Filtering skill for turbulent signals for a suite of nonlinear and linear extended Kalman filters

Citation for published version:
Branicki, M, Gershgorin, B & Majda, AJ 2012, 'Filtering skill for turbulent signals for a suite of nonlinear and linear extended Kalman filters', *Journal of Computational Physics*, vol. 231, no. 4, pp. 1462-1498.
https://doi.org/10.1016/j.jcp.2011.10.029

Digital Object Identifier (DOI):
10.1016/j.jcp.2011.10.029

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Early version, also known as pre-print

Published In:
Journal of Computational Physics

Filtering skill for turbulent signals for a suite of Nonlinear and Linear Extended Kalman Filters

M. Branicki, B. Gershgorin, A. J. Majda

Department of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, NY 10012, United States

Abstract

The filtering skill for turbulent signals from nature is often limited by model errors created by utilizing an imperfect model for filtering. Updating the parameters associated with unresolved or unknown processes in the imperfect model "on the fly" through stochastic parameter estimation is an efficient way to increase filtering skill and model performance. Here, a suite of filters implementing stochastic parameter estimation is examined on a nonlinear, exactly solvable, stochastic test model mimicking turbulent signals in regimes ranging from configurations with strongly intermittent, transient instabilities to laminar behavior. Stochastic Parameterization Extended Kalman Filter (SPEKF) systematically corrects both multiplicative and additive biases in the observed dynamics and it involves exact formulas for propagating the mean and covariance including the unresolved parameters in the test model. The remaining filters use the same nonlinear test model but they introduce additional model error through different moment closure approximations and/or linear tangent approximation used for computing the second-order statistics in the stochastic forecast model. A comprehensive study of filter performance is carried out in the presence of various sources of model error as the observation time and observation noise levels are varied. In particular, regimes of filter divergence for the linear tangent filter are identified. The estimation skill of the unresolved stochastic parameters by various filters is also discussed and it is shown that the linear tangent filter, despite its popularity, is completely unreliable in many dynamical regimes. The results presented here provide useful guidelines for filtering turbulent, high-dimensional, spatially extended systems with significant model errors. They also provide unambiguous benchmarks for the capabilities of linear and nonlinear extended Kalman filters on a stringent, exactly solvable test bed.

Keywords:
Stochastic parameter estimation, filtering turbulence, model error, Gaussian closure filter, Extended Kalman Filter, data assimilation.

1. Introduction

A central problem in many contemporary applications in science and engineering lies in developing techniques for real-time statistical estimation of a state of a natural system based on partial observations and an imperfect model. This process is referred to as *filtering* and is crucial for making accurate predictions of the future state of the system. Development of efficient, accurate and robust filtering algorithms obviously poses a problem with significant practical impact in a number of strategic areas. Important contemporary examples include tracking and positioning systems, real-time filtering of multiscale turbulent signals for prediction of weather and climate, as well as the spread of hazardous plumes or pollutants.

A major difficulty in accurate filtering of noisy turbulent signals with many degrees of freedom is model error [46, 31]; the fact that the signal from nature is processed through an imperfect model where important physical processes are parameterized due to inadequate numerical resolution or incomplete physical understanding. Virtually all atmosphere, ocean, and climate models with sufficiently high resolution are turbulent dynamical systems with multiple spatio-temporal scales. Similarly, many contemporary engineering applications such as satellite orbit control [28, 7], distributed power generation systems [48, 49] or modern optical devices [42, 39] require real-time filtering with nonlinear models.

Dealing with model error is often difficult since its exact properties are, by its very nature, unknown in any realistic scenario. Various strategies for mitigating model error in nonlinear filtering have been developed and they roughly fall into techniques based on deterministic models, e.g., shadowing filters [24], and techniques combining stochastic models and sequential Bayesian filters. Miscellaneous approaches have been developed in the latter context; examples include particle filters [6, 37, 36, 4, 43], reduced order filters [18, 45, 11] and a myriad of ensemble Kalman filters [10, 5, 1, 2, 44, 22] which draw from the classical Kalman filter framework [25, 26]. Most of these techniques have their own niches of applicability but they generally suffer from the so-called "curse of dimension" (e.g., [4, 43]) and/or the "curse of ensemble size" (e.g., [19, 31]) when applied to realistic systems; consequently, the unifying feature of these methods in the context of filtering high-dimensional turbulent systems is often an unrealistically large computational overhead necessary for a reliable operation.

Filters utilizing stochastic forecast models with 'on the fly' parameter estimation for reducing model error offer a cheap and skillful alternative for filtering turbulent systems with many spatio-temporal scales [13, 14, 21, 31, 17]. The real-time, adaptive parameterization of the unresolved scales used in such filters requires augmentation of the forecast model with the dynamics of hidden stochastic parameters and leads to a nonlinear system for each mode of the signal in the turbulent spectrum. The Stochastic Parameter Estimation Filter (SPEKF, [13]) uses exact nonlinear statistics and it has been successfully tested in a hierarchy of complex models [13, 14, 21, 27, 17]. However, other filter implementations based on the Extended Kalman Filter (EKF, [26]) and various moment closure approximations are possible within the same framework. These filters introduce additional model error through the use of incorrect statistics but are often easier to derive. It is therefore imperative to compare the skill of these filters with SPEKF using an unambiguous test model which allows for disentangling the following two sources of model error:

(i) *Imperfect forecast model used for filtering.* These errors arise when the signal from nature is processed through a forecast model which parameterizes important physical processes due to inadequate numerical resolution or incomplete physical understanding.

