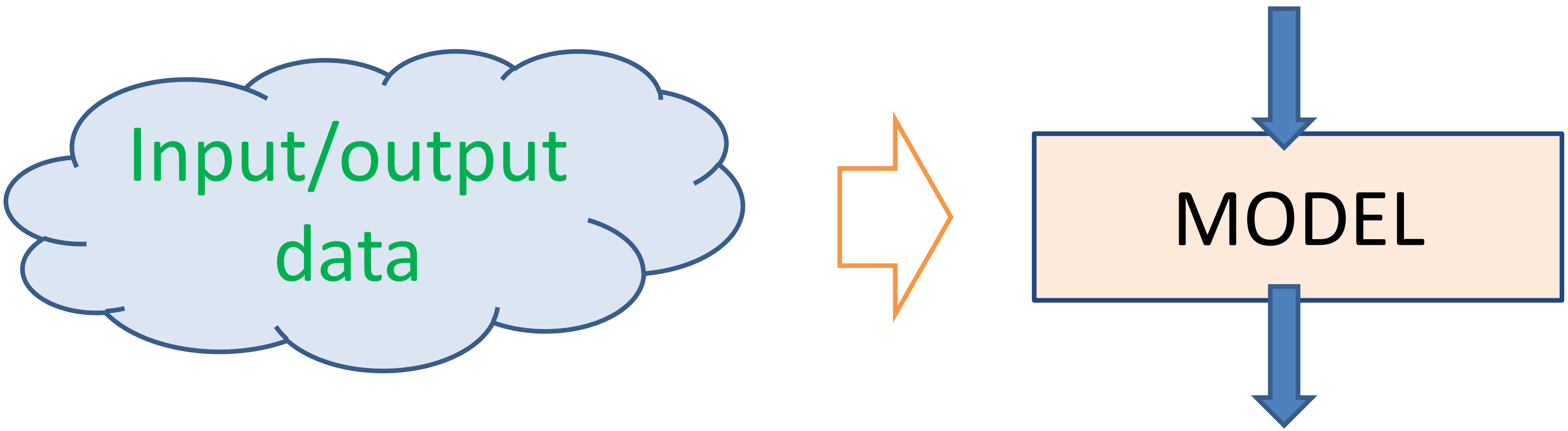
Convex Relaxations in Optimization-Based Identification of Robust Nonlinear Dynamical Models

Alex Megretski, MIT

with M.Tobenkin, I.Manchester, B.Bond, Y.Lin
R.Tedrake, L.Daniel, V.Stojanovic

optimal fitting of rationally parameterized dynamical nonlinear system models

FITTING MODELS TO DATA

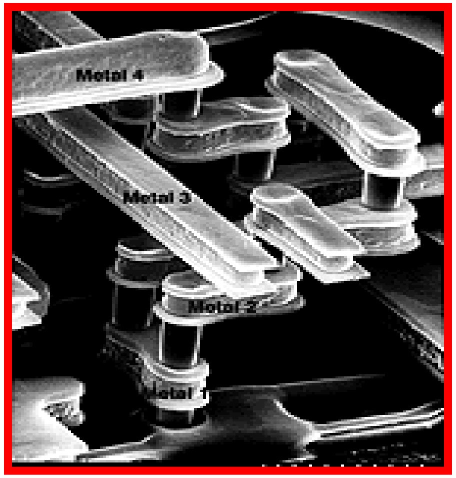


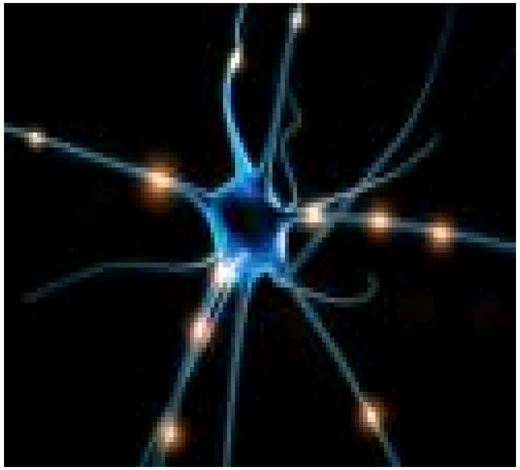
GIVEN:

- ❑ MODEL COMPLEXITY MEASURE
- ❑ DATA MATCHING QUALITY MEASURE
- ❑ INPUT/OUTPUT DATA SET

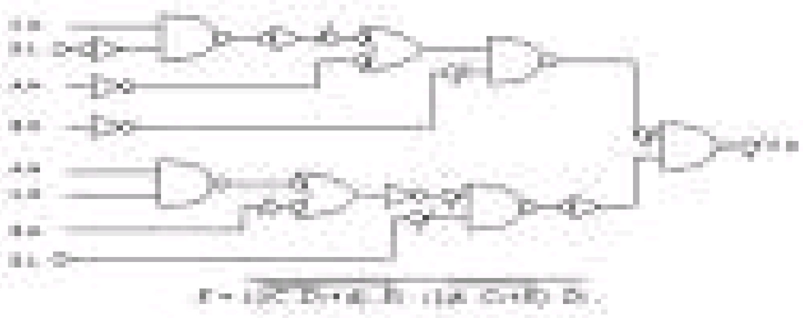
⇒ (SUB)OPTIMAL MODEL OF FIXED COMPLEXITY

