Lagrangian Data Assimilation and Manifold Detection for a Point-Vortex Model

Mid-year Progress Report

David Darmon, AMSC

Kayo Ide, AOSC, IPST, CSCAMM, ESSIC

Today's Presentation

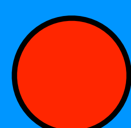
Background

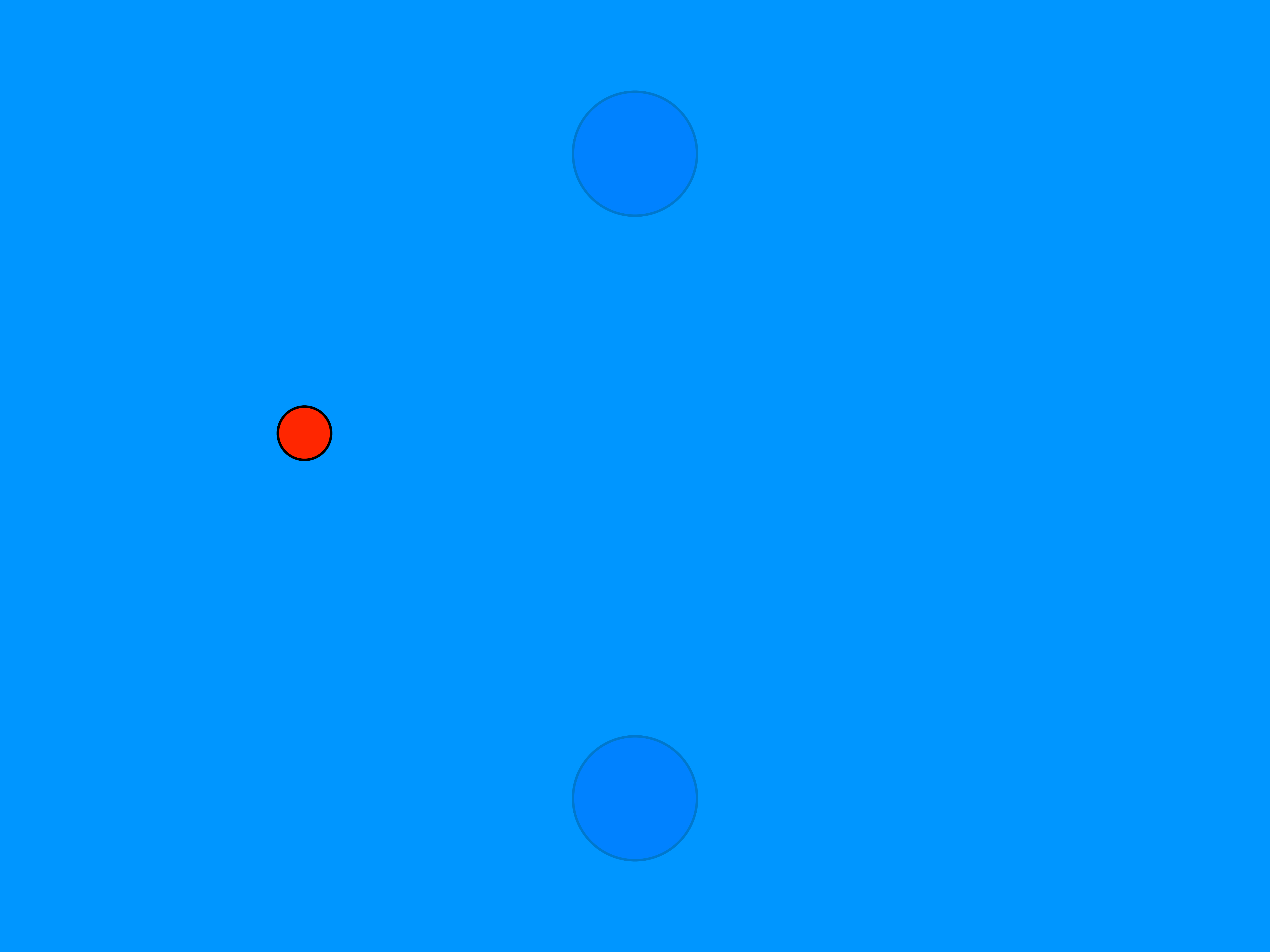
Database Generation

The EKF, Validation and Testing

The EnKF, Validation and Testing

Moving Forward

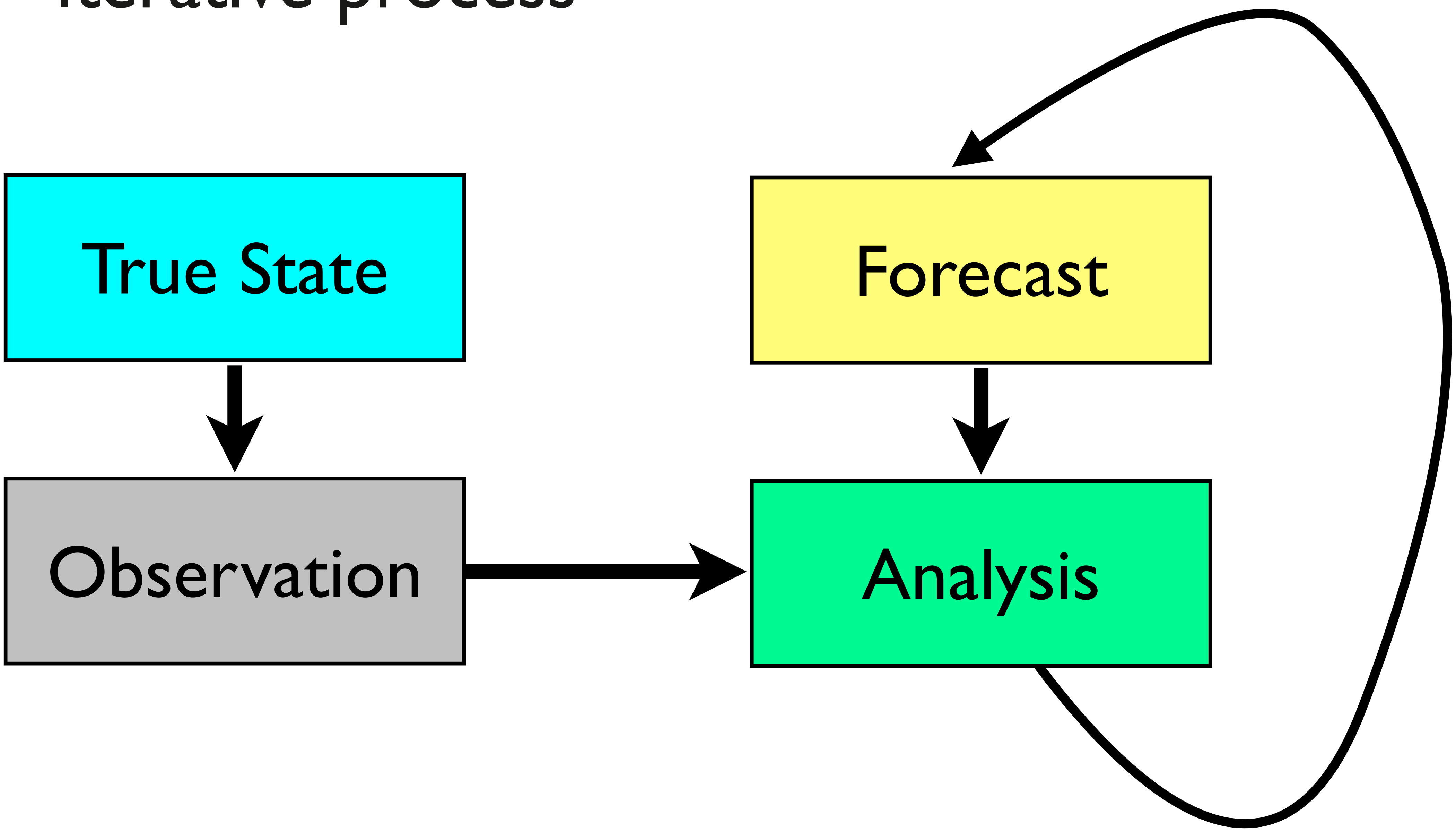




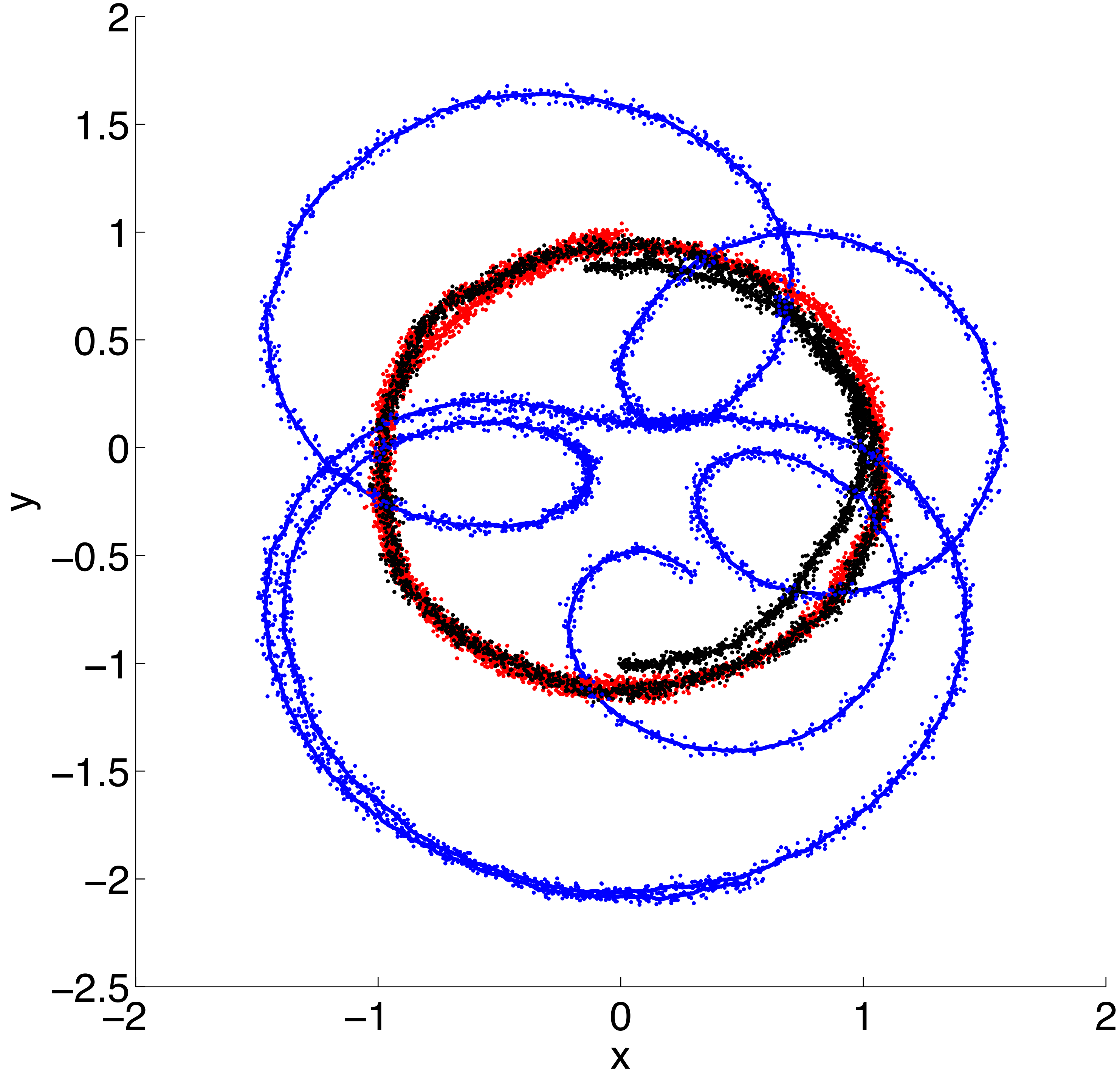
Background

Data Assimilation

Iterative process



Database Generation



Database Generation

True System Dynamics

$$d\mathbf{x}^t = M(\mathbf{x}^t, t)dt + d\boldsymbol{\eta}^t$$

$\mathbf{x}^t$ – system state

$\boldsymbol{\eta}^t$ – Brownian motion

$M(\cdot,\cdot)$ – deterministic evolution operator

Database Generation

Deterministic Dynamics

Point-Vortices

$$\frac{dx_i}{dt} = -\frac{1}{2\pi} \sum_{i'=1, i' \neq i}^{N_v} \frac{\Gamma_{i'}(y_i - y_{i'})}{l_{ii'}^2}$$

$$\frac{dy_i}{dt} = \frac{1}{2\pi} \sum_{i'=1, i' \neq i}^{N_v} \frac{\Gamma_{i'}(x_i - x_{i'})}{l_{ii'}^2}$$

$$l_{ii'}^2 = (x_i - x_{i'})^2 + (y_i - y_{i'}^2)$$

Database Generation

Deterministic Dynamics

Drifters

$$\frac{d\xi_i}{dt} = -\frac{1}{2\pi}\sum_{j=1}^{N_v}\frac{\Gamma_j(\eta_i - y_j)}{l_{ij}^2}$$

$$\frac{d\eta_i}{dt} = \frac{1}{2\pi}\sum_{j=1}^{N_v}\frac{\Gamma_j(\xi_i - x_j)}{l_{ij}^2}$$

$$l_{ij}^2 = (\xi_i - x_j)^2 + (\eta_i - y_j^2)$$

Database Generation

Scalar vs. Vector Case

Stochastic Differential Equation

$$dX(t) = f(X(t), t)dt + g(X(t), t)dW(t)$$
$$X(0) = X_0, 0 \leq t \leq T$$

$$d\mathbf{x}^t(t) = M(\mathbf{x}^t(t), t)dt + d\boldsymbol{\eta}^t(t)$$
$$\mathbf{x}^t(0) = \mathbf{x}_0^t, 0 \leq t \leq T$$

Database Generation

Numerical Solution

Runge-Kutta

$$h(X_j, \tau_j) = \left[f(X_j, \tau_j) - \frac{1}{2} g(X_j, \tau_j) \frac{\partial g(X_j, \tau_j)}{\partial X_j} \right] \Delta\tau + g(X_j, \tau_j)(W(\tau_j) - W(\tau_{j-1}))$$

Database Generation

Numerical Solution

Runge-Kutta

$$K_1 = h(X_j, \tau_j)$$

$$K_2 = h(X_j + \frac{1}{2}K_1, \tau_j + \frac{1}{2}\Delta\tau)$$

$$K_3 = h(X_j + \frac{1}{2}K_2, \tau_j + \frac{1}{2}\Delta\tau)$$

$$K_4 = h(X_j + K_3, \tau_j + \Delta\tau)$$

$$X_{j+1} = X_j + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

Database Generation

Numerical Solution

Runge-Kutta

$$K_1 = h(X_j, \tau_j)$$

$$K_2 = h(X_j + \frac{1}{2}K_1, \tau_j + \frac{1}{2}\Delta\tau)$$

$$K_3 = h(X_j + \frac{1}{2}K_2, \tau_j + \frac{1}{2}\Delta\tau)$$

$$K_4 = h(X_j + K_3, \tau_j + \Delta\tau)$$

$$\boxed{X_{j+1} = X_j + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)}$$

