Entropy in Quantum Information Theory and Condensed Matter Physics

M. B. Hastings
Microsoft Research
Station Q, Santa Barbara

Thanks: P. Hayden, J. Yard, F. Verstraete, S. Bravyi, J. Eisert, L. S. Levitov

Steam engines in the early 1800s

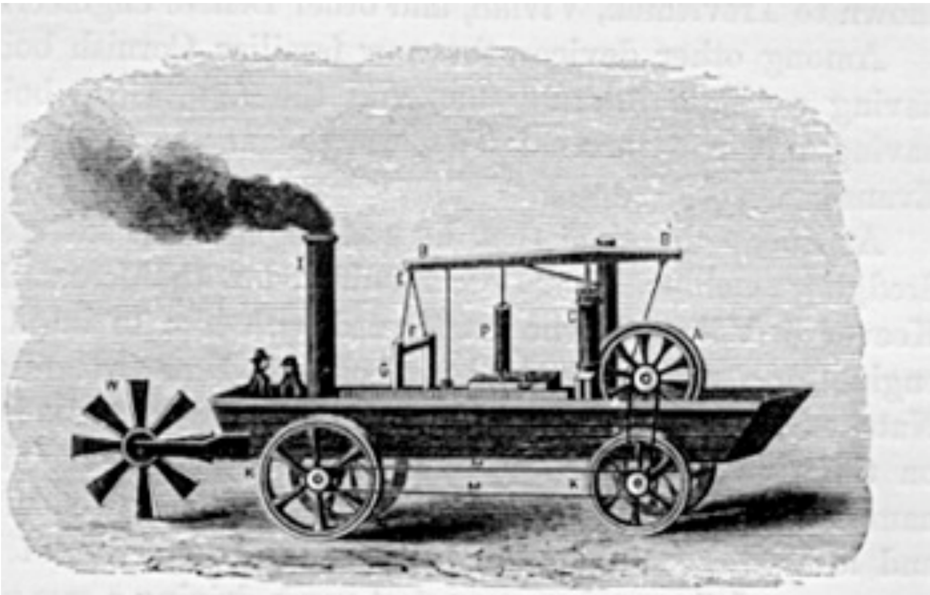


FIG. 47.—Evans's "Oruktor Amphibolis," 1804.

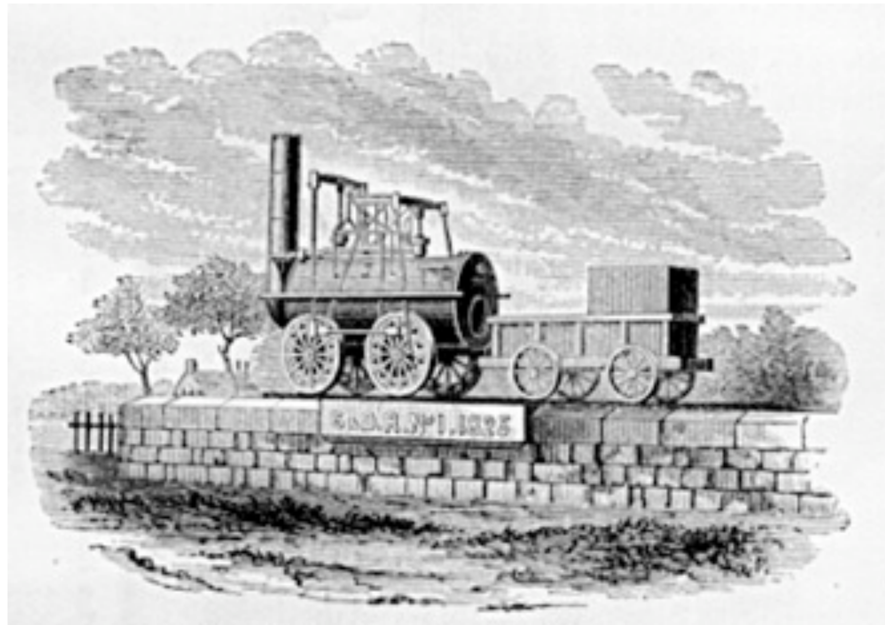


FIG. 52.—Stephenson's No. 1 Engine, 1825.

-from an 1878 book by Thurston

Carnot cycle

How efficient can a heat engine be? ("Reflections on the Motive Power of Fire", 1824)

$$T_H \xrightarrow{Q_H} \boxed{\text{Engine}} \xrightarrow{Q_C} T_C$$

$$\Big\downarrow$$

Work = $Q_H - Q_C$

$$\mathrm{d}Q = T\,\mathrm{d}S \qquad\qquad \mathrm{d}S \geq 0$$

Theoretical limit for efficiency: $1 - \dfrac{T_C}{T_H}$

- Entropy in Quantum Information Theory
    - Classical Information Theory Background
    - Communicating over a Quantum Channel
    - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
    - Entropy from quantum entanglement
    - Thermalization
    - Relaxation of positional order

Classical Information Theory
(from Boltzmann, 1877 to Shannon, 1948)

A sequence of 20 random bits: 00100111010010100111

A sequence of 20 1s: 11111111111111111111

$$S = -\sum_x p(x) \log_2(p(x))$$

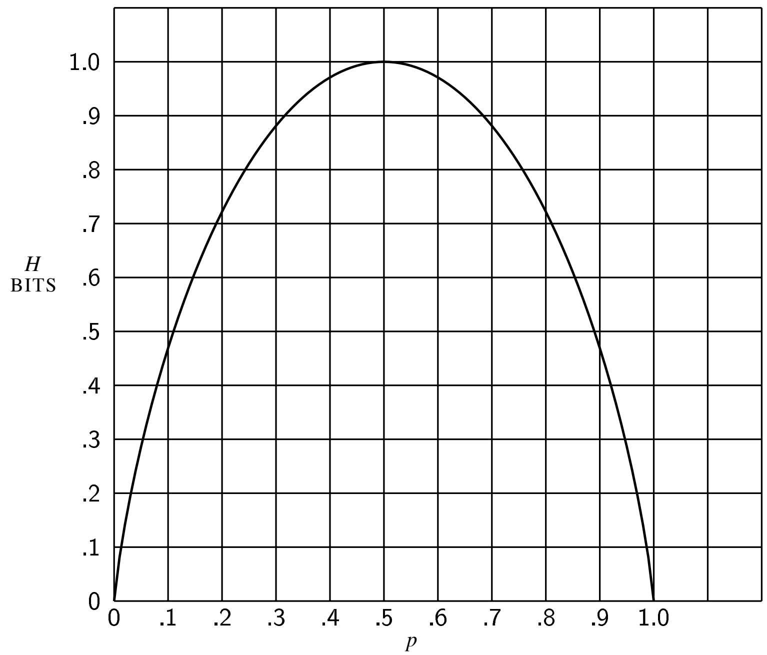


Fig. 7 — Entropy in the case of two possibilities with probabilities p and $(1-p)$.

