Smoldyn – Howto

@ IIT Guwahati's Param-Ishan

Probhonjon Baruah

May 5, 2017

Department of Biosciences And Bioengineering
Indian Institute of Technology Guwahati
Guwahati 781039, Assam, INDIA

Introduction to Smoldyn

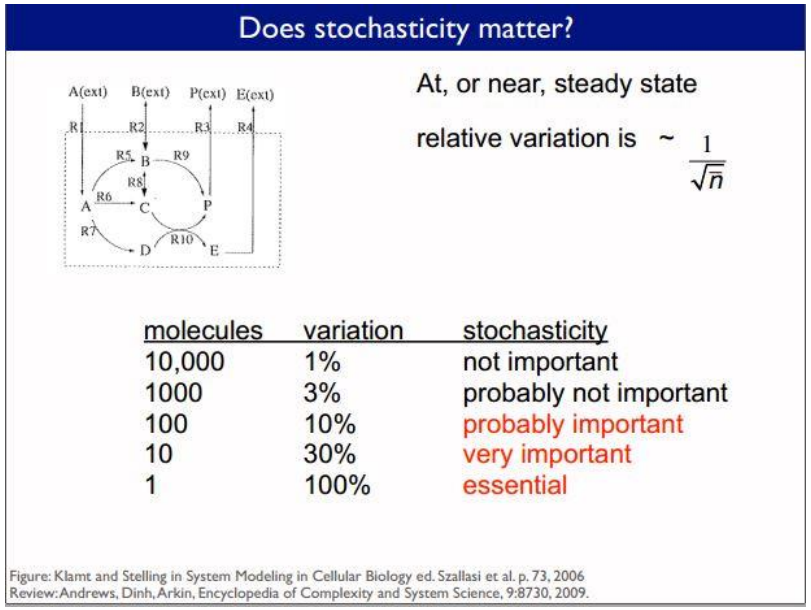
What is smoldyn?

Smoldyn is a computer program for simulating chemical processes on a microscopic size scale. In Smoldyn, each molecule is represented by a single point in 1-, 2-, or 3-dimensional continuous space. In other words, smoldyn uses particle-based modelling.

Homepage : *http://www.smoldyn.org/index.html*
Download Link : *http://www.smoldyn.org/download.html*

Why is smoldyn needed?

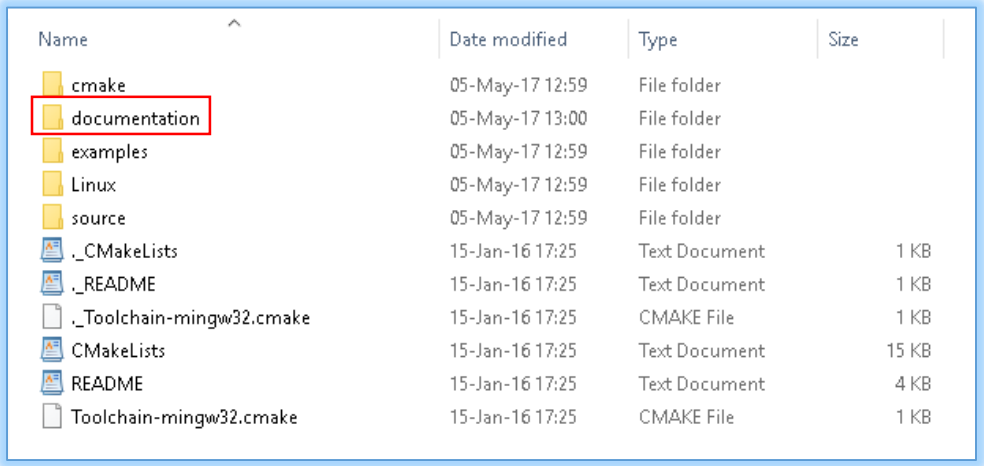
In a cell, the concentration of certain molecular species is very low. As such, behavior of such molecules can be better explained by stochastic models rather than deterministic models. Smoldyn is one of the few softwares that help in simulating cell reaction based on stochastic models.



