Parameter Inference for a Simple Stochastic Hydrological Model

C. Albert*, S. Ulzega

Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

March 9, 2015

In this paper, we study the feasibility of a full Bayesian inference, for simple conceptual rainfall-runoff models with scale-invariant noise. We start with a single linear reservoir and describe the dynamics of its water content, $S(t)$, by means of the stochastic differential equation

$$\dot{S}(t) = r(t) - \frac{1}{K}\left(1 + \frac{\gamma}{2}\right) S(t) + \sqrt{\frac{\gamma}{K}} S(t)\eta(t)\,, \tag{1}$$

where $r(t)$ describes the rain input and $\eta(t)$ denotes white noise, i.e.,

$$\langle \eta(t)\eta(t') \rangle = \delta(t - t')\,. \tag{2}$$

Given rainfall and runoff time-series, we want to infer the two system parameters K and γ, measuring, respectively, the retention time and the noise. The scale-invariance of the noise term renders the noise parameter dimensionless. Due to the state-dependence of the noise term, eq. (1) is ill-defined unless we specify the discretization convention. For computational reasons, we work with the *Stratonovich* convention. Our parametrization is such that, for constant rain input $r(t) = r$ and in the long-time limit, the mean of S converges to the limit of the unperturbed ($\gamma = 0$) system Kr (see below).

According to Lau and Lubensky (2007), the probability, $P(S, T|S_0, t_0)$, of finding the system in a state S at time $t = T$ given that it was in an initial state S_0 at time $t = 0$, can be expressed in the form of a path integral as

$$P(S, T|S_0, t_0) = \int_{S_0}^{S} e^{-\tilde{\mathcal{S}}[\dot{S}(t), S(t)]} \mathcal{D}S\,, \tag{3}$$

where the integral extends over all paths starting at S_0 and ending at S. The path-measure $\mathcal{D}S$ is the formal limit

$$\mathcal{D}S = \lim_{\Delta t \to 0} \prod_i \frac{dS_i}{\sqrt{2\pi \Delta t \gamma / K S_i}}\,, \tag{4}$$

for a discretization $S_i = S(t_i)$ with $t_i - t_{i-1} = \Delta t$. The *action* is a functional defined on paths $S : [0, T] \longrightarrow \mathbb{R}_+$ and reads as

$$\tilde{\mathcal{S}}[\dot{S}(t), S(t)] = \int_0^T dt \left\{ \frac{K}{2\gamma S^2(t)} \left(\dot{S}(t) - r(t) + \frac{1 + \gamma}{K} S(t) \right)^2 - \frac{2 + \gamma}{4K} \right\}. \tag{5}$$

*Email: carlo.albert@eawag.ch

Note that the action includes the Jacobian that arises when changing coordinates from $\eta(t)$ to $S(t)$ and ensures the normalization condition $\int P(S,T|S_0,t_0)dS = 1$.

It is possible to calculate the equilibrium distribution $P_{\rm eq}(S) = \lim_{T\to\infty} P(S,T|S_0,t_0)$ in the simple case of a constant rain input r_0. Systems at thermal equilibrium must fulfill detailed balance, i.e.,

$$P(S_1t_1|S_0t_0)P_{\rm eq}(S_0) = P(S_0t_1|S_1t_0)P_{\rm eq}(S_1)\,. \tag{6}$$

Assuming a solution of the form

$$P_{\rm eq}(S) \propto e^{-\mathcal{H}(S)}\,, \tag{7}$$

equation (6) is rewritten as

$$P(S_1t_1|S_0t_0) = P(S_0t_1|S_1t_0)\exp\left\{-\int_{t_0}^{t_1} dt \frac{dS}{dt}\frac{d\mathcal{H}}{dS}\right\}\,, \tag{8}$$

where $S(t)$ is an arbitrary path connecting S_0 with S_1. After plugging in (3) and (5), and using the Stratonovich property, which allows us to transform a path from S_0 to S_1 into a path from S_1 to S_0 through the transformation $S(t) \to S(-t)$, we get

$$-r_0 + \frac{1+\gamma}{K}S = \frac{\gamma S^2}{2K}\frac{d\mathcal{H}(S)}{dS}\,. \tag{9}$$

The solution to (9) is, up to an additive constant, given by eq.

$$\mathcal{H}(S) = \frac{2Kr_0}{\gamma S} + \frac{2(1+\gamma)}{\gamma}\ln(S)\,. \tag{10}$$

