Easy parameter identifiability analysis with *Copasi*

Supplementary Material

Jörg Schaber

Institute for Experimental Internal Medicine, Medical Faculty

Otto von Guericke University

Magdeburg, Germany

Contents

A worked example

The model

In order to demonstrate the usage of *Copasi*'s hidden features to easily and rapidly conduct a parameter identifiability analysis, we use a simple model whose identifiability has already been described elsewhere (Schaber and Klipp, 2011).

The model scheme and differential equations are depicted in Figure 1.
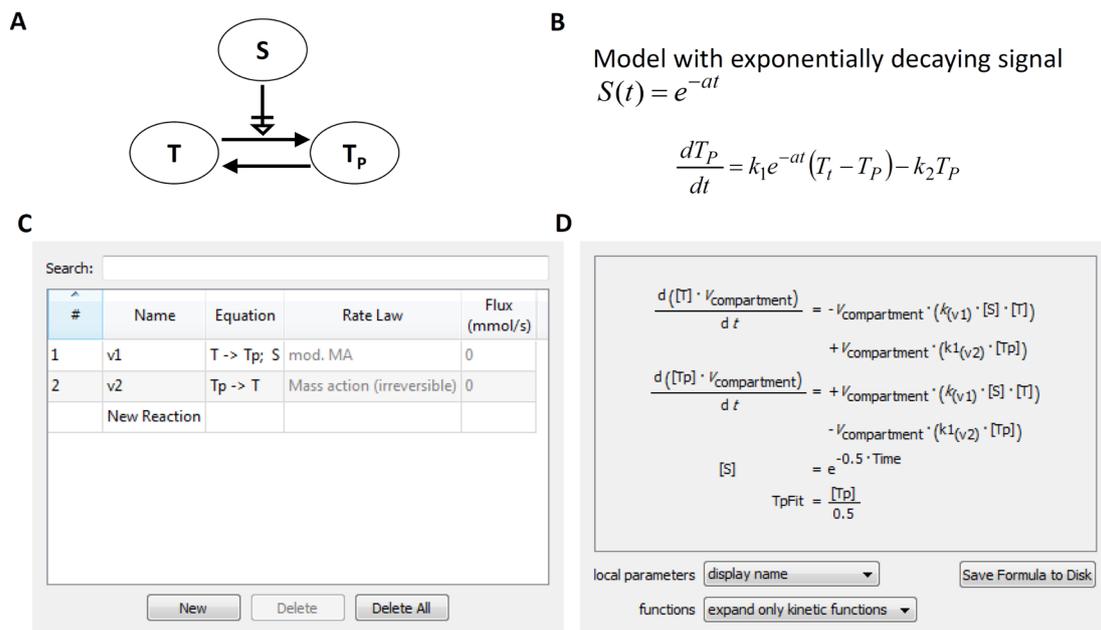
Figure 1: Model scheme and differential equations. **A**: Wiring scheme in SBGN notation. **B**: Corresponding differential equations. $T_p(0) = 0$, $T_t = 1$. **C**: Reactions in the corresponding *Copasi* model. **D**: Differential equations in *Copasi*. Note that the scaled variable *TpFit* was introduced, which is used for fitting the data, which in turn were scaled to 1. The parameter *a* for the decaying signal S was set to 0.5 and volume $V_{compartment}$ was set to 1.

In Figure 2 we display the two data sets that we used to demonstrate the non-identifiable (Data set 1) and the identifiable case (Data set 2), with corresponding model fits. In Table 1 we list the corresponding data.
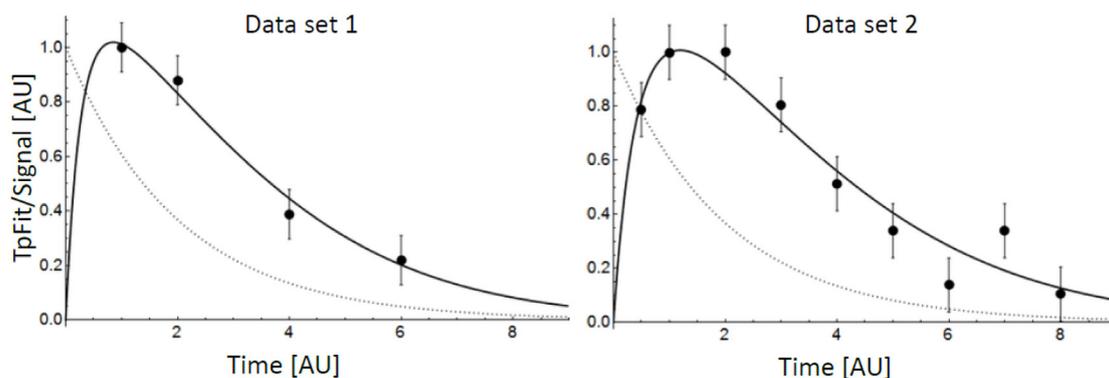
Figure 2: Data sets (dots) and corresponding model fits (black lines). The dotted line indicates the signal *S*.