EXAMPLES/APPLICATIONS





- ❑ Netflix challenge
- ❑ P vs. NP
- ❑ Circuit modeling
- ❑ Live cell modeling



STATIC MODELS

- DATA: $\{(u_k, y_k)\}_{k=1}^N$
- MODEL: $y = g(u)$
- QUALITY: $\sum |y_k - g(u_k)|^2$

DYNAMIC MODELS

- DATA: $\{(u_k, x_k, y_k, x_k^+)\}_{k=1}^N$
- MODEL: $x^+ = f(x, u),\ \ y = g(x, u)$
- QUALITY: $\sum |y_k - g(\bar{x}_k, u_k)|^2$

LINEAR PARAMETERIZATION

$$y_k \approx g(u_k) \quad (k = 1, \ldots, \bar{k})$$

where $$g(u) = \sum_{i=1}^{r} g_i G_i(u)$$

e.g., in kernel methods, $$G_i(u) = G(u - u_i)$$

- ❑ (relatively) cheap optimization
- ❑ r is not a good complexity measure
- ❑ needs regularization, sparsity optimization

RATIONAL PARAMETERIZATION

$$y_k \approx g(u_k) \quad (k = 1, \ldots, \bar{k})$$

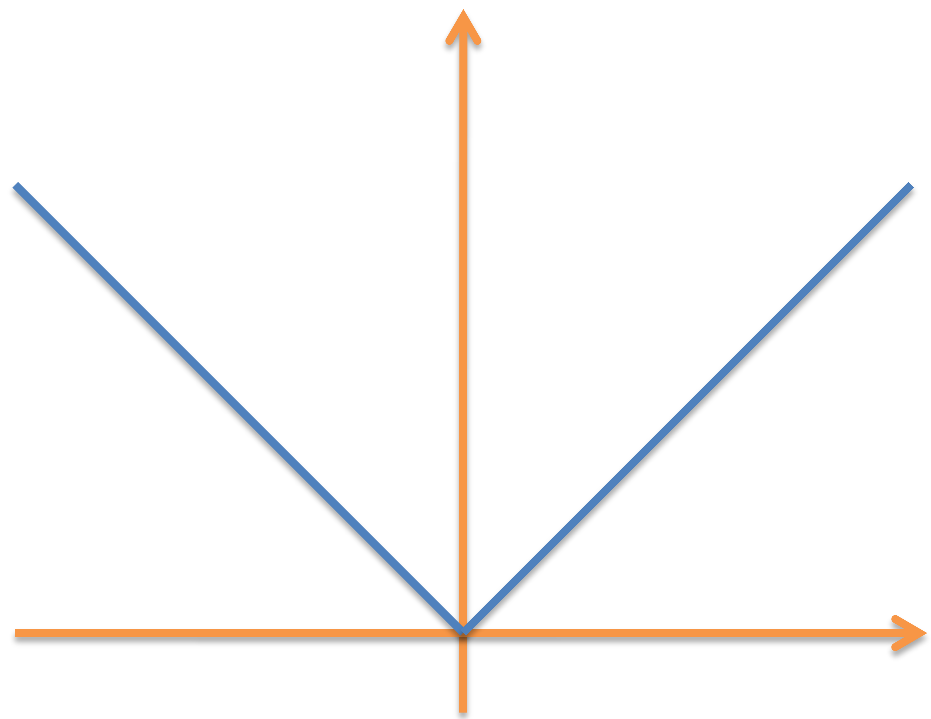
where

$$g(u) = b(u)/a(u)$$
$$b(u) = \sum b_i B_i(u)$$
$$a(u) = \sum a_i A_i(u)$$

- ❑ good: better quality per given complexity
- ❑ bad: tougher to optimize
  - * keeping denominator positive
  - * non-convex setup

EXAMPLE: APPROXIMATING |x|

- consider uniform approximation of $f(x)=|x|$ on $[-1,1]$
- for polynomials of order n, best quality is $O(1/n)$
- for rational functions of order n, best quality is not worse than $3\exp(-\sqrt{n})$



(D. J. Newman, 1963)

ALGEBRAIC PARAMETERIZATION

$$y_k \approx g(u_k) \quad (k = 1, \ldots, \bar{k})$$

where $\quad g(\cdot): \quad h(g(u), u) = 0$

$$h(y, u) = \sum h_i H_i(y, u)$$

- ❑ Linear: $\quad h(y, u) = y - f(u)$
- ❑ Rational: $\quad h(y, u) = a(u)y - b(u)$

difference between equation and output errors!

LINEAR PARAMETERIZATION FOR SYSTEMS

$$y_t = g(x_t, u_t), \quad x_t = \begin{bmatrix} u_{t-m} \\ \vdots \\ u_{t-1} \end{bmatrix}$$

$$g(u, x) = \sum g_i G_i(u, x)$$

e.g. “Volterra Series” (no feedback)