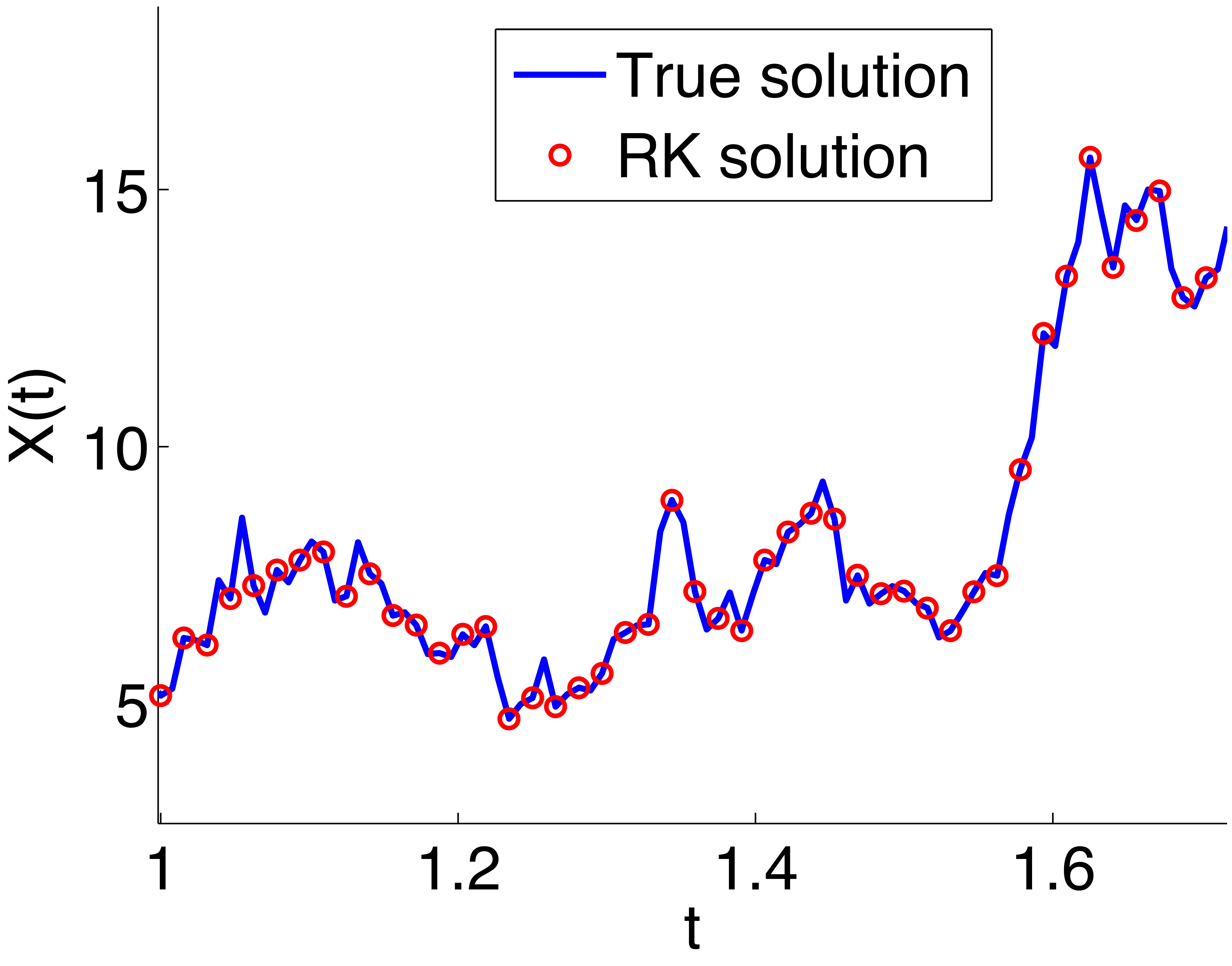
Database Generation

Simple Scalar Case

Validation

$$dX(t) = \lambda X(t)dt + \mu X(t)dW(t)$$
$$X(0) = X_0$$

$$X(t) = X_0 \exp\left[\left(\lambda - \frac{1}{2}\mu^2\right)t + \mu W(t)\right]$$



True solution
RK solution
X(t)
t
5
10
15
1
1.2
1.4
1.6

Database Generation

True Dynamics

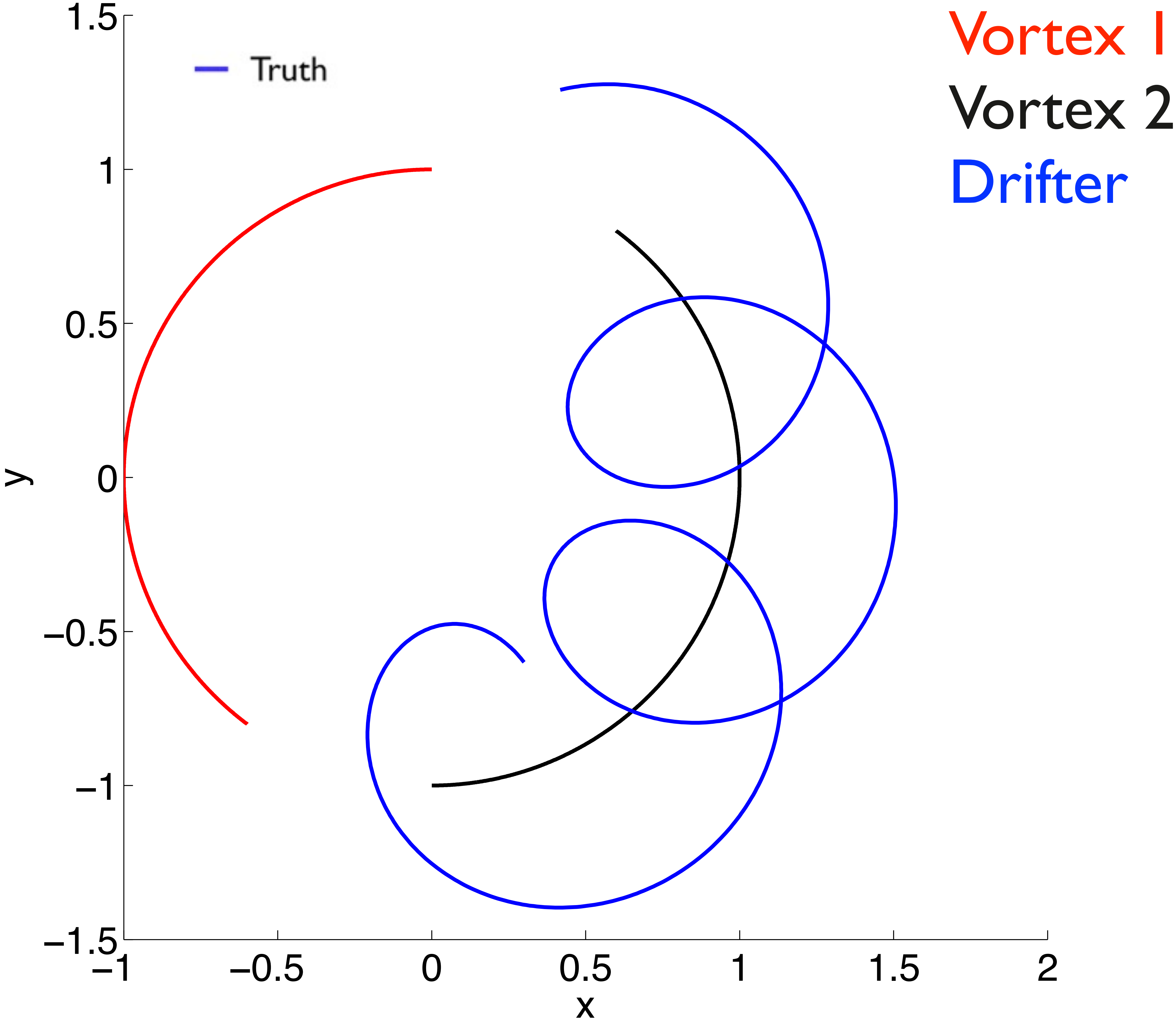
Stochastic Differential Equation

$$d\mathbf{x}^t = M(\mathbf{x}^t, t)dt + d\boldsymbol{\eta}^t$$

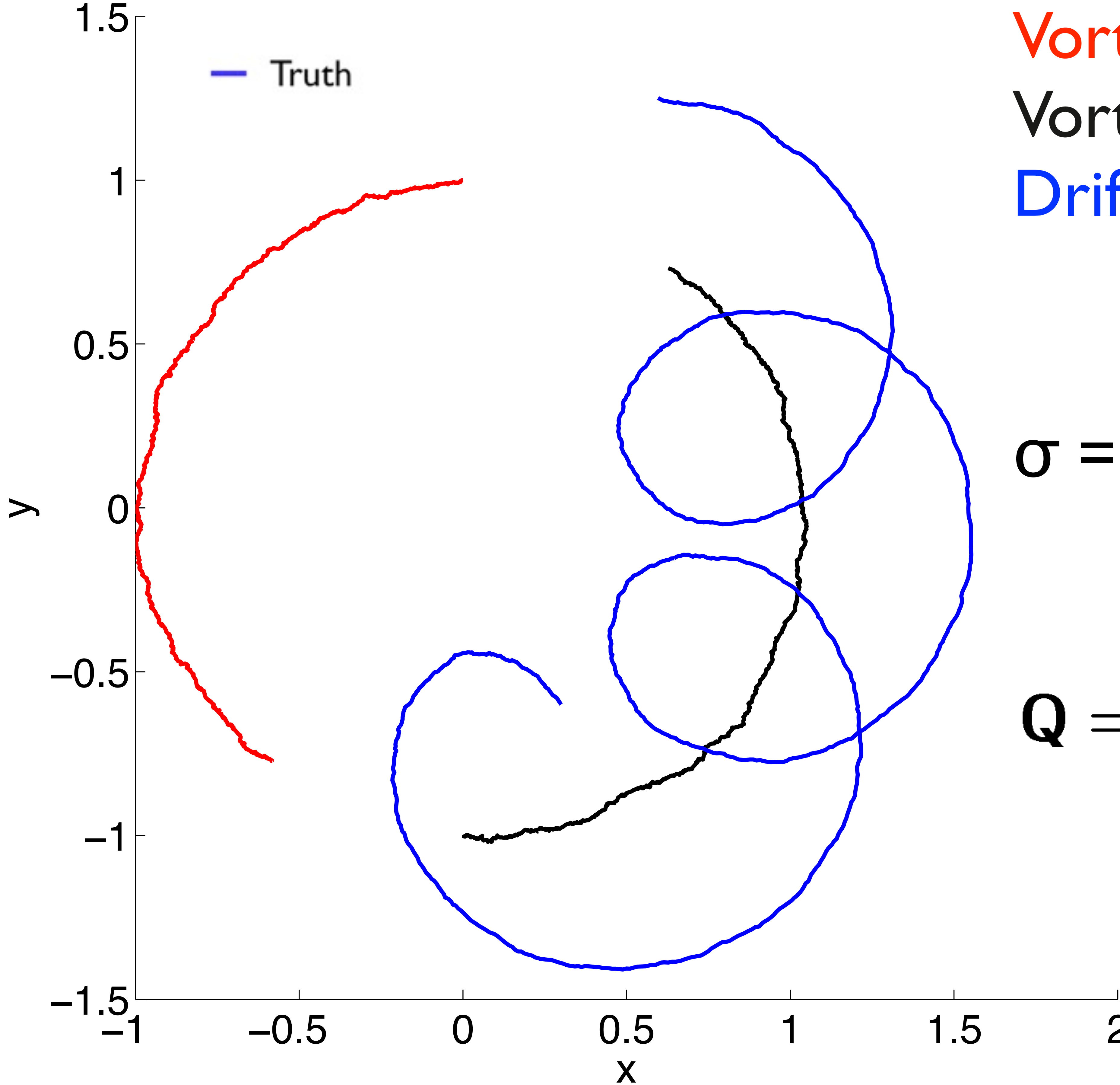
$$d\mathbf{X} = \mathbf{f}(\mathbf{X}, t)dt + \mathbf{g}(\mathbf{X}, t)d\mathbf{W}(t)$$

$$\mathbf{f}(\mathbf{x}^t, t) = M(\mathbf{x}^t, t)$$

$$\mathbf{g}(\mathbf{x}^t, t) = \sqrt{2}\mathbf{Q}^{1/2}$$



Truth
Vortex 1
Vortex 2
Drifter
1.5
1
0.5
0
−0.5
−1
−1.5
y
−1
−0.5
0
0.5
1
1.5
2
x



