Supplementary Material to: S-system parameter estimation for noisy metabolic profiles using Newton-flow analysis

Zoltán Kutalik[1, *], **Warwick Tucker**[2], **Vincent Moulton**[3]

1 Examples of the attractor of the Newton-flow

In the following we provide some examples to indicate that an attractor with the properties defined in the article is likely to exist. For each minimization problem associated with the following S-systems we generated 40 Newton candidates, which are supposed to lie in close vicinity of the attractor curve. The corresponding curves were fitted and R^2-values were computed. In each figure the 2-dimensional projections of the Newton candidates are marked with blue dots, the projection of the global optimum is denoted by a red star, and the attractor curve fitted to the Newton candidates is represented by a dashed green line. All R^2-values were above 0.9.

Example 1. We investigated a 2-dimensional S-system example with four different parameter settings, each of which produced different system behaviors.

$$\begin{aligned} \dot{x}_1 &= 3x_2^{-2} - x_1^{0.5}x_2 \qquad (1)\\ \dot{x}_2 &= x_1^{0.5}x_2 - x_2^{0.5} \\ x_1(0) &= 3 \\ x_2(0) &= 1 \end{aligned}$$

[1]Department of Medical Genetics, University of Lausanne, Rue de Bugnon 27, 1005 Lausanne, Switzerland

[2]Department of Mathematics, Uppsala University, S-75106, Uppsala, Sweden

[3]School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom

[*]To whom correspondence should be addressed. E-mail: zoltan.kutalik@unil.ch

$$\begin{aligned}\dot{x}_1 &= 1.5x_2^{-2} - x_1^{0.5}x_2 \\ \dot{x}_2 &= x_1^{0.5}x_2 - x_2^{0.5} \\ x_1(0) &= 1.5 \\ x_2(0) &= 1.5\end{aligned} \tag{2}$$

$$\begin{aligned}\dot{x}_1 &= .75x_2^{-2} - x_1^{0.5}x_2 \\ \dot{x}_2 &= x_1^{0.5}x_2 - x_2^{0.5} \\ x_1(0) &= 0.75 \\ x_2(0) &= 1\end{aligned} \tag{3}$$

$$\begin{aligned}\dot{x}_1 &= .75x_2^{-2} - x_1^{0.5}x_2 \\ \dot{x}_2 &= x_1^{0.5}x_2 - x_2^{0.5} \\ x_1(0) &= 0.75 \\ x_2(0) &= 1.5\end{aligned} \tag{4}$$

The concentration curves are shown in Figure 1. The reconstructed attractors are illustrated in Figure 2. Note that the four attractors which have not been shown in the figure are extremely similar to the ones in the third column, and we thus omitted them.

Example 2. Figure 3 shows all reconstructed attractor curves for the 4-dimensional example we presented in the paper.

$$\begin{aligned}\dot{x}_1 &= 12x_3^{-0.8} - 10x_1^{0.5} \\ \dot{x}_2 &= 8x_1^{0.5} - 3x_2^{0.75} \\ \dot{x}_3 &= 3x_2^{0.75} - 5x_3^{0.5}x_4^{0.2} \\ \dot{x}_4 &= 2x_1^{0.5} - 6x_4^{0.8}\end{aligned} \tag{5}$$

Example 3. In Figure 4 we present the 2-dimensional projections of three attractors (in $\mathbb{R}^6$) corresponding to Eq. (11), (24), and (27) in the 30-dimensional example presented in the paper. The parameter setup was as follows: In Eq. (1) we set n=30, and take non-zero model parameters $\alpha_i = \beta_i = 1$, $g_{1,14} = -0.1$, $g_{5,1} = 1.0$, $g_{6,1} = 1.0$, $g_{7,2} = 0.5$, $g_{7,3} = 0.4$, $g_{8,4} = 0.2$, $g_{8,17} = -0.2$, $g_{9,5} = 1.0$, $g_{9,6} = -0.1$, $g_{10,7} = 0.3$,