Thus, it turns out that $P_{\rm eq}(S)$ is described by an inverse gamma distribution with scale parameter $2Kr_0/\gamma$ and shape parameter $(2+\gamma)/\gamma$, i.e.,

$$P_{\rm eq}(S) \propto S^{-2(1+\gamma)/\gamma} e^{-2Kr_0/(\gamma S)}\,, \tag{11}$$

whose mean equals the equilibrium solution of the unperturbed system ($\gamma = 0$),

$$\langle S\rangle_{\rm eq} = Kr_0, \tag{12}$$

and whose variance, for $\gamma < 2$, is given by

$$\langle (S - \langle S\rangle_{\rm eq})^2\rangle_{\rm eq} = K^2 r_0^2 \frac{\gamma}{2-\gamma}\,. \tag{13}$$

Note that the variance diverges, for $\gamma \geq 2$. The power-law decay of the inverse gamma distribution is reminiscent of the scale-invariance of the error model.

Let us consider now the general case of a time dependent rain input $r(t)$ with the following time-series of *observed* outputs,

$$y_i = \frac{1}{K}S(t_i) + \sigma\epsilon_i\,, \quad i = 1,\ldots,n+1\,, \tag{14}$$

where the times t_i are chosen so that $0 = t_1 < t_2 < \cdots < t_{n+1} = T$ and the ϵ_i are uncorrelated standard normal random variables. The *likelihood* function for the data (14) is, up to a factor that doesn't depend on the model parameters, given by the path-integral

$$f(\mathbf{y}|\boldsymbol{\theta}) \propto \int \exp\left[-\tilde{\mathcal{S}}[\dot{S}, S] - \frac{1}{2}\sum_{i=1}^{n+1} \frac{(y_i - S(t_i)/K)^2}{\sigma^2} \right] \mathcal{D}S(\boldsymbol{\theta})\,, \tag{15}$$

with the parameter vector $\boldsymbol{\theta} = (K, \gamma)^T$. Choosing for the parameters K and γ the improper prior given by

$$f(\boldsymbol{\theta}) = f(K)f(\gamma) = \frac{1}{K\gamma}\,, \tag{16}$$

parameter inference means generating a sufficiently large sample of parameter vectors from the posterior

$$f(\boldsymbol{\theta}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})\,. \tag{17}$$

We replace the variable $S(t)$ and the parameter K by the dimensionless quantities, $q(t)$ and β, respectively, by means of the transformations

$$\beta = \sqrt{\frac{T\gamma}{K}}\,, \quad S(t) = \frac{T\gamma r(t)}{\beta^2} e^{\beta q(t)}\,. \tag{18}$$

In terms of these dimensionless variables and parameters the posterior reads (without prior)

$$f(\boldsymbol{\theta}|\mathbf{y}) \propto \int \exp\left[-\mathcal{S}[\dot{q}, q] - \frac{1}{2}\sum_{i=1}^{n+1} \frac{(y_i - r_i e^{\beta q_i})^2}{\sigma^2} \right] \mathcal{D}q\,, \tag{19}$$

where r_i and q_i represent $r(t)$ and $q(t)$, respectively, at time t_i. In the transformed variables the path-measure takes the form, up to a parameter-independent multiplicative factor,

$$\mathcal{D}q = \prod_t dq(t)\,, \tag{20}$$

while the action is rewritten in the compact form,

$$\mathcal{S}[\dot{q}, q] = \frac{1}{T}\int_0^T dt \left\{ \frac{1}{2}\left(T\dot{q}(t) + \rho(t) - \frac{\beta}{\gamma} e^{-\beta q(t)} \right)^2 - \frac{(2+\gamma)\beta^2}{4\gamma} \right\}\,, \tag{21}$$

with

$$\rho(t) = \frac{T\dot{r}(t)}{\beta r(t)} + \frac{\beta(\gamma+1)}{\gamma} = \frac{T}{\beta}\frac{d\left[\ln r(t)\right]}{dt} + \frac{\beta(\gamma+1)}{\gamma}\,. \tag{22}$$

