Adjoint-method optimization in the presence of eddies

G. Gebbie, Harvard

1. High dimensional optimization due to high-resolution forward model.
Previous ECCO research (Stammer et al, 2002,2003,2004, Fukumori 2002) shows that this is not a fundamental problem.
Still may be some issues with convergence of optimization process.

2. Nonlinearity can lead to an unstable adjoint.
...“adjoint does not tend to a useful sensitivity values; rather it grows exponentially with integration time.” [Lea et al. 2000, on the basis of toy model studies of Gauthier 1992, Miller et al. 1994, Tanguay et al. 1995.]

3. Issues from a combined nonlinear, high-dimensional system.
Many studies (Schroter et al. 1993, Cong et al. 1998) heeded the advice of the toy models and didn't even discuss long time windows.
Not completely clear the link between eddies and nonlinearity of system, direct atmospheric forcing of eddies, episodic instability of the linearized system, etc.

Today, focus on #2.

The initialization problem: Find initial conditions that give a model simulation consistent with observations. Key for NWP.

A toy model example: forced-damped fixed-length single pendulum

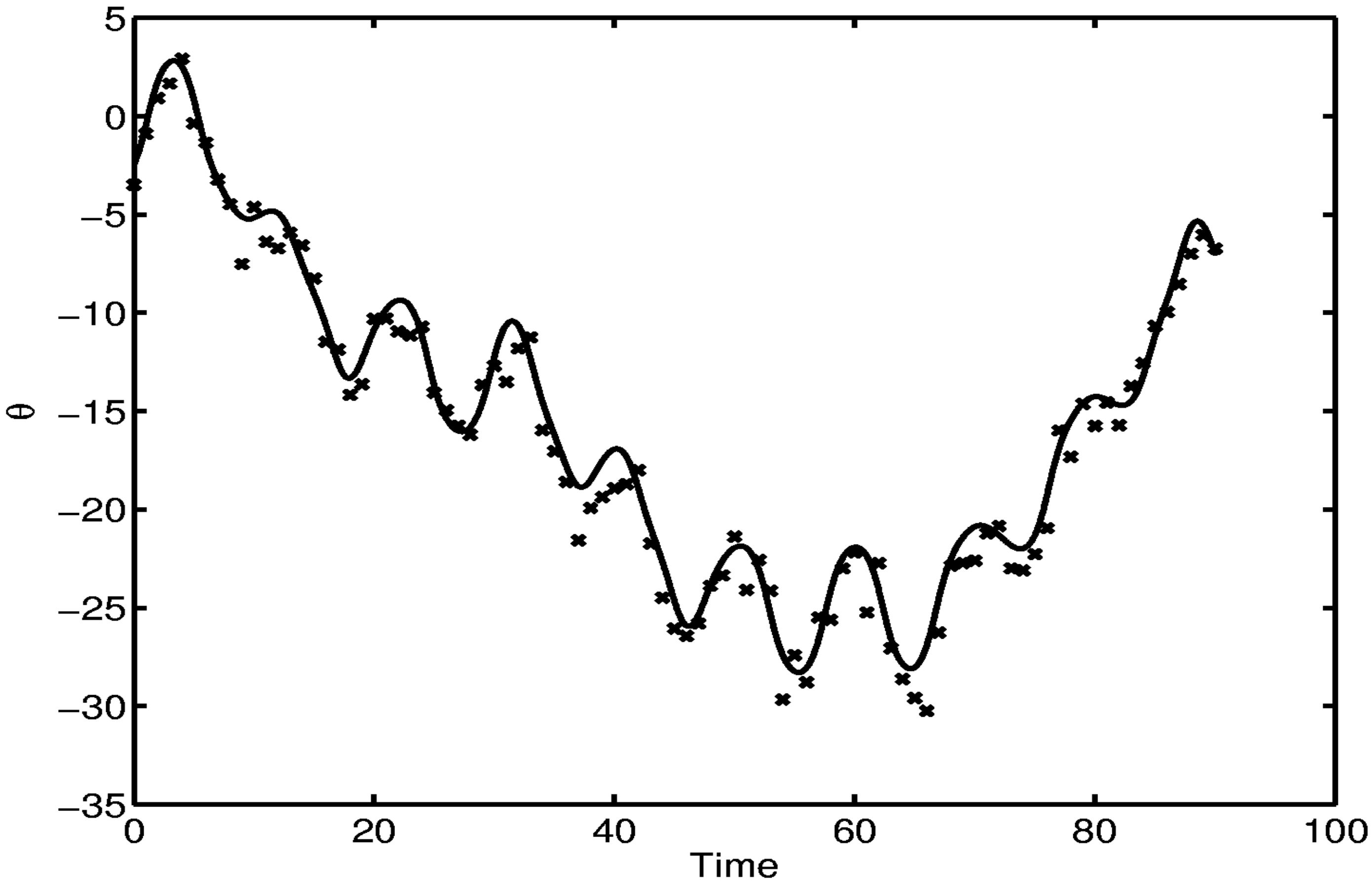
$$\frac{d^2\theta}{dt^2} + q\frac{d\theta}{dt} + \sin\theta = F(t)$$ Where F(t) = g cos (wt)