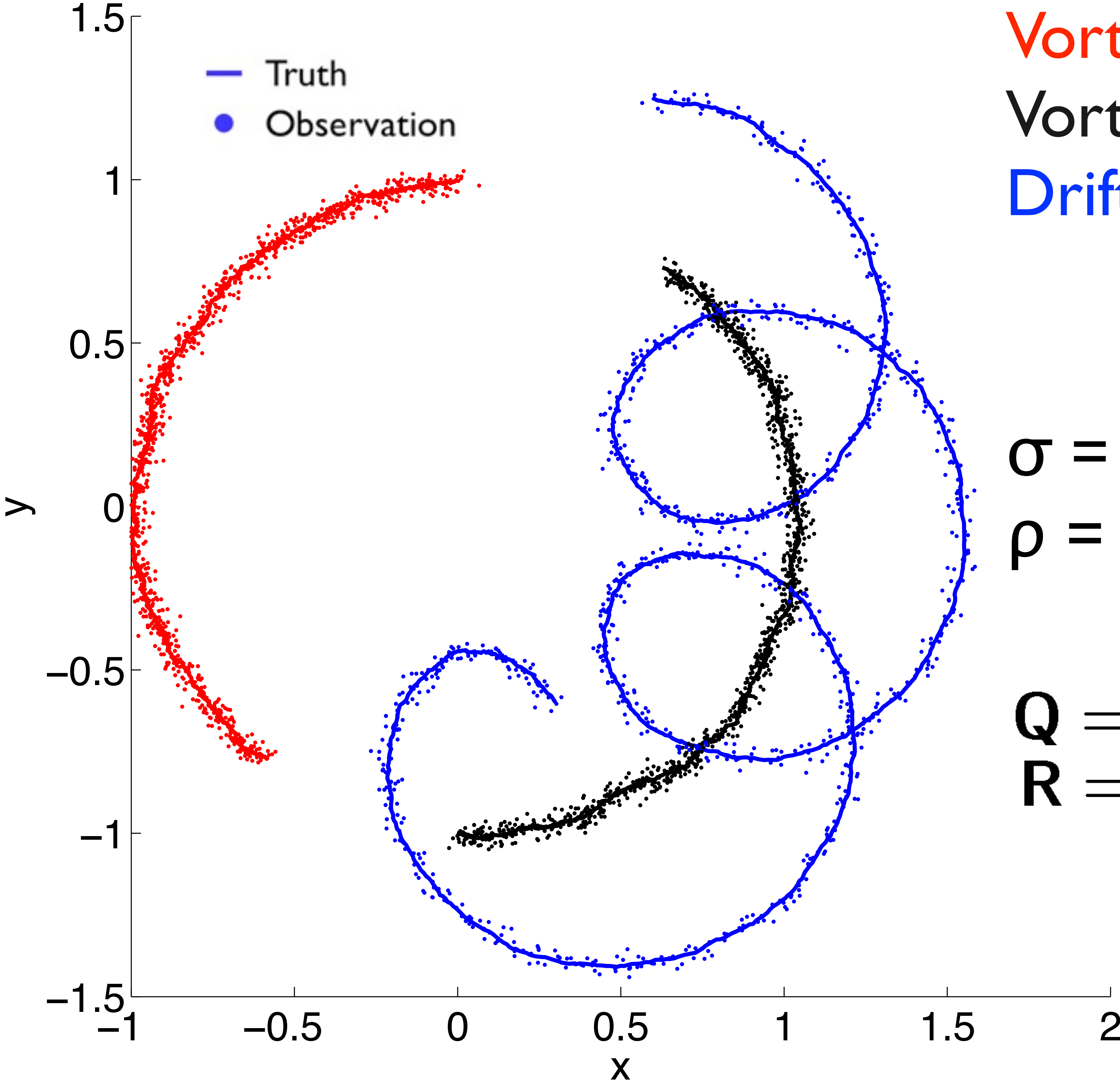
Vortex 1

Vortex 2

Drifter

$\sigma = 0.02$

$$\mathbf{Q} = 2\sigma^2 \mathbf{I}$$

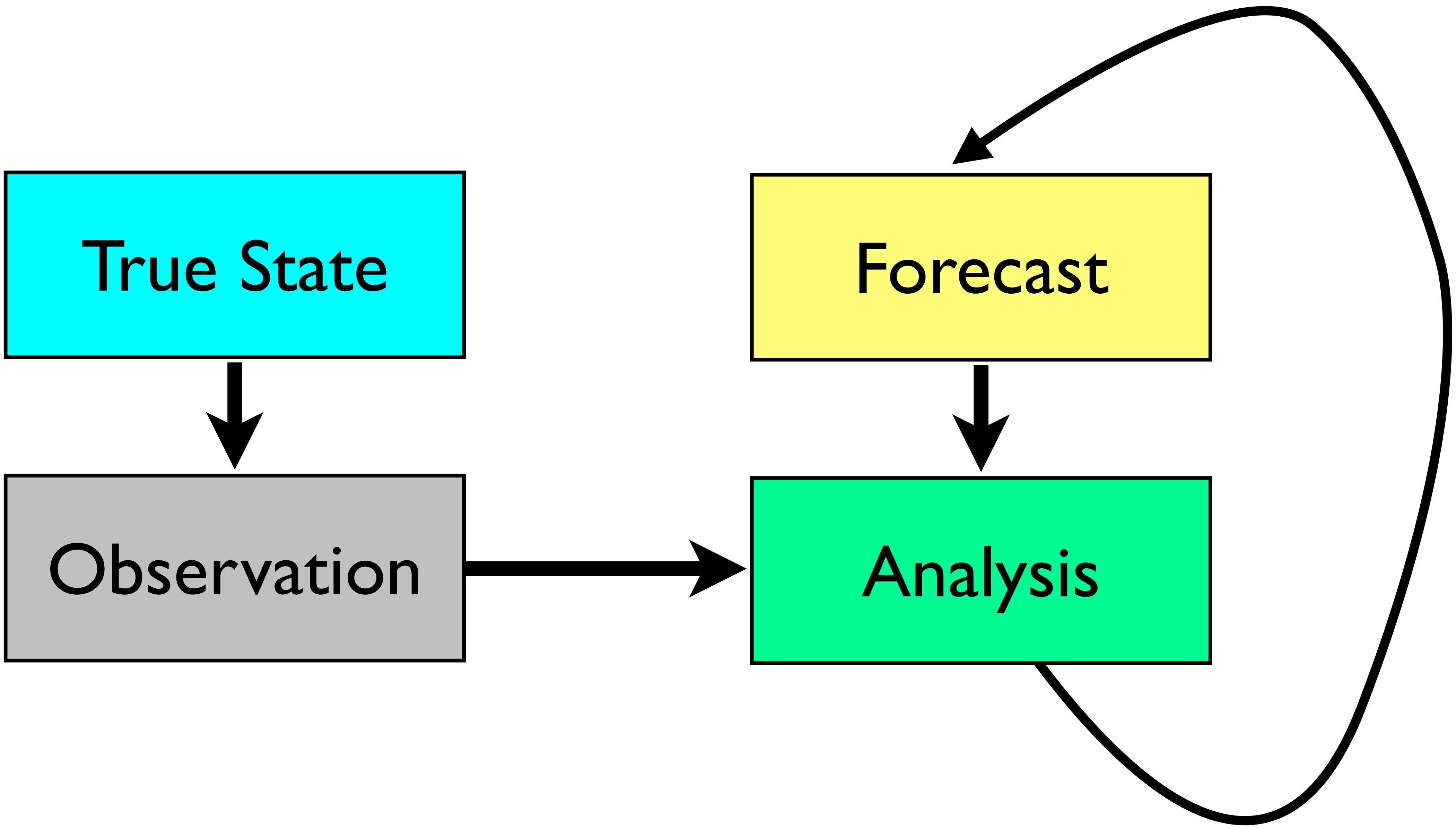


Vortex 1
Vortex 2
Drifter

σ = 0.02
ρ = 0.02

$$\mathbf{Q} = 2\sigma^2\mathbf{I}$$
$$\mathbf{R} = \rho^2\mathbf{I}$$

Data Assimilation



Data Assimilation

Approach 1:
EKF

Approach 1

Extended Kalman Filter (EKF), Forecast

$$\frac{d}{dt}\mathbf{x}^f = M(\mathbf{x}^f, t)$$

$$\frac{d}{dt}\mathbf{P}^f = \mathbf{M}(t)\mathbf{P}^f + \mathbf{P}^f\mathbf{M}^T(t) + \mathbf{Q}$$

$\mathbf{Q}$ – covariance matrix from SDE

$\mathbf{M}(t) = J[M(\mathbf{x}, t)]\big|_{x=x^f}$ - Jacobian of M

Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

$\mathbf{x}_1$ - Reference Trajectory

$\mathbf{x}_2$ - Perturbed Trajectory

Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

$\mathbf{x}_1$ - Reference Trajectory

$\mathbf{x}_2$ - Perturbed Trajectory

$$\mathbf{x}_1(0) = \mathbf{x}^0$$
$$\mathbf{x}_2(0) = \mathbf{x}^0 + \mathbf{y}^0$$

Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

$\mathbf{x}_1$ - Reference Trajectory

$\mathbf{x}_2$ - Perturbed Trajectory

$$\mathbf{x}_1(0) = \mathbf{x}^0$$
$$\mathbf{x}_2(0) = \mathbf{x}^0 + \mathbf{y}^0$$

↑

Tangent Vector (small)

Approach 1

Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

$$\frac{d\mathbf{y}}{dt} = \left.\frac{\partial M(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathrm{x}=\mathrm{x}(t)} \mathbf{y}$$

$$= \mathbf{M}(\mathbf{x}(t))\mathbf{y}$$

Approach 1

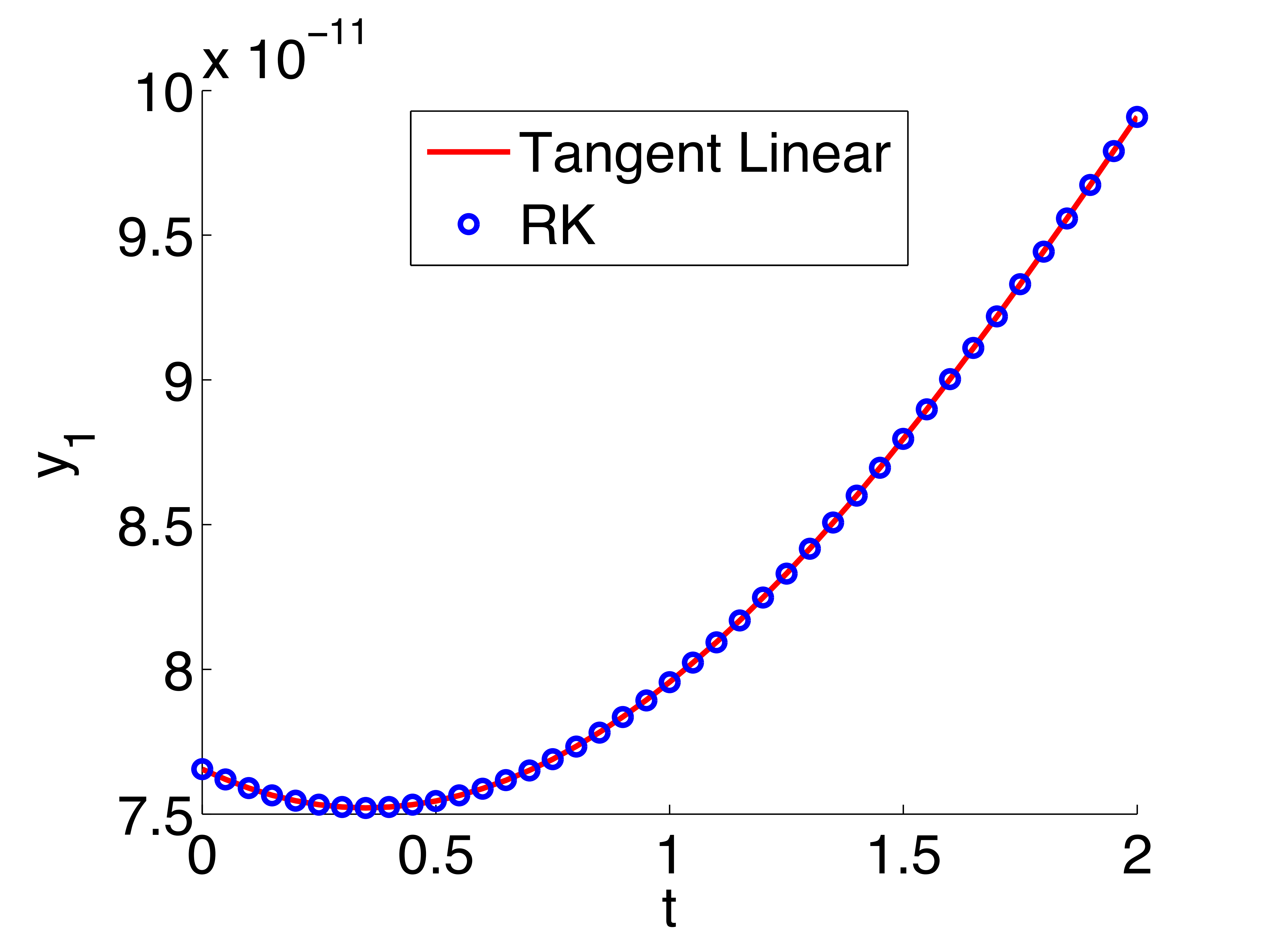
Extended Kalman Filter (EKF)

Validation - Jacobian via Tangent Linear Model

Evolve reference and perturbed trajectories forward using deterministic ODE

Evolve tangent vector forward using Tangent Linear Model

Compare $\mathbf{y}_{\text{perturbed}}$ and $\mathbf{y}_{\text{TL}}$


x 10^−11
10
9.5
9
8.5
8
7.5
y_1
Tangent Linear
RK
0
0.5
1
1.5
2
t

Approach 1

Extended Kalman Filter (EKF), Analysis

$$\mathbf{x}_k^a = \mathbf{x}^f(t_k) + \mathbf{K}_k(\mathbf{y}_k^o - h_k(\mathbf{x}^f(t_k)))$$

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^f(t_k)$$

$$\mathbf{K}_k = \mathbf{P}^f(t_k)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}^f(t_k)\mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}$$

$\mathbf{R}^o$ – covariance matrix from observation

$\mathbf{H}_k = J[h(\mathbf{x}, t_k)]\Big|_{x=x^f}$ - Jacobian of h_k

Approach 1

Extended Kalman Filter (EKF)

Validation - No Noise

$$\mathbf{x}_k^a = \mathbf{x}^f(t_k) + \mathbf{K}_k(\mathbf{y}_k^o - h_k(\mathbf{x}^f(t_k)))$$

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^f(t_k)$$

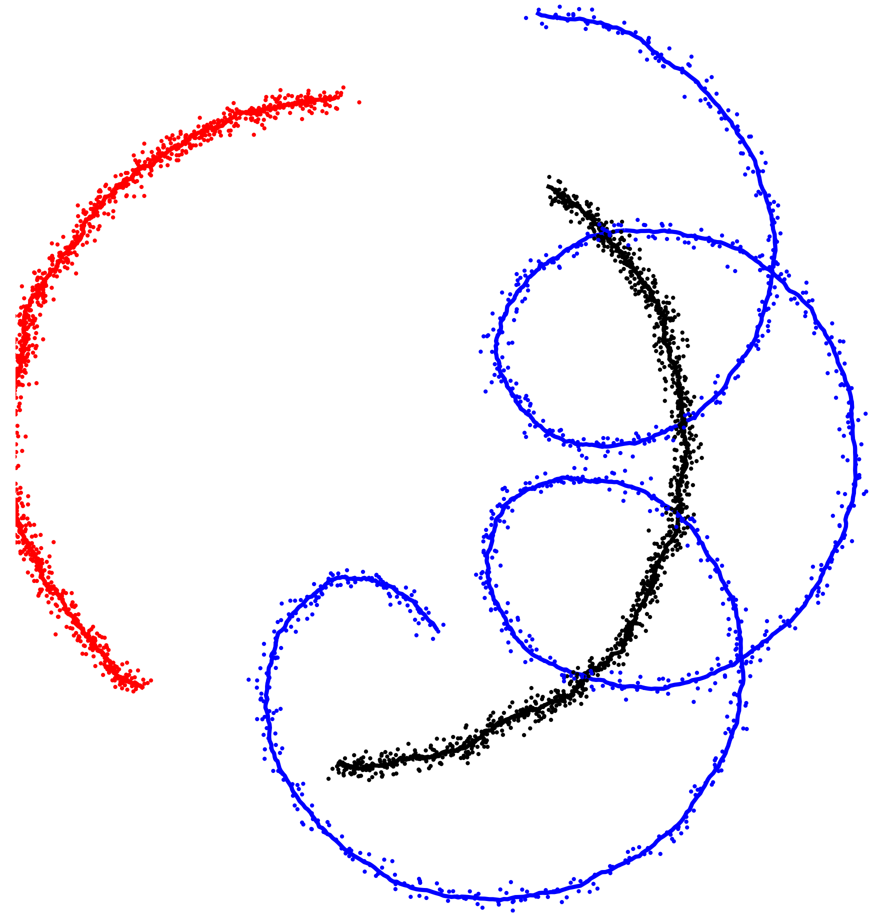
$$\mathbf{K}_k = \mathbf{P}^f(t_k)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}^f(t_k)\mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}$$

