Dynamic Stochastic Superresolution of sparsely observed turbulent systems

Dynamic Stochastic Superresolution of sparsely observed turbulent systems

M. Branicki and A. J. Majda

Department of Mathematics and Center for Atmosphere Ocean Science,
Courant Institute of Mathematical Sciences, New York University, USA

Abstract

Real-time capture of the relevant features of the unresolved turbulent dynamics of complex natural systems from sparse noisy observations and imperfect models is a notoriously difficult problem. The resulting lack of observational resolution and statistical accuracy in estimating the important turbulent processes, which intermittently send significant energy to the large-scale fluctuations, hinders efficient parameterization and real-time prediction using discretized PDE models. This issue is particularly subtle and important when dealing with turbulent geophysical systems with an vast range of interacting spatio-temporal scales and rough energy spectra near the mesh scale of numerical models. Here, we introduce and study a suite of general Dynamic Stochastic Superresolution (DSS) algorithms and show that, by appropriately filtering sparse regular observations with the help of cheap stochastic exactly solvable models, one can derive stochastically 'superresolved' velocity fields and gain insight into the important characteristics of the unresolved dynamics, including the detection of the so-called *black swans*. The DSS algorithms operate in Fourier domain and exploit the fact that the coarse observation network aliases high-wavenumber information into the resolved waveband. It is shown that these cheap algorithms are robust and have significant skill on a test bed of turbulent solutions from realistic nonlinear turbulent spatially extended systems in the presence of a significant model error. In particular, the DSS algorithms are capable of successfully capturing time-localized extreme events in the unresolved modes, and they provide good and robust skill for recovery of the unresolved processes in terms of pattern correlation. Moreover, we show that DSS improves the skill for recovering the primary modes associated with the sparse observation mesh which is equally important in applications. The skill of the various DSS algorithms depends on the energy spectrum of the turbulent signal and the observation time relative to the decorrelation time of the turbulence at a given spatial scale in a fashion elucidated here. The DSS technique exploiting a simple Gaussian closure of the nonlinear stochastic forecast model emerges as the most suitable trade-off between the superresolution skill and computational complexity associated with estimating the cross-correlations between the aliasing modes of the sparsely observed turbulent signal. Such techniques offer a promising and efficient approach to constraining unresolved turbulent fluxes through stochastic superparameterization and a subsequent improvement in coarse-grained filtering and prediction of the next generation Atmosphere-Ocean System (AOS) models.

1 Introduction

This study is motivated by a common and topical problem in atmosphere-ocean and climate science which arises in forecasting multi-scale turbulent geophysical systems that are sparsely observed in both space and time [19, 34, 33, 22, 41, 11]. The resulting lack of resolution in the noisy observations of important turbulent processes which nontrivially interact with the large-scale dynamics through turbulent fluxes hinders efficient parameterization, filtering and real-time prediction using imperfect PDE models. Similar issues arise in turbulent engineering flows. From the practical standpoint, the demand for operationally practical and statistically accurate data assimilation and filtering methods escalates as the resolution of the imperfect discretized forecast models is substantially increased relative to the observation grid scale; in such cases the model error at the unresolved scales can be particularly detrimental to the predictive

capability of the model. This issue becomes particularly subtle when dealing with imperfect models of complex geophysical systems with an enormous range of interacting spatio-temporal scales and rough turbulent energy spectra near the mesh scale of the numerical models with intermittent energy transfers to the large-scale fluctuations. Important examples from the coupled atmosphere-ocean system include mesoscale convection, gravity waves and boundary layer dynamics in the atmosphere, or unstable currents and mesoscale eddies in near-surface ocean dynamics which cause problems in estimating transport from eddy-permitting ocean models (e.g., [21, 22]) and the spread of pollutants (e.g., [36, 3]). In all such cases, the sparseness of the noisy observations relative to the model grid scale implies that some information about the unresolved dynamics is 'mixed up' with the resolved dynamics at the observed scales.

Our approach to extracting the 'hidden' information pertaining to the unresolved dynamics is an extension of the framework developed in [32, 19, 20, 34] for real-time filtering of sparsely observed turbulent signals and actually exploits the spatial sparseness of the observations by utilizing the fact that a coarse observation network aliases high-wavenumber information into the resolved waveband. Here, we study a suite of general Dynamic Stochastic Superresolution (DSS) algorithms and show that, by appropriately filtering sparse observations with the help of cheap stochastic exactly solvable models, one can derive stochastically 'superresolved' velocity fields and gain insight into the important characteristics of the unresolved turbulent processes. Such superresolved information is important for constraining the turbulent backscatter from the unresolved fluxes into the large scale dynamics, as well as detecting the time-localized large amplitude events/triggers in the unresolved dynamics commonly referred to as the *black swans* [40, 39]. While in most geophysical applications the operationally acquired observational data is recorded on a highly irregular network, it is commonly interpolated onto a regular grid. Here, in view of future applications in stochastic multi-scale modeling and prediction, as well as for mathematical simplicity, we assume the regular periodic observation grid; such a framework can be achieved for irregular observations through appropriate interpolation which avoids introducing artificial correlations in the interpolated observations [16]. This approach is particularly suited for further development of stochastic superparameterization techniques [24, 42, 26, 15, 14] which is a subject of an ongoing research to be reported soon elsewhere. A number of important issues arise in this context which are addressed in this work:

(A) How well can we estimate the unresolved/aliased modes from sparse noisy observations?

(B) When is active filtering of the unresolved modes through the stochastic parameterization beneficial for filtering the primary modes associated with the sparse observation network?

(C) Which features of the unresolved processes can be captured through a DSS algorithm?

(D) How robust are these estimates with regard to uncertainties in the parameters of the DSS algorithm?

There exist at least two different ways in which one could address the above issues and we consider both of these approaches since they are often equally important in applications; the measures of skill of the DSS algorithms used here are based on:

(i) DSS skill in the crude statistical sense by means of time-averaged RMS errors and temporal pattern correlation. High skill in this sense is important in many applications as it can significantly improve time-averaged flux estimates, etc. (e.g., [22, 17, 33]).

(ii) DSS skill in the path-wise sense. High path-wise skill of DSS is important for detection of extreme time-localized events as shown later in this study.

Information-theoretic generalizations of these path-wise notions are discussed in [6] in the context of filtering.

Similar to the filtering context discussed in [32, 19, 20], the DSS algorithms consist of two main components: (i) the stochastic forecast model for the dynamics of all the aliased modes in the turbulent signal

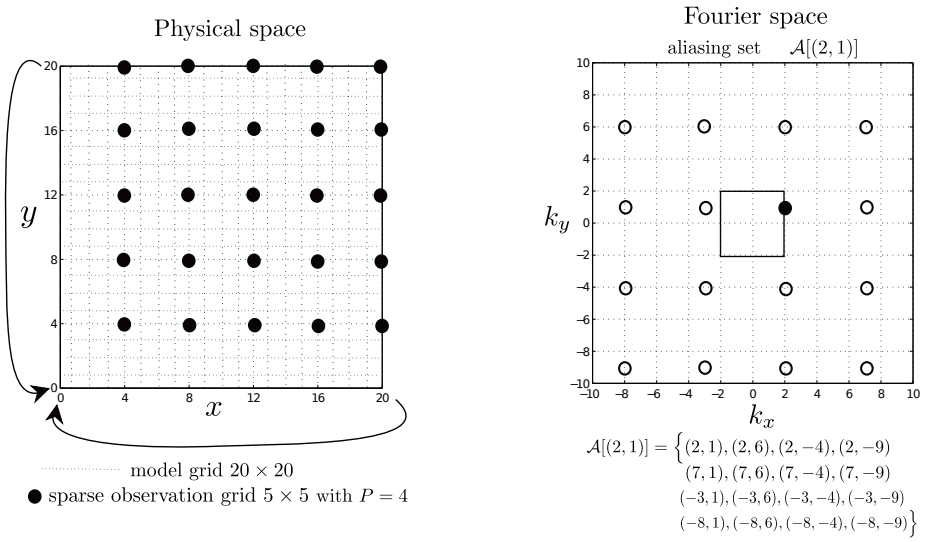

Figure 1: Schematic illustration of aliasing in Fourier space due to sparse observations in the physical space; here the 5×5 sparse observation grid is a regular subset of the 20×20 model mesh so that every $P = 4$ model mesh node is observed (left). The aliasing set $\mathcal{A}(\boldsymbol{l})$ of wavenumber $\boldsymbol{l} = (2, 1)$ is shown in Fourier space (right). For systems with quadratic, advection type nonlinearities none of the modes in the aliasing set are directly coupled (see (1) and Appendix A). The sparse observations lead to aliasing of the smaller scale Fourier modes resolved by the model with the primary Fourier modes associated with the coarse observation grid; in this example all the primary modes lie inside the region $-2 \leqslant k_x, k_y \leqslant 2$.

and (ii) appropriate filtering technique used to simultaneously estimate the resolved and unresolved modes by combining the model forecasts with the available sparse noisy observations. The distinct feature of DSS is that it is aimed at 'superresolving' the aliased modes from the scrambled observations of the large scale waveband rather than focusing on recovering the primary modes associated with the sparse observation network, as usually done in filtering turbulent systems with rapidly decaying energy spectrum [19, 20]. The forecast model used in all the DSS algorithms is obtained from the cheap, exactly solvable SPEKF-type forecast model [13, 12, 20, 5] in which the stochastic bias corrections are learned adaptively 'on the fly' from the observations through Kalman filtering in Fourier space (SPEKF here stands for Stochastic Paremeterization Extended Kalman Filter developed in [13, 12]). The 'judicious' model error introduced in these models with stochastic parameterization helps avoid the so-called curse of dimensionality while preserving the essential structure of the turbulent signal. The SPEKF methodology has been extensively validated for filtering with plentiful observations in a number of stringent idealized test problems [13, 12, 5]. Furthermore, Harlim and Majda [20] tested a suite of radical turbulent filtering algorithms with sparse regular observations in quasigeostrophic simulations mimicking baroclinic turbulence in the atmosphere and the ocean which showed high skill of the SPEKF-based techniques in the atmospheric and oceanic regimes compared with contemporary ensemble-based data assimilation schemes [1, 2]. The use of cheap stochastic models with adaptive stochastic bias correction terms to forecast the true nonlinear dynamics combined with appropriate Kalman filtering are the two main powerful features of DSS which lead to low-computational cost and the robust filtering skill of this approach.

According to the general framework developed in [32, 19, 33], the first step in filtering spatially extended systems in Fourier domain involves the discretization of the stochastic forecast model on a $\tilde{N} \times \tilde{N}$ physical-space mesh with observations on a $\tilde{M} \times \tilde{M}$ regular mesh with $\tilde{M} \leqslant \tilde{N}$ (see figure 1 for an illustration). This problem can be converted to $\tilde{M}^2$ independent filtering problems in Fourier domain [32] for the state vector containing all the Fourier modes within each disjoint aliasing set indexed by the modes resolved by

the observation grid. As we show later, non-trivial mode aliasing occurs when the observations are sparse, i.e., $\tilde{M} < \tilde{N}$, and it implies that the modes resolved by the model but contained in the same aliasing set cannot be distinguished based on the sparse observations alone. The need for active filtering of the modes unresolved by the observation grid in each of these filtering problems raises the issue of the nature and strength of correlations between the modes in each aliasing set. This is important in DSS because all the aliased modes are scrambled together by the sparse noisy observations; intuitively, the stronger the couplings between the aliased modes the more difficult the process of unscrambling them from the sparse noisy observations. However, the aliased Fourier modes in geophysical systems or engineering flows are expected to be relatively weakly correlated, as we show in Appendix A. In such systems the quadratic nonlinearities do not directly couple the Fourier modes contained in the same aliasing set; that is, if mode $\boldsymbol{k}$ is in the aliasing set $\mathcal{A}$ (see the example in figure 1), the quadratic couplings in the dynamics of $u_{\boldsymbol{k}}$ have the form

$$\frac{\mathrm{d}u_{\boldsymbol{k}}}{\mathrm{d}t} \propto \sum u_{\boldsymbol{l}}\, u_{\boldsymbol{m}}, \qquad \boldsymbol{k} \in \mathcal{A},\ \ \boldsymbol{l}, \boldsymbol{m} \notin \mathcal{A}. \tag{1}$$

This fact helps explain the skill of DSS algorithms in filtering sparsely observed multilayer geophysical flows in [22, 20]; it also influences the choice of the turbulent test models used for evaluating the skill of the DSS algorithms in this paper. Here, we extensively analyze these algorithms based on three widely different classes of turbulent signals obtained from systems with either a spatially decaying energy spectrum, e.g., $k^{-5/3}$, or equipartition of energy; these test signals are generated from the following:

- Synthetic signals obtained by superimposing a collection of path-wise solutions of the SPEKF model (see [13, 12, 5] and §2.2,3) from its different turbulent regimes (§2.2.1).
- Spatially extended models with weakly correlated modes in the aliasing set:
  - Truncated Burgers-Hilbert equation (TBHi) with turbulent dispersive dynamics and equipartition of energy except for the first two Fourier modes (see §4.1.1 and [4, 35, 23, 31, 25]);
  - The MMT model for dispersive wave turbulence with forcing and dissipation and a decaying energy spectrum (see §4.1.2 and [27, 8, 7, 43]) in a regime with breaking solitons and a strong inverse cascade of energy.
- Stochastic triad models. This last class of dynamics is included for completeness of the analysis; this extreme case scenario is given by a three-dimensional stochastic system with quadratic nonlinearity and strong couplings between the modes unlike the typical situation in (1).

Here, we show that the DSS framework is capable of successfully capturing the extreme time-localized events in the unresolved modes, and that it provides a good and robust skill for recovery the unresolved processes in terms of pattern correlation. Moreover, we show that DSS improves the skill for recovering the primary modes associated with the sparse observation network which is equally important in applications.

The first application of such stochastic superresolution techniques to a realistic oceanic situation can be found in [22] where significantly improved real-time estimates of poleward eddy heat transport was obtained using low-resolution, noisy observations of the upper ocean velocity field mimicking satellite altimetry data. The authors in [22] employed a DSS algorithm which uses analytical updates for the mean dynamics and Monte Carlo (MC) estimates of cross-correlations between the aliasing modes in the covariance matrix used in the Kalman filter. In this study we examine this and other DSS algorithms, some of them incorporating more accurate estimates for the cross-correlations, in order to understand the dynamical regimes of high skill for various approximations for the cross-correlations in the DSS algorithms. In particular, we show that the proposed techniques have comparable skill when the time interval of the data assimilation cycles exceeds the decorrelation time of the unresolved aliased modes. Moreover, we provide evidence for superiority of a DSS technique based on a Gaussian moment closure of the stochastic forecast model [5] which is easy to implement and avoids the MC estimates for the cross-correlations between the aliasing modes.

The paper is organized as follows. We start by introducing the general framework for Dynamic Stochastic Superresolution in Section 2. The details of the stochastic forecast model used in DSS and its various dynamical regimes are discussed in Section 2.2. The details of different algorithms for DSS are presented in Section 2.3; the main differences between the DSS algorithms lie in the approximations adopted for the cross-correlations between the aliasing modes. In Sections 3-5 we provide an extensive discussion of the performance of the DSS algorithms in different turbulent dynamical regimes. Section 3 deals with highly tunable synthetic examples which are used, in particular, to illustrate the skill of DSS for capturing the black swans represented by extreme time-localized events in the unresolved modes. In section 4 we analyze the performance of our DSS algorithms on two spatially extended nonlinear turbulent systems and illustrate a good DSS skill in various unambiguous realistic situations. Finally, Section 5 is concerned with the extreme case of dynamics with strongly correlated modes in the aliasing set; as it turns out, even in this unrealistically tough test case, the dynamics of the most energetic mode can be captured reasonably well and certainly better than from the observations alone. We conclude in Section 6 with a summary of the main findings and some remarks for further developments of this promising framework.

2 The stochastic framework for dynamic superresolution algorithms

Here, we generalize the Stochastic Parameterization Extended Kalman Filter (SPEKF) technique [13, 12, 19, 34] to the DSS setting and filter the sparse regular noisy observations to extract the information about the unresolved modes contained in the aliased signal (see figure 2) in order to derive the stochastically superresolved estimates.

There are firm mathematical arguments [32, 33] for this approach in the idealized test models involving the filtering of noisy turbulent signals from linear stochastic equations with either plentiful or sparse regularly spaced observations. The SPEKF methodology avoids the curse of ensemble size by committing a judicious model error through replacing the underlying nonlinear dynamics in the transformed (Fourier) domain by cheap exactly solvable stochastic models with the augmented dynamics of stochastic bias correction terms; the model error is corrected through 'on the fly' stochastic parameter estimation which was shown to be robust and skillful in a number of exhaustive tests [13, 12, 5].

Below, we first briefly summarize the general framework for DSS which is based on earlier developments [32, 19] in the context of filtering turbulent spatially extended systems from sparse regularly spaced observations. This setup involves the discretization of the forecast PDE model on a fine computational mesh with partial observations of the true turbulent signal on a coarse regular mesh which is a subset of the model mesh. We then utilize the results of Theorem 3 of [32] for sparse regular observations to obtain a number of independent filtering problems in Fourier space for each aliasing set of wavenumbers. This stochastic framework for dynamic superresolution has the following attractive properties:

- Ability to capture the relevant properties of the unresolved turbulent dynamics with the help of cheap exactly solvable models with stochastic parameterization of dynamic biases which are estimated "on the fly" from sparse observations.
- Flexibility and ease of implementation of stochastic bias correction terms. In particular, the stochastic estimation of biases in the phase, damping and forcing proves crucial for a skillful DSS technique in spatially extended systems with dispersive dynamics and intermittency.
- Low-computational overhead and judicious model error avoids the so-called curse of dimensionality while preserving the essential structure of the turbulent signal.
- Robustness to parameter uncertainties in the forecast models used for computing the prior estimates.

An example of a synthetic signal which aliases two unresolved modes into the primary mode associated with the sparse observation network is shown in figure 2. The stochastic forecast model for the dynamics

of the modes in each aliasing set is introduced in §2.2. In the forecast model the aliasing modes are assumed to be coupled only via the observations and this judicious model error is corrected 'on the fly' via stochastic parameter estimation. Finally, a suite of DSS algorithms exploiting various approximations of the cross-correlations between the aliasing modes is discussed in §2.3.

2.1 Superresolution of sparse regular observations

Majda and Grote [32] and Harlim and Majda [19] describe a set of off-line stability criteria and a novel filtering algorithm for sparsely observed, spatially extended turbulent systems that produces a "superresolved" estimate of the system state, $\boldsymbol{u}(x,y,t) \in I\!R^s$, with a higher nominal resolution than that of the observing network. Following these ideas, we assume that the spatially extended system representing the truth is given by a system of (stochastic) PDEs discretized via standard finite difference approximations on an equispaced $\tilde{N}\times\tilde{N}$ grid with the mesh points $(x_i,y_j)=(ih,jh)$, $1\leqslant i,j\leqslant\tilde{N}$, such that $\tilde{N}h=2\pi$. We assume canonical linear observations with the observation operator G and Gaussian noise, $\boldsymbol{\sigma}^{obs}\in I\!R^s$,

$$\boldsymbol{v}_{(k,l)\,m+1} = G\,\boldsymbol{u}\big(x_k,y_l,(m+1)\Delta t\big)+\boldsymbol{\sigma}^{obs}_{m+1},\quad 1\leqslant k,l\leqslant\tilde{M}, \tag{2}$$

at the nodes of a sparse regular $\tilde{M}\times\tilde{M}$ observation network with $\tilde{M}<\tilde{N}$. Here, for simplicity of exposition we assume that the nodes of the observation network coincide with every $P=\tilde{N}/\tilde{M}$ model grid point. As shown in [32, 19], the above problem in physical space reduces to $\tilde{M}^2$ independent filtering problems in Fourier space for $(P^2\times s)$-dimensional vectors $\hat{\boldsymbol{u}}_{(\boldsymbol{l})}$ which contain all the Fourier modes from the aliasing set of the primary wavenumber $\boldsymbol{l}=(l_x,l_y),\mathbb{Z}\ni|l_x|,|l_y|\leqslant\tilde{M}/2$ which is defined as

$$\mathcal{A}(\boldsymbol{l})=\Big\{\boldsymbol{k}=(k_x,k_y):\ \ \mathbb{Z}\ni|k_x|,|k_y|\leqslant\tilde{N}/2,\quad k_x=l_x+\tilde{M}q_x,\quad k_y=l_y+\tilde{M}q_y,\quad q_x,q_y\in\mathbb{Z}\Big\}. \tag{3}$$

Since the aliasing sets, $\mathcal{A}(\boldsymbol{l})$, for each primary mode $\boldsymbol{l}$ associated with the sparse observation network are disjoint, the $\tilde{N}^2$-dimensional filtering problem in the physical space can be converted to independent $(P^2\times s)$-dimensional filtering problems for each of the $\tilde{M}^2$ primary modes associated with the sparse observation grid (see figure 1 for an example). The prior mean and covariance, $\hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m}$ and $R_{(\boldsymbol{l})\,m+1|m}$, in the Fourier domain filter are updated based on the second-order statistics of the forecast model, i.e.,

$$\begin{aligned} a)\qquad \hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m} &= \mathcal{F}_{m+1}[\hat{\boldsymbol{u}}_{(\boldsymbol{l})\,m|m}],\\ b)\qquad R_{(\boldsymbol{l})\,m+1|m} &= \mathcal{C}_{m+1}[(R_{(\boldsymbol{l})\,m|m}], \end{aligned} \tag{4}$$

where $\mathcal{F}_{m+1}$ and $\mathcal{C}_{m+1}$ denote, respectively, the mean and covariance of the statistically exactly solvable forecast model described in the next section; the exact solvability of the forecast model advocated earlier in [13, 12, 5] removes the need for tangent approximations commonly used in the Extended Kalman Filter which is particularly prone to divergence when dealing with turbulent systems [5]. Note that contrary to filtering with plentiful observations ($P=1$), the covariance matrix $\mathcal{C}_{m+1}$ contains the cross-correlation terms between all the modes in the aliasing set $\mathcal{A}(\boldsymbol{l})$.

The posterior estimate for the mean and covariance, $\hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m+1}$ and $R_{(\boldsymbol{l})\,m+1|m+1}$, in the superresolution algorithm employs the standard Kalman filter steps applied to the $\tilde{M}^2$ disjoint aliasing sets $\mathcal{A}(\boldsymbol{l})$,

$$\hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m+1}=\hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m}+\boldsymbol{K}_{(\boldsymbol{l})\,m+1}\big(\bar{\boldsymbol{v}}_{(\boldsymbol{l})\,m+1}-\boldsymbol{G}_p\cdot\hat{\bar{\boldsymbol{u}}}_{(\boldsymbol{l})\,m+1|m}\big), \tag{5}$$

$$R_{(\boldsymbol{l})\,m+1|m+1}=R_{(\boldsymbol{l})\,m+1|m}-\Lambda(R_{(\boldsymbol{l})\,m+1|m},r^{obs})R_{(\boldsymbol{l})\,m+1|m}(\boldsymbol{G}_P\otimes\boldsymbol{G}_P^T)R_{(\boldsymbol{l})\,m+1|m}, \tag{6}$$

where

$$\boldsymbol{v}_{(\boldsymbol{l})\,m+1}=\boldsymbol{G}_P\cdot\hat{\boldsymbol{u}}_{(\boldsymbol{l})\,m+1|m+1}+\sigma^{obs}_{(\boldsymbol{l})\,m+1}/(2M+1)^2,\qquad \boldsymbol{G}_P=(1,1,\ldots,1)^T\in I\!R^P, \tag{7}$$

$$\boldsymbol{K}_{(\boldsymbol{l})\,m+1}=\Lambda(R_{(\boldsymbol{l})\,m+1|m},r^{obs})R_{(\boldsymbol{l})\,m+1|m}\,\boldsymbol{G}_P,\quad \Lambda(R_{(\boldsymbol{l})\,m+1|m},r^{obs})=(\boldsymbol{G}_P\cdot R_{(\boldsymbol{l})\,m+1|m}\,\boldsymbol{G}_P+r^{obs})^{-1}. \tag{8}$$

It is important to stress here that, unlike the Reduced Fourier Domain Kalman Filter [19], the filters employed for 'superresolving' sparsely observed turbulent signals need to filter more than just the primary Fourier mode in each aliasing set $\mathcal{A}(\boldsymbol{l})$; in the DSS setting all the aliased modes are coupled through a single observation [32, 19].