With parameters q= 2, g=1.5 , w=2/3 , growing sensitivity to initial conditions

Identical twin experiment:
use pendulum equation to get “observations”.

Observational error = 1

Assumes that the investigator has a model which is formulated correctly

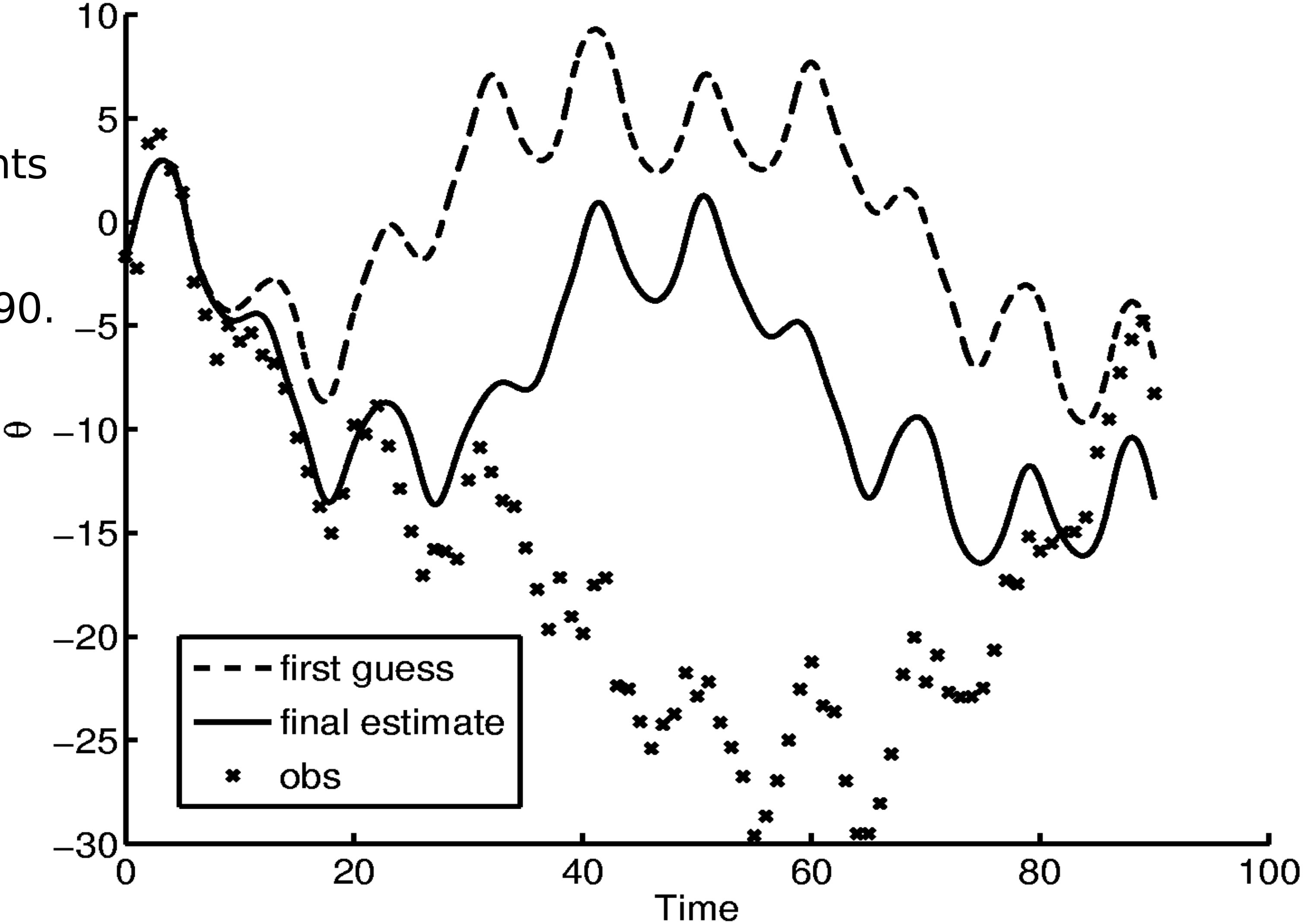


Optimization of the initialization problem

