INTERPOLATING IRREGULARLY SPACED OBSERVATIONS FOR FILTERING TURBULENT COMPLEX SYSTEMS

JOHN HARLIM

Abstract. We present a numerically fast reduced filtering strategy, the Fourier domain Kalman filter with appropriate interpolations to account for irregularly spaced observations of complex turbulent signals. The design of such a reduced filter involves: (i) interpolating irregularly spaced observations to the model regularly spaced grid points, (ii) understanding under which situation the small scale oscillatory artifact from such interpolation won’t degrade the filtered solutions, (iii) understanding when the interpolated covariance structure can be approximated by its diagonal terms when observations are corrupted by independent Gaussian noise, and (iv) applying a scalar Kalman filter formula on each Fourier component independently with an approximate diagonal interpolated covariance matrix.

From the practical point of view, there is an emerging need to understand the effect of (i) toward the filtered solutions, for e.g., in utilizing the data produced from various interpolation techniques that merge multiple satellite measurements of atmospheric and ocean dynamical quantities. To understand point (iii) above and to see how much of non-diagonal terms are effectively ignored in (iv), we compute a ratio $\Lambda$ between the largest non-diagonal components and the smallest diagonal components of the interpolated covariance matrix. We find that for piecewise linear interpolation, this ratio, $\Lambda$, is always smaller than that of the alternative interpolation schemes such as trigonometric and cubic spline for any irregularly spaced observation networks. When observations are not so sparse with small noise, we find that the small scale oscillatory artifact in (ii) above is negligible when piecewise linear interpolation is used whereas for the other schemes such as the nearest neighbor, trigonometric, and cubic spline interpolation, the oscillatory artifact degrades the filtered solutions significantly. Finally, we also find that the reduced filtering strategy with piecewise linear interpolation produces more accurate filtered solutions than conventional approaches when observations are extremely irregularly spaced (such that the ratio $\Lambda$ is not so small) and very sparse.

Key words. interpolations, irregularly spaced observation, Kalman filter, data assimilation.

AMS subject classifications. 93E11, 62M20, 62L12, 65C20

1. Introduction. Given noisy observations from nature, filtering is the process of finding the best statistical estimate of the true signal. When observations are sampled at discrete time, the filtering problem consists of a two-step predictor-corrector scheme that adjusts a prior forecast distribution to be more consistent with the current observations. The revised probability distribution is then fed into the model as an initial condition for the future time prediction. This approach of generating initial conditions is also known as data assimilation.

In practice, the demand for operationally practical filtering methods escalates as the model resolution is significantly increased. For example, in the coupled atmosphere-ocean system, the current operational models for weather and climate predictions are discretized in space and time and the effects of unresolved processes are parametrized according to various recipes. The resulting dynamics of such models are extremely unstable and chaotic with several billion degrees of freedom. They typically have many spatiotemporal scales, rough turbulent energy spectra in the solutions near the mesh scale, and a very large dimensional state space.

One attractive filtering approach is to use ensemble based strategy [2, 3, 4, 11, 36, 23, 15, 16] but this approach relies on a large computational overhead in gen-

*The author is partially supported by the NC State University startup grant and the Office of Naval Research Grant N00014-11-1-0310.
†Department of Mathematics, North Carolina State University, BOX 8205, Raleigh, North Carolina, 27695-8205 (jharlim@ncsu.edu).
erating individual ensemble members through the forward dynamical operator [21]. Furthermore, all these finite ensemble methods are very sensitive to model resolution, observation frequency, and the nature of the turbulent signals when a practical limited ensemble size (typically less than 100) is used. Alternatively, the author and his collaborators have developed a class of fast and accurate Fourier based diagonal filtering strategies for turbulent complex systems [5, 18] and implemented them to filter spatiotemporal chaos [17, 19, 20] with model errors. In their recent development, they also advocated a systematic stochastic parameterization strategy to estimate model errors “on-the-fly” [13, 12]. However, all these diagonal Fourier based reduced filtering techniques are limited to regularly spaced observations (see the review article [30]) that are often not available in many applications. As an example, in the climate and weather prediction there is an emerging need to account for the recently available satellite data especially to uncover the oceanic region where observations are sparse. In this application, various interpolation techniques are typically used to merge multiple satellite data set to reconstruct high resolution estimate for ocean circulation [9, 27]. Therefore it is valuable to understand the effect of various interpolation techniques of the noisy data on the filtered solutions.

In this paper, we use a canonical toy model for turbulent complex systems to systematically clarify the effect of classical interpolation schemes on the filtered solutions. Secondly, we study the potential of using an approximate reduced filtering technique analogues to the diagonal Fourier based filter [30] with appropriate interpolation strategy to filter irregularly spaced observed complex turbulent systems. In particular, we will check the filtered solutions resulting from interpolating irregularly spaced observations to the model grid points with nearest neighbor, piecewise linear, cubic spline interpolations, and compare these approaches with the conventional strategies that interpolate the model state variable into the observation state space with linear interpolation and trigonometric interpolations [2, 29]. The design of such a reduced strategy requires rigorous understanding of the interpolated error covariance matrix when observations are corrupted by independent Gaussian noise. We will consider the spatially correlated noise in a future report. In particular, we will compute a ratio $\Lambda$ between the largest non-diagonal components and the smallest diagonal components of the interpolated covariance matrix and use it to determine how much of cross covariance terms are effectively ignored when a diagonal Fourier based filter as in [30] is used only with the diagonal components of the interpolated covariance for various observation networks and interpolation techniques. Secondly, when observations are not so sparse, the proposed interpolation strategy will introduce artificial small scale oscillations in the filtered solutions. Therefore, it is important to understand for which interpolation scheme and under which regime this oscillatory artifact won’t damage the solutions. Finally, we need to check how robust is the approximate filter when model errors are significant and when observations are extremely irregularly spaced and very sparse.

The paper is organized as follows. In section 2, we briefly review a simple prototype model for complex turbulent systems that we use for testing our filter and the reduced filters in [5, 18, 30]. In Section 3, we discuss the effect of various interpolation schemes on spatiotemporally independent error covariance matrix. We show numerical results in Section 4 and provide a short summary in Section 5.

2. Test model for filtering complex turbulent systems. In this paper, we consider a simple stochastic model for turbulent fluctuations with linear dissipation and white noise forcing to mimic rapid energy transfer due to nonlinear interaction
Interpolating Irregularly Spaced Observations

As a canonical example, the turbulent true signals which will be filtered are determined by solutions of an $s$-dimensional stochastically forced PDE

$$\frac{\partial u(x, t)}{\partial t} = P(\partial_x) u(x, t) - \gamma(\partial_x) u(x, t) + f(x, t) + \sigma(x) \dot{W}(t),$$  \hfill (2.1)

where $u(x, t) \in \mathbb{R}^s$ represents any vector field of interest. For simplicity, we will only consider scalar $u$ on a one-dimensional spatial coordinate $x$. One can extend the results below to two- or three-dimensional spatial coordinates. Here the initial data $u_0(x)$ is a Gaussian random field with nonzero covariance. In (2.1), the term $P(\partial_x) u$ represents the dispersion relation that is typically obtained by linearizing the nonlinear dynamical operator around a constant background field. The external forcing $f(x, t)$ is assumed to be known, damping term $-\gamma(\partial_x) u$ and spatially correlated noise,

$$\sigma(x) \dot{W}(t) \equiv \sum_{k = -\infty}^{\infty} \sigma_k \dot{W}_k(t) e^{ikx},$$  \hfill (2.2)

are added to represent small scale unresolved turbulent fluctuations and nonlinear interaction. In (2.2), $\sigma_k \in \mathbb{R}$ is the noise strength of Fourier mode $k$ and $W_k(t)$ is the corresponding complex Wiener process with independent real and imaginary components with variance $t/2$. We non-dimensionalize (2.1) on a $2\pi$-periodic domain such that infinite Fourier series,

$$u(x, t) = \sum_{k = -\infty}^{\infty} \hat{u}_k(t) e^{ikx}, \quad \hat{u}_k = \hat{u}_k^*, \quad \hat{u}_0 \in \mathbb{R}$$  \hfill (2.3)

where $\hat{u}_k(t) \in \mathbb{C}$ for $k > 0$ can be utilized in analyzing (2.1) and the related finite approximations.

