Simulated annealing for Graph Golf

Toru Koizumi

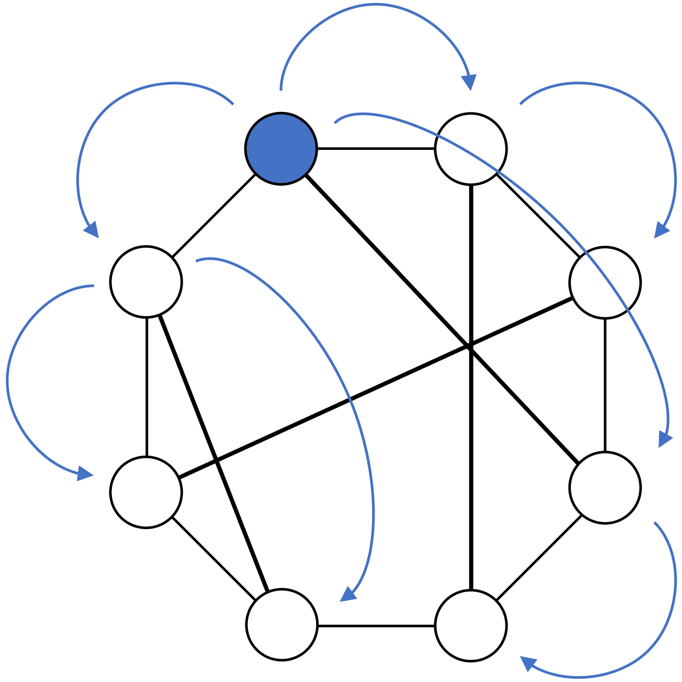
Graduate School of Information Science and Technology
The University of Tokyo

Introduction and my result

Graph Golf

- Given number of vertex (n) and degree (d),
  - find the small diameter graph,
  - and minimize Average Shortest Path Length (ASPL, l).

$n = 8, d = 3$



Diameter = 2
ASPL=1.571 (opt.)

Graph Golf

- Average Shortest Path Length (ASPL, l) has a theoretical lower bound (L).
  - "ASPL gap" ($l - L \geq 0$)

- For general graph with small n, we may be able to find a optimal solution ($l = L$).

- I am a beginner → Target to 'Deepest Improvement'

General graph with small n

- I focused on these three problem.

Order n	Degree d	Length r	Diam. k
72	4		4
256	5		5
256	10		3

- I thought (n=72,d=4) is the easiest but actually (n=256, d=10) is the easiest.

I got optimal solution for (n=256, d=10)