$g_{11,4} = 0.4$, $g_{11,7} = -0.2$, $g_{11,22} = 0.4$, $g_{12,23} = 0.1$, $g_{13,8} = 0.6$, $g_{14,9} = 1.0$, $g_{15,10} = 0.2$, $g_{16,11} = 0.5$, $g_{16,12} = -0.2$, $g_{17,13} = 0.5$, $g_{19,14} = 0.1$, $g_{20,15} = 0.7$, $g_{20,26} = 0.3$, $g_{21,16} = 0.6$, $g_{22,16} = 0.5$, $g_{23,17} = 0.2$, $g_{24,15} = -0.2$, $g_{24,18} = -0.1$, $g_{24,19} = 0.3$, $g_{25,20} = 0.4$, $g_{26,21} = -0.2$, $g_{26,28} = 0.1$, $g_{27,24} = 0.6$, $g_{27,25} = 0.3$, $g_{27,30} = -0.2$, $g_{28,25} = 0.5$, $g_{29,26} = 0.4$, and $g_{30,27} = 0.6$, while $h_{i,j}$ was defined to be -1 if $i = j$, and 0 otherwise. Parameter bounds were also taken as suggested in [1]: α_i, β_i were assumed to be in [0, 3], and $g_{i,j}$, $h_{i,j}$ were restricted to lie in [-3, 3].

Example 4. To investigate a less sparse example, a 7-dimensional example was constructed where the first equation includes 9 non-zero parameters. The corresponding Newton-flow has an attractor in $\mathbb{R}^9$. The projections of this attractor are depicted in Figure 5. The first equation was defined by the following equation.

$$\dot{x}_1 = 5x_1^{0.6}x_2^{0.75}x_5^{0.3}x_7^{-0.4} - 7x_3^{0.9}x_4^{0.8}x_6^{0.5} \tag{6}$$

From these examples we can see that the attractor is less stable in cases where $g_{ij}h_{ij} \neq 0$ (see Example 1). Also, we observed that the more non-zero parameters occur in the equations the harder it is to identify the attractor (see Example 4). Finally, when we have parameters of smaller magnitude (relative to other parameters of the S-system) the corresponding co-ordinate of the attractor tend to become more difficult to estimate (Example 3 and 4).

2 Theoretical noise

As we stated in the discussion section of our paper the error of optimal parameter is proportional to the amount of relative noise in the data. This is a direct consequence of the following theorem:

Theorem 1. *Let σ denote the relative noise of the measurement values, and* $D = diag(x_1^2(t_1), x_1^2(t_2), \ldots, x_n^2(t_N), 2\dot{x}_i^2(t_1), 2\dot{x}_i^2(t_2), \ldots, 2\dot{x}_i^2(t_N))$, $S = \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{p}_i} f_j^\top\right)^{-1} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{X}} f_j^\top\right)$. *Then* $cov(\Delta \mathbf{p}_i) \approx \sigma^2 S \cdot D \cdot S^\top$

Proof. By defining $\mathbf{x} = (x_{11}, x_{12}, \ldots, x_{nN})^\top$, $\mathbf{dx}_i = (dx_{i1}, x_{i2}, \ldots, x_{iN})^\top$, $\mathbf{X} = (\mathbf{x}^\top, \mathbf{dx}_i^\top)^\top$ and $\mathbf{X}_0 = (x_1(t_1), x_1(t_2), \ldots, x_n(t_N), \dot{x}_i(t_1), \dot{x}_i(t_2), \ldots, \dot{x}_i(t_N))^\top$ we obtain

$$f(\mathbf{p}_i) = \sum_{j=1}^N \left(dx_i(j) - \alpha_i \prod_{k=1}^n x_k(j)^{g_{i,k}} + \beta_i \prod_{k=1}^n x_k(j)^{h_{i,k}} \right)^2 = \sum_{j=1}^N f_j^2(\mathbf{x}, \mathbf{dx}_i, \mathbf{p}_i) = \sum_{j=1}^N f_j^2(\mathbf{X}, \mathbf{p}_i)$$

Let $\mathbf{p}_i^*$ denote the true underlying parameter vector, $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}_0$ and $\Delta\mathbf{p}_i = \mathbf{p}_i - \mathbf{p}_i^*$. Since $f_j(\mathbf{X}_0, \mathbf{p}_i^*) = 0$ for all i, j the first order Taylor approximation yields

