MCMC_CLIB

Analysis of a large Sample

Andrei Kramer

December 17, 2013

This Supplement provides some illustration for the chosen model. It shows, how a parameter sample is related to a model fit. This model has been chosen as an example for a hard problem: the model is large and does not fit all of the observations; the measurements are sparse, uncertain, and are provided in arbitrary units. Note: The reason for the missmatch will not be discussed. The data might have been obtained under erroneous assumptions which are consequently not reflected in the modeling setup or alternatively, the model is incorrect.

1 Two different ways to sample the provided model

The sample was obtained in $t_\mathrm{s} = 267385\,\mathrm{s}$. Considering the auto-correlation $\tau_{\mathrm{int.},l} = 1516(250)$, the effective sampling speed is:

$$v = \frac{N}{2\tau_{\mathrm{int.},l}t_\mathrm{s}} = 12(2) \times 10^{-4}\,\mathrm{s}^{-1}\,. \tag{1}$$

The properties recorded during sampling are summarised in Table 1. The sample is plotted in Figure 1 using parallel corrdinates (each sampled parameter vector is shown component wise, as a line). Figure 2 shows the output trajectories obtained from forward simulations using the sampled parameters. There is partial agreement between model and data (black errorbars).

1.1 Alternative sampling approach

Just for demonstration purposes, we could ignore the source of the data and pretend, that the measurements are absolute (gauged). The reference experiment has been included as input line u_0. Although this is wrong, the problem size is the same and the model can still fit some of the data. Figure 3 shows this alternative fit. Note that

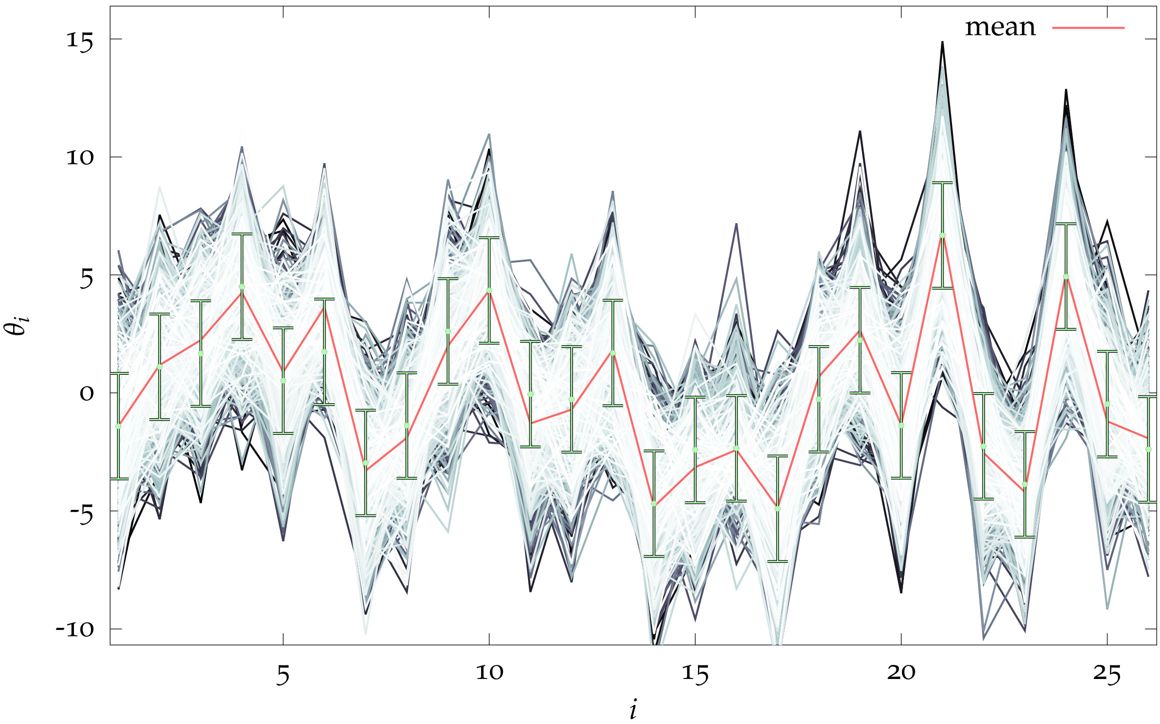

Figure 1: The Posterior Sample of size $N = 1 \times 10^6$, as obtained from the `MCMC_CLIB`; thinned out using the integrated autocorrelation length $\tau_{\text{int.},l} = 1516(250)$ (estimated using the log-posterior l as observable). The red line indicates the maximum posterior estimate, while the green errorbars indicate the used prior distribution (Gaussian). Here, we have incorporated our prior knowledge about the parameter values in the prior parameter density. Yet, we have chosen to let the prior span several orders of magnitude.

