Offline test criteria for filtering complex dynamical systems

IPAM, Feb 15, 2007

John Harlim

(collaborators: Emilio Castronovo and Andy Majda)

Center for Atmospheric Ocean Science

Courant Institute of Mathematical Sciences, NYU

Motivations

- Curse of dimensionality in making prediction for spatio-temporal chaos.
- Recent development in prediction involves a reduced order filtering strategy, an Ensemble Kalman Filter (for series of EnKF papers, see e.g. Evensen 1994, Houtekamer and Mitchell 1998, Bishop et al. 2001, Anderson 2001, Ott et al. 2001, etc).
- A new problem arises from ensemble approach, that is: how large is the ensemble size needed for stable and accurate filtering ?
- In this presentation, we focus on possibility of using large time steps to obtain stable and accurate filtering. In this sense, we can then afford a large ensemble size.
- Failure of standard (Cohn and Dee 1988) finite difference observability criteria (Grote and Majda, PNAS 2006).

Outline of the talk:

A. Offline Testing Criteria (Majda-Grote, PNAS 2007)

1. analogue of Von-Neumann
2. Information criteria

B. Test of theory

1. Complex scalar Ornstein-Uhlenbeck process
2. Scalar wave equation with weak damping

C. Filtering the “poorman” stochastic model for Lorenz-96 model

Offline Testing Criteria (M-G, PNAS 2007): The general setting

The simplest canonical test problem for modeling a turbulent flow is to be given as follow:

$$\frac{\partial}{\partial t}u = P\left(\frac{\partial}{\partial x}\right)u - \gamma\left(\frac{\partial}{\partial x}\right)u + \sigma(x)\dot{\tilde{W}}(t),\ u(x,0) = u_o$$

where $u = u(x,t) \in \mathbb{R}$, x is 2π-periodic, and a canonical observations that available at discrete time t_m and space x_j.

$$v(x_j, t_m) = Gu(x_j, t_m) + \sigma^o,\ \ \sigma^o \sim \mathcal{N}(0, r^o)$$

In this talk, we focus on uniformly distributed observations that available on every discretized model grid point: $\{x_j = jh : 0 \leq j \leq 2N, (2N+1)h = 2\pi\}$.

Offline Testing Criteria: The Fourier domain analog

In Fourier space, the filtering problem for each wave number k is an independent scalar complex problem:

$$
\begin{aligned}
\mathrm{d}\hat{u}_k &= -\lambda_k \hat{u}_k \mathrm{d}t + \sigma_k \mathrm{d}\dot{W}_k \\
\hat{v}_k &= G\hat{u}_k + \hat{\sigma}_k^o, \text{ with } \hat{\sigma}_k^o \sim \mathcal{N}(0, \frac{r^o}{2N+1})
\end{aligned}
$$

with $\hat{u_k}, \hat{v_k} \in \mathbb{C}$, and $\lambda_k = \gamma_k + \mathrm{i}\omega_k$.

Remarks:

- Each SDE (Ornstein-Uhlenbeck process) has climatological energy spectra: $E_k = \frac{\sigma_k^2}{2\gamma_k}$,
- and correlation time $T_{corr} = \gamma_k^{-1}$.
- In each filtering problem, we are interested for the case observation time $T_{obs} = \Delta t$, where Δt is the discrete time step.

"True" signal: The exact solution of the SDE is given by:

$$\hat{u}_{m+1} = \mathrm{e}^{-\lambda \Delta t} \hat{u}_m + \Theta_{m+1}, \text{ where } \Theta_{m+1} \sim \mathcal{N}(0, r),$$

with system noise variance

$$r = \frac{\sigma^2}{2\gamma}(1 - \mathrm{e}^{-2\lambda \Delta t})$$

"True" filter: We define prior state $\hat{u}_{m+1|m}$ and posterior state $\hat{u}_{m|m}$ to distinguish the filtered solution and the true signal $\hat{u}_{m+1}$. So the true filter is the solution of the following pair:

$$\begin{aligned} \hat{u}_{m+1|m} &= \mathrm{e}^{-\lambda \Delta t} \hat{u}_{m|m} + \Theta_{m+1} \\ \hat{v}_{m+1} &= G \hat{u}_{m+1} + \hat{\sigma}^o \end{aligned}$$

Approximate filters: Filtered solution with both temporal and spatial discretized numerical approximation of the scalar SDE

- λ will be affected by different choice of spatial discretization (e.g. upwind).
- Temporally, we check forward Euler, backward-Euler, and trapezoidal. In general, we write the approximate filters with:

$$\hat{u}_{h,m+1|m} = F_h \hat{u}_{h,m|m} + \sigma_{h,m+1}, \text{ where } \sigma_{h,m+1} \sim \mathcal{N}(0, r_h),$$

  with F_h denotes the approximate dynamical operator and r_h denotes the system noise variance.
- We are interested in the unstable case ($|F_h| > 1$) with forward Euler and the stable case ($|F_h| < 1$) with backward Euler and trapezoidal.

Reviewing Kalman filter equation Given our scalar filtering problem:

$$\begin{aligned}\hat{u}_{m+1|m} &= F\hat{u}_{m|m} + \sigma_{m+1},\\ \hat{v}_{m+1} &= G\hat{u}_{m+1} + \hat{\sigma}^o,\end{aligned}$$

The basic Kalman filter solution is:

$$\begin{aligned}\text{offline:}\quad & K_m = r_{m|m-1}G(G^2 r_{m|m-1} + \langle\hat{\sigma}^o\rangle)^{-1}\\ & r_{m|m} = (1 - K_m G)r_{m|m-1}\\ & r_{m|m-1} = F^2 r_{m-1|m-1} + \langle\sigma^2_{m+1}\rangle\\ \text{online:}\quad & \hat{u}_{m|m} = (1 - K_m G)\hat{u}_{m|m-1} + K_m\hat{v}_m\end{aligned}$$

Offline Criteria: The stability, observability, and controllability

Given our scalar filtering problem

$$\begin{aligned}\hat{u}_{m+1|m} &= F\hat{u}_{m|m} + \sigma_{m+1}, \\ \hat{v}_{m+1} &= G\hat{u}_{m+1} + \hat{\sigma}^o.\end{aligned}$$

The classical Stability of filtering (see Anderson and Moore 1979, p.77) is satisfied (i.e. there exists a limiting Kalman gain K_∞ and $|F(1 - K_\infty G)| < 1$).

- for a time invariant and stable dynamics $|F| < 1$.
- for a time invariant dynamics that is not necessarily stable but fully observable and controllable.

Definition:

- observable means $|F| \neq 0$ and this is ignorable,
- and controllable means $r_{m+1} = \langle\sigma^2_{m+1}\rangle \neq 0$.

Offline Criteria: Mean model error

By model errors, we mean the errors due to finite difference schemes. The mean model errors

$$y_m = \mathsf{E}(\hat{u}_{h,m|m-1} - \hat{u}_m)$$

is computed iteratively with

$$y_{m+1} = (\mathrm{e}^{-\lambda\Delta t} - F_h)\mathsf{E}[\hat{u}_m] + F_h(1 - k_{h,m}G)y_m$$

where

$$\mathsf{E}[\hat{u}_m] = \mathrm{e}^{-\lambda m\Delta t}\mathsf{E}[\hat{u}_0].$$

Remark: M-G (PNAS, 2007) shows that for $|F_h| > 1$,

$$|y_m| \le |(\mathrm{e}^{-\lambda\Delta t} - F_h)||\mathsf{E}[\hat{u}_0]| \sum_{\ell=0}^{m-1} \mathrm{e}^{-\lambda\ell\Delta t}|F_h|^{-(m-1-\ell)}$$

This inequality explains why a stable filtering is possible even with unstable difference approximation.

Summary on Offline Testing

In the offline testing, we check several quantities:

- The amplitude of the dynamical operator $|F_h|$
- the limiting Kalman gain K_∞,
- the limiting stability amplification factor $|F_h(1 - K_\infty G)|$,
- the mean model error y_m, and
- Controllability (nonzero system noise variance $r_h \neq 0$).

Information Criteria:

- In the previous exposition, we have the true system variance r and each difference approximation provides a straightforward time discretized system noise variance r_h.

- M-G (PNAS, 2007) suggests an alternative way to choose system noise, that is with information criteria. The idea is to minimize the relative entropy

$$S(p, p_h) = \int p \log(p/p_h)$$

where p and p_h are density functions of the true and approximate filters, respectively.

- This suggests us to choose r_h such that $K_\infty = K_{h,\infty}$.
When $|F_h| > 1$ and $K_\infty \leq 1 - |F_h|^{-2}$, we choose $r_h = 0$.