$$\begin{aligned} f(\mathbf{p}_i) &= \sum_{j=1}^N f_j^2(\mathbf{x}, \mathbf{dx}_i, \mathbf{p}_i) \approx \sum_{j=1}^N \left(\frac{\partial}{\partial \mathbf{X}} f_j(\mathbf{X}_0, \mathbf{p}_i^*)^\top \Delta\mathbf{X} + \frac{\partial}{\partial \mathbf{p}_i} f_j(\mathbf{X}_0, \mathbf{p}_i^*)^\top \Delta\mathbf{p}_i \right)^2 = \\ &= \Delta\mathbf{X}^\top \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{X}} f_j(\mathbf{X}_0, \mathbf{p}_i^*) \frac{\partial}{\partial \mathbf{X}} f_j(\mathbf{X}_0, \mathbf{p}_i^*)^\top \right) \Delta\mathbf{X} \\ &+ 2\Delta\mathbf{p}_i^{*\top} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j(\mathbf{X}_0, \mathbf{p}_i^*) \frac{\partial}{\partial \mathbf{X}} f_j(\mathbf{X}_0, \mathbf{p}_i^*)^\top \right) \Delta\mathbf{X} \\ &+ \Delta\mathbf{p}_i^{*\top} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j(\mathbf{X}_0, \mathbf{p}_i^*) \frac{\partial}{\partial \mathbf{p}_i} f_j(\mathbf{X}_0, \mathbf{p}_i^*)^\top \right) \Delta\mathbf{p}_i \end{aligned}$$

Thus the minimization of $f(\mathbf{p}_i)$ is locally equivalent to

$$\min_{\Delta\mathbf{p}_i} 2\Delta\mathbf{p}_i^{*\top} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{X}} f_j^\top \right) \Delta\mathbf{X} + \Delta\mathbf{p}_i^{*\top} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{p}_i} f_j^\top \right) \Delta\mathbf{p}_i$$

which yields the solution

$$\Delta\mathbf{p}_i \approx -\left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{p}_i} f_j^\top \right)^{-1} \left(\sum_{j=1}^N \frac{\partial}{\partial \mathbf{p}_i} f_j \frac{\partial}{\partial \mathbf{X}} f_j^\top \right) \Delta\mathbf{X}$$

Using the notations defined in the theorem the covariance matrix of the error of the parameter estimation is

$$\mathrm{cov}(\Delta\mathbf{p}_i) = \mathrm{E}[\Delta\mathbf{p}_i \Delta\mathbf{p}_i^\top] \approx S \cdot \mathrm{cov}(\Delta\mathbf{X}) \cdot S^\top = \sigma^2 S \cdot D \cdot S^\top \tag{7}$$

as required. □

3 Full results for the 30-dimensional example

In the Table below we present the complete results for the 30-dimensional example in the paper.