Table 1: Data used in Figure 2

Data set 1			Data set 2		
time	**data**	**sd**	**time**	**data**	**sd**
1	1	0.09	0.5	0.79	0.1
2	0.88	0.09	1	1	0.1
4	0.39	0.09	2	1	0.1
6	0.22	0.09	3	0.81	0.1
			4	0.51	0.1
			5	0.34	0.1
			6	0.14	0.1
			7	0.34	0.1
			8	0.10	0.1

sd: standard deviation

The model was fitted to the two data sets with *Copasi*. In Figure 3 we display a screen-shot of the corresponding parameter estimation task using data set 1.
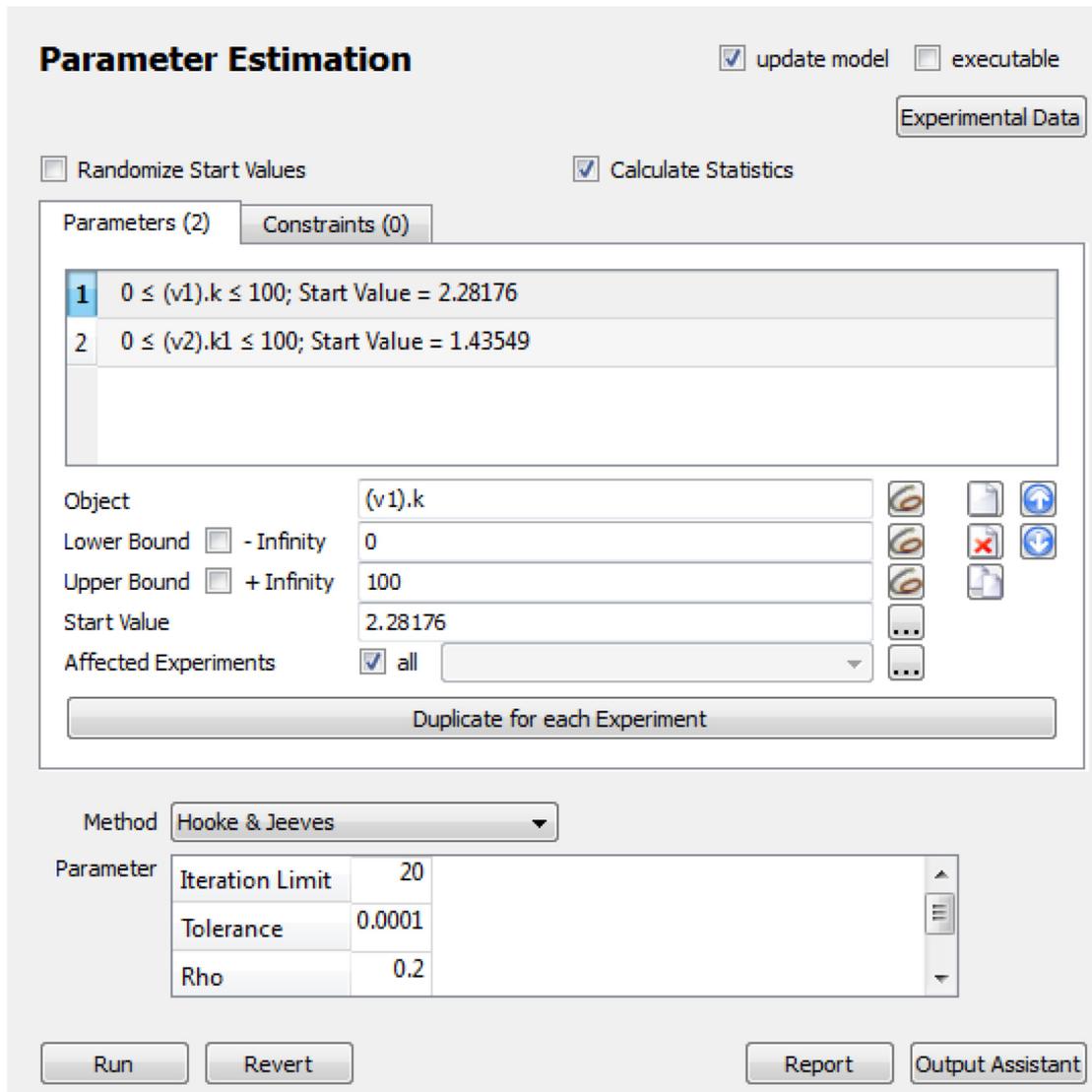
Figure 3: Parameter estimation task interface in *Copasi*. Here, the fitted parameter values for data set 1 are shown as start values, because the check-box 'update model' was checked. Actual start values for the parameter estimation were 0.1 for both parameters.