Very inefficient: e.g. $y_t = \sin(y_{t-1} + u_t)$

RATIONAL PARAMETERIZATION (SYSTEMS)

$$h(x_t, u_t, y_t) = 0, \quad x_t = \begin{bmatrix} y_{t-d} \\ \vdots \\ y_{t-1} \\ u_{t-m} \\ \vdots \\ u_{t-1} \end{bmatrix}$$

$$h(x, u, y) = a(x, u)y - b(x, u)$$

EQUATION ERROR VS. OUTPUT ERROR

Having small equation error does not guarantee that the output error is small,

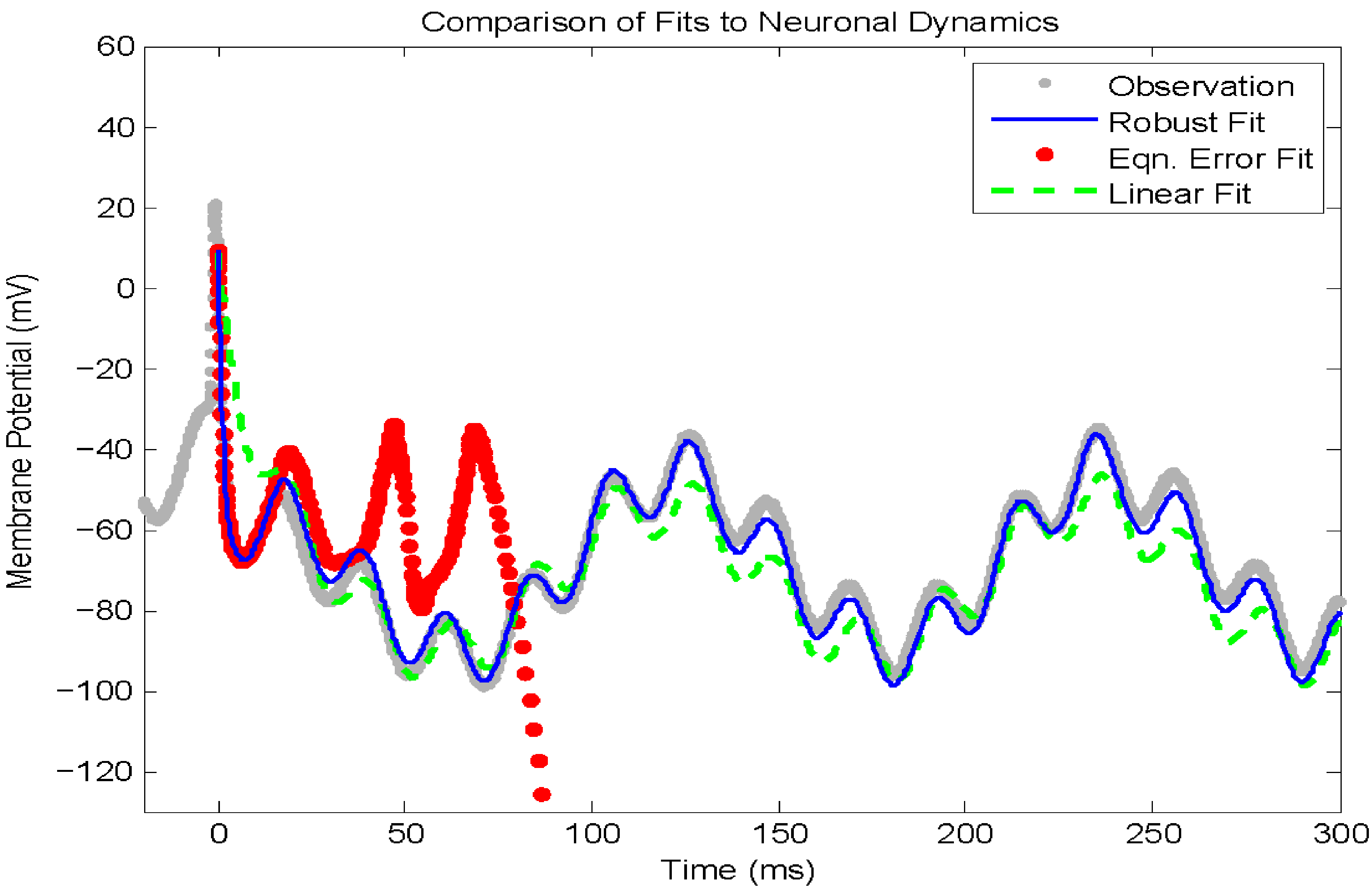
even when

$$h(x, u, y) = y - f(x, u)$$

unless there is no feedback, as in

$$y_t = f(u_t, u_{t-1}, \ldots, u_{t-m})$$

EXAMPLE: OUTPUT ERROR RUN-OFF



OUTPUT ERROR MINIMIZATION

EXISTING OPTIONS:

- ❑ local optimization
- ❑ assume true model is in the class,
  assume number of samples is "large enough"

THIS TALK:

- ❑ robust identification error :
  an instantaneous measure of output error
- ❑ convex upper bound for RIE
- ❑ convex parameterization of robust models
- ❑ minimization of cumulative RIE bounds

