

MA3227 : Assignment #3

Due on Wednesday, 4 Apr 2014

**Chen Yu
A0077976E
Engineering Science Programme**

Homework Part I

(1)

Case 1:

$$\hat{V}_{B1} = \frac{\# \text{hits in B}}{\# \text{hits in A}} V_A$$

Let $\hat{\theta}_1 = \frac{\# \text{hits in B}}{\# \text{hits in A}}$, which follows a binomial distribution with $N = n$, $p = \frac{V_B}{V_A}$
Then

$$\sigma_1 = \frac{\theta_1(1 - \theta_1)}{n} V_A^2 = \frac{V_A V_B (V_A - V_B)}{n}$$

Case 2:

$$\hat{V}_{B2} = \frac{\# \text{hits in } B \setminus D}{\# \text{hits in } A \setminus D} (V_A - V_D) + V_{B \setminus D}$$

Let $\hat{\theta}_2 = \frac{\# \text{hits in } BD}{\# \text{hits in } A \setminus D}$, which follows a binomial distribution with $N = n$, $p = \frac{V_{B \setminus D}}{V_{A \setminus D}}$
Then

$$\begin{aligned} \sigma_2 &= \frac{\theta_2(1 - \theta_2)}{n} (V_A - V_D)^2 \\ &= \frac{(V_A - V_D) V_{B \setminus D} (V_A - V_{B \setminus D} - V_D)}{n} \\ &= \frac{(V_A - V_D) (V_B - V_{B \cap D}) (V_A - V_{D \setminus B} - V_B)}{n} \end{aligned}$$

Since

$$\begin{aligned} V_A &> V_A - V_D \\ V_B &> V_B - V_{B \cap D} \\ V_A - V_B &> V_A - V_B - V_{D \setminus B} \end{aligned}$$

we get $\sigma_1 > \sigma_2$

(2)

a) In this method, the estimator follows a binomial distribution with $n = N$ and $p = \frac{I}{1} = I$. Hence the variance is

$$\frac{1}{N}I(1 - I) = \frac{1}{N}(I - I^2)$$

b) The variance of this Monte Carlo Integration is

$$\begin{aligned}\text{Var}\left[\frac{1}{N} \sum_{i=1}^n f(U_i)\right] &= \frac{1}{N^2} N \text{Var}[f(U_i)] \\ &= \frac{1}{N} (\text{E}[f(U_i)^2] - I^2)\end{aligned}$$

Since

$$\forall x \in [0, 1], f(x) \in (0, 1)$$

we have

$$f(x)^2 < f(x)$$

, thus

$$\text{E}[f(U_i)^2] < \text{E}[f(U_i)]$$

Therefore the method in b) has a smaller variance.

(3)

$$\begin{aligned}
(n-1)S_n^2 &= \sum_{i=1}^n (X_i - \bar{X})^2 \\
&= \sum_{i=1}^n (X_i^2 - 2\bar{X}X_i + \bar{x}^2) \\
&= \sum_{i=1}^n X_i^2 - 2\bar{X} \sum_{i=1}^n X_i + n\bar{x}^2 \\
&= \sum_{i=1}^n X_i^2 - 2n\bar{X}^2 + n\bar{x}^2 \\
&= \sum_{i=1}^n X_i^2 - n\bar{X}^2
\end{aligned}$$

$$\begin{aligned}
E[S_n^2] &= \frac{1}{n-1} E\left[\sum_{i=1}^n X_i^2 - n\bar{X}^2\right] \\
&= \frac{1}{n-1} (nE[X_i^2] - nE[\bar{X}^2]) \\
&= \frac{n}{n-1} (E[X_i^2] - E[\bar{X}^2])
\end{aligned}$$

$$E[X_i^2] = (E[X_i])^2 + \text{Var}[X_i]$$

$$\begin{aligned}
E[\bar{X}^2] &= (E[\bar{X}])^2 + \text{Var}[\bar{X}] \\
&= (E[\frac{1}{n} \sum_{i=1}^n X_i])^2 + \text{Var}[\frac{1}{n} \sum_{i=1}^n X_i] \\
&= (E[X_i])^2 + \frac{1}{n} \text{Var}[X_i] \quad (\text{i.i.d})
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[S_n^2] &= \frac{n}{n-1} ((E[X_i])^2 + \text{Var}[X_i] - (E[X_i])^2 - \frac{1}{n} \text{Var}[X_i]) \\
&= \text{Var}[X_i]
\end{aligned}$$