tuning duration	1000 points
target acceptance a_0	0.50
observed acceptance a	0.27
auto-correlation length $\tau_{\text{int.},l}$	1516(250)
failed Likelihood evaluations `[CVODES]`	16

Table 1: Sample Properties.

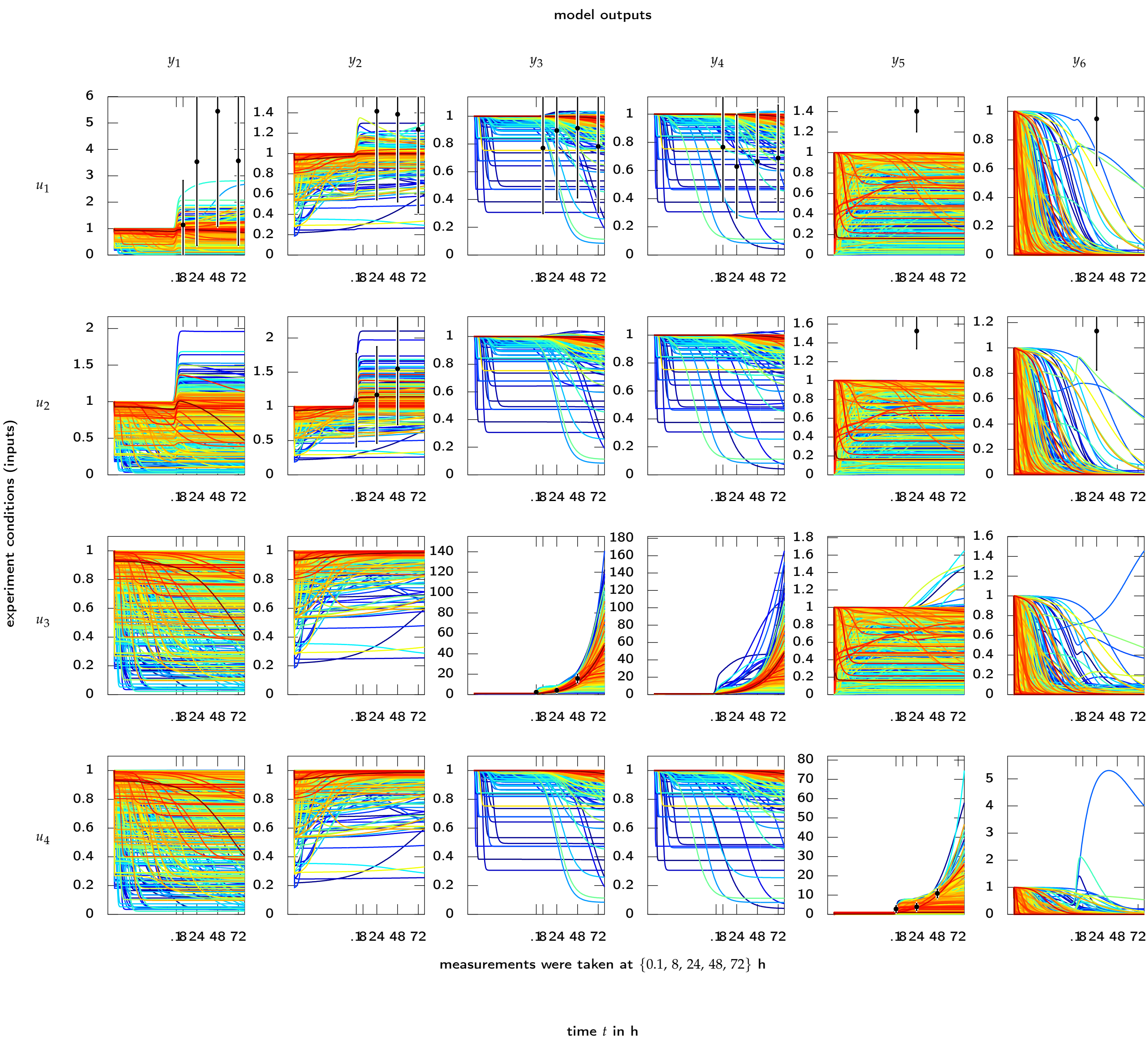

Figure 2: output trajectories per parameter sample member; to reduce the number of lines these trajectory lines are once again plotted for every 3534th $(2(\tau_{\text{int.},l} + \delta_\tau))$ parameter in the parameter sample. The green, vertical errorbars represent the data and its measurement error estimation. The data is sparse (not all outputs were observed at all measurement instances. The model does not fit all data points.)