As the model has only two parameters the likelihood profile is a surface in the three dimensional space, which can be displayed as a contour plot. In Figure 4 we display the likelihood contours of the likelihood function that, in this case, is the weighted sum of squared residuals (*SSR*), as used by *Copasi*.
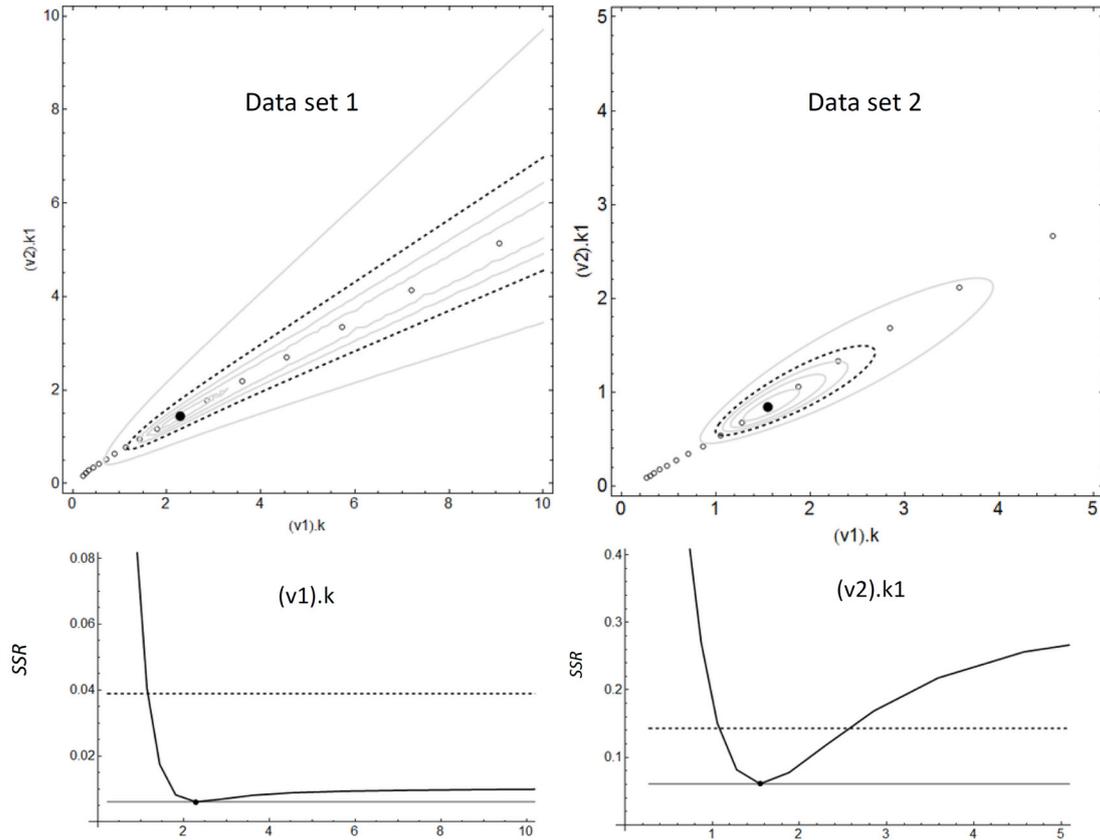
Figure 4: Likelihood contours and profiles for the likelihood function for the two parameters of the model and the two data sets, respectively. In the upper panel, likelihood contours are displayed for the two different data sets, respectively. Grey lines indicate contours for the 50, 80,90 and 99 % confidence levels (in outward direction). The dotted lines indicate 95% confidence levels. The black dot is the fitted parameter vector and the grey circles indicate the parameter vectors that have been sampled in the parameter scan, described below. In the lower panel, one-dimensional likelihood profiles are displayed corresponding to the likelihood contours above, respectively. In the left lower panel the one-dimensional likelihood profile for the parameter (v1).k (k_1) is displayed for the likelihood function with data set 1. In the right lower panel the one-dimensional likelihood profile for the parameter (v2).k1(k_2) is displayed for the likelihood function with data set 2. The dotted lines indicate 95% confidence levels, corresponding to the dotted lines in the upper panel. The grey lines in the lower panel indicate the fitted minimum. The 95% confidence level was calculated according to the

likelihood-based confidence region $\left\{ p : SSR(p) \leq SSR(\hat{p}) \left(1 + \dfrac{m}{n-m} F_{m,n-m}^{\alpha} \right) \right\}$ where

$F_{m,n-m}^{\alpha}$ is the α quantile of the F-Ratio -distribution with *m*=2 the number of parameters and *n*=4 or *n*=9 the number of data points depending on the data set.

In Figure 4, we see that in case of data set 1, the parameter k_1 is not identifiable, because the likelihood-based confidence region for the fitted parameters extends to infinity (upper left panel in Figure 4), as least for the 95% confidence limit (dotted lines in Figure 4). In case of the data set 2, the 95% confidence limit is a confined region and therefore the parameters are well identifiable (upper right panel in Figure 4). The one-dimensional likelihood profiles displayed in the lower panel of Figure 4

are calculated for discrete points for the respective parameters and are indicated as grey circles in the upper panel of Figure 4. The two parameters k_1 and k_2 are highly correlated and the confidence region is symmetric, as can be seen from the likelihood contours in Figure 4. Therefore, the likelihood profiles are almost identical for the two parameters for each likelihood function and we display only the likelihood profile for one parameter for data set 1 (k_1, lower left panel) and data set 2 (k_2, lower right panel). In the one dimensional likelihood profile it can be clearly seen that for data set 1 the likelihood profile does not exceed the 95% confidence limit within the tested interval (lower left panel), whereas for data set 2 the likelihood profile quickly reaches the 95% confidence limit (lower left panel), indicating an identifiable parameter.

Calculating one-dimensional likelihood contours with Copasi

Plots of one-dimensional likelihood contours, as displayed in the lower panel of Figure 4, can easily be produced with *Copasi* (Hoops, et al., 2006). The general procedure is the following:

1) Define a new report and/or plot in the 'Output Specification' section, selecting the following items:

   a. The parameter of interest using the standard pop-up interface

   b. The objective function value using the 'expert-mode' from the pop-up interface, selecting COPASI->ModelList[]->Root->TaskList[]->Parameter Estimation->Parameter Estimation->Best Value.

2) Delete the parameter of interest from the parameter estimation task.

