

## Lecture 18: Sampling Monomer-Dimer Configurations

### 1 Introduction

We continue the discussion of polynomial time approximate computing of the partition function defined in the previous lecture. To achieve this ultimate goal, we first consider the problem of sampling from the monomer-dimer systems according to a distribution that probability of each system is proportional to its energy. To begin with, we review the definition of partition function, the setup for the problem of approximate sampling from monomer-dimer systems and the method of canonical path.

### 2 Review of the problem setup

A monomer-dimer system is a cover of the vertices of a graph  $G = (V, E)$  by monomers (molecules covering one vertex) and dimers (molecules covering a pair of adjacent vertices). Note that a monomer-dimer configuration with  $k$  dimers is equivalent to a matching of size  $k$ . The *weight* or *energy* of a monomer-dimer system is  $\lambda^k$ , where  $\lambda$  is a given non-negative constant and  $k$  is the number of dimers. For a given graph  $G = (V, E)$ , let  $m_k$  be the number of monomer-dimer systems with  $k$  dimers. Let  $Z(\lambda) = \sum m_k \lambda^k$  (known as the *partition function*) be the sum of the weights of the monomer-dimer systems. The partition function can also be viewed as  $Z(\lambda) = \sum_{M: M \text{ is a Matching}} \lambda^{|M|}$ .

It is known that computing  $Z(\lambda)$  is  $\#\mathbf{P}$  hard. Our ultimate goal is to approximately count  $Z(\lambda)$ . However, our immediate goal is to sample from the monomer-dimer systems, such that the probability of a monomer-dimer system of weight  $k$  is  $\frac{\lambda^k}{Z(\lambda)}$ . We will then discuss the details of going from approximate sampling to approximate counting. Since the dimers in a monomer-dimer system form a matching, the object is to sample from the matchings, where a matching of size  $k$  has probability  $\frac{\lambda^k}{Z(\lambda)}$ .

We defined a Markov chain whose states are the matchings in  $G = (V, E)$ . The transitions from a state  $M$  to a new state  $M'$  are made as follows:

1. With probability  $\frac{1}{2}$ ,  $M' = M$ .
2. Select an edge  $e = \{u, v\}$  uniformly at random from  $E$

$$M' = \begin{cases} M - e & \text{if } e \in M \\ M + e & \text{if both } u \text{ and } v \text{ are unmatched in } M \\ M + e - e' & \text{if exactly one of } u \text{ and } v \text{ is matched in } M \text{ and } e' \text{ is the matching edge} \\ M & \text{otherwise} \end{cases}$$

3. Go to  $M'$  with probability  $\min\left\{1, \frac{\lambda^{|M'|}}{\lambda^{|M|}}\right\}$ .

It was an exercise to check that the chain is reversible with stationary distribution  $\pi(M) = \frac{\lambda^{|M|}}{Z(\lambda)}$ .

We now want to use the method of Canonical Paths, which is a very useful tool to obtain bounds on the mixing time. The method of Canonical Paths associates with each pair of states  $(x, y)$  a *canonical path*  $\Gamma_{xy} = x_0, x_1, x_2, \dots, x_{|\Gamma_{xy}|}$ , such that  $x_0 = x$ ,  $x_{|\Gamma_{xy}|} = y$  and  $P_{x_i x_{i+1}} > 0$  for all  $i = 0, 1, \dots, |\Gamma_{xy}| - 1$ .

We defined,

$$\bar{\rho} = \max_{(M, M') : Q(M, M') > 0} \frac{1}{Q(M, M')} \sum_{\Gamma_{xy} \ni (M, M')} \pi(x)\pi(y)|\Gamma_{xy}| \quad (1)$$

Here,  $|\Gamma_{xy}|$  denotes the length of the canonical path associated with the pair of states  $(x, y)$ . We also had the following result that shows for all  $x$  and  $\epsilon > 0$  we have,

$$\tau_M(\epsilon) \leq \bar{\rho} \ln \left( \frac{1}{\epsilon \pi(M)} \right) \quad (2)$$

By bounding the quantities of the above inequalities, we will show that

**THEOREM 1**

*The mixing time of this chain satisfies:*

$$\tau_M(\epsilon) \leq 4n|E|\lambda^2 \left( n \ln n + n \ln \lambda + \ln \frac{1}{\epsilon} \right)$$

This result shows that we get polynomial mixing time, even though the number of matchings may be super-exponential. This result will then be used in counting  $Z(\lambda)$ .

### 3 Definition of the canonical paths and the encoding function

We try to first bound  $\bar{\rho}$  from equation (1) as it governs the mixing time of the chain. We will use the Encoding Function technique. But we first need to define the *canonical paths* associated with the pair of monomer-dimer systems  $(X, Y)$ .

