Importance Sampling and Metropolis-Hastings' algorithm

Morten Hjorth-Jensen Email morten.hjorth-jensen@fys.uio.no[1]

Department of Physics and Center fo Computing in Science Education, University of Oslo, Oslo, Norway[1]

February 13, 2026

Overview of week February 9-13, 2026

Topics

- Short repetition from last week
- Mathematical and computational details of importance sampling and Fokker-Planck and Langevin equations

- Video of lecture at https://youtu.be/HIus9-Ry9qQ
- Whiteboard notes at https://github.com/CompPhysics/ComputationalPhysics2/blob/gh-pages/doc/HandWrittenNotes/2026/FYS4411week4.pdf

Reading suggestions

To read more about Metropolis, Markov Chains, importance sampling, Fokker-Planck and Langevin equations, we recommend Becca and Sorella's text, sections 3.4-3.9 and 4.2-4.4.

Importance sampling and overview of what needs to be coded, reminder from last week

For a diffusion process characterized by a time-dependent probability density $P(x,t)$ in one dimension the Fokker-Planck equation reads (for one particle /walker)

$$\frac{\partial P}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - F\right)P(x,t),$$

where F is a drift term and D is the diffusion coefficient.

Importance sampling

The new positions in coordinate space are given as the solutions of the Langevin equation using Euler's method, namely, we go from the Langevin equation

$$\frac{\partial x(t)}{\partial t} = DF(x(t)) + \eta,$$

with η a random variable, yielding a new position

$$y = x + DF(x)\Delta t + \xi\sqrt{\Delta t},$$

where ξ is gaussian random variable and Δt is a chosen time step. The quantity D is, in atomic units, equal to $1/2$ and comes from the factor $1/2$ in the kinetic energy operator. Note that Δt is to be viewed as a parameter. Values of $\Delta t \in [0.001, 0.01]$ yield in general rather stable values of the ground state energy.

Importance sampling

The process of isotropic diffusion characterized by a time-dependent probability density $P(\mathrm{x}, t)$ obeys (as an approximation) the so-called Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \sum_i D \frac{\partial}{\partial \mathrm{x_i}} \left(\frac{\partial}{\partial \mathrm{x_i}} - \mathrm{F_i} \right) P(\mathrm{x}, t),$$

where $\mathrm{F_i}$ is the i^{th} component of the drift term (drift velocity) caused by an external potential, and D is the diffusion coefficient. The convergence to a stationary probability density can be obtained by setting the left hand side to zero. The resulting equation will be satisfied if and only if all the terms of the sum are equal zero,

$$\frac{\partial^2 P}{\partial \mathrm{x_i}^2} = P \frac{\partial}{\partial \mathrm{x_i}} \mathrm{F_i} + \mathrm{F_i} \frac{\partial}{\partial \mathrm{x_i}} P.$$

Importance sampling

The drift vector should be of the form $\mathrm{F} = g(\mathrm{x})\frac{\partial P}{\partial \mathrm{x}}$. Then,

$$\frac{\partial^2 P}{\partial \mathrm{x_i}^2} = P\frac{\partial g}{\partial P}\left(\frac{\partial P}{\partial \mathrm{x}_i}\right)^2 + Pg\frac{\partial^2 P}{\partial \mathrm{x}_i^2} + g\left(\frac{\partial P}{\partial \mathrm{x}_i}\right)^2.$$

The condition of stationary density means that the left hand side equals zero. In other words, the terms containing first and second derivatives have to cancel each other. It is possible only if $g = \frac{1}{P}$, which yields

$$\mathrm{F} = 2\frac{1}{\Psi_T}\nabla\Psi_T,$$

which is known as the so-called *quantum force*. This term is responsible for pushing the walker towards regions of configuration space where the trial wave function is large, increasing the efficiency of the simulation in contrast to the Metropolis algorithm where the walker has the same probability of moving in every direction.

Importance sampling

The Fokker-Planck equation yields a (the solution to the equation) transition probability given by the Green's function

$$G(y, x, \Delta t) = \frac{1}{(4\pi D\Delta t)^{3N/2}} \exp\left(-(y - x - D\Delta t F(x))^2/4D\Delta t\right)$$

which in turn means that our brute force Metropolis algorithm

$$A(y, x) = \min(1, q(y, x))),$$

with $q(y, x) = |\Psi_T(y)|^2/|\Psi_T(x)|^2$ is now replaced by the Metropolis-Hastings algorithm as well as Hasting's article,

$$q(y, x) = \frac{G(x, y, \Delta t)|\Psi_T(y)|^2}{G(y, x, \Delta t)|\Psi_T(x)|^2}$$

Code example for the interacting case with importance sampling

We are now ready to implement importance sampling. This is done here for the two-electron case with the Coulomb interaction, as in the previous example. We have two variational parameters α and β. After the set up of files

```
# Common imports
import os

# Where to save the figures and data files
PROJECT_ROOT_DIR = "Results"
FIGURE_ID = "Results/FigureFiles"
DATA_ID = "Results/VMCQdotImportance"

if not os.path.exists(PROJECT_ROOT_DIR):
    os.mkdir(PROJECT_ROOT_DIR)

if not os.path.exists(FIGURE_ID):
    os.makedirs(FIGURE_ID)

if not os.path.exists(DATA_ID):
    os.makedirs(DATA_ID)

def image_path(fig_id):
    return os.path.join(FIGURE_ID, fig_id)

def data_path(dat_id):
```