Homework Part II(1)

Let $P_{x,y}$ be the transition probability function. The transition probability function generated by the new Metropolis algorithm is

$$Q_{x,y} = \min\{1, \frac{P_{y,x}f(y)}{P_{x,y}f(x)}\}P_{x,y}$$

$$\begin{aligned} f(x)Q_{x,y} &= f(x) \min\{1, \frac{P_{y,x}f(y)}{P_{x,y}f(x)}\}P_{x,y} \\ &= \min\{f(x)P_{y,x}, f(y)P_{y,x}\} \end{aligned}$$

$$\begin{aligned} f(y)Q_{y,x} &= f(y) \min\{1, \frac{P_{x,y}f(x)}{P_{y,x}f(y)}\}P_{y,x} \\ &= \min\{f(x)P_{y,x}, f(y)P_{y,x}\} \end{aligned}$$

$$\implies f(x)Q_{x,y} = f(y)Q_{y,x}$$

, thus satisfying the detailed balance, which ensures that the Metropolis-Hasting chain generated has invariant distribution.

Computer Project Part I

Code

```

clear;
close all;
graphics_toolkit("gnuplot");

5 function I = repI(i)
    I = [];
    for j = 1:i
        I = [I, 'I'];
    end
10 end

nTrials=2.^(1:20); m=length(nTrials);

vEst = cell(3,1);
15 for i = 1:3
    vEst{i} = zeros(1,m);
end

rand('state',0);
20 for i=1:m
    n=nTrials(i);
    x = rand(1,n);
    vEst{1}(i) = sum(4*sqrt(1-x.^2))/n;
    x = -sqrt(4 - 3*rand(1,n)) + 2;
25 vEst{2}(i) = sum(12*sqrt(1-x.^2)./(4-2*x))/n;
    x = -sqrt(1 - rand(1,n)) + 1;
    vEst{3}(i) = sum(4*sqrt(1-x.^2)./(2-2*x))/n;
end

30 err = cell(3,1);
for i = 1:3
    err{i} = abs(vEst{i} - pi);
end

35 equation = cell(3,1);
equation{1} = '$g(x) = 1$';
equation{2} = '$g(x) = \frac{4-2x}{3}$';
equation{3} = '$g(x) = 2-2x$';

40 for i = 1:3
    the_plot = figure(i);
    subplot(2,1,1);
    semilogx(nTrials,vEst{i}, 'LineWidth', 2);
    title(['Method ',repI(i),': ', equation{i}]);
45
    subplot(2,1,2);
    hold on;
    loglog(nTrials, err{i}, 'LineWidth', 2);
    loglog(nTrials, 1./sqrt(nTrials),'r:', 'LineWidth', 2);
50 legend('$\abs{Err}$ vs number of trials',