	2% NOISE						5% NOISE					
i	α_i	$g_{i,j}$			β_i	$h_{i,i}$	α_i	$g_{i,j}$			β_i	$h_{i,i}$
1	0.0077	0.0174	–	–	0.0094	0.0059	0.0455	0.0791	–	0	0.0658	0.0414
2	0.0091	–	–	–	0.0115	0.0075	0.0404	–	0	–	0.06	0.0386
3	0.0109	–	–	–	0.014	0.0089	0.0657	–	0	–	0.0872	0.0511
4	0.0045	–	–	–	0.0064	0.0042	0.0441	–	0	–	0.0648	0.0384
5	0.1331	0.235	–	–	0.1188	0.0064	0.1753	0.2599	–	0	0.1557	0.0395
6	0.1335	0.2403	–	–	0.1126	0.0079	0.0573	0.0354	–	0	0.0611	0.0390
7	0.0059	0.0066	0.005	–	0.0085	0.0053	0.2931	0.5807	1.019	–	0.265	0.0465
8	0.1338	0.1503	0.1372	–	0.0997	0.0094	0.1815	1.2076	0.2895	–	0.1606	0.0516
9	0.009	0.0049	0.0135	–	0.0103	0.0068	0.1669	0.1855	0.0152	–	0.1552	0.2676
10	0.0085	0.0078	–	–	0.0118	0.0082	0.1656	0.8379	–	0	0.1385	0.0386
11	0.0148	0.0086	0.0084	0.012	0.0155	0.0106	0.0302	0.017	0.0169	0.0256	0.0491	0.0344
12	0.1277	0.076	–	–	0.102	0.0053	0.0488	0.0737	–	0	0.0638	0.0387
13	0.0085	0.0073	–	–	0.0084	0.0061	0.0553	0.0505	–	0	0.0562	0.0355
14	0.0089	0.0054	–	–	0.0095	0.0057	0.0798	0.0457	–	0	0.0814	0.0474
15	0.0049	0.0146	–	–	0.0054	0.0033	0.0591	0.0639	–	0	0.0791	0.0495
16	0.0112	0.0102	0.0151	–	0.0135	0.0083	0.0528	0.0462	0.0417	–	0.0674	0.0420
17	0.0097	0.0087	–	–	0.0139	0.0091	0.1644	0.633	–	0	0.1437	0.0411
18	0.0131	–	–	–	0.0169	0.0108	0.0434	–	0	–	0.0612	0.0384
19	0.0115	0.0346	–	–	0.014	0.0093	0.0432	0.0804	–	0	0.0614	0.0356
20	0.0097	0.0061	0.0082	–	0.014	0.0087	0.0608	0.0427	0.029	–	0.0724	0.0447
21	0.1372	0.4424	–	–	0.1051	0.0113	0.0558	0.0438	–	0	0.0674	0.0414
22	0.0105	0.0094	–	–	0.0127	0.0081	0.067	0.0522	–	0	0.0814	0.0485
23	0.1333	1.2453	–	–	0.0975	0.009	0.1737	0.3687	–	0	0.1437	0.0444
24	0.0082	0.0068	0.0167	0.0087	0.0099	0.0062	0.0536	0.0278	0.066	0.0503	0.0704	0.0453
25	0.0093	0.0077	–	–	0.0112	0.0071	0.052	0.0551	–	0	0.0625	0.0395
26	0.0113	0.011	0.0223	–	0.014	0.0085	0.0449	0.0234	0.0948	–	0.0567	0.0353
27	0.2531	0.7511	0.466	0.0588	0.237	0.279	0.0414	0.0323	0.0253	0.0219	0.0519	0.0350
28	0.0118	0.0109	–	–	0.0128	0.0073	0.0397	0.0343	–	0	0.0494	0.0318
29	0.0131	0.011	–	–	0.0149	0.0089	0.1675	0.8255	–	0	0.1295	0.0375
30	0.1367	0.4356	–	–	0.0959	0.0106	0.0733	0.0514	–	0	0.0883	0.0538
	10% NOISE						20% NOISE					
i	α_i	$g_{i,j}$			β_i	$h_{i,i}$	α_i	$g_{i,j}$			β_i	$h_{i,i}$
1	0.1704	0.0845	–	–	0.2447	0.1379	0.4872	0.3167	–	–	0.7432	0.3469
2	0.1636	–	–	–	0.2393	0.1361	0.4755	–	–	–	0.7173	0.3381
3	0.1904	–	–	–	0.2673	0.1465	0.4706	–	–	–	0.7180	0.3417
4	0.1393	–	–	–	0.2144	0.1249	0.3966	–	–	–	0.6231	0.3100
5	0.2453	0.1349	–	–	0.2532	0.1404	0.6607	0.3350	–	–	0.6479	0.3292
6	0.2582	0.1363	–	–	0.2753	0.1521	0.7509	0.3441	–	–	0.7515	0.3463
7	0.1815	0.0994	0.1146	–	0.2379	0.1375	0.5943	0.3196	0.2871	–	0.7148	0.3402
8	0.1963	0.1527	0.0969	–	0.2656	0.1454	0.4706	0.3749	0.1906	–	0.6570	0.3199
9	0.2325	0.1314	0.0705	–	0.2315	0.1322	0.6315	0.3199	0.1282	–	0.6394	0.3236
10	0.2203	0.2099	–	–	0.2735	0.149	0.4383	0.2868	–	–	0.6512	0.3221
11	0.1551	0.1006	0.0853	0.1012	0.2053	0.1213	0.5234	0.3017	0.1967	0.2759	0.6829	0.3313
12	0.1676	0.1326	–	–	0.2346	0.1289	0.4503	0.3933	–	–	0.6785	0.3314
13	0.2094	0.1535	–	–	0.2398	0.138	0.4719	0.3215	–	–	0.6200	0.3166
14	0.2452	0.1391	–	–	0.2384	0.1331	0.6829	0.3347	–	–	0.6529	0.3202
15	0.1874	0.1549	–	–	0.2569	0.1462	0.5458	0.3639	–	–	0.7879	0.3619
16	0.1691	0.1343	0.0641	–	0.2177	0.1269	0.5228	0.3626	0.2062	–	0.6737	0.3358
17	0.1851	0.1375	–	–	0.2427	0.1362	0.5719	0.3902	–	–	0.7128	0.3358
18	0.1496	–	–	–	0.2185	0.1261	0.5165	–	–	–	0.7717	0.3550
19	0.1453	0.1894	–	–	0.2113	0.1234	0.5468	0.5373	–	–	0.7659	0.3556
20	0.1811	0.1172	0.093	–	0.2147	0.1306	0.5896	0.2940	0.2435	–	0.7044	0.3330
21	0.1832	0.1342	–	–	0.2254	0.1323	0.5200	0.3221	–	–	0.6678	0.3292
22	0.1989	0.1565	–	–	0.2303	0.1284	0.5223	0.3500	–	–	0.6581	0.3181
23	0.1442	0.1216	–	–	0.2165	0.1282	0.5127	0.3551	–	–	0.7542	0.3521
24	0.1875	0.1167	0.156	0.1527	0.2601	0.152	0.5327	0.2607	0.2520	0.3575	0.7427	0.3466
25	0.1709	0.145	–	–	0.2157	0.1265	0.5867	0.3759	–	–	0.7729	0.3498
26	0.1424	0.0945	0.1181	–	0.2063	0.119	0.5064	0.3087	0.3521	–	0.7465	0.3455
27	0.2094	0.1381	0.1076	0.0620	0.2468	0.142	0.5359	0.3220	0.2626	0.1552	0.6390	0.3254
28	0.1893	0.15	–	–	0.2409	0.137	0.5280	0.3363	–	–	0.7015	0.3398
29	0.1726	0.1279	–	–	0.2428	0.139	0.5182	0.3310	–	–	0.7094	0.3348
30	0.1951	0.1405	–	–	0.2462	0.1412	0.5579	0.3364	–	–	0.7000	0.3356