Let  $(X, Y)$  be a pair of monomer-dimer systems. Let  $Z = X \oplus Y$  be the symmetric difference of  $X$  and  $Y$ . The definition of  $X$  and  $Y$  implies that  $Z$  is decomposable to the union of some alternating paths and some alternating simple cycles. We can define a linear order  $\alpha$  over the alternating paths and the alternating cycles. We also need to specify a *designated end point* for each path and a *designated vertex* for each cycle (For example, see figure 1.)

The canonical paths  $\Gamma_{X, Y}$  should use the markov chain moves to get us from  $X$  to  $Y$ . The edges in  $X \cap Y$  are remained fixed when we traverse the path, but as for the edges appearing in the symmetric difference between  $X$  and  $Y$ , they will be gradually removed, changed, or added to the system in order to get us from  $X$  to  $Y$ . We will define the operation of unwinding an alternating path or alternating simple cycle. The paths  $\Gamma_{X, Y}$  is obtained

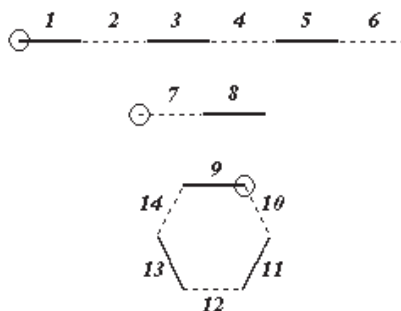


Figure 1: The symmetric difference between two monomer-dimer systems,  $X$  and  $Y$ . Solid lines were used for  $X$  and dashed-dotted lines were used for  $Y$ .

by unwinding the connected components of  $X \oplus Y$  in the order determined by the linear order  $\alpha$ . To unwind a path, we start from the designated end point of the path and carry out the following sequence of transitions (as shown in figure 1, assume that solid lines were used for  $X$  and dashed-dotted lines were used for  $Y$ ):

1. delete the first edge if it is solid
2. replace the second edge by the first edge if the first edge is dashed-dotted.
3. add the first edge if it is dashed-dotted and there is no second edge.

For example, figure 2 shows the unwinding operations on one of the simple paths of figure 1.

To unwind a simple alternating cycle, the following sequence of transitions are carried out (see figure 3):

- delete the solid edge incident with the designated vertex.
- Unwind the simple path that remains, rotating at its non-designated end point. The reason for this will be explained later.

For a fixed pair of monomer-dimer systems  $(M, M')$ , we will define a one-one function

$$\eta_{M, M'} : \{(X, Y) \mid (M, M') \in \Gamma_{X, Y}\} \longrightarrow \Omega$$

where  $\Omega$  is the space of all possible monomer-dimer systems. Given  $M$  and  $M'$ ,  $\eta_{M, M'}(X, Y)$  can be viewed as a "code" identifying the pair  $X, Y$ . We also know that the existence of this one-one function implies that the size of the domain of this function is at most  $|\Omega|$ .

The mapping is defined as follows:

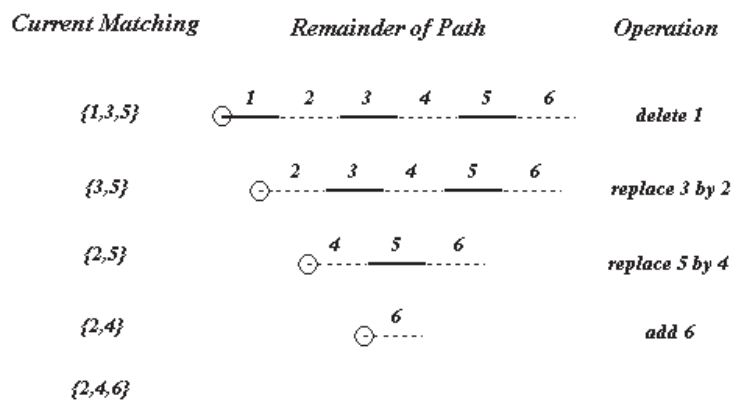


Figure 2: The unwinding operations for a simple alternating path.