```

```

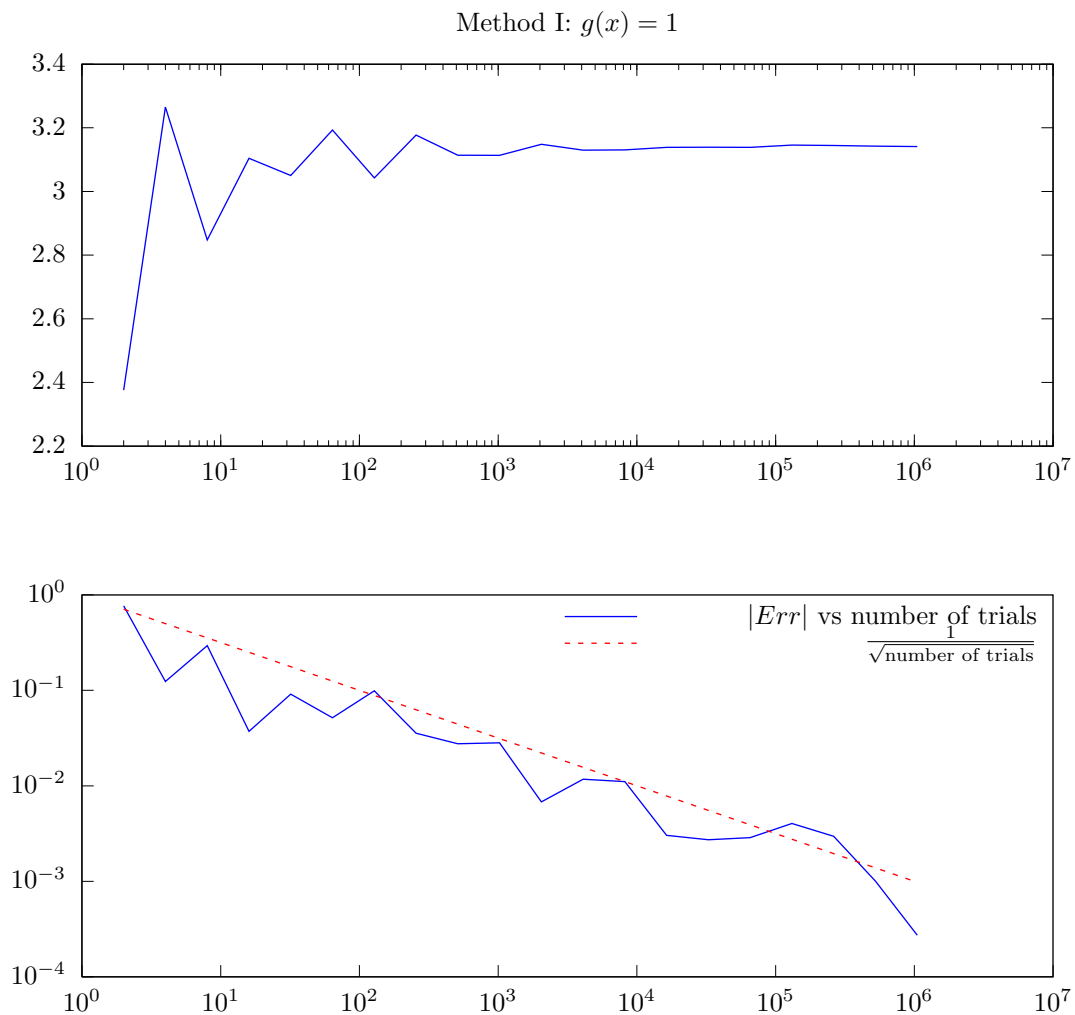
        '$\frac{1}{\sqrt{\text{number of trials}}}$');
    legend boxoff;

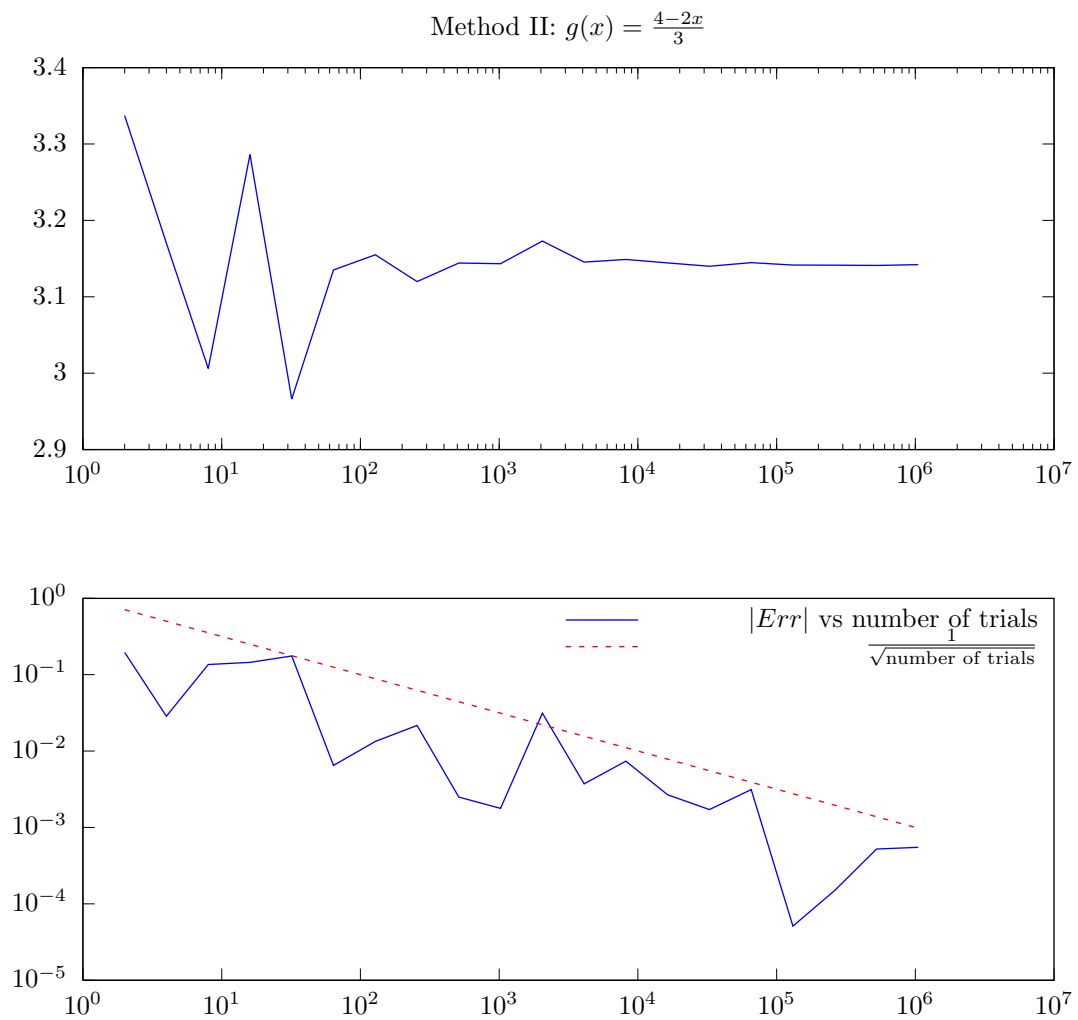
    print(the_plot, ['Method_', repI(i), '.tex'], '-S470,420', '-dtex')