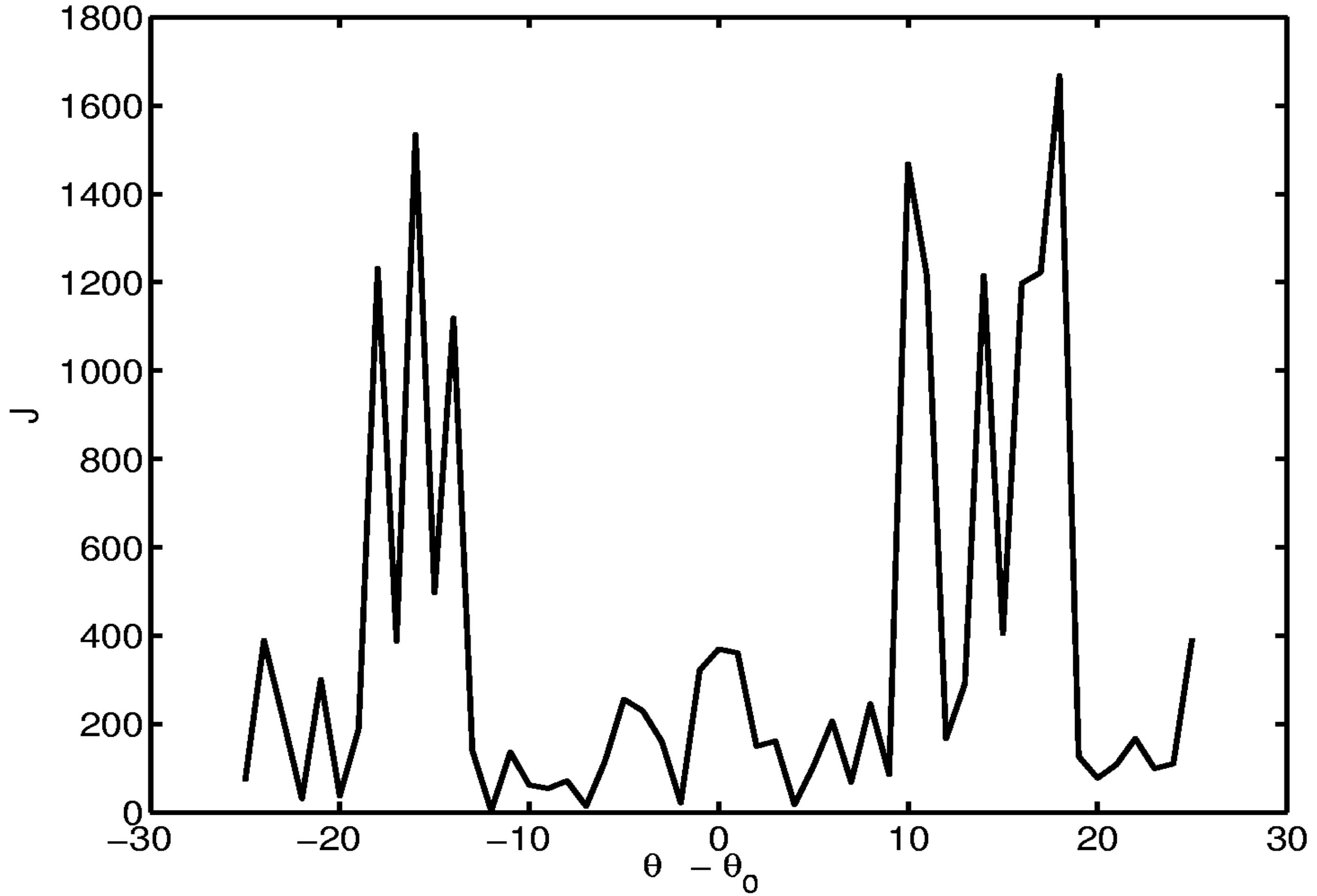
1. Use observation at t=0 to get first-guess initial conditions.
2. Run first-guess forward model and adjoint model (hand-coded in MATLAB)
3. Use gradient descent method (canned MATLAB quasi-Newton BFGS method) to update the initial condition
4. Stop at convergence to solution or stalled optimization procedure