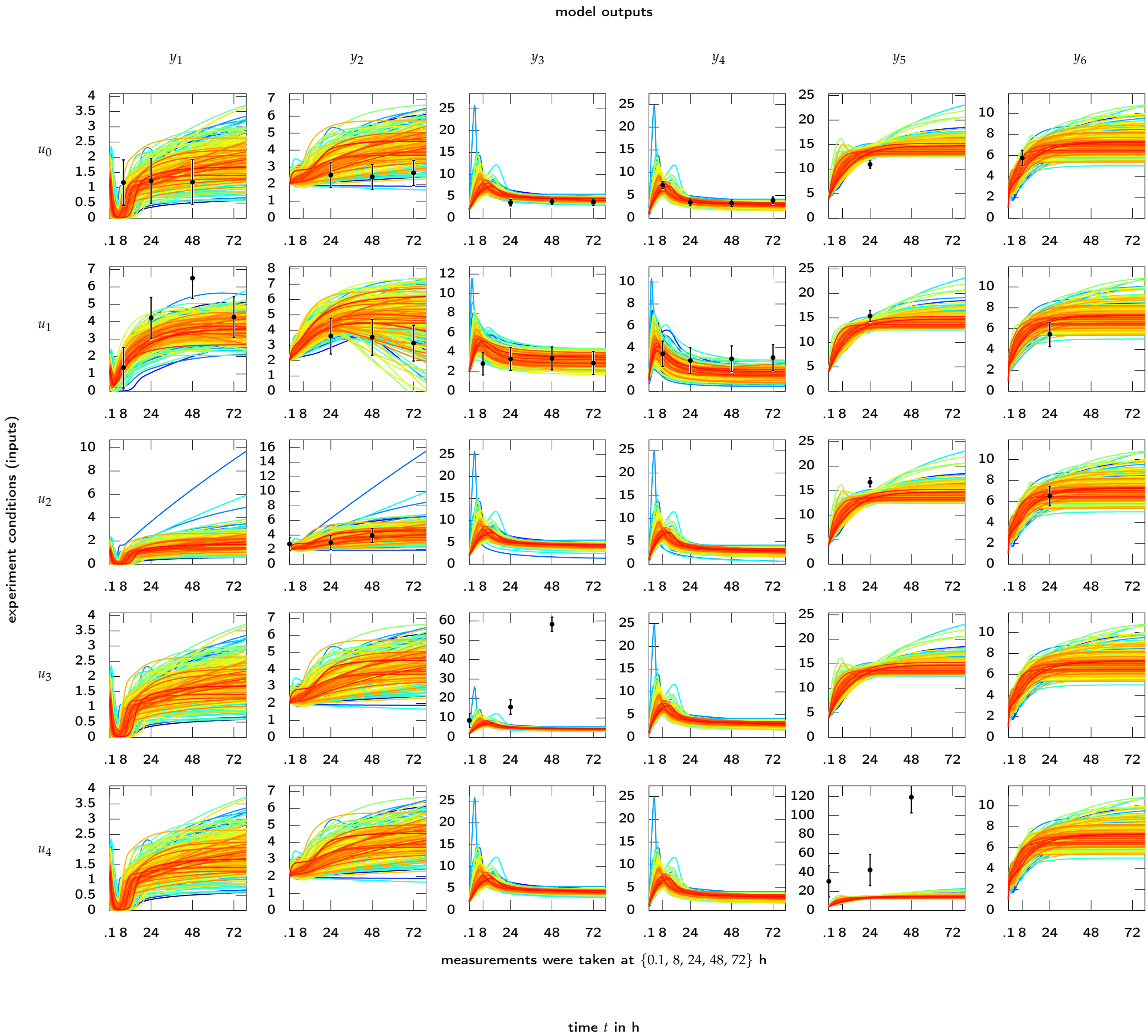

Figure 3: Treating the data as absolute, the fit changes (u_0 is the setup for the reference experiment). Some data points get an improved fit, some system behaviour cannot be reproduced. Note that this treatment of the data is incorrect and is done for illustration purposes; we show the difference between fitting absolute and relative data. To obtain these Trajectories, we have used a different, uninformative, iid, Gaussian