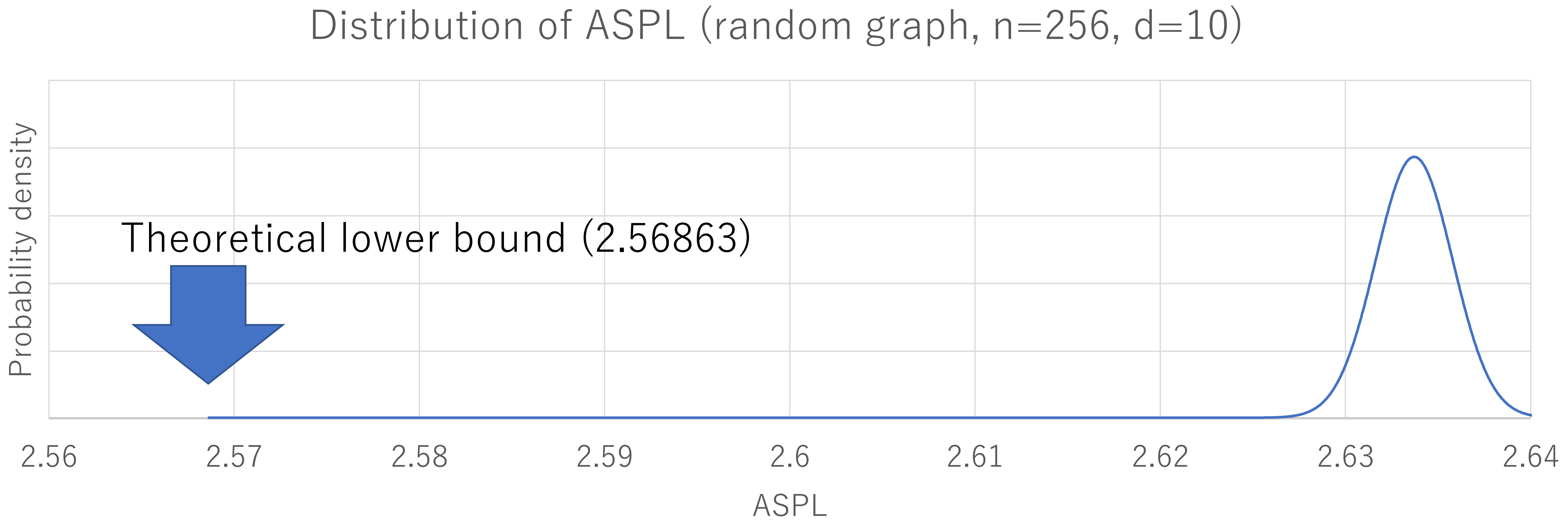
256 nodes, degree 10

	Rank	Author	Diam. k	ASPL l	Diam. gap	ASPL gap	Info.	Date (UTC)	Week
	1	Haruishi masato	3	2.56863	0	0.00000		2018-08-13 09:05:21	33
	1	thai9cdb	3	2.56863	0	0.00000		2018-07-29 13:08:27	30
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-12 01:36:40	24
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-12 01:33:30	24
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-11 14:52:50	24
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-11 14:50:46	24
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-11 14:48:06	24
★	1	Toru Koizumi	3	2.56863	0	0.00000		2018-06-11 14:42:56	24
★	1	Teruaki Kitasuka, Masahiro Iida	3	2.56863	0	0.00000	>_	2018-05-15 09:40:41	20
★	1	Masahiro Nakao	3	2.56863	0	0.00000		2018-05-15 05:11:20	20

My approach

Random Graph

- Random graph has a ASPL that is not bad but far away from optimal solution.



Search for optimal graph

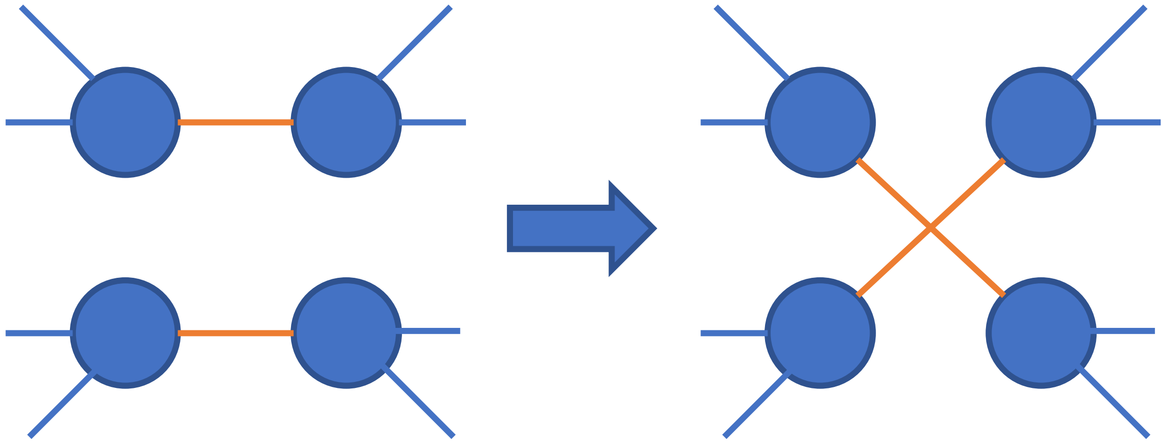
- Random search cannot achieve optimal graph
  - Probability is $\sim\mathbf{10^{-530}}$(n=256, d=10)
- Adopt simulated annealing
  - ? The Neighbors of a state
  - ? Probability distribution
  - ? Transition probabilities
  - ? Cooling schedule

