Hands-on activity worksheet for Thursday (July 4th) afternoon: Stochastic Modelling and the Stochastic Multi-Cloud Model

5th Summer School on Theory, Mechanisms and Hierarchical Modelling of Climate Dynamics: Convection and Clouds

July 3, 2024

1 Theoretical teasers:

- Let T_1, T_2 be two independent exponentially distributed random variables with rates $\lambda > 0, \mu > 0$, respectively.

  (a) Show that $S = \min(T_1, T_2)$ is an exponential random variable with rate $\lambda + \mu$.

  (b) Show that $P\{T_1 < T_2\} = \frac{\lambda}{\lambda+\mu}$.

  (c) Show that if $\mu = \lambda$ then $T_1 + T_2$ is a Gamma random variable with parameters $\alpha = 2$ and λ and in general if $T_1, T_2, \cdots, T_n$ are n independent and exponentially distributed random variables, with the same rate λ, then the sum $S_n = T_1 + T_2 + \cdots T_n$ is Gamma distributed with parameters n and λ; $f_{S_n}(x) = x^{n-1}\lambda^n e^{-\lambda x}/(n-1)!$.

- Write down the forward and backward equations for a bounded birth death process with birth rates λ_n and death rates μ_n where $\mu_0 = 0$ and $\lambda_k = 0$ for $k \geq 1$. Give the infinitesimal generator matrix.

- Find the transition probabilities for a birth only process, i.e, a birth-death process for which $\lambda_n > 0$ and $\mu_n = 0$ for all n. Start with the case $\lambda_n = \lambda$, i.e, the birth rate is independent of n.

2 Experiments with the random walk

It is known in probability theory that the symmetric random walk in 2d (on an infinite lattice, i.e, on $\mathbb{Z}^2$) is recurrent, i.e, if the process is started at the origin then the probability for it to return to the same point (some time in the far future) is unity. This is not the case for the random walk in three dimensions and higher. For the 3d symmetric random walk the probability of return is approximately 0.35. Write a computer code[1] to simulate the random paths of the symmetric random walk in 2d, using Monte Carlo sampling, and try to demonstrate the above result. To do so you can make a sufficiently large ensemble of runs, i.e simulated paths, all of the same length, call it $Nstep$, each starting at the origin. For clarity let us call the ensemble size $Nsample$. Compute the frequency of returns to the origin for the given ensemble by counting how many of the ensemble members have actually returned in Nstep steps. Consider and increasing sequence of experiments corresponding to a sequence of $(Nstep, Nsample)$ values to see whether the frequencies tend towards unity. Better results are obtained with larger Nstep values while Nsample can be kept relatively small. As a starting point set Nstep= 1000 and vary Nsample. At first you can try increasing values of Nsample (like 10, 100, 1000) while Nsample is kept the same. Then, increase Nstep and repeat the experiment. Conclude.

Hint: The experiments tend to be more effictive when Nsample is small and Nstep is large.

Other useful games to play with the random walk code:

1. The random walk is known to be a diffusive process and in fact a good model for the phenomenon of diffusion and known to converge to the heat equation when properly scaled. The distribution of the position of the process $X(t)$ at any given time $t > 0$, quickly converges to a Gaussian distribution as t increases. For the symmetric Random walk, the mean of $X(t)$ remains always zero and the standard deviation increases as $t^{1/2}$, i.e, $\sigma^2(t) \sim t$. The continuous time version of the Random walk is in fact the Brownian motion and there are specific algorithms on how to sample this process.

The features above and other similarities with the Brownian motion can in principle be easy demonstrated using numerical experiments. One important feature of the Brownian motion is that it has independent increments and that $X(t_2) - X(t_1)$ is a Gaussian distributed with mean zero and variance $\sim (t_2 - t_1)^{1/2}$. This also can be demonstrated.

2. Incidentally, asymmetric Random walks converge to a Brownian motions with a drift, i.e, the similar process except that the mean is not zero; it drifts/varies linearly with time.

3. Curious to experiment with the 3D case? You can also adopt the code to 3d and demonstrate non recurrence for the symmetric random walk in 3d. Convincingly establishing convergence of the return frequency to the number 0.35 can be hard though!

[1] Python code provided

3 Standalone SMCM code

Experiment with the standalone SMCM Python code provided to you. The file `driver_stochastic_local_int.py` contains the main driver and all intrernal dependencies/routines. It is a good idea to first look through the code and try to get familiar with its structure and understand which part of the code does what. In particular you may identify some key parameters such as

- the external controls (CAPE, low level CAPE, atmospheric dryness),
- internal dynamics parameters such as the transition time scales (τ_{kl}), and the local interaction matrix J_0
- the resolution parameters consisting of the size of the microscopic lattice n and the coarsening parameter q
- the integration time, $Tend$, and the time step, dt used to solve the mean field equations.

To run the code you must download all the dependencies and type python3 `driver_stochastic_local_int.py`. After the program finishes running it will produce five figures and a text file. They are Meanfield Plot.png, Micro MultiState Process Grid.png, Micro MultiState Process Time.png, Coarse Grained Plot Grid.png, Coarse Grained Plot Time.png, stochasticlocal.txt

Now that you are familiar with the code, there are many ways to play and experiment with the code. You can for instance try to see how sensitive the model is to any of the key parameters listed above.

- You can begin by playing with the parameters τ_{kl}. You can for instance try to anticipate which τ_{kl} should be changed and how (up or down) in order to promote a certain desire, such as a dominant cloud fraction of one particular cloud type (say deep or maybe stratiform).
- Similarly you can try to see how sensitive the model is to the entries of J_0. You can see how certain configurations of J_0 lead to more or less clumpy-ness (self- organization) of one or more cloud types. **The matrix J_0 can be used as a medium to incorporate into the SMCM the effect of land-sea (temperature and or moisture) contrast on convective organization or how a background shear can organize convection in the form of MCS's. The latter is the subject of at least one project for Week 3, with the goal of using observed (i.e reanalysis) shear profiles during MCS activity to inspire a choice of J_0 that would lead to an MCS like behaviour in the standalone SMCM.**

- It is also interesting, from the modelling point of view, to see how changing the number of lattice sites and level of coarsening affects, one way or another, the behaviour of the model.

- Finally, perhaps the ultimate goal is to explore to what extent the external controls actually control the behaviour of the small scale dynamics, say the area fraction distribution and make diagrams like in the figure below. This particular task can be extended to a project for Week 3 (if there is interest). The project would consists in first produce the equivalent diagrams from observations (using radar data for the rain-type area fractions as a proxy for the cloud type area fractions combined with corresponding controls from reanalysis data).