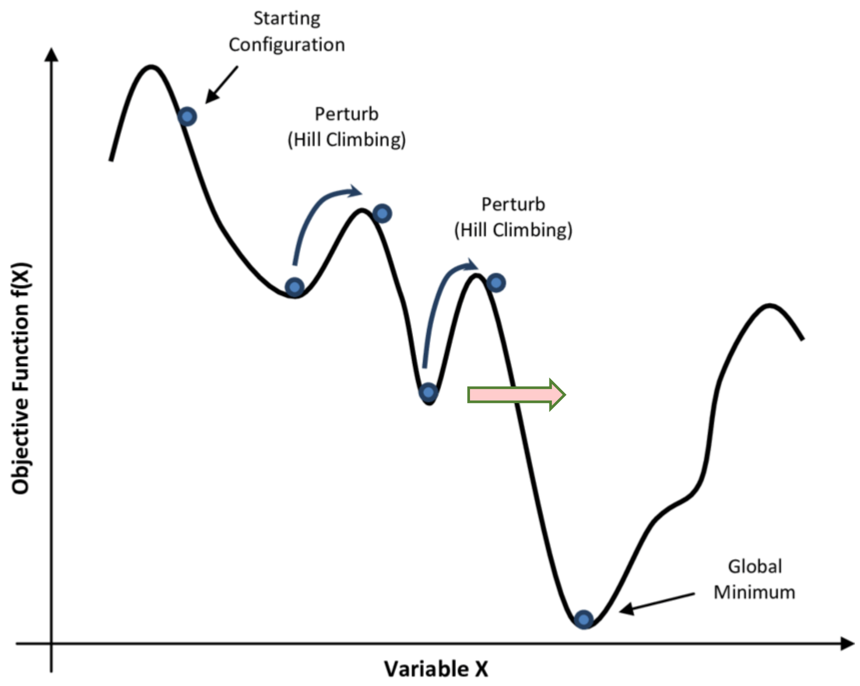
Fast Slow

Example illustrating the effect of cooling schedule on the performance of simulated annealing. The problem is to rearrange the pixels of an image so as to minimize a certain potential energy function, which causes similar colours to attract at short range and repel at a slightly larger distance. The elementary moves swap two adjacent pixels. These images were obtained with a fast cooling schedule (left) and a slow cooling schedule (right), producing results similar to amorphous and crystalline solids, respectively.

S. Geman; D. Geman (1984). "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". IEEE Transactions on Pattern Analysis and Machine Intelligence. 6 (6): 721–741.

Hill Climbing vs Tunneling

- The slow dynamics of simulated annealing is also related to the time to climb over the energy barriers.
- It may be more efficient to "dig a tunnel" penetrating the energy barriers?
- Quantum tunneling is proposed.
- Quantum tunneling is a phenomenon unique to quantum mechanics. This is the next topic.



10.3 Quantum Tunneling: de Broglie waves

- In 1924 a French physicist, Louis de Broglie (pronounced “de broy”), proposed that particles may, in some situations, behave like waves.
- A free particle with mass *m*, moving with speed *v*, should have a wavelength related to its momentum:

Planck’s constant

De Broglie wavelength of a particle $$\lambda = \frac{h}{p} = \frac{h}{mv}$$ Particle’s speed

Particle’s momentum   Particle’s mass

- A particle’s frequency is related to its energy in the same way as for a photon:

Planck’s constant

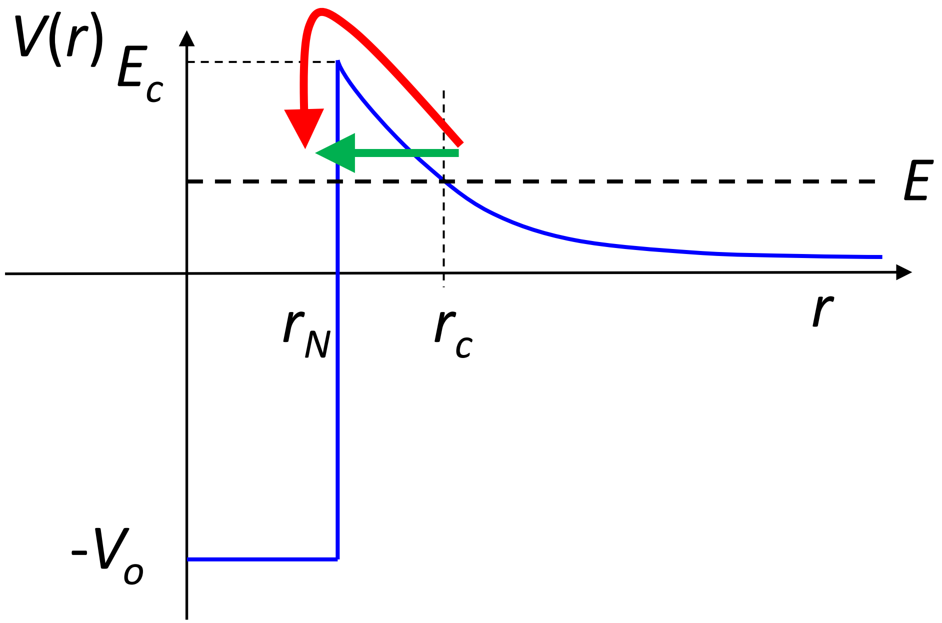
Energy of a particle $$E = hf$$ Frequency

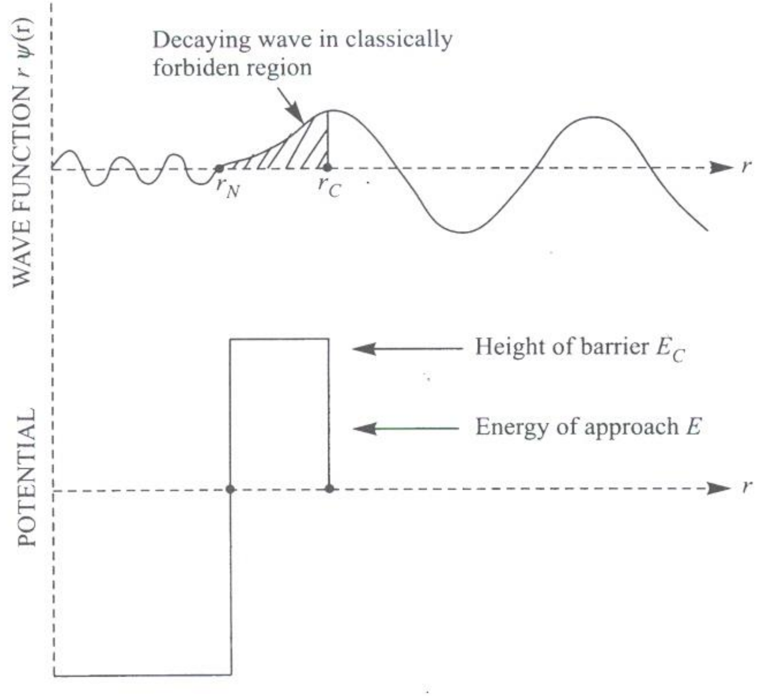
The Most Important Tunneling in the Solar System

Due to electric repulsion of the protons, an energy barrier prevents protons to get close enough to start nuclear fusion.

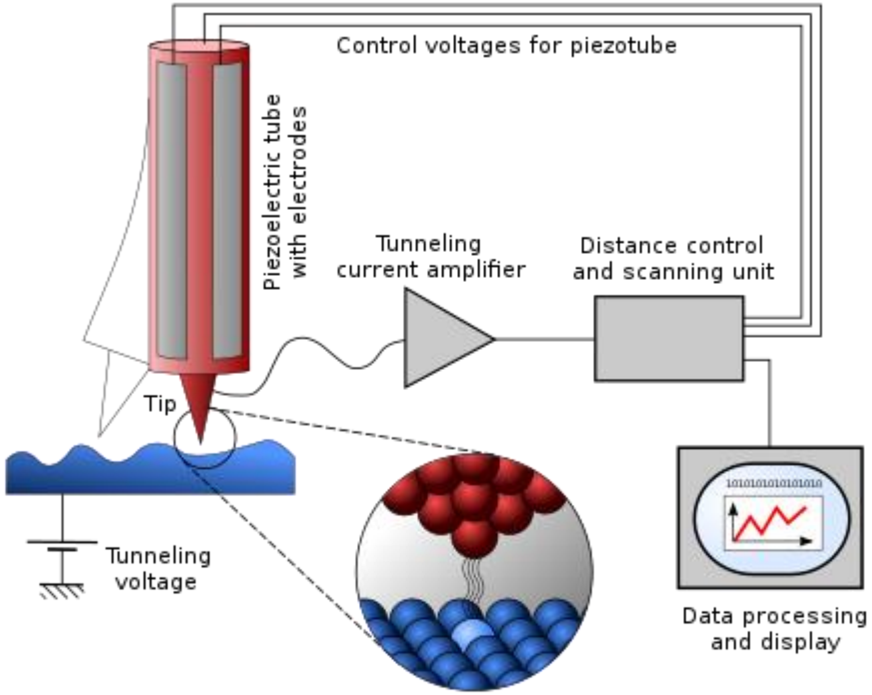
Due to the wave nature of protons, they can penetrate the barrier by quantum tunneling.