In order to ease the numerical computations, it is useful to introduce the time-dependent Hamiltonian,

$$\mathcal{H}_q(q, t) = \frac{1}{\gamma} e^{-\beta q} + q\rho(t)\,, \tag{23}$$

and rewrite the action (21) as,

$$\mathcal{S}[\dot{q},q] = \frac{1}{T}\int_0^T dt\left\{\frac{1}{2}T^2\dot{q}^2(t) + \frac{1}{2}\left(\rho(t) - \frac{\beta}{\gamma}e^{-\beta q(t)}\right)^2 - T\frac{\partial\mathcal{H}_q(q,t)}{\partial t} - \frac{(2+\gamma)\beta^2}{4\gamma}\right\}$$
$$+ \mathcal{H}_q(q(T),T) - \mathcal{H}_q(q(0),0)$$
$$= \frac{1}{T}\int_0^T dt\left\{\frac{1}{2}T^2\dot{q}^2(t) + \frac{1}{2}\left(\rho(t) - \frac{\beta}{\gamma}e^{-\beta q(t)}\right)^2 - Tq(t)\dot{\rho}(t) - \frac{(2+\gamma)\beta^2}{4\gamma}\right\}$$
$$+ \frac{1}{\gamma}e^{-\beta q(T)} + q(T)\rho(T) - \frac{1}{\gamma}e^{-\beta q(0)} - q(0)\rho(0)\,. \quad (24)$$

Since the integral on the r.h.s. of (19) is very expensive to calculate, we will sample simultaneously parameter vectors $\boldsymbol{\theta}$ and discretized system realizations $q(t)$ directly from an appropriate discretization of the kernel of (19). Since this distribution is very high dimensional and strongly correlated, we employ the method called *Hamiltonian Monte Carlo* (HMC) introduced by Duane et al. (1987). This method interprets the exponent of the kernel of the posterior as the potential of a one-dimensional statistical mechanical system. Each degree of freedom, $q(t)$ and $\boldsymbol{\theta}$ in our case, is paired with a conjugate variable, $p(t)$ and $\boldsymbol{\pi}$, respectively, and the system is defined by the Hamiltonian

$$\mathcal{H}_{\mathrm{HMC}}(q,\boldsymbol{\theta};p,\boldsymbol{\pi}) = K(p,\boldsymbol{\pi}) + V(q,\boldsymbol{\theta})\,, \quad (25)$$

where

$$K(p,\boldsymbol{\pi}) = \int_0^T \frac{p^2(t)}{2m_q}dt + \sum_{\alpha=1}^{2}\frac{\pi_\alpha^2}{2m_\alpha}\,, \quad (26)$$

where $V(q,\boldsymbol{\theta})$ is the negative logarithm of the kernel of (19). The HMC method, which is a combination of the *Metropolis algorithm* by Metropolis et al. (1953) and *molecular dynamics* methods **[ref?]**, iterates the following steps

1. Momenta are sampled from the Gaussian distribution defined by (26).
2. A volume-preserving and time-reversible solution to a discretization of Hamilton's equations is calculated numerically (leap-frog algorithm).
3. The discretization error on the energy preservation due to the previous step is corrected by a Metropolis acceptance/rejection step.

The last step is the standard Metropolis algorithm, while the first two steps allow to make relatively large jumps in phase space while maintaining a relatively large acceptance rate.

Let us first write down the proper discretization of the path-integral (19), with the action (24). Therefore, let us assume that the measurement time points $\{y_i\}_{i=1,\ldots,n+1}$ of the time series (14) are equidistantly distributed on the time interval $[0,T]$, with $t_1 = 0$ and $t_{n+1} = T$. Each interval between two consecutive data points is further partitioned into j bins, such that we have a total of $nj+1 = N >> 1$ discretization points. The discretized versions of $K(p,\boldsymbol{\pi})$ and $V(q,\boldsymbol{\theta})$ read

$$K(p,\boldsymbol{\pi}) \approx \sum_{i=1}^{N} \frac{p_i^2}{2m_q}\Delta t + \sum_{\alpha=1}^{s} \frac{\pi_\alpha^2}{2m_\alpha}\,, \tag{27}$$

$$V(q,\boldsymbol{\theta}) \approx \frac{\Delta t}{T}\sum_{i=2}^{N}\left\{\frac{1}{2}T^2\dot{q}_i^2 + \left(\frac{1}{2}\left(\rho_i - \frac{\beta}{\gamma}e^{-\beta\bar{q}_i}\right)^2 - T\bar{q}_i\dot{\rho}_i\right)\right\} \tag{28}$$

$$- \frac{(2+\gamma)\beta^2}{4\gamma} + \frac{1}{\gamma}e^{-\beta q_N} + q_N\rho_{N-1} - \frac{1}{\gamma}e^{-\beta q_1} - q_1\rho_1$$

$$+ \sum_{s=1}^{n+1}\frac{(y_s - r_{(s-1)j+1}e^{\beta q_{(s-1)j+1}})^2}{2\sigma^2}\,, \tag{29}$$

with

$$\dot{q}_i = \frac{q_i - q_{i-1}}{\Delta t}\,, \quad \bar{q}_i = \frac{q_i + q_{i-1}}{2}\,, \tag{30}$$

and

$$\rho_i = K\frac{\ln(r_{i+1}/r_i)}{\Delta t} + \gamma + 1\,, \quad \dot{\rho}_i = \frac{\rho_i - \rho_{i-1}}{\Delta t}\,. \tag{31}$$

Note that we've neglected terms of order $\mathcal{O}(1/N)$ in the action (29) and that we've assumed $r(t)$ to be at least twice differentiable.