Table 1: Median relative error of the parameters for all noise levels. The $g_{i,j}$ column is filled with the non-zero parameters appearing in the given equation. For instance, in line 8 (2% relative noise) the relative errors for parameters $g_{8,4}$, $g_{8,17}$ can be read in the $g_{i,j}$ column: 0.1503 and 0.1372, respectively.

4 Pseudo code for the topology searching algorithm

We denote our algorithm as function ALG, whose input is the DATA and the network topology ($\overrightarrow{N}$) and the output is the parameter estimates ($\mathbf{p}$), the residuals ($f(\mathbf{p})$) and the R^2 value obtained for the attractor fitting. In the following algorithm N_i denotes all the networks whose edges point to vertex i.

TOPOLOGY SEARCH ALGORITHM (INPUT : DATA, vertex i,OUTPUT : $\overrightarrow{N}_{j^*}$, p_{j^*})

FOR i=1 TO n

  $\mathcal{N}^s = \mathcal{N}_i$

  $residual = \infty$

  WHILE $\mathcal{N}^s \neq \emptyset$

    $\mathcal{N}^* = \{\overrightarrow{N} \in \mathcal{N}^s : \overrightarrow{N} \text{ is minimal in } \mathcal{N}^s\} = \{\overrightarrow{N}_j : j = 1, \ldots, J\}$

    $\mathcal{N}^s = \mathcal{N}^s \setminus \mathcal{N}^*$

    FOR $j = 1$ TO J

      $(p[j], res[j], R^2[j]) =$ ALG(DATA, $\overrightarrow{N}_j$)

      IF $R^2[j]) < 0.9$

        $res[j] = \infty$

        $\mathcal{N}^s = \mathcal{N}^s \setminus \{\overrightarrow{N} : \overrightarrow{N}_j \prec \overrightarrow{N}\}$