Below we often use the terminology for the modes in the aliasing set where the primary mode associated with the sparse observation network is called the observed or resolved mode and the aliased modes are the unobserved or unresolved modesthrough the observations. Note that under appropriate circumstances skillful filters can be developed that further simplify (5)-(8) in various ways and can actively filter the primary mode [19, 33].

2.2 The stochastic forecast model

The fundamental approach to constructing the stochastic forecast model for the DSS algorithms is the same as that developed for radical filtering strategies of turbulent signals in the Fourier domain [19, 20, 33]. However, the forecast model in the DSS framework has a more complex structure than in the previously considered filtering techniques since it accounts for the dynamics of the aliasing modes. Here, we follow the results of [32] and introduce two types of judicious model error into the forecast model. First, we assume that the dynamics of the Fourier modes in each aliasing set is independent of the dynamics in the remaining sets; this approximation was already used in §2.1 and [32], and validated in [19, 20, 22]. Secondly, unlike the true dynamics, the forecast model in each aliasing set consists of a collection of independent SPEKF-type systems [13, 12] for the aliased modes which are coupled only through the sparse observations. The idea for replacing the complex nonlinear interactions between different Fourier modes by multiplicative stochastic damping/frequency corrections and additional stochastic forcing is familiar from stochastic modeling of shear turbulence [38, 9]. Further justification of this judicious model error is given in Appendix A where we show that Fourier modes of geophysical systems with quadratic, advection type nonlinearity are not directly coupled with the other modes in the same aliasing set. Consequently, in DSS the stochastic exactly solvable forecast model for the dynamics of each Fourier mode $\hat{u}_k, k = 1, ..., P$ in each aliasing set is given by the following six-dimensional system:

$$\begin{aligned}
&(a) \quad \mathrm{d}\hat{u}_k(t) = \big[(-\gamma_k(t) + i\omega_k)\hat{u}_k(t) + b_k(t) + F_k(t)\big]\mathrm{d}t + \sigma_{u_k}\mathrm{d}W_{u_k}(t),\\
&(b) \quad \mathrm{d}b_k(t) = \big[(-d_{b_k} + i\omega_{b_k})(b_k(t) - \hat{b}_k)\big]\mathrm{d}t + \sigma_{b_k}\mathrm{d}W_{b_k}(t),\\
&(c) \quad \mathrm{d}\gamma_k(t) = -d_{\gamma_k}(\gamma_k(t) - \hat{\gamma}_k)\mathrm{d}t + \sigma_{\gamma_k}\mathrm{d}W_{\gamma_k}(t),\\
&(d) \quad \mathrm{d}\omega_k(t) = -d_{\omega_k}(\omega_k(t) - \hat{\omega}_k)\mathrm{d}t + \sigma_{\omega_k}\mathrm{d}W_{\omega_k}(t),
\end{aligned} \tag{9}$$

where W_{u_k}, W_{b_k} are independent complex Wiener processes and $W_{\gamma_k}, W_{\omega_k}$ are real Wiener processes. For each mode k of the signal there are eleven parameters in the system (9): three damping parameters d_{b_k}, $d_{\gamma_k}, d_{\omega_k} > 0$, one phase parameter ω_{b_k}, three stationary mean terms $\hat{b}_k$, $\hat{\gamma}_k$ and $\hat{\omega}_k$ and noise amplitudes $\sigma_{u_k}, \sigma_{b_k}, \sigma_{\gamma_k}, \sigma_{\omega_k} > 0$; F_k is a deterministic forcing. Due to the independent evolution of the aliased modes assumed in the forecast model (9), nontrivial cross-correlations between the aliasing modes $\hat{u}_k$ and their stochastic bias correction terms γ_k, ω_k, b_k are induced through the assimilation of the observations. Moreover, the cross-correlations between the different aliasing modes and their bias corrections, $(\hat{u}_k, \gamma_k, \omega_k, b_k)$ and $(\hat{u}_l, \gamma_l, \omega_l, b_l), k \neq l$, decay in time from the initial values generated at the data assimilation cycles. Accurate estimates of these cross-correlations are generally needed for skillful superresolution of the aliased turbulent signal, as will be discussed in the following sections. In the suite of the DSS algorithms introduced in §2.3 different approximations for these cross-correlation terms will be considered, resulting in different stochastic superresolution skill.

We note that more elaborate and prohibitively expensive statistical closures for turbulence are available in the literature where the statistical effects of the nonlinear interactions at each Fourier mode are determined from global space-time nonlinear interactions across scales. In particular, O'Kane and Frederiksen

[37] have implemented one of these elaborate closures as the prediction step operator in the filtering of atmospheric blocking events with high skill for plentiful observations. Our approach relies on direct exploitation of sparse observations combined with computationally cheap filtering algorithms which 'learn' the net statistical effects of the nonlinear interactions in the dynamics of each filtered Fourier mode 'on the fly' and mimic the effect of more accurate, yet prohibitively expensive closures; this approach is further justified by the weak coupling between the aliasing modes in geophysical-type flows, as discussed in the Appendix A.

2.2.1 Dynamical regimes of the forecast system

The nonlinear stochastic system (9) has a number of attractive properties as a test model in our analysis. Firstly, it has a surprisingly rich dynamics mimicking turbulent signals in various regimes of the turbulent spectrum, including intermittently positive finite-time Lyapunov exponents [5]. Secondly, due to the particular structure of the nonlinearity in (9), exact path-wise solutions and exact second-order statistics of this non-Gaussian system can be obtained analytically [13]. A number of physically relevant dynamical regimes of the model (9) characterized by stability of the mean dynamics, $\chi = -\hat{\gamma} + \sigma_\gamma^2/2d_\gamma^2 < 0$, were determined in [5] and are listed below:

(I) *Regime of plentiful, short-lasting transient instabilities in the resolved component $u(t)$ with fat-tailed marginal equilibrium PDF.* This type of dynamics is characteristic of the turbulent energy transfer range and in the model (9); it occurs for $\sigma_\gamma, d_\gamma \gg 1$, $\sigma_\gamma/d_\gamma \sim \mathcal{O}(1)$ and positive $\hat{\gamma}$ sufficiently large so that $\chi < 0$. This is a regime of rapidly decorrelating damping fluctuations $\gamma(t)$ and relatively weak damping of fluctuations in $u(t)$ with decorrelation time given approximately by $1/\hat{\gamma}$.

(II) *Regime of intermittent large-amplitude bursts of instability in $u(t)$ with fat-tailed marginal equilibrium PDF.* This regime is characteristic of turbulent modes in the dissipative range and it occurs for small σ_γ, d_γ, with $\sigma_\gamma/d_\gamma \sim \mathcal{O}(1)$ and $\hat{\gamma}$ sufficiently large so that $\chi < 0$. The decorrelation time of the damping fluctuations $\gamma(t)$ is long here but the decorrelation time of fluctuations in $u(t)$ can vary widely, as in regime (I).

(III) *Laminar regime with nearly Gaussian equilibrium PDF.* Here, the transient instabilities in $u(t)$ occur very rarely. This type of dynamics occurs in (9) for $\sigma_\gamma^2/2d_\gamma^2 \gg 1$, $\sigma_\gamma \sim \mathcal{O}(1)$ and $\hat{\gamma}$ large enough so that $\chi < 0$. In the extreme case when $\hat{\gamma} \gg \sigma_\gamma^2/2d_\gamma^2$ there are almost surely no transient instabilities in the dynamics of $u(t)$.

The path-wise solutions from the above three turbulent regimes of (9) are used in figure 2 to generate the synthetic signal with two unresolved modes aliased into the primary mode associated with the sparse observation network according to (7) with $P = 3$.

2.3 The Dynamic Stochastic Superresolution algorithms

Here, we introduce three algorithms for Dynamic Stochastic Superresolution (DSS) of turbulent signals which all exploit the cheap exactly solvable stochastic model (9) and nonlinear extended Kalman filtering described in the previous section in order to capture the relevant properties of the unresolved turbulent dynamics from the sparse regular observations. These algorithms are based on the earlier work of Majda and Grote [32] and Harlim and Majda [20] developed for filtering sparsely observed turbulent signals. In the DSS context the spatial sparseness of the observations is exploited for superresolving the turbulent signal by utilizing the fact that a coarse observation network aliases high-wavenumber information into the resolved waveband.

Recall from §2.1 that, given a regular sparse observation network which coincides with every P-th point of the discretization mesh of the forecast model, the DSS technique involves active filtering of P^2 Fourier modes in each aliasing set with a $6P \times 6P$ covariance matrix $\mathcal{C}$. The cross-correlations between

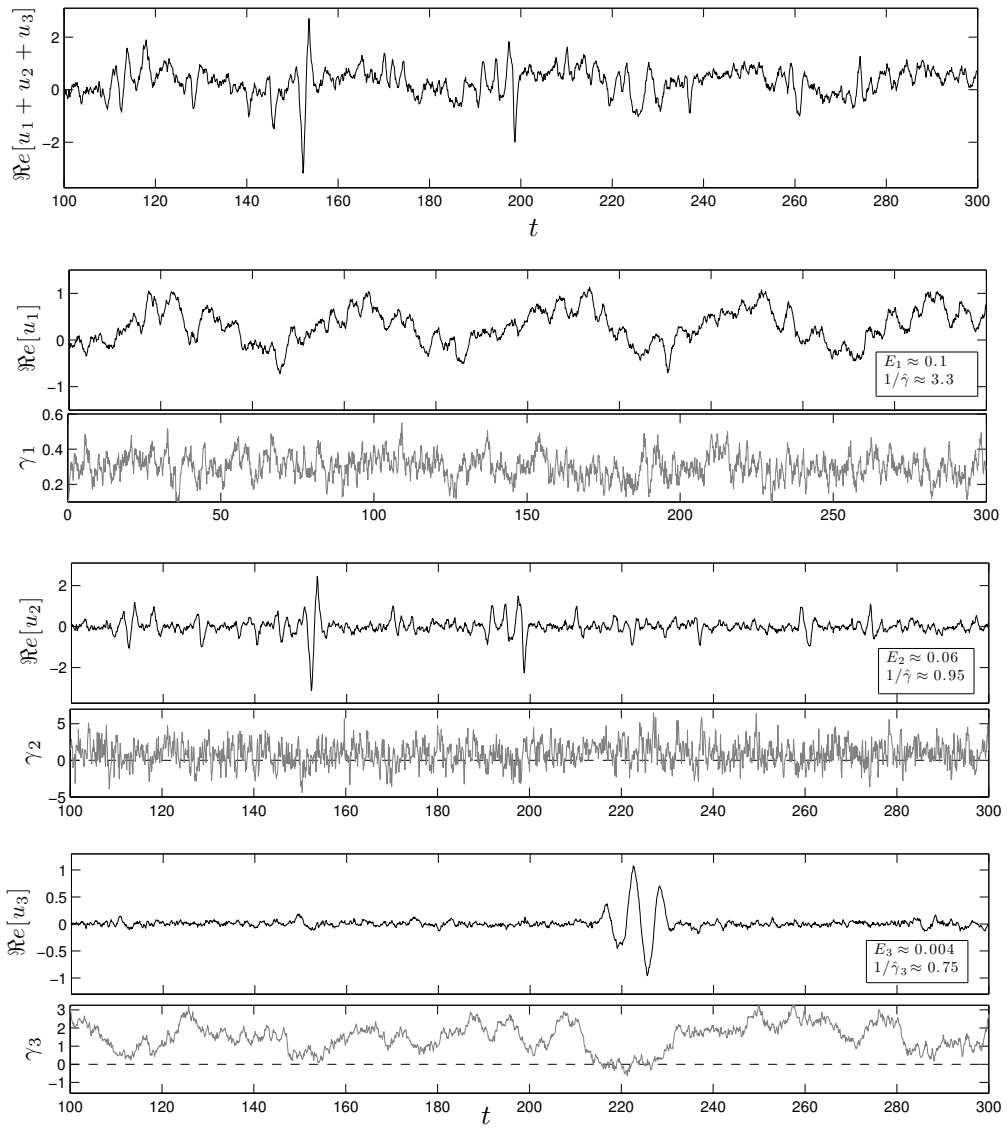

Figure 2: Example of an aliased signal composed of three Fourier modes with decreasing energy and contained in the same aliasing set (see §2.1); in this configuration two unresolved Fourier modes, u_2 and u_3, are aliased into the primary mode u_1 associated with the sparse observation grid. The three sample paths are obtained from the SPEKF system (9) in regimes I II, III of mean-stable dynamics (see §2.2.1) which mimic different regimes of the turbulent spectrum.
Parameters of the SPEKF systems:
$\hat{\gamma}_1 = 8$, $\sigma_{\gamma_1} = 1$, $d_{\gamma_1} = 0.3$, $\sigma_{u_1} = 1$, $\omega_{u_1} = 2$, $A_{f_1} = 6$, $\omega_{f_1} = 0.15$, $\hat{b} = 0, \sigma_{b_1} = d_{b_1} = 0.4, \omega_{b_1} = 1$
$\hat{\gamma}_2 = 1.95$, $\sigma_{\gamma_2} = 10$, $d_{\gamma_2} = 10$, $\sigma_{u_2} = 0.1$, $\omega_{u_2} = 1$, $A_{f_2} = 0$, $\omega_{f_2} = 0$
$\hat{\gamma}_3 = 1.2$, $\sigma_{\gamma_3} = 1$, $d_{\gamma_3} = 1$, $\sigma_{u_3} = 0.05$, $\omega_{u_3} = 2$, $A_{f_3} = 0$, $\omega_{f_3} = 0$

different aliasing modes and different stochastic bias corrections, $\{\hat{u}_k, \gamma_k, \omega_k, b_k\}$ and $\{\hat{u}_l, \gamma_l, \omega_l, b_l\}, k \neq l$, in the forecast model are induced by the sparse observations which scramble these modes together at every analysis step; these terms can be derived analytically in a similar fashion to the cross-correlations between each mode $\hat{u}_k$ and its bias correction terms γ_k, ω_k, b_k derived in [13] in the context of filtering with plentiful observations ($P = 1$). Instead of deriving these additional exact formulas, we consider here a simplified but more flexible approach based on various approximations of the covariance matrix of the forecast model. The suite of DSS algorithms introduced below includes the algorithm with Monte Carlo estimates for the cross-correlations between the aliasing modes used in [22], and two other algorithms which all use the stochastic forecast model (9) described in §2.2. The differences between these algorithms lie in the treatment of the cross-correlations between the aliasing modes. We compare the skill of these three different DSS algorithms in the following sections in order to establish the regimes of applicability of various approximations in the covariance matrix. As we show in §4.2, accurate estimates of these cross-correlations are needed for skillful superresolution of the aliased turbulent signal, except for DSS with sufficiently long observation times. In this context the technique based on a simple Gaussian moment closure of the forecast model (9) and described below (GCSSF) emerges as the most suitable trade-off between the superresolution skill and computational complexity associated with estimating the cross-correlations between the aliasing modes.

Algorithms for Dynamic Stochastic Superresolution

- **GCSSF** This Gaussian Closure Stochastic Superresolution Filter algorithm uses the approximate statistics of the forecast model (9) which is obtained via the simple Gaussian moment closure known from turbulence theory and used successfully for filtering turbulent signals in [5]. The $6P \times 6P$ covariance matrix in GCSSF, whose size depends on the number of the aliasing modes in each aliasing set, includes all cross-correlations between different modes and their parameters, $\{u_m, \gamma_m, \omega_m, b_m\}$ and $\{u_n, \gamma_n, \omega_n, b_n\}$. As discussed earlier in [5], for systems with quadratic nonlinearities such as (9) this closure correctly accounts for the turbulent backscatter in the evolution of the mean but it neglects the third order moments of fluctuations in the evolution of the covariance; this approximation performed very well for filtering with plentiful observations ($P = 1$) in [5]. As discussed in the following sections, GCSSF emerges as the most suitable trade off between the superresolution skill and the computational complexity associated with estimates of the cross-correlations between the aliasing modes. We derive the dynamical system for the evolution of the approximate statistics in Appendix B.

  We note that possibilities exist for further approximations of the covariance matrix within the Gaussian closure approximations. These can, in principle, be obtained by judiciously neglecting some of the cross-correlations in the covariance matrix. This procedure, however, frequently leads to violation of the positive semi-definiteness of the approximate covariance matrices.

- **MC mSPEKF** This multi-SPEKF forecast model uses analytical updates for the mean of all quantities but incorporates Monte Carlo estimates for the cross-correlations between the P Fourier modes and their parameters $\{u_k, \gamma_k, \omega_k, b_k\}$, $k = 1, 2, ...P$ within each aliasing set and with analytical updates for the mean. In the simulations $N = 100$ samples are used, similar to [22]. We will show in the following sections that while this approach has well defined regimes of good performance, it looses skill when observations are assimilated frequently relative to the typical decorrelation time of the dominant unresolved modes.

- **crude mSPEKF** This is the most simplified DSS algorithm with a block-diagonal covariance matrix obtained by neglecting all the cross-correlations between different modes in the aliasing set (see figure 3 for an intuitive sketch). The remaining cross-correlations between the modes u_k and their stochastic parameters, b_k, γ_k, ω_k, $k = 1, 2, ...P$, are computed using the exact analytical formulas derived in [13]. In the tests carried out in the subsequent sections we impose another simplification by assuming constant phases, $\omega_k = const.$, in the forecast model of the crude mSPEKF.

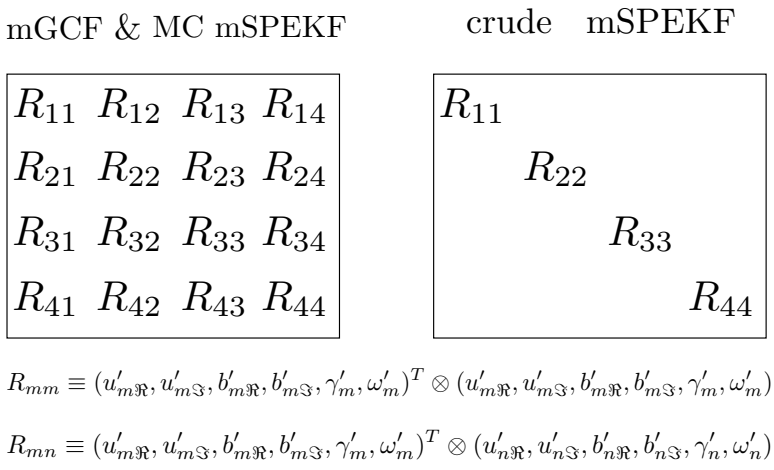

Figure 3: Schematic illustration of the covariance matrices R in the DSS algorithms (§2.3); this example corresponds to $P = 4$ wavenumbers in the aliasing set $\mathcal{A}$ (see §2.1).

The following sections discuss a series of examples of increasing complexity where we examine the skill of the different algorithms in different dynamical situations.

3 Synthetic examples for testing DSS algorithms

We begin the analysis of the DSS algorithms by discussing the stochastic superresolution of synthetic signals which are generated by superposing different turbulent modes u_k from the stochastic test model (9). Figure 2 shows an example of such an aliased signal involving three modes from different dynamical regimes of (9) with rapidly decaying energies. This idealized setting provides an unambiguous and highly tunable testbed for highlighting the main features of the DSS algorithms. The main questions addressed in this section concern the following general themes

- How well can we estimate the unresolved modes aliased into the observed modes?
- Which features of the unresolved modes can be captured through a suitable DSS algorithm? Which features can be recovered path-wise and which features are captured in the statistical sense (e.g., through the pattern correlation)?

Here, we illustrate the performance of the DSS algorithms on three-mode combinations from different dynamical regimes of the turbulent test model (9) with (i) decaying energy spectrum and (ii) equipartition of energy. This configuration mimics DSS in a single aliasing set for a spatially extended one-dimensional system observed on a sparse regular network which coincides with every third point of the discretized mesh of the forecast model (see §2.1). We expect all the DSS algorithms to perform similarly in this setting since the aliased modes in the synthetic signal are uncorrelated and different approximations of the cross-correlations between the modes are a secondary issue here. In particular, we show that even our simplest DSS technique, crude mSPEKF (see §2.3), has a good skill for recovering the unresolved modes in terms of RMS error and pattern correlation, and a very good pathwise-skill for capturing time-localized extreme events in the unresolved modes which we refer to hereafter as the *black swans* following [40, 39]. The skill of our DSS algorithms in realistic spatially extended turbulent systems with weak correlations between the aliasing modes (see also Appendix A) is discussed in the next section where differences between different approximations in the suite of the DSS algorithms begin to emerge.

The average statistical skill is assessed both in terms of the *Root Mean Square Error* (RMSE) and pattern correlation $\langle corr \rangle$; the RMS error between the truth u and its DSS estimate u^{DSS} is defined in a standard fashion as

$$RMSE(u, \bar{u}^{\text{DSS}}) = \sqrt{\frac{1}{M}\sum_{m=1}^{M} |\bar{u}^{\text{DSS}}_{m|m} - u_m|^2}, \tag{10}$$

where $u_m \equiv u(t_m)$ and $\bar{u}^{\text{DSS}}_{m|m}$ is the posterior estimate at time t_m obtained from filtering in the DSS algorithm (see §2.1). The RMS error is computed after discarding the initial transient onto the attractor. The pattern correlation between the (real-valued) truth x and its DSS estimate x^{DSS} is defined as

$$\big\langle corr \big\rangle(x, \bar{x}^{\text{DSS}}) = \frac{\sum_{m=1}^{M} x_m x^{\text{DSS}}_{m|m}}{\sqrt{\sum_{m=1}^{M} x_m^2 \sum_{m=1}^{M} (x^{\text{DSS}}_{m|m})^2}}, \tag{11}$$

and for complex valued truth, u, and its DSS estimate, u^{DSS}, it is given by

$$\big\langle corr \big\rangle(u, \bar{u}^{\text{DSS}}) = \frac{1}{3}\Big(\big\langle corr \big\rangle(\Re[u], \Re[\bar{u}^{\text{DSS}}]) + \big\langle corr \big\rangle(\Im[u], \Im[\bar{u}^{\text{DSS}}]) \Big). \tag{12}$$

Note that if two signals are close to each other, the RMSE of their difference is close to zero and their correlation pattern is close to one. For two different signals RMSE diverges from zero and the pattern correlation approaches zero.

Two path-wise examples of DSS are shown in figures 4-5, illustrating the high skill of GCSSF for recovering the two most energetic modes (u_1, u_2) and for capturing the black swans generated by the unresolved modes; high skill for capturing these extreme events is present for both the decaying energy case (figure, 4) and for the equipartitioned configuration in figure 5.