3) Create a new 'Parameter Scan' task selecting the parameter of interest from the standard pop-up interface and define the interval and its discretization to be tested.

4) Select 'Parameter Estimation' as a subtask.

5) Select the above defined new report as the report in the 'Parameter Scan' interface.

6) Save the model in a new file and run the 'Parameter Scan' task.

Here, we exemplify the procedure to produce the plot in the lower left panel of Figure 4, i.e. the likelihood profile for parameter k_1 for data set 1. For the other parameter and data set, the procedure works accordingly.

Step 1: Define a new report

Having the model defined and having fitted parameters from the 'Parameter Estimation', as in Figure 3, the new report can be defined on the 'Output Specification' section. Extending the 'Output Specification' menu and clicking on 'Reports' and then 'new report', the report definition interface is displayed (Figure 5).
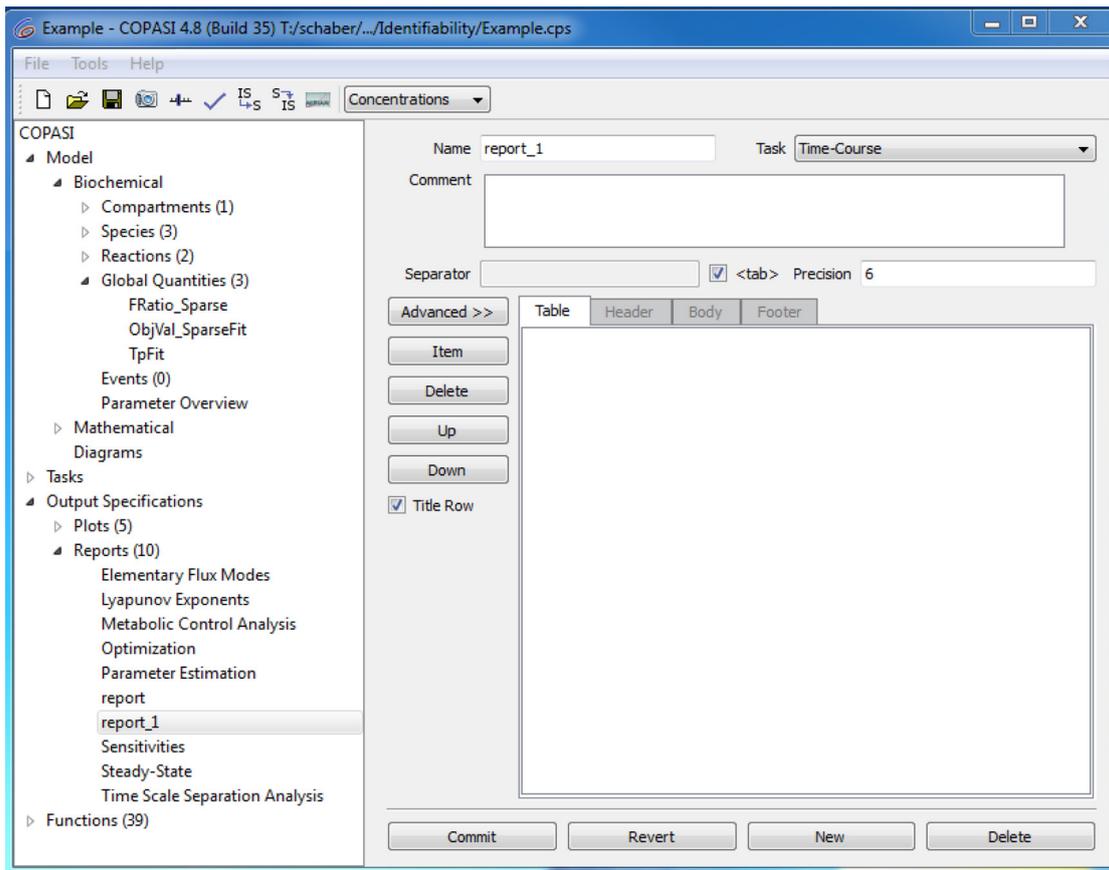
Figure 5: *Copasi* report definition interface.

