EE 638: Assignment

Verification of the asymptotic behaviour of MLE via Monte Carlo simulation

Archishman Biswas, 180070009

October 10, 2021

Contents

1 Basics of Monte Carlo Simulations

The Monte Carlo Simulation or Monte Carlo Method(MCM) can be defined as a procedure that uses randomly generated numbers to solve a given problem [2]. Often, in statistics, MCM is used to solve problems that are not analytically tractable. They are also often used to determine the power of statistical tests for which analytical derivation is not possible. Hence, in general we can define Monte Carlo Methods as techniques that models the uncertainty in output when uncertainty in input and the model(input-output relations) are given[3, 1].

General steps to be followed in MCM:

- Generate data as per given input random variable distribution
- Pass the data generated through the defined system and store the output
- Repeat the previous steps multiple times so that a histogram of output random variable can be obtained and also parameters of the distribution of output can be estimated reliably

In order for MCMs to work, we need to ensure that a large number of inputs are used and also the inputs are taken truly at random from the given input distributions.

2 Method Followed

2.1 Background Theory

The problem given involves the estimation of the deterministic parameter A, with the data coming from the following process:

$$x[n] = A + w[n] \; n \in \{0, 1, 2, ..., N-1\} \tag{1}$$

with $w[n] \sim \mathcal{N}(0, A)$. The Fisher information for the above case is calculated to be:

$$\mathbb{I}(A) = \frac{N}{A^2}(\frac{1}{2} + A) \tag{2}$$

The MLE estimator for the above case is given be the following expression:

$$\hat{A}_{MLE} = -\frac{1}{2} + \sqrt{\frac{\sum_{n=0}^{N-1} x_n^2}{N} + \frac{1}{4}} \tag{3}$$

The MLE estimator $\hat{A}_{MLE}$ is biased but is a consistent estimator, so we have $\hat{A}_{MLE} \to A$ as $N \to \infty$. In general the distribution of $\hat{A}_{MLE}$ asymptotically approaches a normal distribution with the parameters as given below:

$$\hat{A}_{MLE} \sim \mathcal{N}(A, \mathbb{I}^{-1}(A))$$

$$\hat{A}_{MLE} \sim \mathcal{N}(A, \frac{A^2}{N(\frac{1}{2} + A)}) \tag{4}$$

However, this is difficult to show analytically, hence we will use MCM to verify the fact that the MLE asymptotically achieves normal distribution with its mean equal to A(the unknown parameter) and variance equal to $\mathbb{I}^{-1}(A)$(the inverse of Fisher information).

2.2 Steps Followed

The steps followed to obtain the results are:

- Generate N data points $x[n]$ as per equation 1.
- Calculate the $\hat{A}_{MLE}$ for the given data $x[n]$ as per equation 3.

- Repeat the calculation of $\hat{A}_{MLE}$ for K such realizations of the data $x[n]$ and store the estimates.
- Estimate the mean and variance of the $\hat{A}_{MLE}$ random variable from the K samples/realizations of $\hat{A}_{MLE}$.
- Observe the convergence of the sample mean and sample variance of $\hat{A}_{MLE}$ to A and $\mathbb{I}^{-1}(A)$ for different choices of N and K.

Also, we need to check the plots for histogram of $\hat{A}_{MLE}$ to verify it is similar to the PDF of $\mathcal{N}(A, \mathbb{I}^{-1}(A))$, i.e. confirming the Gaussian shape of the distribution.

3 Codes with Explanation

This section describes the **MATLAB** codes along with the explanations.

3.1 Basic Functions Used

In order to generate a N length data vector, we use the following function:

```
function X = generate_data(A,N)
%generates a vector data X of Gaussian distribution,
%mean = A, variance = A and length = N
X = normrnd(A,sqrt(A),N,1);
end
```

Then, for calculation of an $\hat{A}_{MLE}$, we use the function:

```
function A_MLE = calc_MLE(X)
%calculates the MLE from the given N length data X
N = size(X,1);
X_norm_by_N = dot(X,X)/N;
A_MLE = -0.5 + sqrt(X_norm_by_N + 0.25);
end
```