Figure 6 shows the average statistical skill, in terms of RMS errors and pattern correlation, for superresolving the synthetic signals as functions of the observation time, Δt^{obs}, and observation noise variance, r^{obs}; decaying energy spectrum was imposed on the aliasing modes. The skill dependence on the observation time is illustrated for the observation noise variance smaller than the energy tertiary mode u_3 in the aliased signal; deterioration of the skill for recovery of the unresolved modes u_2, u_3 can be seen for Δt^{obs} approaching the decorrelation time of u_3.

Finally, the robustness of DSS is illustrated in figures 7-8 where either incorrect mean damping parameters $\hat{\gamma}_{1,3}$ or incorrect deterministic forcing parameters are used in the DSS algorithms to superresolve the turbulent signal involving three aliased modes generated from different dynamical regimes of (9). Note, in particular, that the errors in the mean damping parameters imply underestimating/overestimating the decorrelation time of the respective aliased mode. Clearly, the DSS skill is largely insensitive and robust to parameter uncertainties in the DSS forecast model for a wide range of parameter values. The skill sensitivity to errors in the remaining parameters is even weaker and not shown. This skill robustness of the DSS algorithms, which is tested also the realistic spatially extended systems in §4, is reminiscent of the robustness of SPEKF filtering of turbulent spatially extended systems with plentiful observations [13, 12, 5]. The additional complication in the DSS setting arises due to the mode coupling through sparse noisy observations which makes the robustness of the DSS algorithms not immediately obvious from the previously studied setting.

Based on the results illustrated in figures 4-8 we stress the following points:

- GCSSF has a very good overall skill for filtering the synthetic signals both in terms of RMSE and pattern correlation. However, the reduced GCSSF and crude mSPEKF which neglect cross-correlations between the aliasing modes perform similarly in this setting.
- Good *black swan detection* (i.e., extreme time-localized events in the unresolved modes). GCSSF and even the crude mSPEKF can successfully detect and resolve the large amplitude bursts in the unobserved modes, as well as resolve the time-mean trends in these modes.

- The DSS for capturing the black swans is similar for both the decaying and equipartitioned energy spectrum (e.g., figures 4 and 5).
- Robust superresolution with respect to the parameters in the DSS algorithm
  - For correct parameters GCSSF and crude mSPEKF perform similarly well for the capture of the black swans in a wide range of parameter values and observation times, and for various turbulent dynamical regimes (e.g., figure 4).
  - For incorrect forcing GCSSF performs significantly better especially for the recovery of the unresolved modes (figures 7-8).

  The desirable lack of sensitivity in the DSS skill to variations of the intrinsic parameters in the algorithm is due to the specific design of the SPEKF forecast model in which the model error is learned adaptively 'on the fly' through the Kalman update; similar robustness with regard to the model parameters were reported earlier for filtering in [13, 12].
- The main differences between GCSSF and crude mSPEKF are associated with the resolution of the black swans; neglecting the cross correlations between the aliasing modes in crude mSPEKF is the most detrimental for the DSS skill in such situations.
- In regimes with few large amplitude instabilities in the unresolved modes the DSS algorithms perform similarly well. The good performance of the crude mSPEKF in this configuration is a further confirmation of the ideas used in constructing the Strongly Damped Approximate Filter (SDAF) [19] in cases with very weak cross-correlations between the aliasing modes.

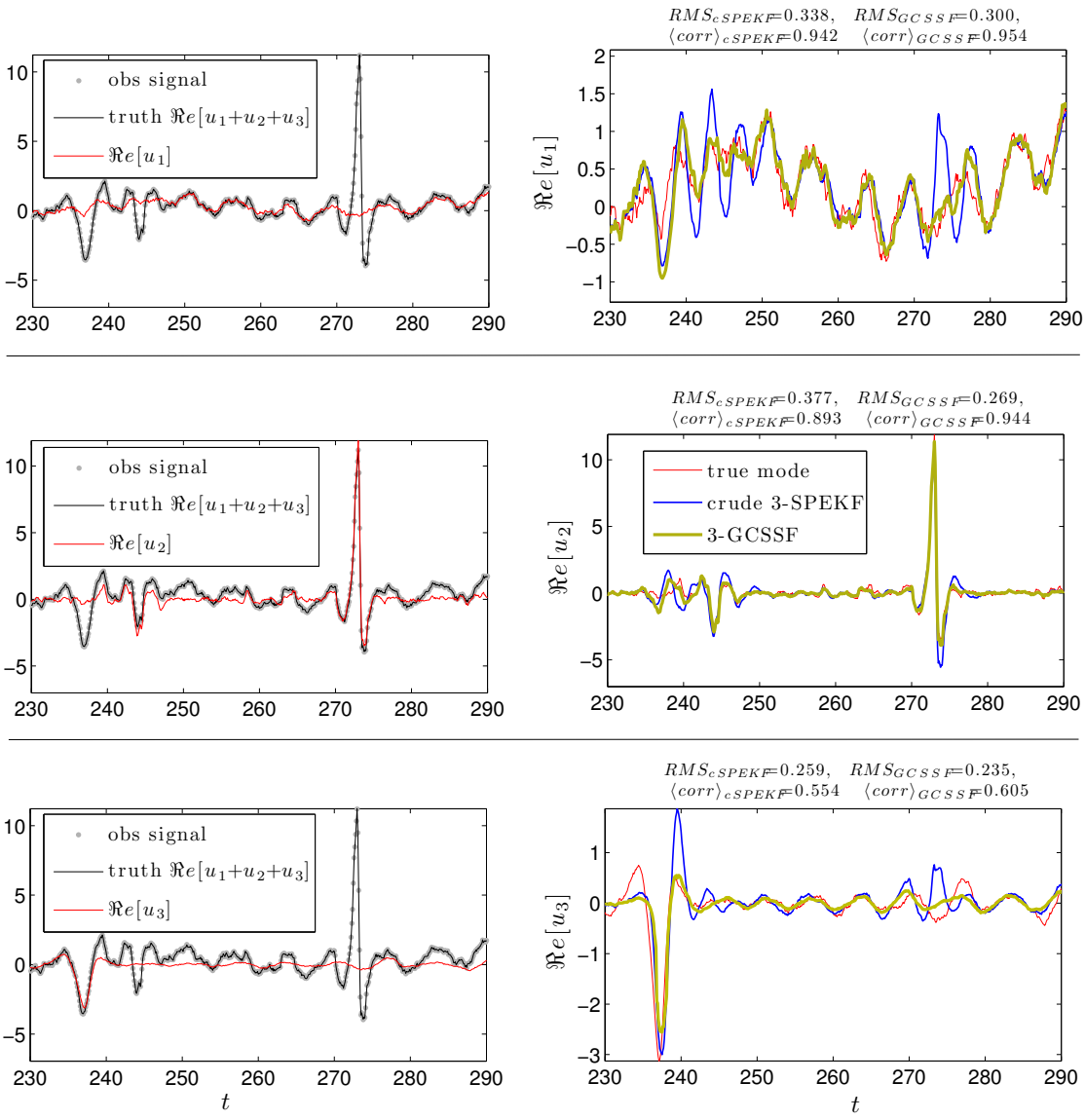

Figure 4: **Synthetic three-mode aliased signal with decaying energy.** Path-wise example of DSS for a synthetic three-mode signal (regimes (III)+(I)+(II) of the system (9)) with mode energies $\boldsymbol{E_1 \approx 0.15, E_2 \approx 0.06, E_3 \approx 0.004}$; the results shown are obtained from two DSS algorithms, 2-GCSSF and crude 2-SPEKF (see §2.3). DSS parameters are as in the truth modes (see figures 6-8 for the average statistical skill for incorrect DSS parameters).

True signal parameters: $[u_1]$ $\hat{\gamma} = 0.3, d_\gamma = 4.7, \sigma_\gamma = 5, \sigma_u = 0.3, \omega_u = 1, d_b = 0.4, \sigma_b = 0.1, \omega_b = 1$;
$[u_2]$ $\hat{\gamma} = 1.05, d_\gamma = 1.2, \sigma_\gamma = 1.2, \sigma_u = 0.35, \omega_u = 2, d_b = \sigma_b = 0.4, \omega_b = 1$;
$[u_3]$ $\hat{\gamma} = 1.35, d_\gamma = 0.25, \sigma_\gamma = 0.45, \sigma_u = 0.1, \omega_u = 1, d_b = \sigma_b = 0.1, \omega_b = 1$;
Forcing: $F_{u_1}(t) = 1 + 1.5\exp(i0.15t)$, $F_{u_2}(t) = 0$; *Observations*: $\Delta t^{obs} = 0.05$, $r^0 = 0.05$.

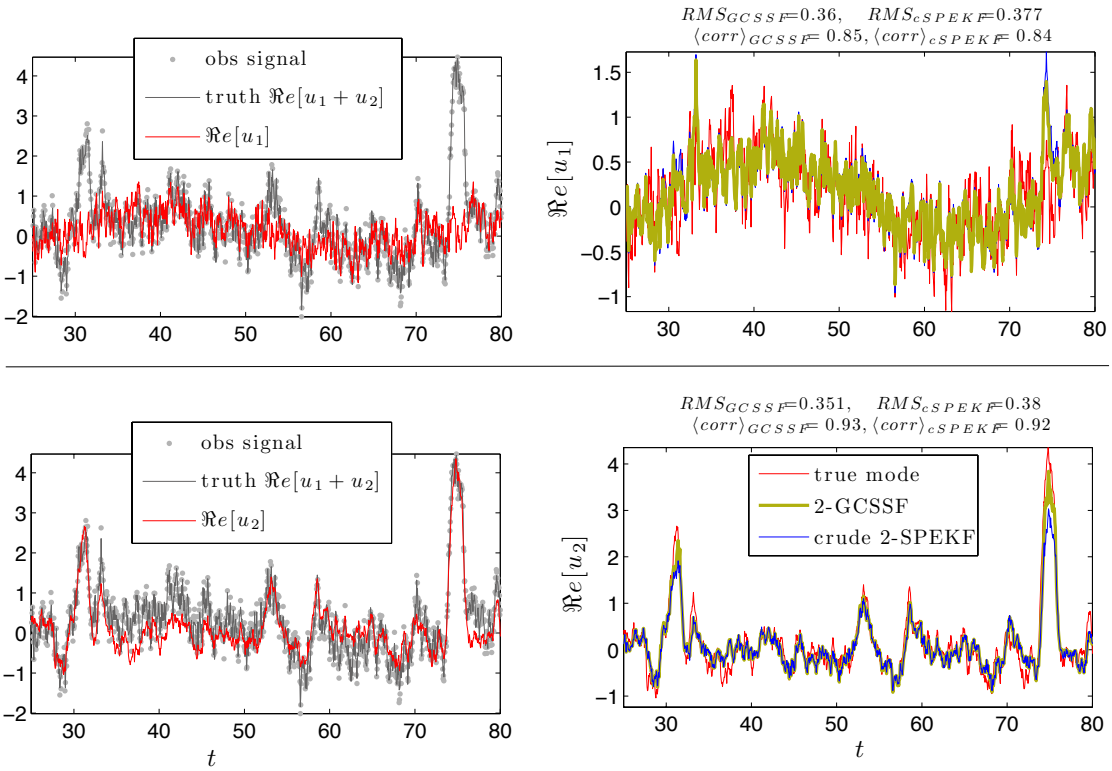

Figure 5: **Synthetic signal; equipartition of energy.** Path-wise example of DSS for a synthetic two-mode signal (regimes (III)+(II) of the system (9)) with mode energies $\boldsymbol{E_1 = E_2 = 0.187}$; the results shown are obtained from two DSS algorithms, 2-GCSSF and crude 2-SPEKF (see §2.3). DSS parameters are as in the truth modes.

True signal parameters: $[u_1]$ $\hat{\gamma} = 6, d_\gamma = 0.3, \sigma_\gamma = 1, \sigma_u = 1.5, \omega_u = 2, d_b = \sigma_b = 0.4, \omega_b = 1$;
$[u_2]$ $\hat{\gamma} = 1.1, d_\gamma = 1.2, \sigma_\gamma = 1.2, \sigma_u = 0.64, \omega_u = 1, d_b = \sigma_b = 0.4, \omega_b = 1$;

Forcing: $F_{u_1}(t) = 1 + 3\exp(i0.15t)$, $F_{u_2}(t) = 0$; *Observations*: $\Delta t^{obs} = 0.1$, $r^0 = 0.05$.

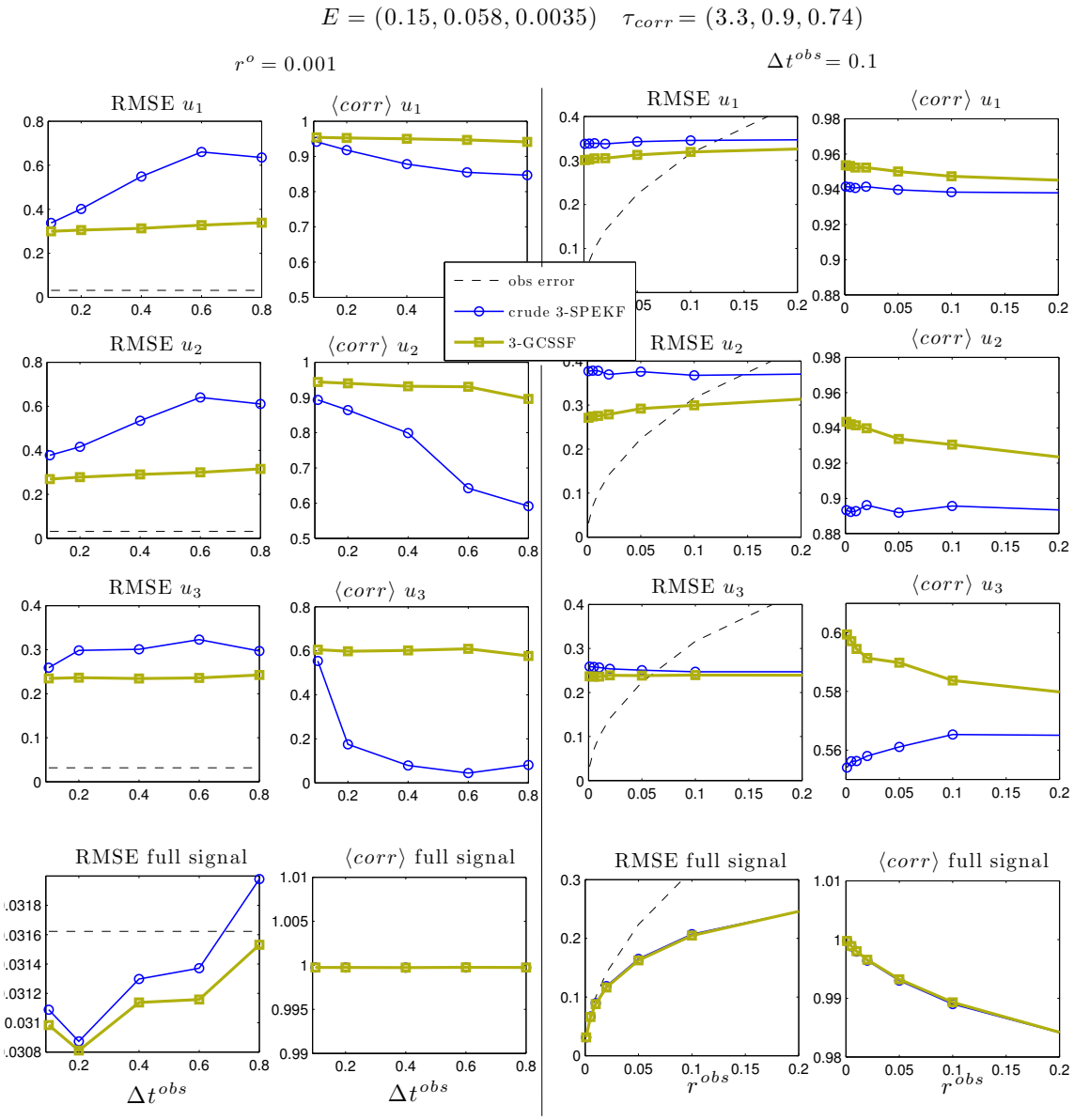

Figure 6: **Synthetic three-mode aliased signal with decaying energy.** The DSS skill in terms of RMS error and pattern correlation as a function of the observation time and the observation noise variance. The truth signal parameters are as in figure 4.

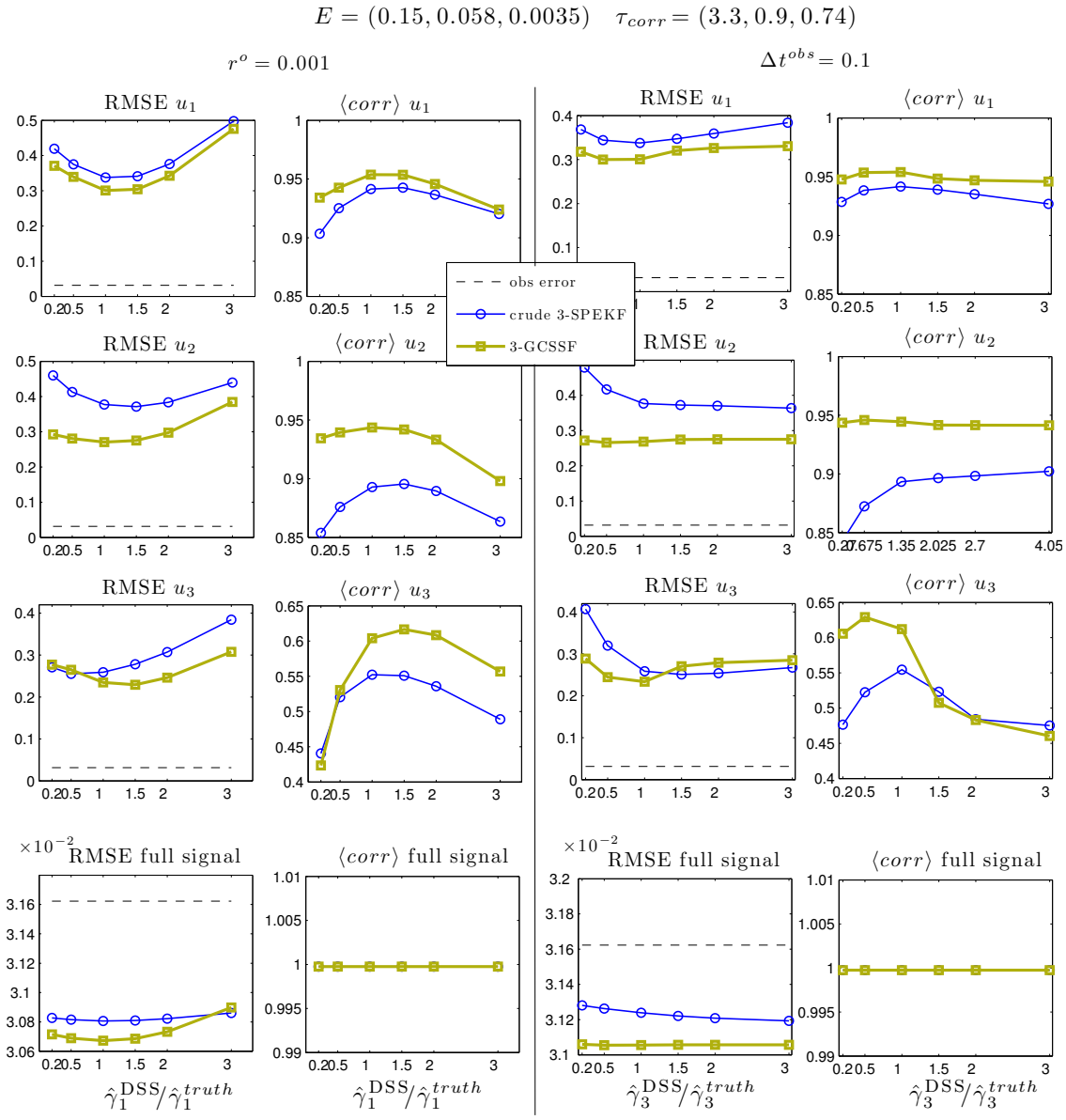

Figure 7: **Skill robustness of DSS algorithms.** The DSS skill in terms of RMS error and pattern correlation as a function of the mean damping parameters $\hat{\gamma}_1^{\mathrm{DSS}}, \hat{\gamma}_3^{\mathrm{DSS}}$ used in the DSS algorithms. Synthetic three-mode aliased signal with decaying energy and parameters as in figure 4 was used in the computations. The skill sensitivity to errors in the remaining parameters is even weaker and not shown.

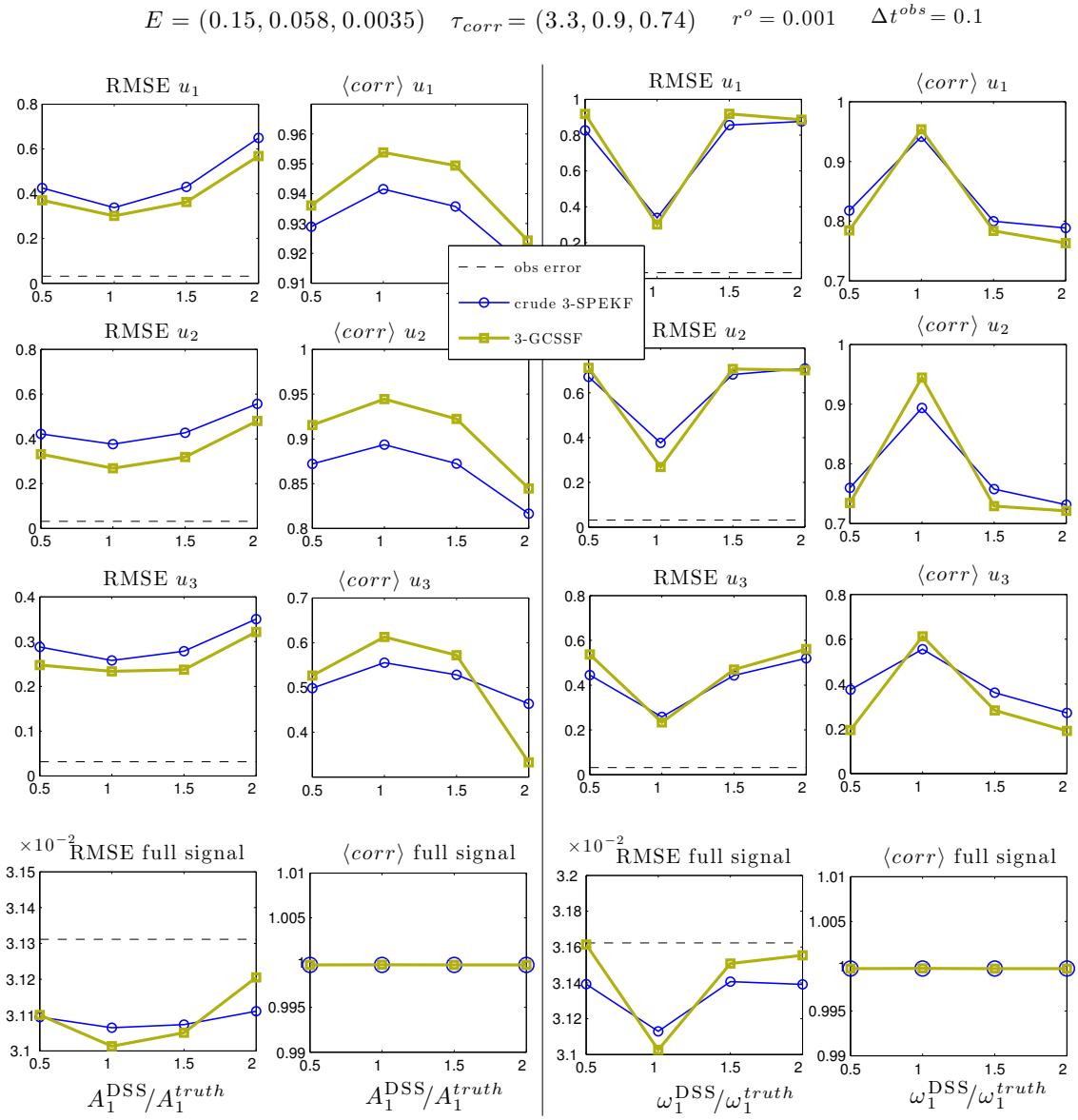

Figure 8: **Skill robustness of DSS algorithms.** The DSS skill in terms of RMS error and pattern correlation as a function of the forcing parameters of the primary mode u_1 used in DSS ($F_{u_1}^{\mathrm{DSS}} = C_1^{\mathrm{DSS}} + A_1^{\mathrm{DSS}} \exp(i\omega_1^{\mathrm{DSS}} t)$). Synthetic three-mode aliased signal with decaying energy and parameters as in figure 4 was used in the computations. The skill sensitivity to errors in the constant component, C_1^{DSS}, are even weaker and not shown.