For strongly varying rain input or high-frequency output data, Δt needs to be chosen small. Thus, the harmonic part of the Hamiltonian becomes stiff and, consequently, the time-step of MD needs to be chosen small, which leads to slow convergence. To solve this problem, we change coordinates, such that the harmonic part of the Hamiltonian becomes (partly) diagonal and the particle interaction moves to the potential. One option would be to use the Fourier modes of $\eta(t)$ as configurational variables. Calculating the potential requires then an inversion of a Toeplitz matrix. This approach is interesting because, in our case, it allows for an approximate analytical integration of high-frequency modes (see Doll et al. (1985)).

We choose another approach, the so-called *staging* as described in Tuckerman et al. (1993), and diagonalize the harmonic part only in between the measurement points. We first note that the Hamiltonian $\mathcal{H}_{\mathrm{HMC}}(q,\boldsymbol{\theta};p,\boldsymbol{\pi})$ (25) can be seen as the classical Hamiltonian of a polymer chain with harmonic bonds between neighbouring beads, described by the first harmonic term of $V(q,\boldsymbol{\theta})$, in an external field described by the other non-harmonic terms of $V(q,\boldsymbol{\theta})$. Furthermore, the model parameters are viewed as additional degrees of freedom that couple with all other degrees of freedom. Sampling system realizations, each defined by a set of coordinates $\{q_i\}$ and parameters $\boldsymbol{\theta}$, is equivalent to sampling states in the polymer dynamics.

The discretized harmonic part of the action is rewritten as

$$\sum_{i=2}^{N}\frac{T}{2\Delta t}(q_i - q_{i-1})^2$$

$$= \frac{T}{2}\sum_{s=1}^{n}\left\{\frac{(q_{(s-1)j+1} - q_{sj+1})^2}{j\Delta t} + \sum_{k=2}^{j}\frac{k}{(k-1)\Delta t}(q_{(s-1)j+k} - q^*_{(s-1)j+k})^2\right\}\,, \tag{32}$$

with

$$q^*_{(s-1)j+k} = \frac{(k-1)q_{(s-1)j+k+1} + q_{(s-1)j+1}}{k}\,. \tag{33}$$

Following Tuckerman et al. (1993) we apply the coordinate transformation

$$u_{sj+1} = q_{sj+1}\,, \quad u_{sj+k} = q_{sj+k} - q^*_{sj+k}\,, \tag{34}$$

with its inverse given by

$$q_{sj+1} = u_{sj+1}\,, \quad q_{sj+k} = \sum_{l=k}^{j+1} \frac{k-1}{l-1} u_{sj+l} + \frac{j-k+1}{j} u_{sj+1}\,, \tag{35}$$

or equivalently by the recursive relation

$$q_{sj+k} = u_{sj+k} + \frac{k-1}{k} q_{s+k+1} + \frac{1}{k} q_{sj+1}\,. \tag{36}$$

This transformation decouples fast degrees of freedom from slow ones and allows us to split the Hamiltonian $\mathcal{H}_{HMC}$ into components with different scaling behaviour in the potentially large numbers n and N. We write

$$\mathcal{H}_{HMC} = \mathcal{H}_N + \mathcal{H}_n + \mathcal{H}_1\,, \tag{37}$$

where

$$\mathcal{H}_N = \frac{1}{2}\sum_{s=1}^{n}\sum_{k=2}^{j}\left\{\frac{\Delta t}{m_q} p^2_{(s-1)j+k} + \frac{Tk}{\Delta t(k-1)} u^2_{(s-1)j+k}\right\}\,, \tag{38}$$

$$\mathcal{H}_n = \frac{1}{2}\sum_{s=1}^{n+1}\left\{\frac{\Delta t}{m_q} p^2_{(s-1)j+1} + \frac{T}{j\Delta t}(u_{(s-1)j+1} - u_{sj+1})^2 - \frac{(y_s - r_{(s-1)j+1} e^{u_{(s-1)j+1}})^2}{\sigma^2}\right\}\,, \tag{39}$$