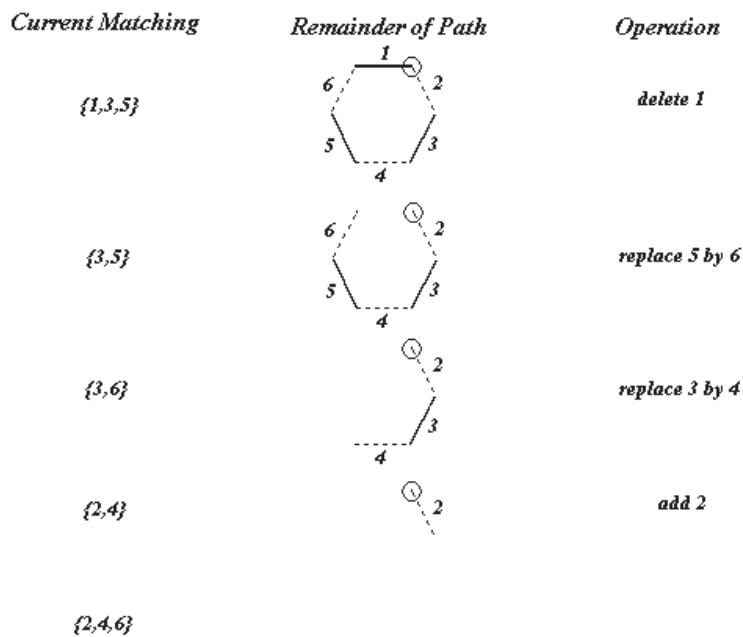


Figure 3: The unwinding operations for a simple alternating cycle.

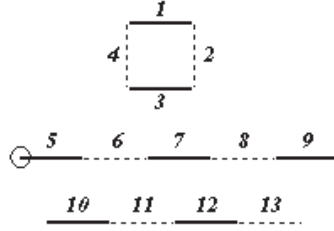


Figure 4: Example 1.

$$\eta_{M,M'}(X, Y) = \begin{cases} (X \oplus Y \oplus (M \cup M')) - e & \text{if the current component of } M \text{ being unwound is a cycle} \\ & \text{and the first solid edge to be deleted is } e \\ X \oplus Y \oplus (M \cup M') & \text{otherwise} \end{cases}$$

**Example 1** In figure 4), consider the step just after the cycle has been unwound. We have:

$$M = \{2, 4, 5, 7, 9, 10, 12\}, M' = \{2, 4, 7, 9, 10, 12\}$$

$$\eta_{M,M'}(X, Y) = X \oplus Y \oplus (M \cup M') = \{1, 3, 6, 8, 11, 13\}$$

**Example 2** Another example is shown in figure 5): Consider the step after edge 1 has been deleted. We have:

$$M = \{3, 5\}, M' = \{3, 6\}, \eta_{M,M'}(X, Y) = \{1, 2, 4\} - \{1\} = \{2, 4\}$$

Reconstructing X and Y:

$$M \cup M' \cup \eta_{M,M'} = \{2, 3, 4, 5, 6\}$$

The fact that the path is being unwound from its non-designated end point tells us to close the cycle by adding edge 1 to the path 2,3,4,5,6.

## 4 A Lemma

In this section , we prove a Lemma which will be useful in bounding  $\bar{\rho}$  from (1).

LEMMA 2

For any  $\lambda \geq 1$ ,

$$\pi(X)\pi(Y) \leq \lambda^2 \min(\pi(M), \pi(M'))\pi(\eta_{M,M'}(X, Y))$$

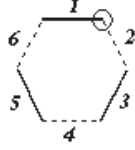


Figure 5: Example 2.

PROOF: For the proof, one can note that

$$\pi(X) = \frac{\lambda^{|X|}}{Z(\lambda)}, \pi(Y) = \frac{\lambda^{|Y|}}{Z(\lambda)}, \text{ etc.}$$

So, it suffices to prove that

$$|X| + |Y| \leq |M| + \eta_{M,M'}(X, Y) + 2 \quad (3)$$

This can be proved by a case analysis. We consider the hardest case, in which

$$\eta = (X \oplus Y \oplus (M \cup M')) - \hat{e}$$

and

$$M' = M + e - e'$$

We can therefore conclude that:

$$(\eta + \hat{e}) \oplus (M \cup M') = X \oplus Y$$

and

$$M \cup M' \subset X \oplus Y$$

The case analysis, shown in figure 6), proves that

$$|X| + |Y| \leq |\eta + \hat{e}| + |M \cup M'| \quad (4)$$

Each row indicates a possible case for membership of an edge in the sets  $X$ ,  $Y$ ,  $\eta + \hat{e}$  and  $M \cup M'$ :

In each case, the sum of the last two columns is at most the sum of the first two, yielding (4). Now,

$|\eta + \hat{e}| \leq |\eta| + 1$  and  $|M \cup M'| = 1 + \min(|M|, |M'|)$ , so (3) is proved.  $\square$

$X$	$Y$	$\eta + \hat{e}$	$M \cup M'$
0	0	0	0
0	1	1	0
0	1	0	1
1	0	1	0
1	0	0	1
1	1	0	0
1	1	1	1