The Neighbors of a state

- The neighbors of a optimal solution must be good solutions

- Should(?) consider the constraint of degree
  - Every vertex has d edges

- 2-opt is the simplest choice



Probability distribution

- Most popular choice: canonical ensemble
- Preparation: normalize graph related values
  - Unit of ASPL $k \coloneqq \frac{2}{n(n-1)}$
  - ASPL gap $\mathrm{E}(G) \coloneqq \mathrm{ASPL}(G) - \mathrm{ASPL}_{\text{lower bound}}$
    - $\frac{\mathrm{E}(G)}{k} \in \mathbb{Z}_{\geq 0}$

- The probability distribution

$$\mathrm{p}(T; G) \propto e^{-\frac{\mathrm{E}(G)}{kT}}$$

Transition probability

- Use Metropolis-Hastings algorithm
- One of the transition probability for realizing $\mathrm{p}(T; G)$

$$\mathrm{A}(T; G, G') = \begin{cases} 1 & for\ \ \mathrm{E}(G) \geq \mathrm{E}(G') \\ e^{-\frac{\mathrm{E}(G') - \mathrm{E}(G)}{kT}} & for\ \ \mathrm{E}(G) < \mathrm{E}(G') \end{cases}$$

Cooling schedule

- Most popular way: Exponential cooling
- This is a bad choice.
- Why?