Example 1: Testing Ornstein-Uhlenbeck process:

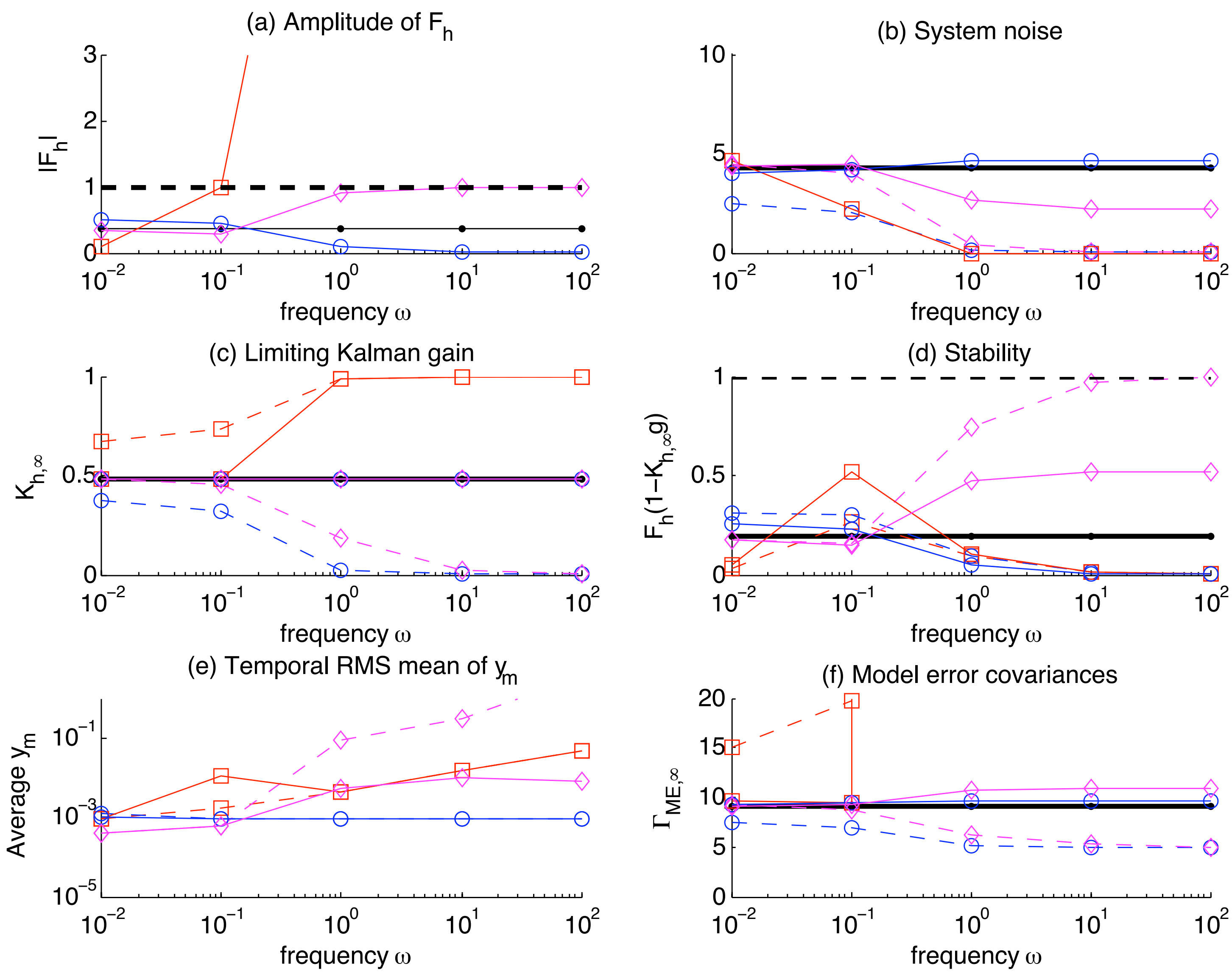
$$\begin{aligned} \mathrm{d}\hat{u} &= -\lambda\hat{u}\mathrm{d}t + \sigma\mathrm{d}\dot{W} \\ \hat{v} &= G\hat{u} + \hat{\sigma}^o, \text{ with } \hat{\sigma}^o \sim \mathcal{N}(0, \frac{r^o}{2N+1}) \end{aligned}$$

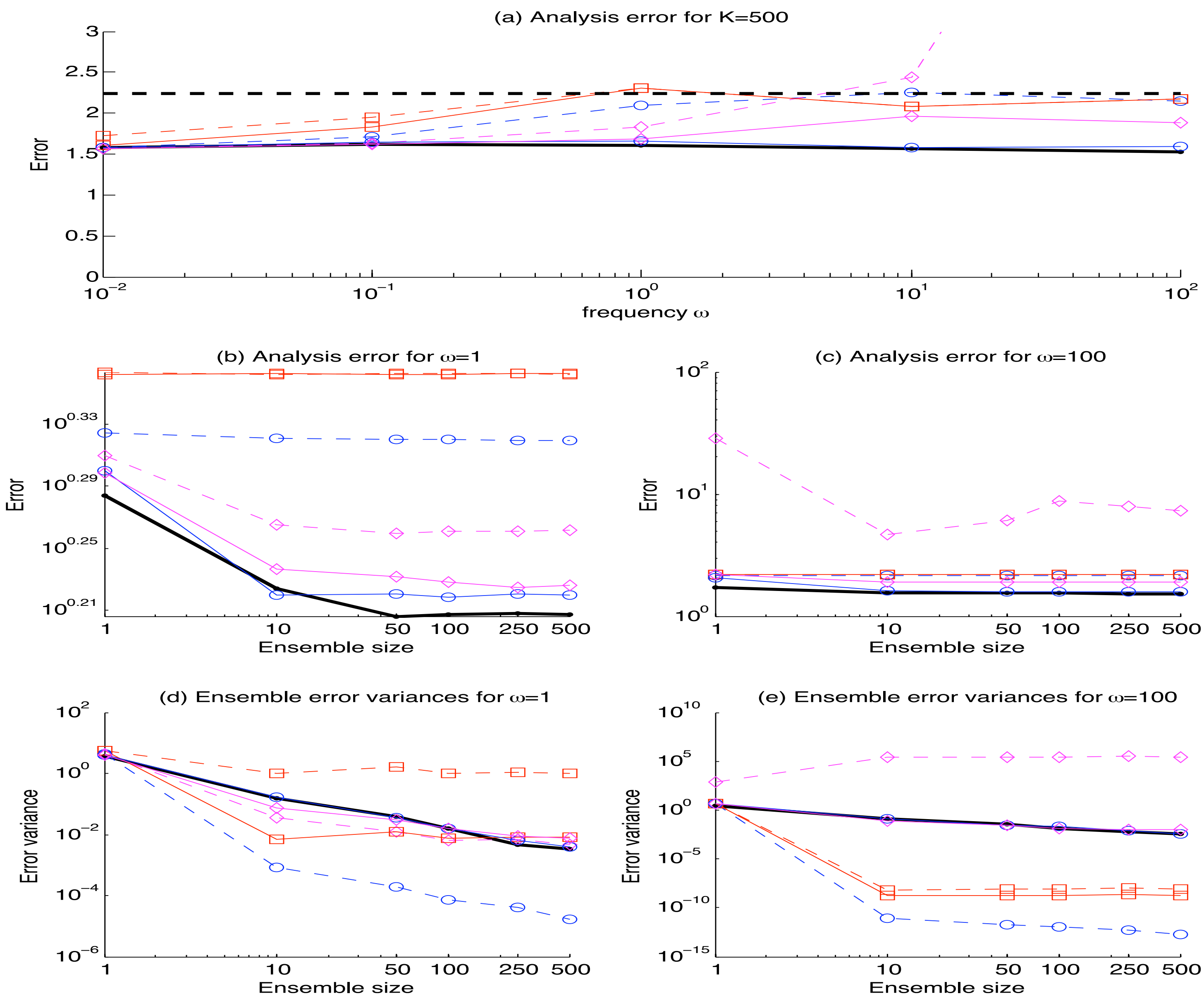
where $\lambda = \gamma + \mathrm{i}\omega$.

- For this problem, we fix $r^o = E = \sigma^2/2\gamma = 5$, observations time $\Delta t = T_{corr} = 1/\gamma$ with $\gamma = 1$, and vary the frequency $10^{-2} \leq \omega \leq 10^2$.
- We also filter the solution using ensemble of size $N = 1, 10, 50, \ldots, 500$. For filter performance, we check the average analysis error (in RMS sense) and the ensemble error variance (temporal average of ensemble spread).
- We choose $G = 1$ and we start the filtering process with random initial states $u^n_{0|0}, n = 1, \ldots, N$.

Testing Ornstein-Uhlenbeck process:

- The unstable forward Euler trust the observations since $K_{h,\infty} = 1$.
- The backward-Euler with no information criteria has $|F_{h,k}| = 0$ and $r_h = 0$ for large frequency ω. This parameter set indicates a bad filtering since the system is zero all the time. Here $|F_h(1 - K_{h,\infty}G)| = 0$ is meaningless.
- In the backward-Euler with information criteria, by restoring back the controllability, we satisfy stability condition (2) and this predicts a good filter.
- The trapezoidal with no information theory. $|F_h| = 1$ and $r_h = 0$ for large frequency. In this case, stability condition (2) is not satisfied. Hence the filter diverges.
- The trapezoidal with information criteria restores the controllability, so even with $|F_h| = 1$, by stability condition (2), the filter will perform well.