(ii) *Incorrect statistics* used in the filtering algorithm due to a particular moment closure approximation applied to the forecast model.

In this study we assess the performance of various linear and nonlinear filters employing the stochastic parameter estimation for filtering turbulent systems on a stringent test bed mimicking various prototypical phenomena in an unambiguous mathematical framework. Our test model is derived from the system developed in [13] for filtering multiscale turbulent signals in the presence of significant model error associated with unknown processes inducing transient instabilities. Here, we focus on understanding the consequences of incorrect statistics introduced through various moment closures on the filtering skill. An attractive feature of our nonlinear test model is that its path-wise solutions and its non-Gaussian statistics, including the augmented stochastic parameters, can be obtained analytically. Moreover, as the parameters in the test model are varied, it can be tuned to mimic a wide range of realistic turbulent signals; the dynamical regimes we focus on here, all with overall ensemble mean stability, include: (i) regimes with plentiful transient instabilities, (ii) regimes with intermittent, large-amplitude bursts of transient instability interleaved with quiescent phases and (iii) laminar behavior.

We will show below how combining the analytical tractability of the test model with an appropriate choice of the "synthetic truth" signals can be used for disentangling the effects of various errors on filter performance. Results discussed in the subsequent sections include:

(i) A comprehensive understanding of the effects of model error on the filtering skill for a suite of linear and nonlinear extended Kalman filters implementing stochastic parameter estimation for different modes in the turbulent spectrum.

(ii) Identification of filters in our test suite which provide the best and the most robust real-time estimation of hidden parameters in the forecast model for a range of modes in the turbulent spectrum.

Besides benchmarking the performance of our suite of filters, these findings should be very useful in developing cheap techniques for filtering spatially extended systems involving multiple spatio-temporal scales and sparse observations [15, 16, 17, 21, 27].

The analysis presented here is structured as follows: The model and its different dynamical regimes mimicking turbulent signals are discussed in section 2. The suite of filters utilizing "on the fly" stochastic parameter estimation and different moment closure approximations are discussed in section 3. The skill of the filtering algorithms, as well as the effects of different sources of model error, are discussed in sections 5-9. Finally, section 10 contains the concluding discussion and describes future goals for research in this topic.

2. The test model and its diverse regimes of mean-stable dynamics

The nonlinear test model which we use here for filtering turbulent signals is given by the following stochastic system introduced in [13, 14]

$$
\left.\begin{aligned}
&(a) \qquad \mathrm{d}u(t) = \big[(-\hat{\gamma} - \gamma(t) + i\omega)u(t) + b(t) + f(t)\big]\mathrm{d}t + \sigma_u \mathrm{d}W_u(t),\\
&(b) \qquad \mathrm{d}b(t) = \big[(-\gamma_b + i\omega_b)(b(t) - \hat{b})\big]\mathrm{d}t + \sigma_b \mathrm{d}W_b(t),\\
&(c) \qquad \mathrm{d}\gamma(t) = -d_\gamma \gamma(t)\mathrm{d}t + \sigma_\gamma \mathrm{d}W_\gamma(t),
\end{aligned}\right\} \tag{1}
$$

where W_u, W_b are independent complex Wiener processes and W_γ is a real Wiener process. There are nine parameters in the system (1): two damping parameters γ_b, d_γ, two oscillation frequencies ω and ω_b, two stationary mean terms $\hat{b}$ and $\hat{\gamma}$ and noise amplitudes $\sigma_u, \sigma_b, \sigma_\gamma$; f is a deterministic forcing.

Here, we regard $u(t)$ as representing one of the resolved modes in a turbulent signal where the nonlinear mode-interaction terms are replaced by a stochastic drag $\gamma(t)$ and an additive noise term $b(t)$, as is often done in turbulence models [29, 30, 8]. The augmented system (1) was introduced in [13] for filtering multiscale turbulent signals where the stochastic parameters, $\gamma(t)$ and $b(t)$, are estimated adaptively 'on the fly' in order to improve the filtering skill. The novel feature of this approach is that the augmented dynamics of the hidden bias correction terms (1b,c) is modelled via the Ornstein-Uhlenbeck processes with finite decorrelation times.

The nonlinear system (1) has a number of attractive properties as a test model in our analysis. Firstly, the system (1) has a surprisingly reach dynamics mimicking turbulent signals in various regimes of the turbulent spectrum. Secondly, due to the particular structure of the nonlinearity in (1), exact path-wise solutions and exact second order statistics of this non-Gaussian system can be obtained analytically, as discussed in [15, 16, 13] and recapitulated in Appendix A. The mathematical tractability of this model and its rich dynamical behavior provides a perfect test bed for analyzing effects of model error in a suite of filters introduced in §3.

2.1. Regimes of mean stable dynamics

We focus here on describing the most interesting dynamical regimes of the system (1). The analytical formulas for path-wise solutions of (1) and its exact second-order statistics, derived previously in [15, 16, 13], are briefly recapitulated in Appendix A. All of the dynamical regimes discussed here are characterized by the mean-stability of their solutions in the sense defined below:

Definition 1.[*Global mean stability*] Given the solutions

$$\boldsymbol{x}(t,t_0) = (u(t,t_0), b(t,t_0), \gamma(t,t_0))^T$$

of the system (1) with initial condition $\boldsymbol{x}_0 \equiv (u_0, b_0, \gamma_0)^T$, the dynamics of (1) is said to be globally *mean stable* if there exists a finite constant, C, depending on $\boldsymbol{x}_0$ and t_0 such that

$$\max_{t\in[t_0,\,\infty)} \big|\,\overline{\boldsymbol{x}(t,t_0)}\,\big| < C,$$

for all times t_0; the overbar here and below denotes ensemble average.