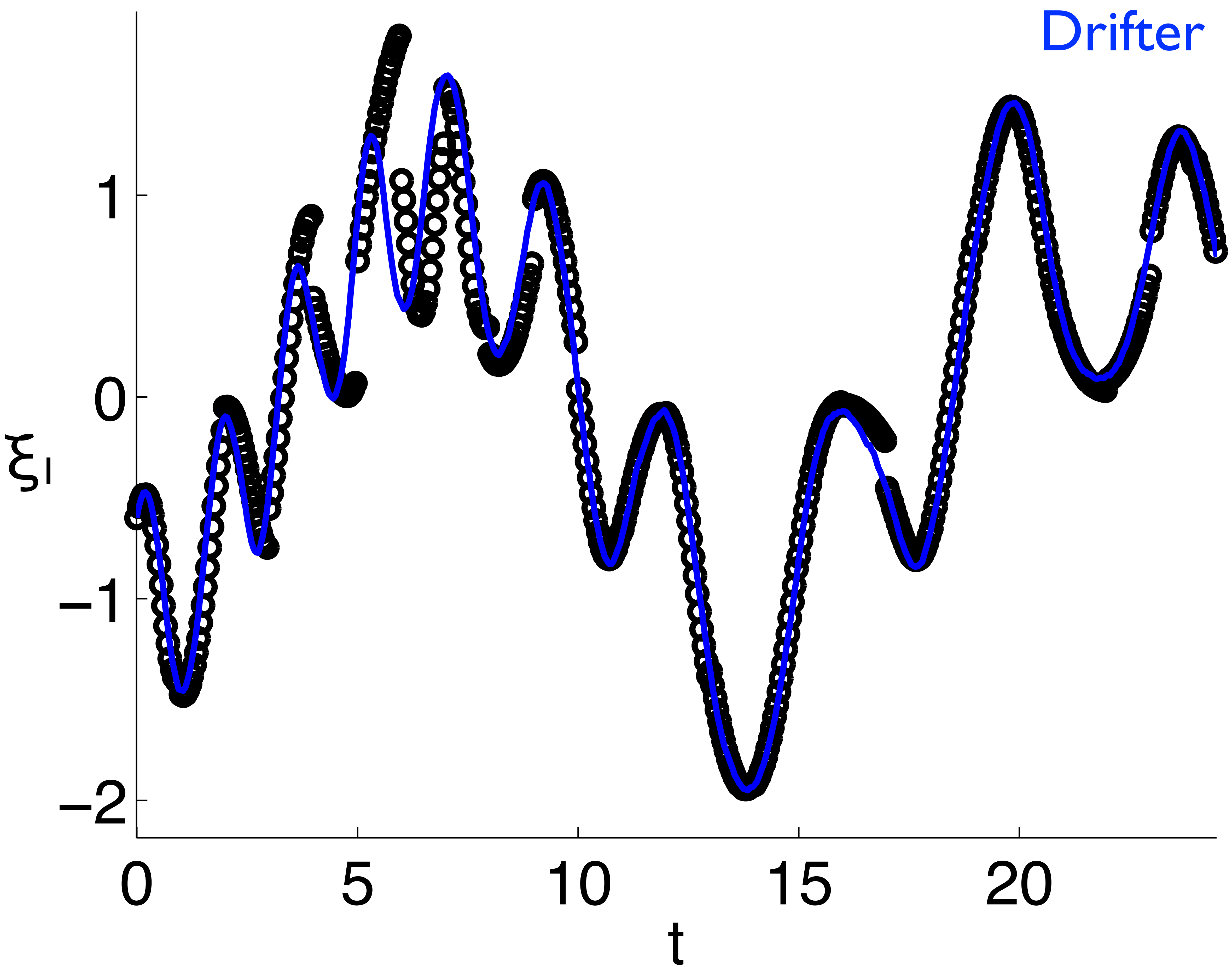
Approach 1

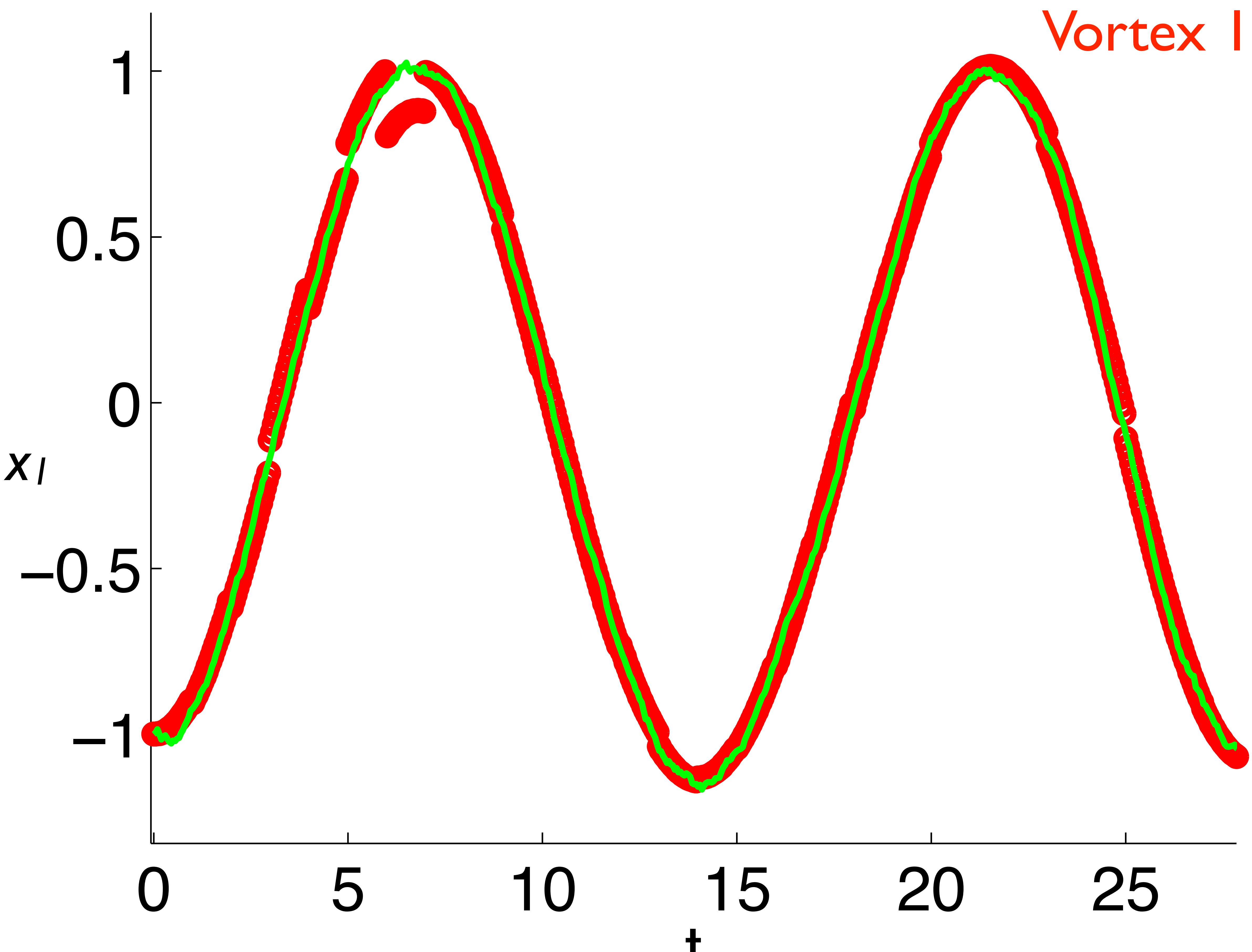
Extended Kalman Filter (EKF)

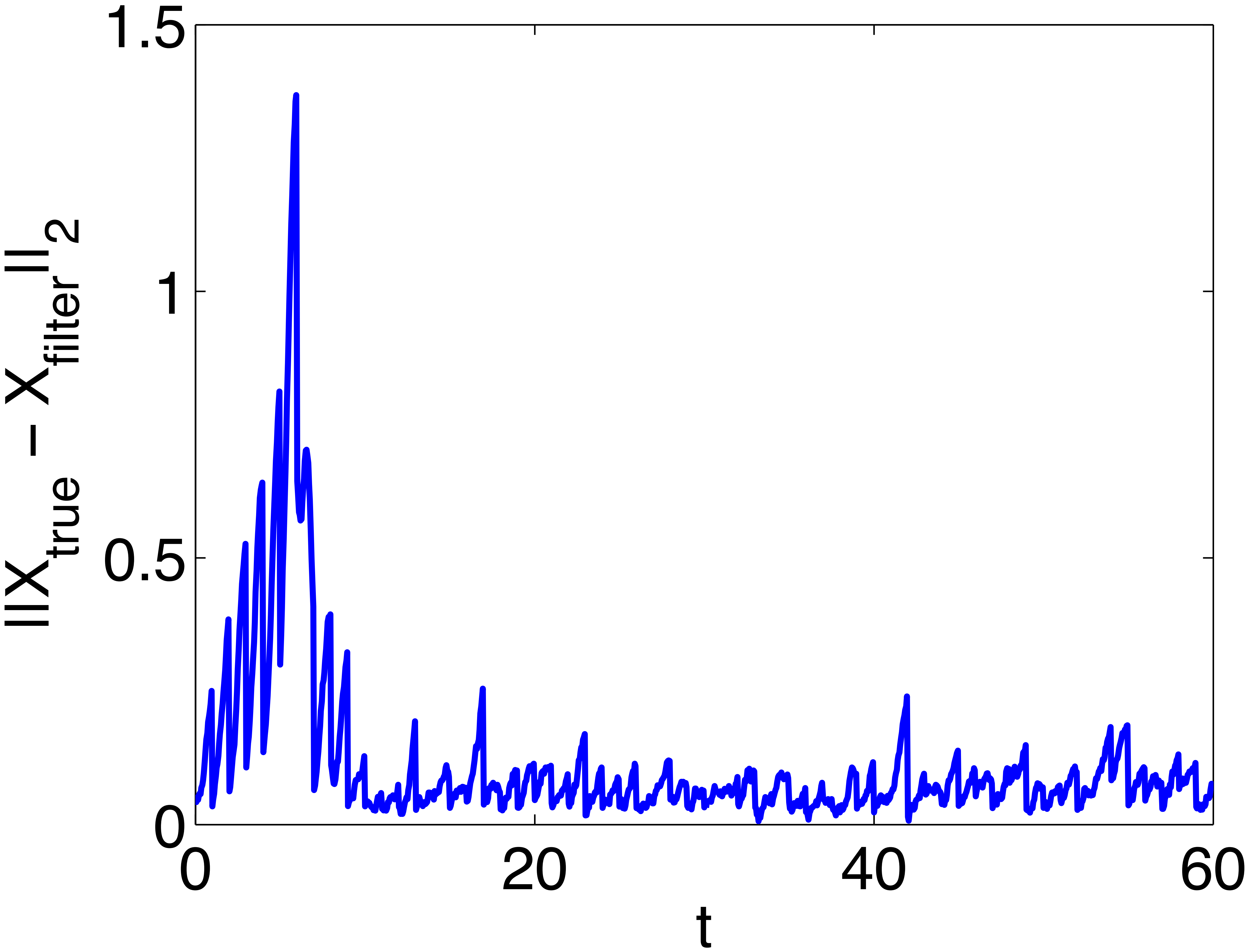
Validation - Full Observation, Imperfect Data

Observe vortices and drifter under observational noise




Drifter
1
0
−1
−2
0
5
10
15
20
t
ξ


Vortex 1
1
0.5
0
−0.5
−1
x_I
0
5
10
15
20
25
t



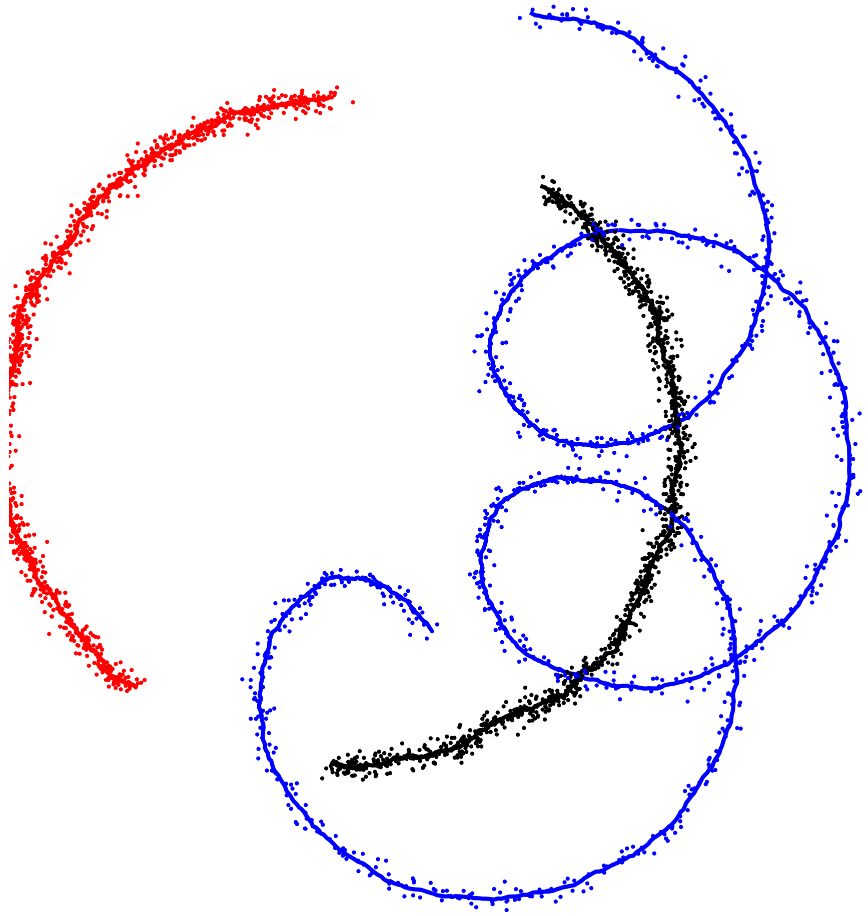
1.5
1
0.5
0
0
20
40
60
t
$||X_{true} - X_{filter}||_2$

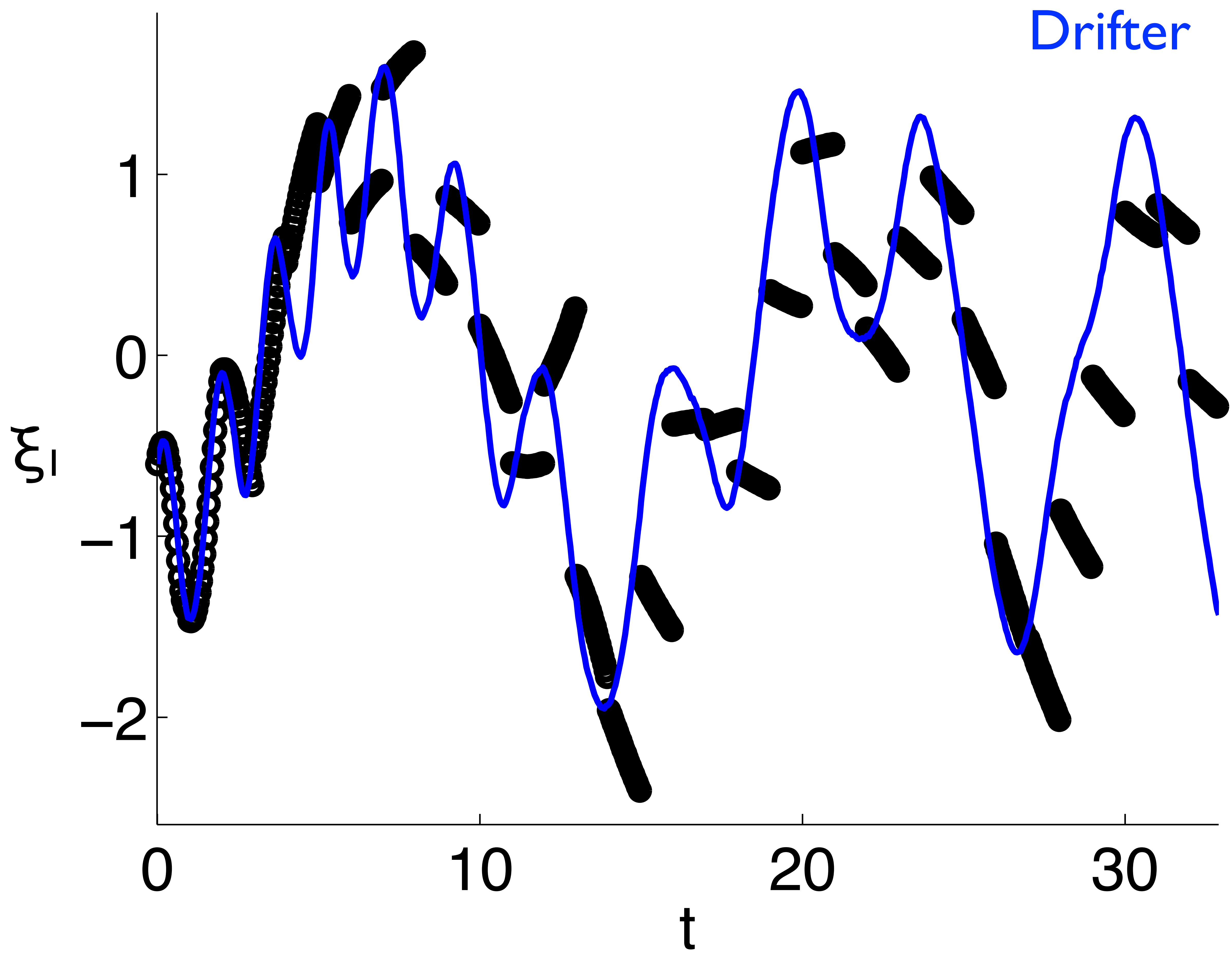
Approach 1

Extended Kalman Filter (EKF)