The probability of penetration is dramatically improved from 10^{-400} to 10^{-10}.





Scanning Tunneling Microscope



Schematic view of an STM

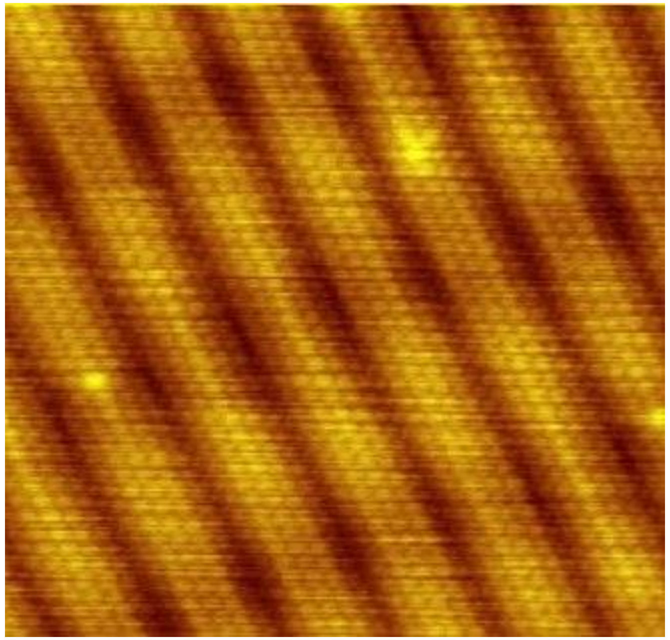
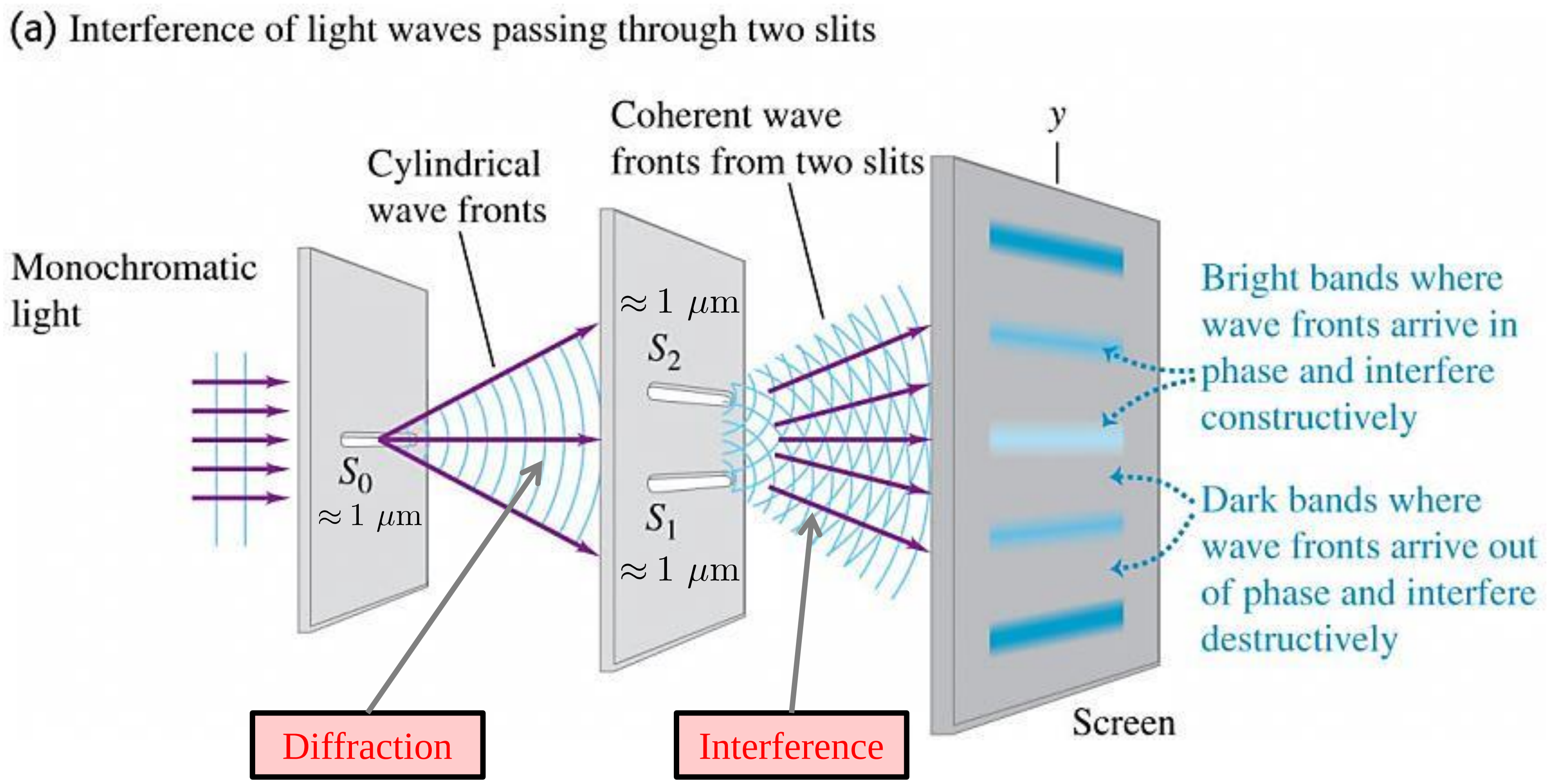


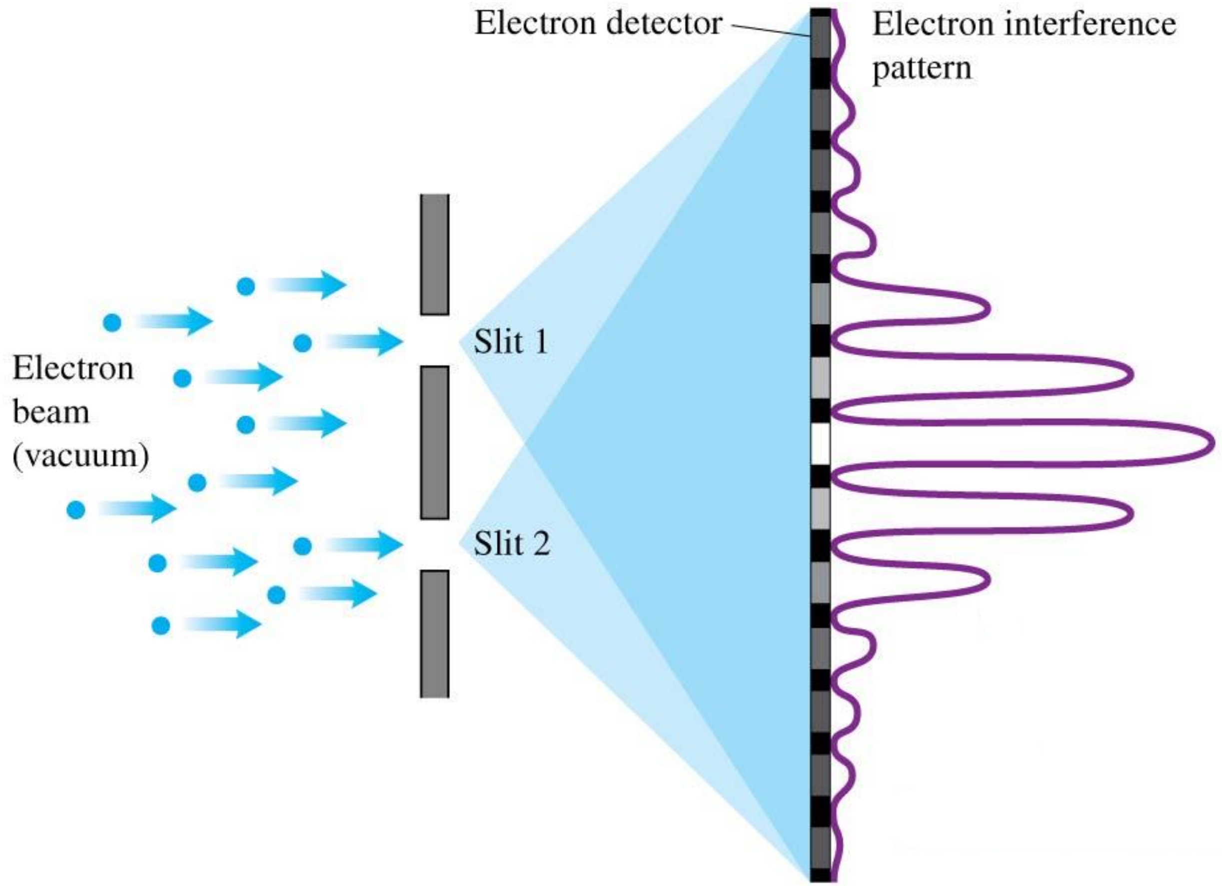
Image of reconstruction on a clean (100) surface of gold.