Proposition 1.[*Mean stability of system* (1)] The dynamics of the stochastic system (1) is mean stable provided that

$$\chi = -\hat{\gamma} + \frac{\sigma_\gamma^2}{2d_\gamma^2} < 0, \tag{2}$$

where $\hat{\gamma}$ is the mean damping in $u(t)$, d_γ is the damping in $\gamma(t)$, and σ_γ is the noise variance in (1c).

Remark. The mean stability of the system (1) is controlled by the mean damping, $\hat{\gamma}$, in $u(t)$ and the dynamical properties of fluctuations about the mean damping represented by $\gamma(t)$. The additive noise term, $b(t)$, has no effect on the mean stability.

A simple proof of Proposition 1 is given in Appendix A. The proof exploits the exact formulas for the first moments of (1) and does not rely on the linearization of (1).

Definition 2.[*Decorrelation time*] Since we study only one Fourier mode here, we define the decorrelation time of each component in (1) as the time it takes for that component to decorrelate in the equilibrium statistical steady state. Mathematically, the decorrelation time of a scalar solution $x(t)$ is defined as the integral on the positive half-line (from 0 to ∞) of the absolute value of the correlation function

$$\mathcal{R}(\tau) \equiv \overline{\big(x(t) - \overline{x}\big)\big(x(t+\tau) - \overline{x}\big)^*},$$

where the overbar denotes equilibrium ensemble average and "$*$" is the complex conjugate.

Using Definition 2 and the path-wise solutions of the system (1) given in Appendix A, it is straightforward to check that the decorrelation time of $\gamma(t)$ is given by $1/d_\gamma$, and the decorrelation time of $b(t)$ is given by $1/\gamma_b$. Computation of the decorrelation time of $u(t)$ is more involved and not necessary for our purposes; it is sufficient here to use the approximate decorrelation time, $1/\hat{\gamma}$, based on the mean damping in $u(t)$.

Based on the mean-stability criterion (2), it is possible to distinguish the following:

Regimes of mean-stable dynamics of the system (1)

(I) $\sigma_\gamma, d_\gamma \gg 1$, $\sigma_\gamma/d_\gamma \sim \mathcal{O}(1)$ and $\hat{\gamma} > 0$ sufficiently large so that $\chi < 0$.

This is a regime of rapidly decorrelating $\gamma(t)$. The dynamics of $u(t)$ is dominated by frequent, short-lasting transient instabilities (see figures 1, 2). Decorrelation time of $u(t)$ is approximately $1/\hat{\gamma}$ and can vary widely. This type of dynamics is characteristic of the turbulent energy transfer range.

(II) $\sigma_\gamma, d_\gamma \sim \mathcal{O}(1)$ small, $\sigma_\gamma/d_\gamma \sim \mathcal{O}(1)$ and $\hat{\gamma} > 0$ sufficiently large so that $\chi < 0$.

In this regime the decorrelation time of $\gamma(t)$ is long. The dynamics of $u(t)$ is characterized by intermittent bursts of large-amplitude, transient instabilities followed by quiescent phases (see figure 10). This regime is characteristic of the turbulent modes in the dissipative range. Similarly to (I), decorrelation time of $u(t)$ can vary widely in this regime.

(III) $\sigma_\gamma^2/2d_\gamma^2 \gg 1$, $\sigma_\gamma \sim \mathcal{O}(1)$ and $\hat{\gamma} \gg 1$ sufficiently large so that $\chi < 0$. This regime is characteristic of the laminar modes in the turbulent spectrum (see figure 23). Here, $u(t)$ decorrelates rapidly compared to $\gamma(t)$ and the transient instabilities occur very rarely. 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 $u(t)$.

---

3. Suite of filters

We focus here on computationally cheap algorithms for filtering turbulent signals with multiple spatio-temporal scales which employ stochastic turbulence models for the unresolved scales. As pointed out in §2, the augmented stochastic forecast model (1) for filtering individual Fourier modes of the physical system is nonlinear. A standard technique for reconstructing the observed mode $u(t)$ and estimation of the hidden stochastic parameters, γ and b, involves the linear tangent approximation of the forecast model and Kalman filtering. However, the use of the Extended Kalman Filter (EKF, [26]) is rarely justified in practice and it often leads to divergent solutions on turbulent signals, as will be shown in §7-9.

One alternative approach to this problem, besides deriving exact model statistics as in SPEKF below, is to avoid the linearization in propagating the prior statistics in the filter through a moment closure approximation applied to the nonlinear forecast model. Despite the relative simplicity of filters obtained in this way, they are not uniquely optimal in the same sense as Kalman filtering for linear systems. One potential deficiency, common to all nonlinear extensions of the Kalman filter, is that the filter update rules only involve the second-order statistics. Clearly, the gaussianity of the signal can be lost due to the nonlinearity and such filters may not be optimal anymore. Moreover, if the moment closures are used in derivation of the second order-statistics in the nonlinear filters, additional model error is introduced to the problem.