55 end

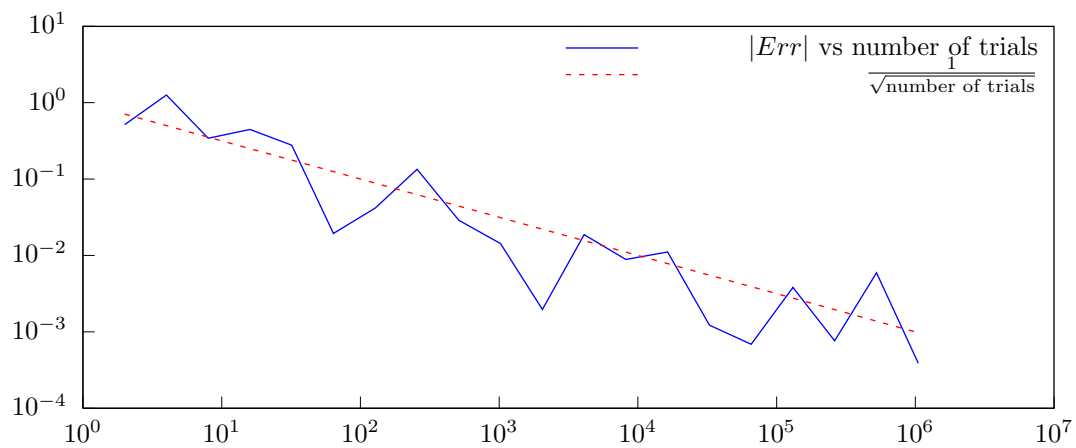
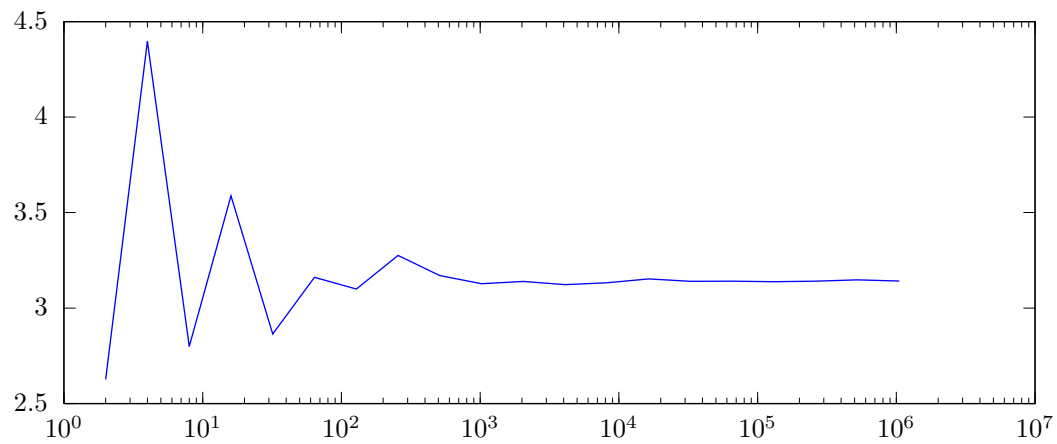
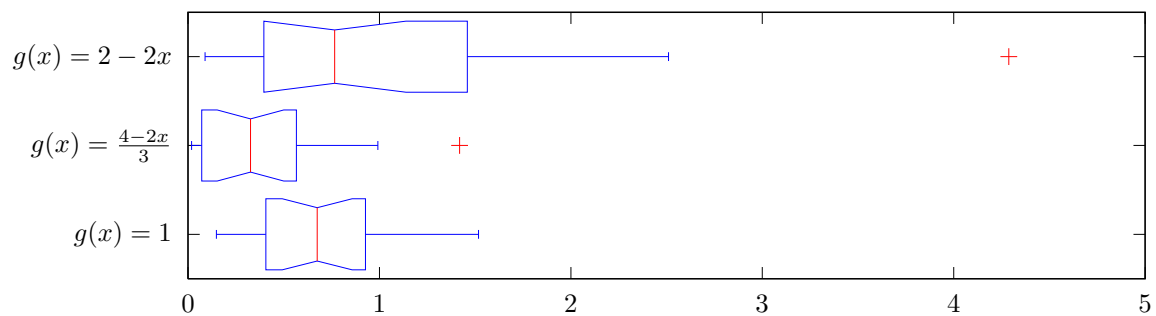
err_plot = figure(4);
err_diff = cell2mat(err).*sqrt(nTrials);
boxplot(err_diff', 1, '+', 0);
60 title(['$\abs{Err} \times \sqrt{\text{number of trials}}$',
        ' for different methods'])
set(gca (), 'ytick', [1 2 3], 'yticklabel', equation)
print(err_plot, 'comparison', '-S470,150', '-dtex')