Waves Only: Diffraction and Interference

Two-Source Interference of Light: Young's Experiment

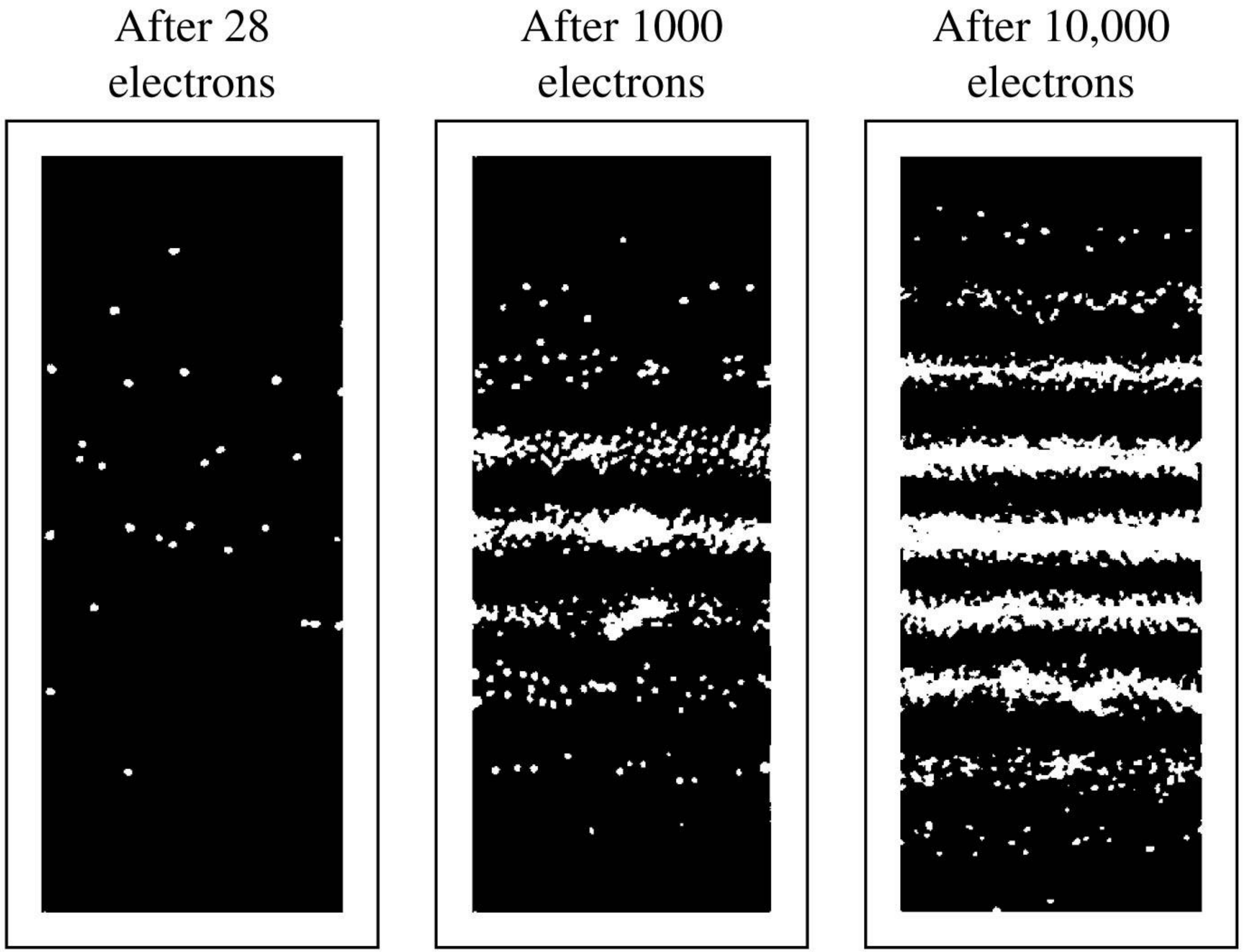


A two-slit interference experiment for electrons



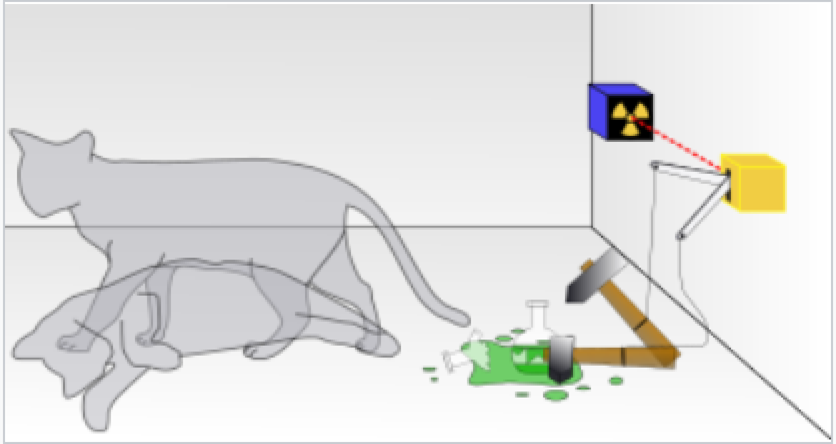
We can send the electron once a time through the slits. Each electron wave interferes with itself!!!

A two-slit interference experiment for electrons



Copenhagen Interpretation of QM

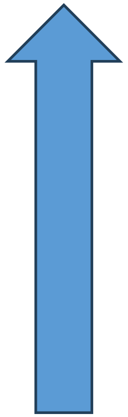
- Material objects, on a microscopic level, generally do not have definite properties prior to being measured.
- Quantum mechanics can only predict the probability distribution of a given measurement's possible results.
- Wave function collapse: The act of measurement affects the system, causing the set of probabilities to reduce to only one of the possible values immediately after the measurement.



Schrödinger's cat: a cat, a flask of poison, and a radioactive source are placed in a sealed box. If an internal monitor (e.g. Geiger counter) detects radioactivity (i.e. a single atom decaying), the flask is shattered, releasing the poison, which kills the cat. The Copenhagen interpretation of quantum mechanics implies that after a while, the cat is *simultaneously* alive *and* dead. Yet, when one looks in the box, one sees the cat *either* alive *or* dead, not both alive *and* dead. This poses the question of when exactly quantum superposition ends and reality resolves into one possibility or the other.

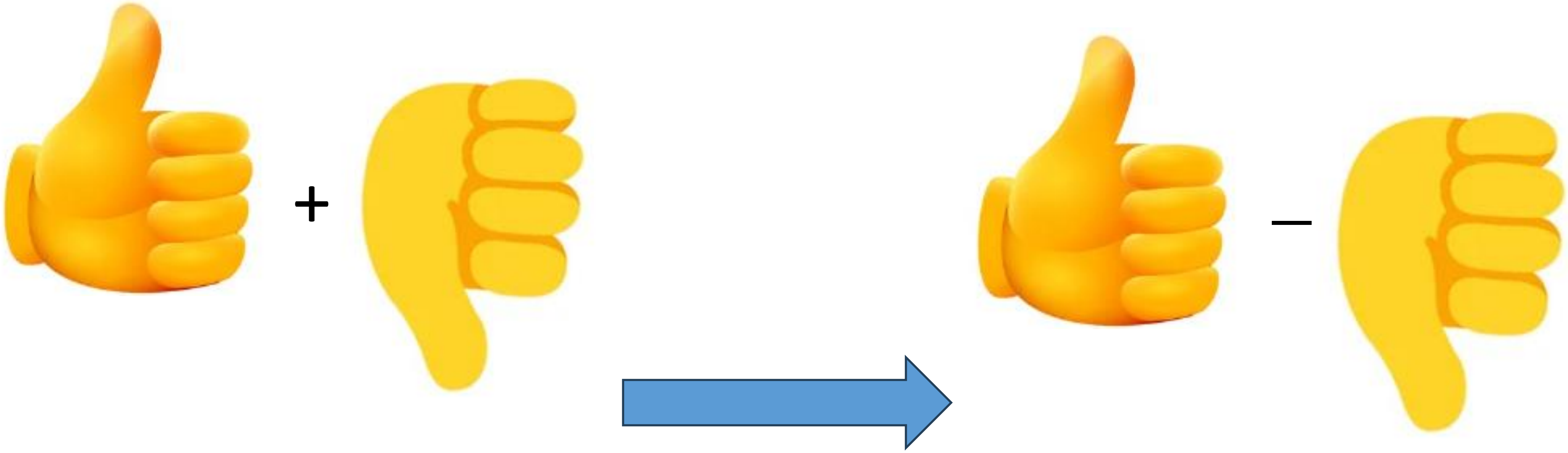
Example: A Quantum Spin

- A quantum spin in a magnetic field B_z of z-direction has two stable states (eigenstates): spin-up state and spin-down state.
- If B_z is positive, the spin-up state is more stable.