In this section we first outline the properties of basic Kalman filtering with a linear tangent approximation, leading to the Tangent Extended Kalman Filter (TEKF). We then develop a general framework for filtering with quadratic models and introduce four nonlinear filters based on different moment closures. Stochastic Parameterization Extended Kalman Filter (SPEKF), which we use as a benchmark here, exploits the particular structure of the forecast model (1) and uses exact formulas for the second-order statistics without any moment closures. Performance of these filters and the effects of various model errors will be tested and discussed in the subsequent sections 7-9.

3.1. Basic filtering for linear models

The classical discrete Kalman filter [25] is a two-step, predictor-corrector method which incorporates noisy observations of a physical system at a discrete sequence of times, $T_M =$

$\{t_1, t_2, ..., t_M\}$, in order to adjust the model prediction for the state of the system at the same times T_M. In its original formulation the model dynamics is described by a linear stochastic process, all uncertainties and initial conditions have Gaussian distributions, and the filtering process can be described uniquely in terms of the mean state and the covariance matrix. In such a case the model forecast, $\{\boldsymbol{x}_m\}_{m=1,...,M}$, and the observations, $\{\boldsymbol{v}_m\}_{m=1,...,M}$, both recorded at the same sequence of times T_M can be written as:

$$\textit{model forecast}: \quad \boldsymbol{x}_{m+1} = F_{m+1}\,\boldsymbol{x}_m + \Phi_{m+1} + \boldsymbol{\sigma}_{m+1}, \tag{3}$$

$$\textit{observation}: \quad \boldsymbol{v}_{m+1} = G\,\boldsymbol{x}_{m+1} + \boldsymbol{\sigma}^o_{m+1}, \tag{4}$$

where $\boldsymbol{x}_{m+1}$ represents the n-dimensional state of the system at time t_{m+1}, F_{m+1} is a linear deterministic operator that maps $\boldsymbol{x}_m$ forward in time, Φ_{m+1} is the deterministic forcing at t_{m+1}, and $\boldsymbol{\sigma}_{m+1}$ is an n-dimensional white Gaussian vector at time t_{m+1}. For simplicity in exposition the observations $\{\boldsymbol{v}_m\}_{m=1,...,M}$ of the true state are modelled here by a linear transformation with the *observation operator* G and additive white Gaussian noise $\boldsymbol{\sigma}^o$.

The Kalman solution to filtering the linear system (3), (4) at the discrete sequence of times T_M produces an optimal estimate (see [25]) of the *posterior* mean and covariance of the system state at t_{m+1} based on the observation $\boldsymbol{v}_{m+1}$ and the model prediction prior to incorporating the observation. The prior mean and covariance at time t_{m+1} are denoted by $\bar{\boldsymbol{x}}_{m+1|m}$ and $R_{m+1|m}$, while the posterior mean and covariance are denoted by $\bar{\boldsymbol{x}}_{m+1|m+1}$ and $R_{m+1|m+1}$.

The second order statistics in the Kalman filter is updated iteratively as follows:

Initialization:

$$\bar{\boldsymbol{x}}_{0|0} = E[\boldsymbol{x}_0], R_{0|0} = Var[\boldsymbol{x}_0], \tag{5}$$

Prior update (model forecast):

$$\bar{\boldsymbol{x}}_{m+1|m} = F_{m+1}\bar{\boldsymbol{x}}_{m|m} + \Phi_{m+1}, \tag{6}$$

$$R_{m+1|m} = F_{m+1}R_{m|m}F^*_{m+1} + \boldsymbol{\sigma}_{m+1}\boldsymbol{\sigma}^T_{m+1}. \tag{7}$$

Posterior update (observation incorporated):

$$\bar{\boldsymbol{x}}_{m+1|m+1} = \bar{\boldsymbol{x}}_{m+1|m} + K_{m+1}(\boldsymbol{v}_{m+1} - G\,\bar{\boldsymbol{x}}_{m+1|m}), \tag{8}$$

$$R_{m+1|m+1} = (\mathcal{I} - K_{m+1}G)R_{m+1|m}, \tag{9}$$

$$K_{m+1} = R_{m+1|m}G^*\left(GR_{m+1|m}G^* + \boldsymbol{\sigma}^o_{m+1}{\boldsymbol{\sigma}^o}^T_{m+1}\right)^{-1}. \tag{10}$$

The operator K_{m+1} is referred to as the Kalman gain at t_{m+1} and the asterisk, "$*$" in (7) and (10) denotes the complex conjugate.

3.1.1. Tangent EKF

The procedure described below is the simplest and historically the earliest extension of the Kalman filter to deal with nonlinear stochastic forecast models (see [26]). We will refer to such a procedure as the Tangent Extended Kalman Filter (TEKF) in order to differentiate it from other extensions of the Kalman filter to nonlinear models discussed in the next section.

Assume that the continuous-time, imperfect forecast model for the dynamics of a physical system is given by the following nonlinear (Ito) stochastic differential equation (see, e.g., [12, 23])

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}, t)\mathrm{d}t + \mathrm{d}\mathbf{W}(t), \quad \mathbf{W}(t) \sim \mathcal{N}_n(0, \Sigma(t)), \tag{11}$$

where $\boldsymbol{x}$ is the n-dimensional state vector, $\boldsymbol{f}$ denotes the deterministic part of the model, and $\mathbf{W}(t)$ is an n-dimensional Wiener process with covariance Σ. The stochastic noise accounts for the unresolved processes in the forecast model. The sequence of observations, $\{\boldsymbol{v}_m\}_{m=1,\dots,M}$, where each $\boldsymbol{v}_m$ is a k-dimensional vector ($k \leqslant n$), is modelled here for simplicity by a linear transformation

$$\boldsymbol{v}_m = G\,\boldsymbol{x}_m + \mathbf{V}(t_m), \quad \mathbf{V}(t) \sim \mathcal{N}_k\big(0, Q(t)\big), \tag{12}$$

where $\mathbf{V}(t)$ a k-dimensional white Gaussian vector with covariance $Q(t)$. (The observations can be modelled by some nonlinear, time-dependent function but this complication is unnecessary here.)