Now, to get a sense of distribution of $\hat{A}_{MLE}$ using histograms, we will need many such MLE values for different realizations of the data, this is achieved using the function:

```
function A_MLE = get_A_distribution(A,N,K)
%generates a total of K valus of A_MLE for given A and N,
%this K-length vector A_MLE gets the idea of distribution of A_MLE
A_MLE = zeros(K,1);
for j = 1:K
    X = generate_data(A,N);
    A_MLE(j) =  calc_MLE(X);
end
end
```

To plot the distribution of $\hat{A}_{MLE}$ and calculate its sample mean and sample variance, we use the function:

```
function [A_mean_est, A_var_est] = show_results_single_run(A_MLE,N,A)
%given the samples of A_MLE for chosen N and A, this function
%calculates mean and variance of A_MLE distribution
%and plots the histogram along with actual normal PDF of A_MLE

K = size(A_MLE,1);
A_mean_est = mean(A_MLE);  A_var_est = (std(A_MLE))^2;
A_var_act = (A^2)/(N*(0.5+A));

```

```
[f,x] = hist(A_MLE,K/10);
g = (1/sqrt(2*pi*A_var_act))* exp(-((x-A).^2)/(2*A_var_act));
bar(x,f/trapz(x,f)); hold on
plot(x, g, 'r', LineWidth = 2); hold off
xlabel("A_{MLE} Samples"); ylabel("PDF of A_{MLE}");
title("For N = "+num2str(N)+" and K = "+num2str(K));
end
```

3.2 Part 1: Qualitative for Distributions

In order to get the distributions of $\hat{A}_{MLE}$ for various choices of N and K, we use the following wrapper code:

```
clear all; close all;

A = 9; N = 1e3;
K = 1e4; %total number of times we estimate the MLE of A

%% Check the variation of estimated mean and variance for different K
A_var_act = (A^2)/(N*(0.5+A));

K_vals = [1e2,5e2,1e3,5e3,1e4,5e4];
[mean_MLE,var_MLE] = results_for_different_K(N,A,K_vals);
```

The main function used in part 1 is the one obtaining results for different K. It essentially takes the given N, A and an array of K values and returns the distribution $\hat{A}_{MLE}$ along with actual PDF of $\hat{A}_{MLE}$ as per the procedures described earlier and using the functions defined previously. The function codes are:

```
function [mean_MLE,var_MLE] = results_for_different_K(N,A,K_vals)
%For chosen N and A, this function plots the
%histogram along with actual normal PDF for different K

data_len = size(K_vals,2);
mean_MLE = zeros(data_len,1); var_MLE = zeros(data_len,1);

s = get(0, 'ScreenSize');
figure('Position', [0,0,s(3),s(4)]);

for index = 1:data_len
    K_this = K_vals(index);
    A_MLE = get_A_distribution(A,N,K_this);
    subplot(2,data_len/2,index)
    [mean_MLE_this,var_MLE_this] = show_results_single_run(A_MLE,N,A);
    mean_MLE(index) = mean_MLE_this; var_MLE(index) = var_MLE_this;
end
end
```

3.3 Part 2: Quantitative for Parameters

Similarly, a wrapper code for part 2 in order to verify the mean and variance of distribution is used:

```
clear all; close all;

A = 9; %last two digits of roll number = 09
```

```
%% Check the variation of estimated mean and variance for different N
N_vals = round(logspace(0,3,10));

s = get(0, 'ScreenSize');
figure('Position', [0,0,s(3),s(4)]);

subplot(2,2,1);
K = 2e2;
results_for_different_N(N_vals,A,K);

subplot(2,2,2);
K = 3e2;
results_for_different_N(N_vals,A,K);

subplot(2,2,3);
K = 4e2;
results_for_different_N(N_vals,A,K);

subplot(2,2,4);
K = 5e2;
results_for_different_N(N_vals,A,K);
```