The operators $P(\partial_x)$ and $\gamma(\partial_x)$ are defined through unique symbols at a given wavenumber $k$ by

$$P(\partial_x) e^{ikx} = \tilde{p}(ik) e^{ikx}$$  \hfill (2.4)

$$\gamma(\partial_x) e^{ikx} = \gamma(ik) u^{ik},$$  \hfill (2.5)

where $\tilde{p}(ik)$ and $\gamma(ik)$ denote the eigen-solutions of operators $P(\partial_x)$ and $\gamma(\partial_x)$, respectively. In general, $P(\partial_x)$ can be any differential operator which is a combination of odd derivatives while $\gamma(\partial_x)$ is a suitable combination of even derivative with positive $\gamma(ik)$ (see Chapter 1 of [34] for various examples of damping operators). Substituting (2.2), (2.3)-(2.5) into initial value problem in (2.1), we reduce our problem to solving the following uncoupled forced Langevin equations on each Fourier mode

$$d\hat{u}_k(t) = [\tilde{p}(ik) - \gamma(ik)] \hat{u}_k(t) dt + \hat{f}_k(t) dt + \sigma_k dW_k, \quad \hat{u}_k(0) = \hat{u}_{k,0},$$  \hfill (2.6)

where $\hat{f}_k(t)$ is the Fourier coefficient of the deterministic forcing, $f(x, t)$. Note that although this model does not have energy transfer between Fourier modes, the author and his collaborators have shown that this simple diagonal model produces accurate filtered solutions when it is used for generating prior estimates in turbulent regimes [17, 19, 20, 30]. This procedure is also shown to be promising for estimating poleward eddy heat transport through satellite altimetry [26].
We assume that $\tilde{p}(ik) = i\omega_k$ is wave-like solution with $-\omega_k$ the real valued dispersion relation. In our numerical simulation below, we choose linear wave dispersion, $\omega_k = k$, for simplicity (see [13, 12] where $\omega_k = 1/k$ is chosen to mimic the Rossby waves linear dispersion in a one-dimensional periodic domain). In this simple context, the statistical equilibrium distribution for (2.6) exists provided that $\tilde{f}_k = 0$ and the damping is non-negative,

$$\gamma(ik) > 0 \text{ for all } k \neq 0.$$  \hspace{1cm} (2.7)

Under these assumptions, the statistical equilibrium for (2.6) is Gaussian with mean zero and variance (or energy spectrum)

$$E_k = \frac{\sigma_k^2}{2\gamma(ik)}, \quad 1 \leq k < +\infty.$$  

Mathematically, one needs to require $\sum E_k < \infty$ to define the stochastic solution of (2.1) correctly with a similar requirement on the Gaussian initial data in $u_0(x)$. However, there is genuine physical interest in situations with an even rougher turbulent spectrum such as white noise where $E_k$ is constant or equipartition. Under this situation and in many standard discretizations, we truncate the sum in (2.3) to the smallest meshsize. In the numerical experiment in Section 4, we truncate up to mode $|k| \leq N = 220$ with decaying spectrum $E_k = k^{-3}$. In other words, we discretize (or resolve) the SPDE in (2.1) at $2N + 1$ equally spaced grid points.

In this paper, we consider non-steady $\gamma_k \equiv \gamma(ik)$ [13, 12], that is, we allow the damping coefficients $\{\gamma_k\}_{|k|\leq 20}$ to alternate between stable and unstable regimes. The stable regime is characterized by positive damping and corresponds to the loss of energy whereas the unstable regime is characterized by negative damping and corresponds to the gain of energy. Numerically, we allow the first 10 modes to alternate between a weaker damping $\gamma_w$ in the stable regime and stronger damping $\gamma_s$ in the unstable regime and the next 10 modes to alternate between a positive damping $\gamma^+$ in the stable regime and negative damping $\gamma^-$ in the unstable regime (see Table 2.1 for the damping strength).

<table>
<thead>
<tr>
<th>Table 2.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Damping coefficients for simulating the true signal</strong></td>
</tr>
<tr>
<td><strong>modes</strong>&lt;br&gt;</td>
</tr>
<tr>
<td>1-10</td>
</tr>
<tr>
<td>11-20</td>
</tr>
<tr>
<td>21-220</td>
</tr>
</tbody>
</table>

As in [13, 12], we consider an exponentially distributed waiting time between these two regimes. That is, the time the system spends in the stable regime before it switches to the unstable regime is a random variable $T_{st}$ with cumulative distribution function given by

$$P(T_{st} < \tau) = 1 - e^{-\nu \tau},$$

where $\nu$ is a rate of change from stable to unstable regime. Similarly, for the random time $T_{un}$ that the system spends in the unstable regime, we have

$$P(T_{un} < \tau) = 1 - e^{-\mu \tau},$$
where $\mu$ is a rate of change from unstable to stable regime. In our numerical model, we set $\nu = 0.1$ and $\mu = 0.2$ such that the system spends on average $2/3$ of the time in the stable regime and $1/3$ of the time in the unstable regime. The choice of damping coefficients in Table 2 is such that the average damping is constant $\bar{\gamma} = 1.5$ for each mode,

$$\bar{\gamma} = \frac{\nu \gamma^- + \mu \gamma^+}{\nu + \mu} = \frac{\nu \gamma_w + \mu \gamma_s}{\nu + \mu} = 1.5.$$  

Thus, the equilibrium distribution still exists and it is Gaussian with energy spectrum $E_k = \sigma_k^2/2\bar{\gamma}$. The remaining modes, $|k| > 20$, are assumed to have a constant damping $\gamma_k = 1.5$ in both regimes. With this damping coefficient, we choose a reasonably long observation time $\Delta t = 1/2$ yet shorter than the system correlation time, $\gamma_k^{-1} = 2/3$.

We also assume a resonance periodic forcing to mimic extreme events as in [18] for the first 20 modes. That is, we set $\hat{f}_k = \exp(-i\omega_k t)$. This nonzero forcing will be a stringent test to the filter on the modes that are not fully observed (see [18] where practical observability condition is important when the system has nonzero forcing and also [25, 1, 6] where observability condition is absolutely required for filter stability when the filter model is not stable). The true spectrum of this model is shown in Figure 2.1 (see the solid thick line). Here, we see a large excursion from $k^{-3}$ line on the intermediate modes which characterizes the instability transition.

In previous work [5, 18], we considered two special observation networks:

**2.1. Plentiful Observations.** Consider noisy observations, $v_{j,m} = u(x_j, t_m) + \sigma_{j,m}$, at every model grid point $x_j$ and discrete time $t_m$. That is, in physical space we have $2M + 1$ observations and they coincide with the $2N + 1$ regularly spaced model grid points, i.e., $M = N$. Here, the noise $\sigma_{j,m}$ is spatially and temporally uncorrelated Gaussian with mean zero and variance $r^2$. This observation model can be written in Fourier space as follows:

$$\hat{v}_{k,m} = \hat{u}_{k,m} + \hat{\sigma}_{k,m}, \quad |k| \leq N,$$  

(2.8)
where \( \hat{v}_{k,m} \) denotes the \( k \)th Fourier coefficient of observations \( \{v_{j,m}\}_{j=1}^{2M+1} \); \( \hat{\sigma}_o^{\ell,m} \) denotes the \( k \)th Fourier coefficient of noise \( \sigma_o^{\ell,m} \). Note that \( \hat{\sigma}_o^{\ell,m} \) is still Gaussian and independent with mean zero and variance \( r_o/(2N+1) \) in the frequency space \([29, 5]\). As it is advocated in \([5]\), we can therefore apply the Kalman filter formula on each Fourier mode and solve \( N \)-independent complex valued scalar filtering problems in \((2.6)\) and \((2.8)\), as opposed to solving the \((2N+1)\)-dimensional filtering problem in \((2.1)\) the corresponding observation in the real space. The latter involves propagating and inverting \((2N+1)\times(2N+1)\) covariance matrices, see \([1, 6]\), which is numerically expensive when \( N \) is large. Readers that are not familiar with Kalman filter can consult \([1, 6]\).