The model forecast step in TEKF algorithm is obtained by linearizing the model $\boldsymbol{f}(\boldsymbol{x}, t)$ about the posterior mean $\bar{\boldsymbol{x}}_{m|m}$ at t_m and integrating the resulting tangent model between the successive observations, generating the prior $\bar{\boldsymbol{x}}_{m+1|m}$ at $t_{m+1} = t_m + \Delta t^{obs}$. Consequently, the forecast step in TEKF follows the general steps (6)-(10) with

$$F_{m+1} = e^{A_{m|m}\Delta t^{obs}}, \tag{13}$$

$$R_{m+1|m} = e^{A_{m|m}\Delta t^{obs}}\left(R_{m|m} + \int_{t_m}^{t_{m+1}} e^{-A_{m|m}\,s}\,Q(s)\,e^{-A^T_{m|m}\,s}\,\mathrm{d}s\right)e^{A^T_{m|m}\Delta t^{obs}}, \tag{14}$$

where $A_{m|m} = \nabla\boldsymbol{f}(\bar{\boldsymbol{x}}_{m|m})$ is the the Jacobian of $\boldsymbol{f}$ at t_m evaluated at the posterior mean $\boldsymbol{x}_{m|m}$. Details of implementation of TEKF for the forecast model (1) are discussed in Appendix B.1.

3.2. Nonlinear filters for quadratic models

We introduce here general equations for the evolution of the first two moments of the stochastic system (11) which provide a setting for deriving extensions of the Kalman filter for nonlinear forecast models through various moment closure approximations. The discussion is restricted here to quadratic models, such as the system (1), but this framework readily generalizes to any finite dimension.

The moment closure approximations discussed here introduce additional errors into the filtering algorithms through incorrect statistics. It is important to understand when these errors are large enough to compete with errors introduced by incorrectly modelled dynamics or inadequate parameterization. We will show later that the effects of these two sources of model error can be disentangled by using the same system for both generating the synthetic "truth" signal and for filtering.

To this end, consider an imperfect quadratic forecast model (11) with the deterministic part, $\boldsymbol{f}$, in the form

$$\boldsymbol{f}(\boldsymbol{x}, t) = \hat{L}(t)\boldsymbol{x} + \mathbf{B}(\boldsymbol{x}, \boldsymbol{x}, t) + \bar{\mathbf{F}}(t), \tag{15}$$

where $\boldsymbol{x}$ is an n-dimensional state vector, $\hat{L}$ is a linear operator, $\mathbf{B}$ is a bilinear function, and $\bar{\mathbf{F}}$ is a deterministic forcing. In what follows we will skip the explicit dependence on time in $\boldsymbol{f}$ in order to simplify the notation.

It can be easily shown (e.g., [23]) that by adopting an analogue of the average Reynolds decomposition of the state vector, $\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{x}'$, such that $\overline{\boldsymbol{x}'} = 0$ and $\overline{\bar{x}_i x'_j} = 0$, the evolution of the mean $\bar{\boldsymbol{x}}$ and covariance $R \equiv \overline{\boldsymbol{x}'\,\boldsymbol{x}'^T}$ of the process $\boldsymbol{x}$ satisfying (11) is given by

$$\begin{cases} a) \quad \dot{\bar{\boldsymbol{x}}} = \boldsymbol{f}(\bar{\boldsymbol{x}}) + \overline{\mathbf{B}(\boldsymbol{x}', \boldsymbol{x}')}, \\[2ex] b) \quad \dot{R} = R\,A^T(\bar{\boldsymbol{x}}) + A(\bar{\boldsymbol{x}})R + \Sigma + \overline{\boldsymbol{x}'\,\mathbf{B}^T(\boldsymbol{x}', \boldsymbol{x}') + \mathbf{B}(\boldsymbol{x}', \boldsymbol{x}')\,\boldsymbol{x}'^T}, \end{cases} \tag{16}$$

where A is the Jacobian of $\boldsymbol{f}$ evaluated at $\bar{\boldsymbol{x}}$, i.e., $A(\bar{\boldsymbol{x}}) \equiv \nabla\boldsymbol{f}(\bar{\boldsymbol{x}})$ and the overbar denotes an ensemble average.

For a general quadratic forecast model, solving the equations (16) requires the knowledge of the probability density $p(\boldsymbol{x}, t)$ associated with the process $\boldsymbol{x}$ satisfying (11). If $p(\boldsymbol{x}, t)$ is unknown, some type of moment closure is usually required. In our framework this is equivalent to making certain assumptions in (16) about the terms

$$\overline{\mathbf{B}(\boldsymbol{x}', \boldsymbol{x}')} \quad \text{and} \quad \overline{\boldsymbol{x}'\,\mathbf{B}^T(\boldsymbol{x}', \boldsymbol{x}') + \mathbf{B}(\boldsymbol{x}', \boldsymbol{x}')\,\boldsymbol{x}'^T},$$

which include the second and third moments, respectively. It is important to understand when errors associated with these approximations are important enough to compete with errors due to an imperfect model. A very attractive feature of the test model (1) is that its mean and covariance can be derived exactly due to the particular form of nonlinearity. In sections 6-9 we will argue that this property allows for unambiguous analysis of the effects of incorrect statistics and imperfect models on filter performance.