4 DSS in nonlinear turbulent spatially extended systems

Here, we focus on the inherently difficult and important practical issue of capturing the relevant characteristics of the unresolved processes in nonlinear turbulent spatially extended systems which are observed on a sparse regular grid, as outlined in §2.1. It is particularly important to examine the skill of DSS in this realistic setting since, contrary to the cheap stochastic forecast models used in the DSS algorithms (§2.3), the true nonlinear dynamics implies correlations between different Fourier modes of the observed signal. As shown below, this model error introduced in the DSS algorithms is not detrimental to the DSS skill since the aliasing Fourier modes are expected to be weakly correlated in realistic geophysical systems with a large number of degrees of freedom, as already argued in the Introduction and Appendix A. In order to elucidate these subtle issues in a realistic setting, we study the performance of the DSS algorithms on two radically different turbulent spatially extended nonlinear systems. The first system is given by the one-dimensional truncated Burgers-Hilbert (TBHi) equation [4, 10] with quadratic nonlinearity which belongs to a class of spectrally truncated energy-conserving Burgers-Hopf type systems [35, 23, 31, 25] with a remarkable statistical behavior. The TBHi system offers a very attractive testbed for DSS since one of its turbulent dynamical regimes is characterized by the presence of a large-scale traveling wave with small-scale turbulent waves 'riding' the large-scale wave (see figure 9). The second system, referred to hereafter as MMT [27, 8, 7, 43], is a one-dimensional model for forced dissipative and dispersive wave turbulence. The desirable characteristics of the MMT dynamics for our purposes include (i) the existence of turbulent cycles induced by the interaction of spatially localized coherent structures with dispersive wave radiation (see figure 10) creating strong direct and inverse cascades [7] and (ii) the presence of the cubic nonlinearity which implies significant model error when filtering this system with the forward models derived from the quadratic model (9).

The main points highlighted in this section include:

- Significant gain in skill for recovering the primary modes can be achieved when superresolving all modes in the aliasing sets which are above the noise level.
- There exist important skill differences between the DSS algorithms introduced in §2.3 which are associated with the various approximations of the cross-correlations between the modes in the aliasing sets.
- GCSSF has the best skill for DSS skill, both path-wise and in terms of the RMS error and pattern correlation, for recovering the most energetic unresolved modes in the aliasing sets.
- While all the DSS algorithms introduced in §2.3 are robust to parameter uncertainties in their respective forecast models, GCSSF retains the best skill.
- Inaccuracies in estimating the cross-correlations between the aliasing modes via the Monte Carlo sampling in MC mSPEKF are particularly detrimental to its skill at short observation times (relative to the shortest decorrelation time of unresolved modes in the aliasing set). For increasing observation times the skill of all the DSS algorithms in our suite (§2.3) becomes comparable.
- The use of the stochastic non-Gaussian forward models (9) is crucial for the skill and robustness of DSS. The use of a Gaussian Mean Stochastic Model (MSM) results in divergent filtering algorithms when filtering with Δt^{obs} beyond the decorrelation times of aliased modes.

4.1 The nonlinear spatially extended turbulent test models

The two nonlinear turbulent spatially extended systems used for testing the DSS algorithms introduced in §2.3 are briefly outlined below. Application of the DSS algorithms which exploit the quadratic stochastic forward model (9) for signals generated by these two systems with, respectively, quadratic and cubic nonlinearity is discussed in §4.2.

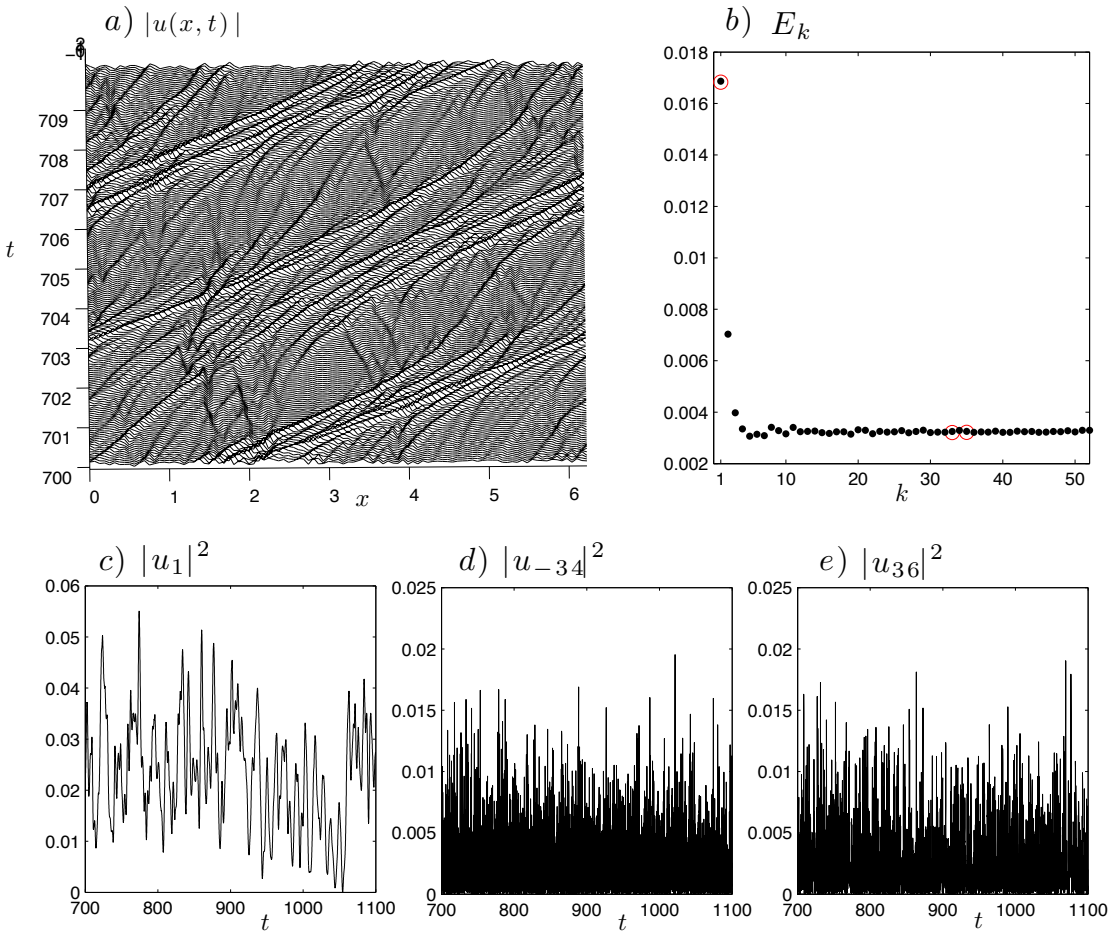

Figure 9: Typical dynamics of the Truncated Burgers-Hilbert (TBHi) system (§4.1.1) with $\Lambda = 52$. a) Evolution of the solution $u(t, x)$ in the physical space, b) Near-equipartition of the Fourier modes in the truncated system except for the few large scale modes, c)-d) Evolution of three aliasing Fourier modes in the aliasing set $\mathcal{A}(1)$ associated with a sparse observation network with $P = 3$ and $\tilde{N} = 105, \tilde{M} = 35$ (see §2.1). The initial condition with energy $E_0 = 0.17$ is $u_0(x) \propto \sin(3x) + 0.2\sigma$, where σ is a Gaussian random vector with zero mean and unit variance.

4.1.1 Truncated Burgers-Hilbert equation

The model studied here is the Galerkin truncated spectral approximation to the inviscid Burgers-Hopf equation with suitable dispersive terms added via the Hilbert transform

$$u_t + \frac{1}{2}(u^2)_x = \alpha H[u], \tag{13}$$

where α is a constant parameter and the operator H denotes the spatial Hilbert transform defined in the Fourier space by

$$\hat{H}[f](k) = -i\,\mathrm{sgn}(k)\hat{f}(k). \tag{14}$$

For $\alpha = 0$ equation (13) becomes the well known Burgers-Hopf model whose truncations (TBH) were introduced and studied earlier in [35, 23]. The Burgers-Hilbert equation with $\alpha \neq 0$ in (13) was studied in detail in [4] as a model for unidirectional constant-frequency dispersive waves. The truncated form of (13) is constructed by projecting the dynamics of (13) on a finite number of Fourier modes in periodic

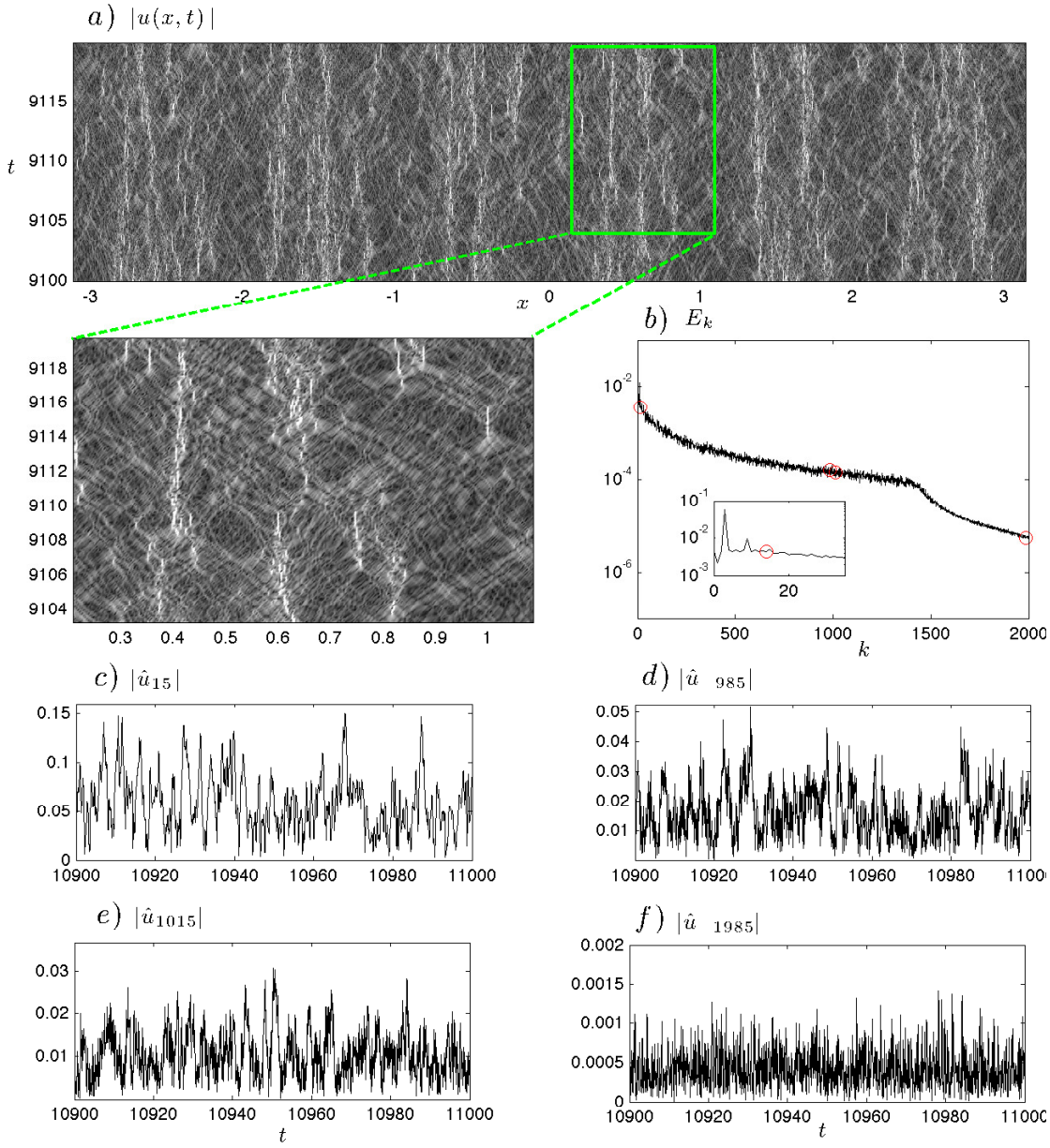

Figure 10: Typical dynamics of the MMT system (§4.1.2) in the focusing case solved using the pseudospectral method on a mesh with $\tilde{N} = 4000$ nodes. a) Evolution of the solution $|u(t,x)|$ in the physical space, b) Energy spectrum corresponding to the forcing and dissipation specified in §4.1.2, (c)-(d) Evolution of four aliasing Fourier modes in the aliasing set $\mathcal{A}(15)$ associated with a sparse observation network with $P = 4$ and $\tilde{N} = 4000, \tilde{M} = 1000$ (see §2.1).

geometry which leads to

$$(u_\Lambda)_t + \frac{1}{2} P_\Lambda (u_\Lambda^2)_x = \alpha H[u_\Lambda], \tag{15}$$

where P_Λ is a projection operator that restricts the wave number to values with absolute values not exceeding Λ. The truncated dynamics of (15) exhibits many of the desirable properties found in more complex systems but has the virtue of allowing a relatively complete analysis of statistical properties and extensive numerical studies. As shown in [31, 10] the truncated dynamics of (15) is deterministic but chaotic and mixing, it is ergodic on suitably defined equi-energy surfaces, and the time correlations for the various degrees of freedom obey a simple scaling law with a slower decay of correlations for the large-scale features of the system. It can be also easily shown that (15) has the Liouville property and conserves both the linear momentum and energy so that it has a Gaussian equilibrium measure [31, 25, 10].

Here, we consider the dynamics of the truncated Burgers-Hilbert model (TBHi) which provides an unambiguous test for DSS in turbulent spatially extended systems with non-trivial interactions between unequilibrated large scale traveling-wave modes and smaller scale thermalized turbulent modes [10] (see figure 9a,b). The TBHi dynamics is obtained by setting $\alpha = 1$ in (15); alternative normalization can be obtained by normalizing (15) with the energy of the solutions in which case the TBH dynamics with no traveling modes is approached in the limit of large energies.

There exist two distinct dynamical regimes of (15) which are controlled by the energy E of the solutions with E_α the critical energy separating the two regimes [4, 10]. It is well known that nontrivial smooth solutions of the full system (13) with sufficiently large energies, $E > E_\alpha$, develop discontinuities in finite time and thus exhibit a transfer of energy from large scales to small scales. Once this transfer develops in a general solution for the truncated system (15), the conservation of energy constraint forces an inverse cascade to the large scale modes. The important feature of this process for our purposes is that it does not radically alter the large-scale profile of the solutions; instead the redistribution of energy results in appearance of small amplitude turbulent waves riding on a weakly chaotic large-amplitude wave. In figures 9c-e we show a few snapshots of the solutions of (15) in the turbulent regime; as shown in figure 9b, energy equipartition holds approximately for all but the first few large scale modes associated with the traveling wave solutions. Moreover, the decorrelation times effectively lie on the hyperbolae $\tau_k^{corr} = \sqrt{\Lambda/E}\, k^{-1}$ [35, 23, 10] with the large scale modes having the longest decorrelation time. For $E < E_\alpha$ the TBHi dynamics remains smooth.

In the numerical examples discussed in the next section we used a pseudospectral method to numerically integrate both the truncated Burgers-Hilbert equation (15) with $\Lambda = 52$ and random initial conditions with energy $E_0 \approx 0.18$. The spatial derivatives and the Hilbert transform were computed spectrally and the time integration was carried out by means of a MATLAB's 'ode45' routine which is similar to the fourth-order Runge-Kutta scheme with adaptive time stepping. In the computations the energy was conserved to within a relative error of less than 10^{-5} for the duration of the integration which was carried out until the statistically steady state was reached.

4.1.2 The MMT equation

The MMT model [27, 8, 7, 43] for a one-dimensional dispersive wave turbulence is given by

$$iu_t = |\partial_x|^{\frac{1}{2}} u + \lambda |u|^2 u - iAu + F, \tag{16}$$

where the dissipation is controlled by the operator A defined in the Fourier space below, and F is the forcing. This one-dimensional forced dissipative model with cubic nonlinearity is drastically different from the energy-conserving dynamics of the TBHi system with quadratic nonlinearity and near-equipartition of energy. In MMT excitations at one spatial scale (usually a long spatial scale) are injected into the system by the external forcing F and are removed by dissipation which is restricted to a different much smaller spatial scale. At the intermediate scales the solutions of (16) satisfy a conservative Hamiltonian system and wave-wave interactions transport the excitations, setting up a statistically steady turbulent

state from the injection to the dissipation scales. The rich mathematical structure of this model allowed the authors in [27, 8, 7] to study the validity of the direct and inverse spectra of the weak turbulence theory and determine new dynamical regimes with the so-called MMT spectrum, as well as transitions between these spectra depending upon details of nonlinearity, forcing, and dissipation.

The sign of the parameter $\lambda = \pm 1$ in (16) is important for the nature of the turbulent dynamics. Here, we consider the more interesting case with the *focusing* nonlinearity, $\lambda = -1$, which induces spatially coherent 'solitonic' excitations at random spatial locations (see figure 10a). Spatial collapse of these coherent structures initiates inverse cascade of energy by transferring energy from small scales to large scales via wave radiation [7]. These fast processes transfer excitations from long spatial scales to shorter ones, and initiate turbulent cycles with coexisting direct and inverse cascades within these cycles. This type of forced, dissipative and dispersive dynamics lends itself as another tough test model for DSS in nonlinear turbulent spatially extended systems, which is very different and complementary to the TBHi dynamics discussed in the previous section with large scale traveling wave and nearly-thermalized smaller scale turbulent modes. For our purposes, there are two particularly desirable characteristics of this model dynamics (i) the existence of turbulent cycles (figure 10a) induced by the interaction of spatially localized coherent structures with dispersive wave radiation, and (ii) the presence of cubic nonlinearity which implies significant model error when filtering with the forward models derived from the quadratic system (9).

The numerical examples illustrating the skill of DSS in this system discussed in the next section were obtained via the pseudospectral method with deterministic forcing at the long spatial scales with $F = 0.75\sin(3x)$, random initial conditions with energy $E_0 \approx 0.02$, and selective damping on short and large scales; the damping is specified in the Fourier space as

$$\hat{A}_k = \begin{cases} a_k & \text{for } k = 1, 2, \\ b_k\big(|k_{thr}|^2 - |k|^2\big) & \text{for } k > k_{thr}, \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

where the coefficients $a_k = 1$, $b_k = 10^{-4}$ and $k_{thr} = 1400$. The total number of Fourier modes in the computations used in the examples below is 4000. The time integration was carried out by means of a MATLAB's 'ode45' routine, an adaptive step-size Runge-Kutta integrator. The equilibrium spectrum is shown in figure 10b, alongside the four modes (figure 10c-f) in the aliasing set $\mathcal{A}(15)$ used for testing the DSS algorithms in the next section.

4.2 Performance of DSS algorithms on the spatially extended test models

Here, we discuss the skill of the DSS algorithms (§2.3) for superresolving noisy turbulent signals from the sparsely observed nonlinear spatially extended test models, TBHi and MMT, discussed above. The dynamics of the aliasing modes in a single aliasing set (see §2.1) is considered, similarly to the case of synthetic signals discussed in §3.

The initial step in DSS involves setting the parameters in the forecast model (9) and it is related to the case of filtering with SPEKF [13, 12, 20, 22]. Following a similar procedure to that discussed in [20], we use the Mean Stochastic Model (MSM) for each aliasing mode

$$\mathrm{d}u_k^{\mathrm{MSM}} = -(\gamma_k^{\mathrm{MSM}} + i\omega_k^{\mathrm{MSM}})u_k^{\mathrm{MSM}}\mathrm{d}t + \sigma_k^{\mathrm{MSM}}\mathrm{d}W_k^{\mathrm{MSM}}, \tag{18}$$

in order to estimate the mean values for the multiplicative and additive bias corrections for each of the aliased modes in the forecast models of the DSS algorithms (§2.3), i.e., we set

$$\hat{\gamma}_k = \gamma_k^{\mathrm{MSM}}, \quad \hat{\omega}_k = \omega_k^{\mathrm{MSM}}, \quad \hat{b}_k = 0, \tag{19}$$

in (9). In this way the DSS algorithms adaptively correct the MSM parameterization through the observations of the aliased noisy signal. In principle, the MSM parameters are chosen to reproduce the energy

spectrum and the decorrelation times of the truth signal and are obtained via

$$\hat{\gamma}_k^{\text{MSM}} + i\omega_k^{\text{MSM}} = \int_0^\infty \langle u_k(t)u_k^*(t+\tau)\rangle \mathrm{d}\tau, \quad \sigma_k^{\text{MSM}} = \sqrt{2E_k\hat{\gamma}_k^{\text{MSM}}}, \tag{20}$$

which can be easily manipulated to obtain (see [18])

$$\hat{\gamma}_k = \frac{\hat{\gamma}_k^{\text{MSM}}}{(\hat{\gamma}_k^{\text{MSM}})^2 + (\hat{\omega}_k^{\text{MSM}})^2}, \quad \hat{\omega}_k = -\frac{\hat{\omega}_k^{\text{MSM}}}{(\hat{\gamma}_k^{\text{MSM}})^2 + (\hat{\omega}_k^{\text{MSM}})^2}. \tag{21}$$

At a first glance, there are two potential limitations of this procedure in the DSS framework which are, however, remedied by the subsequent 'on the fly' adaptive parameter estimation in the DSS algorithms. Firstly, the long-time climatological statistics required in (20) has to be estimated from time series on a finite time interval; consequently, these statistics will be only partially or imperfectly known in practice, resulting in biases in (20). Secondly, given that the available time series contain the scrambled information about the primary and the aliased modes, some additional physical information is needed for estimating the above parameters for the aliased modes (i.e., estimate of the energy spectrum and decorrelation times). However, it turns out only rough estimates for these parameters are needed in our algorithms to achieve nearly optimal DSS skill, as we show below based on numerical examples from two very different spatially extended systems. The remaining parameters in the DSS algorithms associated with the dynamics of the bias correction terms in the augmented forecast model (9), i.e., $d_{\gamma_k}, d_{b_k}, \omega_{b_k}, \sigma_{\gamma_k}, \sigma_{b_k}$, do not have direct interpretation in terms of the statistical characteristics of the turbulent physical system. However, similarly to the results for filtering with the single SPEKF model [13, 12, 5], there exist wide regimes of parameter values for which the DSS skill is near-optimal and robust, as discussed later in this section. Unless explicitly stated, the values of these additional parameters in the DSS algorithms are set to

$$d_{\gamma_k} = d_{b_k} = 0.1\hat{\gamma}_k, \quad \omega_{b_k} = \omega_k, \quad \sigma_{\gamma_k} = \sigma_{b_k} = 4\sigma_k, \tag{22}$$

which correspond to strengthening the role of the stochastic noise and weakening the damping in the dynamics of the stochastic bias correction terms.