DISCRETE TIME STATE SPACE MODELS

MODEL:

$$e(x_{t+1}) = f(x_t, u_t)$$
$$h(x_t, u_t, y_t) = 0$$

WELL-POSEDNESS:

$$\forall \; x, u \; \exists! \; v, y :$$
$$e(v) = f(x, u)$$
$$h(x, u, y) = 0$$

STABILITY:

$$u_t \equiv \bar{u}_t \Rightarrow$$
$$\sum |y_t - \bar{y}_t|^2 < \infty$$

OUTPUT ERROR

DATA:

$$(\tilde{x}_k, \tilde{u}_k, \tilde{x}_k^+, \tilde{y}_k)_{k=1}^{\bar{k}}$$

OUTPUT ERROR:

$$\bar{\mathcal{E}} = \sum |\tilde{y}_t - \bar{y}_t|^2$$

where

$$e(\bar{x}_{t+1}) = h(\bar{x}_t, \tilde{u}_t)$$

$$h(\bar{x}_t, \tilde{u}_t, \bar{y}_t) = 0$$

$$\bar{x}_0 = \tilde{x}_0$$

LINEARIZED OUTPUT ERROR

$$\bar{\mathcal{E}}^o = \sum |\delta_t|^2$$

where

$$E(x^+)\Delta^+ = F(\tilde{x}, \tilde{u})\Delta + \epsilon_x, \quad \Delta_0 = 0$$

$$H_x(\tilde{x}, \tilde{u}, \tilde{y})\Delta + H_y(\tilde{x}, \tilde{u}, \tilde{y})\delta + \epsilon_y = 0$$

$$\epsilon_x = f(\tilde{x}, \tilde{u}) - e(\tilde{x}^+)$$

$$\epsilon_y = h(\tilde{x}, \tilde{u}, \tilde{y})$$

ROBUST LINEARIZED OUTPUT ERROR

$$\mathcal{E}^o_Q(x, u, v, y)$$

is the minimal upper bound of

$$|F\Delta + \epsilon_x|^2_Q - |E\Delta|^2_Q + |\delta|^2$$

subject to

$$H_x\Delta + H_y\delta + \epsilon_y = 0$$

LEMMA 1: for Q=Q'>0

$$\bar{\mathcal{E}}^o \leq \sum \mathcal{E}_Q^o(\tilde{x}_k, \tilde{u}_k, \tilde{x}_k^+, \tilde{y}_k)$$

LEMMA 2: models satisfying

$$\mathcal{E}_Q^o(x, u, v, y) < \infty \quad \forall \; x, u, v, y$$

for some Q=Q'>0 are well-posed and stable

CONVEX UPPER BOUND FOR $\mathcal{E}_Q^o$:

since the conditions

$$-|E\Delta|^2_{P^{-1}} \leq |M\Delta|^2_P - 2\Delta' M' E\Delta$$

$$2\delta'(H_x\Delta + H_y\delta + \epsilon_y) = 0$$

are always satisfied, the convex upper bound $\widehat{\mathcal{E}}_Q^o(x, u, v, y)$

can be defined as the minimal upper bound of

$$|F\Delta + \epsilon_x|^2_Q + |M\Delta|^2_P - 2\Delta' M' E\Delta$$
$$+ |\delta|^2 - 2\delta'(H_x\Delta + H_y\delta + \epsilon_y)$$

ANALYSIS: THE LINEAR CASE

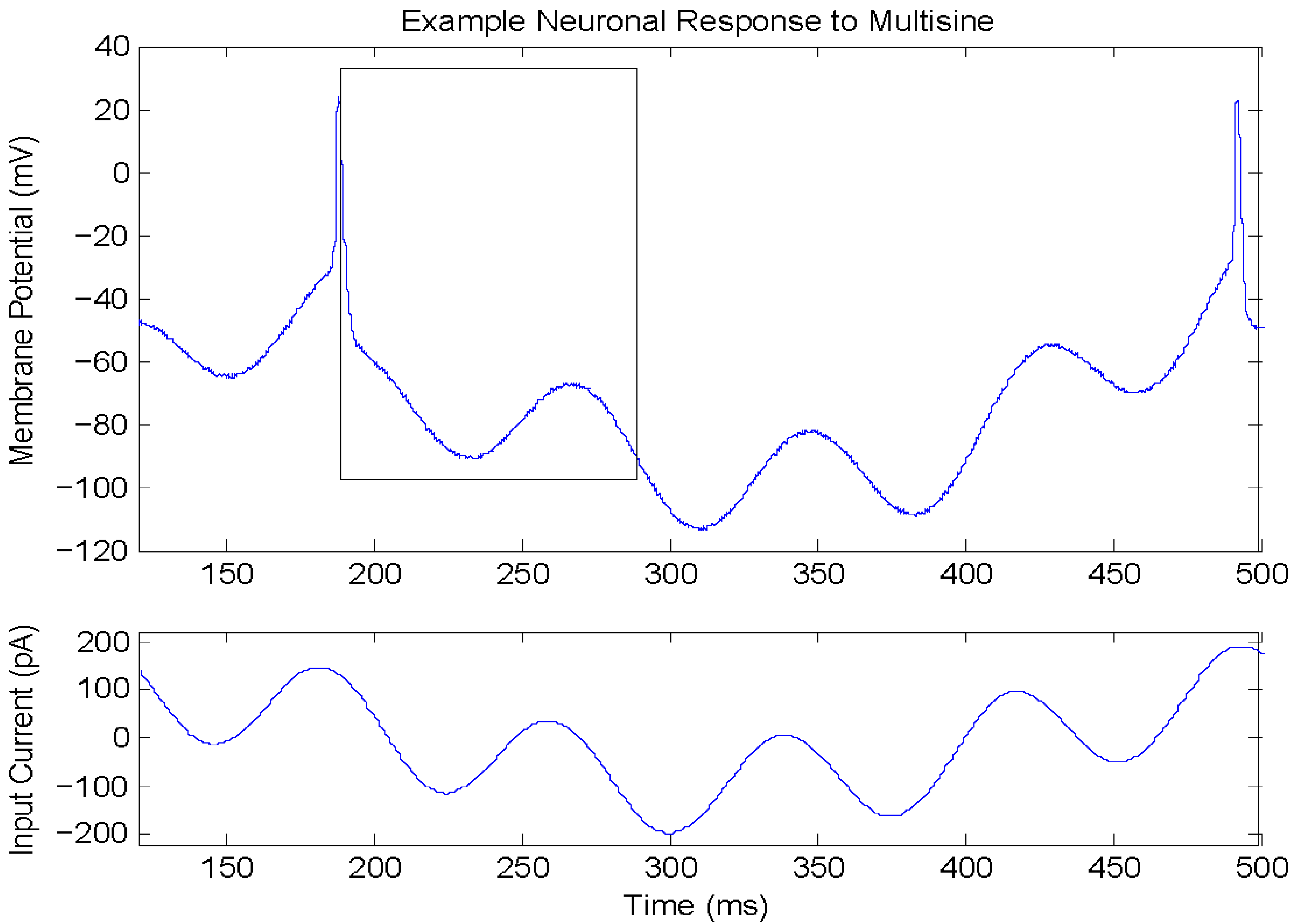
Model:

$$e(v) = Ev,$$
$$f(x,u) = Fx + Lu$$
$$h(x,u,y) = y - Cx - Du$$

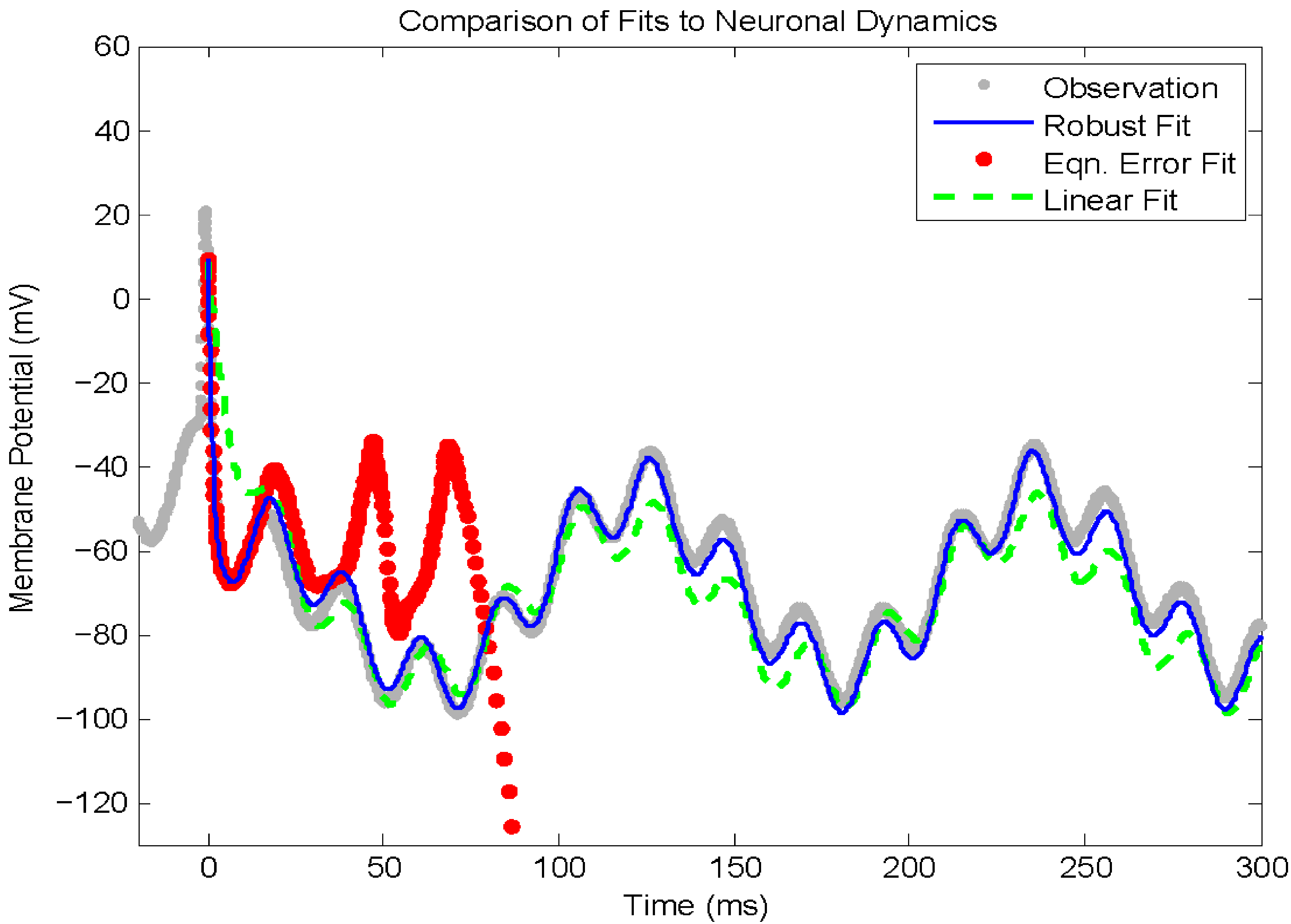
Stability:

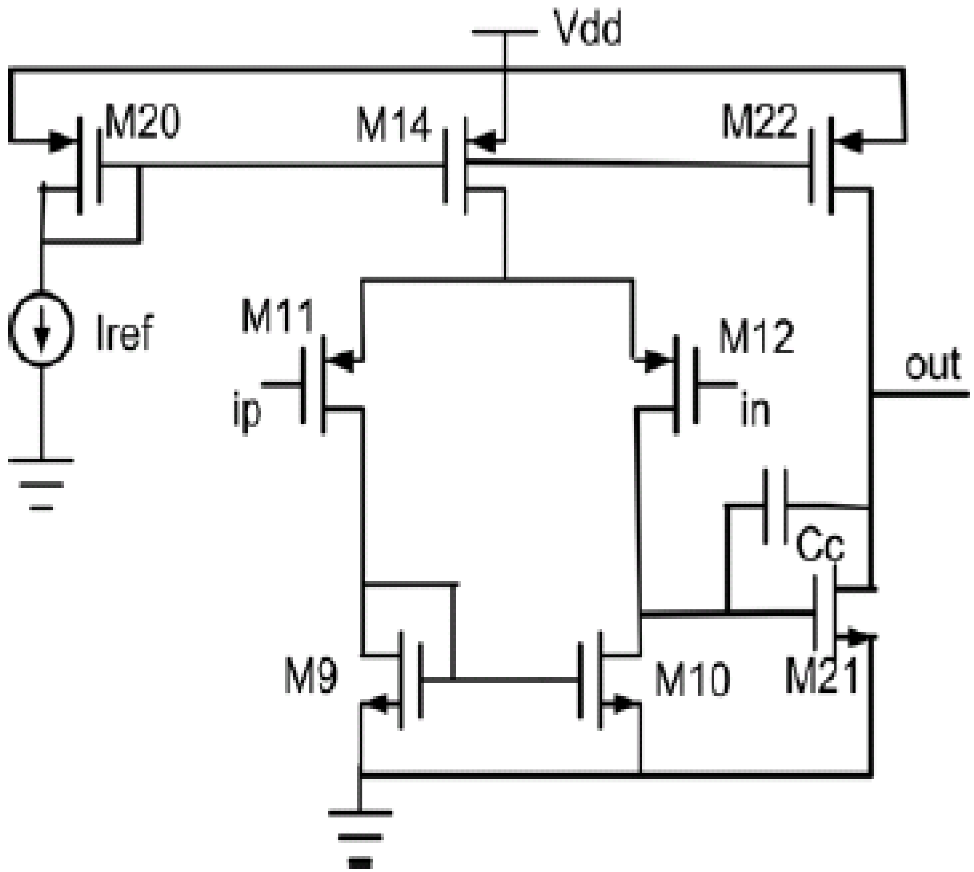
$$F'P^{-1}F + C'C + M'PM < M'E + E'M$$

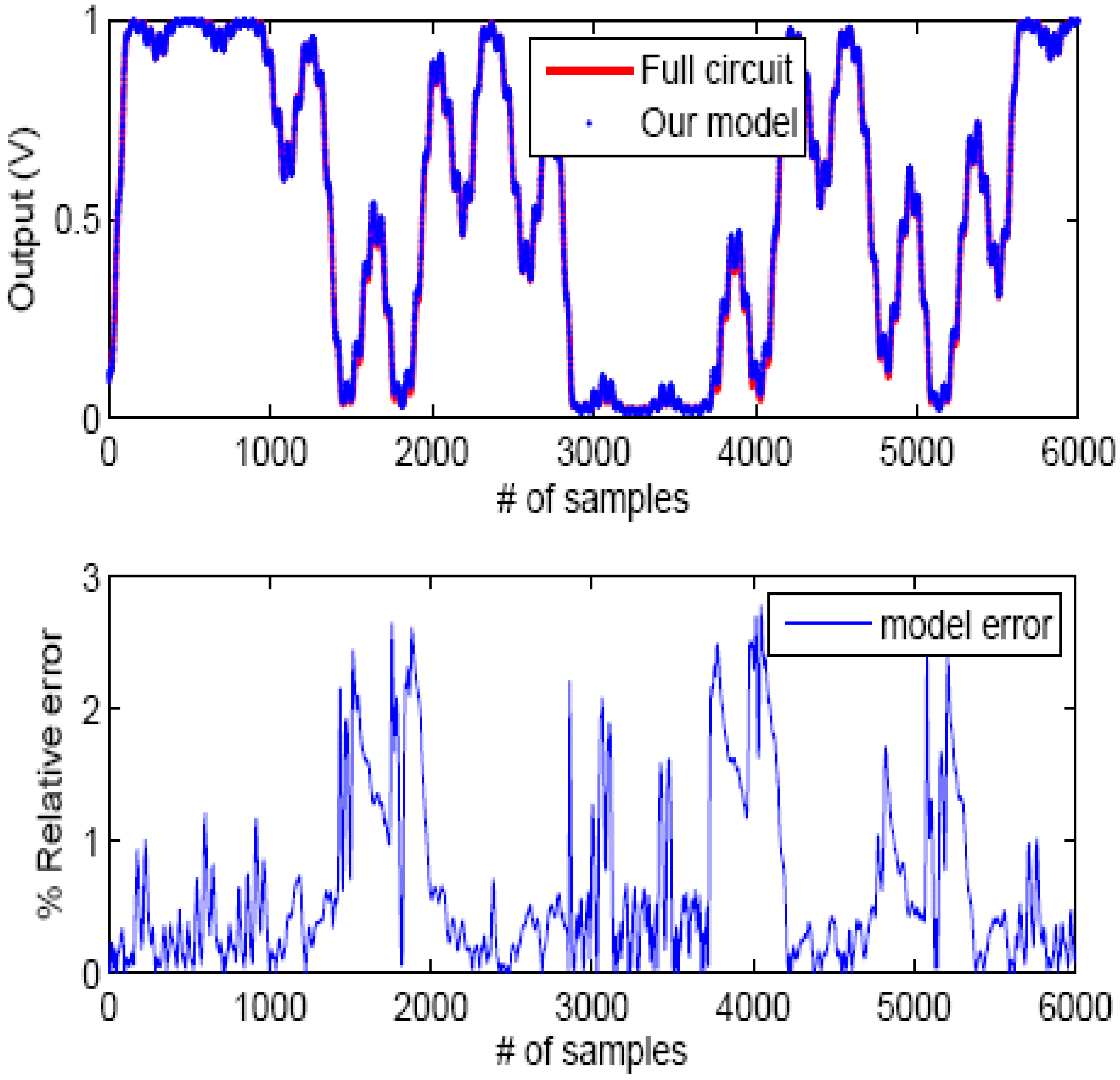
LEMMA: for every C, A (Schur), and M (invertible) there exist P=P'>0,F,E satisfying the stability condition