We focus here on four nonlinear filters which will be compared with TEKF in the subsequent sections. All of these filters, except for SPEKF discussed first, introduce additional model error due to various moment closure approximations applied to (16).

3.2.1. Stochastic Parameterization Kalman Filter (SPEKF)

This filter was extensively discussed in [13, 14, 21]. SPEKF uses the exact second-order statistics including the augmented stochastic parameters with no moment closure approximations. The exact analytical formulas for the mean and covariance in SPEKF were derived in [13]; they can be found due to the particular structure of the quadratic nonlinearity in the forecast model (1). The iteration of the SPEKF algorithm is analogous to the steps discussed for the Kalman filter except that an equivalent of the full system (16) is used for propagating the prior statistics.

3.2.2. Gaussian Closure Filter (GCF)

The quasi-Gaussian closure approximation implies neglecting the third and higher moments of the probability density $p(\boldsymbol{x}, t)$ associated with the process $\boldsymbol{x}$ satisfying (11). For quadratic models only the third moments have to be neglected in (16b), i.e., GCF assumes

$$\overline{\boldsymbol{x}'\,\mathbf{B}^T(\boldsymbol{x}',\boldsymbol{x}') + \mathbf{B}(\boldsymbol{x}',\boldsymbol{x}')\,\boldsymbol{x}'^T} = 0. \tag{17}$$

The closure (17) results in a fully coupled dynamical system for the second order statistics given by

$$\begin{cases} \dot{\bar{\boldsymbol{x}}} = \boldsymbol{f}(\bar{\boldsymbol{x}}) + \overline{\mathbf{B}(\boldsymbol{x}',\boldsymbol{x}')}, \\[2ex] \dot{R} = R\,A^T(\bar{\boldsymbol{x}}) + A(\bar{\boldsymbol{x}})\,R + \Sigma, \qquad A(\bar{\boldsymbol{x}}) = \nabla \boldsymbol{f}(\bar{\boldsymbol{x}}). \end{cases} \tag{18}$$

The system (18) represents the exact evolution of the second order statistics for any Gaussian process where (17) is satisfied identically. Otherwise, the closure introduces additional error into the filtering procedure.

In GCF the prior mean and covariance, $\bar{\boldsymbol{x}}_{m+1|m},\ R_{m+1|m}$, are given by the solution of (18) evaluated at t_{m+1} with the initial condition: $\bar{\boldsymbol{x}}_{t_m} = \boldsymbol{x}_{m|m}, R_{t_m} = R_{m|m}$. The posterior analysis at each observation time t_m is carried out in the same way as for the Kalman filter using equations (8)-(10). Details of implementation of GCF on the quadratic test model (1) are discussed in Appendix B.4.

3.2.3. Deterministic Mean Filter (DMF)

In the engineering literature this algorithm and the tangent approximation leading to TEKF (see §3.1.1) are referred to, rather confusingly, as the Extended Kalman Filter (e.g., [47, 39, 41, 9]). We refer to the algorithm below as DMF in order to avoid the confusion with TEKF.

In DMF the prior mean and covariance, $\bar{\boldsymbol{x}}_{m+1|m},\ R_{m+1|m}$, are updated by solving

$$\begin{cases} a) \quad \dot{\bar{\boldsymbol{x}}} = \boldsymbol{f}(\bar{\boldsymbol{x}}), \\[2ex] b) \quad \dot{R} = R\,A^T(\bar{\boldsymbol{x}}) + A(\bar{\boldsymbol{x}})R + \Sigma, \qquad A(\bar{\boldsymbol{x}}) \equiv \nabla \boldsymbol{f}(\bar{\boldsymbol{x}}), \end{cases} \tag{19}$$

on the time interval $[t_m,\ t_{m+1}]$ with initial conditions $\boldsymbol{x}(t_m) = \boldsymbol{x}_{m|m},\ R(t_m) = R_{m|m}$.

Similarly to GCF, DMF neglects the third and higher moments in the evolution of the covariance. However, DMF uses only the *deterministic part* of the forecast model (11) to propagate the mean, effectively neglecting correlations between variables by assuming that

$$\overline{\boldsymbol{f}(\boldsymbol{x})} = \boldsymbol{f}(\bar{\boldsymbol{x}}).$$

Within the general framework for quadratic models (16) this ad-hoc closure corresponds to imposing

$$\overline{\boldsymbol{x}'\,\mathbf{B}^T(\boldsymbol{x}',\boldsymbol{x}') + \mathbf{B}(\boldsymbol{x}',\boldsymbol{x}')\,\boldsymbol{x}'^T} = 0, \quad \text{and} \quad \overline{\mathbf{B}(\boldsymbol{x}',\boldsymbol{x}')} = 0, \tag{20}$$

in the equations for the evolution of the prior mean and covariance $\bar{\boldsymbol{x}}_{m+1|m}$ and covariance $R_{m+1|m}$

Remark. (i) The ad-hoc moment closure in (19), assuming that the second moments in the equation for the mean (19a) vanish, is inconsistent with the nontrivial evolution of the covariance in (19b). (ii) The covariance in DMF and GCF satisfies the same evolution equation. However, in GCF the equations for the mean and covariance are coupled through the second moments and the Jacobian, $A(\bar{\boldsymbol{x}})$, in GCF is evaluated at a different mean.