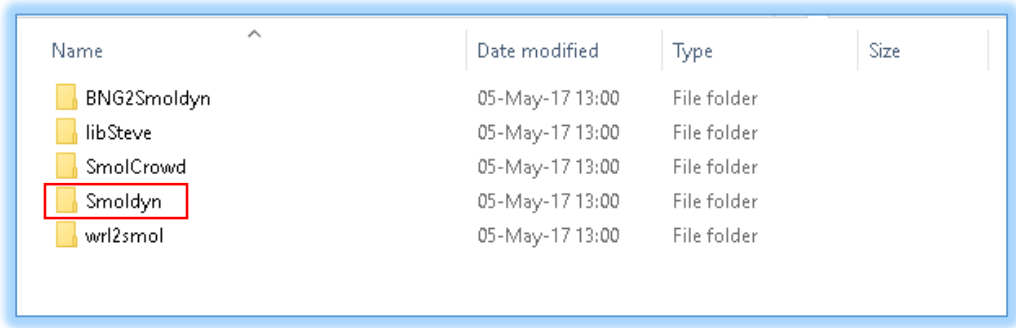
Introduction to Smoldyn

The contents of the smoldyn package

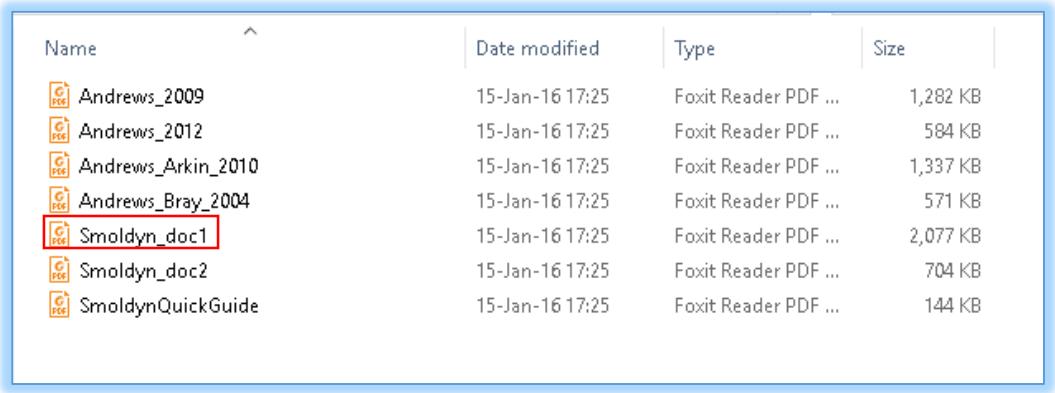
Download and extract the smoldyn package



Name	Date modified	Type	Size
cmake	05-May-17 12:59	File folder	
documentation	05-May-17 13:00	File folder	
examples	05-May-17 12:59	File folder	
Linux	05-May-17 12:59	File folder	
source	05-May-17 12:59	File folder	
._CMakeLists	15-Jan-16 17:25	Text Document	1 KB
._README	15-Jan-16 17:25	Text Document	1 KB
._Toolchain-mingw32.cmake	15-Jan-16 17:25	CMAKE File	1 KB
CMakeLists	15-Jan-16 17:25	Text Document	15 KB
README	15-Jan-16 17:25	Text Document	4 KB
Toolchain-mingw32.cmake	15-Jan-16 17:25	CMAKE File	1 KB



Name	Date modified	Type	Size
BNG2Smoldyn	05-May-17 13:00	File folder	
libSteve	05-May-17 13:00	File folder	
SmolCrowd	05-May-17 13:00	File folder	
Smoldyn	05-May-17 13:00	File folder	
wrl2smol	05-May-17 13:00	File folder	



Name	Date modified	Type	Size
Andrews_2009	15-Jan-16 17:25	Foxit Reader PDF ...	1,282 KB
Andrews_2012	15-Jan-16 17:25	Foxit Reader PDF ...	584 KB
Andrews_Arkin_2010	15-Jan-16 17:25	Foxit Reader PDF ...	1,337 KB
Andrews_Bray_2004	15-Jan-16 17:25	Foxit Reader PDF ...	571 KB
Smoldyn_doc1	15-Jan-16 17:25	Foxit Reader PDF ...	2,077 KB
Smoldyn_doc2	15-Jan-16 17:25	Foxit Reader PDF ...	704 KB
SmoldynQuickGuide	15-Jan-16 17:25	Foxit Reader PDF ...	144 KB

- The documentation folder contains all manuals and research paper related to smoldyn.
- All reference in this presentation are based on Smoldyn_doc1 manual of version 2.39.
- The “examples” folder contains many example script files which help understand how to write configuration file for smoldyn based on different needs.

Introduction to Smoldyn

Recommended scale of smoldyn system

Space : nanometers to micrometres
Time : tens of nanoseconds to tens of minutes
Upto 100,000 molecules

Lower limits:

1. Smoldyn does not represent
   - Excluded volumes
   - Orientation
   - Moments of simulated molecules
2. Spatial resolution : a few molecular radii
3. Temporal resolution : molecules' rotational diffusion time constant

Upper Limits:

1. Computationally intensive
2. Currently not successfully parallelized but GPU implementation under development

Warning!

No portions of Smoldyn ever use any units. Instead, it is up to the user to decide what set of units to use and to stay consistent with them. Not being careful with units may result in wrong simulation and consequent result without smoldyn giving any warnings/errors.

Table 3.2.1: Unit conversion

	Concentration	Diffusion coefficient	Unimolec. reactions	Bimolecular reactions	Adsorption rates
Typical value	10 μM	10 $\mu m^2 s^{-1}$	1 s^{-1}	10^5 $M^{-1}s^{-1}$	1 μm s^{-1}
mks	$6x10^{21}$ m^{-3}	10^{-11} m^2s^{-1}	1 s^{-1}	10^2 $m^3mol^{-1}s^{-1}$ $1.7x10^{-22}$ m^3s^{-1}	10^{-6} m s^{-1}
cgs	$6x10^{15}$ cm^{-3}	10^{-7} cm^2s^{-1}	1 s^{-1}	$1.7x10^{-16}$ cm^3s^{-1}	10^{-4} cm s^{-1}
μm-ms	6000 μm^{-3}	10^{-2} μm^2ms^{-1}	10^{-3} ms^{-1}	$1.7x10^{-7}$ μm^3ms^{-1}	10^{-3} μm ms^{-1}
μm-s	6000 μm^{-3}	10 μm^2s^{-1}	1 s^{-1}	$1.7x10^{-4}$ μm^3s^{-1}	1 μm s^{-1}
nm-ms	$6x10^{-6}$ nm^{-3}	10^4 nm^2ms^{-1}	10^{-3} ms^{-1}	170 nm^3ms^{-1}	1 nm ms^{-1}
nm-μs	$6x10^{-6}$ nm^{-3}	10 $nm^2\mu s^{-1}$	10^{-6} μs^{-1}	0.17 $nm^3\mu s^{-1}$	10^{-3} nm μs^{-1}
px-ms	$6x10^{-3}$ px^{-3}	100 px^2ms^{-1}	10^{-3} ms^{-1}	0.17 px^3ms^{-1}	0.1 px ms^{-1}

