Structural identifiability and observability analysis of the AKT pathway model

In this example we analyse the AKT pathway model from Fujita et al. (*Science Signaling*, 3(132):ra56–ra56, 2010). This model has 9 states, 19 parameters (excluding 3 unknown initial conditions), 3 measured outputs, and one input, EGF. The experimental setup is such that there are 6 experimental conditions, each one with a different constant value of EGF. Several initial conditions are known and equal to zero.

We begin by installing STRIKE-GOLDD:

```
clear
clc
cd('options_files')
copyfile("options_*","../../strike-goldd/STRIKE-GOLDD")
cd('../../strike-goldd/STRIKE-GOLDD')
install
```

```
STRIKE-GOLDD (v3.0) folders added to the path
```

(1) We first analyse the model considering just one experiment with constant EGF (i.e., setting the option: opts.nnzDerU = 0):

```
STRIKE_GOLDD('options_Fujita_1exp_constEGF_zeroICs.m')
```

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the Fujita model...

>>> The model contains:
9 states:
[EGFR; pEGFR; pEGFR_Akt; Akt; pAkt; S6; pAkt_S6; pS6; EGF_EGFR]
3 outputs:
[scaleFactor_pEGFR*(pEGFR + pEGFR_Akt); scaleFactor_pAkt*(pAkt + pAkt_S6); pS6*scaleFactor_pS6]
1 known inputs:
EGF_u
0 unknown inputs:

16 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 8 Lie derivatives
    Calculating derivatives: 1 2 3 4 5 6 7 8
>>> Observability-Identifiability matrix built with 8 Lie derivatives
    (calculated in 1.826865e+01 seconds)
>>> Calculating rank...
    Rank = 21 (calculated in 1.016917e+01 seconds)
>>> Observability-Identifiability matrix built with 9 Lie derivatives
    (calculated in 4.086958e+01 seconds)
>>> Calculating rank...
    Rank = 21 (calculated in 5.224002e+00 seconds)
   => Parameter scaleFactor_pEGFR is structurally unidentifiable
   => Parameter scaleFactor_pAkt is structurally unidentifiable
   => Parameter scaleFactor_pS6 is structurally unidentifiable
   => Parameter EGFR_turnover is structurally unidentifiable
   => Parameter reaction_1_k1 is structurally unidentifiable
   => Parameter reaction_1_k2 is structurally unidentifiable
   => Parameter reaction_2_k1 is structurally unidentifiable
   => Parameter reaction_2_k2 is structurally identifiable
```

```
   => Parameter reaction_3_k1 is structurally identifiable
   => Parameter reaction_4_k1 is structurally identifiable
   => Parameter reaction_5_k1 is structurally unidentifiable
   => Parameter reaction_5_k2 is structurally unidentifiable
   => Parameter reaction_6_k1 is structurally unidentifiable
   => Parameter reaction_7_k1 is structurally identifiable
   => Parameter reaction_8_k1 is structurally identifiable
   => Parameter reaction_9_k1 is structurally unidentifiable
   => State EGFR is unobservable
   => State pEGFR is observable
   => State pEGFR_Akt is observable
   => State Akt is unobservable
   => State pAkt is observable
   => State S6 is unobservable
   => State pAkt_S6 is observable
   => State pS6 is observable
   => State EGF_EGFR is observable

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is structurally unidentifiable.
>>> These parameters are identifiable:
     [reaction_2_k2, reaction_3_k1, reaction_4_k1, reaction_7_k1, reaction_8_k1]
>>> These parameters are unidentifiable:
     [EGFR_turnover, reaction_1_k1, reaction_1_k2, reaction_2_k1, reaction_5_k1, reaction_5_k2, reaction_6_k1, reac

>>> These states are observable (and their initial conditions, if considered unknown, are identifiable):
     [EGF_EGFR, pAkt, pAkt_S6, pEGFR, pEGFR_Akt, pS6]
>>> These states are unobservable (and their initial conditions, if considered unknown, are unidentifiable):
     [Akt, EGFR, S6]
>>> These inputs are known:
     EGF_u
Total execution time: 8.356697e+01
```

The analysis yields that the model is unidentifiable from one experiment with a constant input.

(2) Let us see what happens if we can manipulate the input so that it varies during the experiment:

```
STRIKE_GOLDD('options_Fujita_1exp_tvarEGF.m')
```

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the Fujita model...

>>> The model contains:
9 states:
[EGFR; pEGFR; pEGFR_Akt; Akt; pAkt; S6; pAkt_S6; pS6; EGF_EGFR]
3 outputs:
[scaleFactor_pEGFR*(pEGFR + pEGFR_Akt); scaleFactor_pAkt*(pAkt + pAkt_S6); pS6*scaleFactor_pS6]
1 known inputs:
EGF_u
0 unknown inputs:

16 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 8 Lie derivatives
    Calculating derivatives: 1 2 3 4 5 6 7 8
>>> Observability-Identifiability matrix built with 8 Lie derivatives
```

```
    (calculated in 1.952027e+01 seconds)
>>> Calculating rank...
    Rank = 25 (calculated in 9.242405e+01 seconds)

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is Fully Input-State-Parameter Observable (FISPO):
    All its states are observable.
    All its parameters are locally structurally identifiable.
Total execution time: 1.155617e+02
```

In this case the model becomes FISPO, i.e. structurally identifiable and observable...

... although we have not taken into account that six initial conditions are set to zero in the model definition, which may alter the results.

(3) Let us repeat the analysis, setting said initial conditions to zero:

```
STRIKE_GOLDD('options_Fujita_1exp_tvarEGF_zeroICs.m')
```

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the Fujita model...

>>> The model contains:
9 states:
[EGFR; pEGFR; pEGFR_Akt; Akt; pAkt; S6; pAkt_S6; pS6; EGF_EGFR]
3 outputs:
[scaleFactor_pEGFR*(pEGFR + pEGFR_Akt); scaleFactor_pAkt*(pAkt + pAkt_S6); pS6*scaleFactor_pS6]
1 known inputs:
EGF_u
0 unknown inputs:

16 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 8 Lie derivatives
    Calculating derivatives: 1 2 3 4 5 6 7 8
>>> Observability-Identifiability matrix built with 8 Lie derivatives
    (calculated in 1.938023e+01 seconds)
>>> Calculating rank...
    Rank = 21 (calculated in 1.187657e+01 seconds)
>>> Observability-Identifiability matrix built with 9 Lie derivatives
    (calculated in 4.555061e+01 seconds)
>>> Calculating rank...
    Rank = 23 (calculated in 1.143886e+01 seconds)
>>> Observability-Identifiability matrix built with 10 Lie derivatives
    (calculated in 6.728718e+01 seconds)
>>> Calculating rank...
    Rank = 23 (calculated in 4.163040e+01 seconds)
   => Parameter scaleFactor_pEGFR is structurally identifiable
   => Parameter scaleFactor_pAkt is structurally identifiable
   => Parameter scaleFactor_pS6 is structurally unidentifiable
   => Parameter EGFR_turnover is structurally identifiable
   => Parameter reaction_1_k1 is structurally identifiable
   => Parameter reaction_1_k2 is structurally identifiable
   => Parameter reaction_2_k1 is structurally identifiable
   => Parameter reaction_2_k2 is structurally identifiable
```

```
   => Parameter reaction_3_k1 is structurally identifiable
   => Parameter reaction_4_k1 is structurally identifiable
   => Parameter reaction_5_k1 is structurally unidentifiable
   => Parameter reaction_5_k2 is structurally unidentifiable
   => Parameter reaction_6_k1 is structurally unidentifiable
   => Parameter reaction_7_k1 is structurally identifiable
   => Parameter reaction_8_k1 is structurally identifiable
   => Parameter reaction_9_k1 is structurally identifiable
   => State EGFR is observable
   => State pEGFR is observable
   => State pEGFR_Akt is observable
   => State Akt is observable
   => State pAkt is observable
   => State S6 is unobservable
   => State pAkt_S6 is observable
   => State pS6 is observable
   => State EGF_EGFR is observable

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is structurally unidentifiable.
>>> These parameters are identifiable:
     [EGFR_turnover, reaction_1_k1, reaction_1_k2, reaction_2_k1, reaction_2_k2, reaction_3_k1, reaction_4_k1, reac
>>> These parameters are unidentifiable:
     [reaction_5_k1, reaction_5_k2, reaction_6_k1, scaleFactor_pS6]

>>> These states are observable (and their initial conditions, if considered unknown, are identifiable):
     [Akt, EGFR, EGF_EGFR, pAkt, pAkt_S6, pEGFR, pEGFR_Akt, pS6]
>>> These states are unobservable (and their initial conditions, if considered unknown, are unidentifiable):
     S6
>>> These inputs are known:
     EGF_u
Total execution time: 2.963581e+02
```

Thus, after considering that several initial conditions are zero, four parameters become unidentifiable and one estate becomes unobservable.

(4) Let us now see if the model is FISPO under assumptions that resemble the real experimental conditions:

- In reality, the input cannot be manipulated so as to change continuously during the experiment, so we consider it constant.
- In reality, we know that six initial conditions are zero, so we fix them to zero.
- In reality, we can perform several experiments with different values of the input EGF. Hence, we analyse the model in a multi-experiment setting, starting with two experiments.