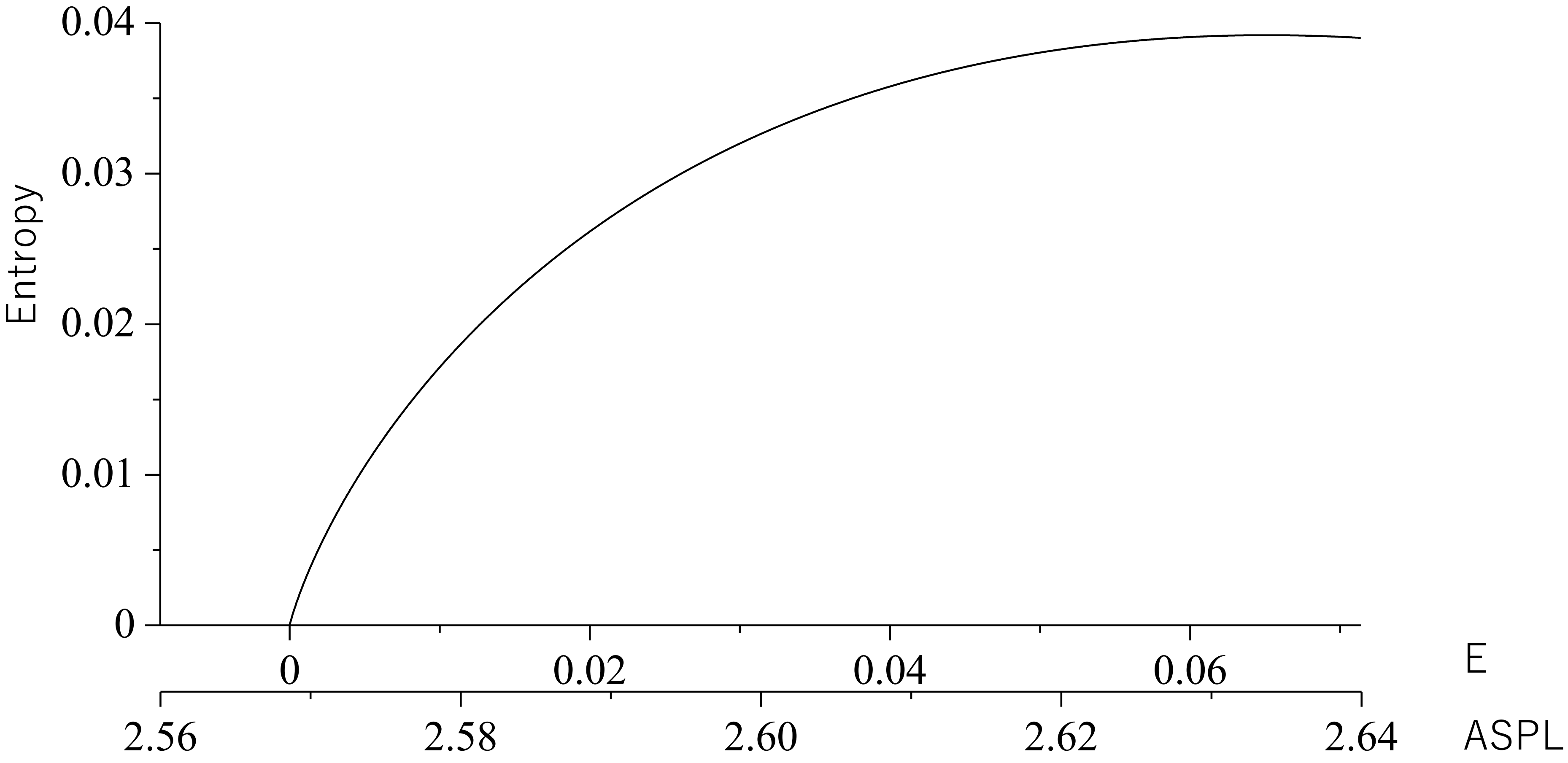
Review of statistical physics

- Number of state $W(E) \coloneqq \#\{G | \mathrm{E}(G) = E\}$
- Entropy $S(E) \coloneqq k \log W(E)$