$$\mathcal{H}_1 = \sum_{\alpha=1}^{2}\frac{\pi^2_\alpha}{2m_\alpha} + \frac{\Delta t}{T}\sum_{i=2}^{N}\left\{\frac{1}{2}\left(\rho_i - \frac{\beta}{\gamma} e^{-\beta\bar{q}_i}\right)^2 - T\bar{q}_i\dot{\rho}_i\right\} \tag{40}$$

$$- \frac{(2+\gamma)\beta^2}{4\gamma} + \frac{1}{\gamma} e^{-\beta q_N} + q_N\rho_{N-1} - \frac{1}{\gamma} e^{-\beta q_1} - q_1\rho_1\,. \tag{41}$$

The harmonic part $\mathcal{H}_N$, for the staging beads, scales linearly with N. The harmonic term $\mathcal{H}_n$, for the boundary beads, as well as the measurement term scale linearly with n. Finally, the potential does not scale with neither n nor N. Thus, depending on the values of N and n, the dynamics happens on very different time scales. Therefore, we use the Trotter formula according to Tuckerman et al. (1992) in order to design a reversible molecular dynamics integrator that takes these different time scales into account. In order to design an appropriate partition of the Hamiltonian, we need to distinguish different regimes.

i $\mathcal{H}_N \sim \mathcal{H}_n >> \mathcal{H}_1$

ii $\mathcal{H}_N >> \mathcal{H}_n \sim \mathcal{H}_1$

iii $\mathcal{H}_N >> \mathcal{H}_n >> \mathcal{H}_1$

In regime (ii) we simply separate the harmonic part of the action, for the staging beads, from the rest and write

$$\mathcal{H}_{HMC} = \mathcal{H}_N + \mathcal{H}' \,. \tag{42}$$

In order to design a reversible integrator, for the associated Hamiltonian equations, we define the Liouville operators

$$iL_N = \{\cdot\,,\, \mathcal{H}_N\}\,, \quad iL' = \{\cdot\,,\, \mathcal{H}'\}\,, \tag{43}$$

where $\{\cdot\,,\,\cdot\}$ denotes the Poisson brackets that are defined on functions on the phase space. Trotter's formula allows us to write the Hamiltonian propagator as

$$e^{i(L_N+L')\tau} = (e^{iL_N(\Delta\tau/2)} e^{iL'\Delta\tau} e^{iL_N(\Delta\tau/2)})^P + \mathcal{O}(P^{-2})\,, \tag{44}$$

for $\tau = P\Delta\tau$. In regime (ii) the outer propagator $\exp[iL_N(\Delta\tau/2)]$ describes a much faster dynamics than the inner one. However, it is the dynamics of uncoupled harmonic oscillators, which we can readily solve. Masses and frequencies of the oscillators are derived from (38) as

$$m = m_q/\Delta t\,, \quad \omega_k = \sqrt{\frac{Nk}{(k-1)m}}\,. \tag{45}$$

The fast outer propagator is then given by eqs.

$$u_{(s-1)j+k}(\Delta\tau/2) = u_{(s-1)j+k}(0)\cos(\omega_k\Delta\tau/2) + \frac{p_{(s-1)j+k}(0)}{m\omega_k}\sin(\omega_k\Delta\tau/2)\,, \tag{46}$$

$$p_{(s-1)j+k}(\Delta\tau/2) = p_{(s-1)j+k}(0)\cos(\omega_k\Delta\tau/2) - m\omega_k u_{(s-1)j+k}(0)\sin(\omega_k\Delta\tau/2)\,, \tag{47}$$

for $k = 2, \ldots, j$.

For the inner, slow propagator, we employ the velocity Verlet algorithm. For the boundary beads, it reads

$$u_{(s-1)j+1}(\Delta\tau) = u_{(s-1)j+1}(0) + \frac{\Delta\tau}{m}p_{(s-1)j+1}(0) + \frac{\Delta\tau^2}{2m}F_{(s-1)j+1}[\mathbf{u}(0), \boldsymbol{\theta}(0)]\,, \tag{48}$$

$$p_{(s-1)j+1}(\Delta\tau) = p_{(s-1)j+1}(0) + \frac{\Delta\tau}{2}(F_{(s-1)j+1}[\mathbf{u}(0), \boldsymbol{\theta}(0)] + F_{(s-1)j+1}[\mathbf{u}(\Delta\tau), \boldsymbol{\theta}(\Delta\tau)])\,, \tag{49}$$