We may name the new report 'Likelihood Profile' and introduce the parameters of interest into the report by clicking on 'Item' and selecting 'Reaction Parameters' from the standard pop-up interface (Figure 6).
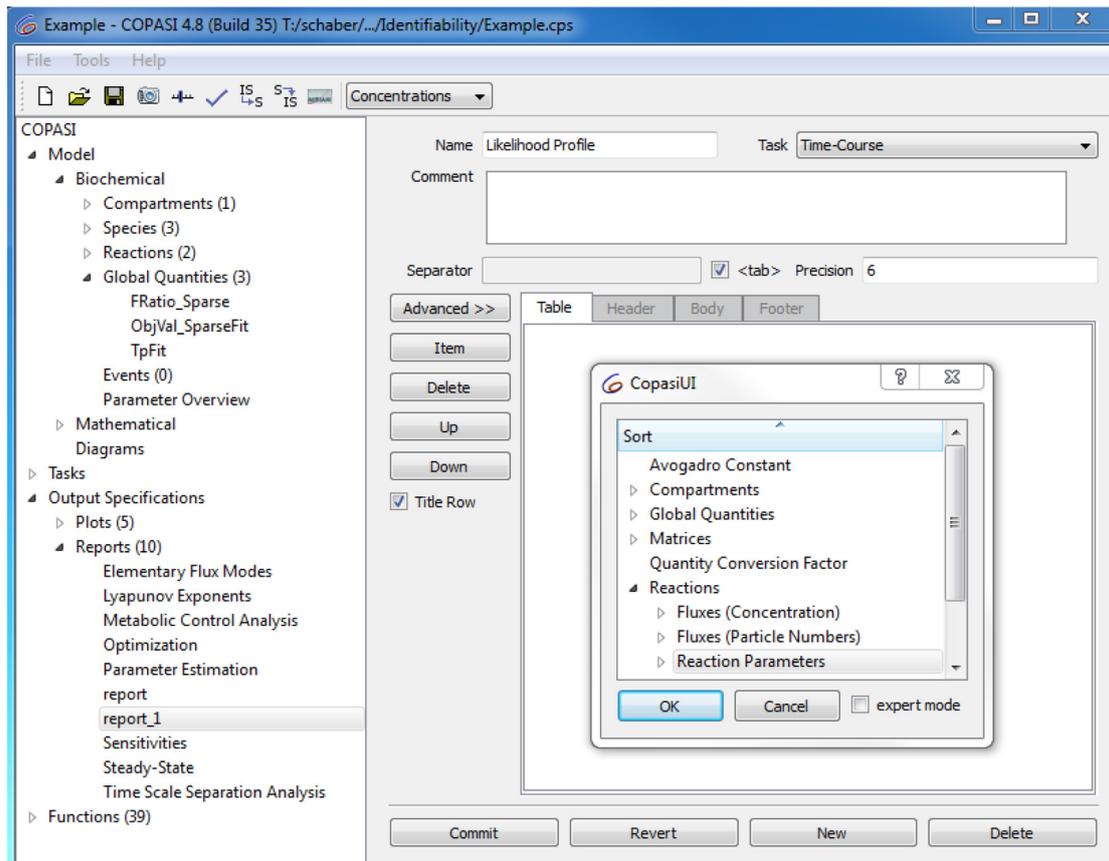
Figure 6: Introducing reaction parameters into the new report.

By clicking 'OK' all reaction parameter are introduced in the report. Now, the corresponding *SSR* has to be introduced in the report. Again, we click on 'Item', but now the standard pop-up interface is extended by selecting the 'expert mode'. From the export-mode interface the likelihood *SSR* is introduced into the report by subsequently selecting COPASI->ModelList[]->Root->TaskList[]->Parameter Estimation->Parameter Estimation->Best Value (Figure 7).
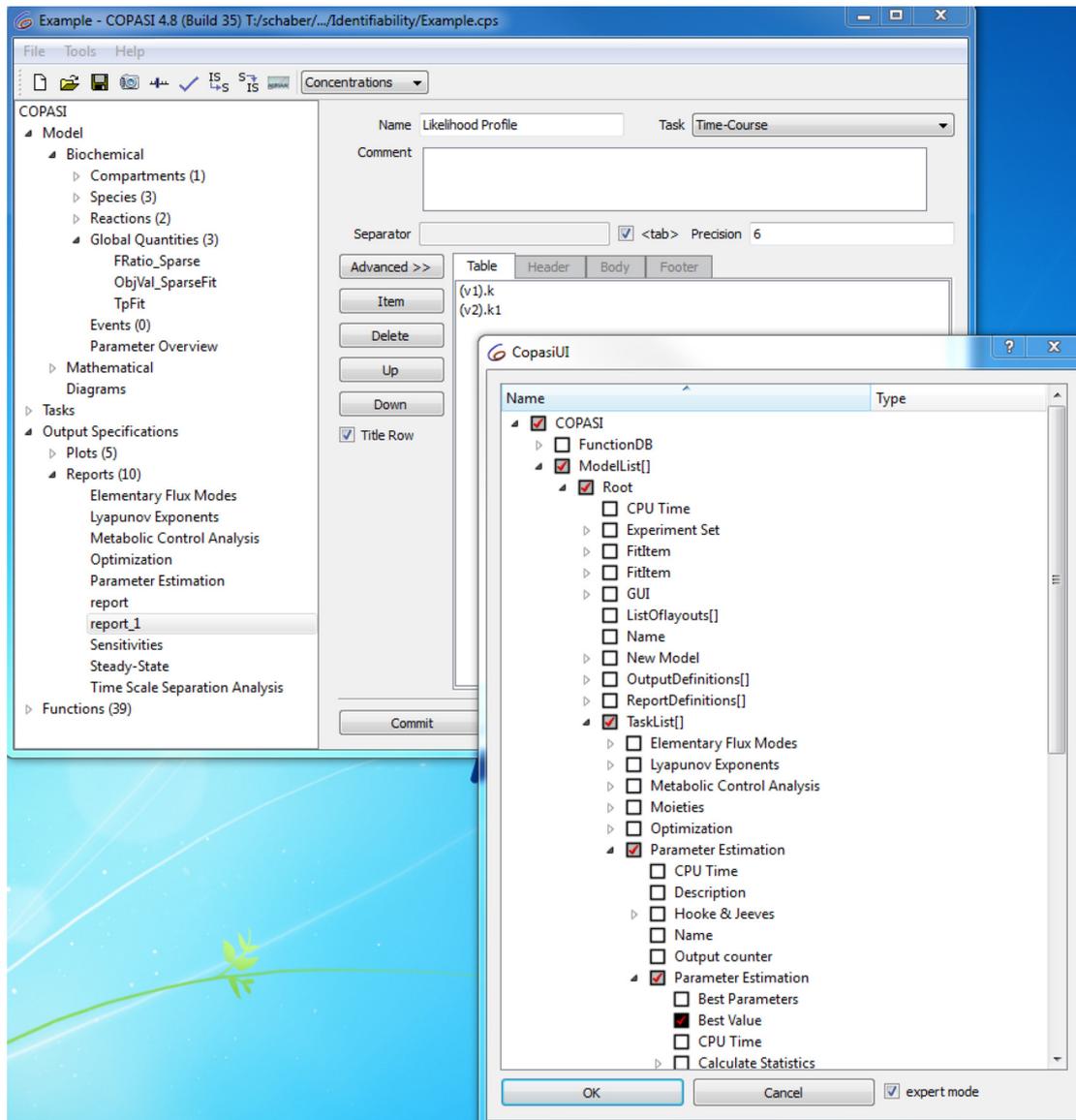
Figure 7: Selecting the *SSR* for the likelihood profile report.