```
STRIKE_GOLDD('options_Fujita_2exp_constEGF_zeroICs.m')
```

```
-------------------------------
>>> STRIKE-GOLDD toolbox 3.0
-------------------------------

Analyzing the Fujita_2Exp model...

>>> The model contains:
18 states:
[EGFRExp1; pEGFRExp1; pEGFR_AktExp1; AktExp1; pAktExp1; S6Exp1; pAkt_S6Exp1; pS6Exp1; EGF_EGFRExp1; EGFRExp2; pEGFR
6 outputs:
```

```
[scaleFactor_pEGFR*(pEGFRExp1 + pEGFR_AktExp1); scaleFactor_pAkt*(pAktExp1 + pAkt_S6Exp1); pS6Exp1*scaleFactor_pS6;
2 known inputs:
[EGF_uExp1; EGF_uExp2]
0 unknown inputs:
[]
16 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 5 Lie derivatives
    Calculating derivatives: 1 2 3 4 5
>>> Observability-Identifiability matrix built with 5 Lie derivatives
    (calculated in 8.054194e-01 seconds)
>>> Calculating rank...
    Rank = 24 (calculated in 4.496741e-01 seconds)
>>> Observability-Identifiability matrix built with 6 Lie derivatives
    (calculated in 1.262137e+00 seconds)
>>> Calculating rank...
    Rank = 28 (calculated in 6.593355e-01 seconds)
>>> Observability-Identifiability matrix built with 7 Lie derivatives
    (calculated in 1.753684e+00 seconds)
>>> Calculating rank...
    Rank = 31 (calculated in 1.087147e+00 seconds)
>>> Observability-Identifiability matrix built with 8 Lie derivatives
    (calculated in 2.718064e+00 seconds)
>>> Calculating rank...
    Rank = 32 (calculated in 2.152553e+00 seconds)
>>> Observability-Identifiability matrix built with 9 Lie derivatives
    (calculated in 4.953649e+00 seconds)
>>> Calculating rank...
    Rank = 32 (calculated in 5.336677e+00 seconds)
   => Parameter scaleFactor_pEGFR is structurally identifiable
   => Parameter scaleFactor_pAkt is structurally identifiable
   => Parameter scaleFactor_pS6 is structurally unidentifiable
   => Parameter EGFR_turnover is structurally identifiable
   => Parameter reaction_1_k1 is structurally identifiable
   => Parameter reaction_1_k2 is structurally identifiable
   => Parameter reaction_2_k1 is structurally identifiable
   => Parameter reaction_2_k2 is structurally identifiable
   => Parameter reaction_3_k1 is structurally identifiable
   => Parameter reaction_4_k1 is structurally identifiable
   => Parameter reaction_5_k1 is structurally unidentifiable
   => Parameter reaction_5_k2 is structurally unidentifiable
   => Parameter reaction_6_k1 is structurally unidentifiable
   => Parameter reaction_7_k1 is structurally identifiable
   => Parameter reaction_8_k1 is structurally identifiable
   => Parameter reaction_9_k1 is structurally identifiable
   => State EGFRExp1 is observable
   => State pEGFRExp1 is observable
   => State pEGFR_AktExp1 is observable
   => State AktExp1 is observable
   => State pAktExp1 is observable
   => State S6Exp1 is unobservable
   => State pAkt_S6Exp1 is observable
   => State pS6Exp1 is observable
   => State EGF_EGFRExp1 is observable
   => State EGFRExp2 is observable
   => State pEGFRExp2 is observable
   => State pEGFR_AktExp2 is observable
   => State AktExp2 is observable
   => State pAktExp2 is observable
   => State S6Exp2 is unobservable
   => State pAkt_S6Exp2 is observable
   => State pS6Exp2 is observable
   => State EGF_EGFRExp2 is observable
```

```
------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is structurally unidentifiable.
>>> These parameters are identifiable:
     [EGFR_turnover, reaction_1_k1, reaction_1_k2, reaction_2_k1, reaction_2_k2, reaction_3_k1, reaction_4_k1, reac
>>> These parameters are unidentifiable:
     [reaction_5_k1, reaction_5_k2, reaction_6_k1, scaleFactor_pS6]