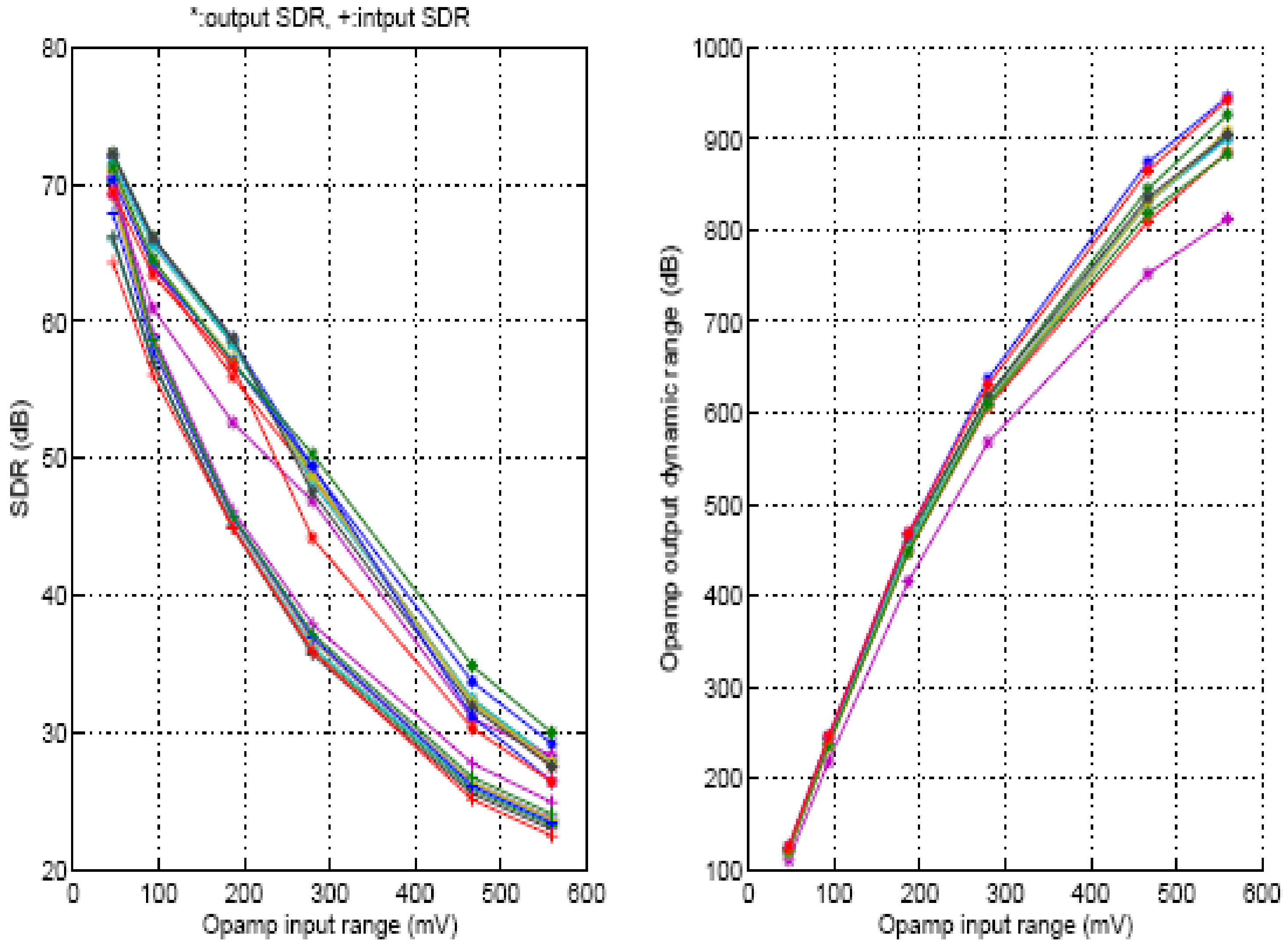
Example Neuronal Response to Multisine
Membrane Potential (mV)
40
20
0
−20
−40
−60
−80
−100
−120
150
200
250
300
350
400
450
500
Input Current (pA)
200
100
0
−100
−200
150
200
250
300
350
400
450
500
Time (ms)