      END

      $j^* = \text{argmin}_j \ res[j]$

      IF $res[j^*] < residual$

        $fin[i] = p[j^*]$

        $residual = res[j^*]$

      END

    END

  END

END

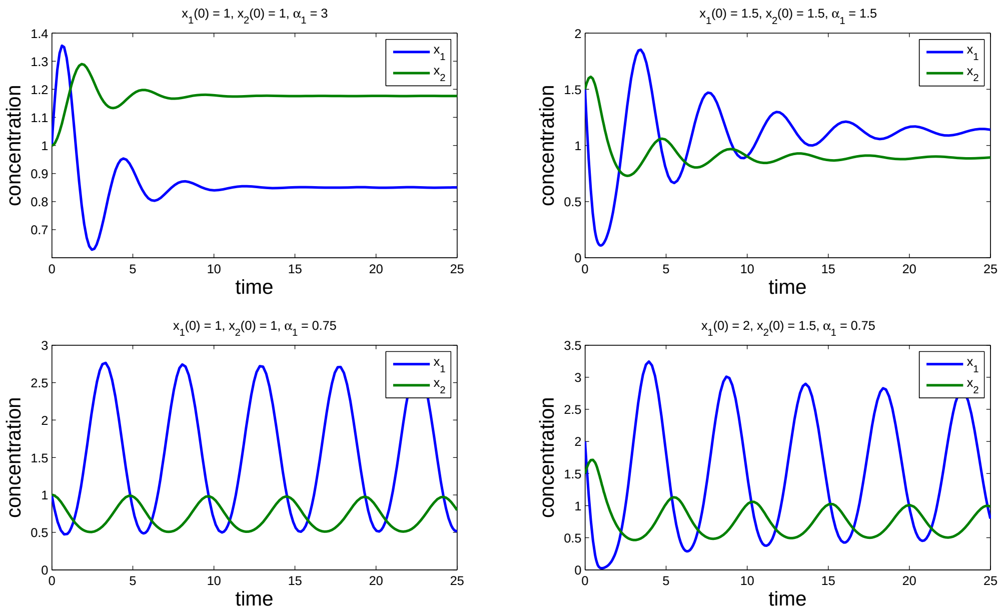

Figure 1: Different behaviours of the 2-dimensional system (Eq.(1-4)) depending on the initial concentrations and α_1.

References

[1] S Kimura, K Ide, A Kashihara, M Kano, M Hatakeyama, R Masui, N Nakagawa, S Yokoyama, S Kuramitsu, and A Konagaya. Inference of s-system models of genetic networks using a cooperative coevolutionary algorithm. *Bioinformatics*, 21:1154–1163, 2005.