A low entropy source can be compressed: .gz, .zip, ...

Communicating over a noisy channel:

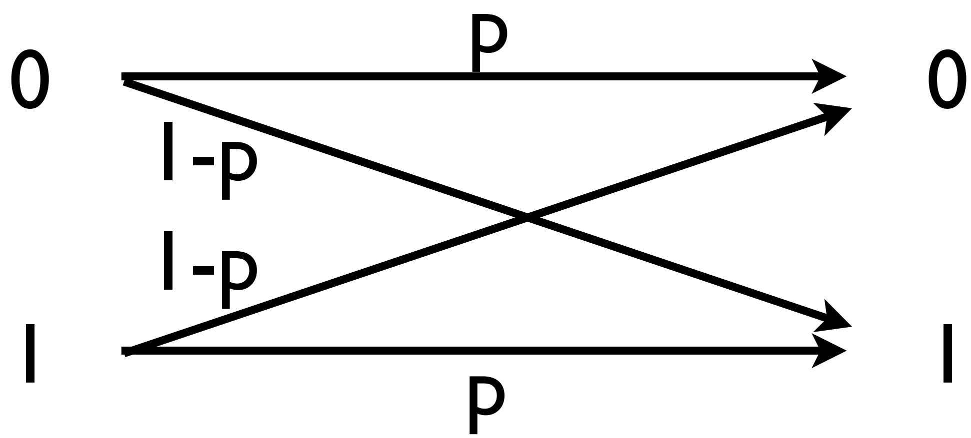
Alice $\longrightarrow$ Bob

input X output Y

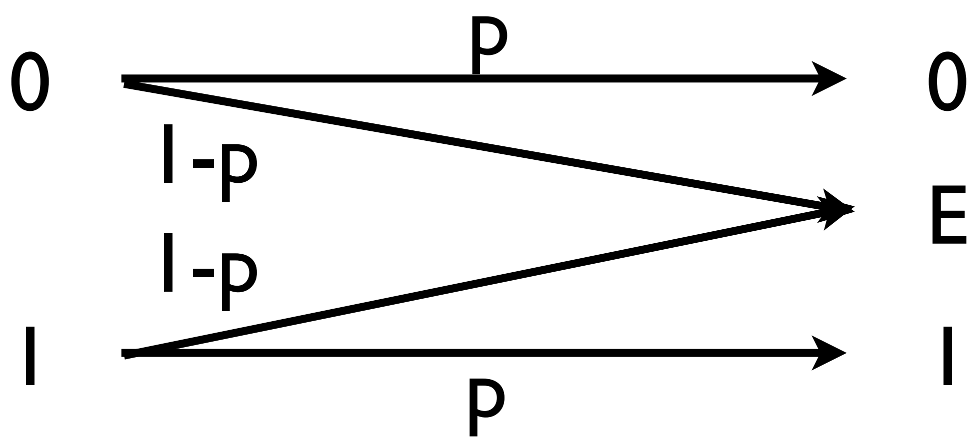
Channel defined by allowed inputs and outputs
and by probability P(Y|X)

Communicating over a noisy channel (examples):

Binary symmetric channel:



Binary erasure channel:



Meaning of entropy (Shannon's noisy channel coding theorem):

Error correction:

k bits — N>k bits — N bits (with errors)

10....1 → Encode → 11.....1 → Channel → 01.....1

k bits

01.....1 → Decode → 10....1

A simple code (repetition code):

1 → 11.....1 → Channel → Pick majority

For p>1/2, error probability exponentially small in N, but we encode at rate k/N=1/N, so rate->0

Communicating over a noisy channel: the capacity

Choose inputs with probability $p(X)$

Output: $p(Y) = \sum_X P(Y|X)p(X)$

conditional information

The capacity: $\max_{p(X)} S(B) - S(B|A)$

$$S(B) = -\sum_Y p(Y)\log_2(p(Y))$$

$$S(B|A) = -\sum_X p(X)\left[\sum_Y P(Y|X)\log_2(P(Y|X))\right]$$

How much noise is output, minus how much noise is due to the channel, equals the information transmitted.

Meaning of entropy (Shannon's noisy channel coding theorem):

Shannon '48

We can encode k bits into N bits, such that the error probability goes to zero as N goes to infinity, with k/N asymptotically approaching C, the capacity of the channel.

This gives a meaning to the capacity of the channel.

Amazing things about channel coding:

- C is not zero!
- C is actually quite large. C for the erasure channel is equal to p.
- We can calculate C.
- We do the calculation by a **single-letter** formula, despite using correlations to correct errors.

Additivity of capacities.

- Entropy in Quantum Information Theory
  - Classical Information Theory Background
  - Communicating over a Quantum Channel
  - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
  - Entropy from quantum entanglement
  - Thermalization
  - Relaxation of positional order

Communicating over a noisy quantum channel:

Channel is a linear map on density matrices.