Comparison of Fits to Neuronal Dynamics
Observation
Robust Fit
Eqn. Error Fit
Linear Fit
Membrane Potential (mV)
60
40
20
0
−20
−40
−60
−80
−100
−120
0
50
100
150
200
250
300
Time (ms)

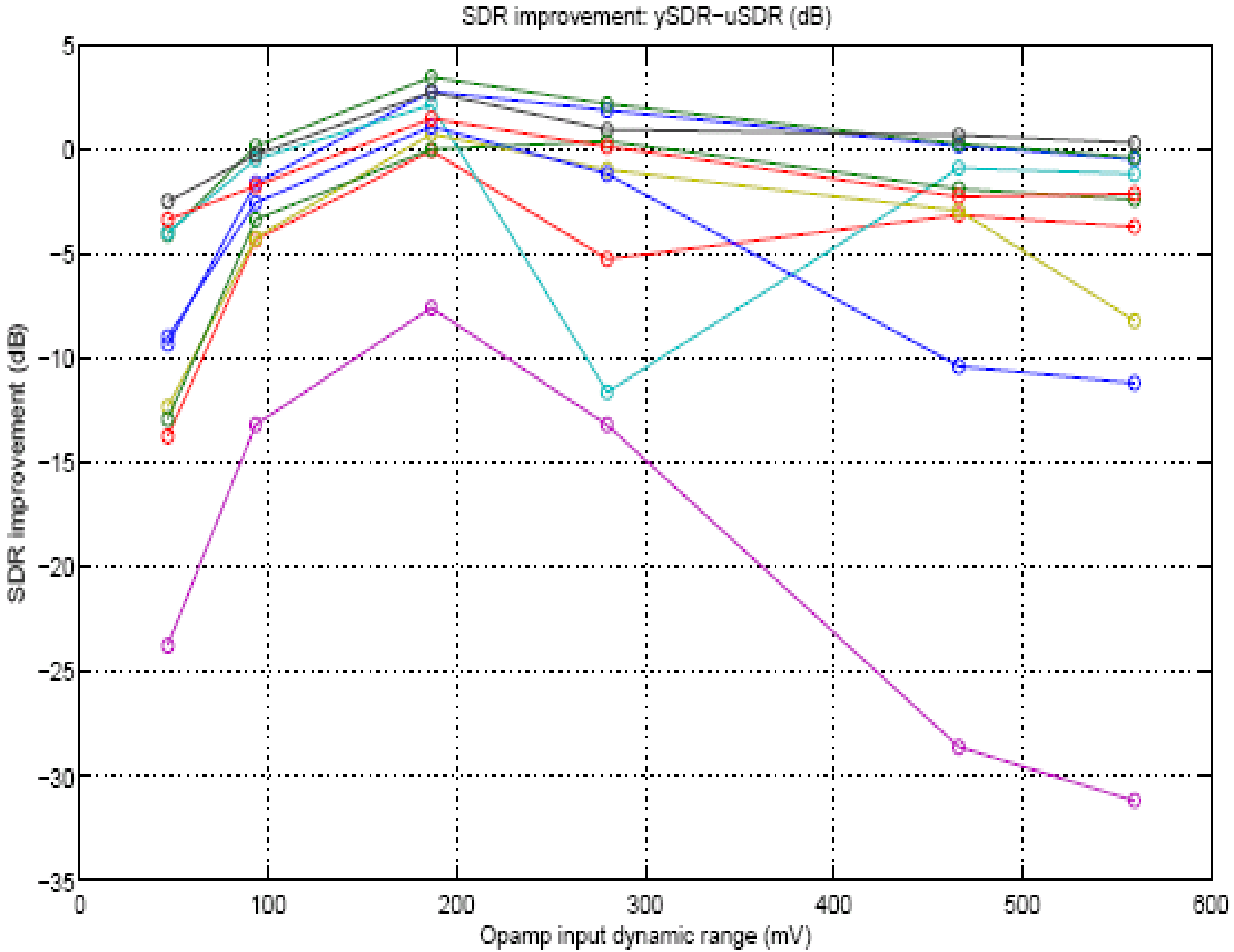

Vdd
M20
M14
M22
M11
M12
Iref
ip
in
out
Cc
M9
M10
M21



Full circuit
Our model
Output (V)
0
0.5
1
0
1000
2000
3000
4000
5000
6000
# of samples
model error
% Relative error
0
1
2
3
0
1000
2000
3000
4000
5000
6000
# of samples



*:output SDR, +:intput SDR
SDR (dB)
Opamp input range (mV)
80
70
60
50
40
30
20
0
100
200
300
400
500
600
Opamp output dynamic range (dB)
1000
900
800
700
600
500
400
300
200
100


SDR improvement: ySDR−uSDR (dB)
5
0
−5
−10
−15
−20
−25
−30
−35
0
100
200
300
400
500
600
SDR improvement (dB)
Opamp input dynamic range (mV)

SUMMARY:

- a framework for handling rationally parameterized models in system id
- convex parameterization of large families of systems with established robustness
- a toolbox for working with algebraic parameterizations and positive polynomials
- excessive conservatism a possible drawback
- alternative parameterizations are developed