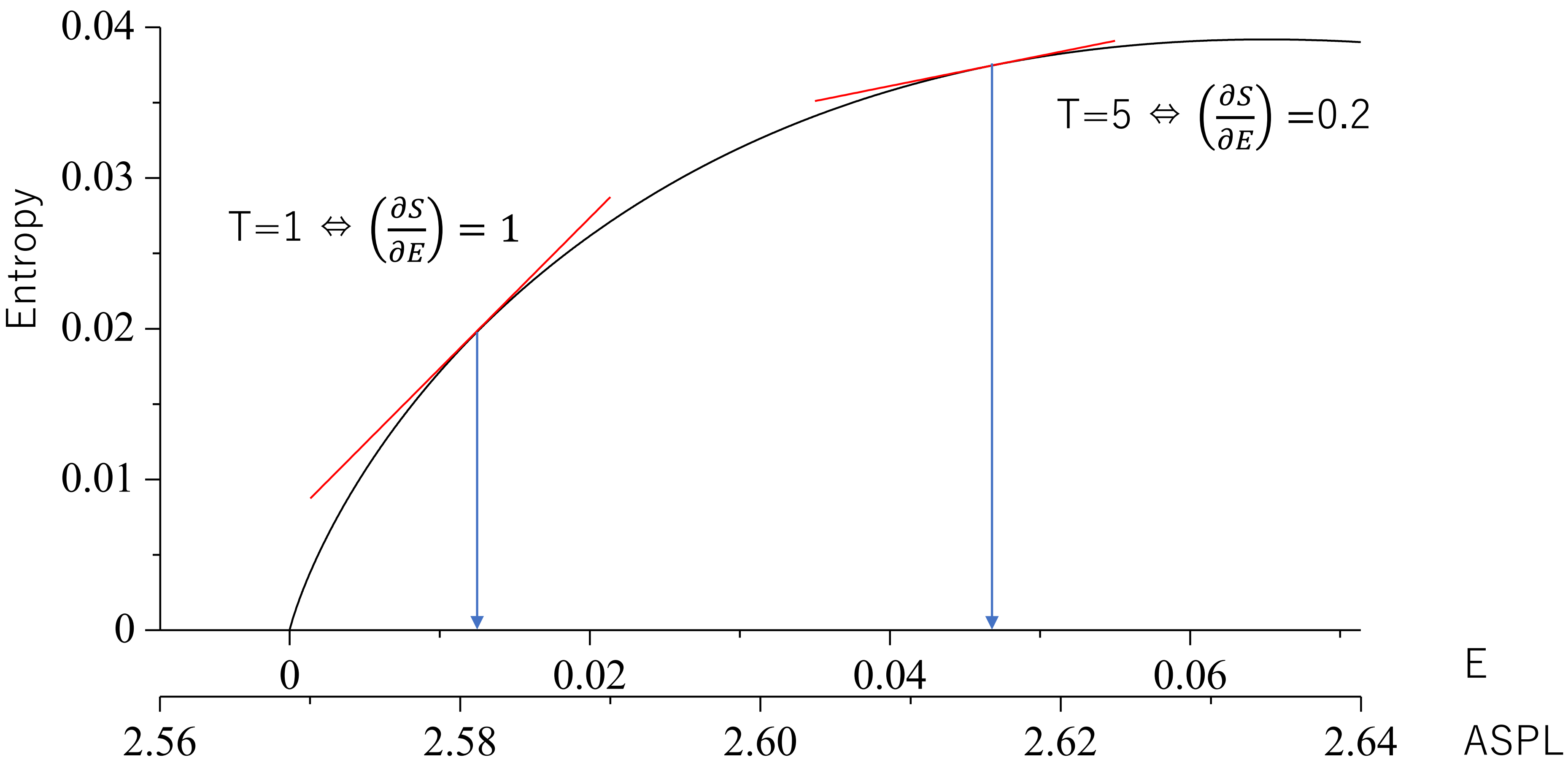
- $\left(\frac{\partial S}{\partial E}\right)_V = \frac{1}{T}$

(V is always constant because we are calculating on canonical ensemble)