In figures 11-18 we illustrate the skill of the DSS algorithms (see §2.3) for recovery of the aliased modes from sparse observations of the TBHi and MMT systems; for both test models, the parameter values in the DSS algorithms are estimated as described above. The skill of DSS is illustrated both in terms of the rough statistical measures (RMS errors and pattern correlation), as well as in the path-wise sense. Despite the significantly different turbulent dynamics of the two spatially extended test models, the results are similarly good, pointing to the versatility of these DSS algorithms combining forecast models with judicious model error and 'on the fly' stochastic parameter estimation.

Here, for the dynamics of TBHi we consider the aliasing set of $k = 1$ observed on a sparse grid with $P = 3$ (see §2.1) which is given by $\mathcal{A}(k{=}1) = \{1, -34, 36\}$ (see figure 9). The primary, large-scale mode in this aliasing set corresponds to the traveling wave solution and does not equilibrate (see figures 9a,b,c); the remaining two modes in the aliasing set have much smaller spatial scales and nearly equal decorrelation times which are significantly shorter than the decorrelation time of the the primary mode u_1. This is a tough test bed for the DSS where the unresolved modes decorrelate much faster compared to the primary mode yet they contain non-negligible energy.

Figure 11 illustrates the DSS skill in terms of RMS error and pattern correlation for the TBHi system as functions of the observation time Δt^{obs} and the observation noise variance r^{obs}. A particularly important property of the GCSSF algorithm is illustrated in figure 12 which shows the RMS errors and pattern correlation for DSS of the three aliased modes and for SPEKF filtering of the primary mode alone. In the latter scenario the aliased modes are not actively filtered and are treated by the filter as noise. Note that active filtering of all the aliasing modes in DSS significantly improves the skill for estimating the

primary mode. The robustness of the estimates obtained from the DSS algorithms for the dynamics of the aliasing modes is illustrated in figures 13 where we show the RMS errors and pattern correlation as functions of the parameters of the forward model (the most relevant parameters are shown for brevity). Finally, figure 14 shows an example of path-wise recovery of the primary and secondary mode from the noisy observations which scramble together the dynamics of all the aliased modes; note that the damping fluctuations $\gamma_{1,2,3}$ and the phases $\omega_{1,2,3}$ are essentially constant. It is also worth pointing out, although not shown here, that attempting DSS with MSM as the forward model instead of the non-Gaussian model (9) leads to divergent algorithms when filtering with observation times, Δt^{obs}, exceeding decorrelation times of the remaining modes in the aliasing sets.

The analysis of the DSS algorithms for the MMT dynamics is illustrated on the aliasing set of $k = 15$ observed on a sparse grid with $P = 4$ (see §2.1) which is given by $\mathcal{A}(k{=}15) = \{15, -985, 1015, -1985\}$ (see figure 10). The primary, large-scale mode in this aliasing set is the most energetic one but it is sufficiently well separated from the energy injection scale at $|k| = 3$, as shown in figure 10b. Here, the energy of the mode u_{-1985} in the aliasing set is below the observation noise variance and we do not actively resolve it. Similar to the TBHi dynamics this is another tough test bed for the DSS where the unresolved modes decorrelate fast compared to the primary mode yet they contain non-negligible energy; we define the shortest decorrelation time of a mode within the aliasing set of wavenumber k as

$$\tau_{\mathcal{A}}^{\min} \equiv \min_{\substack{u_i \in \mathcal{A}(k) \\ Var[u_i] > r^o}} \tau_{u_i}. \tag{23}$$

Figures 15-18 summarize the analysis of the skill of the DSS for the MMT system. In the computations, we used the focusing case of the MMT dynamics (i.e., $\lambda = -1$ in (16)) which was solved using the pseudo-spectral method with 4000 modes. Figure 15 illustrates the DSS skill (RMSE and pattern correlation) as functions of the observation time Δt^{obs} and the observation noise variance r^{obs}. The robustness of the estimates for the aliased modes obtained from DSS algorithms is similar to that shown for the TBHi system. Similarly to the case of TBHi dynamics, active filtering of all the aliasing modes above the noise level significantly improves the skill for estimating the primary mode, as shown in figure 18. Finally, figures 16-17 show an example of path-wise recovery of the primary and secondary mode from the noisy observations scrambling all the aliased modes; note that the damping fluctuations $\gamma_{1,2,3}(t)$ are essentially constant while the phase fluctuations $\omega_{1,2,3}(t)$ show significant variability due to the strongly dispersive dynamics of the MMT system.

Based on the results illustrated in figures 11-18, we stress the following important points about DSS in turbulent spatially extended systems:

- Assume $\tau_{\mathcal{A}}^{\min}$ and $\tau_{\mathcal{A}}^{\max}$ to be, respectively, the smallest and the largest decorrelation times among those modes in the aliasing set $\mathcal{A}$ which are above the noise level; then
  - the skill of the DSS algorithms for $\Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\min}$ satisfies

    GCSSF > crude mSPEKF >MC mSPEKF (ens. size $N \sim \mathcal{O}(100)$);
  - DSS when $\tau_{\mathcal{A}}^{\min} \ll \Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\max}$: all filters are comparable.
- Larger ensemble size in MC mSPEKF than $N = 100$ used here improves its performance but results in a rapidly increasing computational cost.
- GCSSF has the best skill (in terms of RMS error and pattern correlation) for recovering the primary mode when filtering with $\Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\min}$. The pattern correlation remains good for GCSSF and crude SPEKF when superresolving with Δt^{obs} exceeding $\tau_{\mathcal{A}}^{\min}$ (figures 11, 12, 15).
- DSS is robust and near-optimal near the MSM parameter values (e.g., figure 13).

$$\Delta t^{obs} = 0.1, \quad r^o = 0.003, \quad E_1 = (0.017, 0.003, 0.003) \quad \tau_{corr} = (3.5, 0.1, 0.1)$$

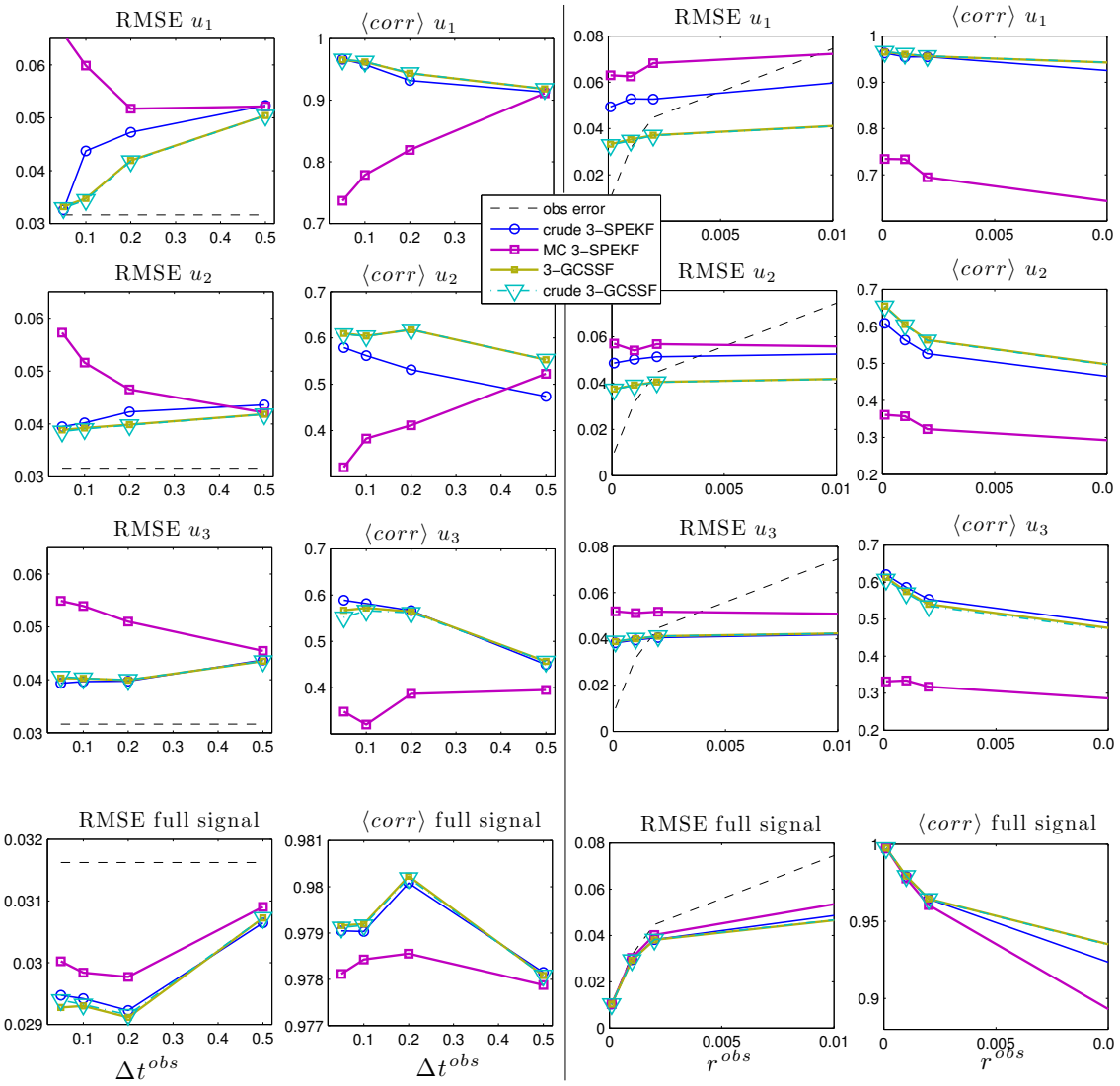

Figure 11: Performance of DSS algorithms in terms of RMS error and pattern correlation for the TBHi system (§4.1.1) as a function of the observation time and observation noise variance; single three-mode aliasing set of the wavenumber $\boldsymbol{l} = 1$ (the traveling wave mode) is considered with $P = 3$ (§2.1).

– Aliased mode energies: $E_1 = 0.017, E_2 = 0.003, E_3 = 0.003$

– Correlation times $\tau_1 = 3.5, \tau_2 \approx 0.1, \tau_3 \approx 0.1$.

– Dampings $\hat{\gamma}_1$=$2.5\gamma_1^{MSM}$, $\hat{\gamma}_2$=γ_2^{MSM}, $\hat{\gamma}_3$=γ_3^{MSM},$d_{\gamma_{(1,2,2)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{b_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{\omega_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$

– Remaining noise amps: $\sigma_{\gamma_{1,2,3}}$=$0.2\sigma_{u_{1,2,3}}$, $\sigma_{b_{(1,2,3)}}$=$0.3\sigma_{u_{(1,2,3)}}$, $\sigma_{\omega_{(1,2,3)}}$=$\sigma_{u_{(1,2,3)}}$,$\sigma_{u_1}$=$2.5\sigma_{u_1}^{MSM}$, σ_{u_2}=$\sigma_{u_2}^{MSM}$, σ_{u_3}=$\sigma_{u_3}^{MSM}$

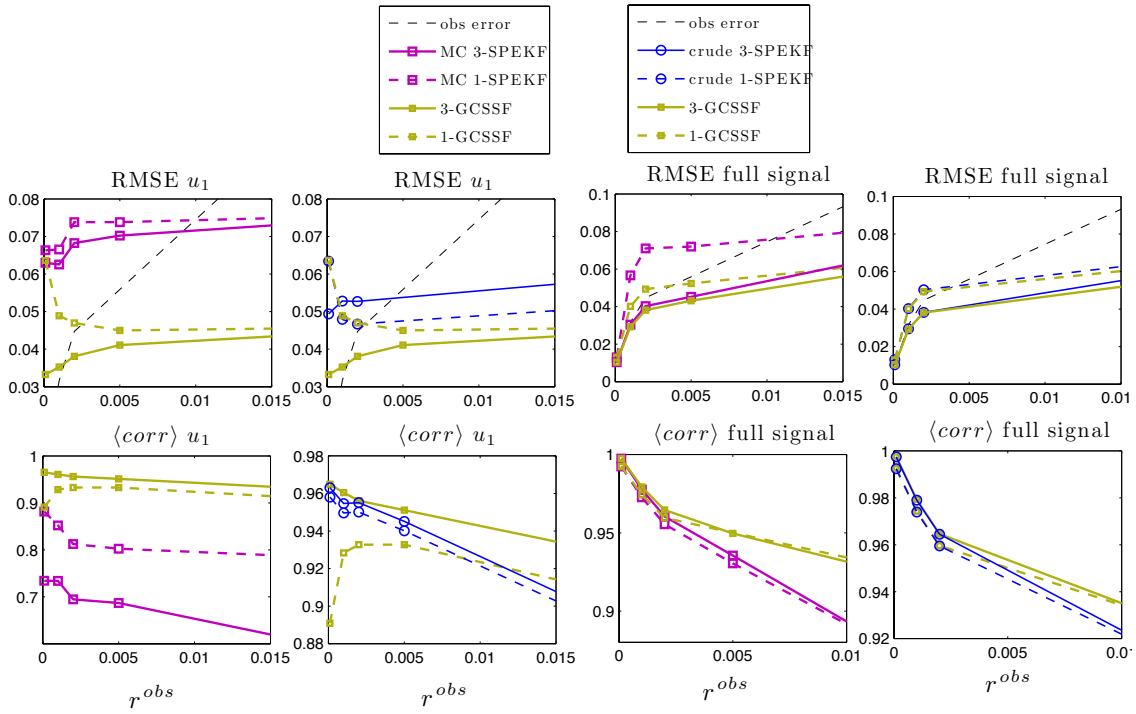

Figure 12: DSS improves the recovery of the primary mode associated with the sparse observations. Comparison of the DSS results and one-mode filtering for recovering mode u_1 of TBHi with $\Lambda = 52$ from an aliased signal involving two other modes u_1, u_{-34}, u_{36} (see figure 9). DSS actively filters all three modes while the filtering algorithms 1-GCSSF, MC 1-SPEKF, and crude 1-SPEKF filter only the primary mode.

– Aliased mode energies: $E_1 = 0.017, E_2 = 0.003, E_3 = 0.003$

– Correlation times $\tau_1 = 3.5, \tau_2 \approx 0.1, \tau_3 \approx 0.1$.

– Dampings $\hat{\gamma}_1$=$2.5\gamma_1^{MSM}$, $\hat{\gamma}_2$=γ_2^{MSM}, $\hat{\gamma}_3$=γ_3^{MSM},$d_{\gamma_{(1,2,2)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{b_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{\omega_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$

– Remaining noise amps: $\sigma_{\gamma_{1,2,3}}$=$0.2\sigma_{u_{1,2,3}}$, $\sigma_{b_{(1,2,3)}}$=$0.3\sigma_{u_{(1,2,3)}}$, $\sigma_{\omega_{(1,2,3)}}$=$\sigma_{u_{(1,2,3)}}$,$\sigma_{u_1}$=$2.5\sigma_{u_1}^{MSM}$, σ_{u_2}=$\sigma_{u_2}^{MSM}$, σ_{u_3}=$\sigma_{u_3}^{MSM}$

$$\Delta t^{obs} = 0.1, \quad r^o = 0.003, \quad E_1 = (0.017, 0.003, 0.003) \quad \tau_{corr} = (3.5, 0.1, 0.1)$$

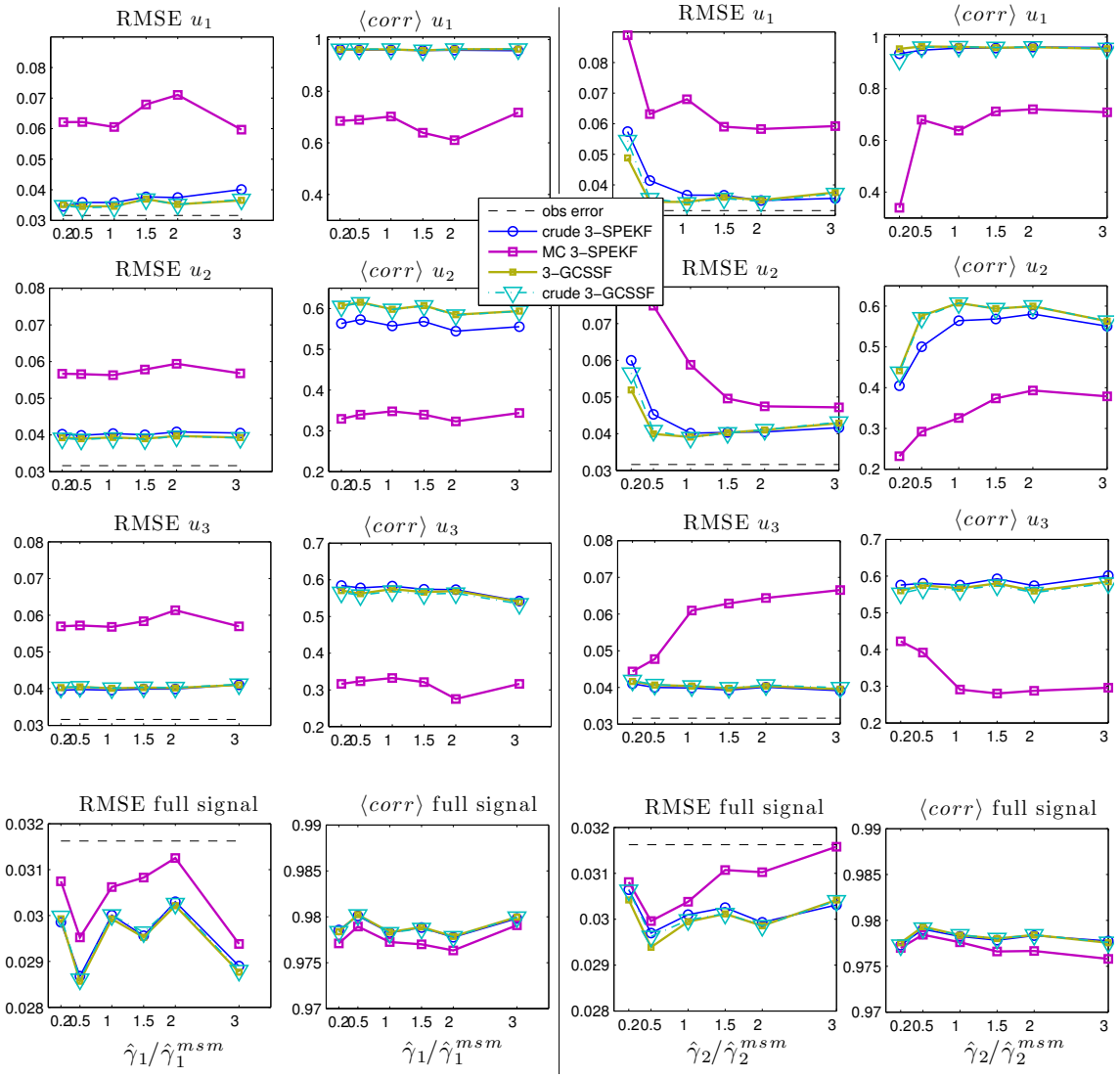

Figure 13: Robustness of DSS algorithms to parameter uncertainties in terms of RMS error and pattern correlation for the TBHi system (§4.1.1) as a function of the mean damping parameters in the forecast models; single three-mode aliasing set, $\mathcal{A}(1) = \{1, -34, 36\}$ (§2.1), is considered with $P = 3$ and $\Lambda = 52$ in (13) .

– Aliased mode energies: $E_1 = 0.017, E_2 = 0.003, E_3 = 0.003$

– Correlation times $\tau_1 = 3.5, \tau_2 \approx 0.1, \tau_3 \approx 0.1$.

– Dampings $\hat{\gamma}_1$=$2.5\gamma_1^{MSM}$, $\hat{\gamma}_2$=γ_2^{MSM}, $\hat{\gamma}_3$=γ_3^{MSM},$d_{\gamma_{(1,2,2)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{b_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$, $d_{\omega_{(1,2,3)}}$=$0.1\hat{\gamma}_{(1,2,3)}$

– Remaining noise amps: $\sigma_{\gamma_{1,2,3}}$=$0.2\sigma_{u_{1,2,3}}$, $\sigma_{b_{(1,2,3)}}$=$0.3\sigma_{u_{(1,2,3)}}$, $\sigma_{\omega_{(1,2,3)}}$=$\sigma_{u_{(1,2,3)}}$,$\sigma_{u_1}$=$2.5\sigma_{u_1}^{MSM}$, σ_{u_2}=$\sigma_{u_2}^{MSM}$, σ_{u_3}=$\sigma_{u_3}^{MSM}$

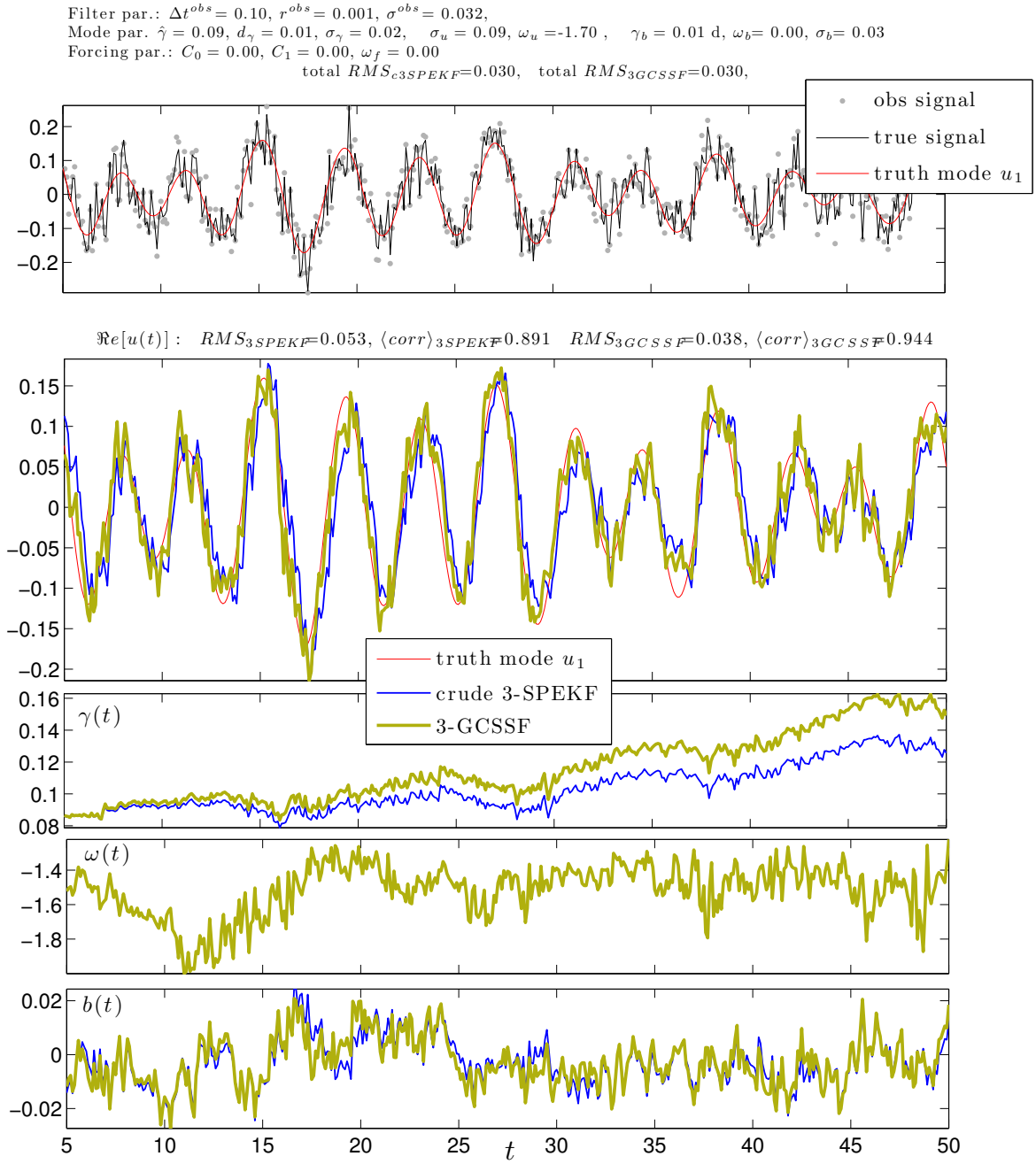

Figure 14: Path-wise example of DSS for the primary mode u_1 recovered from the aliased signal involving u_1, u_{-34}, u_{36} (see figure 9) using two different DSS algorithms 3-GCSSF and crude 3-SPEKF (see §2.3 for details).