### 2.2. Sparse Regularly Spaced Observations

Sparse observations are available at every \( P = (2N+1)/(2M+1) \) model grid points, where \( N > M \). Here, the Fourier coefficient \( \hat{v}_{\ell,m} \) of observations couples Fourier modes in aliasing set \( \mathcal{A}(\ell) = \{ k : |k| \leq N, k = \ell + (2M+1)q, q \in \mathbb{Z} \} \) as follows

\[
\hat{v}_{\ell,m} = \sum_{k \in \mathcal{A}(\ell)} \hat{u}_{k,m} + \hat{\sigma}_o^{\ell,m}, \quad |\ell| \leq M. \tag{2.9}
\]

In \((2.9)\), the observation noise \( \hat{\sigma}_o^{\ell,m} \) stays Gaussian and independent with mean zero and variance \( r_o/(2M+1) \) \([29, 5]\). This is true as long as these sparse observations are regularly spaced. The reduced filtering strategy decouples Fourier modes belonging to different aliasing sets and consists of \( P \)-dimensional filtering problems, coupling the modes in each aliasing set through observations in \((2.9)\).

In the two special observation networks above, we found that ignoring error covariances between appropriate frequency modes not only suppress the computational cost but also improves the accuracy of the filtered solutions (see \([5, 17]\) for details). Unfortunately, such a reduced Fourier Domain Kalman filtering (FDKF) strategy is not available when observations are not regularly spaced. In the next section, we will devise an approximate FDKF that is analogue to the decoupled filters above with appropriate interpolated observations. In all numerical experiments below, we will only apply FDKF with constant damping \( \gamma_k = \bar{\gamma} = 1.5 \) and we commit model errors on modes 10-20 due to the intermittent stability transition. As we mentioned in the introduction, there is a more sophisticated nonlinear stochastic parameterization strategy that can track the stability transition on-the-fly that can also be utilized \([13, 12]\) for more accurate filtered solution but we omit it in this paper since our point here is to understand the effect of various interpolations.

### 3. Interpolating irregularly spaced observations

One approach for filtering irregularly spaced observation in Fourier domain is to use the trigonometric interpolation \([29]\) which naturally couples observations to the Fourier coefficients of the model state. This interpolation technique transforms a diagonal observation error covariance matrix \( r_o I \) into a fully correlated covariance matrix in Fourier space. To see this, consider observations

\[
\hat{v}_{\ell,m} = u(\tilde{x}_\ell, t_m) + \hat{\sigma}_o^{\ell,m}, \quad \ell = 0, \ldots, 2M, \tag{3.1}
\]

at irregularly spaced grid points \( \tilde{x}_\ell \) on a \( 2\pi \)-periodic domain with spatially and temporally uncorrelated (i.i.d) noises, \( \hat{\sigma}_o^{\ell,m} \sim \mathcal{N}(0, r_o) \). Basically, trigonometric interpolation solves linear problem \( Af = \tilde{y} \) for Fourier coefficients \( \tilde{f} = (\tilde{f}_0, \ldots, \tilde{f}_M)^T \). Here, matrix \( A \) has components \( A_{\ell,k} = (e^{ik\tilde{x}_\ell}) \) with \( |k| \leq M \) and components of vector \( \tilde{y} = (y(\tilde{x}_0), \ldots, y(\tilde{x}_{2M}))^T \).
are the irregularly spaced data points. This linear problem has a unique solution (see [24] for an explicit expression of $V \equiv A^{-1}$). Applying $V$ to (3.1), we can write
\[ \hat{v}_{k,m} = \hat{u}_{k,m} + \sum_{\ell} V_{k,\ell} \tilde{\sigma}_{\ell,m} = \hat{u}_{k,m} + \tilde{\sigma}_{k,m}, \quad |k| \leq M, \quad (3.2) \]
where the noise $\tilde{\sigma}_m = (\tilde{\sigma}_{o_{-M,m}}, \ldots, \tilde{\sigma}_{o_{M,m}})$ becomes correlated,
\[ \langle \tilde{\sigma}_m \tilde{\sigma}_m^* \rangle = r^o V V^*, \quad (3.3) \]
where $\langle \cdot \rangle$ denotes expectation. Notice that the computational advantage of the diagonal Fourier domain filters, discussed in Sections 2.1-2.2, diminishes since the filter is now fully coupled through the observation noise covariance matrix. Implementation-wise, we do not need to compute matrix $V$ since we can replace $u(\tilde{x}_\ell, t_m)$ in (3.1) with $A \hat{u}$ that maps all Fourier components of the model state to the irregularly spaced real valued observations.

In the data assimilation community, the state-of-the-art ensemble Kalman filters [2, 4, 10] interpolate the dynamical variables at the model grid points on the observation site, i.e.,
\[ \hat{v}_{\ell,m} = G(u(x_j, t_m)) + \tilde{\sigma}_{\ell,m}^o, \quad (3.4) \]
where operator $G$ is an interpolation function that maps the dynamical variables to the observation space and $x_j$ are the regularly spaced model grid points. In practice, linear interpolation is often used [2]. An even cruder approach is to move observations to the nearest model grid points [22]. The task of designing numerically fast filters as in [5, 18] becomes non-trivial because we need the Fourier coefficients at the model mesh (not at the observation mesh as in (3.4)) and inappropriate spatial interpolation introduces significant correlation between any two Fourier modes as we showed in (3.2). Later, we will numerically show that when trigonometric interpolation is used, the cross correlation is not negligible.

### 3.1. Piecewise Linear Interpolation.
In this section, we analyze the potential of using piecewise linear interpolation to interpolate observations to the model grid points (which is the opposite of the standard approach). In particular, we will study the analytical covariance structure resulting from this interpolation approach.

Let $x_j = jh \in [0, 2\pi), j = 0, \ldots, 2N, (2N+1)h = 2\pi$ be the regularly spaced model grid points and $\tilde{x}_\ell, \ell = 0, \ldots, 2M$ be the irregularly spaced observation locations in increasing ordered, i.e., $\tilde{x}_\ell < \tilde{x}_{\ell+1}$ for all $\ell$. With this configuration, as long as $N > 0$, every grid point $x_j$ is located in between two observation locations. In other words, there exists an $\ell_j$ such that the following inequality, either $\tilde{x}_{\ell_j} < x_j \leq \tilde{x}_{\ell_j+1}$ or $\tilde{x}_{\ell_j} \leq x_j < \tilde{x}_{\ell_j+1}$, is satisfied. Now, let’s partition $[0, 2\pi)$ into subintervals $[\tilde{x}_\ell, \tilde{x}_{\ell+1}]$, for $\ell = 0, \ldots, 2M$, where for the last subinterval, we set $\tilde{x}_{2M+1} = \tilde{x}_0 + 2\pi$ due to the periodicity. On each of these subintervals, let us consider the following linear interpolation:
\[ v(x) = \tilde{v}_{\ell_j} L_{\ell_j}(x) + \tilde{v}_{\ell_j+1} L_{\ell_j+1}(x), \]
where
\[ L_{\ell_j}(x) = \frac{x - \tilde{x}_{\ell_j+1}}{\tilde{x}_{\ell_j} - \tilde{x}_{\ell_j+1}}, \quad L_{\ell_j+1}(x) = 1 - L_{\ell_j}(x) \quad (3.5) \]
are the corresponding linear Lagrange functions. In the following, we denote \( L_x(x_j) \equiv C_j \) for simplicity and notice that \( 0 \leq C_j \leq 1 \). Thus, the linearly interpolated observations at the model grid points are