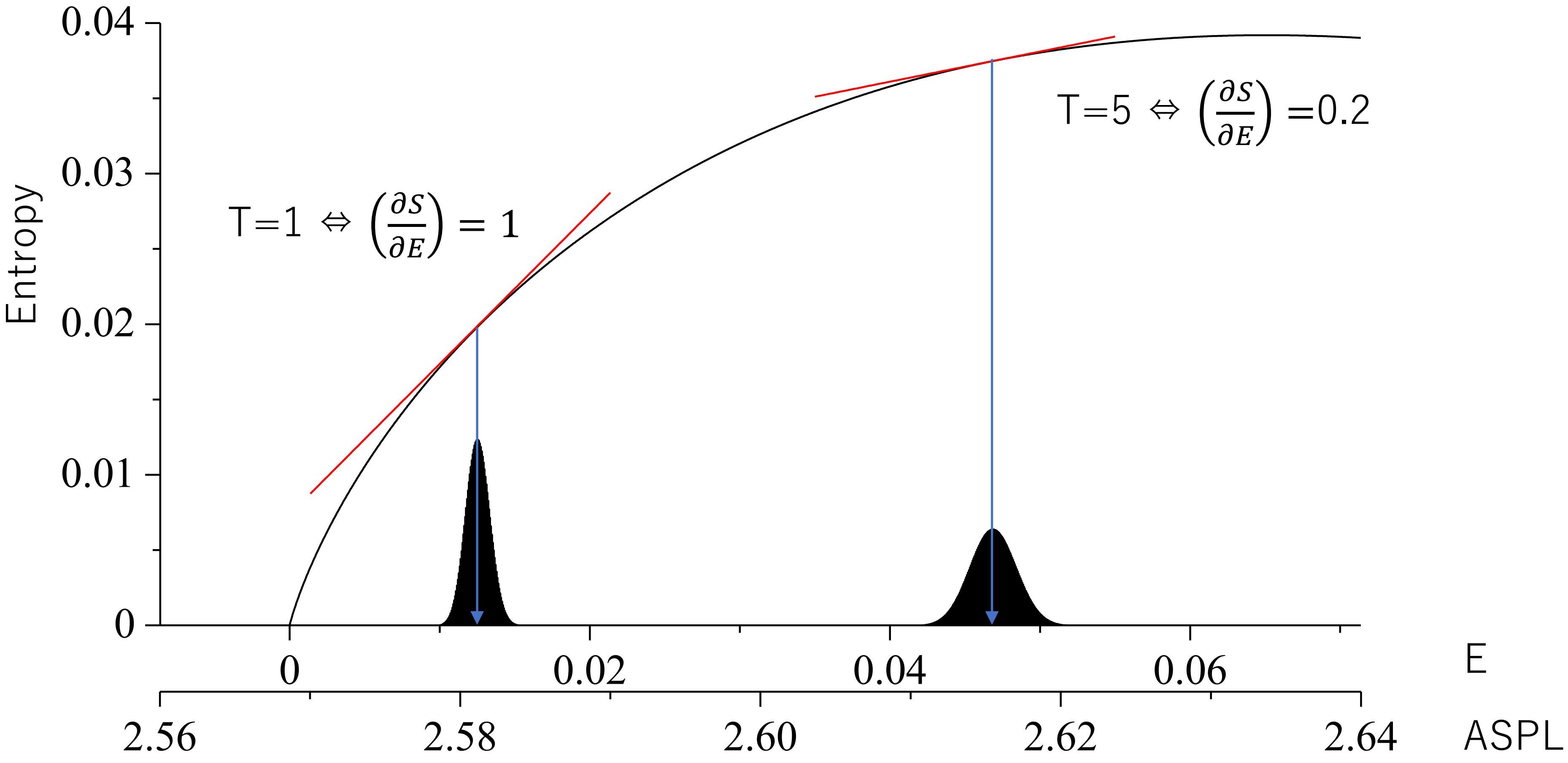
Entropy of graph (n=256, d=10)



Entropy of graph (n=256, d=10)



Entropy of graph (n=256, d=10)