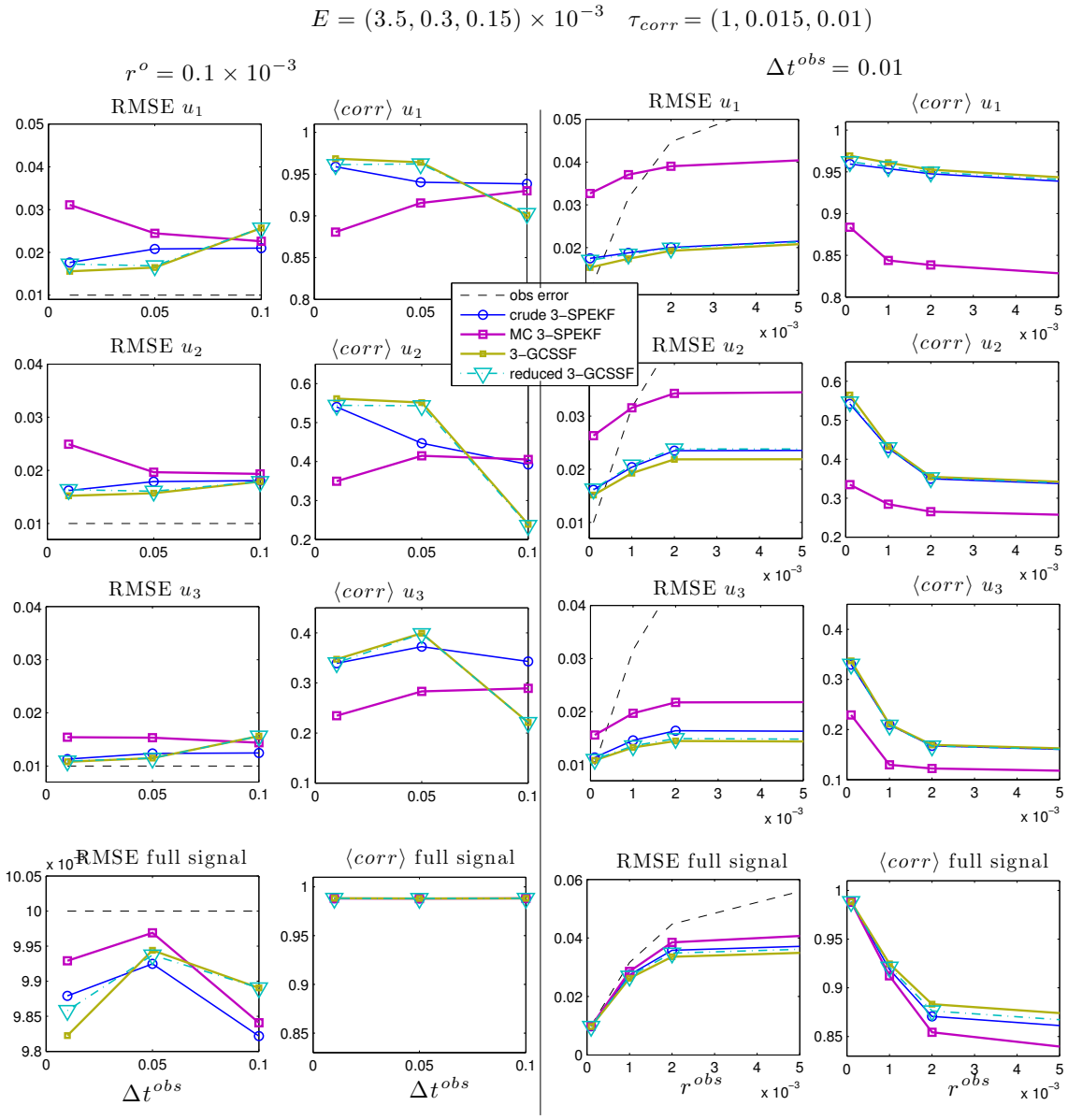

Figure 15: Performance of DSS algorithms in terms of RMS error and pattern correlation for the MMT system (§4.1.2) as a function of the observation time and observation noise variance; single three-mode aliasing set $\mathcal{A}(15)$ of the wavenumber $\boldsymbol{l} = 15$ is considered with $P = 4$ (§2.1) which corresponds to four aliased modes $u_{15}, u_{-985}, u_{1015}, u_{-1985}$ with energies $E = (3.5, 0.3, 0.15, 0.005) \times 10^{-3}$; the mode u_{-1985} is below the noise levels and not actively filtered (see also figure 10).

– Correlation times $\tau_1 = 1, \tau_2 \approx 0.015, \tau_3 \approx 0.01$.
– Dampings $\hat{\gamma}_1 = \gamma_1^{MSM}$, $\hat{\gamma}_2 = \gamma_2^{MSM}$, $\hat{\gamma}_3 = \gamma_3^{MSM}$ $\sigma_{u_1} = \sigma_{u_1}^{MSM}$, $\sigma_{u_2} = \sigma_{u_2}^{MSM}$, $\sigma_{u_3} = \sigma_{u_3}^{MSM}$
– Remaining dampings: $d_{\gamma_{(1,2,2)}} = 0.1\hat{\gamma}_{(1,2,3)}$, $d_{b_{(1,2,3)}} = 0.1\hat{\gamma}_{(1,2,3)}$, $d_{\omega_{(1,2,3)}} = 0.1\hat{\gamma}_{(1,2,3)}$
– Remaining noise amps: $\sigma_{\gamma_{1,2,3}} = 4\sigma_{u_{1,2,3}}$, $\sigma_{b_{(1,2,3)}} = 4\sigma_{u_{(1,2,3)}}$, $\sigma_{\omega_{(1,2,3)}} = 4\sigma_{u_{(1,2,3)}}$,

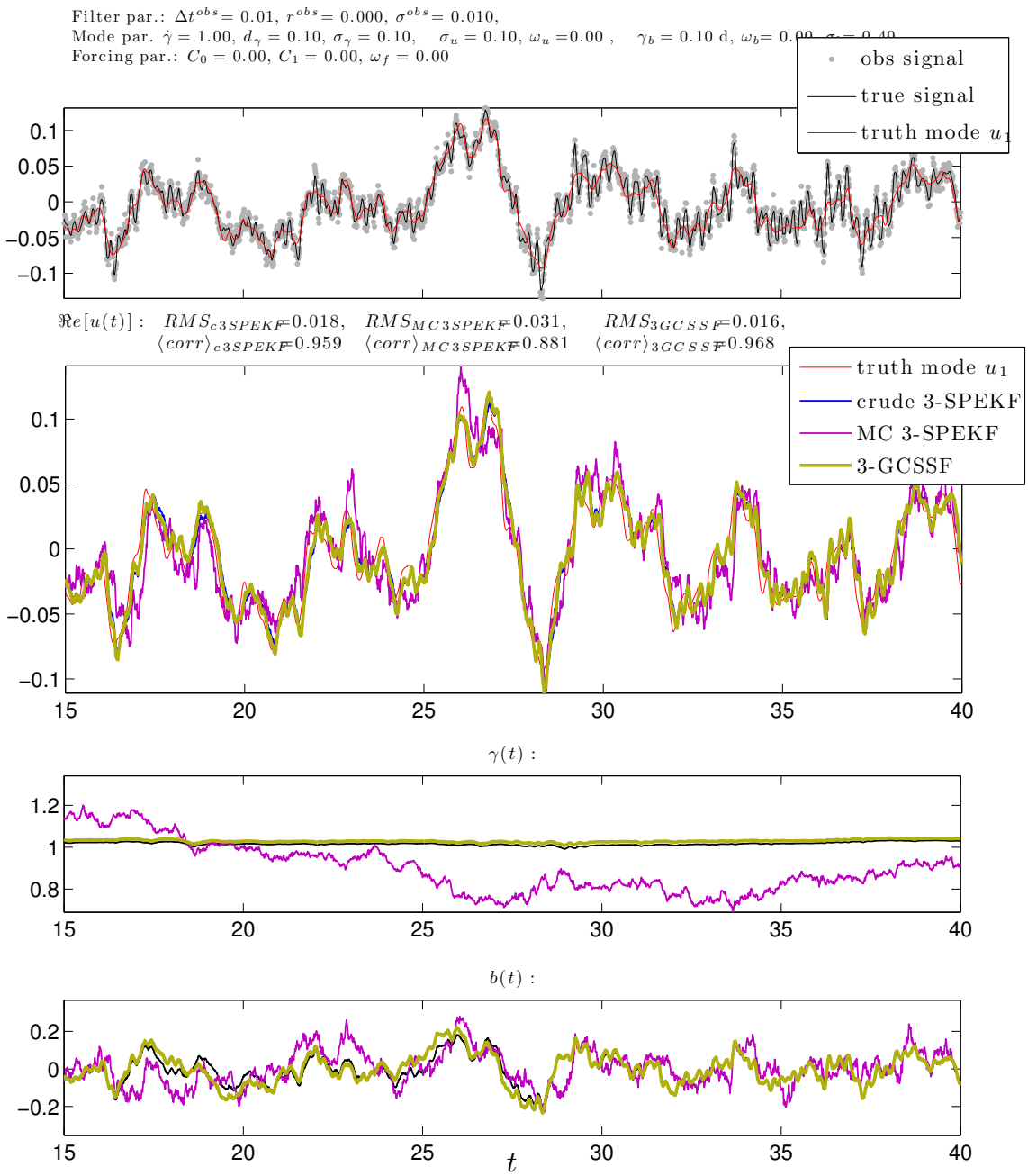

Figure 16: Path-wise example of DSS for the primary mode u_1 recovered from the aliased signal observed from the MMT system (16) and involving four modes $u_{15}, u_{-985}, u_{1015}$ with energies $E = (3.5, 0.3, 0.15, 0.005) \times 10^{-3}$ (see also figure 10) using three different DSS algorithms 3-GCSSF, MC 3-SPEKF and crude 3-SPEKF (see §2.3 for details); the mode u_{-1985} is below the noise levels and not actively filtered even in DSS. The results for the recovery of the secondary mode are shown in figure 17.

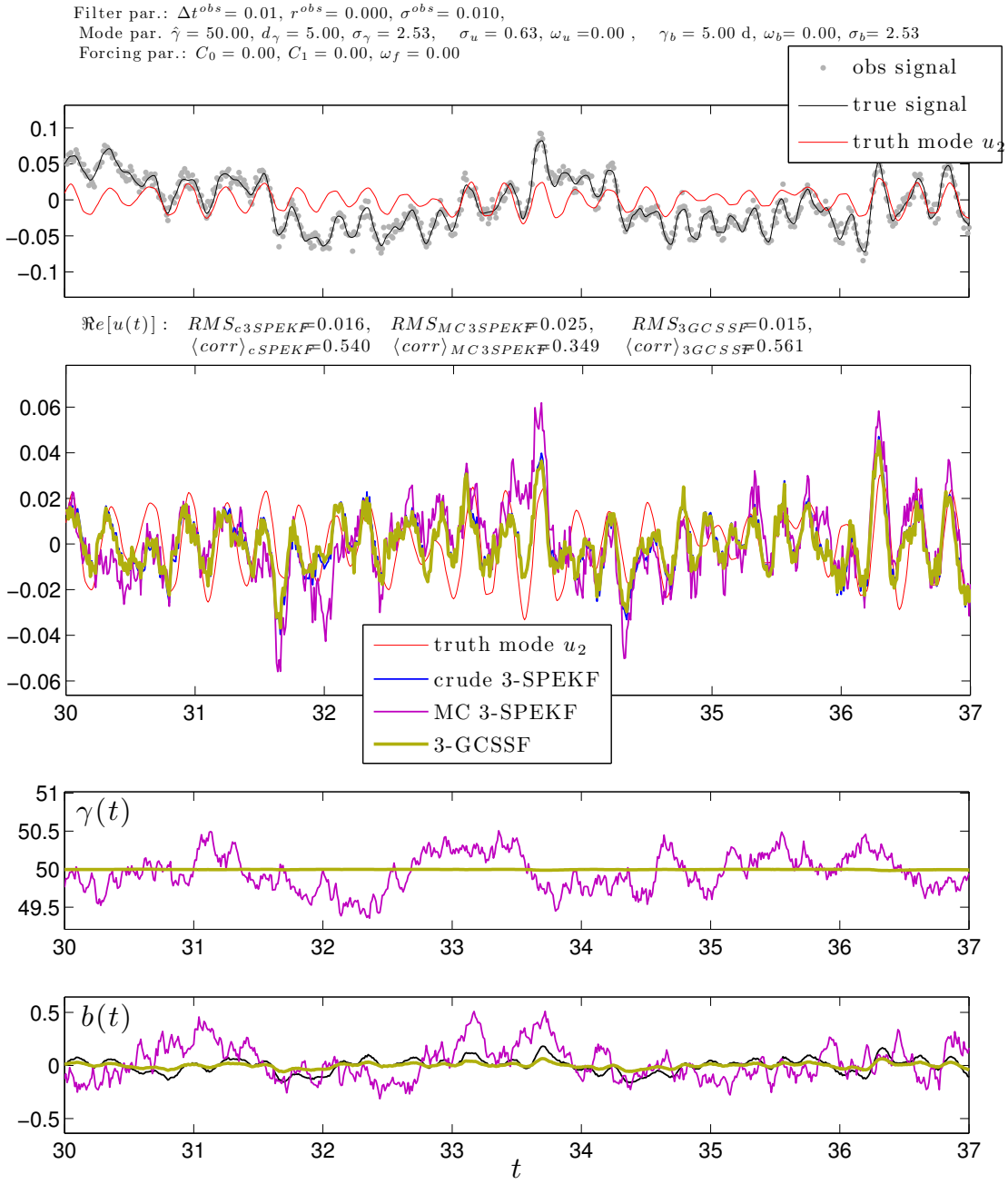

Figure 17: Path-wise example of DSS for the recovery of the secondary mode u_{-985} from the aliased signal observed from the MMT system (16) and involving four modes $u_{15}, u_{-985}, u_{1015}, u_{-1985}$ with energies $E = (3.5, 0.3, 0.15, 0.005) \times 10^{-3}$ (see also figure 10) using three different DSS algorithms 3-GCSSF, MC 3-SPEKF and crude 3-SPEKF (see §2.3 for details); the mode u_{-1985} is below the noise levels and not actively filtered even in DSS. The results for the recovery of the secondary mode are shown in figure 16.

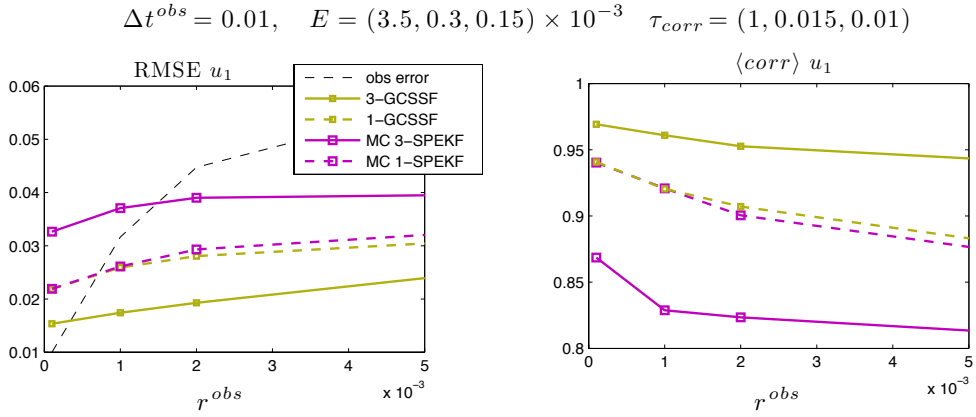

Figure 18: Improved recovery through DSS of the primary modes associated with the sparse observation grid. Comparison of the DSS results (3-GCSSF, MC 3-SPEKF) and one-mode filtering (1-GCSSF, MC 1-SPEKF) for recovering the primary mode u_{15} of the MMT system (§4.1.2) from an aliased signal involving modes $u_{15}, u_{-985}, u_{1015}, u_{-1985}$ (see figure 10); the mode u_{-1985} is below the noise levels and not actively filtered even in DSS; the filtering algorithms, 1-GCSSF and MC 1-SPEKF, filter only the primary mode.

- Significant skill improvement for resolving the primary modes is achieved when actively superresolving the aliasing modes which are above the noise levels (figures 12 and 18).
- Unsurprisingly, the skill for recovering the aliased modes is worse than the skill for estimating the primary mode (figures 11-18); however, for GCSSF the pattern correlation for the secondary mode remains above 0.6 for short to moderate observation times. For the MMT dynamics, pattern correlation of the recovered secondary modes deteriorates faster with increasing Δt^{obs} and r^{obs} than for THBi but it remains above 0.5 for short observation times.
- The skill of MC mSPEKF for recovering the primary mode improves with increasing observation time Δt^{obs} but the skill is poor at short observation times $\Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\min}$ (figures 11, 15).
- The poor performance of full MC mSPEKF when $\Delta t^{obs}/\tau_{\mathcal{A}}^{\min} \lesssim 1$ is due to the 'small ensemble size' errors in the MC estimates for the cross-correlations between the aliased modes. At short observation times, poor estimation of these cross-correlations is worse than neglecting them (MC 3-SPEKF vs crude 3-SPEKF in figure 11, 15; these sampling errors become less important for longer Δt^{obs} when the cross-correlations between the aliasing modes decay.
- MC mSPEKF generates more fluctuating bias correction terms γ, ω, b which do not however improve its performance over GCSSF with more accurate approximates the cross-correlations (figures 16, 17).
- The innovation error for filtering the full signal (the sum of all aliasing modes) is similar for all the DSS algorithms and it is close to the innovation error for filtering with a single SPEKF or GCF filter (figures 11, 12,13,15).

The above results clearly point to GCSSF as the method of choice for DSS in turbulent signals from nature. GCSSF combines high skill with a straightforward implementation which does not require the use of more accurate, but laborious, exact analytical computations of the cross-correlations between the aliasing modes.

5 Extreme case examples: Triad models

Here, for the sake of completeness of the analysis we discuss the DSS in an extremely tough turbulent case with strongly correlated modes in the aliasing set. As already discussed in the Introduction and Appendix A, relatively weak correlations between the aliasing modes should be expected in realistic applications to geophysical problems and the evidence of high DSS skill in such cases was presented in the previous section. Nevertheless, it is useful to gain some insight into the performance of DSS in such tough, exotic regimes.

The test model for this strongly correlated modes within the aliasing set is given by the turbulent triad in the form [30]

$$\begin{aligned} \frac{\mathrm{d}u_1}{\mathrm{d}t} &= -\gamma_1 u_1 + B_1 u_2 u_3 + \sigma_1 \mathrm{d}W_1(t), \\ \frac{\mathrm{d}u_2}{\mathrm{d}t} &= -\gamma_2 u_2 + B_2 u_1 u_3 + \sigma_2 \mathrm{d}W_2(t), \\ \frac{\mathrm{d}u_3}{\mathrm{d}t} &= -\gamma_3 u_3 + B_3 u_1 u_2 + \sigma_3 \mathrm{d}W_3(t), \end{aligned} \tag{24}$$

where $B_1 + B_2 + B_3 = 0$ and W_1, W_2, W_3 are independent scalar Wiener processes. The nonlinear coupling in triad systems is generic of nonlinear coupling between any three modes in larger systems with quadratic nonlinearity [28, 29]. However, the low dimension of this system implies strong couplings between the triad modes which is generally not expected for aliasing modes (see Appendix A). In this section, we consider nearly Gaussian triads with decaying energy spectrum with weak nonlinear coupling and well separated energies.

In particular, we show here the following:

- For nearly Gaussian triads the pattern correlation skill for recovery of the primary modes and some marginal skill even for the recovery of the secondary mode in the aliasing set.
- Despite the strong correlations, there is some skill for 'splitting' the aliased noisy signal in intervals of weaker correlations (figure 21).
- GCSSF better than crude SPEKF and MC mSPEKF (due to strong correlations; MC mSPEKF would need a very large ensemble)
- Unsurprisingly, the robustness to uncertainties in the filter parameters is poor in this case.

As shown in [30], the modes in the degenerate triad with $\gamma_1 = \sigma_1 = 0$ are all Gaussian with energies given by

$$E_1 = -\frac{B_1 \sigma_2^2 \sigma_3^2}{2(B_2 \gamma_2 \sigma_3^2 + B_3 \gamma_3 \sigma_2^2)}, \quad E_2 = \frac{\sigma_2^2}{4\gamma_2}, \quad E_3 = \frac{\sigma_3^2}{4\gamma_3}, \tag{25}$$

provided that the triad parameters lead to $E_1 > 0$. Analogous reasoning for non-degenerate triads implies that the modes are only Gaussian for

$$\frac{\sigma_1^2}{d_1} = \frac{\sigma_2^2}{d_2} = \frac{\sigma_3^2}{d_3}, \tag{26}$$

which corresponds to the case of equipartition of energy $E_1 = E_2 = E_3$. Here, we deliberately violate the energy equipartition constraint (26) in order to impose a decaying energy spectrum but we keep a nearly Gaussian dynamics through imposing weak nonlinearities. In the numerical examples the parameters in the DSS algorithms were chosen in the same fashion as in §4 by setting the equilibrium values for the dampings $\hat{\gamma}_k$, phases $\hat{\omega}_k$, and noise amplitudes σ_k, based on the corresponding MSM parameters. Tuning the remaining parameters in the DSS forecast models requires more care in this strongly-correlated configuration and we discuss the issues of robustness with respect to these parameters below; unless explicitly stated the parameter values used in the computations are set to

$$d_{\gamma_k} = d_{b_k} = 0.1\hat{\gamma}_k, \quad \omega_{b_k} = \omega_k, \quad \sigma_{\gamma_k} = \sigma_{b_k} = 0.2\sigma_k, \tag{27}$$

which correspond to much smaller damping and noise amplitude in the evolution of the augmented stochastic parameters compared to the previous cases of §4.

In figures 19-21 we show the DSS skill for a nearly Gaussian triad by means of the RMS error and the pattern correlation as functions of the observation time interval Δt^{obs} and the observation noise variance r^{obs}. Despite the significant RMS errors for the range of observation times and noise amplitudes examined, it should be noted that the pattern correlation for the primary mode is surprisingly good throughout even in this tough strongly correlated regime. A path-wise result of DSS for the primary mode is shown in figure 20. Another, feature of DSS is shown in figure 21 showing some skill for 'splitting' the aliased noisy signal in intervals between bursts of strong mode correlations.