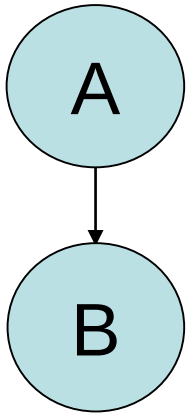
Notes: A pixel, abbreviated px, is defined as a length of 10 nm. In the concentration column, '6' is short for 6.022045. In the bimolecular reactions column, 1.7 is short for 1.660565.

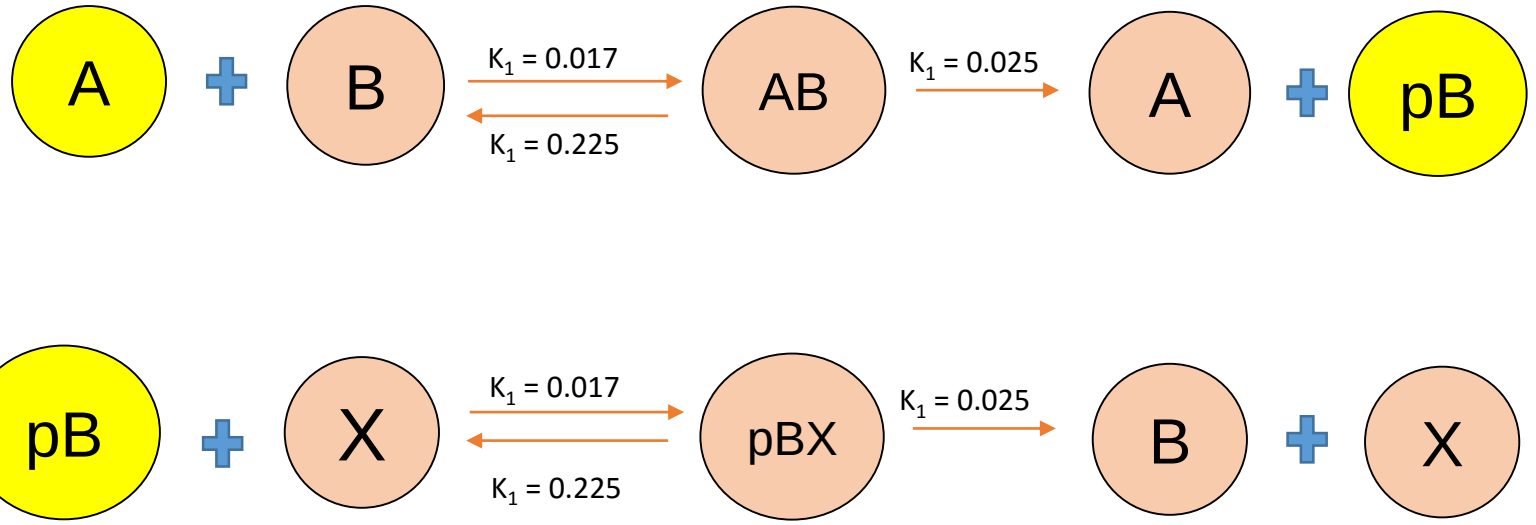
Ref: page no. 12

Using Smoldyn

The Reaction Scheme We will Be Using as Example

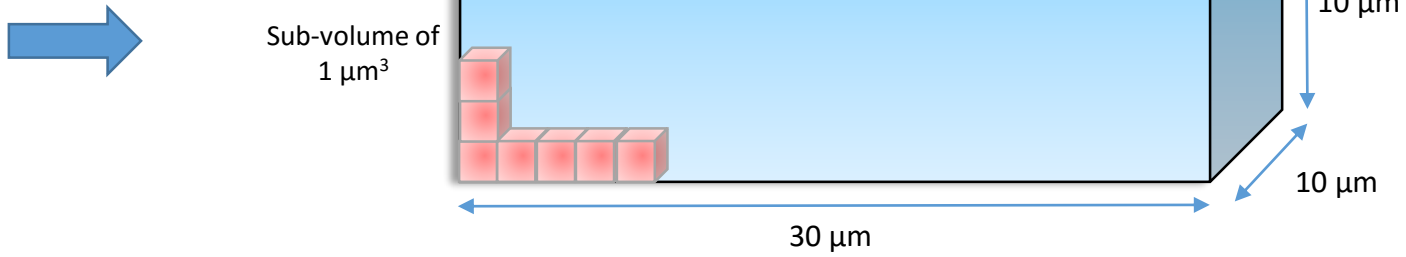
A phosphorylates B





Description of our system

- The cell is modelled as a cuboid
- Our System is 3D
- Dimension : 30 x 10 x 10 μm^3
- Sub-volume : 1μm^3