Alice $\longrightarrow$ Bob

input ρ output $\mathcal{E}(\rho)$

$$\mathcal{E}(\rho) = \sum_{s=1}^{D} A(s)\rho A^\dagger(s) \qquad \sum_{s=1}^{D} A^\dagger(s)A(s) = I$$

Communicating over a noisy quantum channel:

Quantum entropy: $H(\rho) = -\mathrm{tr}\Big(\rho \log_2(\rho)\Big)$

Signal words: input state ρ_i with probability p_i

$$\chi(\mathcal{E}, \{p_i, \rho_i\}) = H\Big(\mathcal{E}(\sum_i p_i \rho_i)\Big) - \sum_i p_i H\Big(\mathcal{E}(\rho_i)\Big)$$

$$\underset{\text{recall S(B)}}{\uparrow} \qquad\qquad \underset{\text{recall S(B|A)}}{\uparrow}$$

Holevo capacity for sending classical information over a quantum channel:

$$\chi_{\max}(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \chi(\mathcal{E}, \{p_i, \rho_i\})$$

"I wish that physicists would
... give us a general expression for the
capacity of a channel with quantum
effects taken into account rather than a
number of special cases."

-J. R. Pierce, 1973, in a retrospective
on Shannon's paper.

Communicating over a noisy quantum channel:

Why the Holevo capacity is hard to evaluate: should we entangle?

$$\chi_{\max}(\mathcal{E}^{\otimes n}) \stackrel{?}{=} n\chi_{\max}(\mathcal{E})$$

Additive: $\chi_{\max}(\mathcal{E}_1 \otimes \mathcal{E}_2) = \chi_{\max}(\mathcal{E}_1) + \chi_{\max}(\mathcal{E}_2)$

Non-Additive: $\chi_{\max}(\mathcal{E}_1 \otimes \mathcal{E}_2) > \chi_{\max}(\mathcal{E}_1) + \chi_{\max}(\mathcal{E}_2)$

The additivity conjecture: the first case is true for all quantum channels.

Equivalence of additivity conjectures (Shor, 2004):

- Additivity of Holevo capacity
- Additivity of minimum output entropy
- Additivity of entanglement of formation
- Strong super-additivity of entanglement of formation

Why Additivity Is Important:

- We can boost capacity using entangled inputs.
- If additivity fails, then we physicists have not answered Pierce's question. It is not practical to compute the capacity maximizing over arbitrary entangled inputs.
- Additivity in the classical case gives meaning to the capacity of a channel.

"On the other hand, there are the big challenges, like the additivity problems ..... Probably every quantum information theorist worth his salt has had a go on that one. Such challenges are landmarks in any field. If you can make serious progress on one of them, you know you have really moved."

-R. F. Werner, in a 2005 review of open problems in quantum information theory

- Entropy in Quantum Information Theory
  - Classical Information Theory Background
  - Communicating over a Quantum Channel
  - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
  - Entropy from quantum entanglement
  - Thermalization
  - Relaxation of positional order

The minimum output entropy conjecture:

$$H^{\min}(\mathcal{E}) = \min_{|\psi\rangle} H(\mathcal{E}(|\psi\rangle\langle\psi|))$$

$$H^{\min}(\mathcal{E}_1 \otimes \mathcal{E}_2) \stackrel{?}{=} H^{\min}(\mathcal{E}_1) + H^{\min}(\mathcal{E}_2)$$

Relation to additivity of Holevo capacity: by reducing the output entropy for a given input state, we can communicate more effectively over the channel.

A Counterexample to Additivity!

MBH '08

Two random channels, related by complex conjugation:

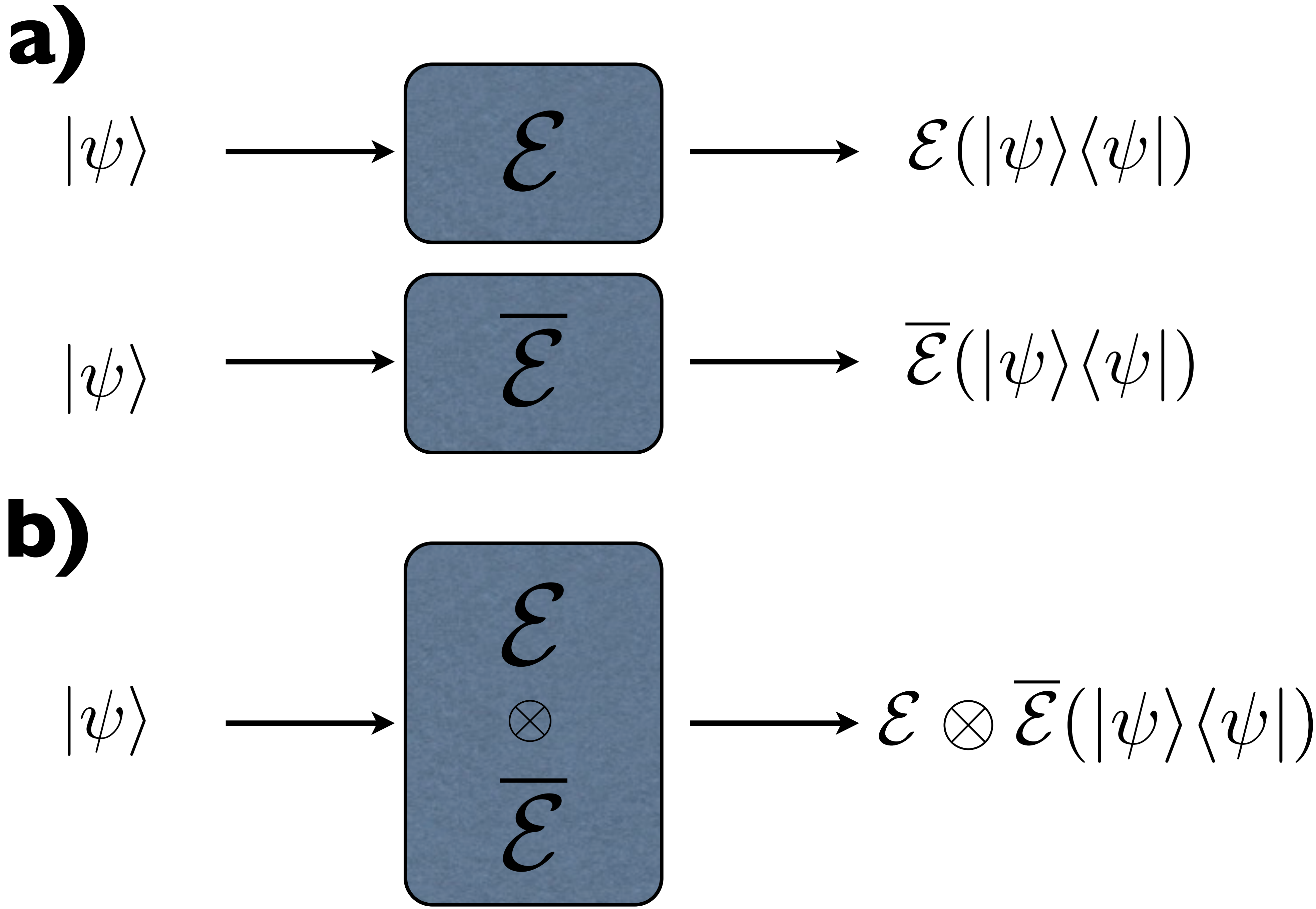
$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i^\dagger \rho U_i$$