(a) Amplitude of F_h
$|F_h|$
frequency ω
(b) System noise
frequency ω
(c) Limiting Kalman gain
$K_{h,\infty}$
frequency ω
(d) Stability
$F_h(1-K_{h,\infty}g)$
frequency ω
(e) Temporal RMS mean of y_m
Average y_m
frequency ω
(f) Model error covariances
$\Gamma_{ME,\infty}$
frequency ω


(a) Analysis error for K=500
Error
frequency ω
(b) Analysis error for ω=1
Error
Ensemble size
(c) Analysis error for ω=100
Error
Ensemble size
(d) Ensemble error variances for ω=1
Error variance
Ensemble size
(e) Ensemble error variances for ω=100
Error variance
Ensemble size

Summary of testing Ornstein-Uhlenbeck process:

- From our results, we conclude that when we use the giant time steps, both the backward-Euler and trapezoidal schemes are the better approximate schemes when the information criteria are used.

- As the information criteria are used, the best scheme among all the stable approximate schemes can be predicted by looking at quantity $|F_h(1 - K_{h,\infty} g)|$, the smaller this quantity implies a more stable scheme.

- We see that, in general, the filter is better when the ensemble size is larger. However, the filter improvement is not significant in the true filter and in all time discretized filters. We found that even with a single realization, we can get a reasonably good filter solution.

- Furthermore, the ensemble error variances decrease as functions of ensemble size. When the filter solely trusts either the observations or the dynamics, the ensemble error variance is small.

Example 2: Testing the scalar wave equation with weak damping

$$\frac{\partial u(x,t)}{\partial t} = -c\frac{\partial}{\partial x}u(x,t) - du(x,t) + \mathcal{F}(x,t), \quad 0 \leq x \leq 2\pi$$

where $c > 0$ is the wave speed propagation and

$$\mathcal{F}(x,t) = \sum_k \sigma_k \mathrm{e}^{\mathrm{i}kx}\dot{W}(t)$$

is the mean zero spatio-temporal forcing with variance $\sum_k \sigma_k^2$. In Fourier space, the truth model is basically the Ornstein-Uhlenbeck processes for each wave numbers with deterministic term.

$$\lambda_k = \gamma_k + \mathrm{i}\omega_k = d + \mathrm{i}ck$$

Now, the additional complexity is in the model resolutions N.

Role of model resolutions and energy spectra to the truth model: In general, the damping and the energy spectra can be written in the form

$$\gamma_k = \gamma_o |k|^\alpha \text{ and } E_k = E_o |k|^{-\beta}, \text{ where } 0 \leq \alpha, \beta < \infty.$$

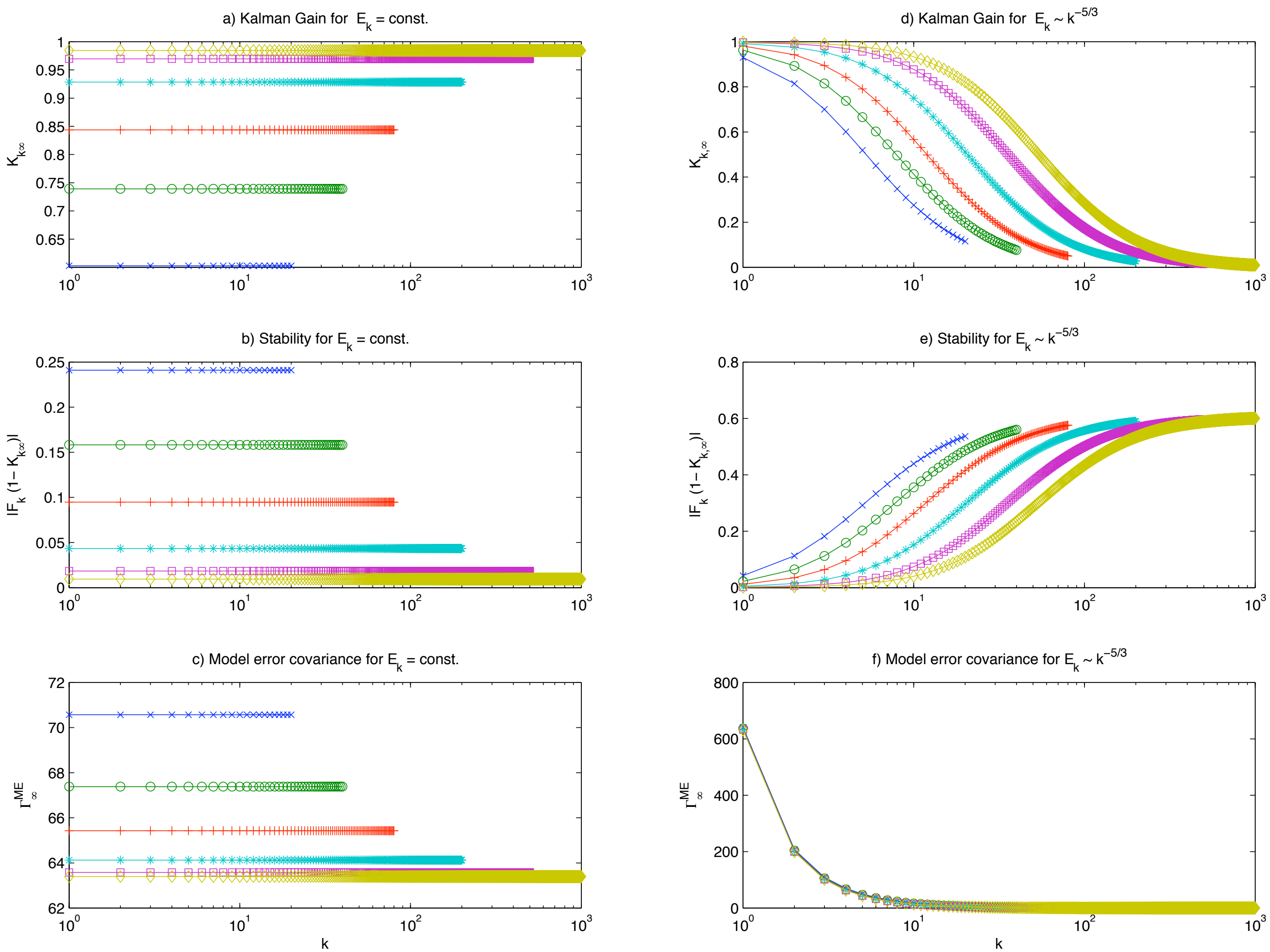
Theorem: There are two regimes of behavior:

1. When $\underline{\beta < 1}$, the asymptotic Kalman gain tends to one uniformly and the filter should primarily <u>trust the observations</u> for $N/2 \leq |k| \leq N$ as $N \to \infty$.

2. When $\underline{\beta > 1}$, the asymptotic Kalman gain tends to zero uniformly and the filter should primarily <u>trust the dynamics</u> for $N/2 \leq |k| \leq N$ as $N \to \infty$.