17/20 experiments did not find an acceptable solution over t=90.

For time window of 45 time units, only 2/20 experiments failed.



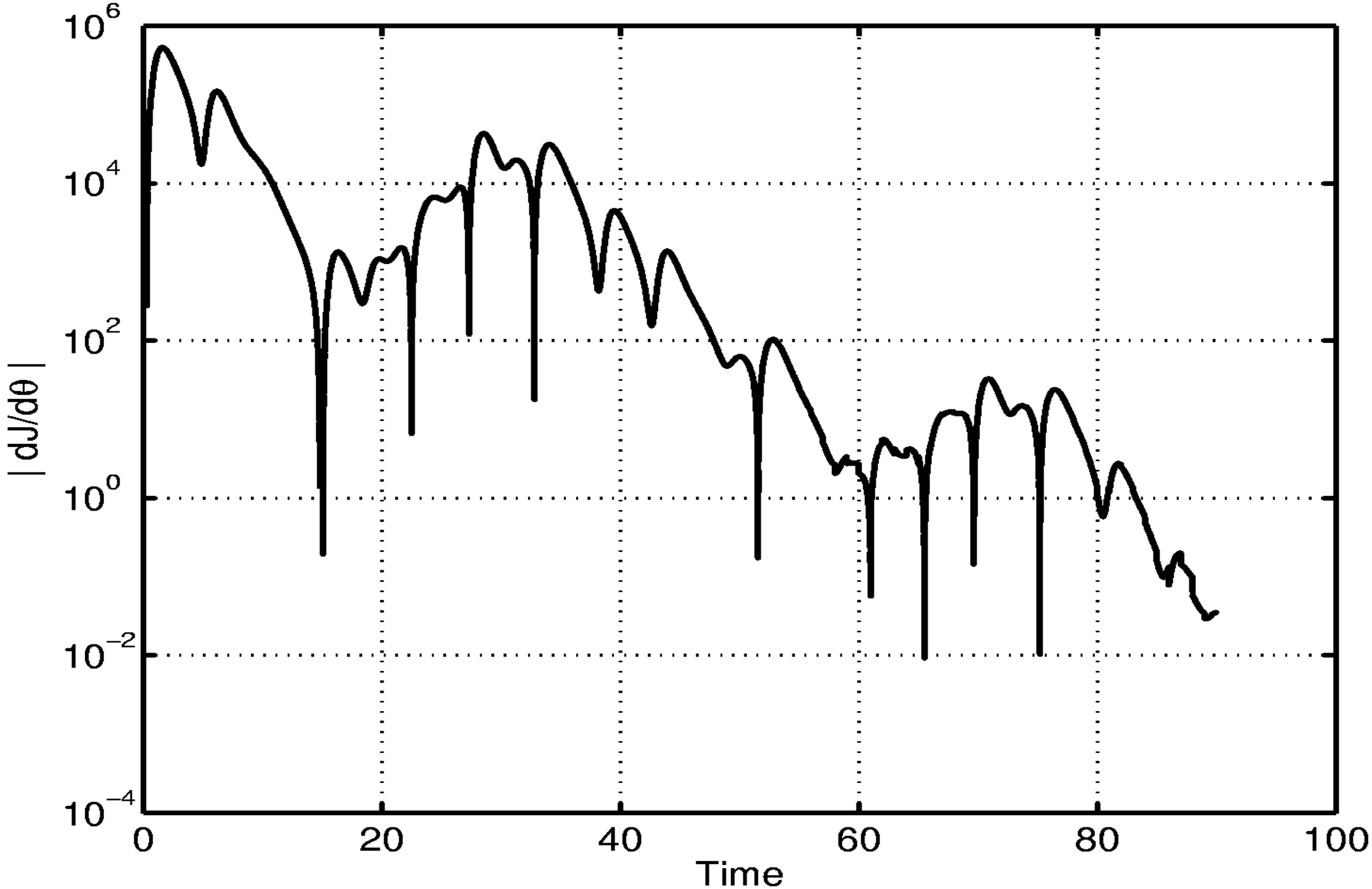
Gradient descent in a complicated topological space