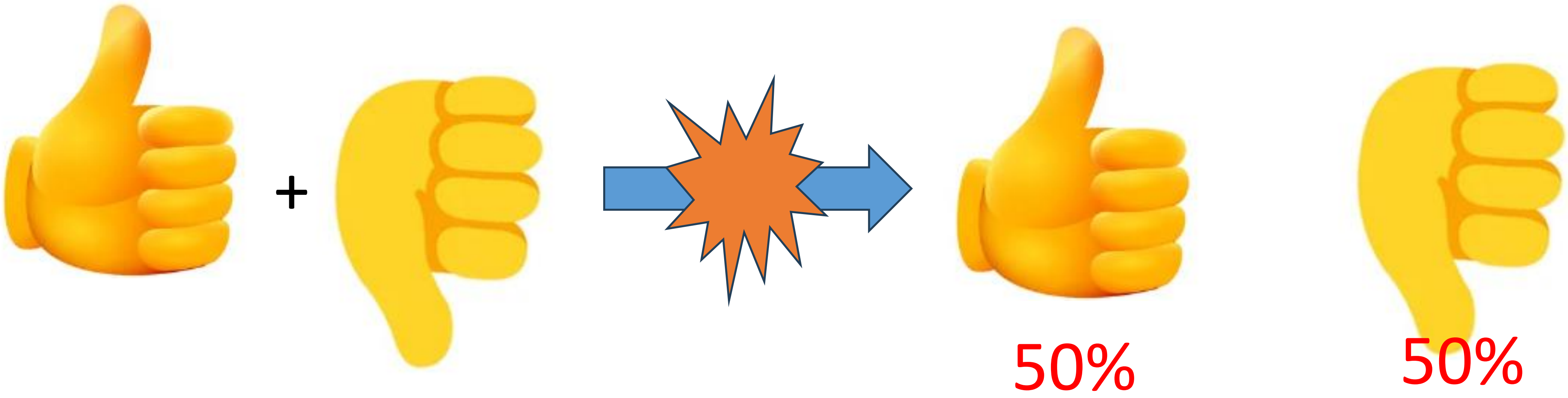
Quantum Spin in a Transverse Field

- When a quantum spin is placed in a transverse magnetic field B_x of x-direction (and measurements will be made in the z-direction subsequently), the pure states  and  are no longer stable. After some time, they degenerate into mixtures of  and  .
- Instead, in the transverse magnetic field, the two stable states (eigenstates) are

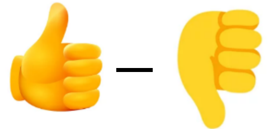


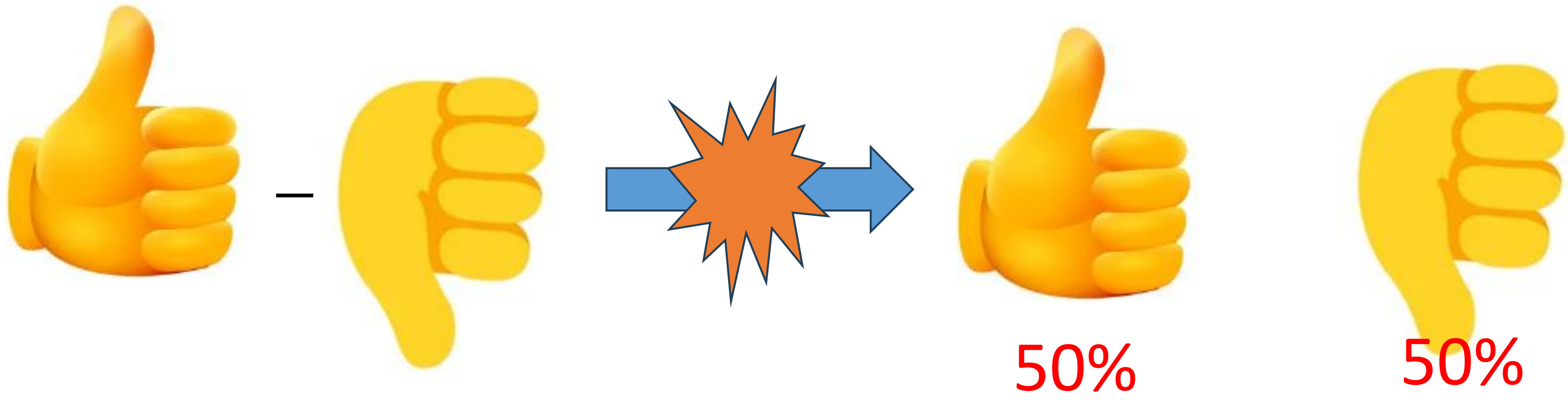
Collapse of Wave Function (1)

- When a measurement is made in the z-direction at time t, the quantum state collapses to the first or second stable state.
- When it collapses to the first stable state  + , there is 50% probability that the measurement result is an up spin, and 50% a down spin.



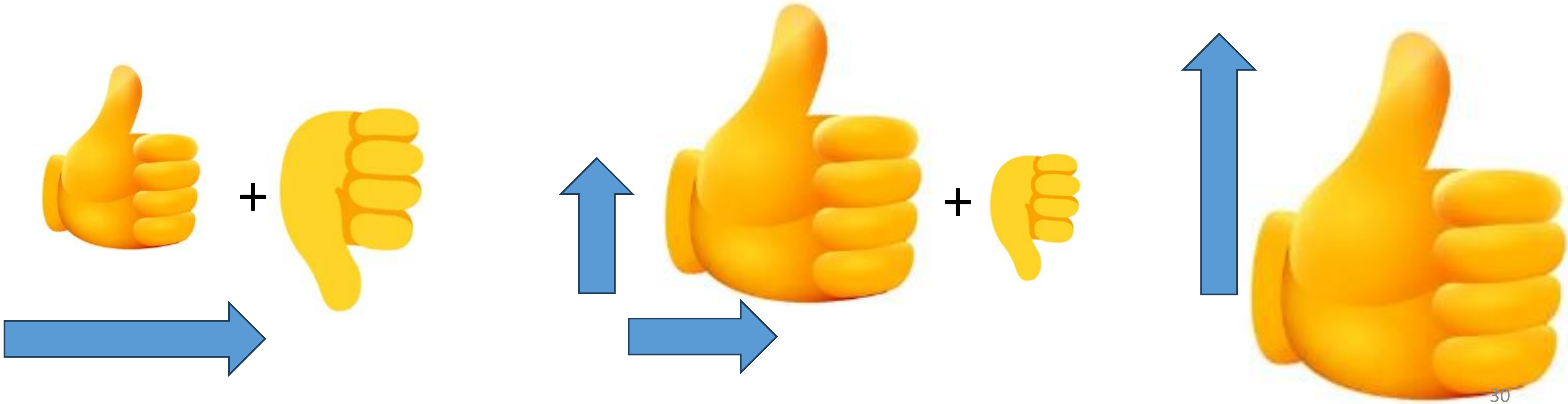
Collapse of Wave Function (2)

- Similarly, when it collapses to the first stable state  , there is 50% probability that the measurement result is an up spin, and 50% a down spin.
- Note that this 50-50 distribution of up and down states resembles the high temperature state of spin glasses.



In a Changing Field

- Suppose the quantum spin is placed in the B_x field, and B_z field is increased gradually.
- The quantum spin starts with the stable state 👍 + 👎. Then it will gradually change into 👍 or 👎 depending on the direction of the B_z field.



10. 4 Quantum Annealing

- Nishimori's team proposed the following quantum annealing scheme:

$$\text{Total Hamiltonian} = \text{Initial Hamiltonian} + \text{Final Hamiltonian}$$

(In QM, Hamiltonians are the same as energy functions.)

- The initial Hamiltonian is the energy of all spins placed in a magnetic field in the x-direction, the field magnitude is a slowly decreasing function $\Gamma(t)$ with time t.
- The final Hamiltonian is the energy of the optimization problem of the spins, with the spins being quantum spins to be measured in the z-direction.

Tadashi Kadowaki and Hidetoshi Nishimori, "Quantum annealing in the transverse Ising model", Phys. Rev. E **58**, 5355-5363 (1998).

The Annealing Process

- Initially, when $\Gamma(t)$ is large, the spins are uniformly distributed in the spin-up and spin-down states. This is similar to the high temperature state in simulated annealing.
- When $\Gamma(t)$ decreases, the final Hamiltonian becomes increasingly important, and the quantum state evolves into the minimum energy state of the final Hamiltonian.
- During the process, the spin states tunnel between the states without the need of hill climbing.