The posterior update in DMF, incorporating observations at t_{m+1}, is carried out in the same fashion as in the Kalman filter (cf. (8)-(10)). Details of the implementation of DMF on the test model (1) are discussed in Appendix B.2.

3.2.4. Split Deterministic Mean Filter (SDMF)

This filter uses the same moment closures as DMF but the propagation of the prior mean and covariance are fully decoupled here. Similarly to DMF, the prior mean, $\bar{\boldsymbol{x}}_{m+1|m}$, in SDMF is updated through the deterministic part of the forecast model by solving

$$\dot{\bar{\boldsymbol{x}}} = \boldsymbol{f}(\bar{\boldsymbol{x}}), \tag{21}$$

on a time interval $[t_m,\ t_{m+1}]$ with initial condition $\boldsymbol{x}(t_m) = \bar{\boldsymbol{x}}_{m|m}$.

The prior estimate of the covariance matrix, $R_{m+1|m}$, at t_{m+1} is computed, as in TEKF, by integrating (19b) between successive observations with the Jacobian $A(\bar{\boldsymbol{x}})$ evaluated at $\bar{\boldsymbol{x}}_{m|m}$, i.e.,

$$R_{m+1|m} = e^{A_{m|m}\Delta t^{obs}}\left(R_{m|m} + \int_{t_m}^{t_m+1} e^{-A_{m|m}s}\hat{Q}(s)e^{-A^T_{m|m}s}\mathrm{d}s\right)e^{A^T_{m|m}\Delta t^{obs}}, \tag{22}$$

where

$$A_{m|m} = \nabla\boldsymbol{f}(\bar{\boldsymbol{x}}_{m|m}), \quad \Delta t^{obs} = t_{m+1} - t_m. \tag{23}$$

The posterior update is carried out as in the Kalman filter through (8)-(10). Details of implementation of SDMF on the test model (1) are discussed in Appendix B.3.

4. Generation of the synthetic "truth" signal

As already mentioned earlier, our main focus here is to understand the effects of incorrect statistics used in the filters employing stochastic parameter estimation on the filter performance (cf. §3). We thus use same system (1) for both generating the synthetic truth and for filtering the observed signal. This approach avoids overshadowing the effects of model error by incorrect dynamics used in the filters. The system parameters used for filtering the observed signal, i.e.,

$$\Lambda^M = \{\omega^M,\ \sigma_u^M, \quad \hat{\gamma}^M,\ d_\gamma^M\ ,\sigma_\gamma^M, \quad \hat{b}^M,\ \gamma_b^M,\ \omega_b^M,\ \sigma_b^M\}, \tag{24}$$

are, in principle, independent of the ones used for generating the truth,

$$\Lambda = \{\omega,\ \sigma_u, \quad \hat{\gamma},\ d_\gamma\ ,\sigma_\gamma, \quad \hat{b},\ \gamma_b,\ \omega_b,\ \sigma_b\}. \tag{25}$$

The numerical realizations of the truth signal are obtained by integrating the exact solutions (A.1-A.3) of (1) which are derived in Appendix A.

We assume that only the 'resolved' variable $u(t)$ is directly observed. Updates of the augmented dynamics of the hidden variables, γ and b, in (1) have to be estimated adaptively during the filtering from the observations of $u(t)$ and from the forecast model. Consequently, the observation operator G used in the posterior update (8)-(10) in all filters has the form

$$\hat{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}. \tag{26}$$

The filtering is carried out for forced systems with the deterministic forcing given by

$$f(t) = A_f e^{i\omega_f t}, \tag{27}$$

where the amplitude is $A_f = 1$ and the frequency is $\omega_f = 0.15$. The number of assimilation cycles is chosen so that in each case the model is run for about 20 periods of the deterministic forcing. A number of tests of the filter performance for both correct and incorrect parameter values are discussed in §7-9.

5. Filtering with perfect model

The performance of a perfect filter, i.e., a filter which is not affected by model error, provides a benchmark for the subsequent analysis of skill of imperfect filters. As discussed in §4, we use the same system for generating the synthetic truth and for filtering the observed signal. Thus, the perfect filter here is given by SPEKF (cf. §3.2.1) which uses the nonlinear model (1) with exact second-order statistics and correctly specified parameters, i.e., $\Lambda^M = \Lambda$ (see (24), (25)).

5.1. Measure of filter skill and parameter estimation

We measure how well the perfect filter reproduces the truth signal (we refer to this property as skill) by means of the root mean square difference (RMS) between the true signal, $\{x_m\}_{m=1,\ldots,M}$, and the filtered solution, $\{\bar{x}_{m|m}\}_{m=1,\ldots,M}$, recorded at the same sequence of times of length M.

We start by introducing some notions used for quantifying filter the performance:

Definition [*RMS error*] The RMS difference (or RMS error) between the truth sequence $\{x_m\}_{m=1,\ldots,M}$ and the filtered solution $\{\bar{x}_{m|m}\}_{m=1,\ldots,M}$ is defined as

$$RMS(x) = \sqrt{\frac{1}{M}\sum_{m=1}^{M} |\bar{x}_{m|m} - x_m|^2}. \tag{28}$$

The RMS error is computed independently for each component of the filtered signal after discarding the initial transient onto the attractor.