Upon clicking 'OK' and committing, the report definition should look like in Figure 8. Note that new report 'Likelihood Profile' should now be in the list of reports.
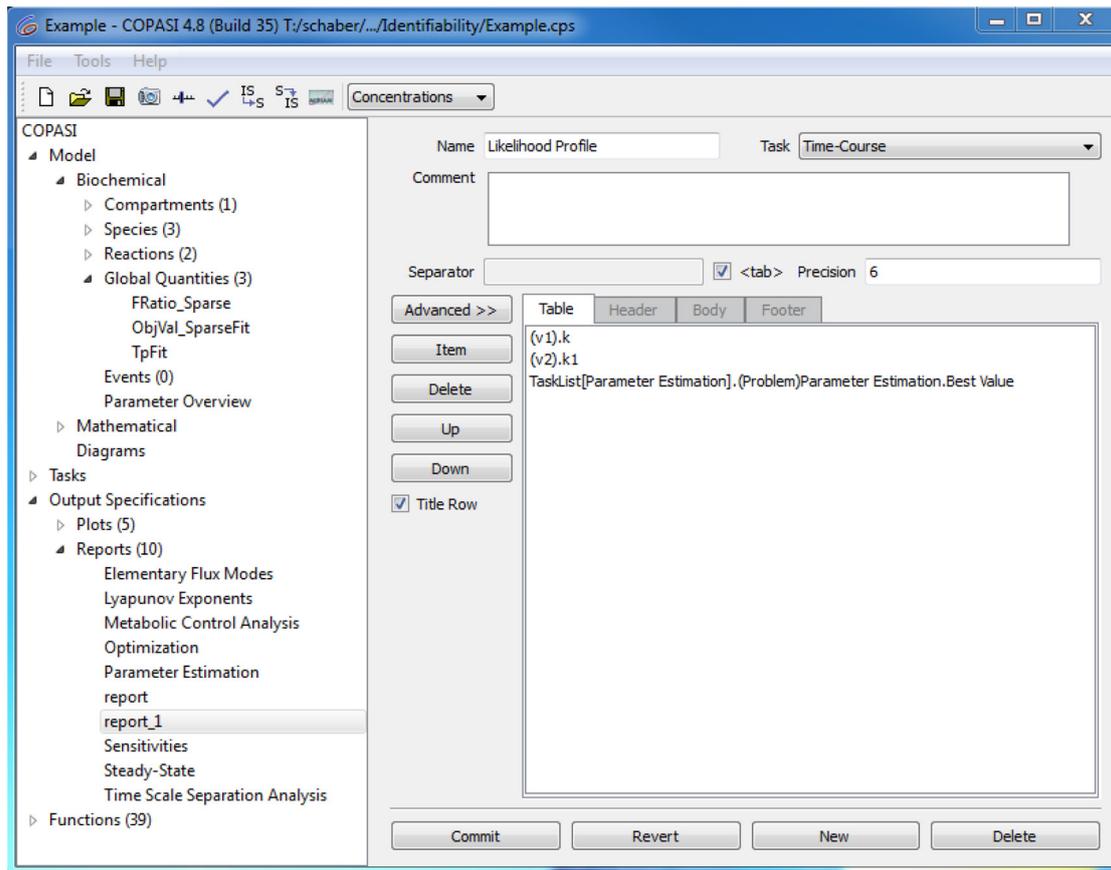
Figure 8: Complete new report definition for likelihood profiles.