Result: Probability of Finding Ground State

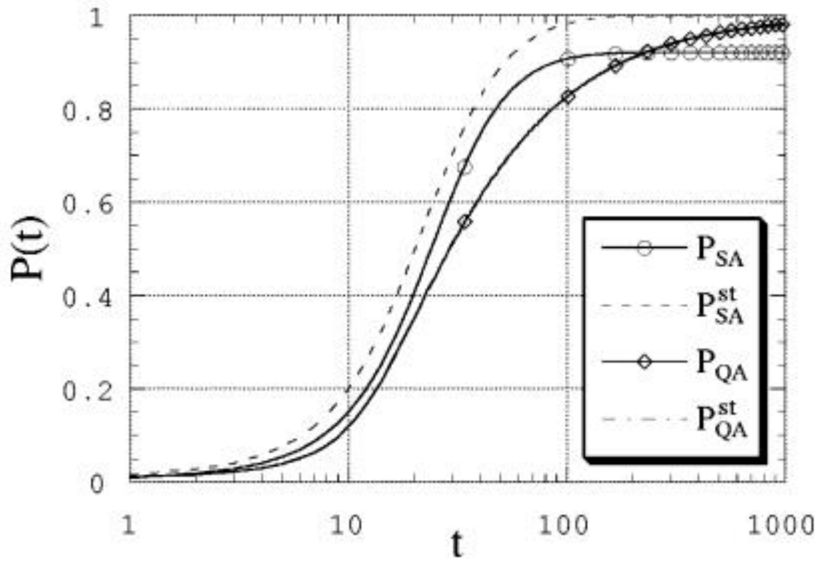


FIG. 2. Time dependence of the overlaps of the ferromagnetic model with $\Gamma(t)=T(t)=3/\sqrt{t}$.

Ferromagnetic model

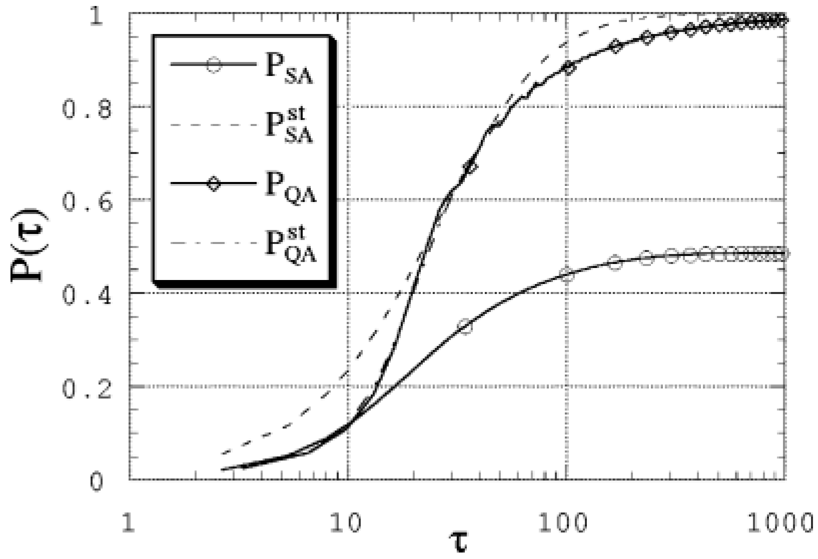


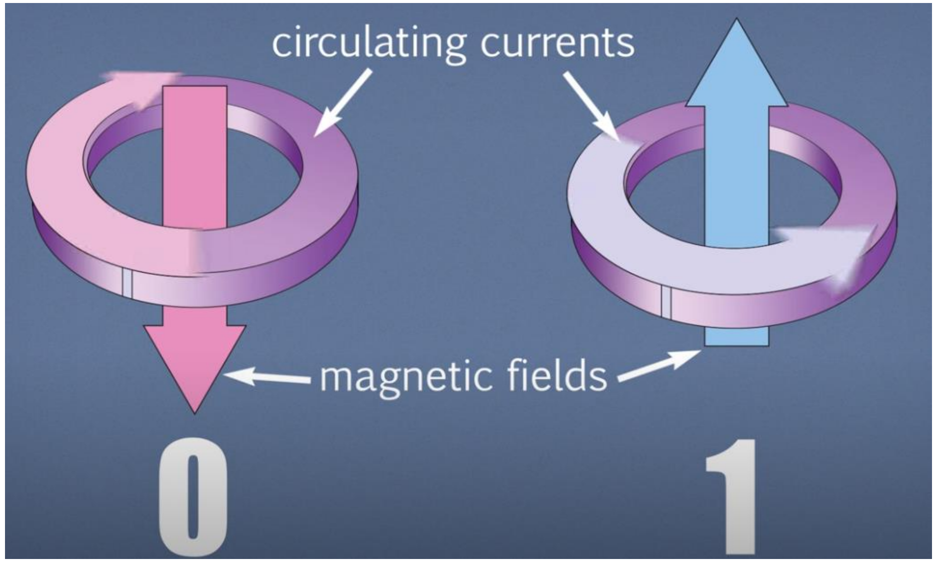
FIG. 7. Time dependence of the overlaps of the frustrated model under $\Gamma(t)=T(t)=3/\sqrt{t}$. Here the time scale τ is normalized by Γ_c and T_c (the points where the correlation functions vanish in Fig. 6).

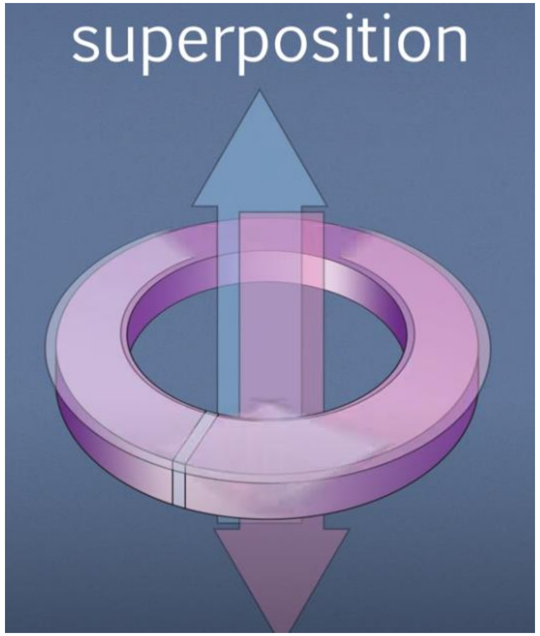
Frustrated model

The probability of finding the ground state in the frustrated model is much higher using quantum tunneling compared with simulated annealing.

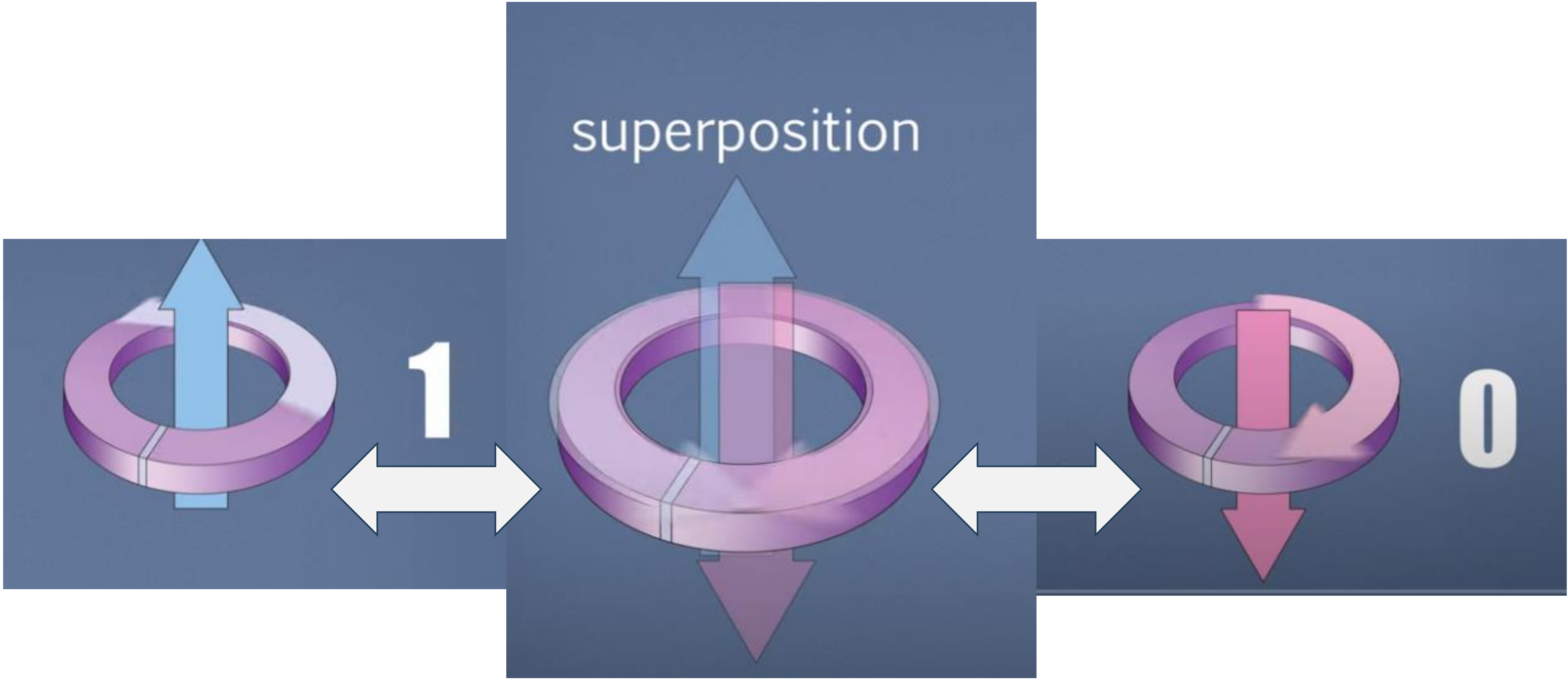
Quantum Annealing in Hardware

- When some metals are cooled to very low temperature, they become superconductors.
- In superconducting rings, the persistent currents can flow in clockwise direction (state 0) or anticlockwise direction (state 1). They correspond to the (right-handed) thumbs down and thumbs up states.
- In superconducting currents, electrons are coupled to form macroscopic quantum states. So they can exist as superpositions of states 0 and 1. Quantum spin is realized.