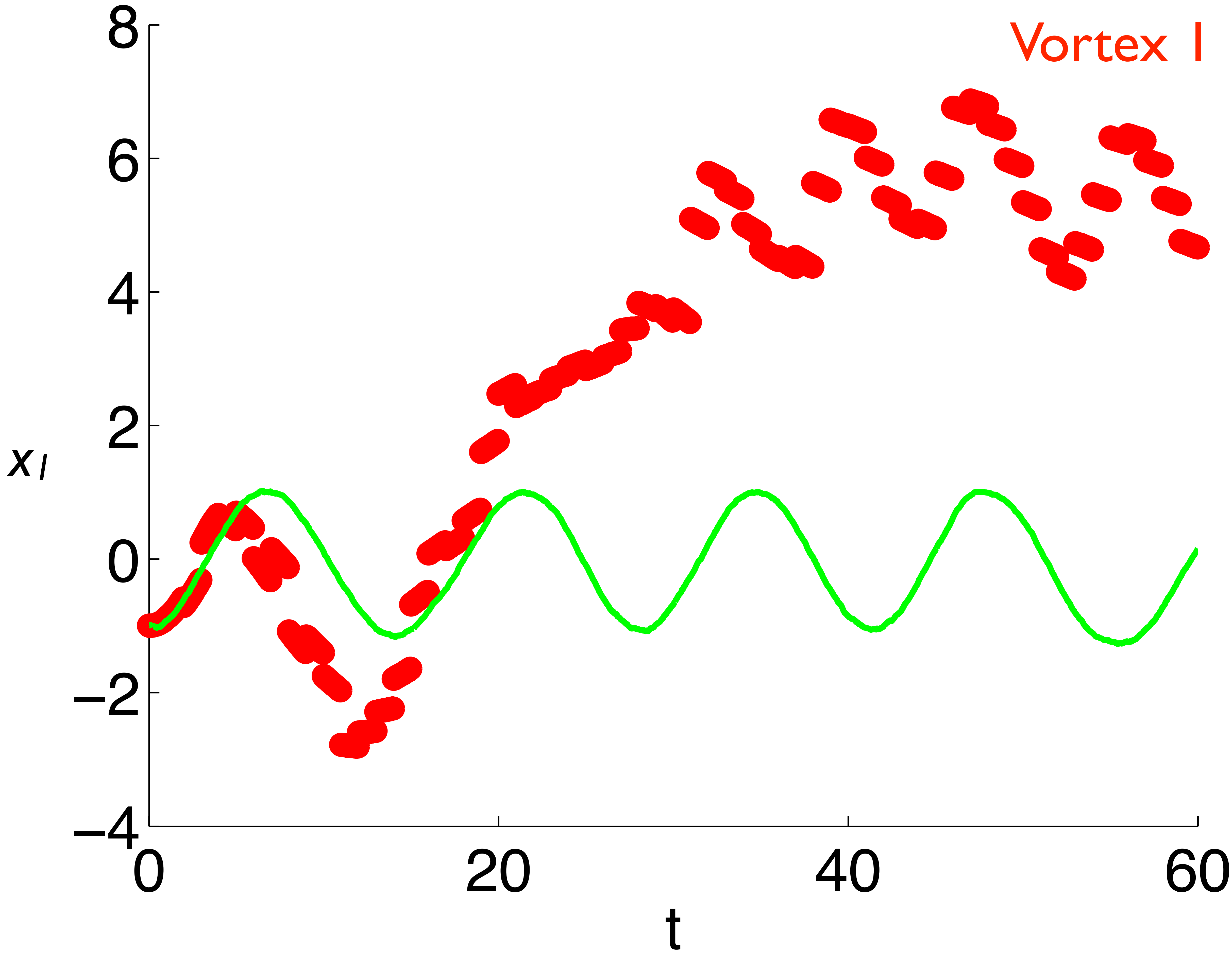
Results - Partial Observation, Imperfect Data

Observe only drifter under observational noise

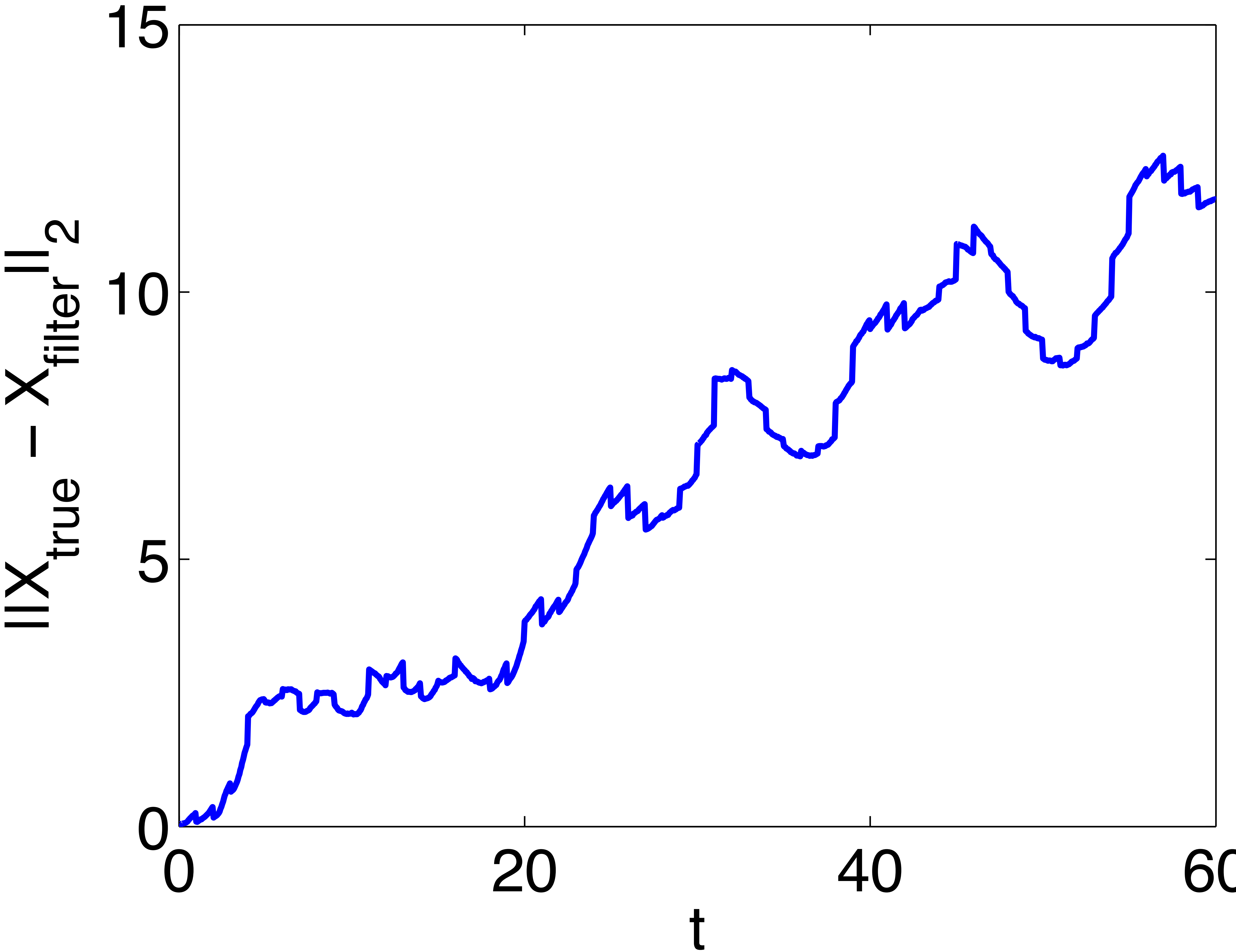




Drifter
1
0
−1
−2
ξ
0
10
20
30
t



Vortex 1
8
6
4
2
0
−2
−4
x_1
0
20
40
60
t



15
10
5
0
0
20
40
60
t
||X_true − X_filter||_2

Data Assimilation

Approach 1:
EnKF

Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

$$\{\mathbf{x}_i^f\}_{i=1}^N$$

Forecast: Evolve particles forward using SDE

Analysis: Perform Kalman-like analysis

Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

$$\{\mathbf{x}_i^f\}_{i=1}^N$$

Forecast: Evolve particles forward using ~~SDE~~ ODE

Analysis: Perform Kalman-like analysis

Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

$$\{\mathbf{x}_i^f\}_{i=1}^N$$

Forecast: Evolve particles forward using ODE

Analysis: Perform Kalman-like analysis

Approach 2a

EnKF with Observational Perturbations

$$\mathbf{y}^o_{k,i} = \mathbf{y}^o_k + \epsilon_i, \qquad i = 1, \dots, n$$

$$\epsilon_i \sim N(\mathbf{0}, \mathbf{R}^o)$$

Approach 2a

EnKF with Observational Perturbations

$$\mathbf{y}^o_{k,i} = \mathbf{y}^o_k + \epsilon_i, \qquad i = 1, \ldots, n$$

$$\epsilon_i \sim N(\mathbf{0}, \mathbf{R}^o)$$

$\{\mathbf{x}^f_i\}_{i=1}^N$ Forecast ensemble

$\{\mathbf{y}^o_{k,i}\}_{i=1}^N$ Observation ensemble

Approach 2a

EnKF with Observational Perturbations

Kalman-like update

$$\mathbf{x}^a_{k,i} = \mathbf{x}^f_i(t_k) + \mathbf{K}_k(\mathbf{y}^o_{k,i} - h_k(\mathbf{x}^f_i(t_k)))$$

$$\mathbf{K}_k = \mathbf{P}^f_e(t_k)\mathbf{H}^T_k(\mathbf{H}_k\mathbf{P}^f_e(t_k)\mathbf{H}^T_k + \mathbf{R}^o_k)^{-1}$$

Approach 2a

EnKF with Observational Perturbations

Kalman-like update

$$\mathbf{x}^a_{k,i} = \mathbf{x}^f_i(t_k) + \mathbf{K}_k(\mathbf{y}^o_{k,i} - h_k(\mathbf{x}^f_i(t_k)))$$

$$\mathbf{K}_k = \mathbf{P}^f_e(t_k)\mathbf{H}^T_k(\mathbf{H}_k\mathbf{P}^f_e(t_k)\mathbf{H}^T_k + \mathbf{R}^o_k)^{-1}$$

Sample Covariance Matrix

$$\mathbf{P}^f_e(t_k) = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}^f_i - \bar{\mathbf{x}}^f)(\mathbf{x}^f_i - \bar{\mathbf{x}}^f)^T$$

Approach 2

Ensemble Kalman Filter (EnKF)

Approximate pdf by an ensemble of particles:

$$\{\mathbf{x}_i^f\}_{i=1}^N$$

Forecast: Evolve particles forward using ~~SDE~~ ODE

Analysis: Perform Kalman-like analysis

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

$$\mathbf{x}_k^a = \mathbf{x}^f(t_k) + \mathbf{K}_k(\mathbf{y}_k^o - h_k(\mathbf{x}^f(t_k)))$$

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^f(t_k)$$

$$\mathbf{K}_k = \mathbf{P}^f(t_k)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}^f(t_k)\mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

$$\mathbf{x}_k^a = \mathbf{x}^f(t_k) + \mathbf{K}_k(\mathbf{y}_k^o - h_k(\mathbf{x}^f(t_k)))$$

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^f(t_k)$$

$$\mathbf{K}_k = \mathbf{P}^f(t_k)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}^f(t_k)\mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}$$

Use sample estimates:

$$\bar{\mathbf{x}}_k^a = \bar{\mathbf{x}}^f(t_k) + \mathbf{K}_{e,k}(\mathbf{y}_k^o - h_k(\bar{\mathbf{x}}^f(t_k)))$$

$$\mathbf{P}_{e,k}^a = (\mathbf{I} - \mathbf{K}_{e,k}\mathbf{H}_k)\mathbf{P}_e^f(t_k)$$

$$\mathbf{K}_{e,k} = \mathbf{P}_e^f(t_k)\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_e^f(t_k)\mathbf{H}_k^T + \mathbf{R}_k^o)^{-1}$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

$$\mathbf{X}^f = \left[\begin{array}{cccc} \mathbf{x}_1^f - \bar{\mathbf{x}} & \mathbf{x}_2^f - \bar{\mathbf{x}} & \ldots & \mathbf{x}_n^f - \bar{\mathbf{x}} \end{array} \right]$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

$$\mathbf{X}^f = \left[\begin{array}{cccc} \mathbf{x}_1^f - \bar{\mathbf{x}} & \mathbf{x}_2^f - \bar{\mathbf{x}} & \dots & \mathbf{x}_n^f - \bar{\mathbf{x}} \end{array}\right]$$

$$\frac{1}{N-1}\mathbf{X}^f (\mathbf{X}^f)^T = \mathbf{P}_e^f$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

$$\mathbf{X}^f = \left[\begin{array}{cccc} \mathbf{x}_1^f - \bar{\mathbf{x}} & \mathbf{x}_2^f - \bar{\mathbf{x}} & \ldots & \mathbf{x}_n^f - \bar{\mathbf{x}} \end{array} \right]$$

$$\frac{1}{N-1}\mathbf{X}^f(\mathbf{X}^f)^T = \mathbf{P}_e^f$$

$$\boxed{\mathbf{X}^a = \mathbf{X}^f \mathbf{W}}$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Define

$$\mathbf{D} = \mathbf{W}\mathbf{W}^T$$

$\mathbf{W}$ is the *matrix square root* of $\mathbf{D}$

Choose

$$\mathbf{D} = (\mathbf{I} + (\mathbf{X}^f)^T \mathbf{H}^T (\mathbf{R}^o)^{-1} \mathbf{H}\mathbf{X}^f)^{-1}$$

Approach 2b

Ensemble Transform Kalman Filter (ETKF)

Analysis Ensemble

$$\mathbf{X}^a = \mathbf{X}^f \mathbf{W}$$

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{X}^f \mathbf{D} (\mathbf{X}^f)^T \mathbf{H}^T (\mathbf{R}^\circ)^{-1} (\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^f)$$