From the report, all the information needed to produce Figure 4 can be extracted, however, *Copasi* can also directly plot the likelihood profiles. To this end, a new plot has to be defined, similar to the new report definition. Here, x- and y- Axis have to be defined from the same selection interface. In Figure 9 a completed plot definition is displayed.
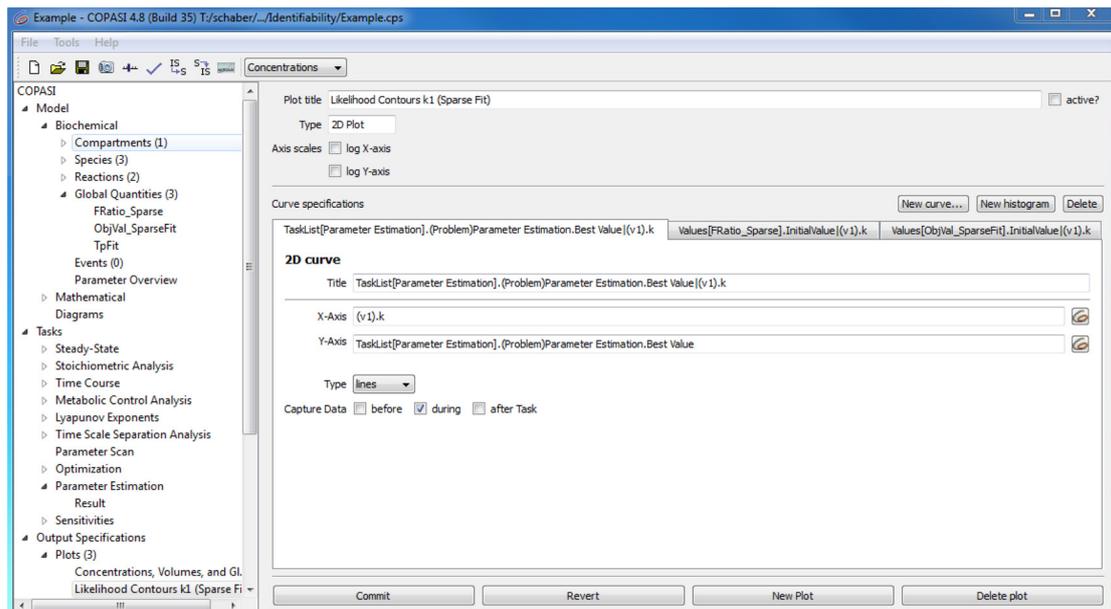
Figure 9: Plot definition of the likelihood contours for parameter k_1.

Note that apart from the likelihood profile, we also plot the objective value from the initial fit using both parameters as well as the 95% confidence level, which was calculated according to the main text or as described in the caption of Figure 4.

Step 2: Delete the parameter of interest from the parameter estimation task

Now, the parameter of interest has to be deleted from the parameter estimation task. Here, parameter (v1).k is removed from the 'Parameter estimation'-interface (Figure 3).

Step 3-6 : The 'Parameter Scan' task

Now, a parameter scan task can be created by selecting 'Parameter Scan' from the task menu, click on '… Create!' and select the parameter of interest from the pop-up menu. Here, we select parameter (v1).k (k_1) within the range of 0.1- and 10-times the fitted value, i.e min=0.228 and max=22.8. Twenty interpolation points are selected on a logarithmic scale in order to get the same number of value above and below the originally fitted value. As a sub-task, 'Parameter Estimation' is selected and the newly defined report 'Likelihood Profile' is selected by clicking on 'Report' (Figure 10). Finally, a file name for the report has to be defined.
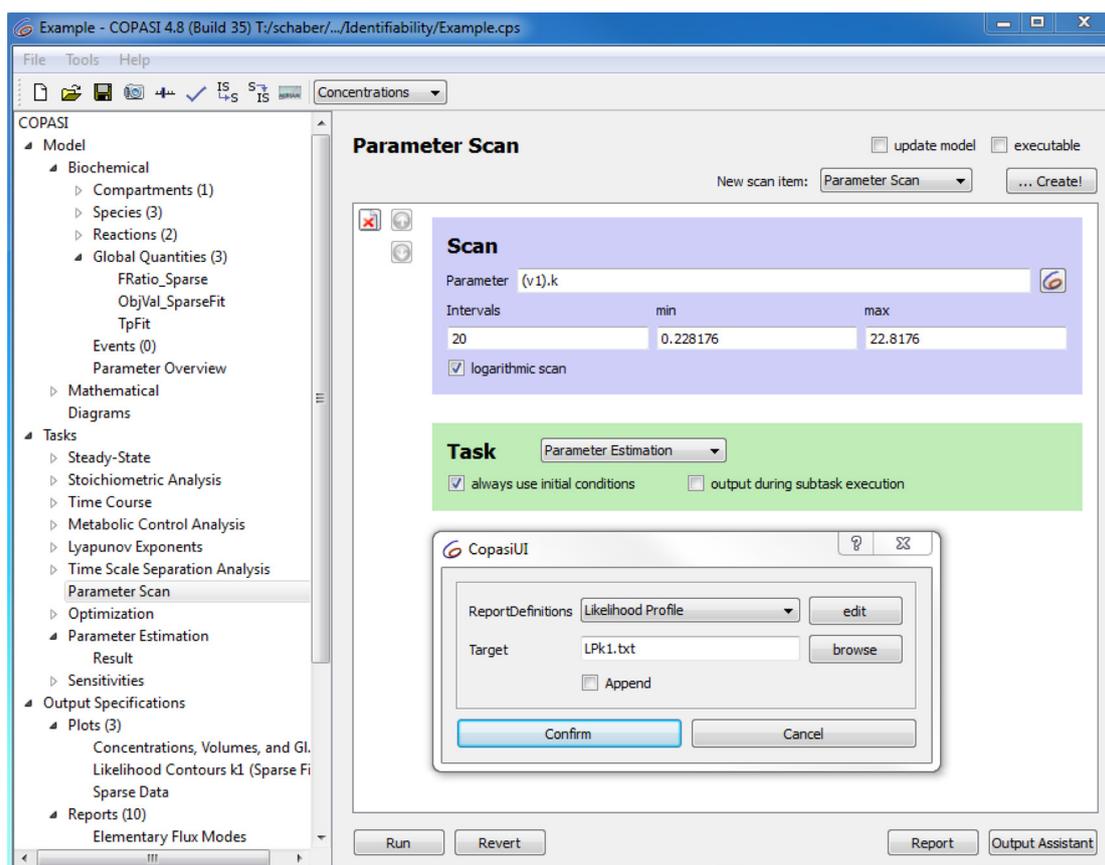
Figure 10: Parameter scan task definition.