\[
v_j \equiv v(x_j) = C_j \hat{v}_{\ell_j} + (1 - C_j) \hat{v}_{\ell_j + 1} \equiv C_{j,1} \hat{v}_{\ell_j} + C_{j,2} \hat{v}_{\ell_j + 1},
\]

where \( C_{j,1} \equiv C_j \) and \( C_{j,2} \equiv 1 - C_j \) (thus \( C_{j,1} + C_{j,2} = 1 \)) with Fourier coefficients

\[
\hat{v}_k = \frac{h}{2\pi} \sum_{j=0}^{2N} v_j e^{-ikj} = \frac{h}{2\pi} \sum_{j=0}^{2N} (C_{j,1} \hat{v}_{\ell_j} + C_{j,2} \hat{v}_{\ell_j + 1}) e^{-ikj}.
\]

For an independent noise \( \tilde{x}_{\ell_j} \) at irregularly spaced \( \tilde{x}_{\ell_j} \) (we drop subscript \( 'm' \) that indicates time and superscript \( 'o' \) that indicates observation noise to simplify the notation) with variance \( r^o \), the \((k,k')\) Fourier component of the covariance matrix is

\[
\langle \hat{\sigma}_k^o \hat{\sigma}_k^{o*} \rangle = \left\langle \left( \frac{h}{2\pi} \sum_{j=0}^{2N} (C_{j,1} \tilde{\sigma}_{\ell_j} + C_{j,2} \tilde{\sigma}_{\ell_j + 1}) e^{-ikj} \right) \left( \frac{h}{2\pi} \sum_{j'=0}^{2N} (C_{j',1} \tilde{\sigma}_{\ell_{j'}} + C_{j',2} \tilde{\sigma}_{\ell_{j'} + 1}) e^{-ik'j} \right)^* \right\rangle
\]

\[
= \frac{h^2 r^o}{4\pi^2} \sum_{j=0}^{2N} \sum_{j'=0}^{2N} (C_{j,1} C_{j',1} \delta_{\ell_j, \ell_{j'}} + C_{j,1} C_{j',2} \delta_{\ell_j, \ell_{j'}+1}
\]

\[
+ C_{j,2} C_{j',1} \delta_{\ell_j+1, \ell_{j'}} + C_{j,2} C_{j',2} \delta_{\ell_j+1, \ell_{j'}+1} \bigg) e^{-i(k-j')h}
\]

\[
= \frac{r^o}{(2N+1)^2} \left[ \sum_{j=0}^{2N} (C_{j,1}^2 + C_{j,2}^2) e^{-i(k-j')h} + \sum_{\substack{j,j' \neq j' \neq \ell_j}} \left( C_{j,1} C_{j',2} \delta_{\ell_j, \ell_{j'}+1} + C_{j,2} C_{j',1} \delta_{\ell_j+1, \ell_{j'}} + C_{j,1} C_{j',1} \delta_{\ell_j, \ell_{j'+1}} + C_{j,2} C_{j',2} \delta_{\ell_j+1, \ell_{j'+1}} \bigg) e^{-i(k-j')h} \right].
\]

In the derivation above, we use the fact that the observation errors are spatially uncorrelated at different grid points. Notice that since \( 1/4 \leq a^2 + (1 - a)^2 \leq 1 \) for \( 0 \leq a \leq 1 \) and \( \{e^{ikj}\} \) is an orthonormal basis of vector space \( \mathbb{C}^{2N+1} \), the first term in the square brackets is nonzero only when \( k = k' \). We also use notation \( \delta_{k,k'} \) which is equal to 1 only if \( k = k' \) and zero otherwise. Therefore, we have

**Lemma 3.1.** If \( \{\tilde{x}_{\ell_j} = \sigma(\tilde{x}_{\ell_j})\} \) are i.i.d. noises with variance \( r^o \), then the covariance matrix of the Fourier coefficients of \( \sigma_j \equiv \sigma(x_j) = C_{j,1} \sigma_{\ell_j} + C_{j,2} \sigma_{\ell_j + 1} \), where \( C_{j,1} + C_{j,2} = 1 \), is as follows

\[
\langle \hat{\sigma}_k^o \hat{\sigma}_k^{o*} \rangle = \frac{r^o}{(2N+1)^2} \left[ \sum_{j=0}^{2N} (C_{j,1}^2 + C_{j,2}^2) \delta_{k,k'} + \sum_{\substack{j,j' \neq j' \neq \ell_j}} \left( C_{j,1} C_{j',2} \delta_{\ell_j, \ell_{j'}+1} + C_{j,2} C_{j',1} \delta_{\ell_j+1, \ell_{j'}} + C_{j,1} C_{j',1} \delta_{\ell_j, \ell_{j'+1}} + C_{j,2} C_{j',2} \delta_{\ell_j+1, \ell_{j'+1}} \bigg) \right].
\]
where the covariance matrix of the Fourier coefficients of \( \sigma \) points for every \( \ell \) the model grid points on \( \mathbb{Z}^2 \) (3.5)) and now the weights are independent of noise is correlated only to its adjacent one. Therefore, the cross terms in the third and fifth lines of (3.7) are zero and we have

\[
(3.6) \quad \text{defined in } \mathbb{R}^{2} \text{ of model grid points}
\]

Noises that are adjacently correlated, this represents pair of grid points with an observation in between them. The third and the fifth terms are contributions from noises that are interpolated with the same pair of adjacent data points, this represents sparse observations. For a special case, we have

**Proposition 3.2.** Suppose that \( \{ \tilde{\sigma}_\ell = \sigma(\tilde{x}_\ell) \}_{\ell=0,\ldots,2M} \) are i.i.d. noises with variance \( \sigma^2 \) and they are regularly spaced but shifted away by \( ch \) from the model grid points for every \( \ell \). In other words, \( \tilde{x}_\ell = x_j + ch \), where \( c \in (0,1) \) and \( x_j = jh \) represents the model grid points on \( 2\pi \) periodic domain such that \( (2M + 1)h = 2\pi \). Then the covariance matrix of the Fourier coefficients of \( \sigma_j \equiv \sigma(x_j) = C_{j,1}\tilde{\sigma}_j + C_{j,2}\tilde{\sigma}_{j+1} \), where \( C_{j,1} + C_{j,2} = 1 \), is diagonal and given in the following form

\[
\langle \tilde{\sigma}_k^0(\tilde{\sigma}_k')^* \rangle = \frac{\sigma^2}{2M + 1} \left[ (\epsilon^2 + (1 - \epsilon)^2) + \epsilon(1 - \epsilon) \cos(kh) \right] \delta_{k,k'}.
\]