>>> These states are observable (and their initial conditions, if considered unknown, are identifiable):
     [AktExp1, AktExp2, EGFRExp1, EGFRExp2, EGF_EGFRExp1, EGF_EGFRExp2, pAktExp1, pAktExp2, pAkt_S6Exp1, pAkt_S6Exp
>>> These states are unobservable (and their initial conditions, if considered unknown, are unidentifiable):
     [S6Exp1, S6Exp2]
>>> These inputs are known:
     [EGF_uExp1; EGF_uExp2]
Total execution time: 6.056221e+01
```

Under these assumptions we obtain similar results as with the analysis number (3), i.e. one experiment with time-varying input. Thus, by going from one to two experiments with constant input(s) we obtain better structural identifiability and observability results, but not FISPO yet (at least with 6 initial conditions being zero).

(5) If we relax the requirement that six initial conditions must be zero, we obtain that the model is FISPO with two experiments with constant inputs:

```
STRIKE_GOLDD('options_Fujita_2exp_constEGF.m')
```

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the Fujita_2Exp model...

>>> The model contains:
18 states:
[EGFRExp1; pEGFRExp1; pEGFR_AktExp1; AktExp1; pAktExp1; S6Exp1; pAkt_S6Exp1; pS6Exp1; EGF_EGFRExp1; EGFRExp2; pEGFR
6 outputs:
[scaleFactor_pEGFR*(pEGFRExp1 + pEGFR_AktExp1); scaleFactor_pAkt*(pAktExp1 + pAkt_S6Exp1); pS6Exp1*scaleFactor_pS6;
2 known inputs:
[EGF_uExp1; EGF_uExp2]
0 unknown inputs:
[]
16 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 5 Lie derivatives
    Calculating derivatives: 1 2 3 4 5
>>> Observability-Identifiability matrix built with 5 Lie derivatives
    (calculated in 7.962068e-01 seconds)
>>> Calculating rank...
    Rank = 34 (calculated in 7.895963e-01 seconds)

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is Fully Input-State-Parameter Observable (FISPO):
    All its states are observable.
    All its parameters are locally structurally identifiable.
Total execution time: 2.851764e+00
```

Thus, the assumption of having 6 initial conditions equal to zero affects the two settings (1 experiment with time-varying input vs 2 experiments with constant inputs) in the same manner.

(6) It can be seen that, as long as the aforementioned initial conditions are zero, we cannot obtain a FISPO model by increasing the number of experiments. Therefore, if we want the model to be FISPO we need to reformulate it. To this end we may fix some parameter(s). It turns out that fixing one parameter is not enough in this case. However, fixing two parameters (reaction_5_k1 and either one of reaction_5_k2, reaction_6_k1, or scaleFactor_pS6) makes the model FISPO:

```
STRIKE_GOLDD('options_Fujita_2exp_constEGF_zeroICs_fix2pars.m')
```

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the Fujita_2Exp model...

>>> The model contains:
18 states:
[EGFRExp1; pEGFRExp1; pEGFR_AktExp1; AktExp1; pAktExp1; S6Exp1; pAkt_S6Exp1; pS6Exp1; EGF_EGFRExp1; EGFRExp2; pEGFR
6 outputs:
[scaleFactor_pEGFR*(pEGFRExp1 + pEGFR_AktExp1); scaleFactor_pAkt*(pAktExp1 + pAkt_S6Exp1); pS6Exp1*scaleFactor_pS6;
2 known inputs:
[EGF_uExp1; EGF_uExp2]
0 unknown inputs:
[]
14 parameters:
[scaleFactor_pEGFR; scaleFactor_pAkt; scaleFactor_pS6; EGFR_turnover; reaction_1_k1; reaction_1_k2; reaction_2_k1;

>>> Building the observability-identifiability matrix requires at least 5 Lie derivatives
    Calculating derivatives: 1 2 3 4 5
>>> Observability-Identifiability matrix built with 5 Lie derivatives
    (calculated in 7.516612e-01 seconds)
>>> Calculating rank...
    Rank = 24 (calculated in 5.010504e-01 seconds)
>>> Observability-Identifiability matrix built with 6 Lie derivatives
    (calculated in 1.191655e+00 seconds)
>>> Calculating rank...
    Rank = 28 (calculated in 7.159455e-01 seconds)
>>> Observability-Identifiability matrix built with 7 Lie derivatives
    (calculated in 1.686557e+00 seconds)
>>> Calculating rank...
    Rank = 31 (calculated in 1.067525e+00 seconds)
>>> Observability-Identifiability matrix built with 8 Lie derivatives
    (calculated in 2.627771e+00 seconds)
>>> Calculating rank...
    Rank = 32 (calculated in 2.157293e+00 seconds)

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is Fully Input-State-Parameter Observable (FISPO):
    All its states are observable.
    All its parameters are locally structurally identifiable.
Total execution time: 8.578743e+00
```