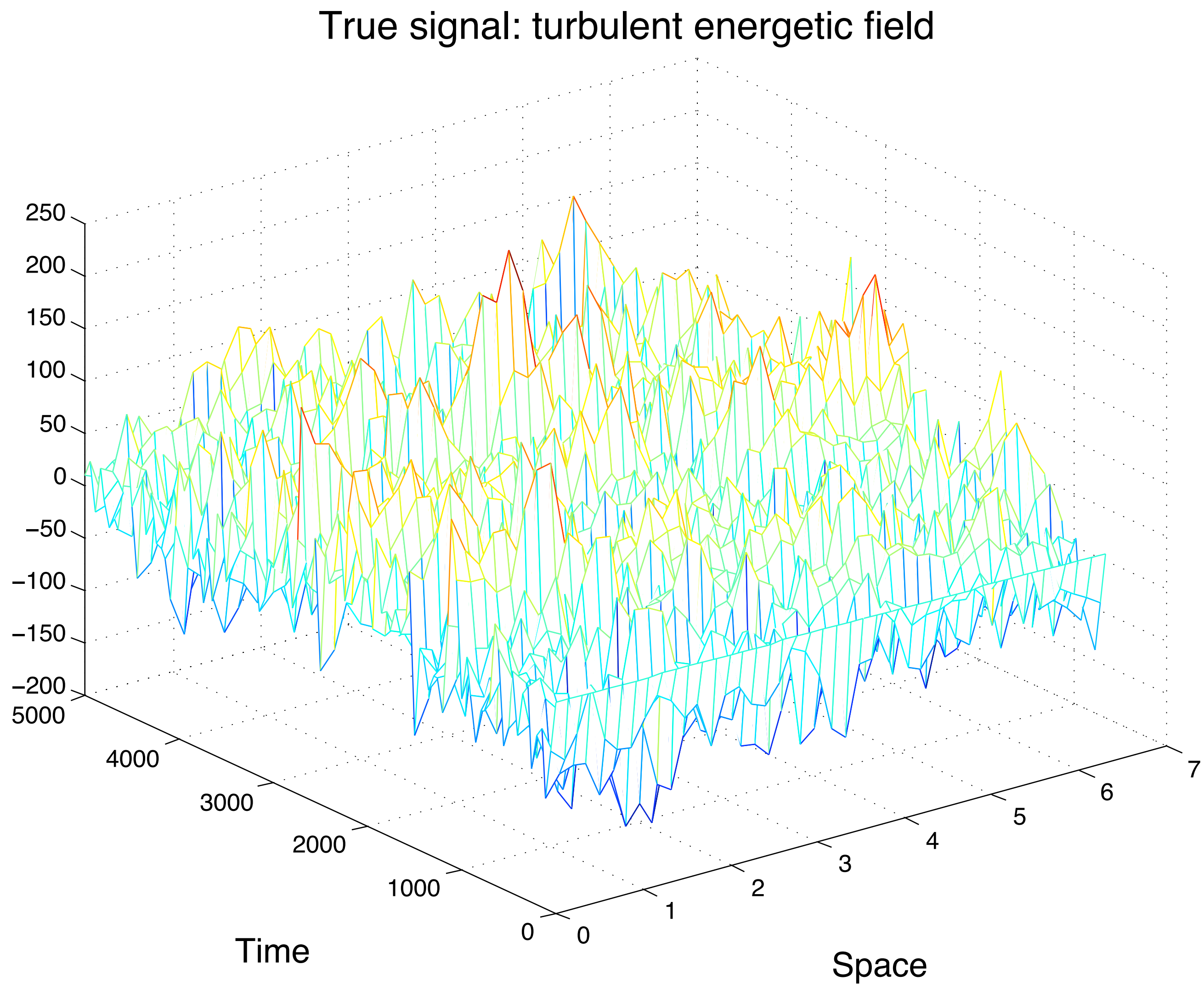
In our example, we have $\alpha = 0$, and two energy spectra: equipartition spectra $\beta = 0$ and fractal energy spectra $\beta = 5/3$.

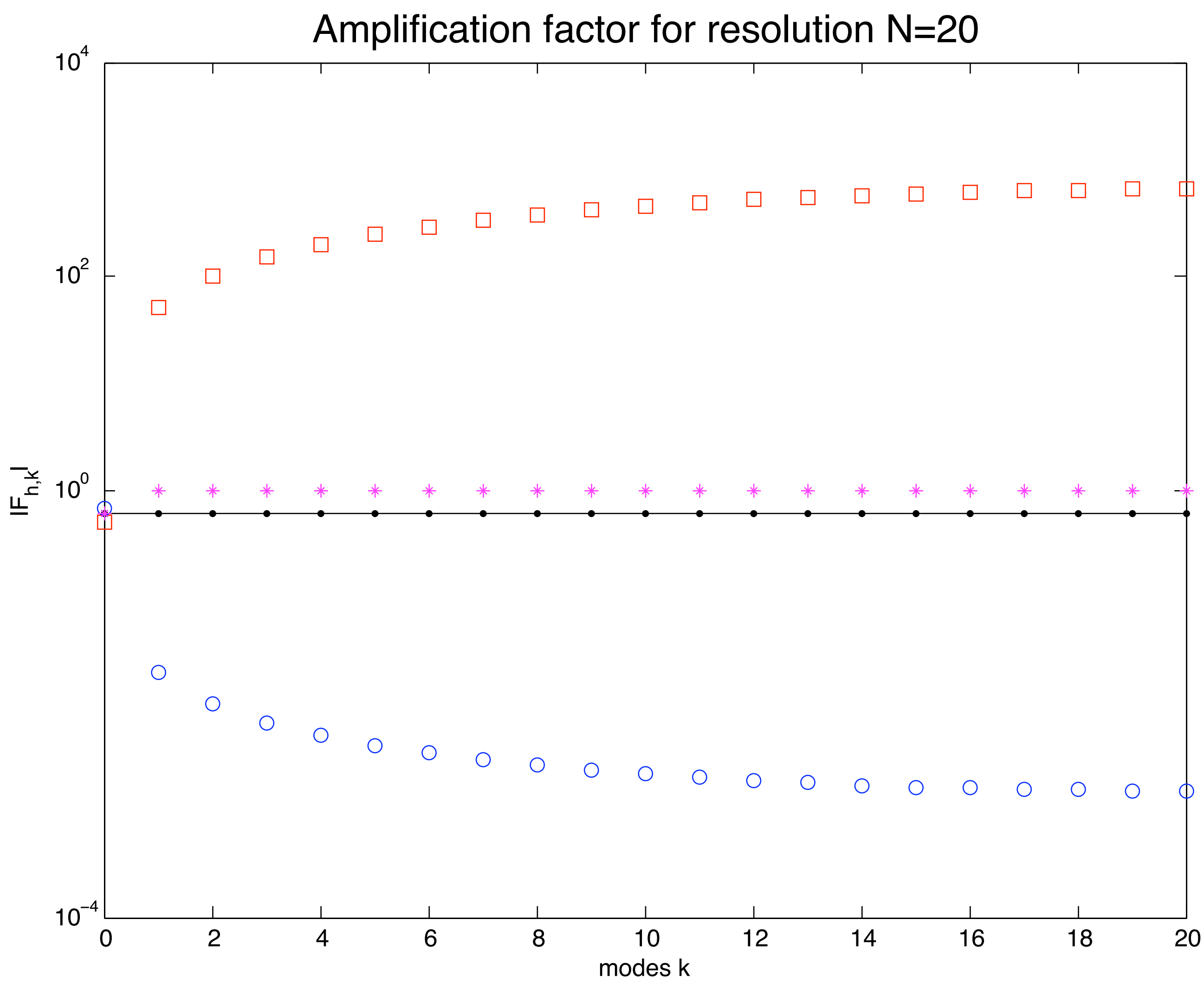
Offline testing true filter



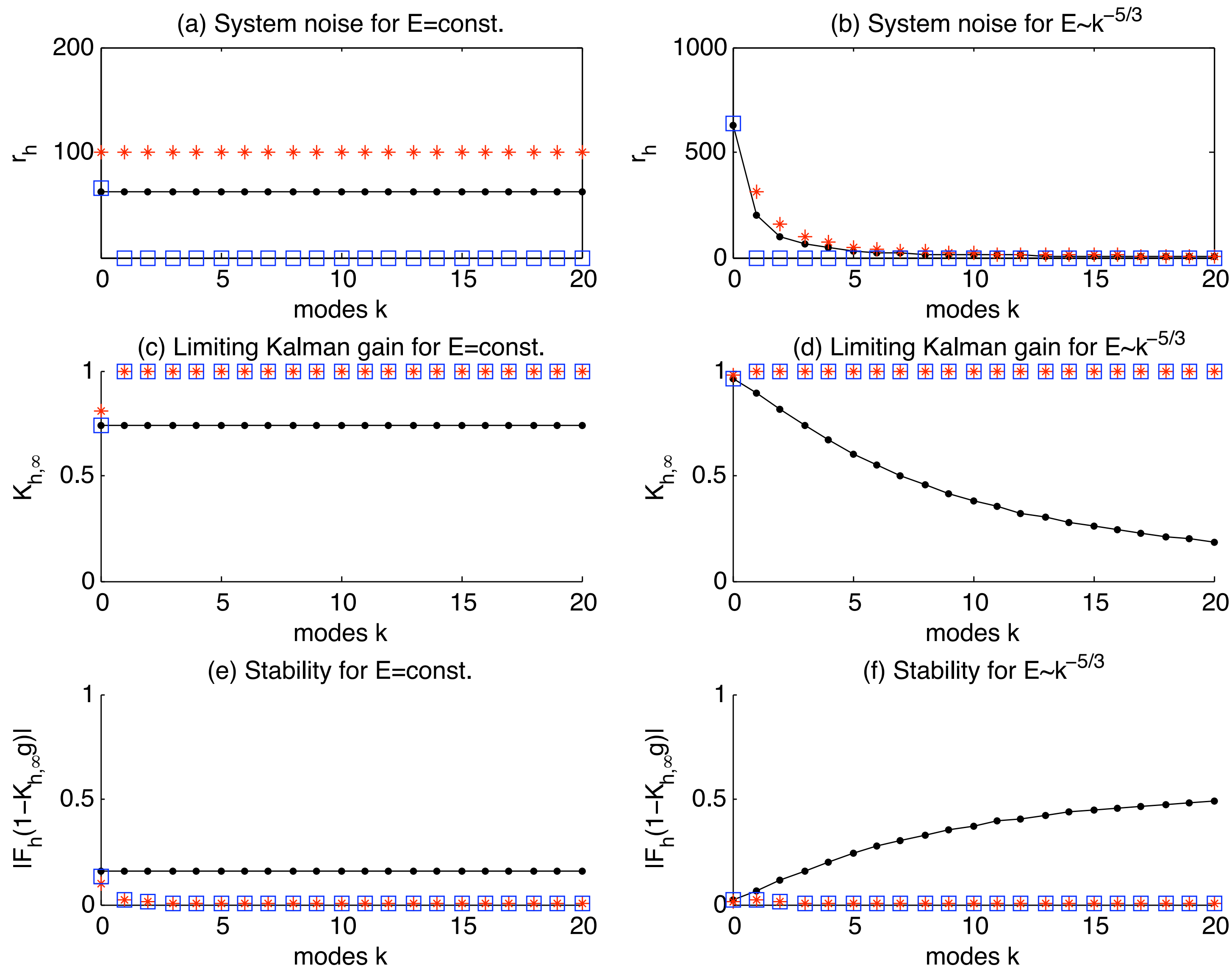
Testing scalar wave eqn. with weak damping:

- For the approximate filter, we use upwind scheme for spatial discretization and forward Euler, backward Euler, and trapezoidal for time differencing.
- We choose energetic spectra $E_k = 100$ and $E_k = 1000k^{-5/3}$. The observations noise variance is chosen to be $r^o = 1000$. Weak damping $\gamma_k = d = 0.01$, non-dimensionalized advective coefficient $c = 1$, and $T_{obs} = \Delta t = T_{corr}/2 = 1/2d = 50$.
- The true state is initiated with $u_0 = \mathrm{e}^{-(x_j-\pi)^2}\cos(x_j - \pi), x_j = jh, 0 \le j \le 2N, (2N+1)h = 2\pi$
- We compare this Fourier domain filter with ensemble transform Kalman filter (Bishop, 2001). Where, we show results with ensemble of size $K = 100$ for resolution $N = 20$ and variance inflation coefficient $r = 10\%$.
- In ETKF, we evolve the model by corresponding difference schemes in real domain and we generate the noise $\mathcal{F}$ in Fourier domain.

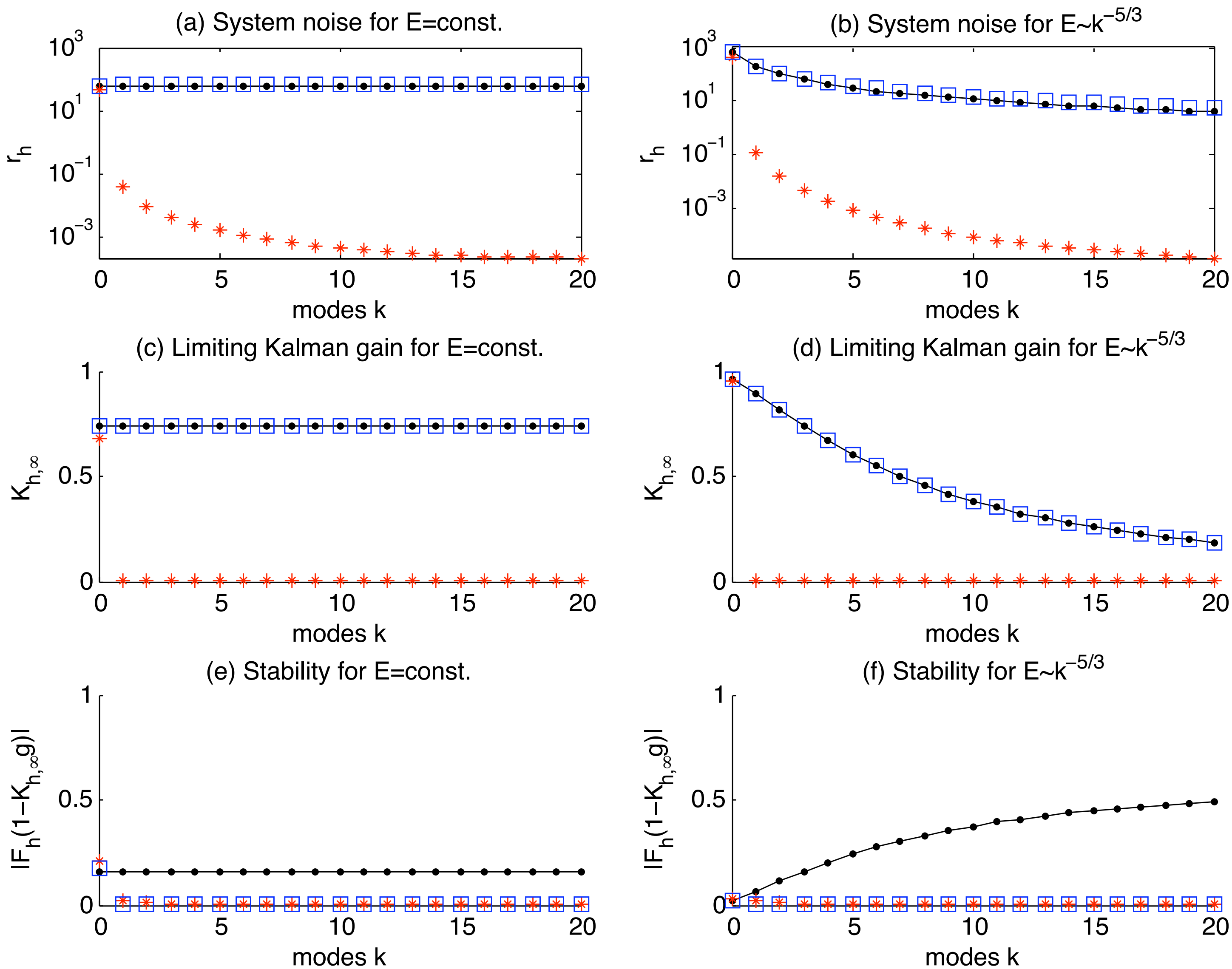

True signal: turbulent energetic field
250
200
150
100
50
0
−50
−100
−150
−200
5000
4000
3000
2000
1000
0
Time
0
1
2
3
4
5
6
7
Space


Amplification factor for resolution N=20
10^4
10^2
10^0
10^{-4}
$|F_{h,k}|$
0
2
4
6
8
10
12
14
16
18
20
modes k