Using Smoldyn

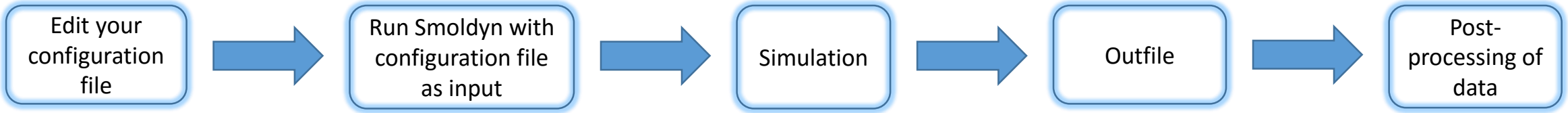
From example to simulation

Smoldyn requires a configuration file as input file

where you define the parameters of the system that is going to be simulated like

- Dimensions
- Molecular species
- Shape of the system
- Reaction
- Reaction rate

Workflow

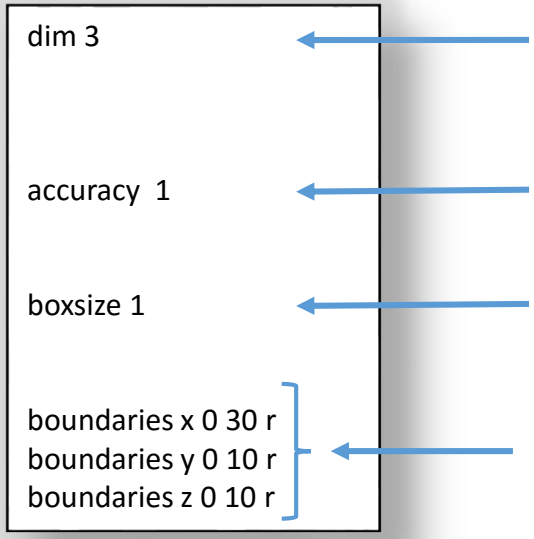


Configuration file is a plain text file. It should end with a .txt extension

Using Smoldyn

How to write configuration file based on our example reaction

Part 1



Define the dimension of the system, dim 2 for 2D. Only 1,2,3 dimension supported

Define the accuracy of the simulation, ranges from 1-10. For qualitative set to 1, for quantitative set to 10.

Define size of the virtual boxes into which the system is divided for speeding up reactions, here 1 µm.

Define dimensions of the system. For 2-D only x,y required.

Format :	boundaries	dimension(x,y,z)	starting_coordinate	ending_coordinate	type_of_system_wall
Here:	boundaries	x	0	30	r (reflective)

What accuracy is?

The accuracy statement sets which neighboring boxes are checked for potential bimolecular reactions.

Ref: Page no. 68

Boundaries may be reflective, transparent, absorbing, or periodic.

- Reflective means that all molecules that diffuse into a boundary will be reflected back into the system.
- Transparent, which is the default type, means that molecules just diffuse through the boundary as though it weren't there.
- With absorbing boundaries, any molecule that touches a boundary is immediately removed from the system.
- Periodic boundaries, which are also called wrap-around or toroidal boundaries, any molecule that diffuses off of one side of space is instantly moved to the opposite edge of space; these are useful for simulating a small portion of a large system while avoiding edge effects.

Ref: Page no. 15

Using Smoldyn

Part 2

```
species A B AB pB X pBX

molecule_lists A_list B_list AB_list pB_list X_list pBX_list

mol_list A A_list
mol_list B B_list
mol_list AB AB_list
mol_list pB pB_list
mol_list X X_list
mol_list pBX pBX_list

color A red
color B yellow
color AB black
color pB blue
color X green
color pBX black
```

Define the species present in the system

Define the names of molecule lists

Define assignment of molecule to a list

Why ?

The separation of the molecules into these two lists speeds up the simulation because all molecules in fixedlist can be safely ignored during diffusion calculations or surface checking

Ref: Page no. 22

Define the colors associated with each species in the system

Using Smoldyn

Part 3

```
difc A 10
difc B 10
difc AB 10
difc pB 10
difc X 10
difc pBX 10

time_start 0
time_stop 1200
time_step 0.01

mol 5000 A u u u
mol 50000 B u u u
mol 0 AB u u u
mol 0 pB u u u
mol 25000 X u u u
mol 0 pBX u u u

reaction R1 A + B -> AB 0.0289
reaction R2 AB -> A + B  0.225
reaction R3 AB -> A + pB 0.025
reaction R4 pB + X -> pBX 0.0289
reaction R5 pBX -> pB + X 0.225
reaction R6 pBX -> B + X 0.025
```

Define the diffusion coefficient of each individual species in the system. **Ref: Page no. 19**

Define the simulation start time, end time and the time step of the simulation. Here, 1200 – 0 = 1200 secs = 20 mins. The unit is seconds.