$$\overline{\mathcal{E}}(\rho) = \sum_{i=1}^{D} p_i \overline{U}_i^\dagger \rho \overline{U}_i$$

U_i are randomly chosen unitaries.

$1 << D << N$

This channel models interaction with a random environment.


a)
$|\psi\rangle$
$\mathcal{E}$
$\mathcal{E}(|\psi\rangle\langle\psi|)$
$|\psi\rangle$
$\overline{\mathcal{E}}$
$\overline{\mathcal{E}}(|\psi\rangle\langle\psi|)$
b)
$|\psi\rangle$
$\mathcal{E}$
$\otimes$
$\overline{\mathcal{E}}$
$\mathcal{E}\otimes\overline{\mathcal{E}}(|\psi\rangle\langle\psi|)$

Why additivity fails:

There is a low entropy input state for the combined channel:

$$|\phi\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} |\alpha\rangle \otimes |\alpha\rangle$$

Note that:

$$U_i^\dagger \otimes \overline{U}_i^\dagger |\phi\rangle = |\phi\rangle$$

So, of the D^2 possible outputs, D of them are the same, when we choose the same unitary for both channels.

$$\begin{aligned} H(\mathcal{E} \otimes \overline{\mathcal{E}}(\phi)) &= \frac{1}{D} \log_2(D) + (1 - 1/D) \log_2(D^2) \\ &= 2 \log_2(D) - \log_2(D)/D \end{aligned}$$

However, for most such channels,

$$H^{\min}(\mathcal{E}) \geq \log_2(D) - \mathrm{const.}/D - O(\sqrt{\ln(N)/N})$$

(proof based on randomness)

Experimental relevance:

- Currently, it is too difficult to manipulate entangled states to expect any practical boost in capacity for any channel.
- However, we may be able to check that certain entangled states decohere less than unentangled states.
- Check simpler claim: that entangled state is more likely to remain unchanged after interacting with environment.
- Need to create large number of entangled pairs (N>>1), and interact in a non-linear way with environment.

The future of additivity?

- Pierce's channel capacity problem remains open.
- I conjecture additivity for channels of the form $\mathcal{E} = \mathcal{F} \otimes \overline{\mathcal{F}}$, giving a **two-letter** formula to solve the capacity problem.
- Can we use these ideas to protect states from decoherence?

- Entropy in Quantum Information Theory
  - Classical Information Theory Background
  - Communicating over a Quantum Channel
  - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
  - Entropy from quantum entanglement
  - Thermalization
  - Relaxation of positional order

Pure states can look locally mixed:

Singlet state:

$$\Psi = |\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle$$

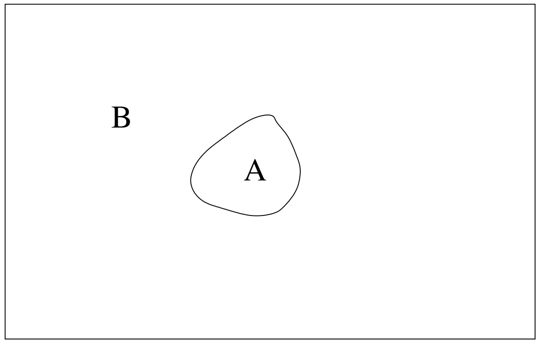
Reduced density matrix:

$$\rho_1 = \frac{1}{2}\Big(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\Big)$$

Each spin looks completely random on its own!

Pure states can look locally mixed (general bipartite system):

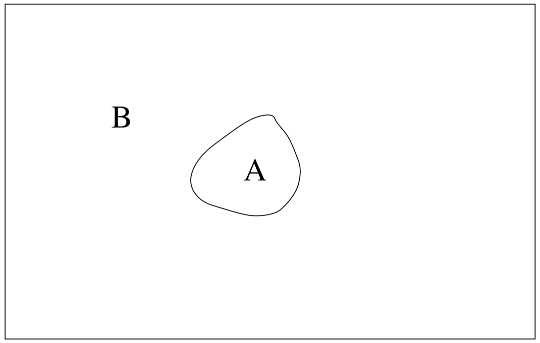
Break space into two regions, A and B



Reduced density matrix: $\rho_A = \mathrm{Tr}_B(\rho)$

For a pure state we can quantify entanglement by the entropy: $H(\rho_A) = H(\rho_B)$

Pure states can look locally mixed:



$$|\Psi_0\rangle = \sum_\gamma A^0(\gamma)|\Psi_A(\gamma)\rangle \otimes |\Psi_B(\gamma)\rangle$$

$$H = -\sum_\gamma |A(\gamma)|^2 \log_2(|A(\gamma)|^2)$$

Singlet has one qubit of entropy.

Relation between entanglement entropy and simulation:

States with low entropy can be simulated efficiently using matrix product states (DMRG by White, TEBD by Vidal, etc...)

F. Verstraete and J. I. Cirac, PRB 73, 094423 (2006).

MBH, JSTAT, 2007

Cost is exponential in entropy.

Most systems in ground state obey "area law", entropy proportional to area, easier to simulate.