where $F_i[\mathbf{u}, \boldsymbol{\theta}]$ denotes the partial derivative of $\mathcal{H}'[\mathbf{u}, \boldsymbol{\theta}]$ w.r.t. u_i. For the model parameters, analogous equations apply. When it comes to the staging beads, however, only the momenta have to be updated, because the associated kinetic term is not part of $\mathcal{H}'$ but of $\mathcal{H}_N$. Thus,

$$p_{(s-1)j+k}(\Delta\tau) = p_{(s-1)j+k}(0) + \frac{\Delta\tau}{2}(F_{(s-1)j+k}[\mathbf{u}(0), \boldsymbol{\theta}(0)] + F_{(s-1)j+k}[\mathbf{u}(\Delta\tau), \boldsymbol{\theta}(\Delta\tau)])\,. \tag{50}$$

Some preliminary results. The results shown in the figures refer to synthetic data obtained assuming a simple sinusoidal model for the rain input. The experimental data were then generated using the "true" parameter values $K = 200$ and $\gamma = 0.2$. We considered $n = 51$ data points, $j = 10$ staging points, and therefore $N = 501$ discretization points. The total time T is ≈ 833 [units of time], so that the discretization points are separated by a time interval $\Delta t \approx 1.6$ [units of time]. Following the RESPA method described above, we used a long time step $\Delta\tau = 0.02$ [units of time] for the evolution of the slow (long-range) potential, further subdivided in 16 smaller steps for the evolution of the fast (short-range) potential.

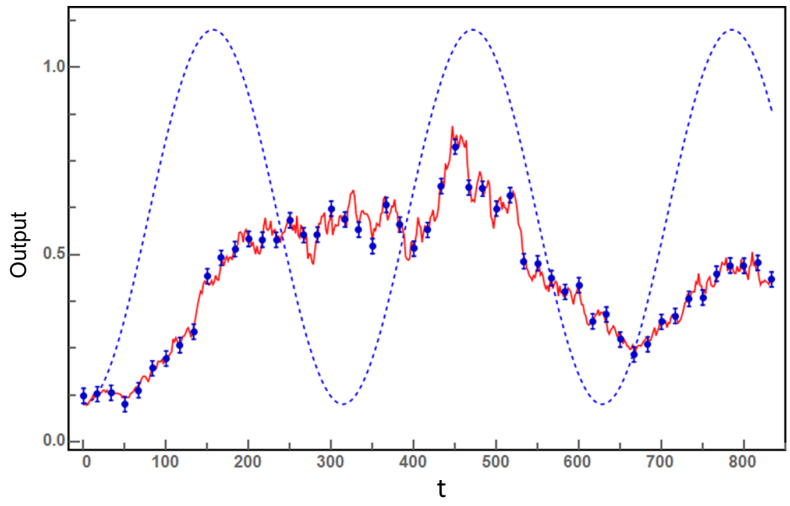

Figure 1: Rain input (dashed), system realization (solid line) and synthetic experimental data. The system output and the observed data were generated using $K = 200$ and $\gamma = 0.2$.