Define the initial number of molecules per species present in the system.
Also the distribution of the species is defined here. Here, "u" means the molecules are randomly spread across the system according to the uniform probability distribution.
Format: mol no. of_molecules_initially species_name distribution_across_the_3_dimensions

Ref: Page no. 19

Define each individual reactions taking place in the system.
Format: reaction reaction_name define_reaction reaction_rate **Ref: Page no. 43 - 63**

Using Smoldyn

Part 4

```
output_files OUTFILE

cmd e molcountspecies pC OUTFILE




end_file
```

Define the name of the output file to which results / observation of the simulation are to be written

Define the command to output the desired information to the output file

Two Parts of the command

cmd e

Commands could be used to measure an aspect of the simulated system at any time

The character following cmd is the command type,

- 'b' for before the simulation,
- 'a' for after the simulation,
- 'e' for every time step during the simulation,
- '@' for a single command execution at time *time*,
- 'n' for every *n*'th iteration of the simulation,

More commands can be found at **page no. 29 & 82**

molcountspecies pC OUTFILE

Describes which aspect of the simulation to observe

- The above specifically outputs the count of molecules of species of "pC" to the file "OUTFILE"
- The whole list of such commands for various situations are present in **page no. 103**

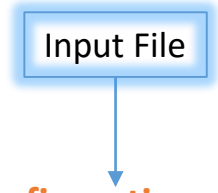
Necessary to acknowledge the end of the configuration file

Using Smoldyn

Complete Configuration File

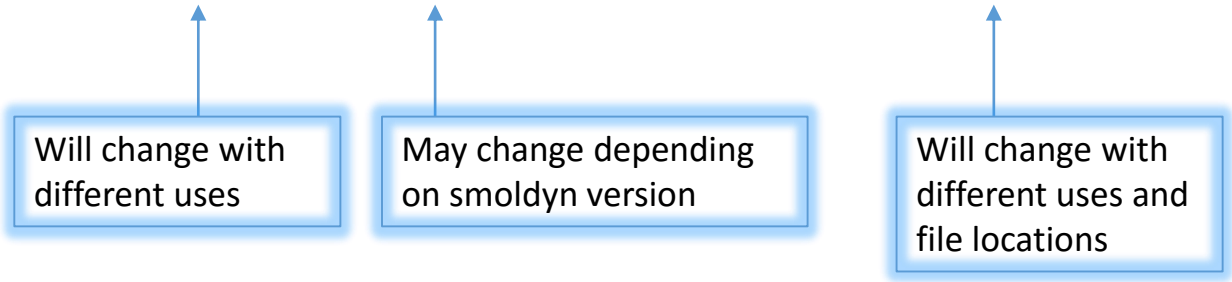
```
dim 3

accuracy  1

boxsize 1

boundaries x 0 30 r
boundaries y 0 10 r
boundaries z 0 10 r

species A B AB pB X pBX

molecule_lists A_list B_list
AB_list pB_list X_list pBX_list

mol_list A A_list
mol_list B B_list
mol_list AB AB_list
mol_list pB pB_list
mol_list X X_list
mol_list pBX pBX_list

color A red
color B yellow
color AB black
color pB blue
color X green
color pBX black
```

```
difc A 10
difc B 10
difc AB 10
difc pB 10
difc X 10
difc pBX 10

time_start 0
time_stop 1200
time_step 0.01

mol 5000 A u u u
mol 50000 B u u u
mol 0 AB u u u
mol 0 pB u u u
mol 25000 X u u u
mol 0 pBX u u u

reaction R1 A + B -> AB 0.0289
reaction R2 AB -> A + B  0.225
reaction R3 AB -> A + pB 0.025
reaction R4 pB + X -> pBX 0.0289
reaction R5 pBX -> pB + X 0.225
reaction R6 pBX -> B + X 0.025

output_files OUTFILE
cmd e molcountspecies pC OUTFILE

end_file
```



- To run smoldyn use the command **smoldyn configuration_file**
- In servers, the complete path to smoldyn executable file as well as the configuration file need to be specified as the smoldyn is not installed on the root folder

  /home/user-name/smoldyn-2.39/smoldyn /home/user-name/configuration_file



- For graphical simulation:
  add **graphics opengl** to the beginning of the configuration file **ref: page no. 25**
- For comments use “#”
- To stop simulation mid-way use **CTRL + C**

Using Smoldyn

Example output file

Time Step Molecule count

```
0 0
0.01 0
0.02 0
0.03 0
0.04 0
0.05 0
0.06 0
0.07 0
0.08 0
0.09 0
0.1 0
0.11 0
0.12 0
0.13 0
0.14 0
0.15 0
0.16 0
0.17 0
0.18 0
0.19 0
0.2 0
0.21 0
0.22 0
0.23 0
```

Example Graphics Window

Example Errors Possible

Errors:
file not found: all files (config file, readfiles) in current directory?
file not found: exact name of file, incl. extensions (.txt) - not always displayed
file cannot be read: line endings - Unix, Windows or Mac

Warnings:
boxsize / molperbox too small: will lead to errors
boxsize / molperbox too large: inefficient
timesteps too large for fast reactions