- Entropy in Quantum Information Theory
  - Classical Information Theory Background
  - Communicating over a Quantum Channel
  - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
  - Entropy from quantum entanglement
  - Thermalization
  - Relaxation of positional order

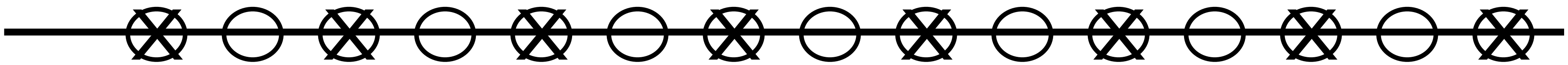
Thermalization:

1. Start with a system in its ground state
2. Make a sudden change in the Hamiltonian
3. Let it evolve without interacting with the environment.

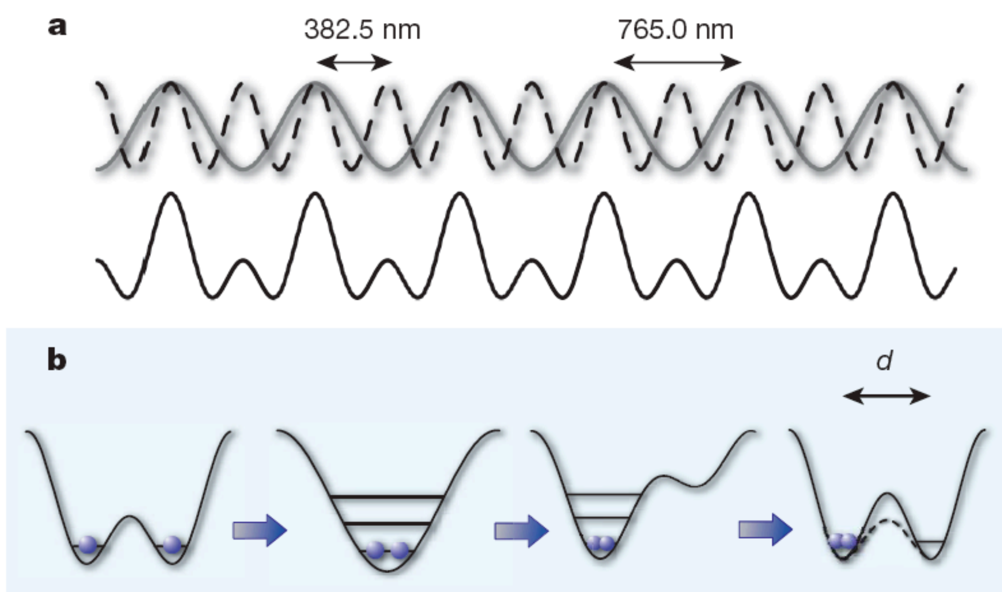
Evolution is unitary, so how does it thermalize?

Proposed experiment (Cramer, Flesch, McCulloch, Schollwock, Eisert, Bloch)

- Prepare atoms in periodic optical lattice
- Add period-2 potential to set up initial alternating particle/ hole/particle/hole/... configuration
- Return to a symmetric potential and let it evolve
- Measure occupation of even and odd sites (order parameter)



1. Bloch's group has shown it is possible to produce these superlattices (Nature '07)



Thermalization:

Evolution is unitary, so how does it thermalize?

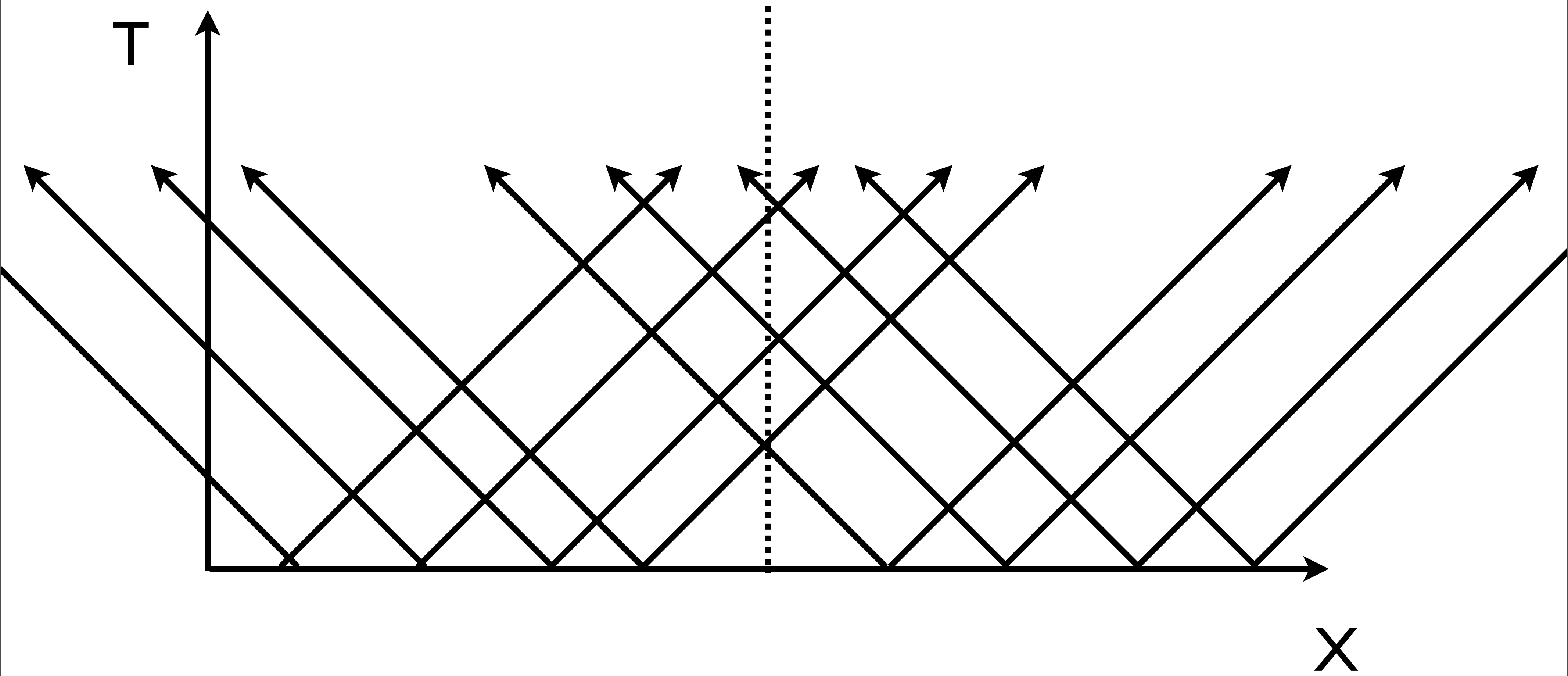
Resolution of the puzzle: if the system is big, each small portion of the system becomes entangled with the rest. This means that the reduced density matrix on part of the system can have non-zero entropy. However, if it is thermal then the entropy is proportional to the volume, not the area.