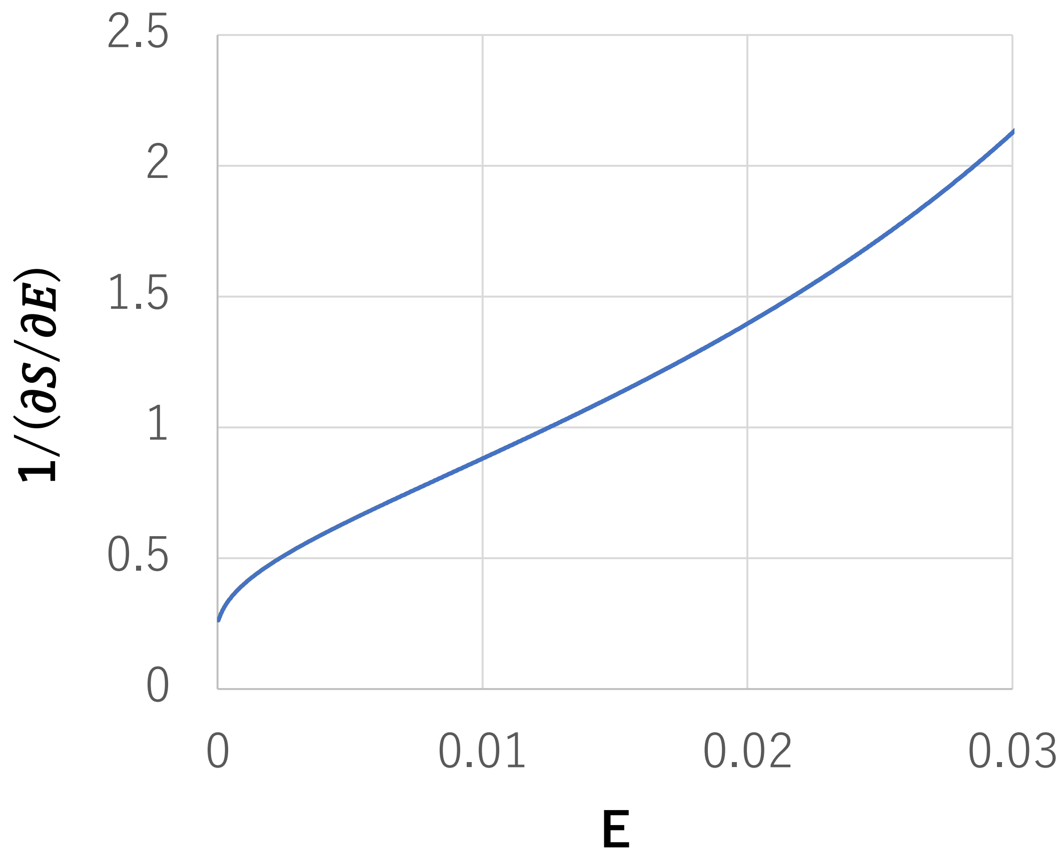


E vs $\left(\frac{\partial S}{\partial E}\right)$ and T (n=256, d=10)

- How far should we lower the temperature?

- $T = \frac{1}{\left(\frac{\partial S}{\partial E}\right)} \sim 0.26$ at $E = 0$