Param-Ishan

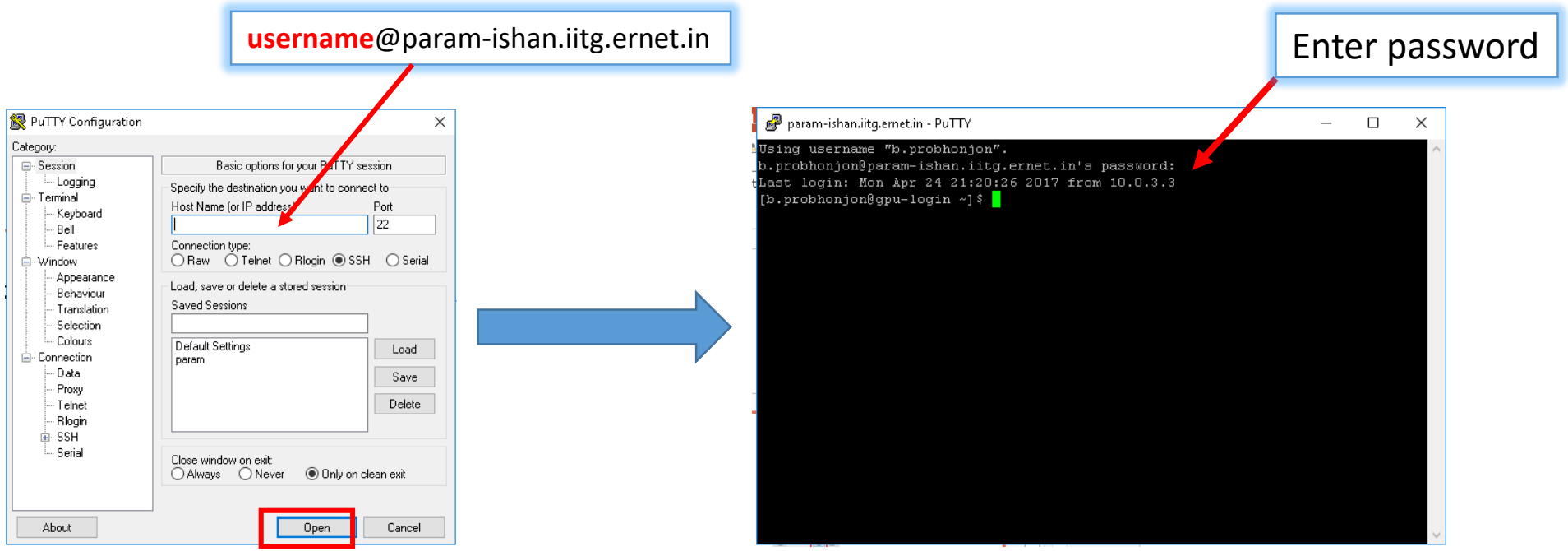
- 'Param-Ishan' is the Supercomputing facility launched at IITG
- ~ 250 Teraflops Peak computing performance
- Request for Param-Ishan account can be made at https://intranet.iitg.ernet.in/param-ishan/user/index.php
- Each user gets 30GB default quota
- The home folder of each user is **/home/user-name**
- Total number of CPUs per user must not exceed 100

Accessing Param-Ishan

What you need:

Username
Password
} *From HPC support*

Software : Putty (download link : http://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html)

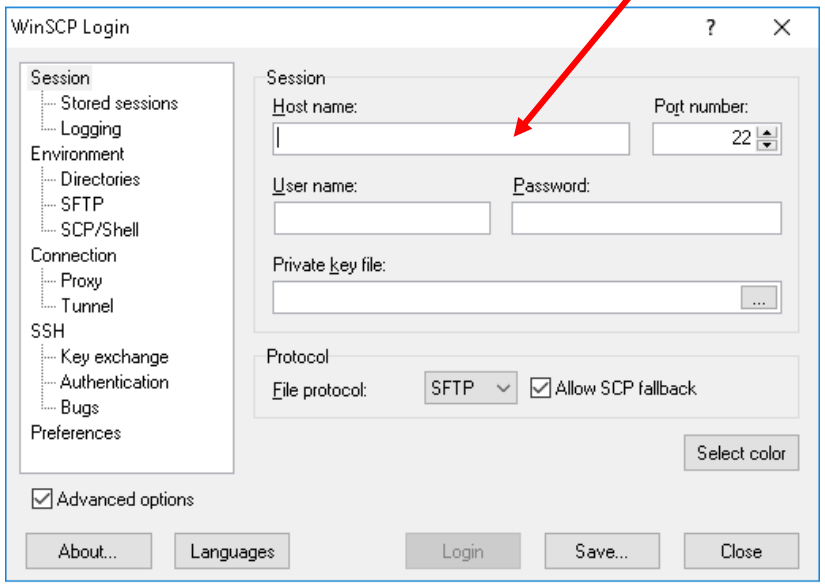


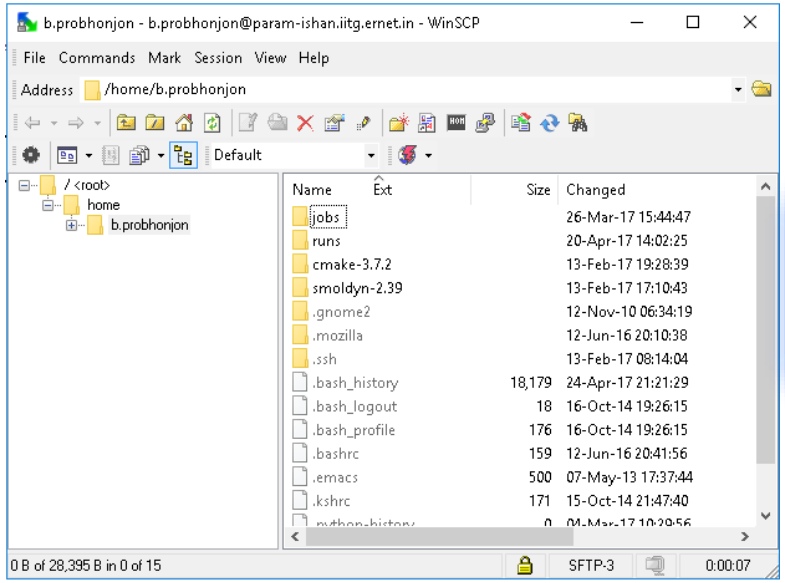
Uploading Files to Param-Ishan

What you need:

Software : WinSCP (download link : https://winscp.net/eng/download.php#download2)

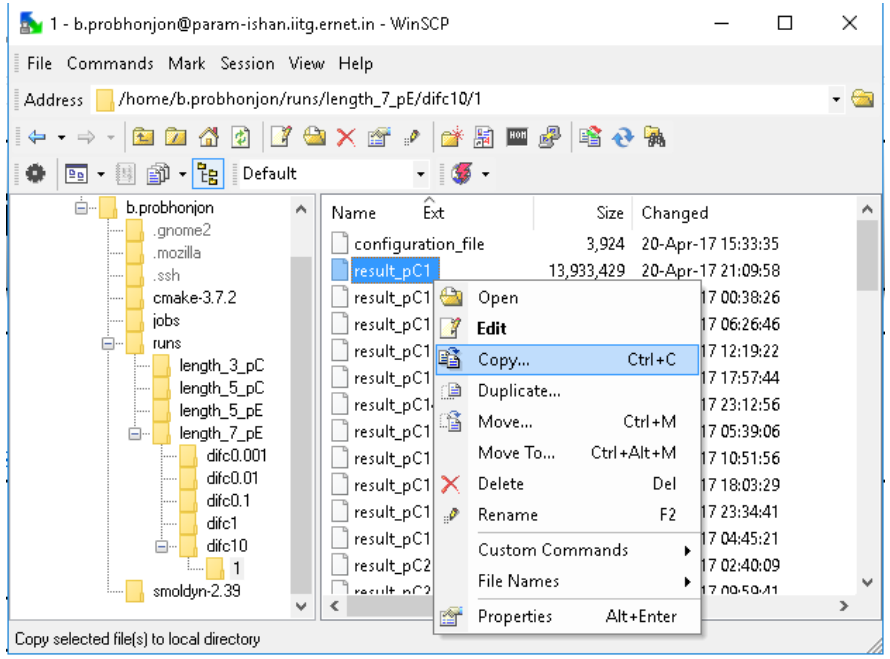
param-ishan.iitg.ernet.in



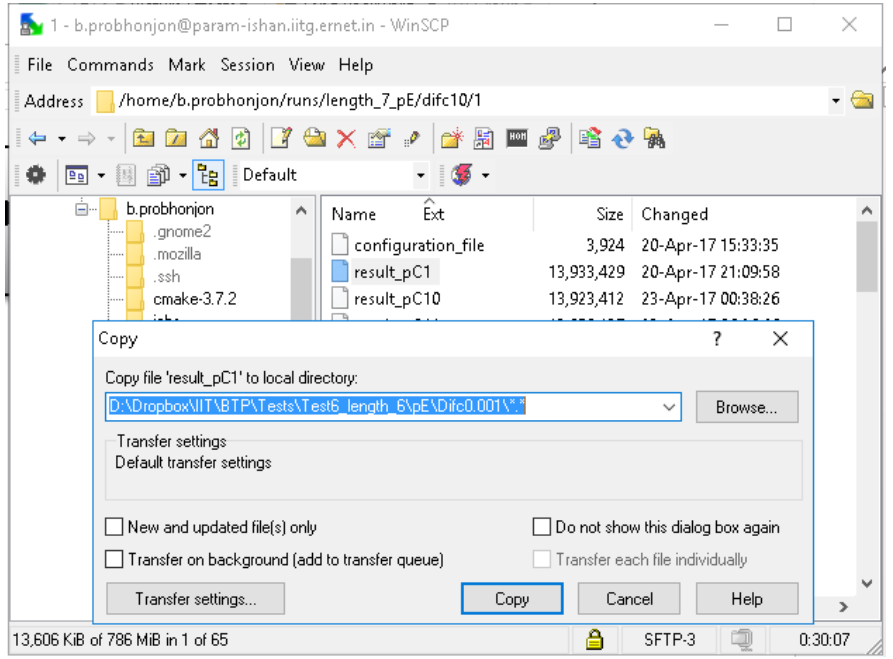


Use drag and drop method to transfer

Downloading Files From Param-Ishan



Right click on the desired file / files and select “copy”



Browse to the folder where you want to copy the file, then click on “copy”

Installing Smoldyn

1. On your local computer *download "smoldyn" (download link: http://www.smoldyn.org/download.html) and extract and upload to home folder of server*
2. On your home folder on the server (or on your local computer for local installation)
   - check if "cmake" is installed on the server (probably not, may be present on Ubuntu 14.04+ on pc) using *cmake - -version*
   - if cmake is not installed else move to step 2.
     - On your Computer download "cmake" (download Link: https://cmake.org/download/, I used cmake-3.7.2) *and extract on your computer*
     - upload extracted "cmake" folder to server *using WinSCP*
     - change directory to cmake folder using the command cd cmake
     - then run the following commands
       - chmod 777 * [1] [2]
       - ./bootstrap
       - make
       - make install=/home/username
3. On your home folder on the server (or on your local computer for local installation)
   - change directory to "smoldyn" folder using the command cd smoldyn-2.39
   - then run the following commands
     - /home/username/cmake-3.7.2 -DOPTION_USE_OPENGL=OFF -DOPTION_USE_LIBTIFF=OFF
     - make
     - make install=/home/username

2.39 in **smoldyn-2.39** depends on software version, check name of folder before using the command