prior during sampling. The resulting parameter sample is shown in Figure 4

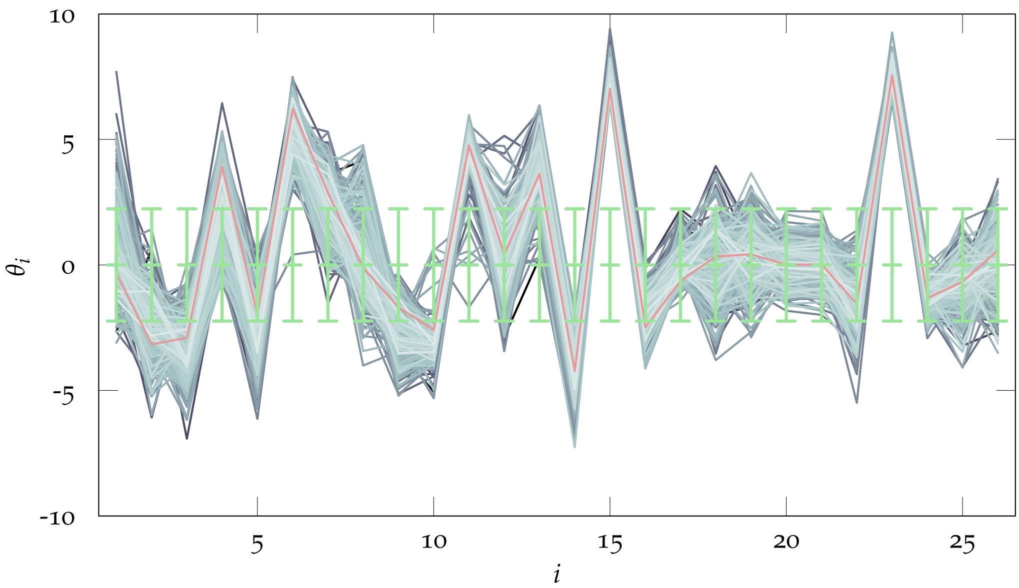

Figure 4: The parameters obtained from the assumption that provided data is absolute (molecule numbers or mols). Here an uninformative prior was used.

the model cannot fit the dynamic behaviour in experiment (u_3, y_4) and (u_4, y_5) while it fits the data very well when arbitrary units, i.e. relative data are treated correctly (see Figure 2).