The main function used in part 2 is the one obtaining results for different N. It essentially takes the given K, A and an array of N values and returns the plot for sample mean and sample variance calculated and compared with the actual means and variance for different values of N:

```
function [mean_MLE,var_MLE] = results_for_different_N(N_vals,A,K)
%For chosen K and A, this function plots the sample mean and
%sample variance of A_MLE along for different N

data_len = size(N_vals,2);
mean_MLE = zeros(data_len,1); var_MLE = zeros(data_len,1);

actual_var = (A^2)./((0.5+A)*N_vals);

for index = 1:data_len
    N_this = N_vals(index);
    A_MLE = get_A_distribution(A,N_this,K);
    A_mean_est = mean(A_MLE); A_var_est = (std(A_MLE))^2;
    mean_MLE(index) = A_mean_est; var_MLE(index) = A_var_est;
end

semilogx(N_vals,mean_MLE,'ro-','LineWidth',2); grid on; hold on;
semilogx(N_vals, A*ones(data_len),'m--','LineWidth',1.5);
legend("Mean of A_{MLE}","Actual Value A")

semilogx(N_vals,var_MLE,'bo-','LineWidth',2,'DisplayName',"Variance of A_{
    MLE}");
semilogx(N_vals, actual_var,'c--','LineWidth',1.5,'DisplayName',"I(A)^{-1}")
    ;
%legend("Mean of A_{MLE}","Actual Value A","Variance of A_{MLE}","I(A)
    ^{-1}");

```

```
hold off;


xlabel("Number of Observations N(in log scale)");
ylabel("Mean and Variance Calculated from MCM");
title("Number of Samples of $$\hat{A}_{MLE}$$ used(K) = "+num2str(K),'Interpreter','Latex');

ylim([-2,12]);

end
```

4 Simulation Plots and Results

In part 1, we verify that the distribution is of normal form and then in part 2, the parameters mean and variance of this distribution is computed and compared with the theoretically known value.

4.1 Part 1: Qualitatively Verifying Distribution

In order to verify the nature of distribution of $\hat{A}_{MLE}$, we first plot the histograms obtained for different sample sizes K and different number of observations N. The plot for some of these cases is shown below:

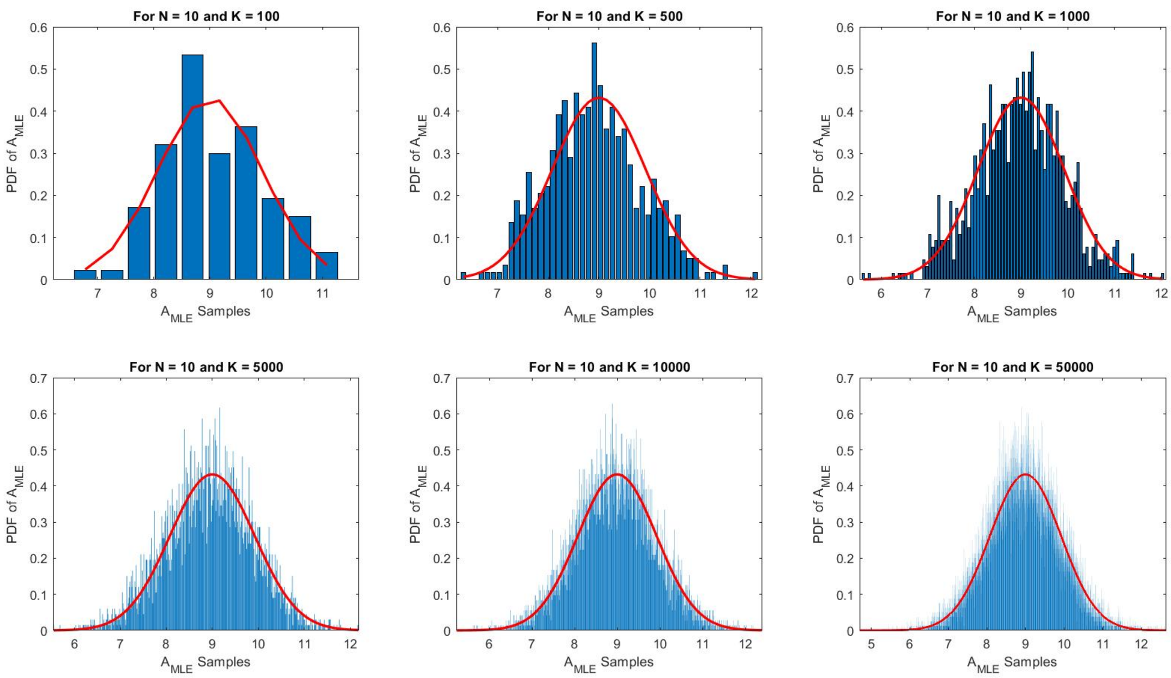

Figure 1: Histograms of $\hat{A}_{MLE}$ for N = 10

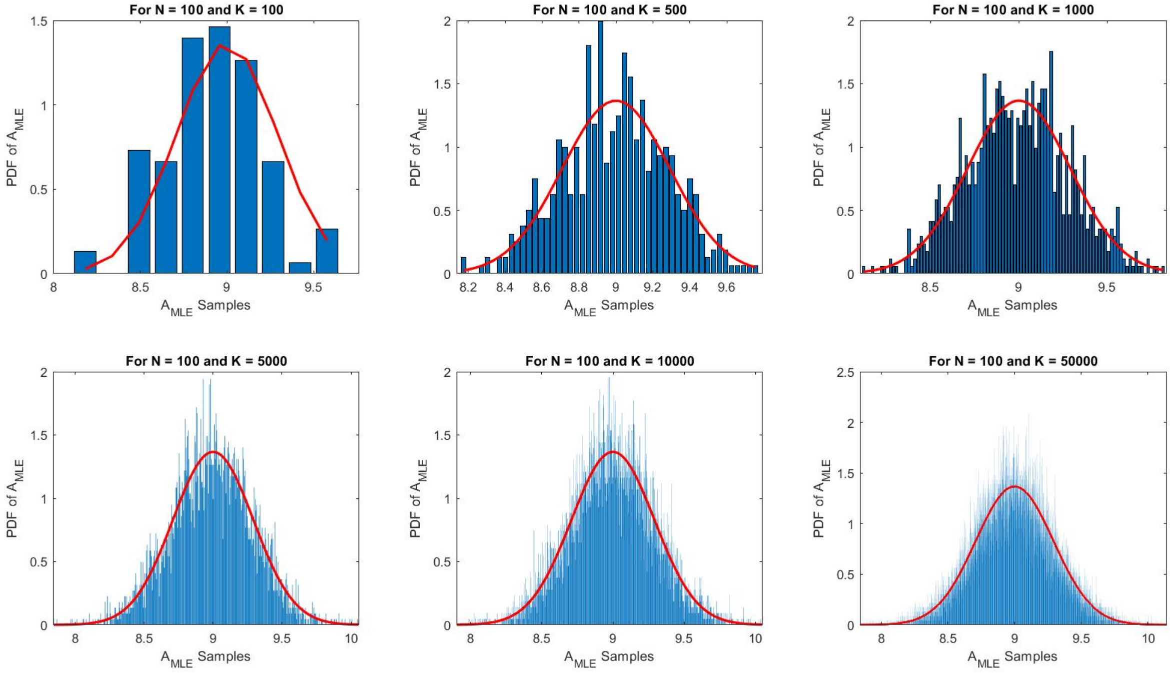

Figure 2: Histograms of $\hat{A}_{MLE}$ for N = 100

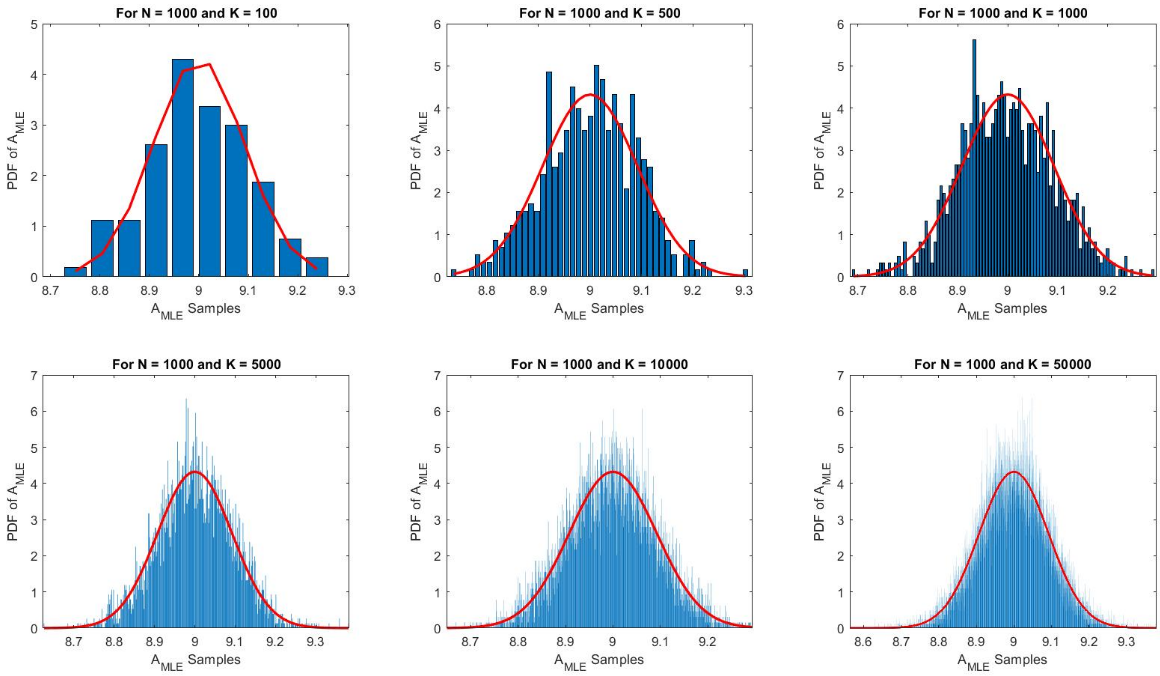

Figure 3: Histograms of $\hat{A}_{MLE}$ for N = 1000

From the above plots, with the actual PDF $\mathcal{N}(A, \mathbb{I}^{-1}(A))$ in red shows that the distribution is indeed very close to a normal distribution.

4.2 Part 2: Quantitatively Verifying Parameters

In order to get a quantitative understanding, we plot the calculated sample mean and variance and compare them with the actual mean and variance for various K and N. In each of the cases with different K, 10 values of N in 1 to 1000 is taken.