**Proof.** First, note that \( C_{j,1} = \epsilon \) and \( C_{j,2} = 1 - \epsilon \) (see the Lagrange function in (3.5)) and now the weights are independent of \( j \). Notice also that each interpolated noise is correlated only to its adjacent one. Therefore, the cross terms in the third and fifth lines of (3.7) are zero and we have

\[
\langle \tilde{\sigma}_k^0(\tilde{\sigma}_k')^* \rangle = \frac{\sigma^2}{(2M + 1)^2} \left[ \sum_{j=0}^{2M} (\epsilon^2 + (1 - \epsilon)^2) \delta_{k,k'} + \epsilon(1 - \epsilon) \sum_{j,j' = 0}^{2M} e^{-i(j-j')kh} \delta_{k,k'} \right]
\]

\[
\quad \quad \quad + \epsilon(1 - \epsilon) \sum_{j,j' = 0, |j-j'|=1}^{2M} e^{-i(kj-k'j')h(1 - \delta_{k,k'})}
\]

\[
= \frac{\sigma^2}{(2M + 1)^2} \left[ \left( \epsilon^2 + (1 - \epsilon)^2 \right)(2M + 1) + \epsilon(1 - \epsilon) \sum_{j=0}^{2M} (e^{-ikh} + e^{ikh}) \right] \delta_{k,k'}
\]

\[
\quad \quad \quad + \epsilon(1 - \epsilon) \left( e^{ikh} + e^{-ikh} \right) \sum_{j=0}^{2M} e^{-i(kj-k'j)h(1 - \delta_{k,k'})}
\]

\[
= \frac{\sigma^2}{2M + 1} \left[ \left( \epsilon^2 + (1 - \epsilon)^2 \right) + 2\epsilon(1 - \epsilon) \cos(kh) \right] \delta_{k,k'}.
\]

In the last line, the non-diagonal terms vanish because \( \{ e^{-ikj} \} \) forms an orthonormal basis of \( \mathbb{C}^{2M+1} \).
In the next proposition, we introduce $\Lambda$, the ratio between the largest off-diagonal and the smallest diagonal components, as a measure that quantifies the importance of the correlation of the noise of the interpolated observations relative to their noise variances. This ratio, $\Lambda$, is a reasonable choice because we know exactly every component in the interpolated error covariance matrix. First, let us compute $\Lambda$ for a simple case where every observation is located exactly at the model grid point except one:

**Proposition 3.3.** Let $\{\sigma_j = \sigma(x_j)\}_{j=0,1,\ldots,2M}$ be i.i.d. noises with variance $r^o$ at regularly spaced grid points $x_j = jh$ such that $(2M + 1)h = 2\pi$. Let us perturb a single observation site $\tilde{x}_j$ by $\delta$, i.e., $\tilde{x}_j = x_j + \delta$. Then the ratio between the largest off-diagonal term and the smallest diagonal term is,

$$\Lambda \equiv \frac{\max_{k \neq k'} |R^o_{k,k'}|}{\min_k |R^o_{k,k}|} \leq \frac{2(\delta^2 + 2\delta h)}{(2M + 1)(\delta + h)^2 - 2(\delta^2 + 2\delta h)}.$$

**Proof.** We first compute the Lagrange weights, $\{C_{j,1}, C_{j,2}\}$, of noise $\sigma_j$ that is affected by the perturbation $\delta > 0$ on interval $[x_{j-1}, x_{j} + \delta]$ (the negative perturbation is a perfect symmetry). Here, we have

$$C_{j,1} = \frac{\delta}{\delta + h}, \quad C_{j,2} = \frac{h}{\delta + h} \quad (3.8)$$

and the interpolated noise at grid point $x_j$ becomes correlated,

$$\sigma_j = \frac{\delta}{\delta + h} \sigma_{j-1} + \frac{h}{\delta + h} \tilde{\sigma}_j. \quad (3.9)$$

Let $j^*$ be the index for the perturbed noise. In the physical space, the $(j, j^*)$ component of the covariance matrix is as follows:

$$\langle \sigma_j \sigma_{j^*} \rangle = r^o \delta_{j,j^*} (1 - \delta_{j,j^*}) + \frac{r^o(\delta^2 + h^2)}{(\delta + h)^2} \delta_{j,j^*} \delta_{j,j^*} + \frac{r^o}{\delta + h} \delta_{j,j^*} \delta_{j,j^*} (1 - \delta_{j,j^*} ) .$$

The first term of the RHS denotes the diagonal term for the unperturbed noise, the second term denotes the diagonal term for the perturbed noise, and the last term denotes the non-diagonal terms that signify correlations due to interpolation. In Fourier space, the covariance matrix can be expressed as follows

$$R^o \equiv \langle \tilde{\sigma}_k^o (\tilde{\sigma}_{k'}^o)^* \rangle = \frac{r^o}{(2M + 1)^2} \left( \frac{\delta^2 + h^2}{(\delta + h)^2} + \frac{2\delta}{\delta + h} \cos(kh) \right) \delta_{k,k'} + \frac{r^o}{(2M + 1)^2} \left[ \frac{\delta}{\delta + h} (e^{-ik'h} + e^{ik'h}) - \frac{2\delta h}{(\delta + h)^2} \right] e^{-(k-k')j^*} (1 - \delta_{k,k'}).$$

The non-diagonal component is bounded above as follows

$$|R^o_{k,k'}| \leq \frac{r^o}{(2M + 1)^2} \left( \left| \frac{2\delta}{\delta + h} \right| + \left| \frac{2\delta h}{(\delta + h)^2} \right| \right) = \frac{r^o}{(2M + 1)^2} \left( \frac{2\delta^2 + 4\delta h}{(\delta + h)^2} \right)$$

where $k \neq k'$, whereas the diagonal term is bounded below as follows

$$|R^o_{k,k'}| \geq \frac{r^o}{(2M + 1)^2} \left( \frac{(2M + 1)(\delta + h)^2 - 2(\delta^2 + 2\delta h)}{(\delta + h)^2} \right). \quad (3.10)$$

Taking fraction of these extrema, we obtain $\Lambda$.  \[\square\]
To confirm the result in Proposition 3.3, we compare it with numerically simulated Λ’s (see circles in Fig. 3.1). These numerically simulated ratios, Λ, are computed with $M = 10$. This result suggests that for a slight perturbation ($\delta/h < 0.5$) from the regularly spaced observations, the covariance matrix can be approximated by a diagonal covariance matrix. The ignored cross-covariance terms are about 5% relative to the smallest diagonal component when a single observation is perturbed. Furthermore, as $\delta/h$ grows, the ratio converges to a value less than 0.2; in contrast, the same ratio with trigonometric interpolation (triangle) increases to about 1.1 when the perturbation size is increased (see Fig. 3.1). When $1 < \delta/h < 2$, the perturbed observation location falls to the right of the next observation location, $x_{j+1} < \tilde{x}_j < x_{j+2}$, yet we consider the linear combination in (3.9) and assume that the noise at $x_{j+1}$ does not correlate with the interpolated noise at $x_j$. In this sense, Λ for linear interpolation case increases and is bounded above since the weight $C_{j,1}$ in (3.8) increases and converges to 1 (similarly $C_{j,2}$ decays to zero) as $\delta/h \to \infty$. In real applications, the condition $\delta/h > 1$ represents the sparse observation case and we only interpolate the closest observations.

When multiple observation are irregularly spaced, the ratio Λ will be larger because the $2M + 1$ factor in (3.10) will be replaced with a number that is much less than $2M + 1$ and the expression becomes cumbersome. In Section 4, we will numerically estimate this ratio and see that the ratio, although not so small for weakly and extremely irregularly spaced observations, is always much smaller than those of cubic spline and trigonometric interpolations. Indeed, we will demonstrate that accurate filtered solutions beyond the standard approaches are obtainable when the diagonal Fourier domain Kalman filter, discussed in Sections 2.1-2.2, is implemented with the piecewise linearly interpolated observations ignoring the not-so-small non-diagonal covariance terms in (3.7). Next, let us generalize Lemma 3.1 and Proposition 3.2 for non-local interpolation function of the form (3.11) without showing the detail proof.