1.2 Model Setup

Listing 1: Model

```
<?xml version="1.0" ?>
<VectorField Name="ODEmodel11S26P4U" Description="A␣model␣for␣testing␣purposes">
<Parameter Name="theta_1" DefaultValue="1.0"/>
<Parameter Name="theta_2" DefaultValue="1.0"/>
<Parameter Name="theta_3" DefaultValue="1.0"/>
<Parameter Name="theta_4" DefaultValue="1.0"/>
<Parameter Name="theta_5" DefaultValue="1.0"/>
<Parameter Name="theta_6" DefaultValue="1.0"/>
<Parameter Name="theta_7" DefaultValue="1.0"/>
<Parameter Name="theta_8" DefaultValue="1.0"/>
<Parameter Name="theta_9" DefaultValue="1.0"/>
<Parameter Name="theta_10" DefaultValue="1.0"/>
<Parameter Name="theta_11" DefaultValue="1.0"/>
<Parameter Name="theta_12" DefaultValue="1.0"/>
<Parameter Name="theta_13" DefaultValue="1.0"/>
<Parameter Name="theta_14" DefaultValue="1.0"/>
<Parameter Name="theta_15" DefaultValue="1.0"/>
```

```
<Parameter Name="theta_16" DefaultValue="1.0"/>
<Parameter Name="theta_17" DefaultValue="1.0"/>
<Parameter Name="theta_18" DefaultValue="1.0"/>
<Parameter Name="theta_19" DefaultValue="1.0"/>
<Parameter Name="theta_20" DefaultValue="1.0"/>
<Parameter Name="theta_21" DefaultValue="1.0"/>
<Parameter Name="theta_22" DefaultValue="1.0"/>
<Parameter Name="theta_23" DefaultValue="1.0"/>
<Parameter Name="theta_24" DefaultValue="1.0"/>
<Parameter Name="theta_25" DefaultValue="1.0"/>
<Parameter Name="theta_26" DefaultValue="1.0"/>
<Parameter Name="u1" DefaultValue="0.0"/>
<Parameter Name="u2" DefaultValue="0.0"/>
<Parameter Name="u3" DefaultValue="0.0"/>
<Parameter Name="u4" DefaultValue="0.0"/>
<Expression Name="logistic" Formula="1.0/(1+exp(-t))"/>
<Expression Name="U1t" Formula="u1*logistic"/>
<Expression Name="U2t" Formula="u2*logistic"/>
<Expression Name="U3t" Formula="u3*logistic"/>
<Expression Name="U4t" Formula="u4*logistic"/>
<Expression Name="S1" Formula="X4+X5"/>
<Expression Name="S2" Formula="(theta_22)*X9*X7"/>
<Expression Name="A1" Formula="(theta_1)*X2"/>
<Expression Name="A2" Formula="(theta_2)*X1"/>
<Expression Name="A3" Formula="(theta_3)*X3*X2"/>
<Expression Name="A4" Formula="(theta_4)*X4"/>
<Expression Name="A5" Formula="(theta_10)*X4"/>
<Expression Name="A51" Formula="(theta_10)*X5"/>
<Expression Name="A6" Formula="(theta_13)*X7"/>
<Expression Name="A7" Formula="(theta_14)*S1*X6"/>
<Expression Name="A8" Formula="(theta_19)*X10*S1"/>
<Expression Name="A9" Formula="(theta_20)*X8"/>
<Expression Name="A10" Formula="(theta_22)*X11*X7"/>
<Expression Name="A11" Formula="(theta_21)*X9"/>
<StateVariable Name="X1"
               DefaultInitialCondition="1000.0"
               Formula="A1-A2+(theta_5)*(0.1+S2)-(theta_7)*X1" />
<StateVariable Name="X2"
               DefaultInitialCondition="1000.0"
               Formula="-A1+A2-(theta_8)*X2" />
<StateVariable Name="X3"
               DefaultInitialCondition="1000.0"
               Formula="-A3+A4+(theta_6)+(theta_12)*U2t-(theta_9)*X3" />
<StateVariable Name="X4"
               DefaultInitialCondition="1000.0"
               Formula="A3-A4-A5" />
<StateVariable Name="X5"
               DefaultInitialCondition="0.0"
               Formula="+(theta_11)*U1t-A51" />
```

```
<StateVariable Name="X6"
               DefaultInitialCondition="1000.0"
               Formula="A6-A7+(theta_15)+(theta_18)*U3t-(theta_16)*X6" />
<StateVariable Name="X7"
               DefaultInitialCondition="1000.0"
               Formula="-A6+A7-(theta_17)*X7" />
<StateVariable Name="X8"
               DefaultInitialCondition="1000.0"
               Formula="+A8-A9-(theta_24)*X8" />
<StateVariable Name="X9"
               DefaultInitialCondition="1000.0"
               Formula="+A9-A11" />
<StateVariable Name="X10"
               DefaultInitialCondition="1000.0"
               Formula="-A8+A10" />
<StateVariable Name="X11"
               DefaultInitialCondition="1000.0"
               Formula="-A10+A11+(theta_23)+(theta_26)*U4t-(theta_25)*X11" />
<Function Name="Y1" Formula="(X4+X5)"/>
<Function Name="Y2" Formula="(X3+X5)"/>
<Function Name="Y3" Formula="(X7+X6)"/>
<Function Name="Y4" Formula="(X6)"/>
<Function Name="Y5" Formula="(X10+X9+X11+X8)"/>
<Function Name="Y6" Formula="X8"/>
</VectorField>
```