$$H \sim V$$

$$H \leq \text{const.} \times At$$

$$t_{thermal} \geq V/A$$

Bravyi, MBH, Verstraete, PRL 06,
Eisert, Osborne PRL 06



After global change of Hamiltonian, the system is in a very excited state of the new Hamiltonian. There are quasi-particle excitations everywhere, which carry entropy across the cut.

Worst case: entropy grows linearly in time.

- Entropy in Quantum Information Theory
    - Classical Information Theory Background
    - Communicating over a Quantum Channel
    - (Non)-Additivity
- Entanglement Entropy in Non-equilibrium Quantum Systems
    - Entropy from quantum entanglement
    - Thermalization
    - Relaxation of positional order

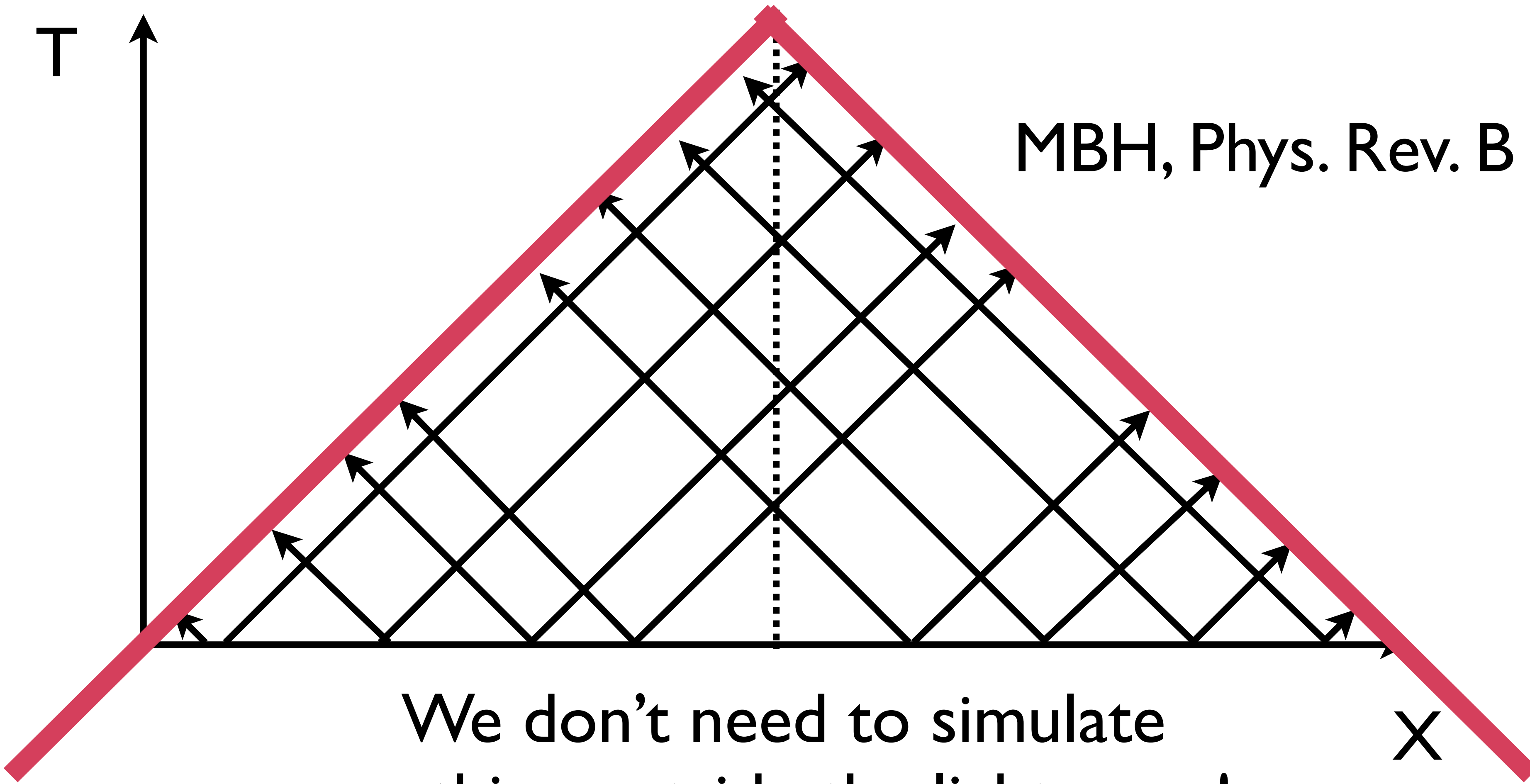
Difficult to simulate far from equilibrium:

Entropy scales linearly in time, so matrix product methods require exponentially growing effort in time.

Brute force simulation of a system of size N requires effort scaling exponentially in N.

Another approach: light-cone (MBH '08), still exponential in time, but a lower exponential.

Light-cone methods



MBH, Phys. Rev. B 08

We don't need to simulate anything outside the light-cone!

Early time dynamics is easier because less entangled.
Result: double the time that can be simulated!