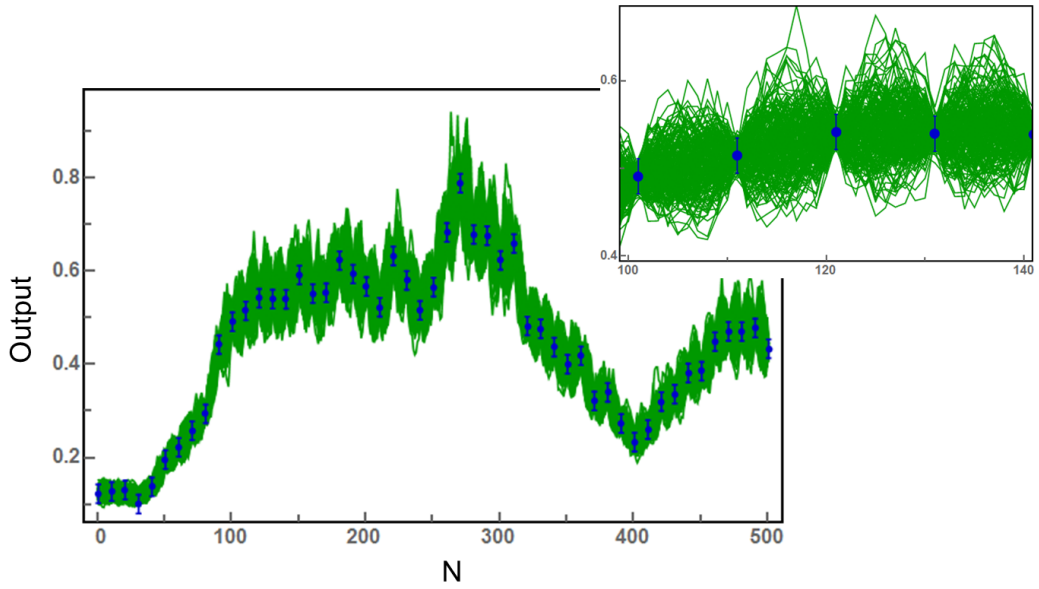

Figure 2: Simulated system realizations and synthetic experimental data obtained assuming a simple sinusoidal model for the rain input. In the inset one may appreciate the different dynamics of the heavy data-point beads and the light staging beads.

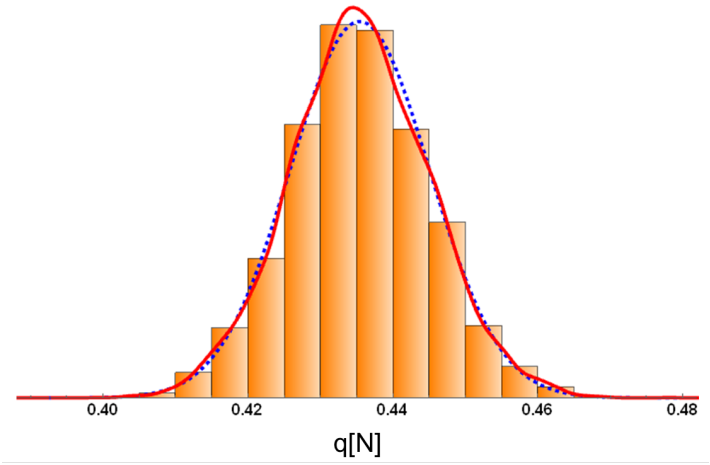

Figure 3: PDF for the system final state expressed via the coordinate set $\{q_N\}$. The dashed curve represents a normal distribution with the same mean and variance. The final state is basically normally distributed.

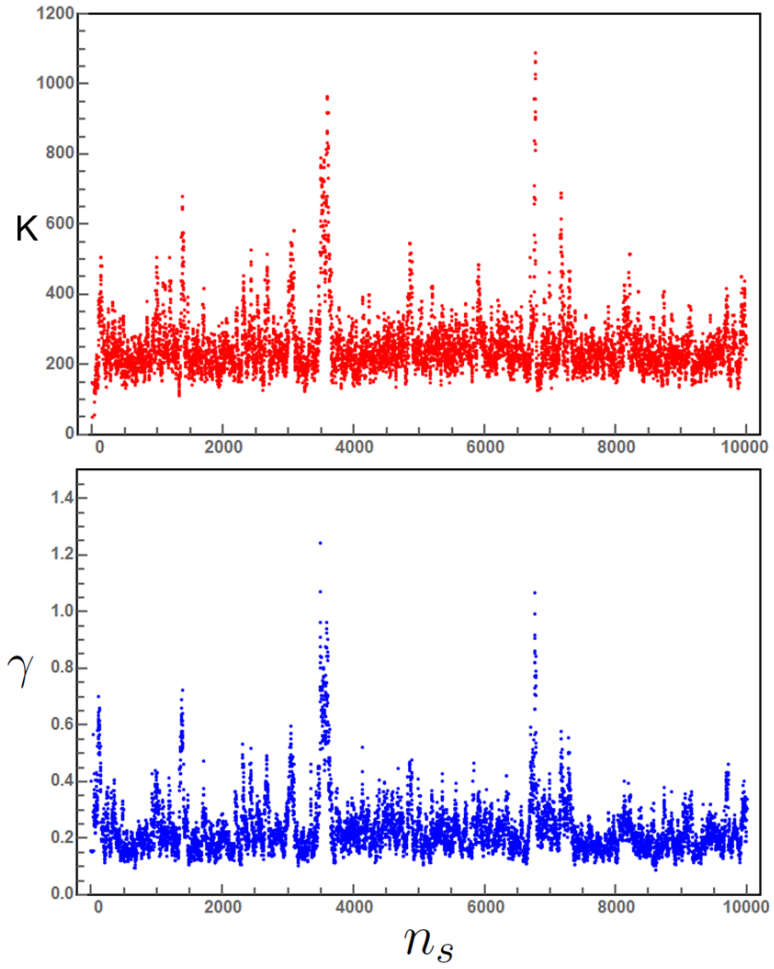

Figure 4: Markov chains for the two parameters K (top) and γ (bottom).