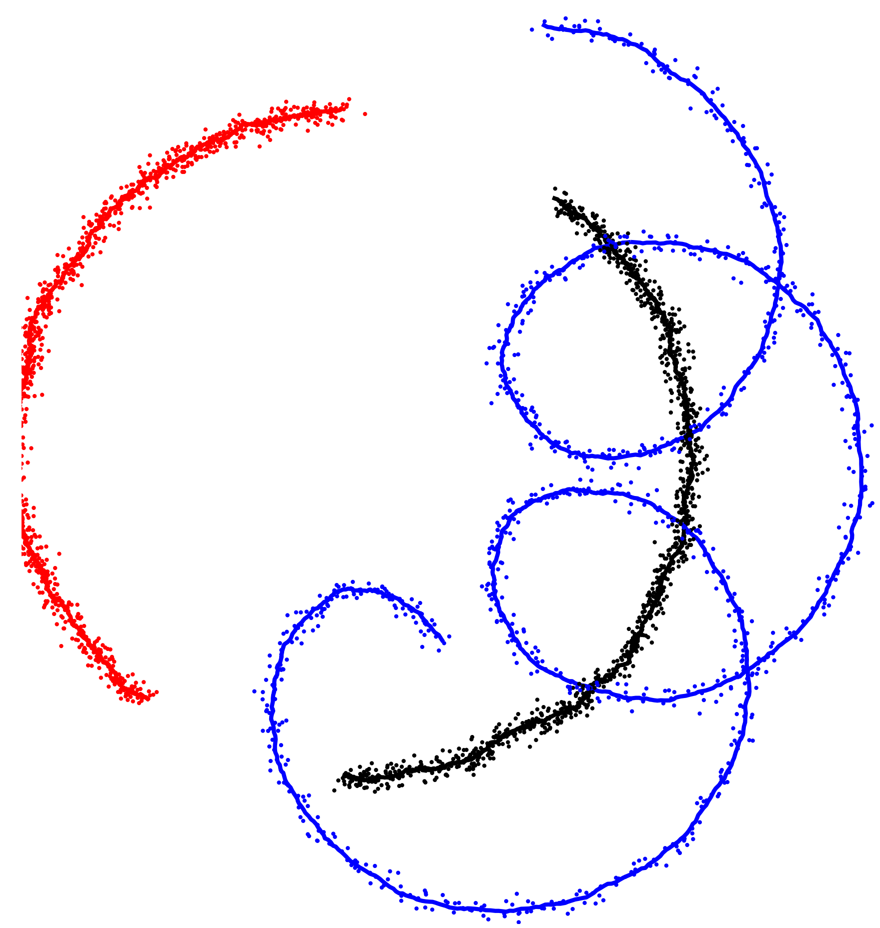
Approach 2

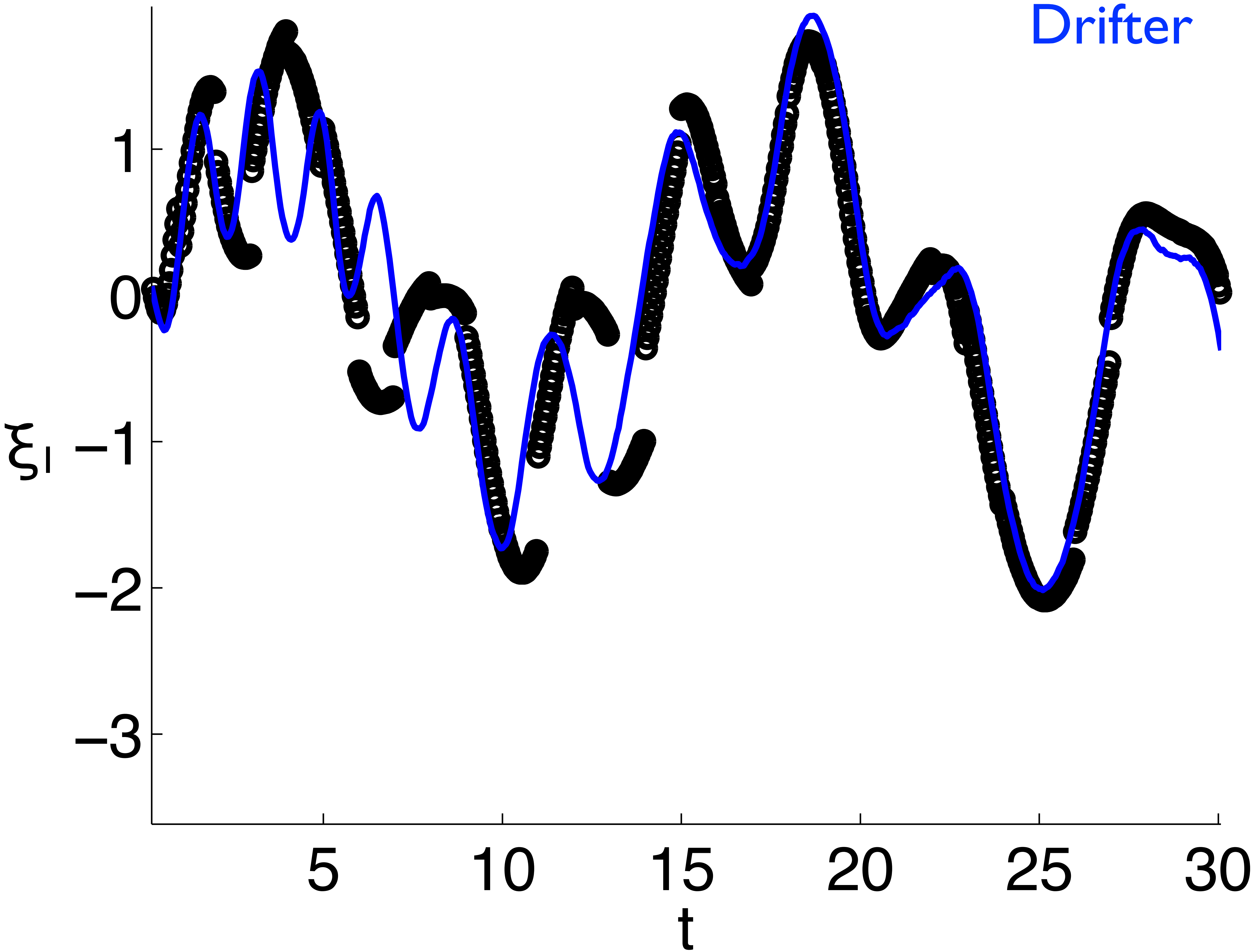
Ensemble Kalman Filter (EnKF)

Validation - Full Observation, Imperfect Data

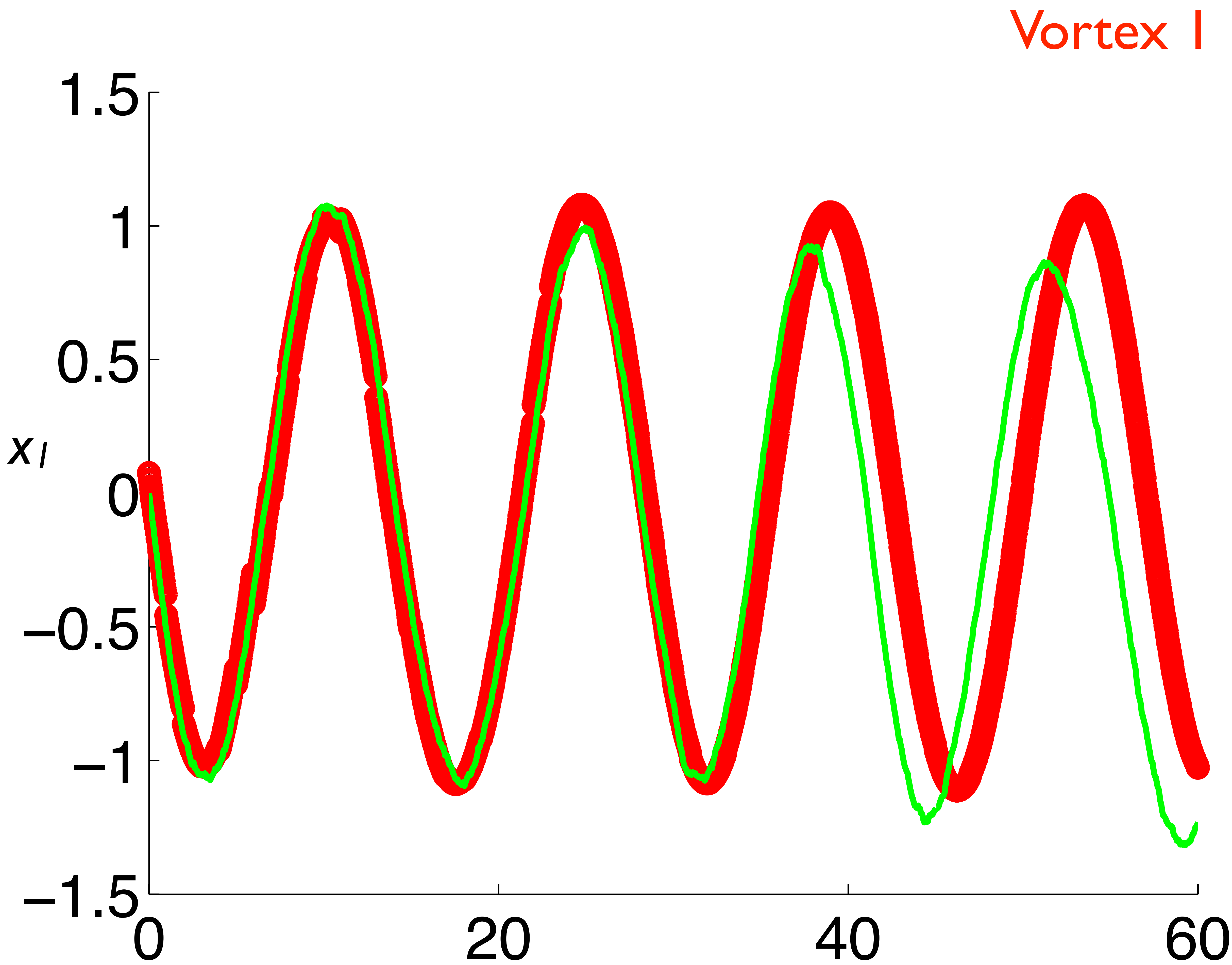
Observe vortices and drifter under observational noise