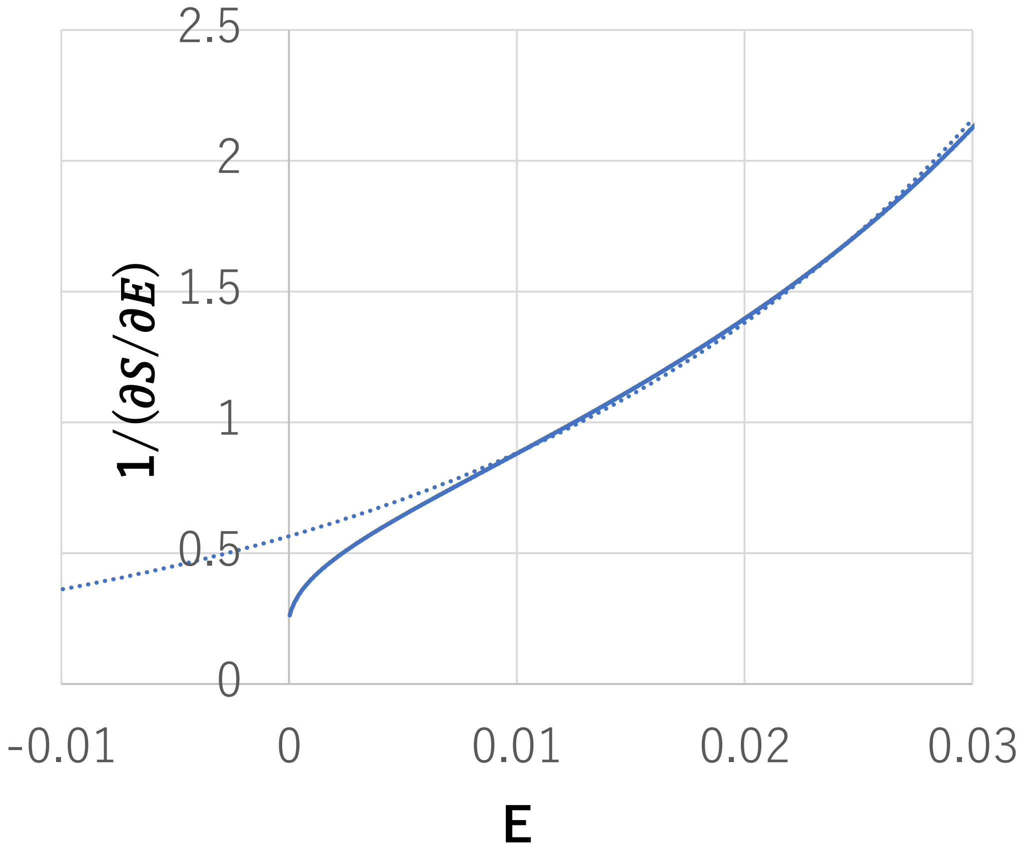
→ calculate with $T \sim 0.26$



Why not exponential cooling

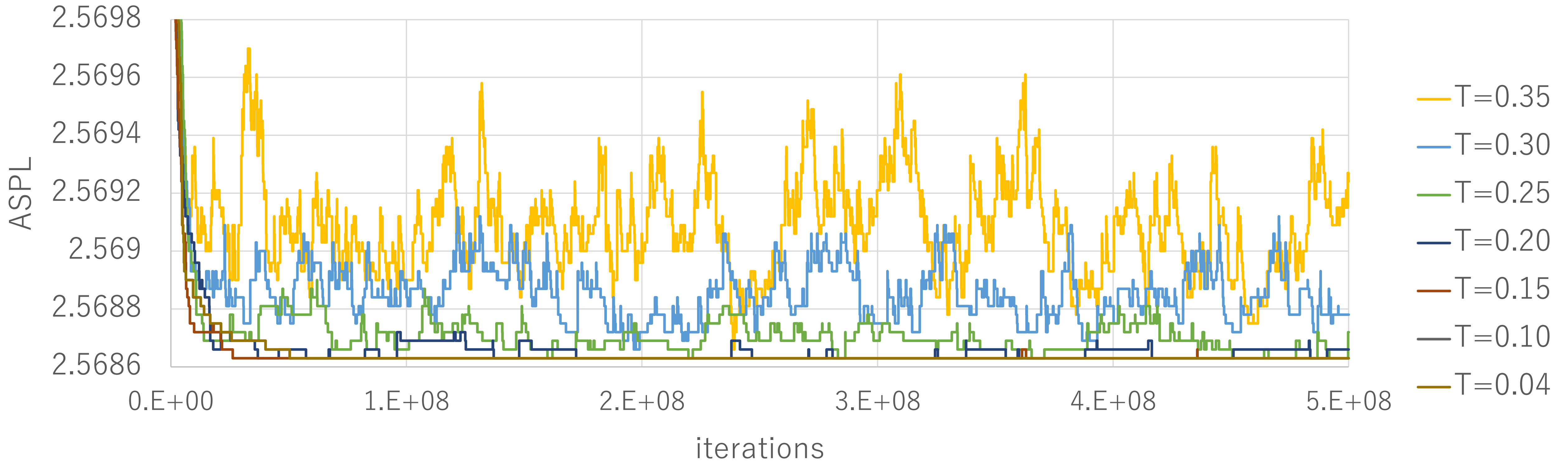
- Using exponential cooling takes very long time to achieve an optimal solution because there are very few E~0 states.

- It is better to quickly cool down when E is small.



Annealing results

- According to theory,
  - $T \geq 0.30$ cannot achieve optimal solution.
  - $T \leq 0.25$ can achieve optimal solution



In fact,

- (n=256, d=10) problem is very easy
- Even greedy method (in other words, simulated annealing with T=0) can achieve optimal solution

Conclusion

- Simulated annealing is simple (but insufficient) way to find optimal solution
- Some configurations can be diverted general purpose one
  - 2-opt
  - Canonical ensemble
  - Metropolis-Hastings algorithm
- It is better to quickly cool down where it is close to optimal solution

Appendix: Entropy of graph (n=72, d=4)