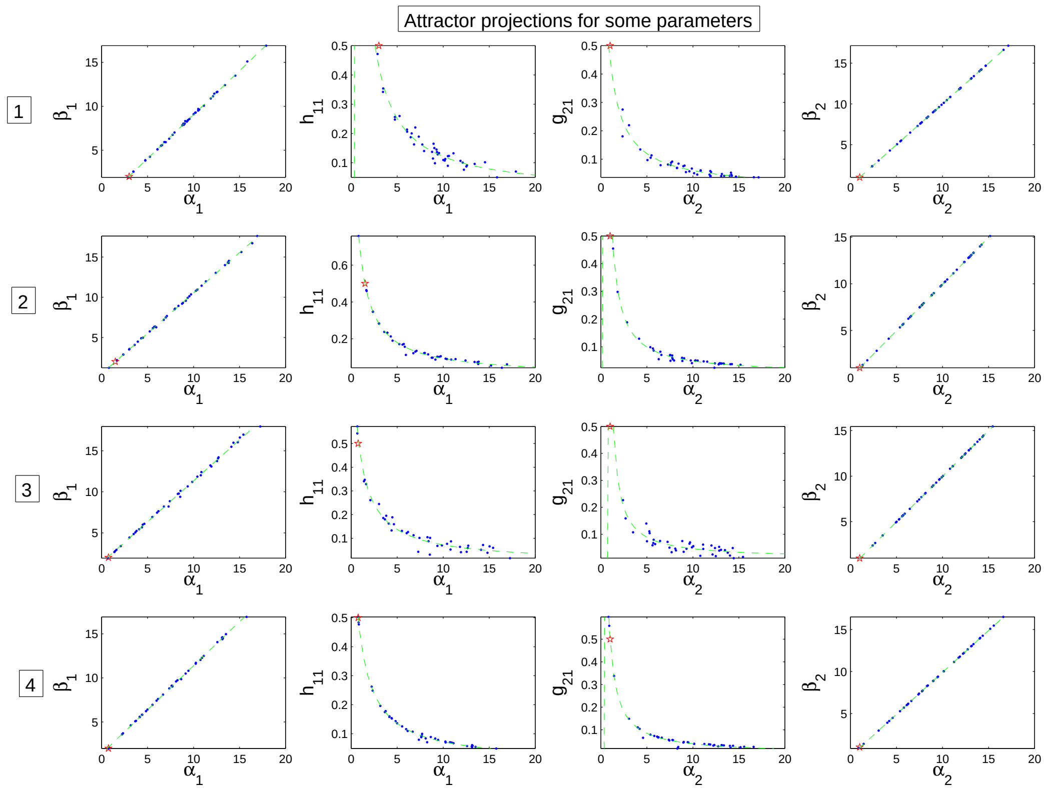

Figure 2: Projections of the Newton candidates and the fitted attractor curve for the 4 different systems described in Eq.(1-4). (1) $\alpha_1 = 3, x_1(0) = 1, x_2(0) = 1$, (2) $\alpha_1 = 1.5, x_1(0) = 1.5, x_2(0) = 1.5$ (3) $\alpha_1 = 0.75, x_1(0) = 1, x_2(0) = 1$, (4) $\alpha_1 = 0.75, x_1(0) = 2, x_2(0) = 1.5$

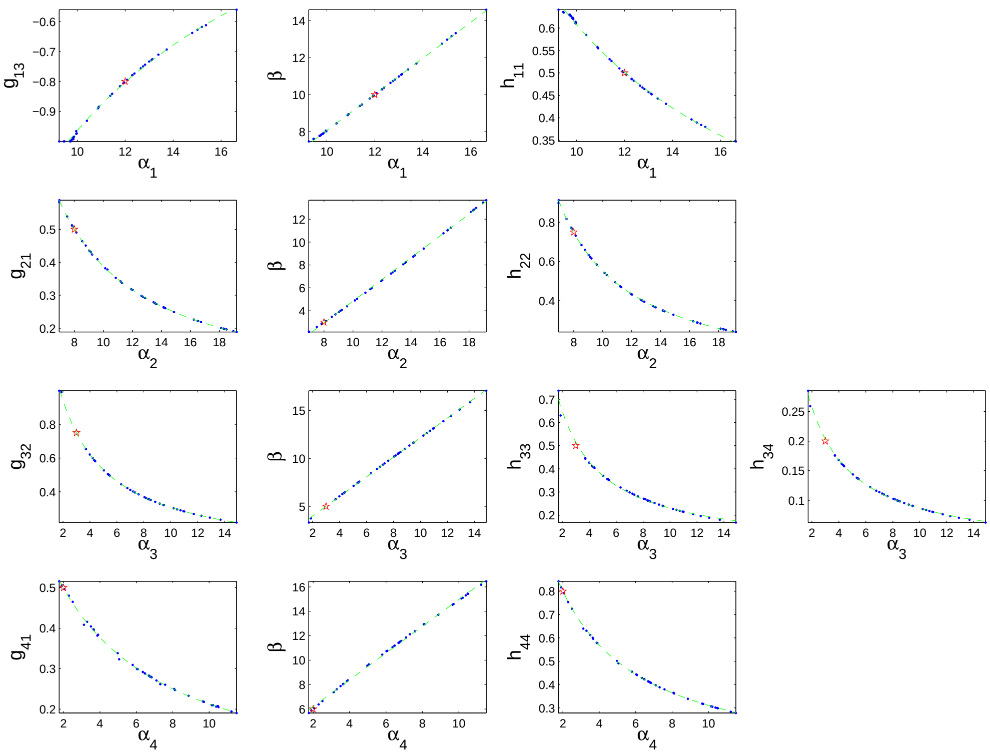

Figure 3: Projections of the Newton candidates and the fitted attractor curves for the 4-dimensional example (Eq.(5)).