Likelihood to get an acceptable solution depends upon:
a) time window for estimation,
b) observational error,
c) observational data stream.

Even with very good observational database, there is a limited time window over which the initialization problem can be solved if nonlinearity is important.

Growth of adjoint state with time



Nonlinearity can lead to an unstable adjoint.

…"adjoint does not tend to a useful sensitivity values; rather it grows exponentially with integration time." [Lea et al. 2000, on the basis of toy model studies of Gauthier 1992, Miller et al. 1994, Tanguay et al. 1995.] Confirmed in the nonlinear pendulum.

The Ocean State Estimation problem...

may be better described as a time-variable boundary value problem.
Find initial conditions and time-variable boundary conditions that give a model simulation consistent with observations.

$$\frac{d^2\theta}{dt^2} + q\frac{d\theta}{dt} + sin\,\theta = F(t)$$

Where F(t) = g cos(wt) + u(t) and u(t) is unknown.

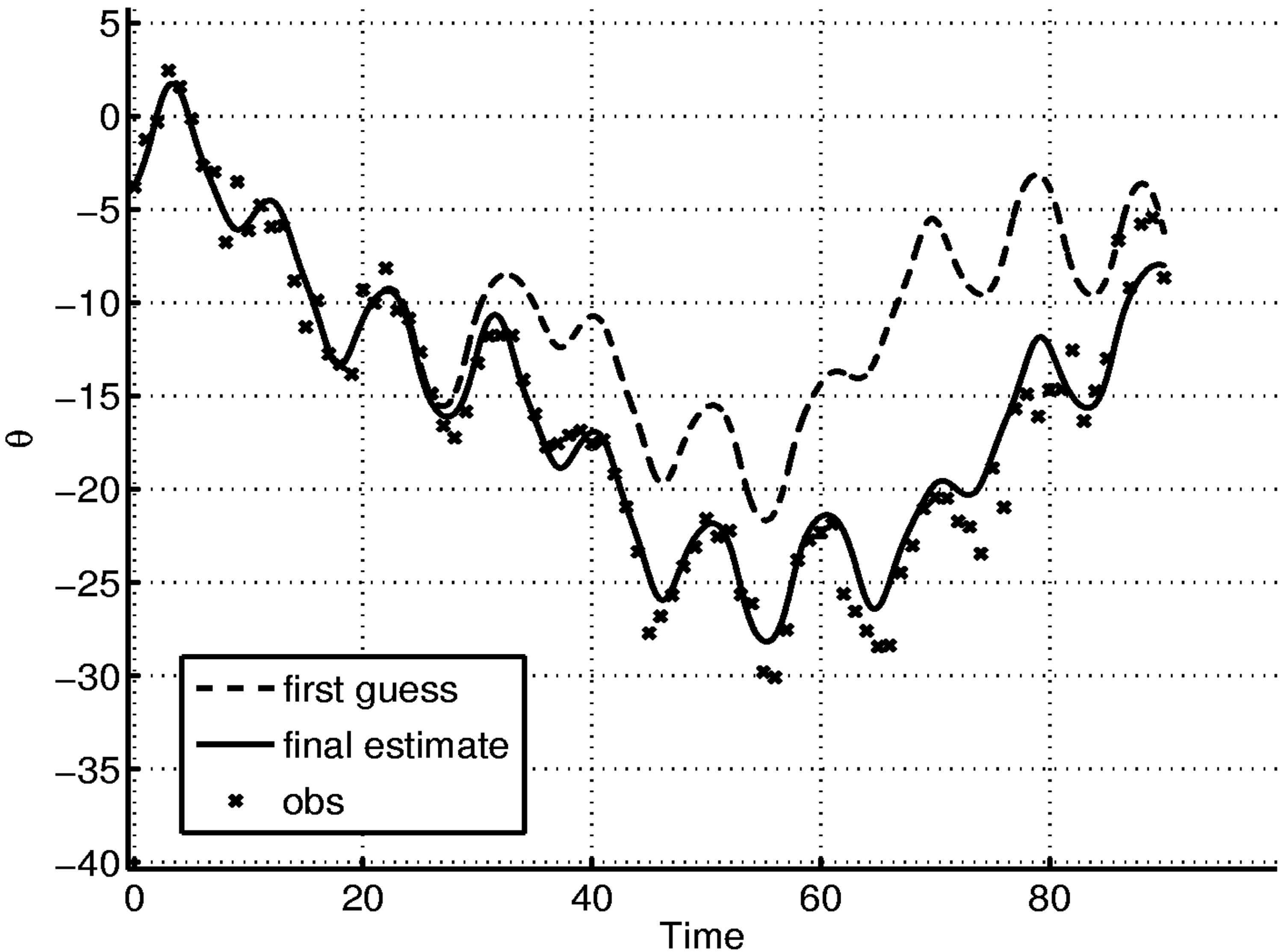
1. Use observation at t=0 to get first-guess initial conditions and assume first-guess forcing is 0.
2. Run first-guess forward model and adjoint model (hand-coded in MATLAB)
3. Use gradient descent method (canned MATLAB quasi-Newton BFGS method) to update the initial condition and the forcing term
4. Stop at convergence to solution or stalled optimization procedure

A better analogy to the uncertain time-variable air-sea fluxes?

The Ocean State Estimation analogy

Over a time window of 90 time units, 16/20 experiments find an acceptable solution to the estimation problem.