Figure 6: Analysis of all possible cases for membership of an edge in the sets  $X$ ,  $Y$ ,  $\eta + \hat{e}$  and  $M \cup M'$

## 5 Bounding the Mixing time

In this section, we bound  $\bar{\rho}$  from equation (1) and consequently bound the mixing time of the chain. Size of the canonical path can be clearly bounded above by  $2n$ . Using this inequality and the above lemma, we have:

$$\begin{aligned} \sum_{X,Y} \pi(X)\pi(Y)|\Gamma_{X,Y}| &\leq 2n \sum_{X,Y} \lambda^2 \min(\pi(M), \pi(M')) \pi(\eta_{M,M'}(X,Y)) = \\ &= 2n\lambda^2 \min(\pi(M), \pi(M')) \sum_{X,Y} \pi(\eta_{M,M'}(X,Y)) \end{aligned}$$

Since  $\eta_{M,M'}(X,Y)$  is one-to-one, and since the probabilities of all states sum to 1, we have:

$$\sum_{X,Y} \pi(X)\pi(Y)|\Gamma_{X,Y}| \leq 2n\lambda^2 \min(\pi(M), \pi(M'))$$

Also,  $Q(M, M') = \min(\pi(M), \pi(M')) \cdot \frac{1}{2|E|}$  and hence  $\bar{\rho} \leq 4|E|n\lambda^2$

The number of matchings is at most  $2^n \cdot n!$ , so  $Z(\lambda) \leq (2\lambda)^n n!$  and thus for any matching  $M$ ,

$$\pi(M) \geq \frac{1}{(2\lambda)^n n!}$$

Now we can easily prove theorem 1 from (2), theorem 1. Theorem 1 proves a polynomial mixing time, even though the number of matchings may be subexponential.

## 6 Approximating the Partition Function $Z(\lambda)$

It is known that computing  $Z(\lambda)$  is  $\#\mathbf{P}$  hard. In this section through an example, we want to show that approximate sampling can be used to do approximate counting. Our goal in this section is to approximate  $Z(\lambda)$  in polynomial time in the size of graph,  $\frac{1}{\epsilon}$  and  $\ln(\frac{1}{\delta})$ . More precisely, we want to find a randomized computation which returns a value  $X$  such that:

$$Pr((1 - \epsilon)Z(\lambda) \leq X \leq (1 + \epsilon)Z(\lambda)) \geq \frac{3}{4}$$

The algorithm is as follows:

- Choose  $\lambda_0 = 0$ ,  $\lambda_1 = \frac{1}{|E|}$  and  $\lambda_2, \lambda_3, \dots, \lambda_r$  such that

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_r = \lambda$$

$$\lambda_i \leq (1 + \frac{1}{n})\lambda_{i-1}$$

and

$$r \approx n \ln(\lambda e)$$

- For  $i = 1, 2, \dots, r$  construct a randomized experiment returning a value  $X_i$  such that

$$Pr((1 - \frac{\epsilon}{r})\frac{Z(\lambda_i)}{Z(\lambda_{i-1})} \leq X_i \leq (1 + \frac{\epsilon}{r})\frac{Z(\lambda_i)}{Z(\lambda_{i-1})}) \geq 1 - \frac{1}{1+r}$$

- Set  $X = X_1 \cdot X_2 \cdot \dots \cdot X_r$

The idea is that:

$$Z(\lambda) = Z(0) \cdot \frac{Z(\lambda_1)}{Z(\lambda_0)} \cdot \frac{Z(\lambda_2)}{Z(\lambda_1)} \cdot \dots \cdot \frac{Z(\lambda_r)}{Z(\lambda_{r-1})}$$

So, noting that  $Z(0) = 1$ , if we can approximate each term very well, we can approximate their products very well. The ration can however found using sampling.

To compute  $X_i$ , perform  $S$  independent sums of the markov chain with stationary distribution  $\pi(M) = \frac{\lambda_i^M}{Z(\lambda_i)}$  from fixed initial state  $M_0$ , where the length of each run is  $\tau_{M_0}(\frac{\epsilon}{r})$  and  $S = \lceil c\epsilon^{-2}r \rceil$ , where  $c$  is a sufficiently large constant ( $c = 130$  will work).

Let the matching produced by the  $j^{th}$  run be  $M_j$ . Then the estimate of  $\frac{Z(\lambda_i)}{Z(\lambda_{i-1})}$  produced by the  $j^{th}$  run is  $(\frac{\lambda_i}{\lambda_{i-1}})^{|M_j|}$ , and  $X_i$  is the average of these estimates. We omit the proof that this is an unbiased estimator and that all the running times will be polynomial.