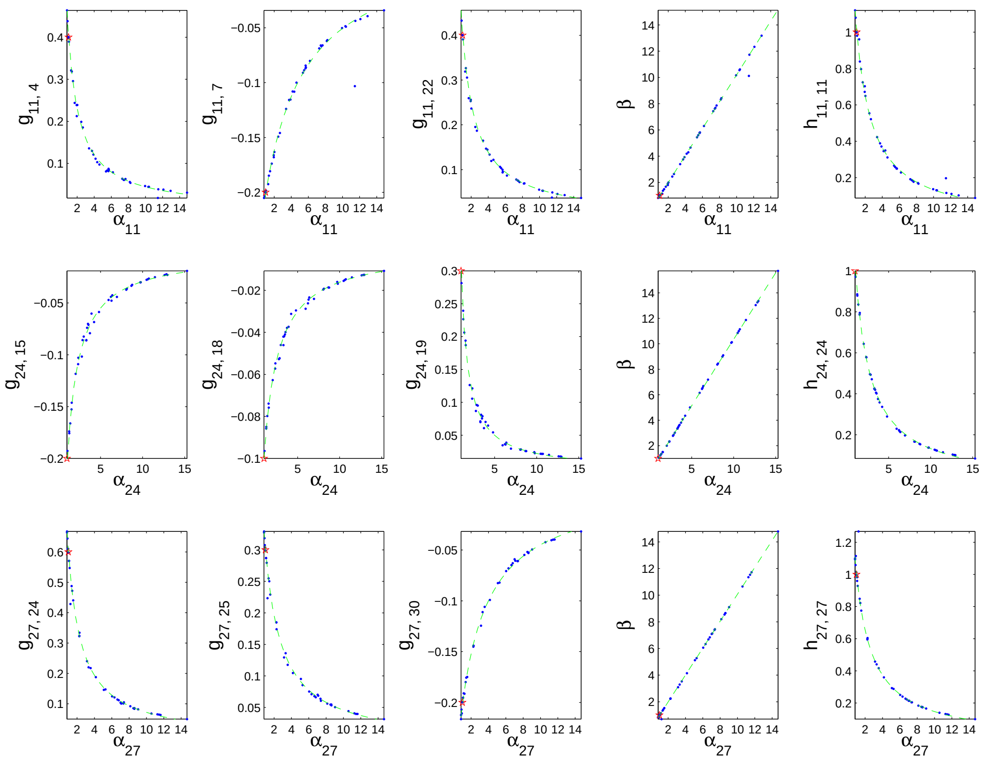

Figure 4: Projections of the Newton candidates and the fitted attractor curve for Eqs. (11), (24), (27) for the 30-dim example.

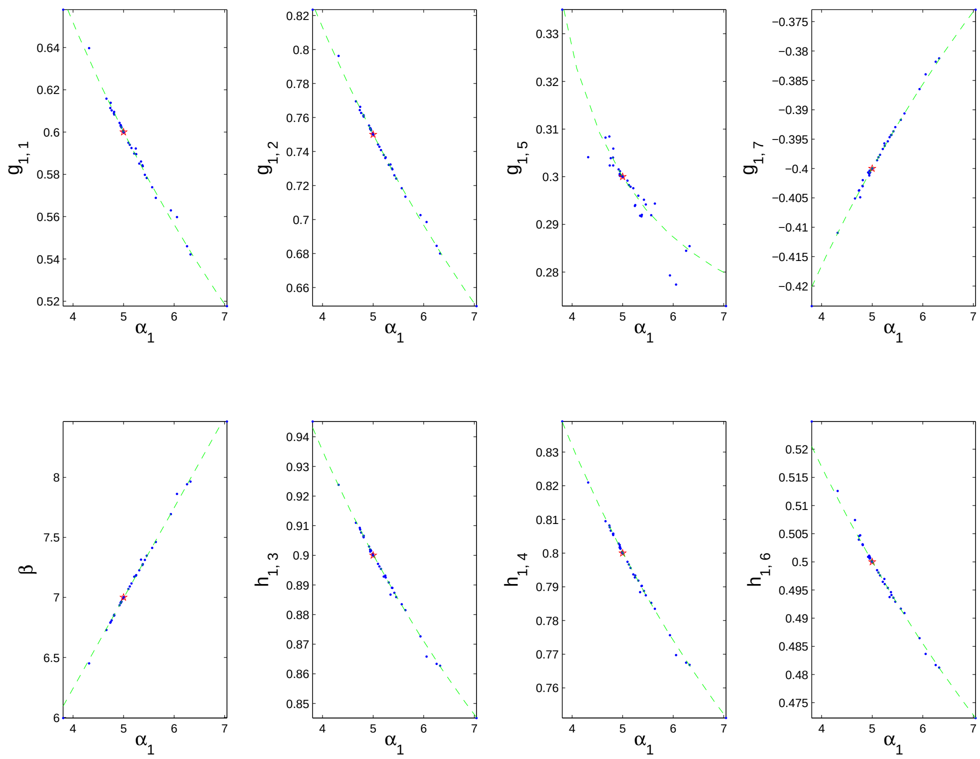

Figure 5: Projections of the Newton candidates and the fitted attractor curves for the algebraic equation (Eq.(6)).