**Proposition 3.4.** Let $\tilde{\sigma}_j \equiv \sigma(\tilde{x}_j)$ be an independent Gaussian noise at irregular grid points $\{\tilde{x}_j\}_{j=0,1,...,2M}$, that is, $\langle \tilde{\sigma}_j \rangle = 0$ and $\langle \tilde{\sigma}_j \tilde{\sigma}_k^* \rangle = \rho^2 I$. If the noise at regular
grid points \( \{ x_j = jh \}_{j=0,1,...,2N} \) with \((2N + 1)h = 2\pi\) is given by

\[
\sigma_j \equiv \sigma(x_j) = \sum_{\ell=0}^{2M} C_{j,\ell} \delta_{\ell}, \tag{3.11}
\]

where \( \sum_{\ell=0}^{2M} C_{j,\ell} = 1 \) for every \( j \), then the covariance matrix of the Fourier coefficients of \( \sigma_j \) is given by

\[
\langle \hat{\sigma}^o_k (\hat{\sigma}^o_{k'})^* \rangle = \frac{r^o}{(2N + 1)^2} \left[ \sum_{j=0}^{2N} \sum_{\ell=0}^{2M} C_{j,\ell}^2 \delta_{k,k'} + \sum_{j,j'=0}^{2N} \left( \sum_{\ell=0}^{2M} C_{j,\ell} C_{j',\ell} \right) e^{-i(j-j')kh} \delta_{k,k'} \right. \\
\left. + \sum_{j,j'=0}^{2N} \left( \sum_{\ell=0}^{2M} C_{j,\ell} C_{j',\ell} \right) e^{-i(j-k)h} (1 - \delta_{k,k'}) \right]. \tag{3.12}
\]

Furthermore, if \( M = N \) and the noise is regularly spaced and shifted by \( \epsilon h \) as in Proposition 3.2, i.e., \( \tilde{x}_j = jh + \epsilon h \), then the covariance is diagonal and it is given as follows:

\[
\langle \hat{\sigma}^o_k (\hat{\sigma}^o_{k'})^* \rangle = \frac{r^o}{(2N + 1)^2} \left[ C_0 + 2 \sum_{n=1}^{N} C_n \cos(knh) \right] \delta_{k,k'}, \quad \text{where} \tag{3.13}
\]

\[
C_n = \sum_{\ell=0}^{2M} C_{j,\ell} C_{j+n,\ell}, \quad n = 0, \ldots, N. \tag{3.14}
\]

As an example, we show in Appendix below that the cubic spline interpolation can be written in the form of (3.11) with \( \sum_{\ell} C_{j,\ell} = 1 \) for every \( j \). The analytical expression for an equivalent result to Proposition 3.4 is too cumbersome (even for cubic spline interpolation), therefore we will only show the numerically estimated ratio \( \Lambda \) in Section 4.

4. **Numerical Simulations.** The main goal of the numerical study below is to check the filtering skill of:

1. *The approximate FDKF with piecewise linear interpolation.* This approach implements FDKF to filter the piecewise linearly interpolated observations and ignores the non-diagonal components of the covariance matrix (3.7).

We will compare this filtering strategy with:

2. *The approximate FDKF with nearest neighbor interpolation.* This approach shifts the irregularly spaced observations to the nearest model grid point and performs the diagonal Fourier domain filter as in Section 2.1.

3. *The approximate FDKF with cubic spline interpolation.* This “non-local” approach implements FDKF to filter the cubic spline interpolated observations and ignores the non-diagonal components of the covariance matrix (3.13).

For completeness, we also compare the filtering schemes above with standard approaches that interpolate the model state variables to the observation state space:

4. *The physical space Kalman Filter with linear interpolation.* This standard approach performs Kalman filter globally in the physical space with a linear interpolation observation operator that maps the model state to the observation state.
5. The coupled FDKF with linear interpolation. Equivalent to 4 but apply the Kalman filter formula in Fourier domain. Implementation-wise, we define an observation operator that maps each Fourier coefficients to the observations.

6. The decoupled FDKF with linear interpolation. This approach is exactly like the approach 5 above but we ignore the non-diagonal observation error covariance matrix and perform the filter on each mode independently.

7. The coupled FDKF with trigonometric interpolation. This approach is described earlier in the beginning of Section 3 (see [29] for detailed discussion). Essentially, we use $A_{t,k} = (e^{ik\tilde{x}_t})$ (defined right after Eq. 3.1 in the Section 3) to map the Fourier coefficients of the model state variable to the irregular spaced observations in the physical space.

8. Diagonal FDKF with trigonometric interpolation. This approach implements FDKF with observation model in (3.2) but ignores the non-diagonal term of the covariance matrix in (3.3).

4.1. Numerical results for weakly irregularly spaced observations. In this section, we consider weakly irregularly spaced observations. In particular, sparse observations of $u(\tilde{x}_\ell, t_m)$ are spatially distributed as follows

$$\tilde{x}_\ell = x_\ell + \frac{\tilde{h}}{2}\lfloor\zeta_\ell\rfloor,$$

where $x_\ell = \ell\tilde{h}$ is equally distributed at every $P = 7$ true signal’s grid point with $(2M + 1)\tilde{h} = 2\pi, M = 31; \zeta_\ell$ is a random number from uniform distribution between interval $(-P/2, P/2)$; the notation $\lfloor\cdot\rfloor$ in (4.1) denotes the nearest integer operator. In total we have $2M + 1 = 63$ observations which are rather sparse relative to the true signal state dimension, $2N + 1 = 441$, but they are not so sparse considering the modes ($k = 10 - 20$) that exhibit instability transition are still observable. With this observation network (see circle in Figure 4.1), the numerically estimated ratio $\Lambda$ is about 0.16 for piecewise linear interpolation and this ratio is rather low relative to 0.64 for cubic spline interpolation, and 1.08 for trigonometric interpolation. The larger ratios for the last two interpolations are not surprising since they both are non-local interpolation techniques.

In Figure 4.1, we show snapshots of filtered solution $u$ for at time 500. First, we see that the most expensive physical space KF with standard linear interpolation produces the smoothest curve with the lowest average RMS error, 0.087, follows by FDKF with piecewise linear interpolation with RMS error, 0.1145. The other interpolation schemes degrade the filter accuracy (we obtained RMS errors of larger than 0.1732) on the smaller scale with high frequency oscillatory artifact. These oscillatory solutions are not so surprising and it can be predicted off-line by looking at the approximate spectra from various interpolation schemes. In Figure 2.1, we show the approximate spectra, computed from noisy observations with small error variance $\hat{\sigma}_o = 2 \times 10^{-5}$, even smaller than the energy spectrum of the smallest resolved wave number, $M^{-3}$. Notice that all except the piecewise linearly interpolated observations overestimate the true spectrum enormously on the smaller scale $k > 20$. The piecewise linear interpolation, in fact, underestimates the spectrum near instability modes, $10 < k < 20$. In this particular numerical experiment, we purposely choose small observation noise and the filter model to have nonzero resonant periodic forcing to increase the role of observations. This is simply to check how different interpolated observations affecting the filtered solutions under these stringent conditions. If the observation noise variance $\hat{\sigma}_o$ is larger, the role of observations decreases and the filter will weight
more toward the model. In this situation, we expect that all of these interpolation strategies to generate an even stronger oscillatory artifact especially if observations are not that sparse as in our example with $P = 7$.