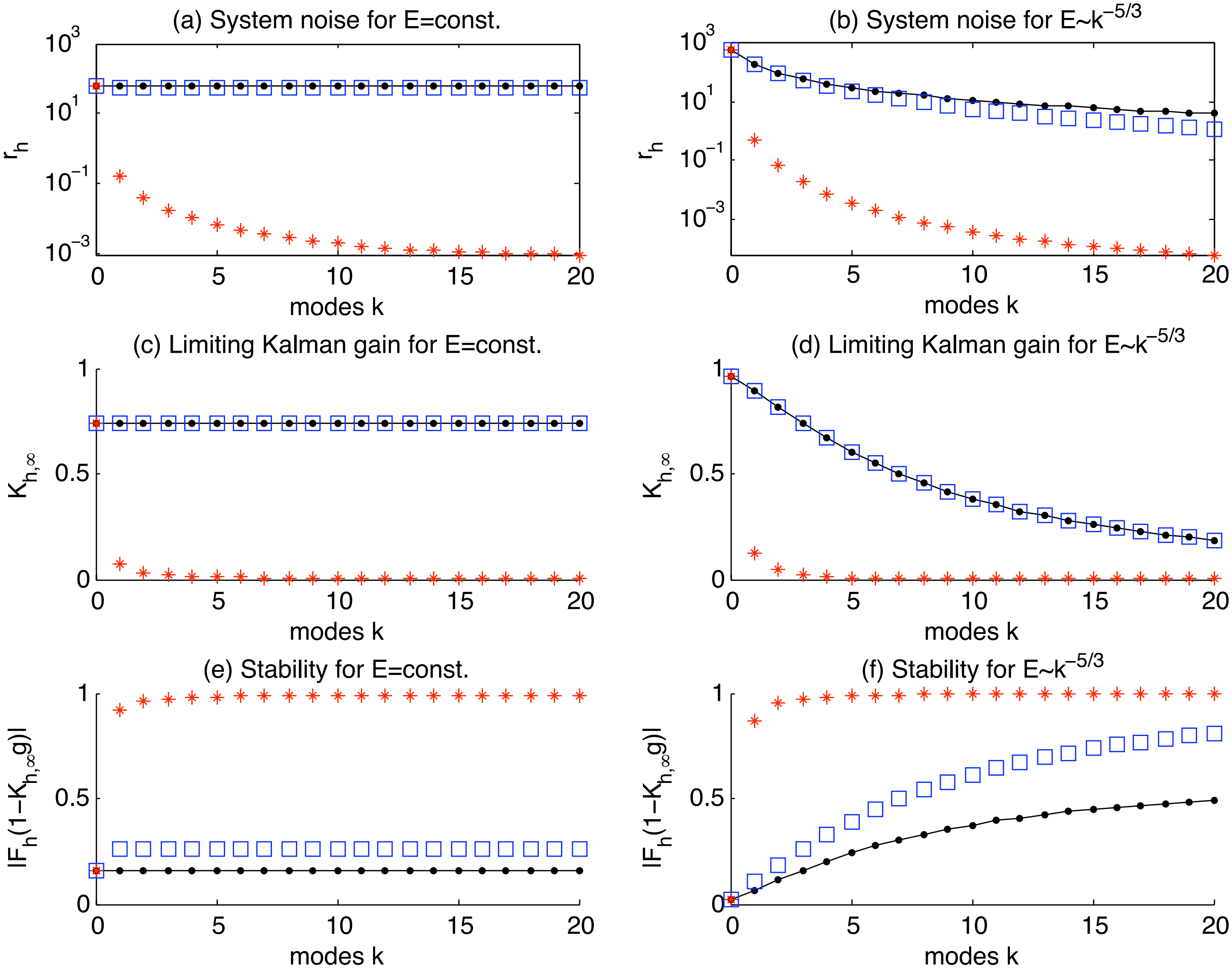
Unstable forward Euler scheme



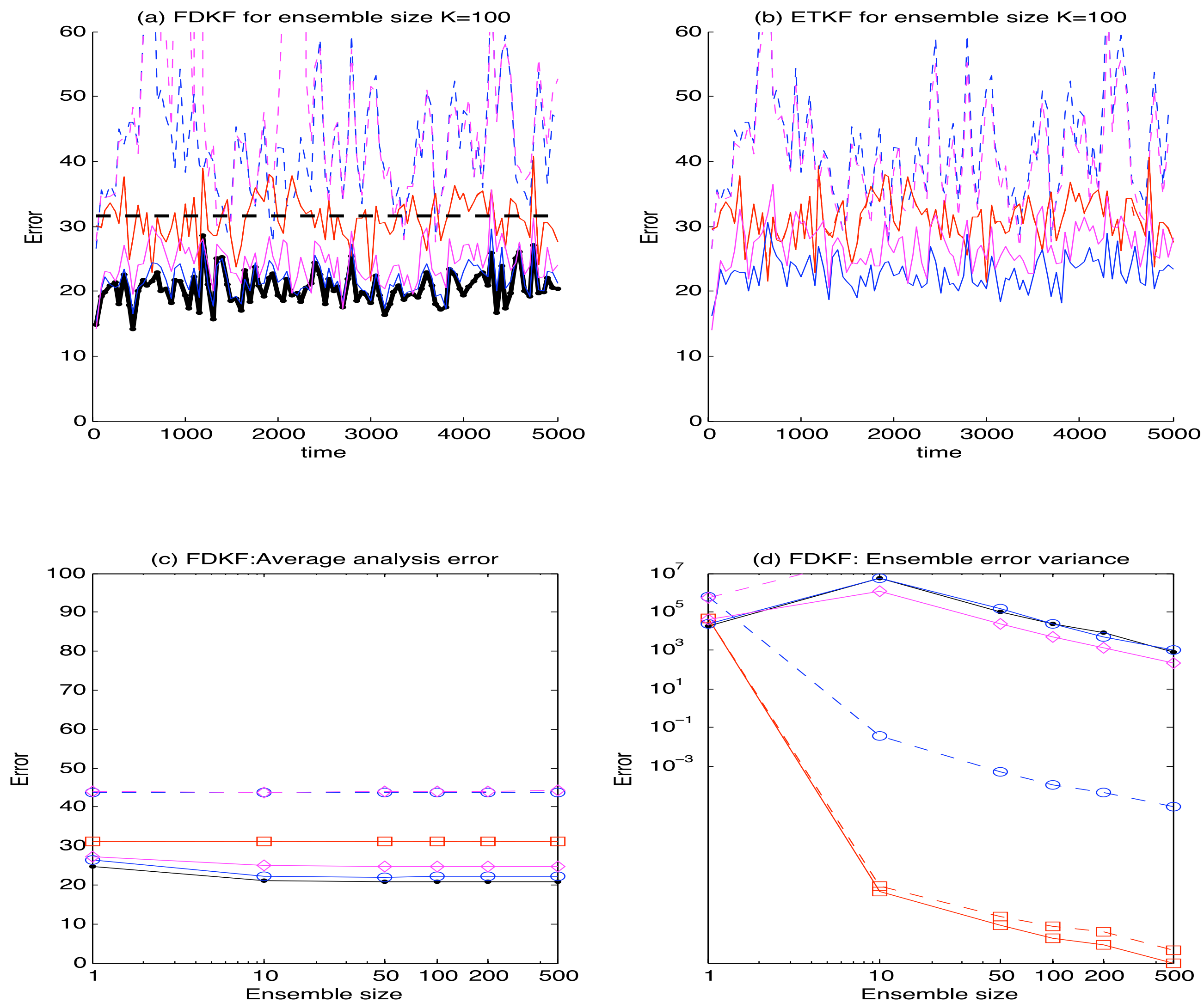
Stable backward Euler scheme



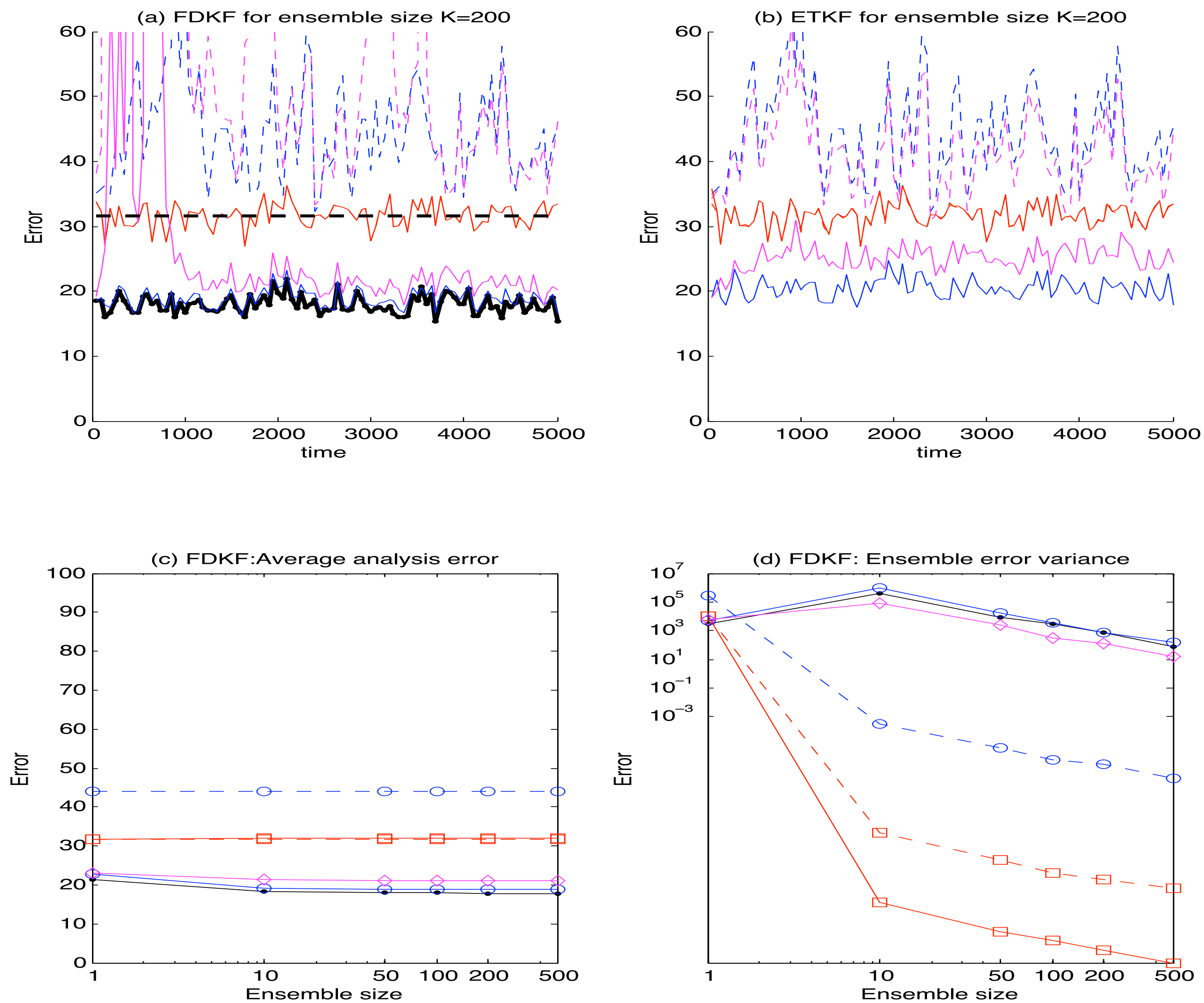
Stable trapezoidal scheme



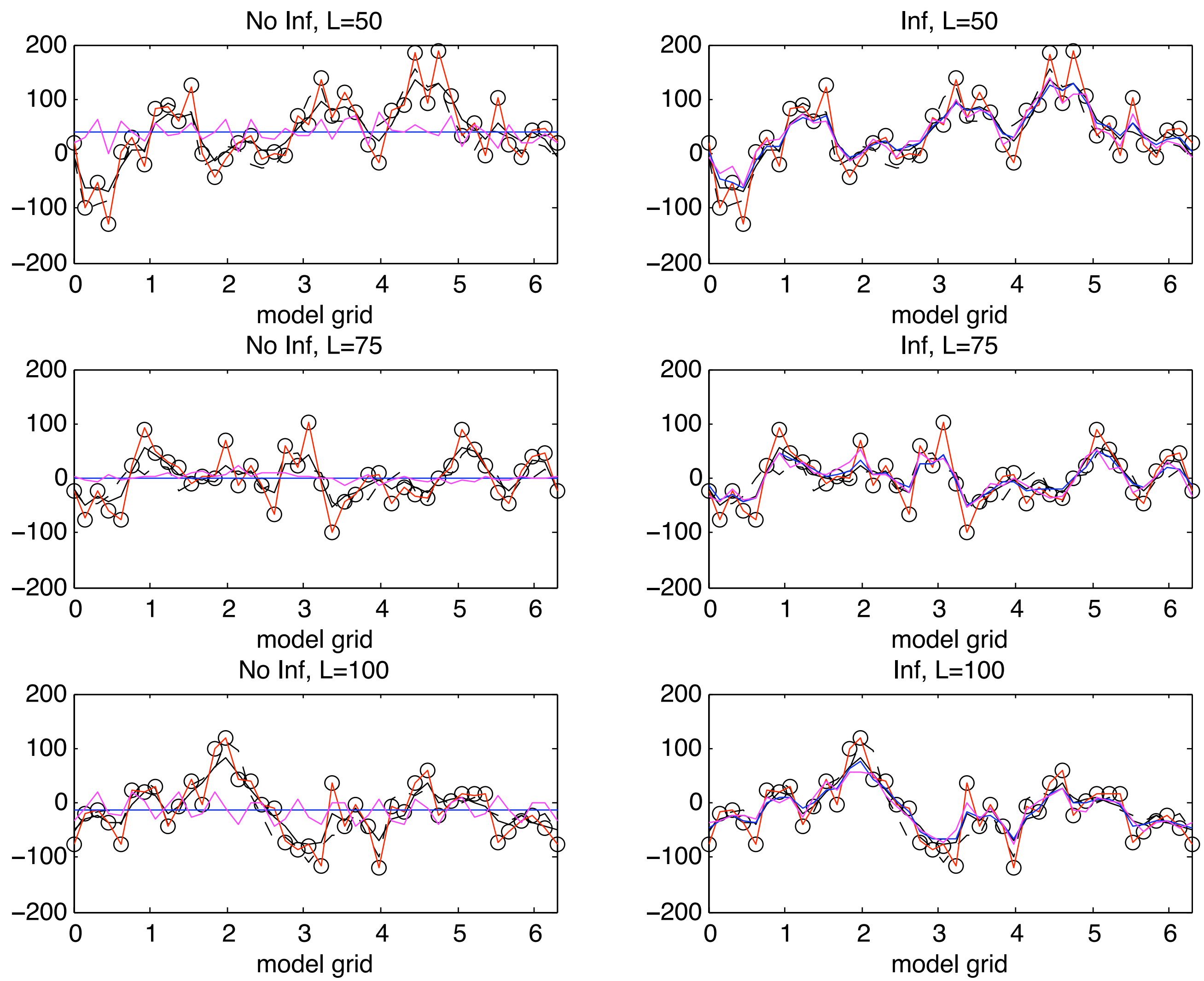
Filtering in Fourier domain vs real domain for N=20



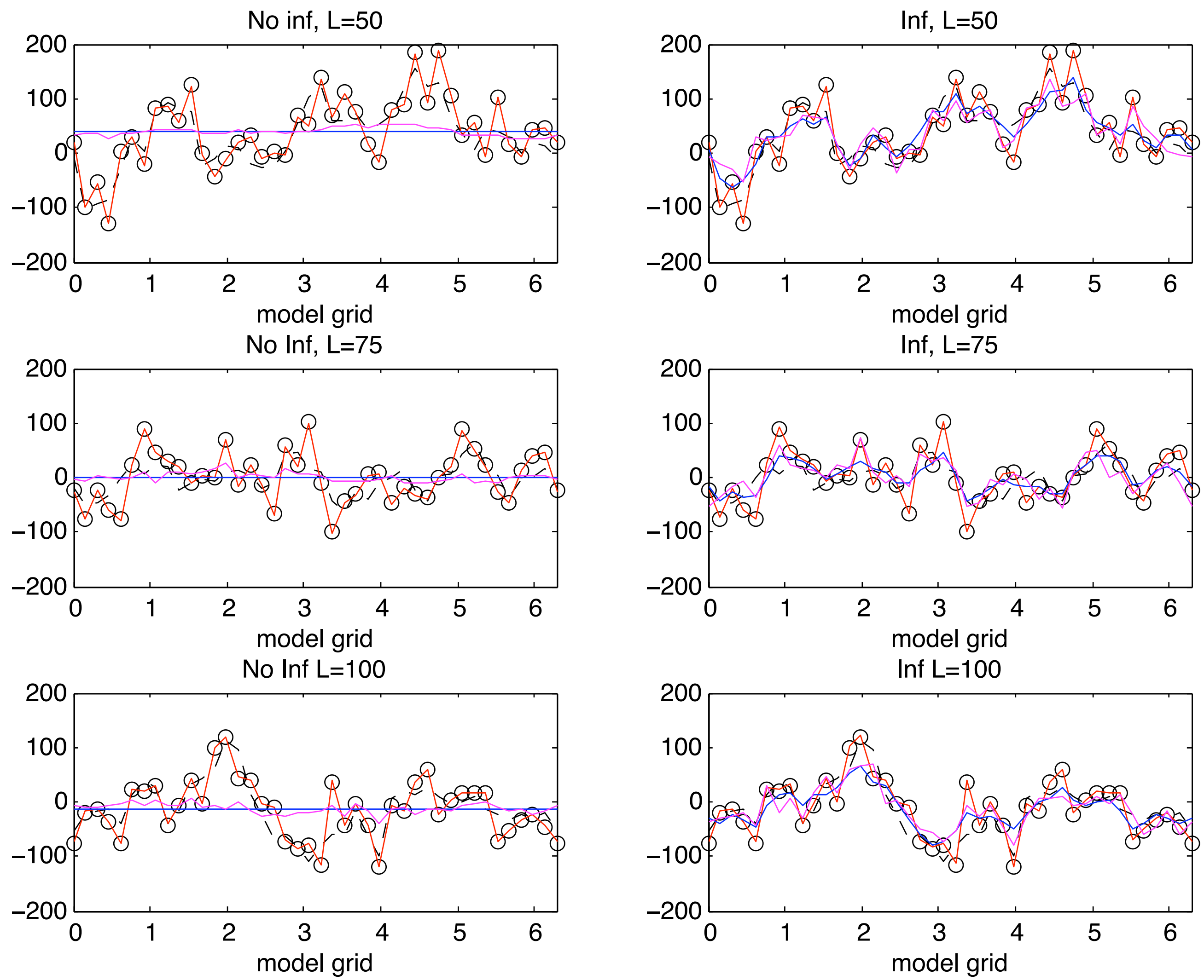
Filtering in Fourier domain vs real domain for N=80



Filtered solution in Fourier domain



Filtered solution in real domain (ETKF)



Summary of testing scalar wave equation with weak damping

From the numerical experiments, we found consistency as that of each scalar Ornstein-Uhlenbeck process and more importantly the offline testing predicts the outcome apriori. The complete summary is as follows:

- The unstable explicit scheme for large time step trusts fully the observations.
- Both the backward-Euler and trapezoidal schemes are significantly better approximate filter for large time step when the information criteria are used.
- The higher the model resolutions may improve the true filtering and hence the approximate filters as well, but not significantly (Theorem is the guideline). However, the error variance reduction is enormous.
- Practically, both implicit filters in the Fourier space are computationally inexpensive with such a giant time step and thus one can afford large ensemble size.
- The scalar Fourier domain filter is not sensitive to the variations of model resolutions, ensemble size, and independent of tunable parameters.

- The omission of accounting the correlation between different modes in our scalar filtering is also found to be insignificant.
- In contrast, an ensemble Kalman filter that mimics the extended Kalman filter (in our experiment, we use ETKF) suggests that the higher the model resolution is, the more realization is needed for filter convergence; it also depends on the variance inflation coefficient.