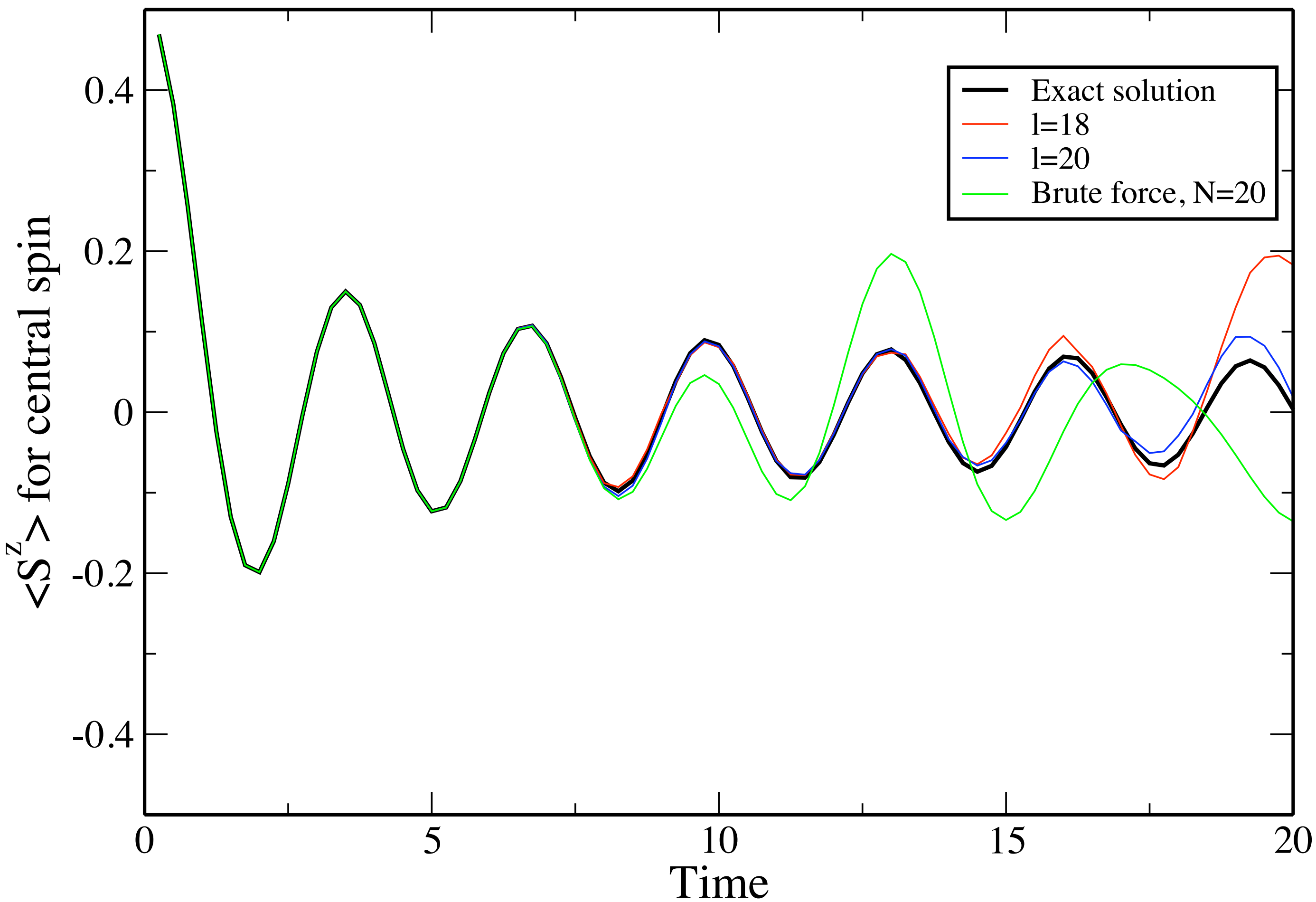
Doubling the time that can be simulated would otherwise require huge increase in resources, due to exponential increase in complexity of problem.

Light cone method is 10^3 times faster, 10^6 times less memory used for this problem size.

$$H = \sum_i S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z$$

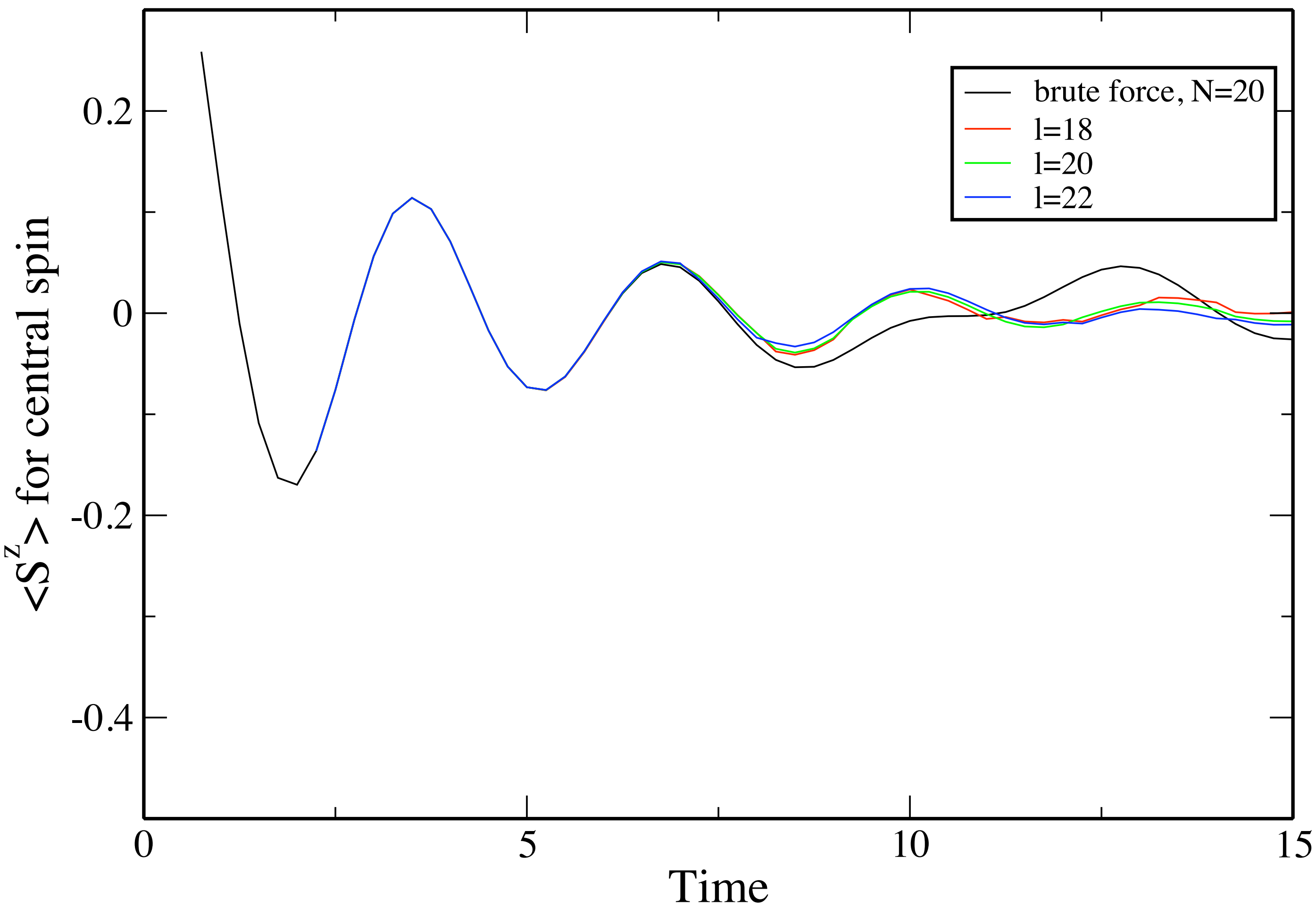
Quench from infinite Delta to finite Delta