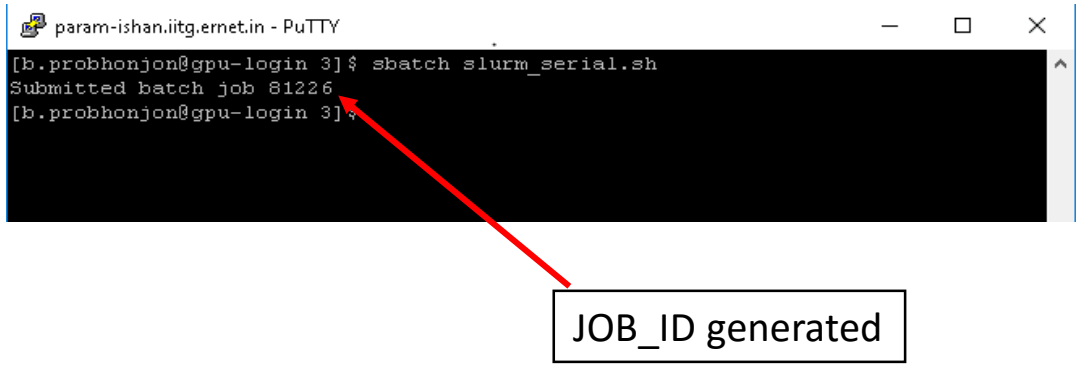
-DOPTION_USE_OPENGL=OFF & -DOPTION_USE_LIBTIFF=OFF are needed because opengl (needed for graphics) is not supported on servers, these commands are not needed in case of local computers and you can just use cmake ..

Submitting a job on a server (Param-Ishan)

Some Basic Commands

All commands follow the following format COMMAND –[OPTIONS] -ARGUMENT

Description	Commands
Submit job	sbatch filename
List all jobs running in the server	squeue
Show the Estimated Start Time of a Job	squeue -j JOB_ID –start
Show all job information for a specific user	sacct -u user-name
Show job account information for a specific job	sacct -j JOB_ID
Users can kill their own jobs	scancel JOB_ID



More details of submitting job can be found here : http://intranet.iitg.ernet.in/param-ishan/files/w-11-16/SLRUM.pdf

Submitting Smoldyn as a job on Param-Ishan

What you need:

1. Configuration File
2. Slurm file : needed for specifying the parameters of job

Example slurm file

```
#!/bin/bash
#SBATCH --job-name="b.probhonjon_exp2"
#SBATCH --nodes=1
#SBATCH --tasks-per-node=8
#SBATCH --partition=standard
#SBATCH --output/home/user-name/smoldyn-2.39/cmake/smoldyn/slurm.%N.%j.out # STDOUT
#SBATCH --error=/home/user-name/smoldyn-2.39/cmake/smoldyn/slurm.%N.%j.err # STDERR
#SBATCH --mail-user=b.probhonjon@iitg.ernet.in
#SBATCH --mail-type=ALL

/home/user-name/smoldyn-2.39/cmake/smoldyn /home/user-name/configuration_file
```

Stick to default if you find it confusing

May vary based

- on the type of project you are doing
- No. of repetitions of the simulation
- The directory in which the simulation is going to be run

Resource	Flag Syntax	Description
partition	--partition=partition name	Partition is a queue for jobs.
time	--time=01:00:00	Time limit for the job.
nodes	--nodes=2	Number of compute nodes for the job.
cpus/cores	--ntasks-per-node=8	Corresponds to number of cores on the compute node.
resource feature	--gres=gpu:2	Request use of GPUs on compute nodes
account	--account=group-slurm-account	Users may belong to groups or accounts.
job name	--job-name="hello_test"	Name of job.
output file	--output=test.out	Name of file for stdout.
email address	--mail-user=username@iitg.ac.in	User's email address
Access	--exclusive	Exclusive acccess to compute nodes.

Max no. of cpu per user = 100 on param-ishan

TIPS

Some tips from experience

- Be careful with units and their conversions.
- For parameters such as accuracy, boxsize, molperbox, etc try a few different values for each and their combinations and check which combinations match expected results before deciding on the final value.
- Smoldyn simulations are slow. If the simulation is to be repeated many times (say 100) it is advisable to split it up into small batches (5 batches of 20 each) and submit each batch as a separate job on the server.
- Initially use the graphical output of the simulation to confirm that the simulation is proceeding as expected. Keep the duration of such simulation short as a personal PC may not be able to handle large simulations.

Smoldyn is good for spatial, stochastic sytems:

- single-molecule studies
- diffusion in restricted environments
- stochastic noise, arising from spatial effects
- intra- or inter-cellular signalling
- effects of localised proteins

Smoldyn is not recommended for

- systems that are always well-mixed
- systems that are big enough for differential equations to be adequate
- systems that cannot be described with elementary interactions

Useful links

- ftp://ftp.ebi.ac.uk/pub/training/2016/In_Silico_Systems_Biology_July_2016/DAY%205%20(Thursday)/Karen_Lipkow/SmoldynWT/Presentations/
- http://intranet.iitg.ernet.in/param-ishan/files/w-11-16/SLRUM.pdf
- http://intranet.iitg.ernet.in/param-ishan/files/w-11-16/slrum_ex.zip
- http://intranet.iitg.ernet.in/param-ishan/files/PARAM-ISHAN_User_Guide_Ver1.0.pdf
- http://intranet.iitg.ernet.in/param-ishan/users.html