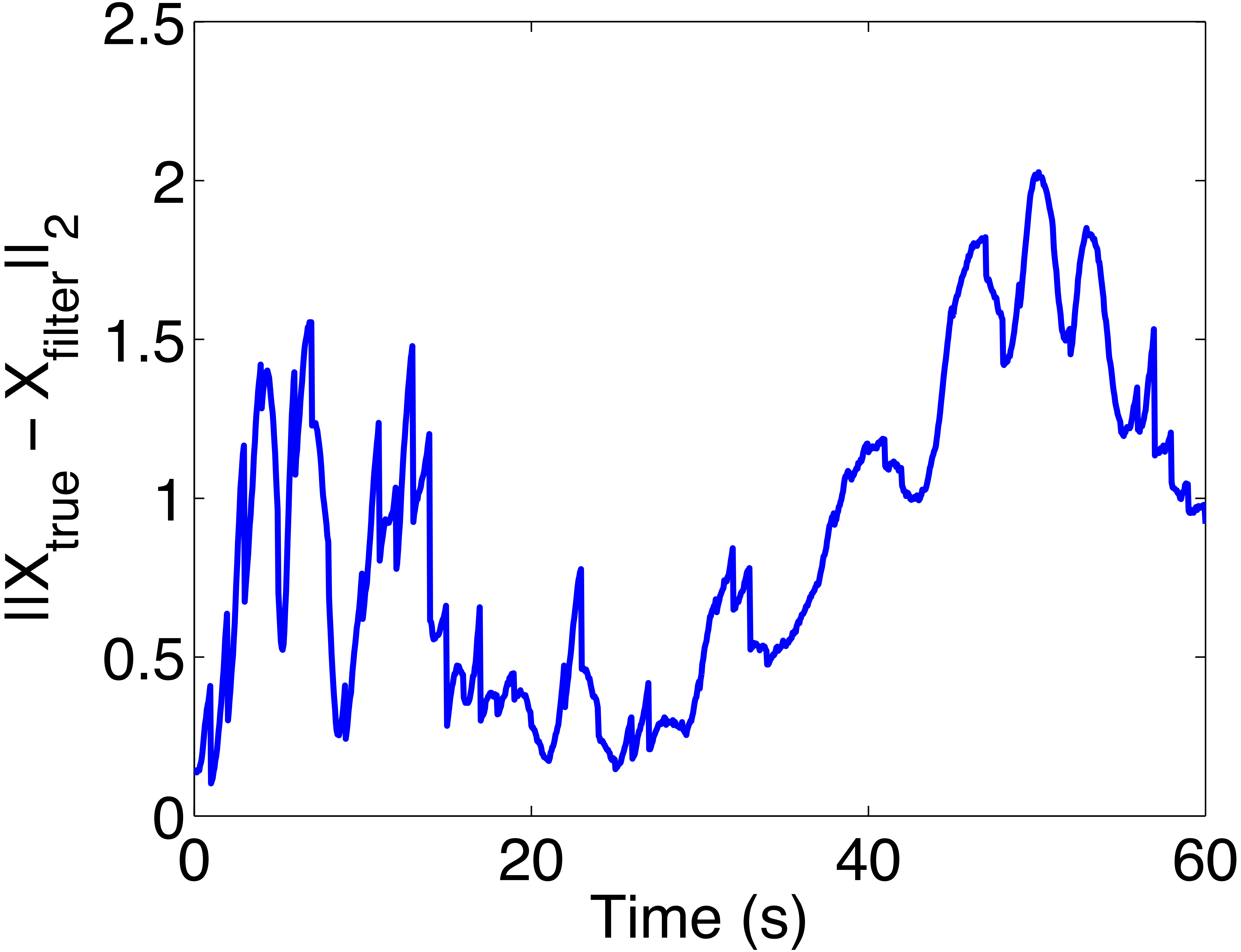
N = 6 ensemble members





Drifter
1
0
−1
−2
−3
ξ_1
5
10
15
20
25
30
t


Vortex 1
1.5
1
0.5
x_l
0
−0.5
−1
−1.5
0
20
40
60



2.5
2
1.5
1
0.5
0
0
20
40
60
Time (s)
||X_true - X_filter||_2

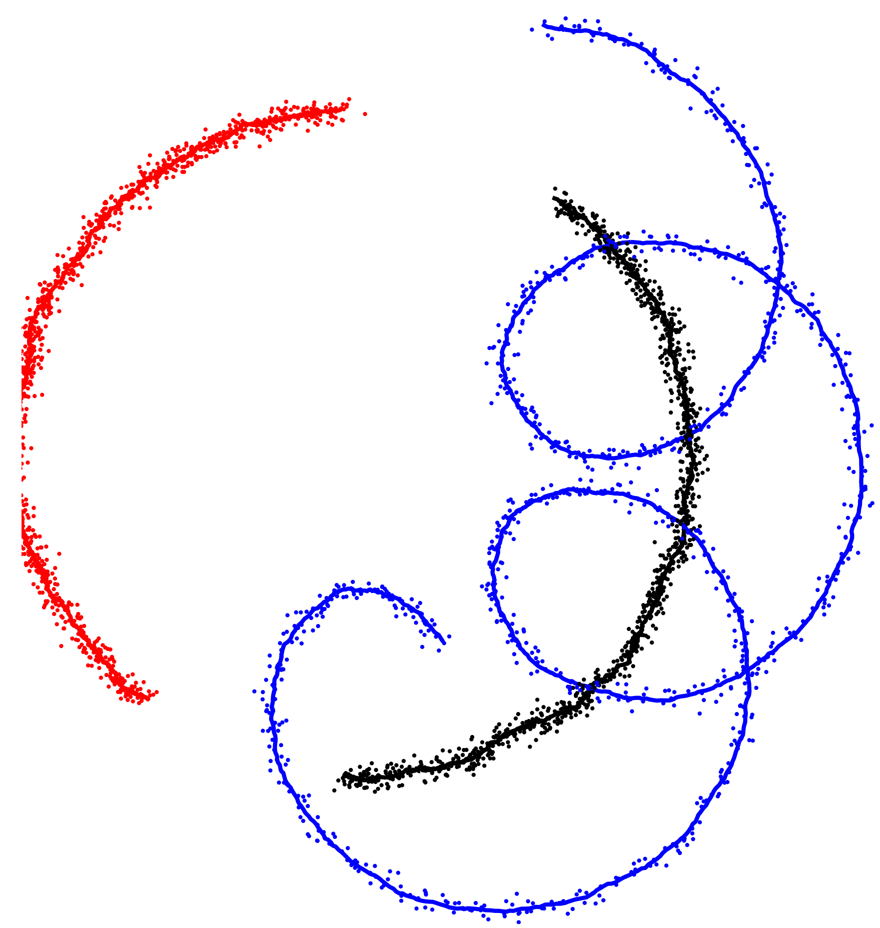
Approach 2

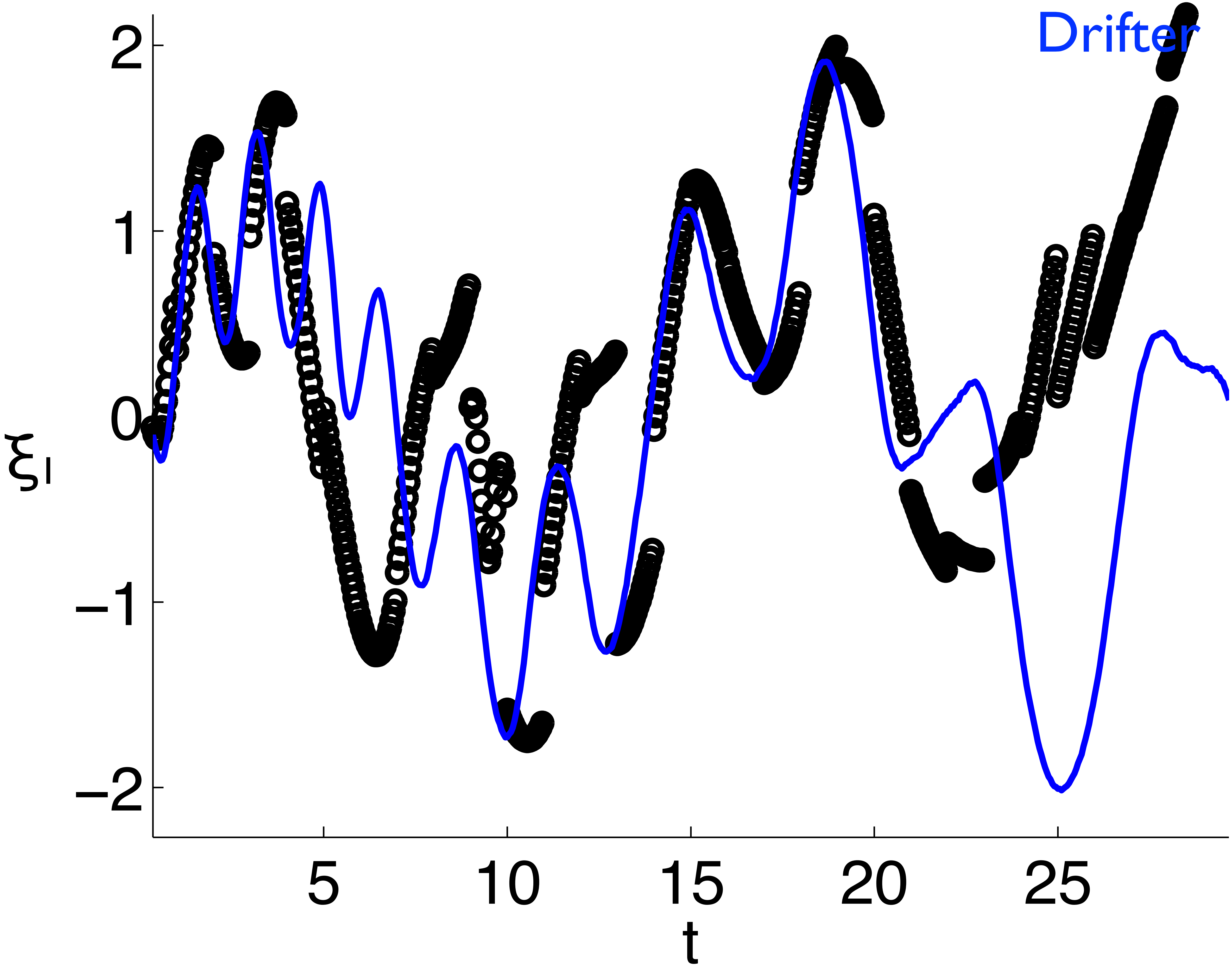
Ensemble Kalman Filter (EnKF)

Results - Partial Observation, Imperfect Data

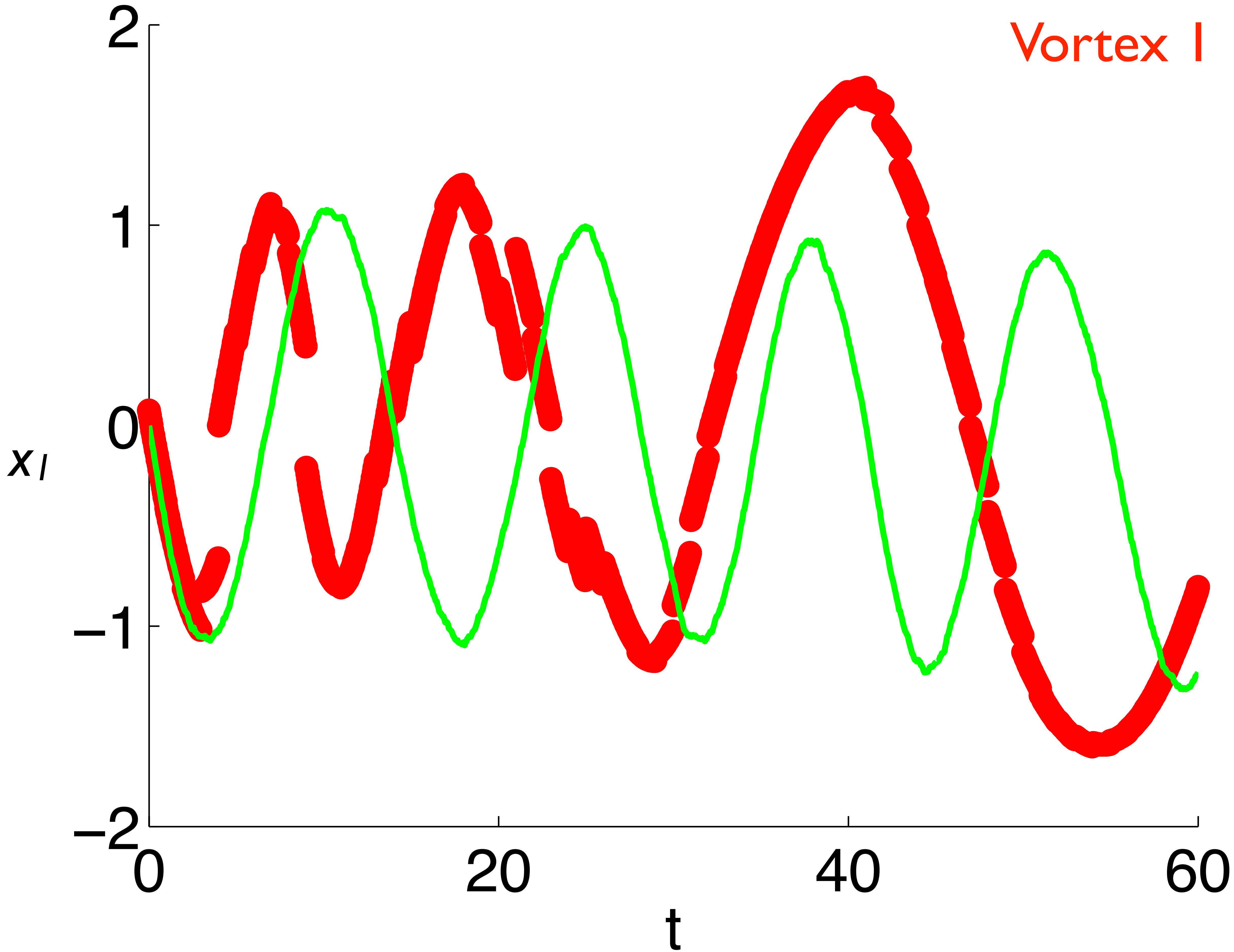
Observe only drifter under observational noise

N = 6 ensemble members

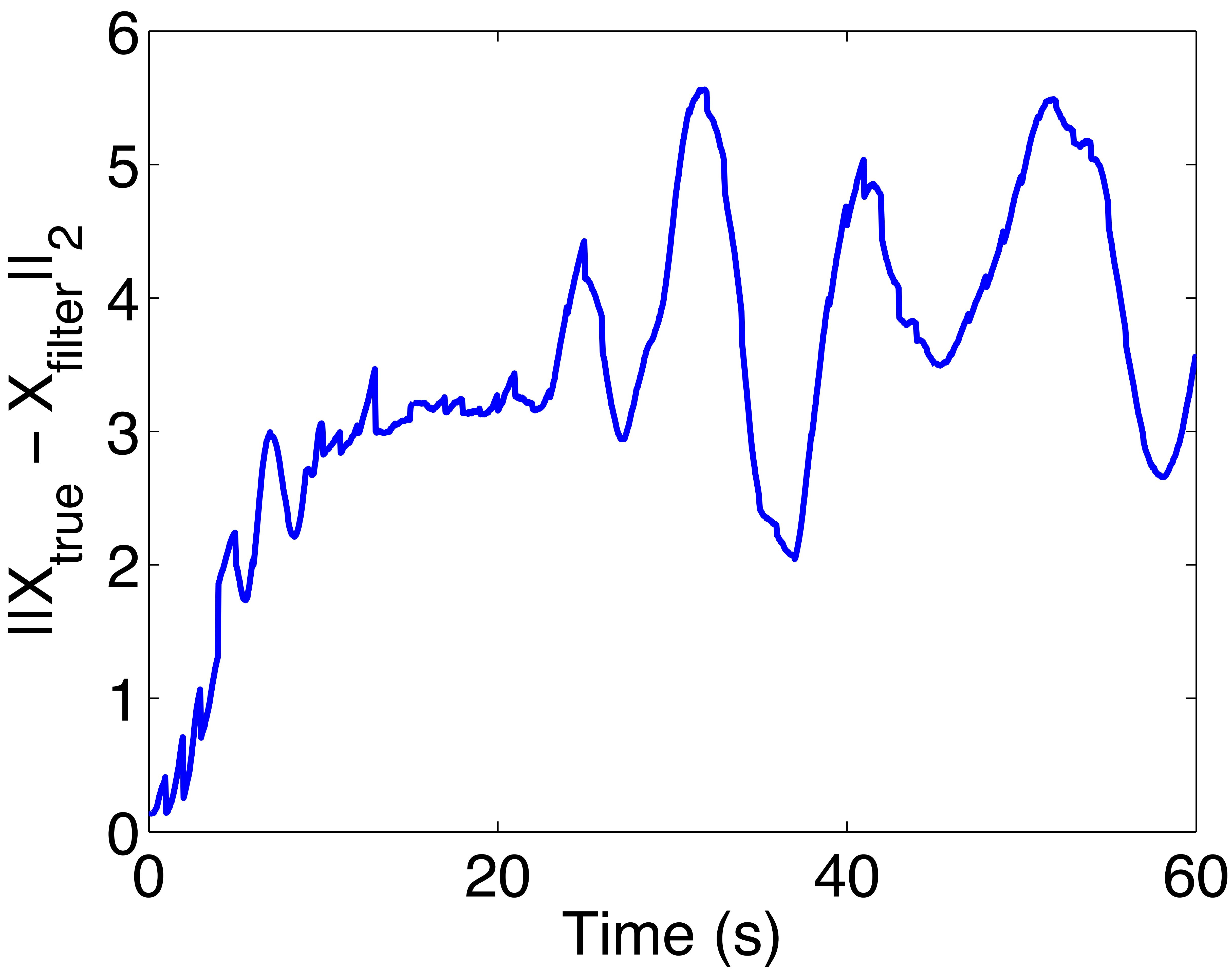




Drifter
2
1
0
−1
−2
ξ_l
5
10
15
20
25
t



Vortex I
2
1
0
−1
−2
x_I
0
20
40
60
t



6
5
4
3
2
1
0
0
20
40
60
Time (s)
$||X_{true} - X_{filter}||_2$

Data Assimilation

Testing

Testing, Phase I

Failure statistic: time to failure

Compare EKF, perturbed observations EnKF and ETKF

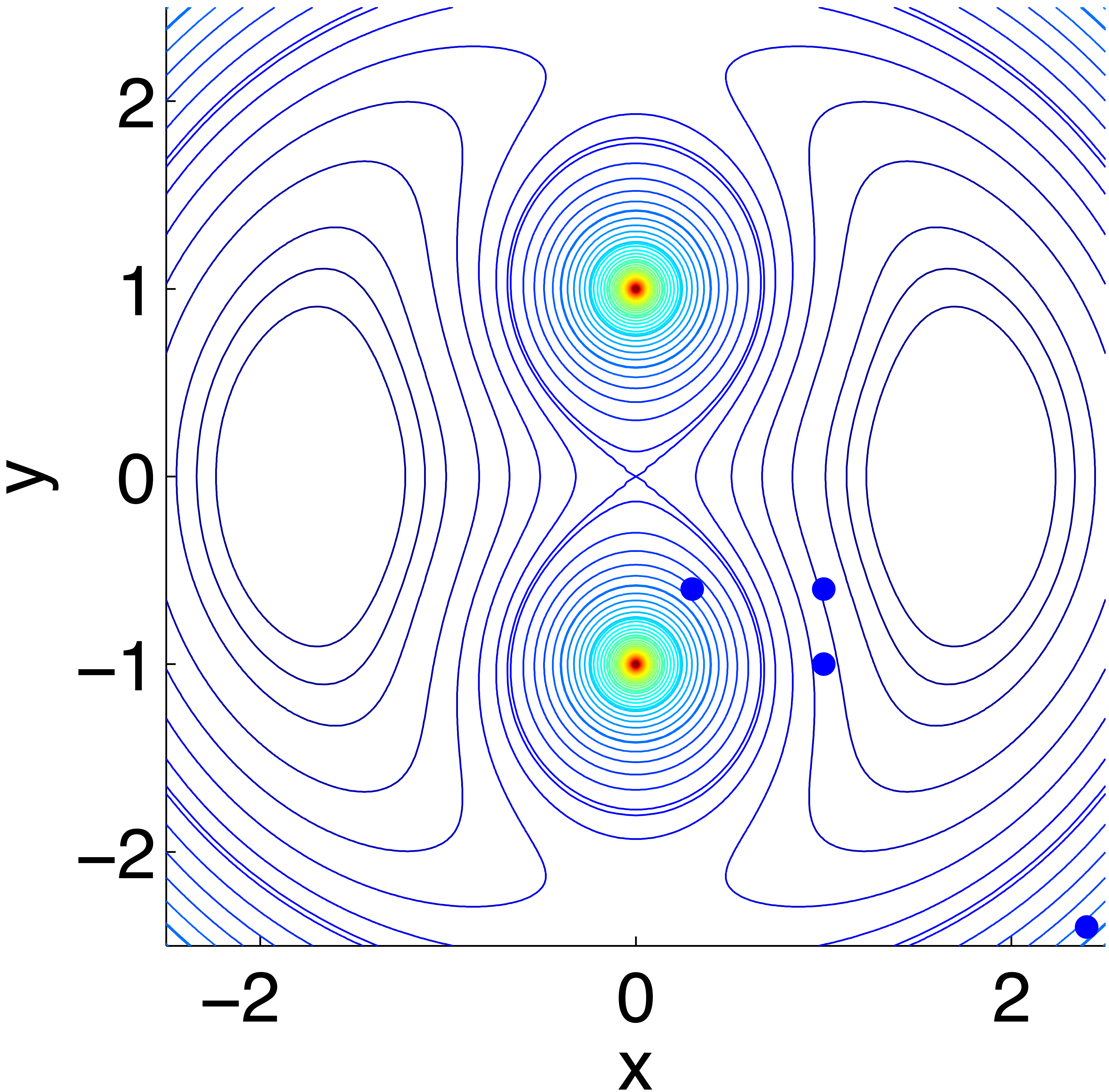
Testing, Phase 1

Failure statistic: time to failure

Generate M = 500 instances of the SDE at each of L = 4 drifter locations

Start vortex 1 at (0, 1)
Start vortex 2 at (0, -1)