Quantum Tunneling in $\boldsymbol{B_x}$ Field



D:Wave
The Quantum Computing Company™

- **D-Wave Systems, Inc.** is a *Canadian* quantum computing company, based in Burnaby, British Columbia, Canada.
- The world's first company to sell computers for quantum annealing.
- D-Wave's early customers include:
  - Lockheed Martin,
  - University of Southern California,
  - Google/NASA
  - Los Alamos National Lab
- See their Youtube videos:
  - How the Quantum Annealing Process Works
  - Physics of Quantum Annealing – Hamiltonian and Eigenspectrum (first part)

Other Quantum Annealing Companies

Aliyun (Alibaba Cloud)[13]	July 30, 2015	Computing/Communication[13][14]	Superconducting	Chinese Academy of Sciences[14][15][16]	Hangzhou, China
Baidu[19]	2018	Computing	Algorithms	University of Technology	Beijing, China
Huawei Noah's Ark Lab[42]		Communication		Nanjing University	Shenzhen, China

Japan: Elyah, Fujitsu, Hitachi, Mitsubishi, NEC, NTT, Toshiba
USA, UK, etc

10.5 Quantum-Inspired Computing

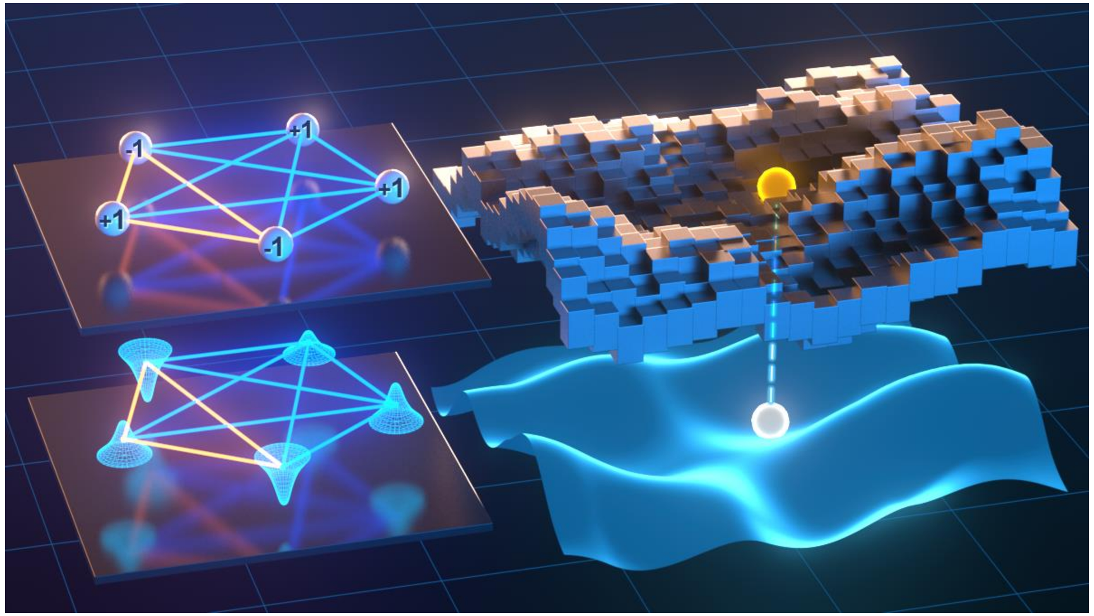
Disadvantages of quantum annealing due to hardware constraints:

- connectivity
- coherence time
- temperature of the qubits

These constraints restrict the size and complexity of the problems that can be solved. Recently, scientists turned to quantum-inspired computing. These machines are based on classical physics, but still they are more scalable and also fast:

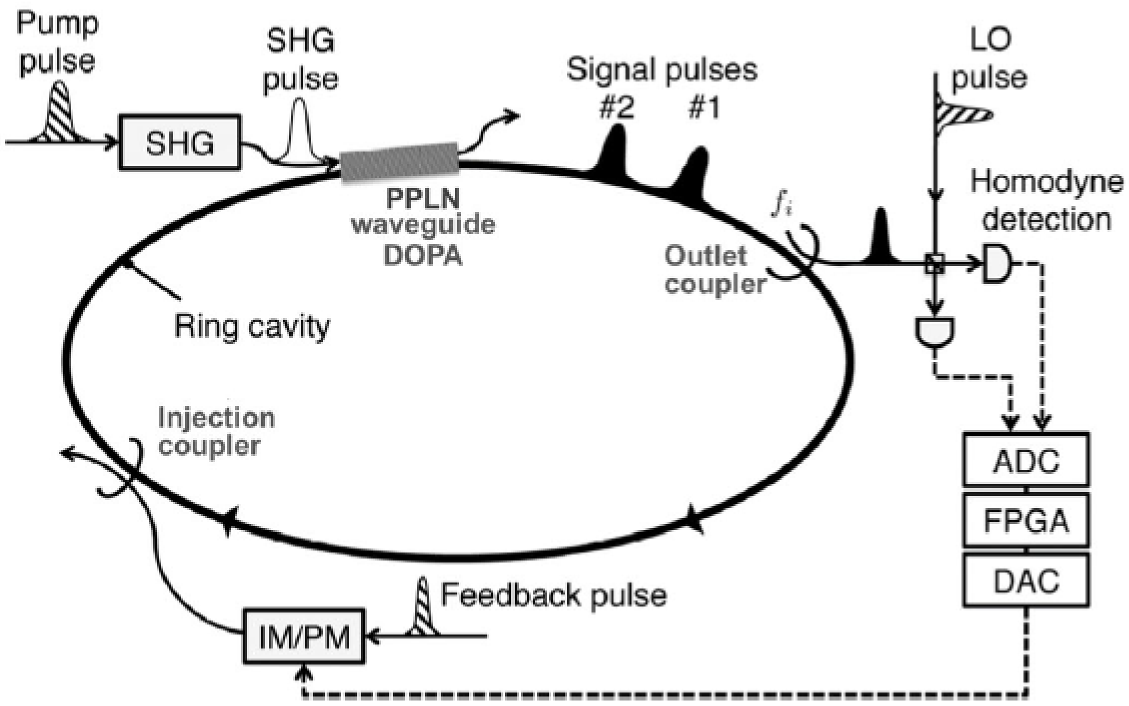
- Coherent Ising machine
- Simulated coherent Ising machine
- Simulated bifurcation (ballistic SB and digital SB)

Continuous Physical Dynamics



- There is a lot of recent interest in using physical dynamics of continuous variables to solve hard combinatorial optimizations of discrete variables.
- Below, we will discuss the coherent Ising machine.

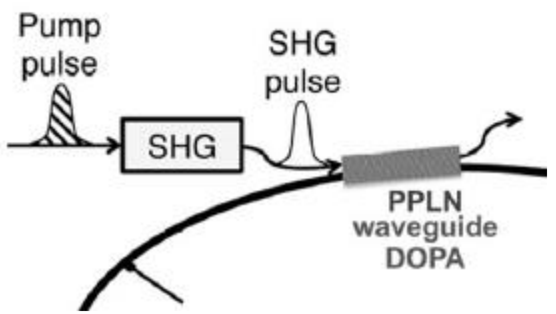
Coherent Ising Machine: Architecture

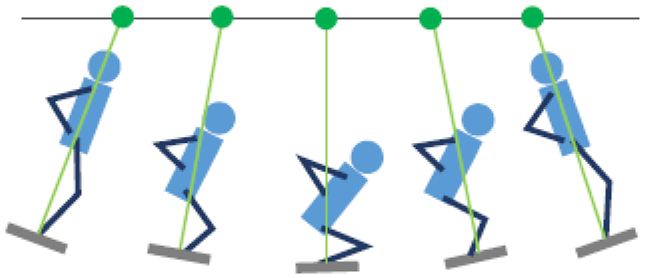


- The main body is called degenerate optical parametric oscillator (DOPO).
- It consists of a loop of optical delay line plus accessories.
- Signal pulses can travel around the loop. Each pulse represents an information bit.