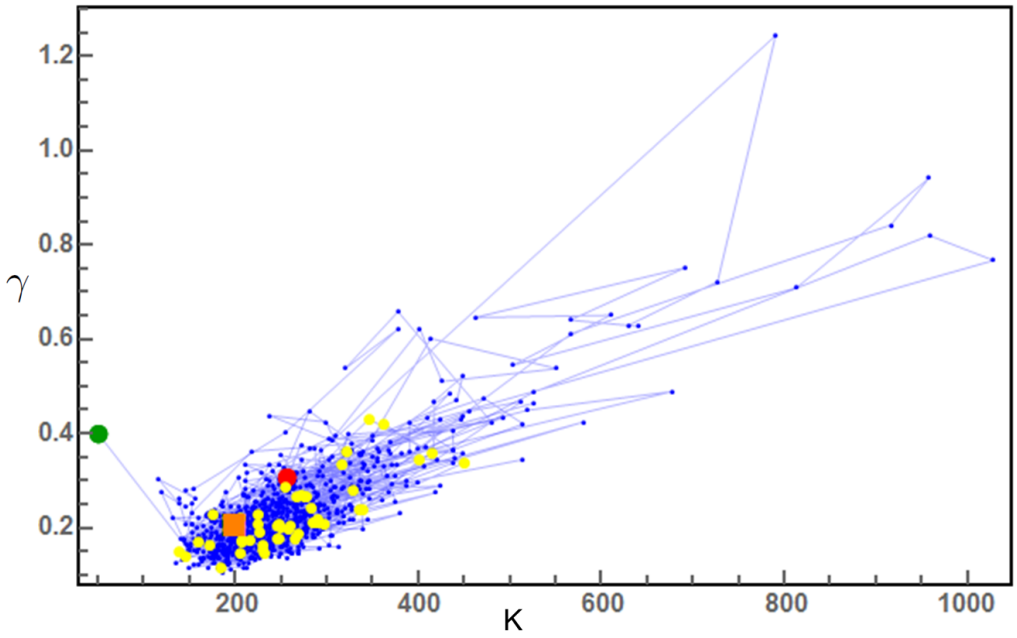

Figure 5: System dynamics in the phase space $K - \gamma$. The green dot is the initial state, the red dot is the final state, the yellow dots are the last 50 states of the HMC chain, and the orange square corresponds to the true parameter values used to generate the data.

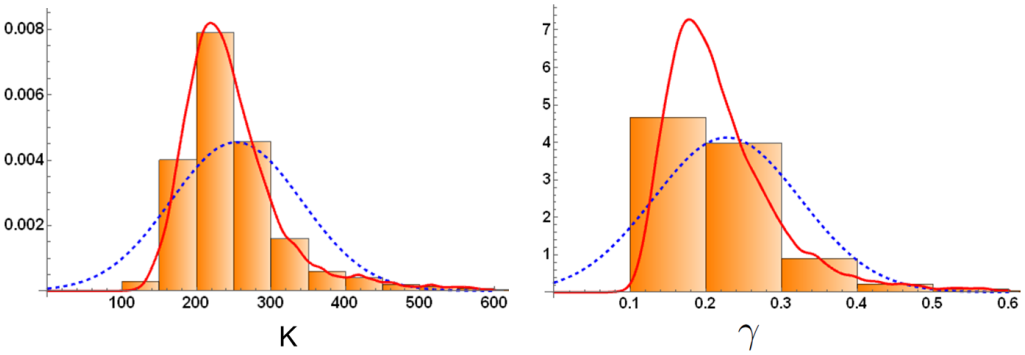

Figure 6: PDF for the parameters K (left) and γ (right). The true values used to generate the data are $K = 200$ and $\gamma = 0.2$. The initial values used in the simulations were $K = 50$ and $\gamma = 0.4$.

References

Doll, J., Coalson, R. D., and Freeman, D. L. (1985). Fourier path-integral monte carlo methods: Partial averaging. *Physical review letters*, 55(1):1.

Duane, S., Kennedy, A., Pendleton, B., and Roweth, D. (1987). Hybrid Monte Carlo. *Phys. Lett. B*, 195(2):216–222.

Lau, A. W. and Lubensky, T. C. (2007). State-dependent diffusion: thermodynamic consistency and its path integral formulation. *Physical Review E*, 76(1):011123.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092.

Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992). Reversible multiple time scale molecular dynamics. *The Journal of chemical physics*, 97(3):1990–2001.