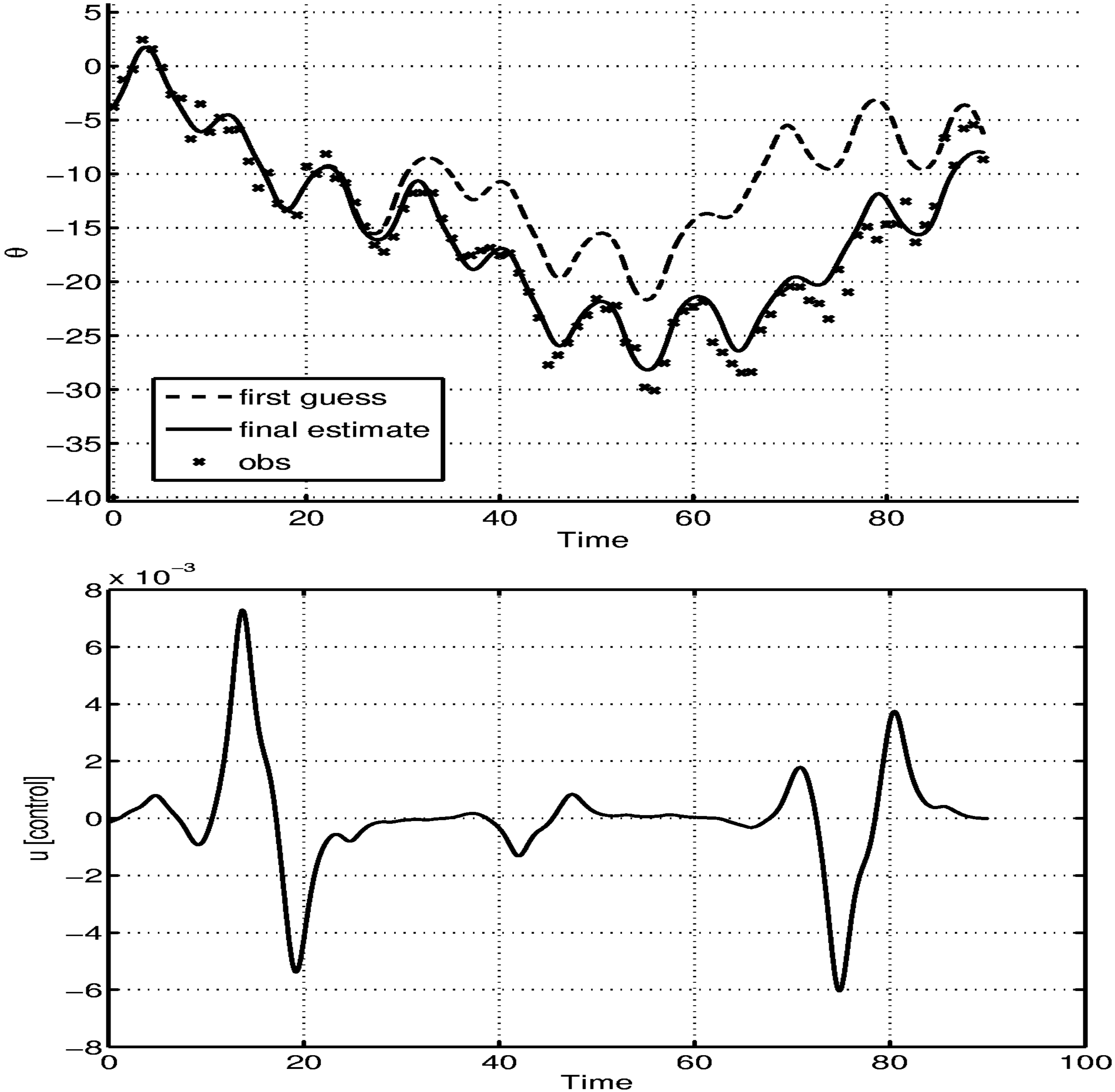
Likelihood of success depends primarily upon the number of controls allowed.