Delta=0, free fermions



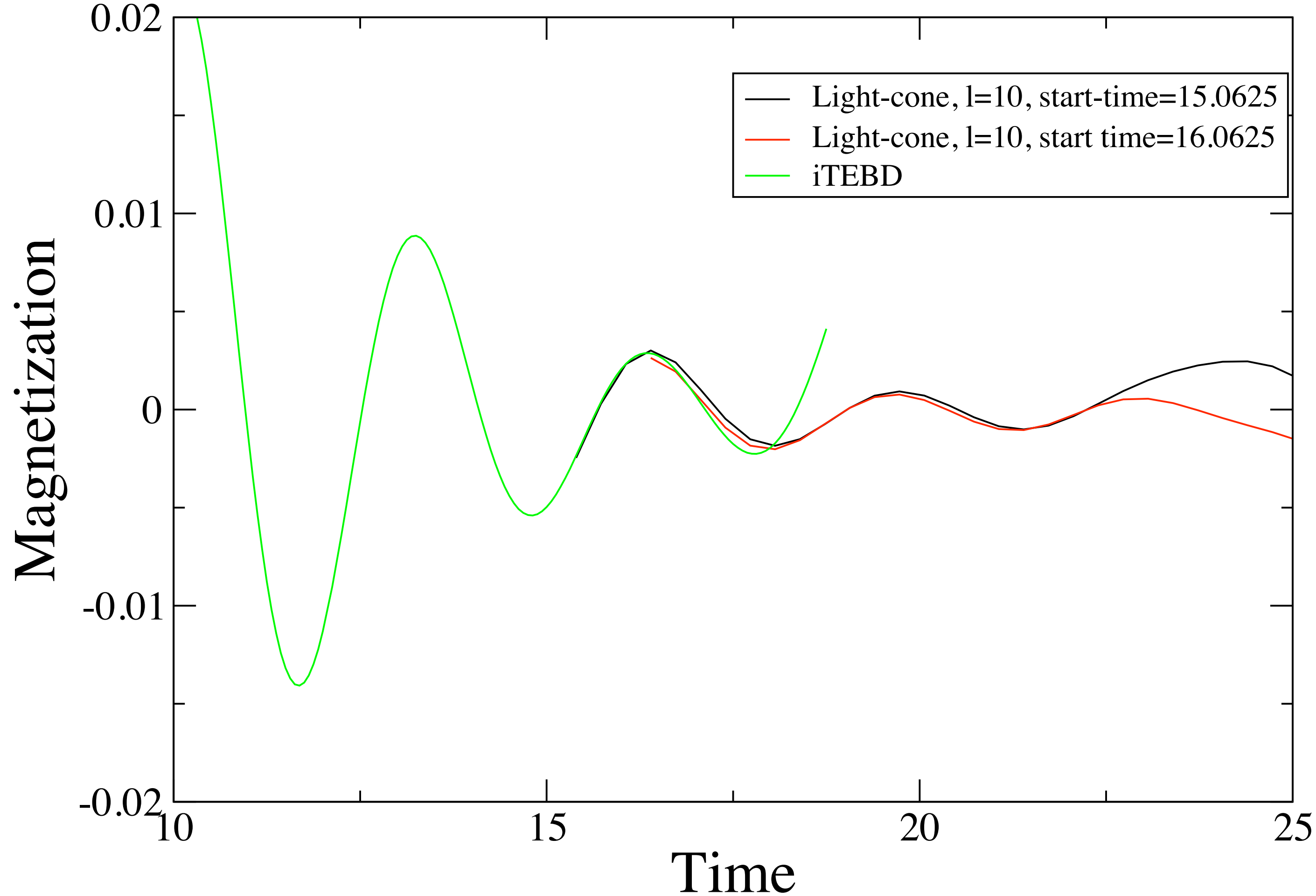
$$H = \sum_i S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z$$

Delta=0.5, interacting



$$H = \sum_i S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z$$

Delta=0.5, interacting



Predictions from Light-Cone Methods:

- Decaying oscillations in order parameter with power-law envelope when no interaction: $cos(2t)/\sqrt{t}$
- Faster decay with interaction. Possible revivals at long time (MBH and L. S. Levitov, 08)
- At very strong interaction, oscillations disappear replaced by kinks

The experiment will probe highly entangled states.

Conclusion

Entangled states improve communication capacity

Thermalization can arise from quantum entanglement

Novel algorithms and predictions

A little history:

- 1824: Carnot efficiency bound. 1890: Rudolf Diesel develops engine, inspired by trying to approach this bound.
- 1948: Shannon's paper. 1993: turbo codes approach Shannon limit. LDPC codes similarly good.
- First topologically protected qubit? Willett et al., 2009. What will come next?