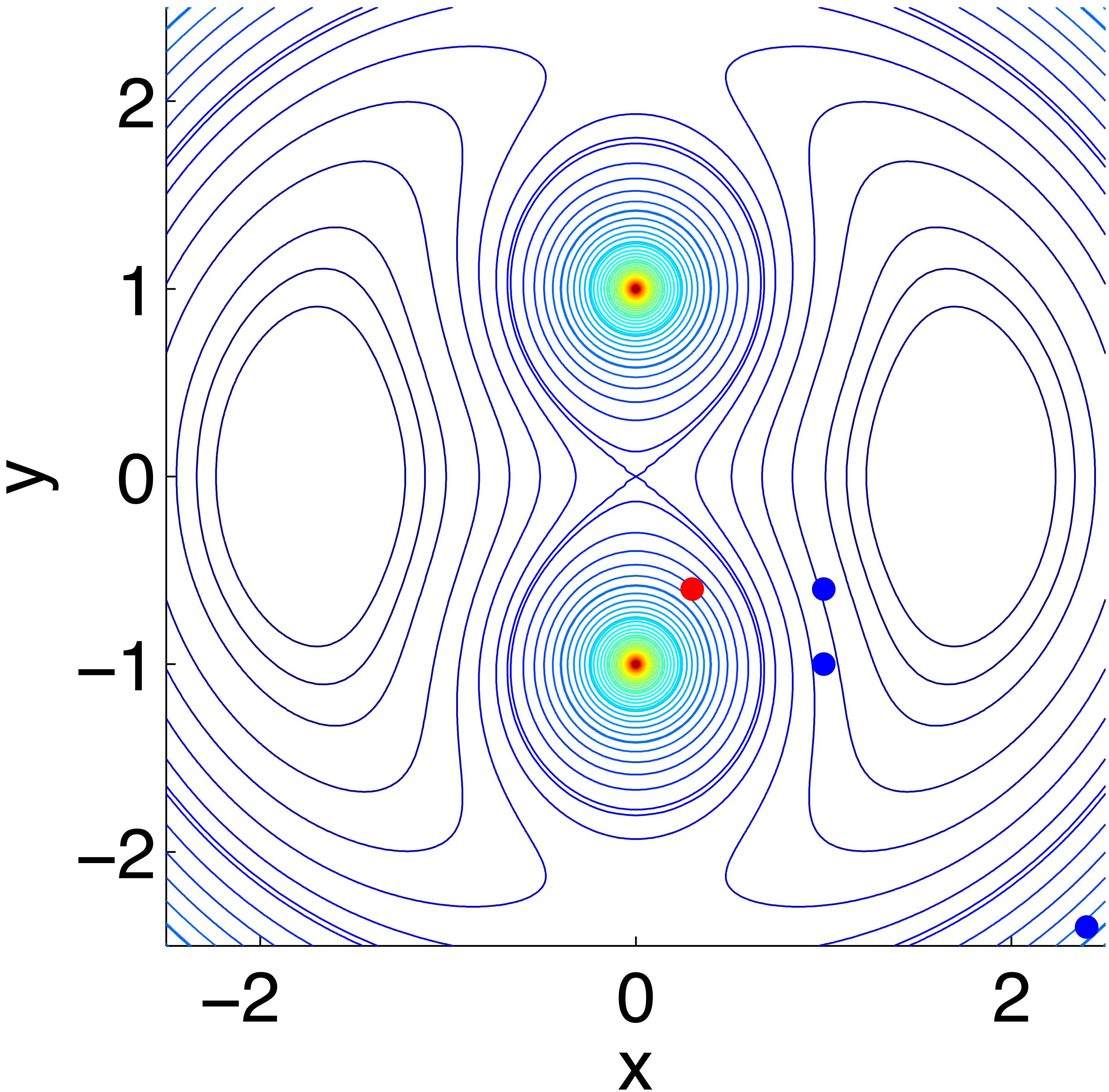
Add observational noise according to the model



2
1
0
−1
−2
y
−2
0
2
x

Testing, Phase I

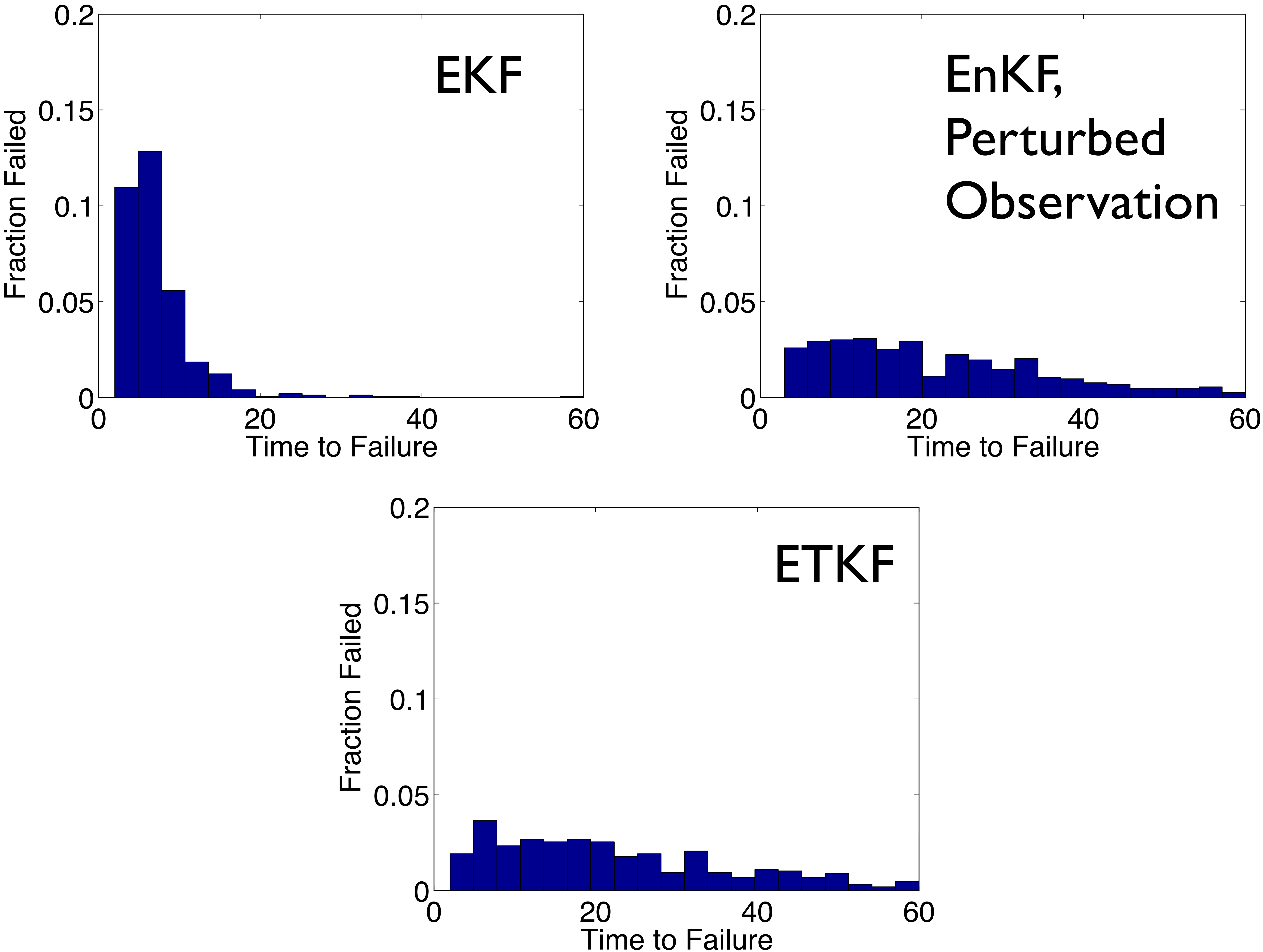
- Failure statistic: time to failure
  - Record when the distance between the analyzed state and the true state for either vortex is greater than 1


2
1
0
−1
−2
y
−2
0
2
x

Testing, Phase I

Failure statistic: time to failure

	EKF	EnKF, Pert	ETKF
Mean	8.23	25.64	25.27
Stdev	9.53	17.17	17.24
Fraction Completed	0.024	0.094	0.084



EKF
EnKF,
Perturbed
Observation
ETKF
Fraction Failed
Time to Failure
0
0.05
0.1
0.15
0.2
0
20
40
60

Moving Forward

Improvements to EnKF

- Covariance Inflation
- Localization
  - Local Ensemble Transform Kalman Filter (LETKF)

Moving Forward

Implement Particle Filter

Approximate pdf by an ensemble of *weighted* particles:

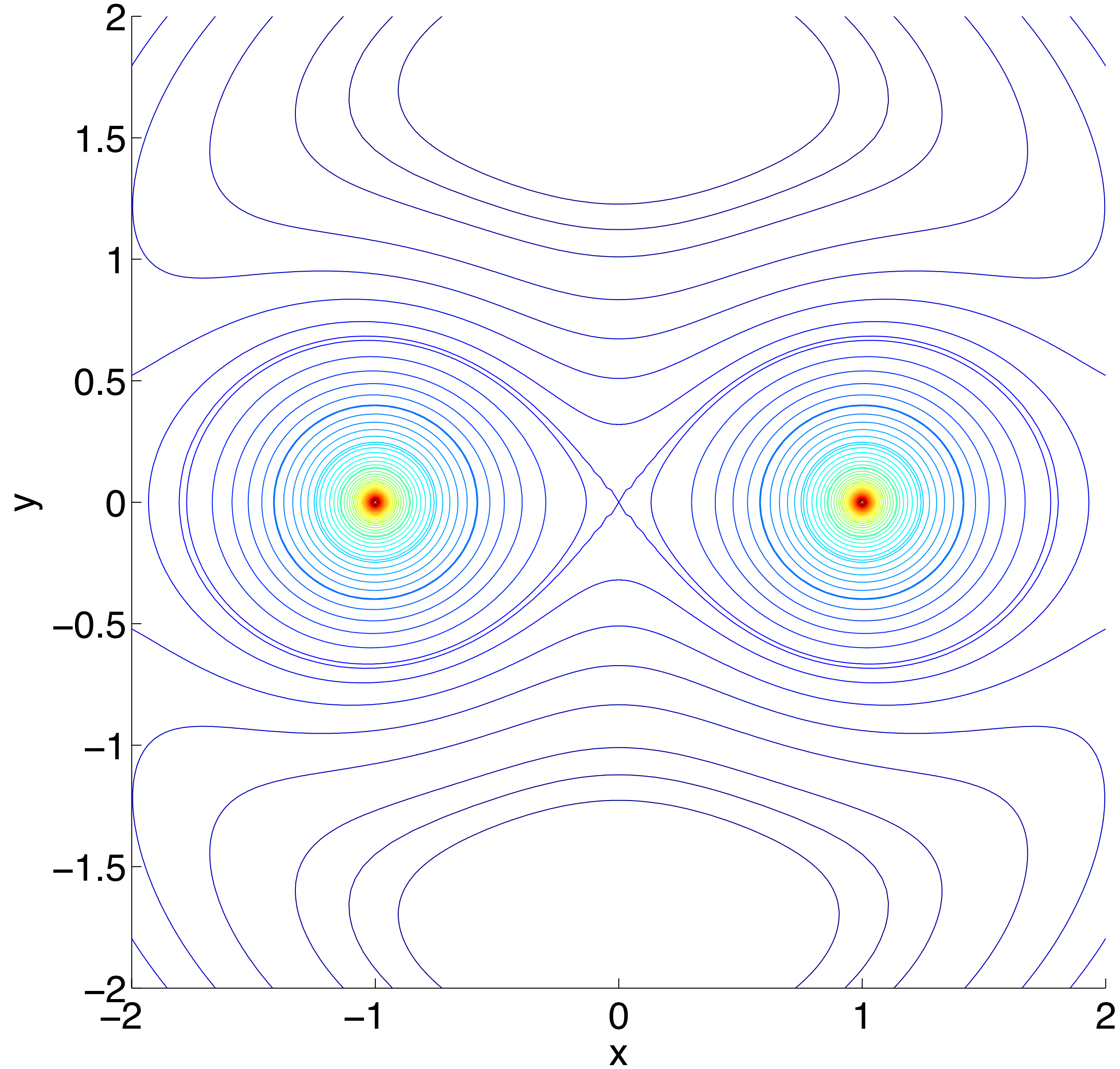
$$\{\mathbf{x}_i^f(t)\}_{i=1}^N \quad \{w_{i,k}\}_{i=1}^N$$

Forecast: Evolve particles forward using SDE

Analysis: Perform full Bayesian update at analysis

Moving Forward

Phase II - Manifold Detection for Observing System Design



Timeline

Phase I

- Produce database: now through mid-October
- Develop extended Kalman Filter: now through mid-October
- Develop ensemble Kalman Filter: mid-October through mid-November
- Develop particle filter: mid-November through end of January
- Validation and testing of three filters (serial): Beginning in mid-October, complete by February

Timeline

Phase I

- Produce database: now through mid-October
- Develop extended Kalman Filter: now through mid-October
- Develop ensemble Kalman Filter: mid-October through mid-November
- Develop particle filter: mid-November through end of January
- Validation and testing of three filters (serial): Beginning in mid-October, complete by February

(Some) References

D.J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM review*, pages 525–546, 2001.

A.H. Jazwinski. *Stochastic processes and filtering theory*. Dover Publications, 2007.

E. Kalnay. *Atmospheric modeling, data assimilation, and predictability*. Cambridge University Press, 2003.

G. Evensen. *Data assimilation: the ensemble Kalman filter.* Springer Verlag, 2009.

T. DelSole and M. Tippett. *The ensemble square root filter*. Accessed: http://mason.gmu.edu/%7Etdelsole/AdvStat/

Questions???