Second Harmonic Generator (SHG)

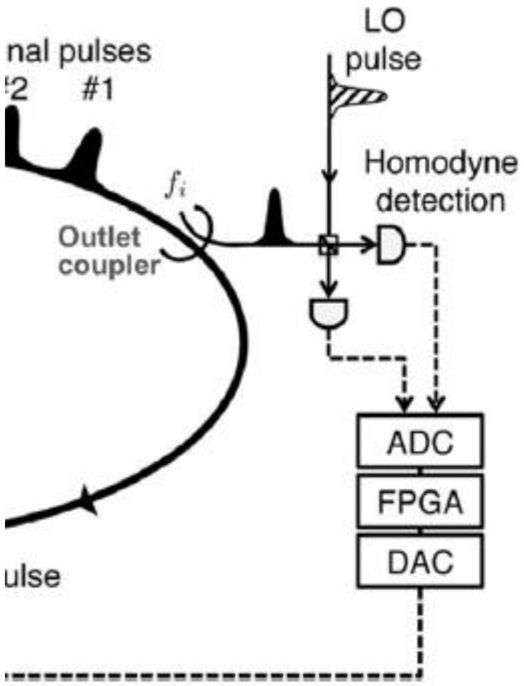
- Optical pulses in the optical delay line is attenuated in the absence of external drives.
- SHG pump energy to the pulses by supplying an input at the frequency 2 times that of the pulse.
- The principle is the same as the amplification process of a swing, in which a person makes a full cycle (up-down-up) during the time a swing makes half a cycle (left-right). This is called parametric excitation.





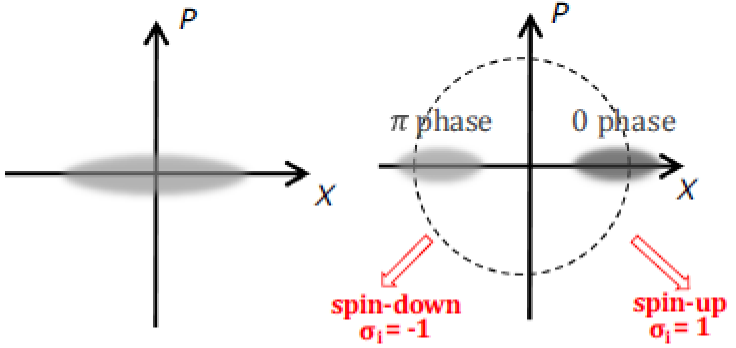
Field-Programmable Gate Array (FPGA)

- A small portion of each DOPO pulse is out-coupled through the output coupler.
- The analog-to-digital converter (ADC) converts the signal to be processed by FPGA.
- FPGA combines the signals weighted by the couplings J_{ij}.
- The digital-to-analog converter (DAC) converts the FPGA outputs $\sum_j J_{ij} S_j$ to be fed back to DOPO.

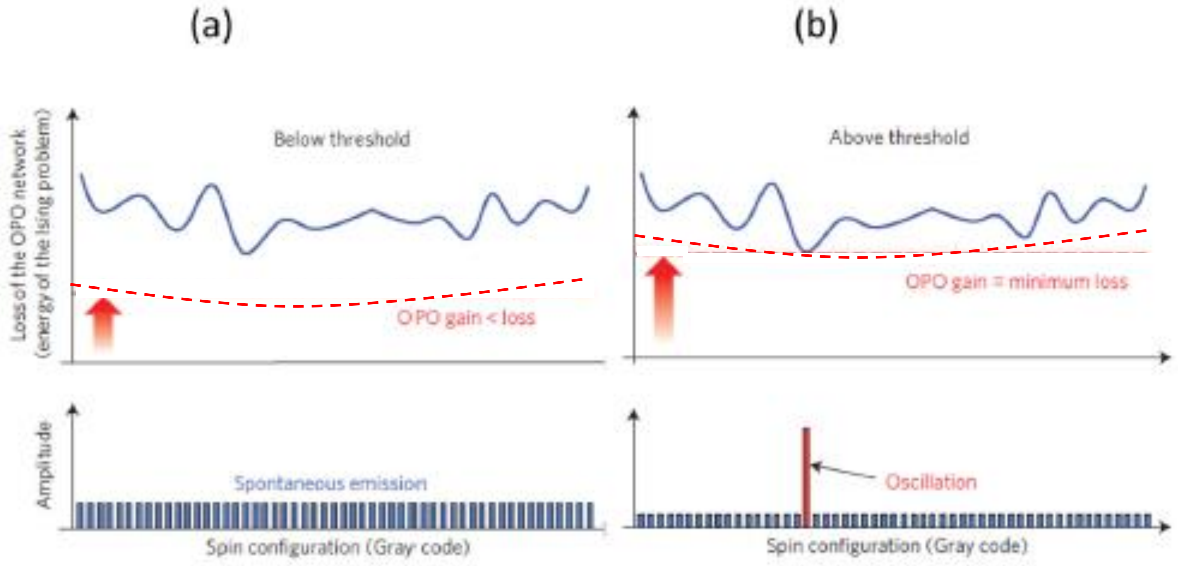


The Pulses

- In quantum optics, the pulses have two components, commonly known as c and s (denoted as x and p in the figure below).
- In DOPO, the pumping amplifies the c field and suppresses the s field.
- When the pump rate is below the threshold, the c field is evenly distributed in the positive and negative values (similar to the paramagnetic phase in combinatorial optimization, or the superposition state in quantum annealing).
- Above the threshold, the c field is centered at 0 phase (equivalent to spin-up state) or π phase (equivalent to spin-down state).



Principle and Sequence of Operations



- The pump rate is gradually increased. Below the threshold the c fields vanish.
- When the pump rate reaches the threshold, it reaches the minimum loss rate of a ground state. A single-mode oscillation occurs at the ground state.

Dynamical Equations

Assuming that the *s* fields are suppressed, the dynamics of the c fields become:

$$\frac{dc_i}{dt} = (p-1)c_i + \xi \sum_{i<j} J_{ij} c_j - c_i^3 + \text{noise.}$$

pump rate   decay rate   interactions   nonlinear response

This is equivalent to the gradient descent of the Hamiltonian

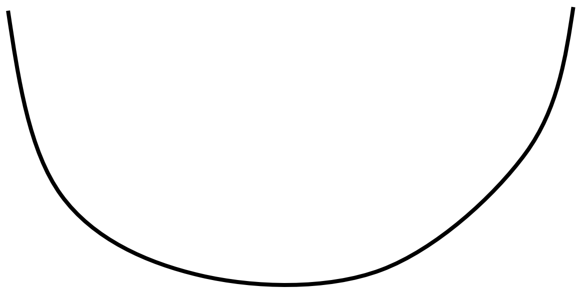
$$H = \sum_i \left[\frac{c_i^4}{4} + (1-p)\frac{c_i^2}{2} \right] - \xi \sum_{i<j} J_{ij} c_j c_j \,.$$

Note that the quadratic coefficient is positive at low pump rate, and negative at high pump rate.

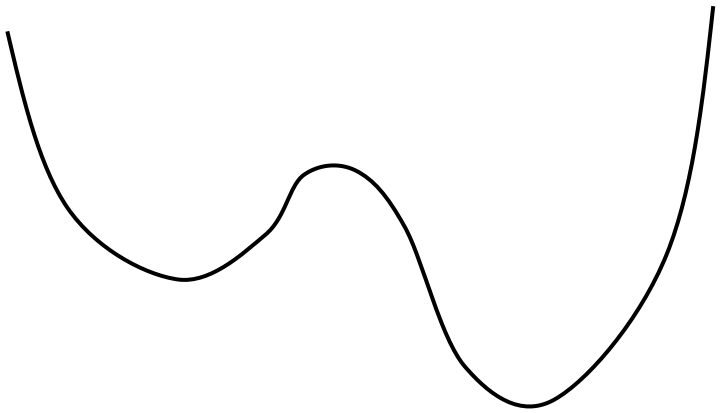
Energy Landscape

The effective energy landscape for a single pulse:

At low pump rate: the quadratic coefficient is positive.



At high pump rate: the quadratic coefficient is negative.



Hence, increasing the pump rate is an annealing process.