In the second test, we consider sparser $2M + 1 = 21$ observations randomly distributed with Eq. (4.1) with $P = 11$. This is a stringent test since we have no information on the intermediate modes that exhibit stability transition. For this particular experiment, we choose observation error variance, $\tilde{\sigma}_o^2 = 0.01$, that is slightly larger than the energy spectrum of the fifth mode, $5^{-3}$. With such a sparse observation network, none of the eight schemes above are able to track smaller scale fluctuations accurately. The FDKF with piecewise linear interpolation produces the most accurate filtered solutions with lowest average RMS error and highest spatial correlation (see Table 4.1). An important aspect we can take from these performance measures is that when we interpolate the model state to the observation space, the filter skill degrades as one ignores the error covariance between different modes (methods 6 and...
8). Secondly, the cubic spline interpolation is not as accurate as the other since this non-local interpolation technique introduces oscillatory interpolated observations especially when the noise is not so small.

| Table 4.1 |

Weakly irregularly spaced observations: Average RMS errors and spatial correlation for numerical experiments with sparse $2M + 1 = 21$ observations and observation noise error $\sqrt{\sigma^2} = \sqrt{(2M + 1)^{p_0}} = 0.4583$.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>RMS error</th>
<th>Spatial corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FDKF with piecewise linear interp</td>
<td>0.3835</td>
<td>0.91</td>
</tr>
<tr>
<td>2. FDKF with nearest nbd</td>
<td>0.4417</td>
<td>0.89</td>
</tr>
<tr>
<td>3. FDKF with cubic spline</td>
<td>0.4184</td>
<td>0.88</td>
</tr>
<tr>
<td>4. Physical space KF with linear interp</td>
<td>0.5136</td>
<td>0.87</td>
</tr>
<tr>
<td>5. Coupled FDKF with linear interp</td>
<td>0.4843</td>
<td>0.88</td>
</tr>
<tr>
<td>6. Decoupled FDKF with linear interp</td>
<td>0.5089</td>
<td>0.87</td>
</tr>
<tr>
<td>7. Coupled FDKF with trig interp</td>
<td>0.4618</td>
<td>0.89</td>
</tr>
<tr>
<td>8. Decoupled FDKF with trig interp</td>
<td>0.5010</td>
<td>0.85</td>
</tr>
</tbody>
</table>

4.2. Numerical results for extremely irregularly spaced observations.

In this section, we consider an extremely irregularly spaced observations. In particular, we choose a sparse observation network such that they are clustered around two local regions and separated almost half way around a $2\pi$-periodic domain (see the circles in Figure 4.2). As before, we set $2M + 1 = 21$ such that we only filter modes $|k| < 10$. For this particular observation network, the ratio $\Lambda$ is 1.55 for piecewise linear interpolation, 72.58 for cubic spline, and in the order of $10^5$ for trigonometric interpolation. Notice that the ratios for the last two methods are much larger than that of the piecewise linear interpolation because both interpolation techniques invert an ill-ranked matrix due to pairs of closely located almost identical observations. From the snapshot in Figure 4.2, one should first convince that there is no reason for any of these filters to perform well near the unobserved locations. Notice that none of the numerically more expensive methods (either the physical space KF or the coupled FDKF, schemes 4, 5, and 7) perform better than the FDKF with piecewise linear interpolation. Here, the long time average RMS error of the FDKF with piecewise linear interpolation is the lowest relative to the remaining seven schemes with non-trivially high spatial correlation, 0.83 (see Table 4.2). This high filtering skill is rather surprising since we ignore completely the non-diagonal observation error covariance terms with large $\Lambda = 1.5$. On the other hand, the standard physical space Kalman filter with linear interpolation (scheme 4) is not skillful at all due to the stiffness in the computation of the Kalman gain matrix; this stiffness is not so surprising since the interpolation observation operator $G$ is ill-conditioned for this extremely irregularly spaced observation network. Furthermore, ignoring the non-diagonal error covariance terms significantly degrades the filter skill when conventional interpolation (see schemes 6 and 8) that maps the model state variable to the observation space is utilized. To close this section, we should mention that more sophisticated interpolation techniques may be needed so that information from the clustered observations that are not adjacent to the model grid points can be accounted and not ignored as in the case here with linear interpolation.

5. Summary. In this paper, we assessed several interpolation techniques for filtering irregularly spaced sparsely observed turbulent complex systems. Our main
Fig. 4.2. Snapshots of \( u \) at time 500 for extremely irregularly spaced observations. In each panel, we plot true signal (dashes), irregularly spaced observations (circle), interpolated observations (square), posterior mean estimate (solid thick).

Table 4.2

Extremely irregularly spaced and sparse observations: Average RMS errors and spatial correlation for numerical experiments with sparse \( 2M + 1 = 21 \) observations and observation noise error \( \sqrt{\frac{\sigma^2}{\sqrt{2M+1}}} = 0.4583 \).

<table>
<thead>
<tr>
<th>Schemes</th>
<th>RMS error</th>
<th>Spatial corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FDKF with piecewise linear interp</td>
<td>0.6774</td>
<td>0.83</td>
</tr>
<tr>
<td>2. FDKF with nearest nbd interp</td>
<td>1.4507</td>
<td>0.61</td>
</tr>
<tr>
<td>3. FDKF with cubic spline</td>
<td>1.0161</td>
<td>0.47</td>
</tr>
<tr>
<td>4. Physical space KF with linear interp</td>
<td>1.5488</td>
<td>0.57</td>
</tr>
<tr>
<td>5. Coupled FDKF with linear interp</td>
<td>0.9160</td>
<td>0.78</td>
</tr>
<tr>
<td>6. Decoupled FDKF with linear interp</td>
<td>3507.9</td>
<td>0</td>
</tr>
<tr>
<td>7. Coupled FDKF with trig interp</td>
<td>0.9198</td>
<td>0.77</td>
</tr>
<tr>
<td>8. Decoupled FDKF with trig interp</td>
<td>1.7558</td>
<td>0</td>
</tr>
</tbody>
</table>
focus is toward the potential of applying the interpolation unconventionally, that is, to map the irregularly spaced observations to the system regularly spaced grid points, and then filter these interpolated observations with numerically fast diagonal Fourier Domain Kalman Filter (FDKF). This study is relevant for practical application such as in atmospheric and ocean weather prediction problem where we typically use the data produced through various interpolation techniques \[9, 27\] relative to raw data. We studied several standard interpolation schemes including the nearest neighbor, piecewise linear, and cubic spline interpolations, and compared them to the standard strategy that linearly and trigonometrically interpolate the model variables into the observation space.

We showed that for a special case where only one grid point is irregularly located, the ratio \( \Lambda \) between the maximum of the non-diagonal components and the minimum of the diagonal components, is bounded (see Fig 3.1) when piecewise linear interpolation is used. We then numerically showed that even if this ratio is not so small when multiple observation locations are extremely irregularly spaced, the filtered solutions with the piecewise linear interpolations are nontrivially accurate relative to the other conventional strategies. On the other hand, if one uses standard approaches that interpolate the model variables to the observation space \[23, 29\], the filtered solutions severely degrade if the non-diagonal components of the covariance matrix are ignored.

In the numerical experiment above, we learn that under a stringent condition where the system is forced with a periodic resonant forcing, all three (nearest neighbor, cubic spline, and trigonometric) interpolations generate small scale oscillatory solutions even when observation noise is extremely small. These smaller scale oscillations are not so apparent with piecewise linear interpolation. In fact, we can predict whether such oscillatory solutions may occur by looking at the approximate spectrum of the interpolated observations. Numerically, we also showed that accurate filtered solutions of the FDKF with piecewise linear are robust even when observations (or at least the interpolated observations) on modes that exhibit instability are not available and the sparse observations are extremely irregularly spaced.

In future work, we will address two-dimensional domain problems with a more concrete example that models geophysical turbulent dynamics \[20\]. We are also investigating a different interpolation technique that does not ignore clustered observations that are not adjacent to the model grid point. Another challenge is to systematically understand a much complicated situation where the observation noises are spatially correlated as often encountered in real problems.