Based on the results illustrated in figures 19-21 we make the following points:

- For Gaussian triads with $E_1 \gtrsim E_2 \gg E_3$ the pattern correlation for the filtered primary and secondary modes is good and robust
- Some skill for 'splitting' the aliased noisy signal in between bursts of strong correlations
- Good recovery of the time-averaged means in the secondary mode (not shown).
- The skill for filtering the composite signal is as good as when filtering with single SPEKF or GCSSF.
- Unsurprisingly, no robustness to filter parameter variations but some robustness in skill in terms of pattern correlation

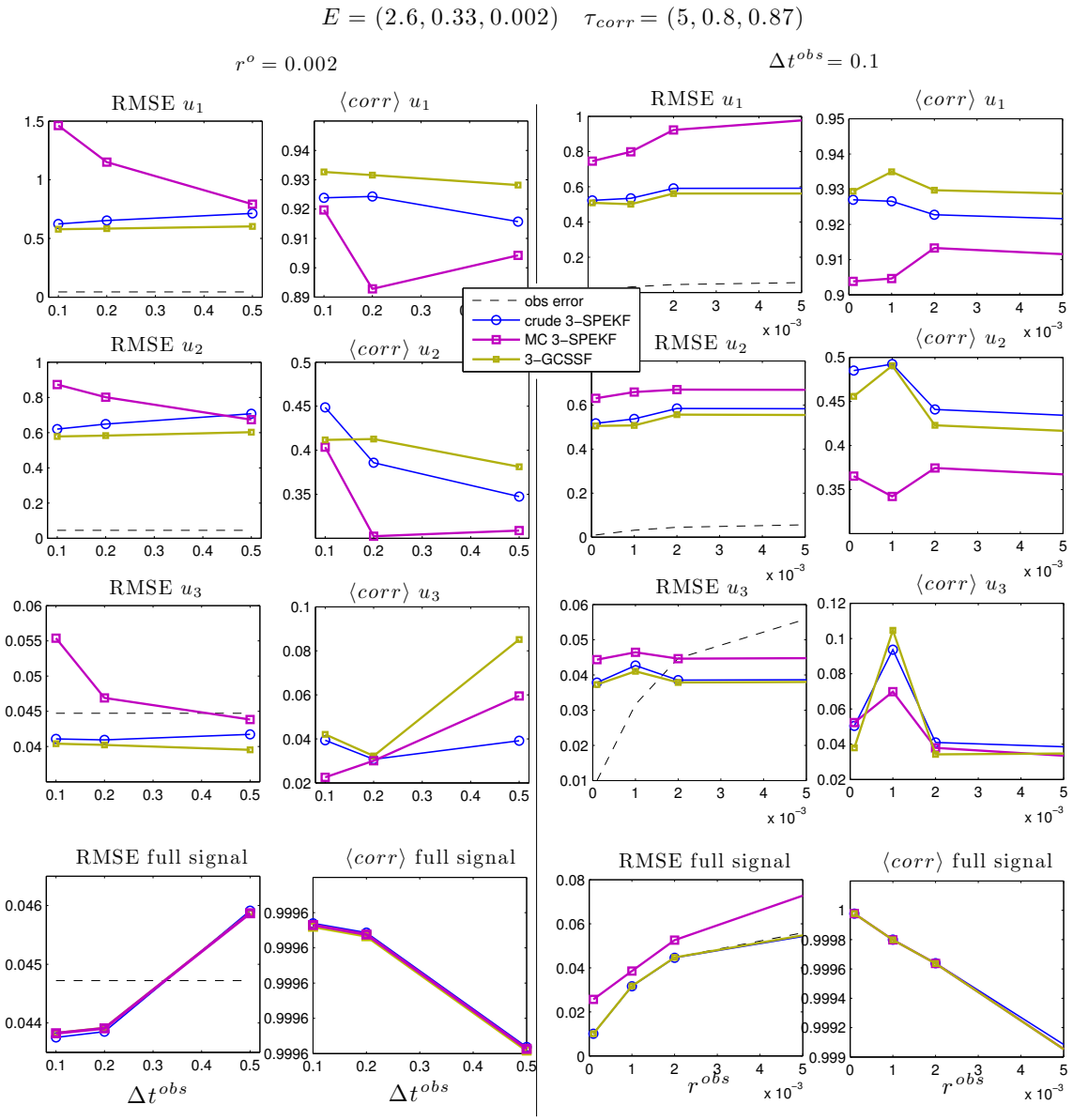

Figure 19: Performance of the DSS algorithms (§2.3) in terms of RMS error and pattern correlation for the triad system (24) with energies $E_1 = 2.6, E_2 = 0.33, E_3 = 0.002$ as a function of the observation time and observation noise variance. The dynamics of the modes in this tough test case due to strong coupling between the modes which is rather unrealistic (see §5 and Appendix A).
– Correlation times $\tau_1 = 5, \tau_2 = 0.8, \tau_3 = 0.87$.
– Dampings in $u_{1,2,3}$ estimated from the ACFs, $\sigma_{1,2,3}$ estimated from the PDFs
– Remaining dampings: $d_{\gamma_{(1,2,3)}} = 4\hat{\gamma}_{(1,2,3)}$, $d_{b_{(1,2,3)}} = 4\hat{\gamma}_{(1,2,3)}$,
– Remaining noise amps: $\sigma_{\gamma_{(1,2,3)}} = 5\sigma_{u_{(1,2,3)}}$, $\sigma_{b_{(1,2,3)}} = 5\sigma_{u_{(1,2,3)}}$,
– System parameters: $B = (0.2, -0.1, -0.1)$, $d = (0.2, 1, 2)$, $\sigma = (1, 0.8, 0.05)$, $f = (0, 0, 0)$

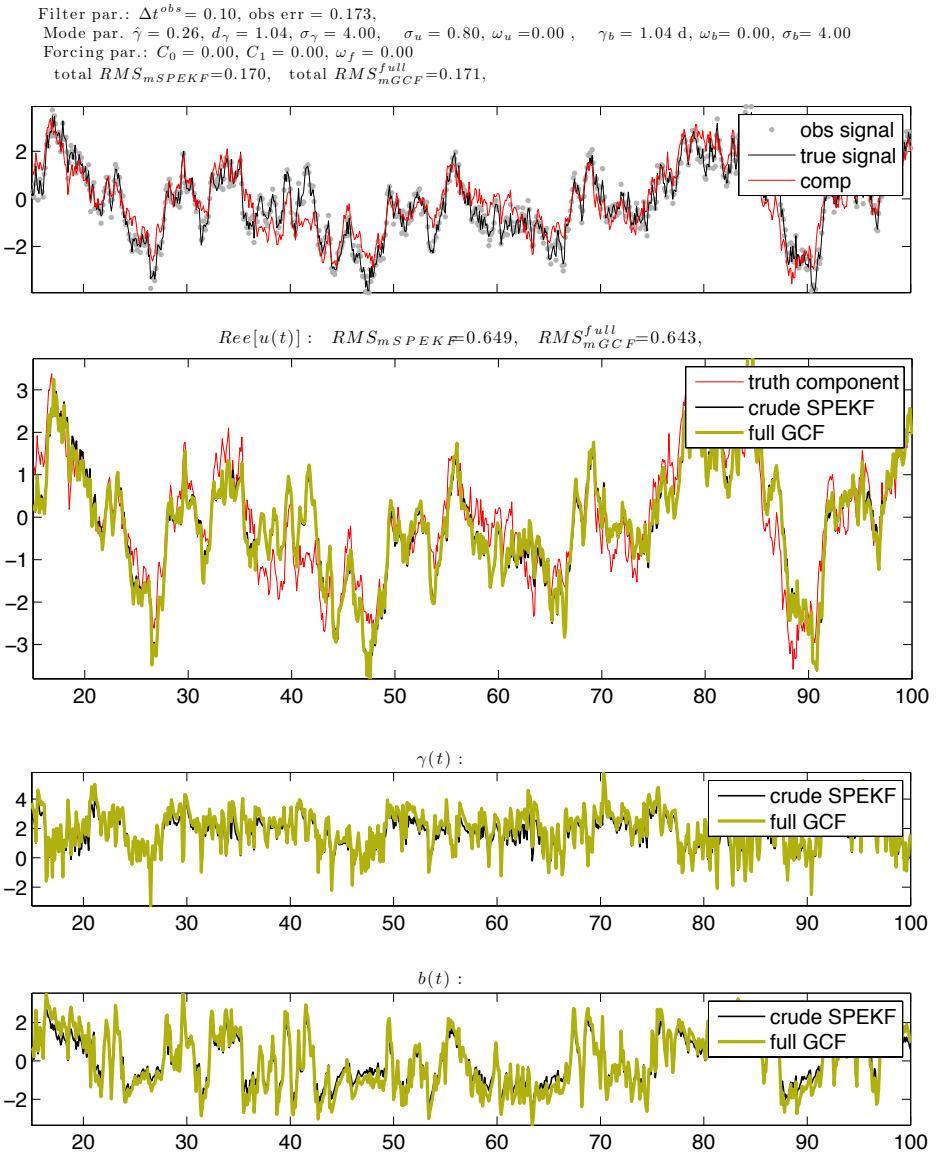

Figure 20: Path-wise example of DSS for the primary mode u_1 recovered from the aliased signal observed from the triad system (24) and involving three modes u_1, u_2, u_3 with energies $E = (2.6, 0.33, 0.002)$ using two different DSS algorithms 3-GCSSF and crude 3-SPEKF (see §2.3 for details). Good pattern correlation for of the recovered primary mode in this unrealistically tough case (see Appendix A) is shown in figure 19.

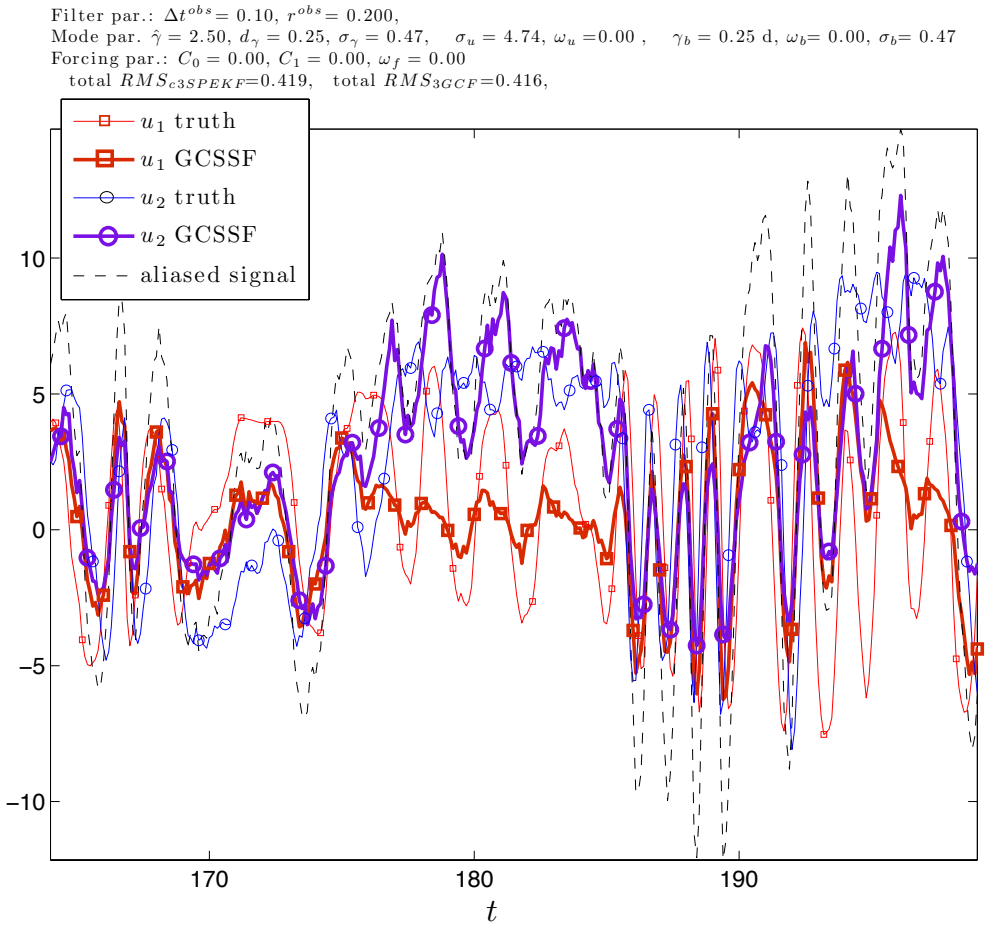

Figure 21: Some skill of DSS for disentangling the primary and secondary modes, u_1, u_2, from the aliased signal generated from the unrealistically tough case of the strongly coupled triad model (24) and involving three modes u_1, u_2, u_3. Some recovery of the aliased modes can be seen between the phases of strong correlations between u_1 and u_2.

6 Conclusions and future work

We introduced and analyzed a suite of Dynamic Stochastic Superresolution (DSS) algorithms for estimating the unresolved, subgrid processes in complex turbulent spatially extended systems from sparse regular observations. In particular, we showed that, by using cheap stochastic forecast models with judicious model error and appropriate filtering with sparse noisy observations, one can derive stochastically superresolved velocity fields and gain insight into important characteristics of the unresolved dynamics. This study was motivated by the ubiquitous problem in atmosphere-ocean and climate science which arises due to sparseness of available noisy observations of turbulent geophysical systems and the associated lack of resolution of important turbulent processes which are nontrivially coupled with the large scale dynamics through turbulent fluxes. These unconstrained turbulent fluxes which intermittently send substantial energy from small to large scales hinder efficient parameterization, filtering and real-time prediction using imperfect PDE models. Our approach to extracting the 'hidden' dynamical characteristics of the unresolved dynamics relied on the framework developed in [32, 19, 20, 33] for real-time filtering of sparsely observed turbulent signals and exploited the spatial sparseness of the observations by utilizing the fact that the coarse observation network aliases high-wavenumber information into the resolved waveband. The forecast model used in the DSS algorithms is obtained from the stochastic parameterization SPEKF forecast model [13, 12, 5, 33] in which the model parameters are learned adaptively from the observations by incorporating them into the state vector and obtaining estimates using the usual Kalman filtering methodology in Fourier space. This judicious model error avoids the curse of dimensionality while preserving the essential structure of the turbulent signal. The use of cheap stochastic models to forecast the underlying nonlinear dynamics combined with the suitable nonlinear Kalman filtering are the two main powerful features of DSS which lead to low-computational cost and robust skill.

The skill of the DSS techniques was analyzed extensively based on three classes of turbulent signals representing systems with different types of turbulent dynamics with either a spatially decaying energy spectrum or equipartition of energy between the modes. The test models included synthetic signals consisting of a superposition of SPEKF signals from its different turbulent regimes, and two nonlinear turbulent spatially extended systems with near-equipartition of energy (TBHi), and with a spatially decaying energy spectrum (MMT) but strong direct and inverse energy cascades. Three-dimensional stochastic triad models were also analyzed in order to study the performance of DSS in this extremely tough scenario with quadratic nonlinearity and strong correlations between the modes. Based on extensive tests, we showed that the DSS framework is capable of successfully capturing the extreme time-localized events in the unresolved modes, and that it provides good and robust skill for recovery the unresolved processes in terms of pattern correlation. Moreover, we provided evidence for superior skill of the DSS technique based on a Gaussian moment closure of the stochastic forecast model which is easy to implement and avoids MC estimates for the cross-correlations between the aliasing modes. The important findings discussed in this work are:

- DSS improves the skill for estimating the primary modes associated with the sparse regular observation network in the presence of aliased modes which are above noise levels (e.g., figures 12 and 18).
- DSS has good skill for recovering the aliased modes in the crude statistical sense (mainly in terms of the pattern correlation but often also in terms of RMSE); (figures 11, 12, 13, 15).
- DSS has a good skill for detecting extreme time-localized events (black swans) in the unresolved dynamics (figures 4, 5).
- Assume $\tau_{\mathcal{A}}^{\min}$ and $\tau_{\mathcal{A}}^{\max}$ to be, respectively, the smallest and the largest decorrelation times among those modes in the aliasing set $\mathcal{A}$ which are above the noise level; then
  - the skill of the DSS algorithms for $\Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\min}$ satisfies

GCSSF > crude mSPEKF >MC mSPEKF (ens. size $N \sim \mathcal{O}(100)$);

- DSS when $\tau_{\mathcal{A}}^{\min} \ll \Delta t^{obs} \lesssim \tau_{\mathcal{A}}^{\max}$: all filters are comparable.

- GCSSF has the best skill (in terms of RMS error and pattern correlation) for recovering the primary mode when filtering with $\Delta t^{obs} \lesssim \tau_{corr}^{\min}$. The pattern correlation remains good for GCSSF and crude SPEKF when superresolving with Δt^{obs} substantially exceeding $\tau_{corr}^{\min}$ (figures 11, 12, 15).
- DSS is robust and near-optimal for the MSM parameter values (e.g., figure 13).
- Larger ensemble size in MC mSPEKF than $N = 100$ used here improves its performance but results in a rapidly increasing computational cost.
- The skill of MC mSPEKF for recovering the primary mode improves with increasing observation time Δt^{obs} but the skill is poor at short observation times $\Delta t^{obs} \lesssim \tau_{min}^{corr}$ (figures 11, 15).
- The innovation error for filtering the full signal (the sum of all aliasing modes) is similar for all the DSS algorithms and it is close to the innovation error for filtering with a single SPEKF or GCF filter (figures 11, 12,13,15).

Some hurdles remain before practical filtering or superparameterization algorithms incorporating these techniques can be developed for use with real observation data. In particular, the assumption of regular simultaneous observations on a Cartesian grid in this study meant that the aliasing condition assumed a particularly simple form; operationally real observational data are often interpolated on such grids. Nevertheless, the implicit interpolation of the observation data from irregular networks onto a regular grid, which is often done operationally, carries its own risks and limitations as discussed recently in [16]. It would be interesting, however, to investigate whether a direct interpolation of the sparse observations on the forecast model grid and using them as the priors in a DSS algorithm leads to improvements in the estimates for the unresolved processes. Despite these caveats, this definitive study combined with the earlier results of [22] provides a strong evidence that computationally cheap DSS techniques offer a promising and efficient approach to constraining unresolved fluxes through stochastic superparameterization [24, 42, 26, 15, 14] and a subsequent improvement in coarse-grained filtering and prediction of the next generation AOS models. Such methods should also be useful for real time filtering of anisotropic turbulent signals in other applications in engineering and environmental sciences.

Acknowledgements

M.B is supported as a postdoctoral fellow on the ONR DRI grant of A.J.M N0014-10-1-0554. The research of A.J.M. is partially supported by National Science Foundation grant DMS-0456713, NSF CMG grant DMS-1025468 and the office of Naval Research grants ONR DRI N0014-10-1-0554 and N00014-11-1-0306. The authors also acknowledge Ian Grooms for his help with the initial version of the code for solving the MMT dynamics.

A The nature of cross-correlations between the Fourier modes of high-dimensional turbulent systems from nature

Here, we demonstrate the feasibility of the general strategy for DSS introduced in §2 for sparsely observed turbulent geophysical systems on the idealized climate model equations for a spectrally truncated barotropic flow on a beta plane. Crucially, we show that the Fourier modes in each aliasing set of the primary wavenumbers associated with the observation network are coupled only indirectly through nonlinear interactions with modes outside the considered aliasing set. This property is a direct consequence of the fact that the nonlinear terms in the dynamics arise from the quadratic advection term.

The damped forced barotropic quasi geostrophic equations in a beta plane are given by (e.g. [31])

$$\frac{\partial \omega}{\partial t} + \nabla^{\perp}\psi \cdot \nabla \omega + \beta \frac{\partial \psi}{\partial x} = \mathcal{D}(\Delta)\psi + \mathcal{F}(t, \boldsymbol{x}), \tag{28}$$

where $\boldsymbol{x} \equiv x\boldsymbol{e}_x + y\boldsymbol{e}_y \in I\!R^2$, $\boldsymbol{\omega} = \boldsymbol{e}_z\omega$ is the vorticity and ψ is the streamfunction. The finite-dimensional truncation of the barotropic equations in (28) is obtained by making a Galerkin approximation where the equations are projected into a finite-dimensional subspace. Consistent with the two-dimensional periodic boundary conditions used in the numerical simulations, the truncation is readily accomplished with the standard Fourier basis. More precisely, we introduce the Fourier series expansion of the truncated small-scale stream function ψ, and the truncated vorticity, ω, in the truncated basis $B_N = \{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \quad \boldsymbol{k} \in \sigma_\Lambda\}$ where $\sigma_{\tilde{N}} = \{\boldsymbol{k} : \ 1 \leqslant |k_x| \leqslant \tilde{N}/2, 1 \leqslant |k_y| \leqslant \tilde{N}/2\}$; these truncated expansion are given by

$$\omega(t, \boldsymbol{x}) = \sum_{\boldsymbol{k}\in\sigma_{\tilde{N}}} \hat{\omega}_{\boldsymbol{k}}(t) e^{i\boldsymbol{x}\cdot\boldsymbol{k}}, \qquad \psi(t, \boldsymbol{x}) = -\sum_{\boldsymbol{k}\in\sigma_{\tilde{N}}} \frac{\hat{\omega}_{\boldsymbol{k}}(t)}{|\boldsymbol{k}|^2} e^{i\boldsymbol{x}\cdot\boldsymbol{k}}. \tag{29}$$

Substitution of (29) into (28) leads to the following Fourier series representation of truncated barotropic quasi-geostrophic equations for each mode $\boldsymbol{k} \in \sigma_{\tilde{N}}$ (see [31])

$$\frac{\mathrm{d}\hat{\omega}_{\boldsymbol{k}}(t)}{\mathrm{d}t} - \frac{i\beta k_x}{|\boldsymbol{k}|^2}\hat{\omega}_{\boldsymbol{k}}(t) + \sum_{\substack{\boldsymbol{m}\in\sigma_{\tilde{N}} \\ |\boldsymbol{m}|\neq|\boldsymbol{k}-\boldsymbol{m}|}} \frac{\boldsymbol{m}^{\perp}\cdot\boldsymbol{k}}{|\boldsymbol{m}|^2}\hat{\omega}_{\boldsymbol{m}}(t)\hat{\omega}_{\boldsymbol{k}-\boldsymbol{m}}(t) = -\frac{\mathcal{D}(-|\boldsymbol{k}|^2)}{|\boldsymbol{k}|^2}\hat{\omega}_{\boldsymbol{k}}(t) + \hat{\mathcal{F}}_{\boldsymbol{k}}(t). \tag{30}$$

We are interested in determining the nature of direct interactions between the modes in the aliasing set. Consider the $\tilde{M} \times \tilde{M}$ observation mesh as in §2.1, which is a sparse regular subset of the model mesh with $\tilde{N} \times \tilde{N}$ points such that $P = \tilde{N}/\tilde{M}$ is an integer.