Filtering the ”poorman” stochastic model for L-96 model

The L-96 model is a given by:

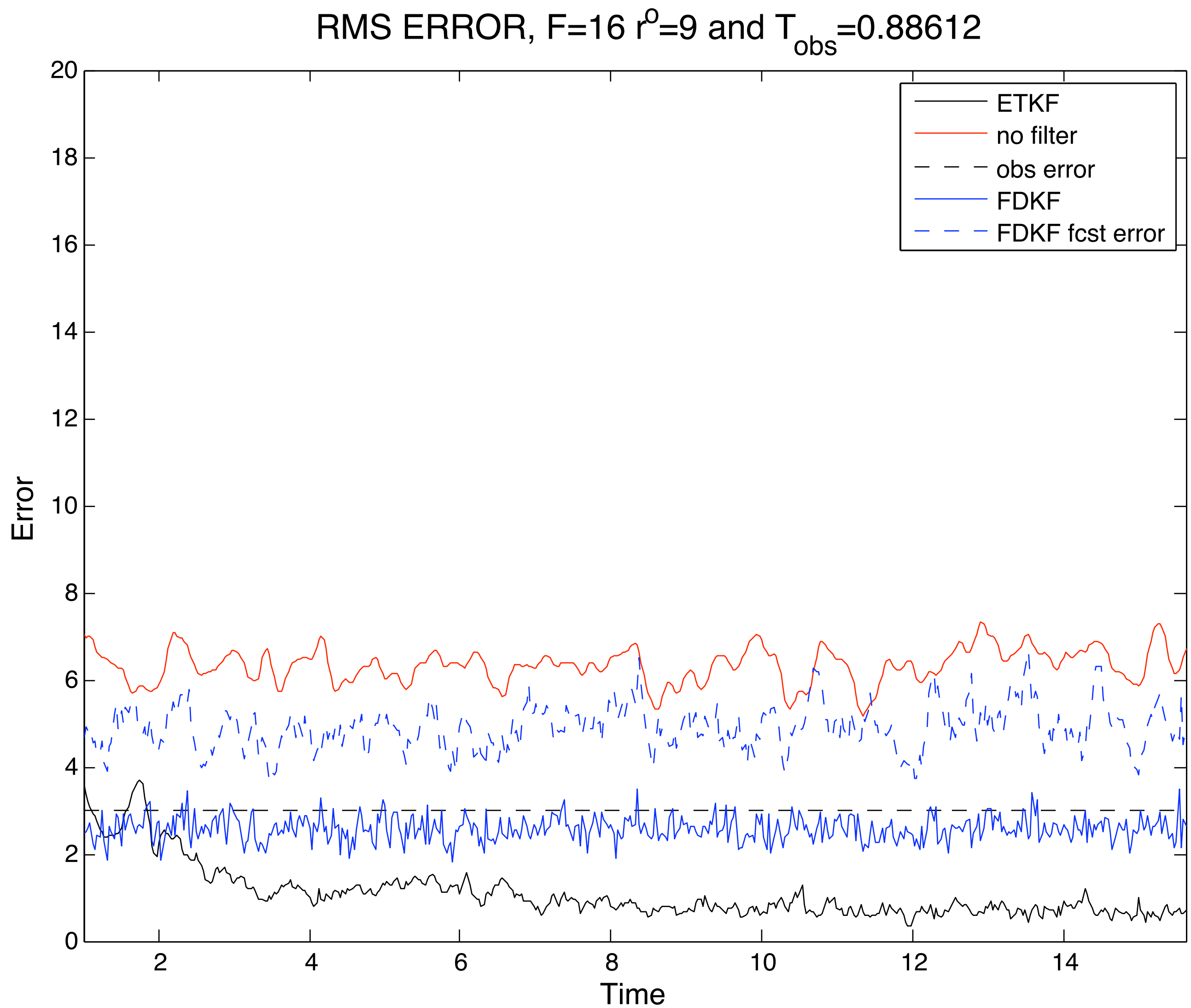
$$\frac{\mathrm{d}u_j}{\mathrm{d}t} = (u_{j+1} - u_{j-2})u_{j-1} - u_j + F,\ j = 0, ..., J-1,$$

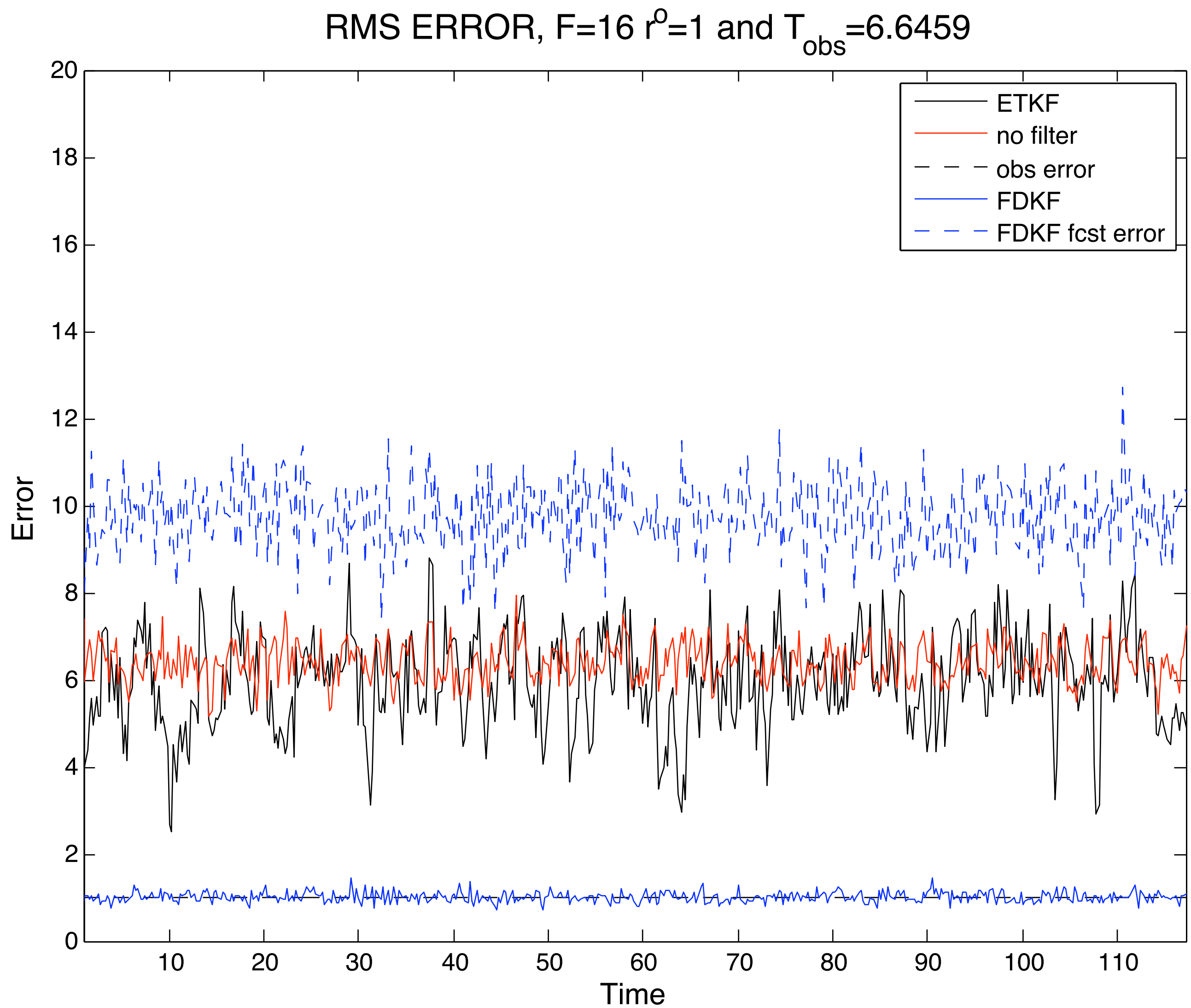
with periodic boundary.

The “poorman” stochastic model is a linearized normalized (Majda and Wang, 2006) model with damping and stochastic forcing. In Fourier space, each mode is as follows:

$$\frac{\mathrm{d}\hat{u}_k(\tilde{t})}{\mathrm{d}\tilde{t}} = \omega(k)\hat{u}_k - d(k)\hat{u}_k(\tilde{t}) + \sigma_k \frac{\mathrm{d}\tilde{W}_k(\tilde{t})}{\mathrm{d}\tilde{t}}$$

We parametrize the damping coefficient $d(k)$ and noise strength σ_k to fit the stationary variance and correlation time.


RMS ERROR, F=16 r^o=9 and T_obs=0.88612
ETKF
no filter
obs error
FDKF
FDKF fcst error
Error
Time
20
18
16
14
12
10
8
6
4
2
0
2
4
6
8
10
12
14



RMS ERROR, F=16 r^o=1 and T_obs=6.6459
Error
Time
0
2
4
6
8
10
12
14
16
18
20
10
20
30
40
50
60
70
80
90
100
110
ETKF
no filter
obs error
FDKF
FDKF fcst error

References:

- Grote and Majda, ”Stable time filtering of strongly unstable spatially extended systems”, PNAS, 2006.
- Majda and Grote, ”Explicit off-line criteria for stable accurate time filtering of strongly unstable spatially extended systems”, PNAS, 2007.
- Castronovo, Harlim, and Majda, ”Mathematical test criteria for filtering complex systems, I: Plentiful observations”, (to be submitted to JCP)
- Castronovo, Harlim, and Majda, ”Mathematical test criteria for filtering complex systems, II: Sparse observations”, (in preparation).
- Harlim and Majda, ”Mathematical criteria for filtering complex dynamical systems”, (in preparation)