**Appendix. Cubic Spline Interpolation.** The goal of this Appendix is to show that the cubic spline interpolation is in the form of \( (3.11) \) and therefore the corresponding interpolated covariance matrix can be written in the form of \( (3.12) \). We begin with a short review of cubic spline interpolation on a periodic domain (see \[7\] for details). In our context, we want to find a cubic spline function \( s(x) \) such that

\[
s(\tilde{x}_i) = y_i, \quad i = 0, 1, \ldots, 2M, \tag{A.1}
\]

where \( \tilde{x}_i \in [0, 2\pi) \) are irregular grid points with periodic boundary domain. For each subinterval \( [\tilde{x}_i, \tilde{x}_{i+1}) \) where \( i = 0, 1, \ldots, 2M \), we would like to find the following cubic spline functions

\[
s(x) = A_i + B_i (x - \tilde{x}_i) + C_i (x - \tilde{x}_i)^2 + D_i (x - \tilde{x}_i)^3,
\]

and so there are totals of \( 4(2M+1) \) unknown coefficients \( \{A_i, B_i, C_i, D_i\} \). When \( i = 2M, \tilde{x}_{2M+1} = \tilde{x}_0 + 2\pi \) and \( s(\tilde{x}_i) = s(\tilde{x}_i + 2\pi) \) due to periodic boundary condition.
For \( s(x) \) to be a cubic spline function, \( s^{(p)}(x) \) has to be continuous for \( p = 0, 1, 2 \), which yields the following constraints

\[
\lim_{x \to \bar{x}_i^+} s^{(p)}(x) = \lim_{x \to \bar{x}_i^-} s^{(p)}(x) \quad (A.2)
\]

for \( p = 0, 1, 2 \) and \( i = 1, \ldots, 2M \). For periodic boundary conditions, we have the following constraints

\[
\lim_{x \to \bar{x}_0} s^{(p)}(x) = \lim_{x \to \bar{x}_{2M+1}} s^{(p)}(x), \quad (A.3)
\]

for \( p = 0, 1, 2 \). Therefore, we have \( 4(2M + 1) \) constraints, accounting conditions in \((A.1), (A.2), \) and \((A.3)\).

The standard approach for constructing \( s(x) \) (see e.g. [7]) is to let \( M_i = s''(\bar{x}_i), \quad i = 0, 1, \ldots, 2M \). Since \( s \) is cubic on each \([\bar{x}_i, \bar{x}_{i+1}]\), we have

\[
s''(x) = \frac{M_i + (x - \bar{x}_i)M_{i+1}}{h_i}, \quad i = 0, 1, \ldots, 2M,
\]

with \( h_i = \bar{x}_{i+1} - \bar{x}_i \) and \( M_{2M+1} = M_0 \). Integrating \( s''(x) \) and inserting conditions \((A.1)\), we have

\[
s(x) = \left( \frac{\bar{x}_{i+1} - x}{6h_i} \right)^3 - \frac{h_i^2}{6} \left( \bar{x}_{i+1} - x \right) M_i + \left( \frac{\bar{x}_{i+1} - x}{6h_i} \right)^3 \frac{h_i^2}{6} (x - \bar{x}_i) M_{i+1} + \frac{\bar{x}_{i+1} - x}{h_i} y_i + \frac{x - \bar{x}_i}{h_i} y_{i+1}.
\]

We then use conditions \((A.2)\) and \((A.3)\) to solve the linear problem \( A\bar{M} = \bar{b} \), for \( \bar{M} = (M_0, M_1, \ldots, M_{2M})^T \) where

\[
A = \begin{pmatrix}
\frac{h_{2M} + h_0}{6} & \frac{h_0}{3} & 0 & \ldots & 0 & \frac{h_{2M}}{6} \\
0 & \frac{h_0 + h_1}{6} & \frac{h_1}{3} & \ldots & 0 & 0 \\
\vdots & & & \ddots & & \vdots \\
\frac{h_{2M}}{6} & 0 & \ldots & 0 & \frac{h_{2M-1}}{6} & \frac{h_{2M-1} + h_{2M}}{3}
\end{pmatrix}
\]

is an \((2M + 1) \times (2M + 1)\) matrix and

\[
\bar{b} = \begin{pmatrix}
w_{1,0} - w_{0,0} & w_{2,0} - w_{0,2} & \ldots & w_{2M,2M} - w_{2M-1,2M-1}
\end{pmatrix}^T.
\]

Not only that \( A \) is symmetric, it is also a positive definite matrix since \( A \) is a diagonally dominant with positive diagonal entries (see [14] p.150) and hence it is invertible.

Suppose now that \( A^{-1} = (\alpha_{i,j})_{i,j=0,\ldots,2M} \), we can write

\[
\bar{M}_i = (\alpha_{i,0}, \alpha_{i,1}, \ldots, \alpha_{i,2M}) \bar{y} = \sum_{j=0}^{2M} w_{i,j} y_j,
\]

where

\[
w_{i,j} = \frac{\alpha_{i,j+1} - \alpha_{i,j}}{h_j} + \frac{\alpha_{i,j-1} - \alpha_{i,j}}{h_j}, \quad i, j = 1, \ldots, 2M - 1
\]

\[
w_{i,0} = \frac{\alpha_{i,1} - \alpha_{i,0}}{h_0} + \frac{\alpha_{i,2M} - \alpha_{i,0}}{h_{2M}}
\]

\[
w_{i,2M} = \frac{\alpha_{i,0} - \alpha_{i,2M}}{h_{2M}} + \frac{\alpha_{i,2M-1} - \alpha_{i,2M}}{h_{2M-1}}.
\]
Therefore, we can write (A.4) as follows

\[
s(x) = \frac{\bar{x}_{i+1} - x}{h_i} y_i + \frac{x - \bar{x}_i}{h_i} y_{i+1} + \left( \frac{\bar{x}_{i+1} - x}{h_i} \right)^2 h_i \left( \frac{\bar{x}_{i+1} - x}{h_i} \right) \sum_{j=0}^{2M} w_{i,j} y_j \\
+ \left( \frac{(x - \bar{x}_i)^3 - h_i^2(x - \bar{x}_i)}{6h_i} \right) \sum_{j=0}^{2M} w_{i+1,j} y_j = \sum_{j=0}^{2M} C_{i,j} y_j,
\]

where

\[
C_{i,j} = \left( \frac{\bar{x}_{i+1} - x}{h_i} \right)^2 h_i \left( \frac{\bar{x}_{i+1} - x}{h_i} \right) w_{i,j} + \left( \frac{(x - \bar{x}_i)^3 - h_i^2(x - \bar{x}_i)}{6h_i} \right) w_{i+1,j} \\
+ \frac{\bar{x}_{i+1} - x}{h_i} \delta_{i,j} + \frac{x - \bar{x}_i}{h_i} \delta_{i+1,j}
\]

for \( i = 0, 1, \ldots, 2M \) and \( w_{2M+1} = w_0 \).

To satisfy Proposition 3.4, we need to verify \( \sum_{j=0}^{2M} C_{i,j} = 1 \) for every \( i = 0, 1, \ldots, 2M \).

\[
\sum_{j=0}^{2M} C_{i,j} = \left( \frac{\bar{x}_{i+1} - x}{h_i} \right)^2 h_i \left( \frac{\bar{x}_{i+1} - x}{h_i} \right) \sum_{j=0}^{2M} w_{i,j} + \left( \frac{(x - \bar{x}_i)^3 - h_i^2(x - \bar{x}_i)}{6h_i} \right) \sum_{j=0}^{2M} w_{i,j} \\
+ \frac{\bar{x}_{i+1} - x}{h_i} + \frac{x - \bar{x}_i}{h_i} = 1,
\]

since the last two terms sum up to one and \( \sum_{j=0}^{2M} w_{i,j} = 0 \) for any \( i \). Therefore, we have shown that the covariance matrix of the interpolated Fourier noise coefficients (via cubic spline) is in the form of (3.12) with \( C_{j,\ell} \) given by (A.5).

REFERENCES