Since for each parameter an extra parameter scan task has to be defined, it is advisable to save the modified model in a new file, and create an extra file for each parameter. By clicking run, the report file is created (not shown) and the defined plot, given, that

it was activated before (Figure 11). The resulting (zoomed) plot is the same as in the lower left panel of Figure 4.
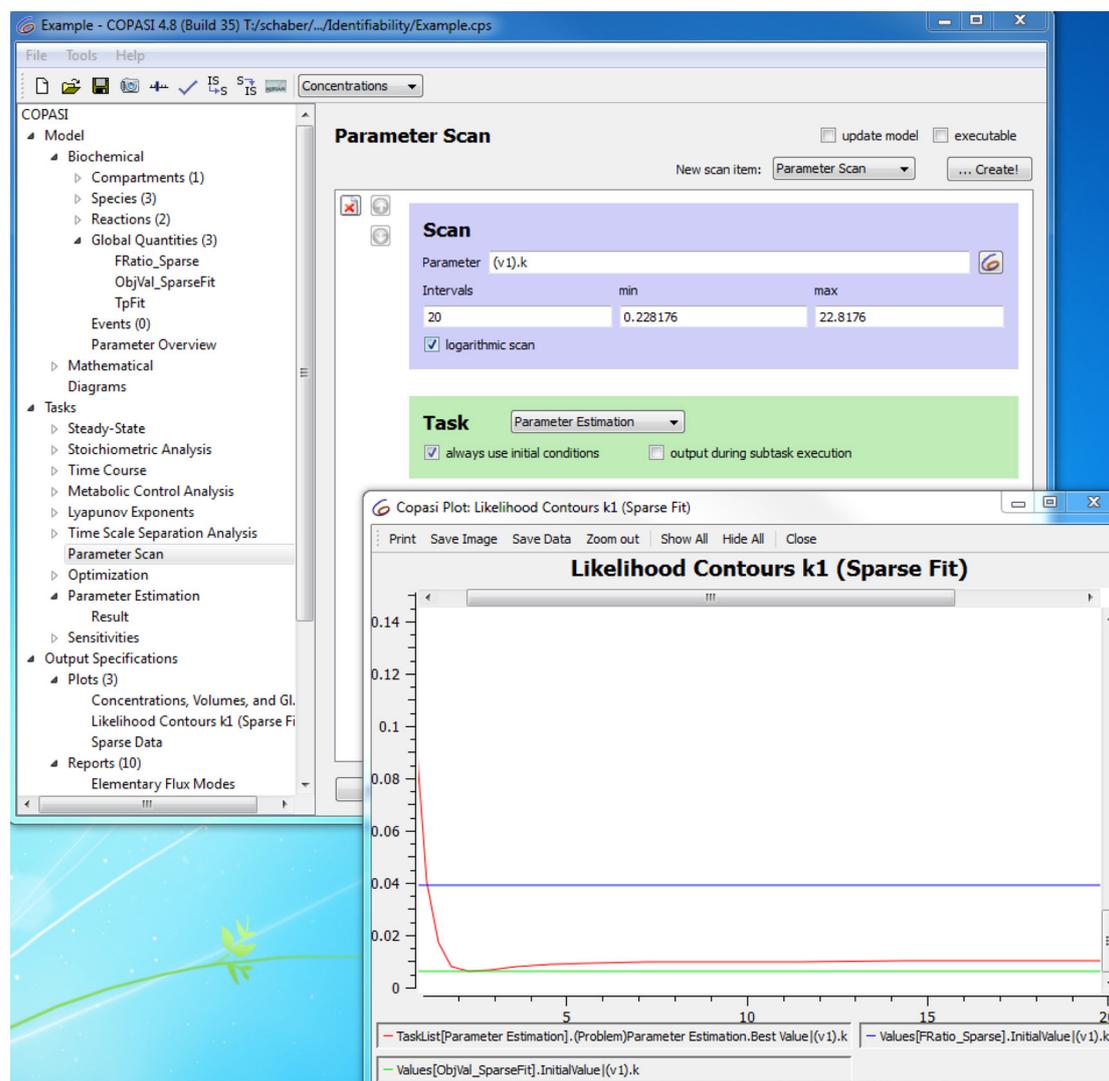
Figure 11: The calculated likelihood profile for parameter (v1).k.

In case such a one-dimensional parameter identifiability analysis is conducted for many parameters, one can define a parameter scan task for each parameter in a separate file, mark the check-box 'executable' in the Parameter Scan-task and run all files automatically on a cluster. Of course, for each file a distinct report file has to be defined. Then, the likelihood profiles for each parameter can be automatically extracted from the respective report files using some other software.

References

Hoops, S.*, et al.* (2006) COPASI--a COmplex PAthway SImulator, *Bioinformatics*, **22**, 3067-3074.

Schaber, J. and Klipp, E. (2011) Model-based inference of biochemical parameters and dynamic properties of microbial signal transduction networks, *Curr Opin Biotechnol*, **22**, 109-116.