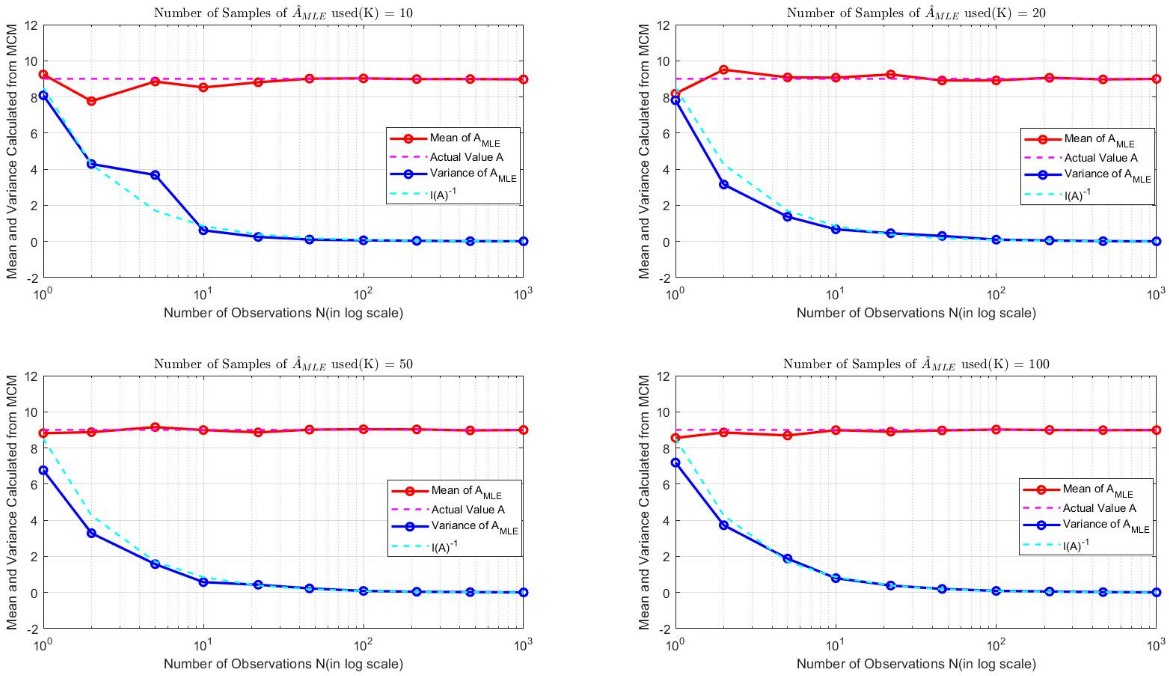

Figure 4: Mean and Variance of $\hat{A}_{MLE}$

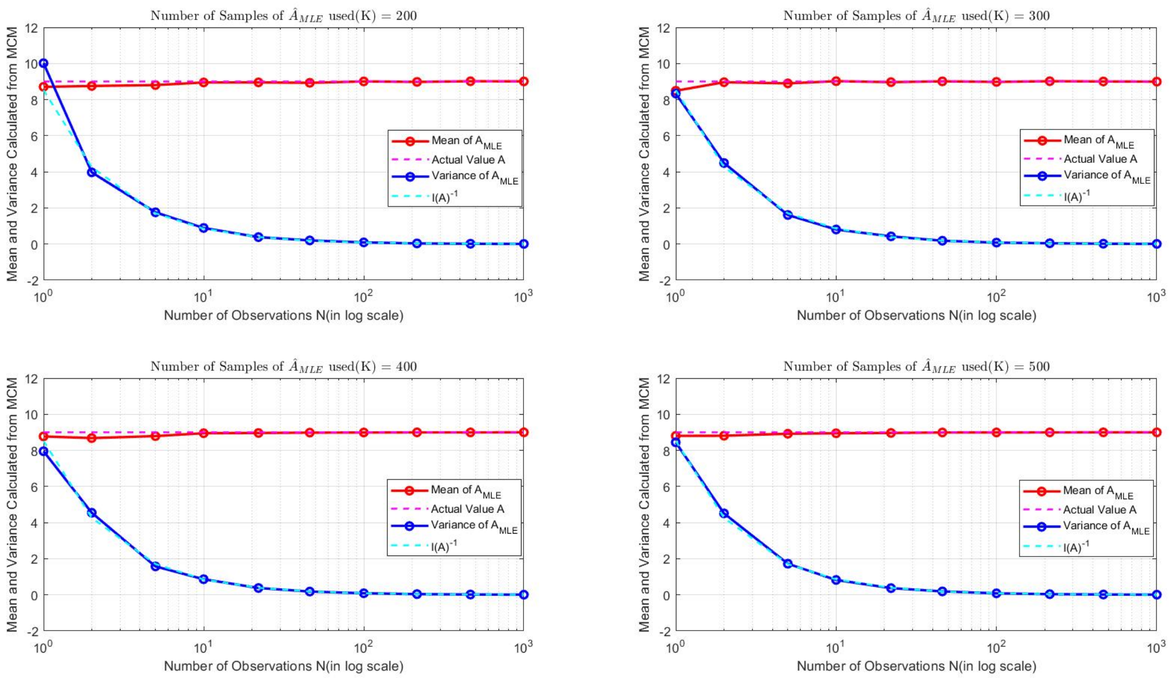

Figure 5: Mean and Variance of $\hat{A}_{MLE}$

In the above plots, we can verify that the parameters of distribution of $\hat{A}_{MLE}$ computed using MCM are very close to the actual values. Also, the calculations improve on having higher values of N and K.

5 Discussions on Simulation Results

We observe the following from the simulation results displayed above:

- The distribution as observed by constructing the histogram of $\hat{A}_{MLE}$ approaches $\mathcal{N}(A, \mathbb{I}^{-1}(A))$ as we keep on increasing N. Also, better histograms are obtained for large number of realizations/samples K.

- As the number of samples N is increased, the sample mean and sample variance of $\hat{A}_{MLE}$ indeed converges to A and $\mathbb{I}^{-1}(A)$.
- Also, for plots of different K, we can observe that the convergence is faster in cases where more realizations/samples K are used.
- Monte Carlo method has been successful in verifying that the distribution of $\hat{A}_{MLE}$ approaches $\mathcal{N}(A, \mathbb{I}^{-1}(A))$ as $N \to \infty$

References

[1] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory.* USA: Prentice-Hall, Inc., 1993. ISBN: 0133457117.

[2] Krishnamurty Muralidhar. "Monte Carlo Simulation". In: *Encyclopedia of Information Systems.* Elsevier, 2003, pp. 193–201. DOI: `10.1016/b0-12-227240-4/00114-3`. URL: `https://doi.org/10.1016/b0-12-227240-4/00114-3`.

[3] Detlev Reiter. "The Monte Carlo Method, an Introduction". In: *Computational Many-Particle Physics.* Springer Berlin Heidelberg, pp. 63–78. DOI: `10.1007/978-3-540-74686-7_3`. URL: `https://doi.org/10.1007/978-3-540-74686-7_3`.