Consider the aliasing set of the observed wavenumber $\boldsymbol{l}$, $|l_x| \leqslant \tilde{M}/2$, $|l_y| \leqslant \tilde{M}/2$, given by

$$\mathcal{A}(\boldsymbol{l}) = \{\boldsymbol{k} : |k_x| \leqslant \tilde{N}/2, |k_y| \leqslant \tilde{N}/2, \ k_x = l_x + \tilde{M}q_x, \ k_y = l_y + \tilde{M}q_y, \ \ q_x, q_y \in \mathbb{Z}\}. \tag{31}$$

It is easy to show that the modes of (30) contained in the aliasing set of $\boldsymbol{l}$ are not directly coupled. In order to see this, consider any three modes in the aliasing set, $\boldsymbol{k}^{ij}, \boldsymbol{k}^{mn}, \boldsymbol{k}^{ab} \in \mathcal{A}(\boldsymbol{l})$, which are not necessarily distinct, and such that

$$\boldsymbol{k}^{ij} \equiv \big(\boldsymbol{l}_x + \tilde{M}q_x^i\big)\boldsymbol{e}_x + \big(\boldsymbol{l}_y + \tilde{M}q_y^j\big)\boldsymbol{e}_y; \tag{32}$$

analogous definitions hold for $\boldsymbol{k}^{mn}, \boldsymbol{k}^{ab}$. Direct couplings between the aliasing modes imply that

$$\frac{\mathrm{d}\hat{\omega}_{\boldsymbol{k}_{ij}}}{\mathrm{d}t} \propto \hat{\omega}_{\boldsymbol{k}_{mn}}\hat{\omega}_{\boldsymbol{k}_{ab}}, \tag{33}$$

where

$$\boldsymbol{k}_{ab} = \boldsymbol{k}_{ij} - \boldsymbol{k}_{mn}, \tag{34}$$

based on (30). It is straightforward to show that the condition (34) cannot be satisfied for the aliasing modes since a direct substitution of (32) into (34) leads to

$$\tilde{M}\big(q_x^i - q_x^m - q_x^a\big) = l_x, \quad \tilde{M}\big(q_y^j - q_y^n - q_y^b\big) = l_y, \tag{35}$$

which cannot be satisfied for any integers q_x, q_y for $|l_x| + |l_y| > 0$ since $|l_x| \leqslant \tilde{M}/2$ and $|l_y| \leqslant \tilde{M}/2$. Consequently, the modes in each aliasing set $\mathcal{A}(\boldsymbol{l})$ are only indirectly coupled through the nonlinear interactions with modes from other aliasing sets. Similar calculations remain valid for multi-layer quasi-geostrophic models.

B Derivation of approximate cross-correlation terms between different aliasing modes for the multi-SPEKF forecast model using Gaussian closure

Here, the goal is to derive the closed system for the approximate second-order statistics of the forecast model (9). This is achieved via the Gaussian moment closure by neglecting odd moments in the equations for the evolution of the second-order statistics associated with the full system (9) with quadratic nonlinearities; see [5] for such a procedure applied to a single SPEKF. We first decompose all variables into the mean and fluctuations around the mean

$$u_k = \bar{u}_k + u'_k, \quad b_k = \bar{b}_k + b'_k, \quad \gamma_k = \bar{\gamma}_k + \gamma'_k, \quad \omega_k = \bar{\omega}_k + \omega'_k, \tag{36}$$

which is the analogue of the Reynolds average decomposition. The resulting equations for the mean and fluctuations are

$$\mathrm{d}\bar{u}_k = \big[(-\bar{\gamma}_k + i\bar{\omega}_k)\bar{u}_k - \overline{u'_k\gamma'_k} + i\overline{u'_k\omega'_k} + \bar{b}_k + f\,\big]\mathrm{d}t, \tag{37}$$

$$\mathrm{d}u'_k = \big[(-\bar{\gamma}_k + i\bar{\omega}_k)u'_k + b'_k - \bar{u}_k\gamma'_k + i\bar{u}_k\omega'_k - u'_k\gamma'_k + iu'_k\omega'_k + \overline{u'_k\gamma'_k} - i\overline{u'_k\omega'_k}\,\big]\mathrm{d}t + \sigma_{u_k}\mathrm{d}W_{u_k}, \tag{38}$$

$$\mathrm{d}\bar{b}_k = (-d_b + i\omega_b)\left(\bar{b}_k - \hat{b}_k\right)\mathrm{d}t, \tag{39}$$

$$\mathrm{d}b'_k = (-d_b + i\omega_b)\,b'_k\mathrm{d}t + \sigma_{b_k}\mathrm{d}W_{b_k}, \tag{40}$$

$$\mathrm{d}\bar{\gamma}_k = -d_{\gamma_k}\left(\bar{\gamma}_k - \hat{\gamma}_k\right)\mathrm{d}t, \tag{41}$$

$$\mathrm{d}\gamma'_k = -d_{\gamma_k}\gamma'_k\mathrm{d}t + \sigma_{\gamma_k}\mathrm{d}W_{\gamma_k}. \tag{42}$$

$$\mathrm{d}\bar{\omega}_k = -d_{\omega_k}\left(\bar{\omega}_k - \hat{\omega}_k\right)\mathrm{d}t, \tag{43}$$

$$\mathrm{d}\omega'_k = -d_{\omega_k}\omega'_k\mathrm{d}t + \sigma_{\omega_k}\mathrm{d}W_{\omega_k}. \tag{44}$$

Based on the above equations and the Ito formula the evolution of the second-order statistics for the system is not closed in the variables

$$\begin{aligned}
y_{m,1} &= \bar{u}_m & y_{mn,10} &= \overline{b'_m b'_n} \\
y_{m,2} &= \bar{b}_m & y_{mn,11} &= \overline{u'_m b'_n} \\
y_{m,3} &= \bar{\gamma}_m & y_{mn,12} &= \overline{u'_m b'^*_n} \\
y_{m,4} &= \bar{\omega}_m & y_{mn,13} &= \overline{u'_m \gamma'_n} \\
y_{mn,5} &= \overline{u'_m u'^*_n} & y_{mn,14} &= \overline{u'_m \omega'_n} \qquad m,n = 1,...,P. \\
y_{mn,6} &= \overline{b'_m b'^*_n} & y_{mn,15} &= \overline{b'_m \gamma'_n} \\
y_{mn,7} &= \overline{\gamma'_m \gamma'_n} & y_{mn,16} &= \overline{b'_m \omega'_n} \\
y_{mn,8} &= \overline{\omega'_m \omega'_n} & y_{mn,17} &= \overline{\gamma'_m \omega'_n} \\
y_{mn,9} &= \overline{u'_m u'_n}
\end{aligned} \tag{45}$$

unless some type of moment closure is used; this can be easily seen, for example, for

$$\mathrm{d}\overline{u'_m u'_n} = \Big[(-(\bar{\gamma}_m + \bar{\gamma}_n) + i(\bar{\omega}_m + \bar{\omega}_n))\overline{u'_m u'_n} + \overline{u'_m b'_n} + \overline{u'_n b'_m} \\ - \bar{u}_n\overline{u'_m\gamma'_n} - \bar{u}_m\overline{u'_n\gamma'_m} + i\bar{u}_n\overline{u'_m\omega'_n} + i\bar{u}_m\overline{u'_n\omega'_m} - \overline{\gamma'_m u'_m u'_n} + i\overline{\omega'_m u'_m u'_n}\Big]\mathrm{d}t, \quad (46)$$

which contains third order moments. Neglecting these odd moments leads to

$$\mathrm{d}\overline{u'_m u'_n} \approx \Big[(-(\bar{\gamma}_m + \bar{\gamma}_n) + i(\bar{\omega}_m + \bar{\omega}_n))\overline{u'_m u'_n} + \overline{u'_m b'_n} + \overline{u'_n b'_m} \\ - \bar{u}_n\overline{u'_m\gamma'_n} - \bar{u}_m\overline{u'_n\gamma'_m} + i\bar{u}_n\overline{u'_m\omega'_n} + i\bar{u}_m\overline{u'_n\omega'_m}\Big]\mathrm{d}t. \quad (47)$$

Application of this simple Gaussian moment closure to the evolution of all the variables y_{mn} in (45) leads to the following dynamical system

$$\dot{y}_{m,1} = (-y_{m,3} + iy_{m,4})y_{m,1} - y_{mm,13} + iy_{mm,14} + y_{m,2} + f_m \quad (48)$$

$$\dot{y}_{m,2} = (-d_{b_m} + i\omega_{b_m})(y_{m,2} - \hat{b}_m) \quad (49)$$

$$\dot{y}_{m,3} = -d_{\gamma_m}(y_{m,3} - \hat{\gamma}_m) \quad (50)$$

$$\dot{y}_{m,4} = -d_{\omega_m}(y_{m,4} - \hat{\omega}_m) \quad (51)$$

$$\dot{y}_{mn,5} = [-y_{m,3} - y_{n,3} + i(y_{m,4} - y_{n,4})]y_{mn,5} + y_{mn,12} + y^*_{nm,12} \\ - y^*_{n,1}y_{mn,13} - y_{m,1}y^*_{nm,13} - iy^*_{n,1}y_{mn,14} + iy_{m,1}y^*_{nm,14} + \delta_{mn}\sigma^2_{u_m} \quad (52)$$

$$\dot{y}_{mn,6} = [-d_{b_m} - d_{b_n} + i(\omega_{b_m} - \omega_{b_n})]y_{mn,6} + \delta_{mn}\sigma^2_{b_m} \quad (53)$$

$$\dot{y}_{mn,7} = -(d_{\gamma_m} + d_{\gamma_n})y_{mn,7} + \delta_{mn}\sigma^2_{\gamma_m} \quad (54)$$

$$\dot{y}_{mn,8} = -(d_{\omega_m} + d_{\omega_n})y_{mn,8} + \delta_{mn}\sigma^2_{\omega_m} \quad (55)$$

$$\dot{y}_{mn,9} = [-y_{m,3} - y_{n,3} + i(y_{m,4} + y_{n,4})]y_{mn,9} + y_{mn,11} + y_{nm,11} \\ - y_{n,1}y_{mn,13} - y_{m,1}y_{nm,13} + iy_{n,1}y_{mn,14} + iy_{m,1}y_{nm,14} \quad (56)$$

$$\dot{y}_{mn,10} = [-d_{b_m} - d_{b_n} + i(\omega_{b_m} + \omega_{b_n})]y_{mn,10} \quad (57)$$

$$\dot{y}_{mn,11} = [-y_{m,3} - d_{b_n} + i(y_{m,4} + \omega_{b_n})]y_{mn,11} + y_{mn,10} - y_{m,1}y_{nm,15} + iy_{m,1}y_{nm,16} \quad (58)$$

$$\dot{y}_{mn,12} = [-y_{m,3} - d_{b_n} + i(y_{m,4} - \omega_{b_n})]y_{mn,12} + y_{mn,6} - y_{m,1}y^*_{nm,15} + iy_{m,1}y^*_{nm,16} \quad (59)$$

$$\dot{y}_{mn,13} = [-y_{m,3} - d_{\gamma_n} + iy_{m,4}]y_{mn,13} + y_{mn,15} - y_{m,1}y_{mn,7} + iy_{m,1}y_{nm,17} \quad (60)$$

$$\dot{y}_{mn,14} = [-y_{m,3} - d_{\omega_n} + iy_{m,4}]y_{mn,14} + y_{mn,16} - y_{m,1}y_{mn,17} + iy_{m,1}y_{mn,8} \quad (61)$$

$$\dot{y}_{mn,15} = (-d_{b_m} - d_{\gamma_n} + i\omega_{b_m})y_{mn,15} \quad (62)$$

$$\dot{y}_{mn,16} = (-d_{b_m} - d_{\omega_n} + i\omega_{b_m})y_{mn,16} \quad (63)$$

$$\dot{y}_{mn,17} = (-d_{\gamma_m} - d_{\omega_n})y_{mn,17} \quad (64)$$

The dynamical system (48) has nine parameters $\hat{b}_m, \hat{\gamma}_m, \omega_{u_m}, d_{b_m}, d_{\gamma_m}, \omega_{b_m}, \sigma_{u_m}, \sigma_{b_m}, \sigma_{\gamma_m}$; f_m is the forcing. We can now use the the nonlinear system in the Kalman Filter to update the prior mean and covariance. The covariance matrix $R_{m+1|m}$ in the Kalman filter of the associated DSS algorithm step is

given by

$$R_{m+1|m} = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1P} \\ R_{21} & R_{22} & \dots & R_{2P} \\ \vdots & \vdots & \vdots & \vdots \\ R_{P1} & R_{P2} & \dots & R_{PP} \end{pmatrix}, \tag{65}$$

where

$$R_{mm} = \begin{bmatrix} \overline{u'^2_{m,\Re}} & \overline{u'_{m,\Re}u'_{m,\Im}} & \overline{u'_{m,\Re}b'_{m,\Re}} & \overline{u'_{m,\Re}b'_{m,\Im}} & \overline{u'_{m,\Re}\gamma'_m} & \overline{u'_{m,\Re}\omega'_m} \\ \overline{u'_{m,\Im}u'_{m,\Re}} & \overline{u'^2_{m,\Im}} & \overline{u'_{m,\Im}b'_{m,\Re}} & \overline{u'_{m,\Im}b'_{m,\Im}} & \overline{u'_{m,\Im}\gamma'_m} & \overline{u'_{m,\Im}\omega'_m} \\ \overline{b'_{m,\Re}u'_{m,\Re}} & \overline{b'_{m,\Re}u'_{m,\Im}} & \overline{b'^2_{m,\Re}} & \overline{b'_{m,\Re}b'_{m,\Im}} & \overline{b'_{m,\Re}\gamma'_m} & \overline{b'_{m,\Re}\omega'_m} \\ \overline{b'_{m,\Im}u'_{m,\Re}} & \overline{b'_{m,\Im}u'_{m,\Im}} & \overline{b'_{m,\Im}b'_{m,\Re}} & \overline{b'^2_{m,\Im}} & \overline{b'_{m,\Im}\gamma'_m} & \overline{b'_{m,\Im}\omega'_m} \\ \overline{\gamma'_m u'_{m,\Re}} & \overline{\gamma'_m u'_{m,\Im}} & \overline{\gamma'_m b'_{m,\Re}} & \overline{\gamma'_m b'_{m,\Im}} & \overline{\gamma'^2_m} & \overline{\gamma'_m\omega'_m} \\ \overline{\omega'_m u'_{m,\Re}} & \overline{\omega'_m u'_{m,\Im}} & \overline{\omega'_m b'_{m,\Re}} & \overline{\omega'_m b'_{m,\Im}} & \overline{\omega'_m\gamma'_m} & \overline{\omega'^2_m} \end{bmatrix} \tag{66}$$

$$R_{mn} = \begin{bmatrix} \overline{u'_{m,\Re}u'_{n,\Re}} & \overline{u'_{m,\Re}u'_{n,\Im}} & \overline{u'_{m,\Re}b'_{n,\Re}} & \overline{u'_{m,\Re}b'_{n,\Im}} & \overline{u'_{m,\Re}\gamma'_n} & \overline{u'_{m,\Re}\omega'_n} \\ \overline{u'_{m,\Im}u'_{n,\Re}} & \overline{u'_{m,\Im}u'_{n,\Im}} & \overline{u'_{m,\Im}b'_{n,\Re}} & \overline{u'_{m,\Im}b'_{n,\Im}} & \overline{u'_{m,\Im}\gamma'_n} & \overline{u'_{m,\Im}\omega'_n} \\ \overline{b'_{m,\Re}u'_{n,\Re}} & \overline{b'_{m,\Re}u'_{n,\Im}} & \overline{b'_{m,\Re}b'_{n,\Re}} & \overline{b'_{m,\Re}b'_{n,\Im}} & \overline{b'_{m,\Re}\gamma'_n} & \overline{b'_{m,\Re}\omega'_n} \\ \overline{b'_{m,\Im}u'_{n,\Re}} & \overline{b'_{m,\Im}u'_{n,\Im}} & \overline{b'_{m,\Im}b'_{n,\Re}} & \overline{b'_{m,\Im}b'_{n,\Im}} & \overline{b'_{m,\Im}\gamma'_n} & \overline{b'_{m,\Im}\omega'_n} \\ \overline{\gamma'_m u'_{n,\Re}} & \overline{\gamma'_m u'_{n,\Im}} & \overline{\gamma'_m b'_{n,\Re}} & \overline{\gamma'_m b'_{n,\Im}} & \overline{\gamma'_m\gamma'_n} & \overline{\gamma'_m\omega'_n} \\ \overline{\omega'_m u'_{n,\Re}} & \overline{\omega'_m u'_{n,\Im}} & \overline{\omega'_m b'_{n,\Re}} & \overline{\omega'_m b'_{n,\Im}} & \overline{\omega'_m\gamma'_n} & \overline{\omega'_m\omega'_n} \end{bmatrix} \tag{67}$$

where $u'_{m,\Re} \equiv \Re e[u'_m]$, $u'_{m,\Im} \equiv \Im m[u'_m]$, etc.

References

[1] J.L. Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129:2884–2903, 2001.

[2] J.L. Anderson. A local least squares framework for ensemble filtering. *Monthly Weather Review*, 131(4):634–642, 2003.

[3] P. Bartello. Using low-resolution winds to deduce fine structure in tracers. *Atmos. Ocean*, 38(2):303–320, 2000.

[4] J. Biello and J. K. Hunter. Nonlinear Hamiltonian Waves with Constant Frequency and Surface Waves on Vorticity Discontinuities. *Comm. Pure Appl. Math*, 63:0303–0336, 2010.

[5] M. Branicki, B. Gershgorin, and A.J. Majda. Filtering skill for turbulent signals for a suite of nonlinear and linear Kalman filters. *J. Comp. Phys*, 231:1462–1498, 2012.

[6] M. Branicki and A. J. Majda. Quantifying Filter Performance for Turbulent Dynamical Systems through Information Theory. 2012. submitted 2012.

[7] D. Cai, A. J. Majda, D. W. McLaughlin, and E. G. Tabak. Dispersive wave turbulence in one dimension. *Physica D*, 152–153:551–572, 2001.

[8] D. Cai, A.J. Majda, D.W. McLaughlin, and E.G. Tabak. Spectral bifurcations in dispersive wave turbulence. *Proc. Nat. Acad. Sci.*, 96:14216–14221, 1999.

[9] T. DelSole. Stochastic Models of Quasigeostrophic Turbulence. *Surveys in Geophys.*, 25(2):107–149, 2004.

[10] F. Gabriel. Statistical mechanics for spectrally truncated equations of burgers-hopf type. Master's thesis, Courant Institute of Mathematical Sciences, New York University, 2009.

[11] A. Ganachaud and C. Wunsch. Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*, 408:453–457, 2000.

[12] B. Gershgorin, J. Harlim, and A.J. Majda. Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation. *J Comp Phys*, 229:32–57, 2010.

[13] B. Gershgorin, J. Harlim, and A.J. Majda. Test models for improving filtering with model errors through stochastic parameter estimation. *J Comp Phys*, 229:1–31, 2010.

[14] W. W. Graboski. Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.*, 58:978–997, 2001.

[15] W.W. Grabowski and P. K. Smolarkiewicz. Crcp: A cloud resolvingconvection parameterization for modeling the tropical convecting atmosphere. *Physica D*, 133:171–178, 1999.

[16] J. Harlim. Interpolating irregularly spaced observations for filtering turbulent complex systems. *SIAM J. Sci. Comput.*, 33(5):2620–2640, 2011.

[17] J. Harlim and A.J.Majda. Test Models for Filtering and Prediction of Moisture-Coupled Tropical Waves. *QJRMS*, 2012. accepted and in press.

[18] J. Harlim and A. Majda. Filtering nonlinear dynamical systems with linear stochastic models. *Nonlinearity*, 21:1281–1306, 2008.

[19] J. Harlim and A.J. Majda. Mathematical strategies for filtering complex systems: Mathematical strategies for filtering complex systems: Regularly spaced sparse observations. *J. Comp. Phys.*, 227:5304–5341, 2008.

[20] J. Harlim and A.J. Majda. Filtering turbulent sparsely observed geophysical flows. *Mon. Wea. Rev.*, 138(4):1050–1083, 2010.

[21] S. R. Jayne and J. Marotzke. The oceanic eddy heat transport. *J. Phys. Oceanogr.*, 32:3328–3345, 2002.

[22] S.R. Keating, A.J. Majda, and K. S. Smith. New methods for estimating poleward eddy heat transport using satellite altimetry. *Mon. Wea. Rev.*, 2011. accepted and in press.

[23] A. Majda and I. Timofeyev. Statistical mechanics for truncations of the Burgers–Hopf equation: a model for intrinsic stochastic behavior with scaling. *Milan J. Math.*, 79(1):39–96, 2002.

[24] A. J. Majda. Multiscale Models with Moisture and Systematic Strategies for Superparameterization. *J. Atmos Sci.*, 64:2726–2734, 2007.

[25] A. J. Majda, R. V. Abramov, and M. J. Grote. *Information Theory and Stochastics for Multiscale Nonlinear Systems*, volume 25 of *CRM Monograph Series*. Americal Mathematical Society, Providence, 2005.

[26] A. J. Majda and M. Grote. Mathematical Test Models for Superparameterization in Anisotropic Turbulence. *PNAS*, 106(14):5470–5476, 2009.

[27] A. J. Majda, D. McLaughlin, and E. Tabak. A one-dimensional model for dispersive wave turbulence. *J. Nonlinear Sci.*, 6:9–44, 1997.

[28] A. J. Majda, I. I. Timofeyev, and E. Vanden Eijnden. Models for stochastic climate predicition. *Proc. Natl. Acad. Sci.*, 96:14687, 1999.

[29] A. J. Majda, I. I. Timofeyev, and E. Vanden Eijnden. A mathematical framework for stochastic climate models. *Comm. Pure Appl. Math*, 54:891–974, 2001.

[30] A. J. Majda, I. I. Timofeyev, and E. Vanden Eijnden. A priori tests of a stochatic mode reduction strategy. *Phys. D*, 170:206–252, 2002.

[31] A. J. Majda and X. Wang. *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*. Cambridge University Press, Cambridge, 2006.

[32] A.J. Majda and M.J. Grote. Explicit off-line criteria for stable accurate time filtering of strongly unstable spatially extended systems. *PNAS*, 104:1124–1129, 2007.

[33] A.J. Majda and J. Harlim. *Filtering Complex Turbulent Systems*. ISBN-13:9781107016668. Cambridge University Press, 2012.

[34] A.J. Majda, J. Harlim, and B. Gershgorin. Mathematical strategies for filtering turbulent dynamical systems. *DCDS*, 27:441–486, 2010.

[35] A.J. Majda and I. Timofeyev. Remarkable statistical behavior for truncated burgers–hopf dynamics. *PNAS*, 97(23):12413–12417, 2000.

[36] J. Neelin, B. Lintner, B. Tian, Q. Li, L. Zhang, P. Patra, M. Chahine, and S. Stechmann. Long tails in deep columns of natural and anthropogenic tropospheric tracers. *Geophysical Res. Lett.*, 37:L05804, 2010.

[37] T. J. O'Kane and J. S. Frederiksen. Comparison of Statistical Dynamical, Square Root and Ensemble Kalman Filters. *Entropy*, 10:684–721, 2008.

[38] R. Salmon. *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 1998.

[39] D. Sornette. Dragon-Kings, Black Swans and the Prediction of Crises. *Int. J. Terraspace Science and Engineering*, 2(1):1–18, 2009.

[40] N.N. Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

[41] C. Wunsch. Where do ocean eddy heat fluxes matter? *J. Geophys. Res.*, 104:13 235 –13 249, 1999.

[42] Y. Xing, A. J. Majda, and W. W. Grabowski. New Efficient Sparse Space-Time Algorithms for Superparamterization on Mesoscales. *Mon. Wea. Rev.*, 137(12):4307–4323, 2009.

[43] V. Zakharov, F. Dias, and A. Pushkarev. One-dimensional wave turbulence. *Physics Reports*, 398(1):1–65, 2004.