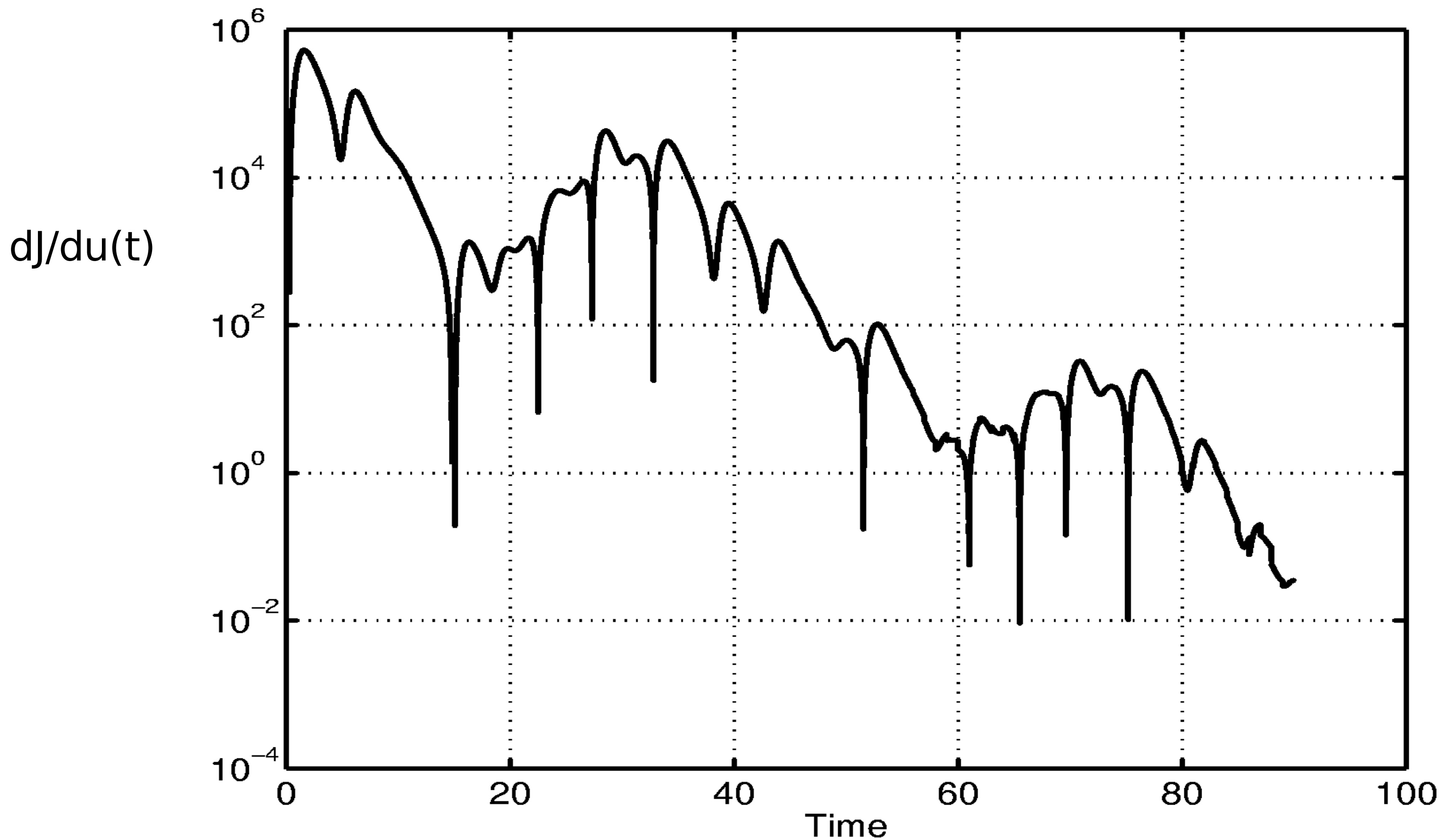
Control of a chaotic system

is sometimes described by engineers as being “easier” than control of other systems (Grebogi and Ott)

because the necessary control adjustments are very small.



Gradient still increases backwards in time



Adjoint is not stabilized but optimization works anyway.

In long time windows, machine overflow is possible.
Because gradient direction is more important than magnitude, could re-normalized and eliminate this problem.

A new criterion for the adjoint method of optimization

Success in the optimization search depends upon the number of controls in a given problem, and is probably very similar to the notion of "controllability."

Controllability (in linear problems) is the ability to move from one arbitrary state to another by control adjustments.

Fukumori et al. (1993) showed that the Kalman filter/smoother was successful when the problem was both controllable and observable.

Evensen (1997) showed that so-called weak-constraint optimization was successful over long time windows. It can be shown that his study is the limiting case where the system is completely controllable (i.e., all estimated quantities are uncertain and are treated as control variables).

Toy models do not rule out the usefulness of adjoint-calculated sensitivities in highly nonlinear models. Instead, controllability, not chaos, appears to be key criterion in success of adjoint method.

In GCM, not clear that ocean state completely controllable by surface boundary conditions - especially deep motions.

Clearly some work to be done to understand the relevance to ECCO...