```

Figures





Method III: $g(x) = 2 - 2x$  $|Err| \times \sqrt{\text{number of trials}}$ for different methods

Computer Project Part II

Code

```

clear; close all;
graphics_toolkit("gnuplot");

N_bins = 101;

5
function M = Metro_Ising(mu, N_steps, sample_rate, N)
    x = ones(N, 1);
    M = zeros(0, N_steps/sample_rate);
    for i = 1:N_steps
10        j = ceil(N*rand);
        if j == 1
            h = exp(-2*mu*x(1)*(x(2)));
        elseif j == N
            h = exp(-2*mu*x(N)*(x(N-1)));
15        else
            h = exp(-2*mu*x(j)*(x(j-1)+x(j+1)));
        end
        U = rand;
        if U <= h
20            x(j) = -x(j);
        end
        if mod(i, 50) == 0
            M(i/50) = sum(x);
        end
25    end
end

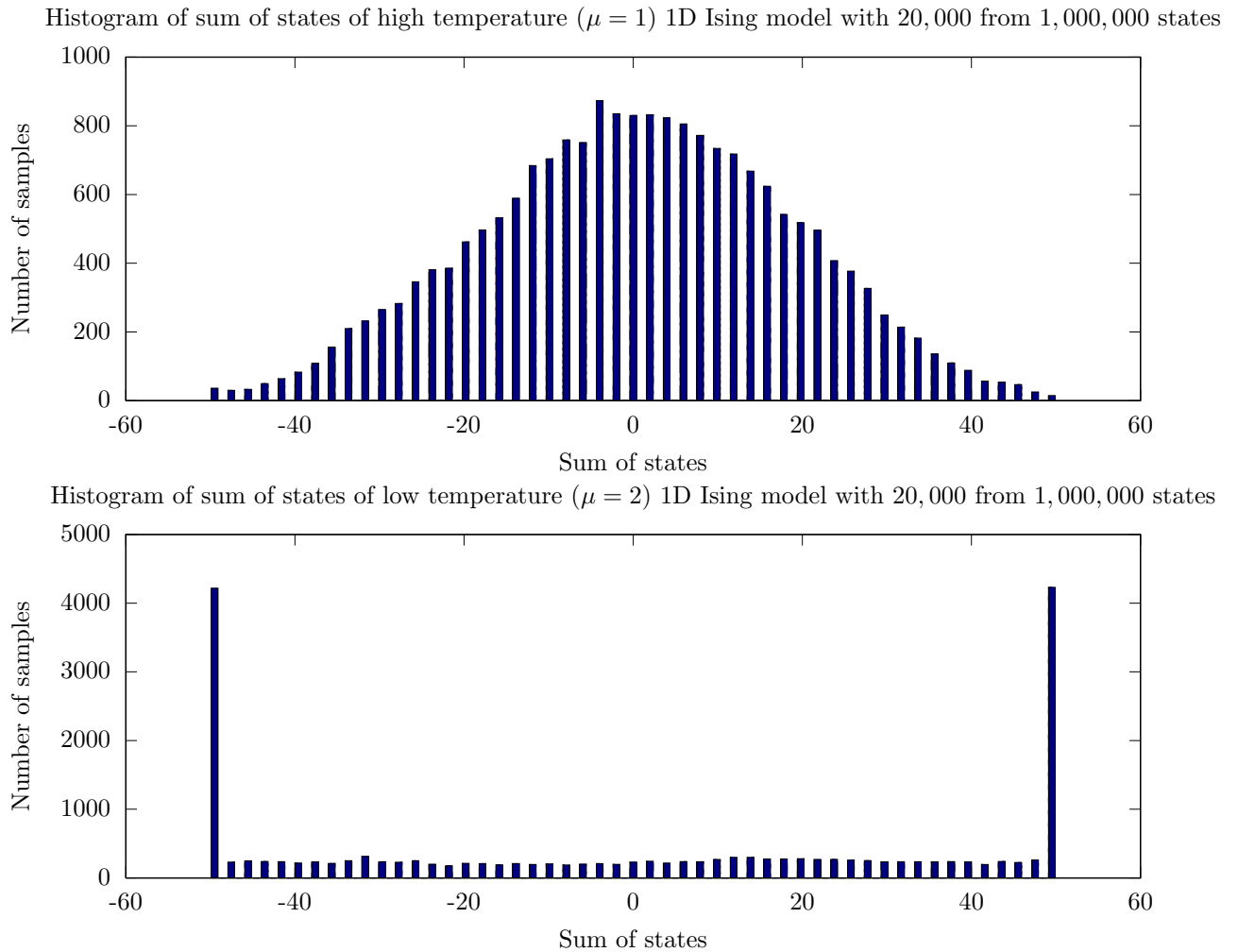
N_steps = 1e6;
sample_rate = 50;
30 N = 50;
%mesh = -N:floor(2*N/(N_bins - 1)):N;
%[n, h] = hist(M, mesh);
the_plot = figure();
M_h = Metro_Ising(1, N_steps, sample_rate, N);
35 M_l = Metro_Ising(2, N_steps, sample_rate, N);

h_plot = subplot(2,1,1);
hist(M_h, 101);
title(['Histogram of sum of states of high temperature ($\mu = 1$)',
40     '1D Ising model with $20,000$ from $1,000,000$ states']);
xlabel('Sum of states');
ylabel('Number of samples');

l_plot = subplot(2,1,2);
45 hist(M_l, 101);
title(['Histogram of sum of states of low temperature ($\mu = 2$)',
     '1D Ising model with $20,000$ from $1,000,000$ states']);
xlabel('Sum of states');
50 ylabel('Number of samples');
```

```
print(the_plot, ['MetropolisIsing', '.tex'], '-S520,400', '-dtex')
```

Figures



Homework Part II(2)

f is the distribution of all microstates in the state space of 2^N possible micro states of (x_1, x_2, \dots, x_N) , where $x_i = -1$ or 1 .
 $P_{x,y} = \frac{1}{N}$ if $x_i = y_i$ for all but one component; which results in a proposed state y where only the spin of one particle is reversed.