Two-Pulse Example

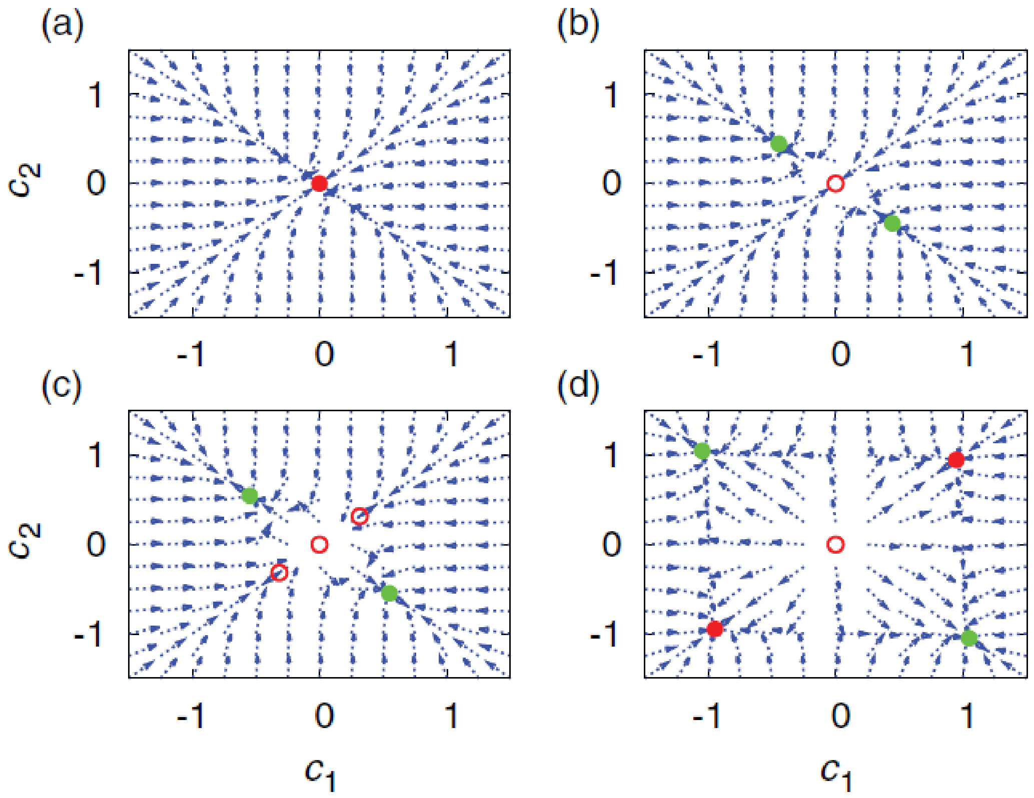
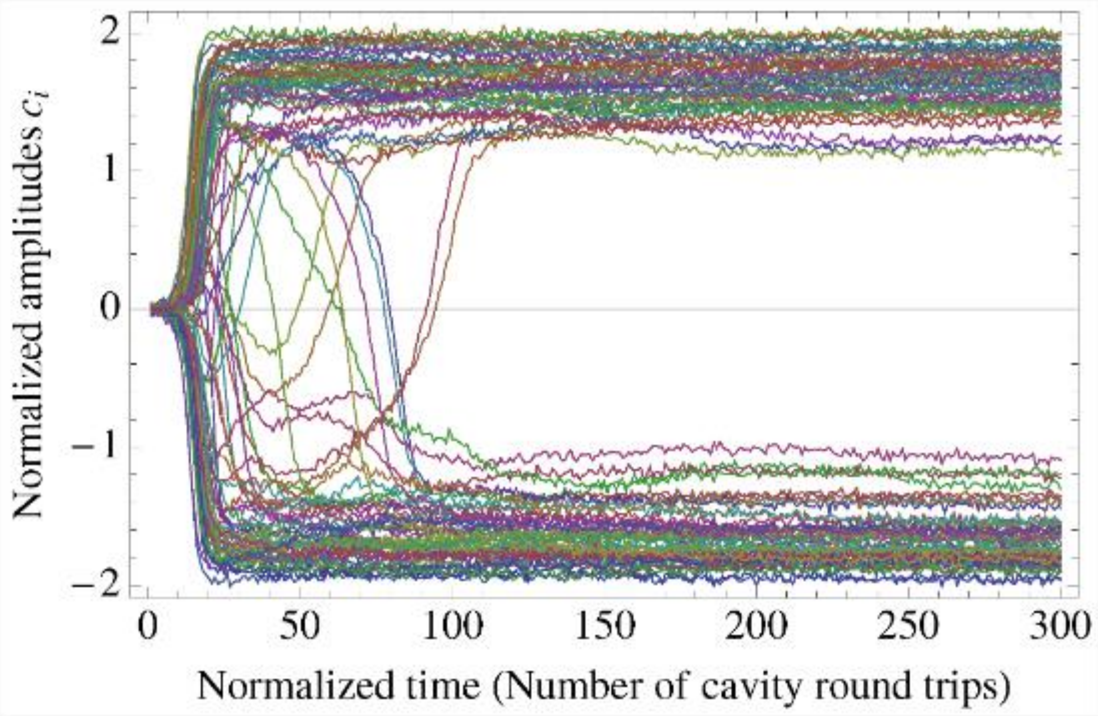


FIG. 2. (Color online) Phase diagrams for the in-phase components of two coupled degenerate OPOs when $\xi = -0.1$. The normalized pump rate for each diagram is (a) $p = p_{th} = 0.9$, (b) $p = 1.1$, (c) $p = 1.2$, and (d) $p = 2.0$. The dots and circles mean stable and unstable steady states, while the colors green (light gray) and red (medium gray) denote correct and incorrect solutions, respectively.

Wang et al, Phys Rev A **88**, 063853 (2013).

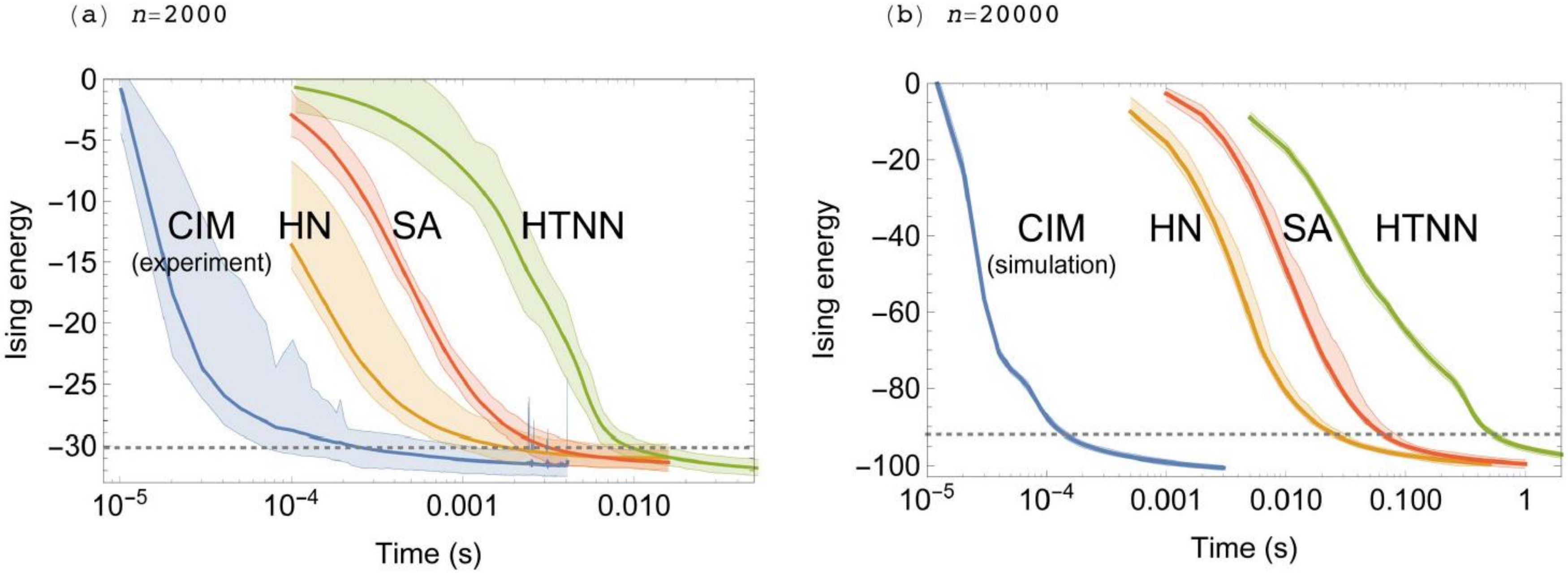
N = 800 Example

Figure 4. Normalized DOPO signal amplitudes when CIM is solving $N = 800$ MAX-CUT is shown. Each trajectory describes a DOPO.



Haribara et al, Entropy **18**, 151 (2016).

Performance evaluation: CIM vs classical neural networks



Haribara et al, Quantum Sci. Technol. 2 (2017) 044002

Figure 3. Energy descent when solving $\{+1, -1\}$-weighted (a) $n = 2000$ and (b) $n = 20\,000$ complete graphs. Each thick line is the ensemble average of 100 trials (except for CIM experiment, which consists of 26 trials), while the lower and upper shaded error bars show the best and worst envelopes for each computational model. The gray dotted line is the target values $-60\,278/n$ and $-1841\,216/n$ for $n = 2000$ and 20 000, respectively, which are obtained by the SDP relaxation algorithm [7]. In the CIM simulation of $n = 20\,000$, the cavity round trip time is assumed to be 10 μs.

Test on MAX-CUT: Hopfield-Tank neural network (HTNN), Hopfield network (HN), simulated annealing (SA)

Tunneling and Entanglement

Quantum tunneling is an effective approach to optimization problems.

However, the quantum nature of the quantum bits (qubits) has not been utilized. These qubits exist in entangled states. For example, the previously mentioned states are entangled:

👍 + 👎 and 👍 – 👎

Manipulating the entangled states is the core of quantum computation nowadays. This will be the topic of the next two lectures by Prof Yi Wang.

Homework (due 11:59 pm, 26 Nov 2023)

Search the website of one of the following companies and describe one application of quantum annealing (200 words or more, on your own words):

- D-Wave Systems
- Toshiba Simulated Bifurcation Machine (Advanced Cases)