Definition[*Filter skill*] Consider filtering a turbulent signal at a single spatial location based on a sequence of observations $\{\boldsymbol{v}_m\}_{m=1,\ldots,M}$ at times $\{t_1, \ldots, t_M\}$ and assume the truth $\{x_m\}_{m=1,\ldots,M}$ at the same time sequence is known. We say that the filter has a *skill* for filtering the observed component $u(t)$ in (1) when the RMS error between the truth and the filtered solution is smaller than the observation error.

The accuracy with which the dynamics of the unobserved components is estimated is judged by comparing the RMS errors between the estimated components and the synthetic truth.

Definition[*Filter divergence*] We define *filter divergence* as the outcome of the filtering procedure when the average RMS error in estimated signal exceeds the observation error.

Remark. Note that in our special case for a single Fourier mode $u(t)$ in (1) filter divergence means that simply trusting the observations provides a better estimate of the true signal than when running a divergent filter. When filtering a spatially extended turbulent system with sparse observations in the spatial domain the above definitions do not apply. Even if the RMS error of the filtered solution is large, there are not enough points on the physical grid for reconstructing the signal everywhere in the domain solely from the observations; in such a case filters can have significant skill while exceeding (pointwise) the observation error [15, 16, 31].

5.2. Performance of the perfect filter in various dynamical regimes

The filtering skill of the perfect model and parameter estimation in each dynamical regime outlined in §2.1 depends on the observation time Δt^{obs}, the observation noise variance r^o, and the characteristics of the dynamics of $\gamma(t)$. We distinguish four general cases characterized by different performance of the perfect filter. Each case is associated with different ratio between the decorrelation times of $u(t)$ and $\gamma(t)$. (The decorrelation time of $u(t)$ is given approximately by $1/\hat{\gamma}$, and the decorrelation time of $\gamma(t)$ is $1/d_\gamma$, as discussed in §2.1).

Case 1: $\boldsymbol{1/d_\gamma < \Delta t^{obs} \ll 1/\hat{\gamma}}$ (Regimes I, II). *Observation time small compared to the decorrelation time of $u(t)$ but longer than that for $\gamma(t)$.*

- Good skill in filtering the observed signal $u(t)$.
- Mean values of $\gamma(t)$ and $b(t)$ recovered. Extrema of instability (i.e., dominant minima of $\gamma(t)$) detected.

In this case filtering with perfect model has a high skill. The observation time here is too long to resolve the dynamics of $\gamma(t)$. However, the mean values of the parameters are recovered correctly. Accurate estimation of γ is not essential for skillful filtering of u as long as the extrema of instability are recovered (see figures 1, 3).

For extremely small observation noise the filter trusts the observations and the improvement of the filtering skill due to a good model is negligible. For sufficiently large observation noise levels the filter trusts the model and fails to detect many unstable events but it still outperforms the observations (see figure 2).

Case 2: $\boldsymbol{\Delta t^{obs} < 1/d_\gamma < 1/\hat{\gamma}}$ or $\boldsymbol{\Delta t^{obs} < 1/\hat{\gamma} < 1/d_\gamma}$ (Regimes I-III). *Observation time step shorter than decorrelation times of both $\gamma(t)$ and $u(t)$.*

- Good skill in filtering the observed signal $u(t)$.
- Mean value of $\gamma(t)$ and its dominant minima recovered well. $b(t)$ estimated well if its decorrelation time is sufficiently long (i.e., when $\Delta t^{obs} \ll 1/\gamma_b$).

This is similar to Case 1, but here the signal is sampled sufficiently frequently so that the filter learns enough about $\gamma(t)$ to capture its most important features (see figures 10-12).

Case 3: $\boldsymbol{1/\hat{\gamma} < \Delta t^{obs} \ll 1/d_\gamma}$ (Regime II-III). *Observation time step small compared to the decorrelation time of* $\gamma(t)$ but longer than that of $u(t)$.

- Only low frequency modulation recovered in $u(t)$.
- Mean value of $\gamma(t)$ and $b(t)$ recovered.

In this case the observed signal decorrelates between the successive observations and filtering does not provide substantial improvement over the observations.

Case 4: $\boldsymbol{\Delta t^{obs} \gg 1/\hat{\gamma} > 1/d_\gamma}$ or $\boldsymbol{\Delta t^{obs} \gg 1/d_\gamma > 1/\hat{\gamma}}$ (Regimes I-III). *Observation time step long compared to decorrelation times of both $\gamma(t)$ and $u(t)$.*

- Only low frequency modulation due to periodic forcing is recovered in $u(t)$.
- The estimation of $\gamma(t)$ and $b(t)$ is completely unreliable.

In this case the observed signal decorrelates between the subsequent observations and the filter dampens whatever information is acquired from the observations by the next analysis time. Consequently, only the low frequency modulation due to the deterministic forcing is recovered in the filtered solution. The flow of information from the observations to the dynamics of γ and b is weak and the parameter estimation is poor.

6. Filtering with model error

An unambiguous assessment of the effects of model error on the filtering skill is often difficult since its properties are unknown almost by definition. In the case of Bayesian filters using stochastic parameter estimation, such as the filters discussed here, the sources of model error in imperfect filters arise from:

- *Imperfect forecast models used for filtering.* These errors arise when the signal from nature is processed through a forecast model where important physical processes are parameterized due to inadequate numerical resolution or incomplete physical understanding. These model errors affect, in principle, all of the filters discussed here (i.e., SPEKF, GCF, DMF, SDMF, TEKF).

- *Incorrect statistics* used in the filters due to particular moment closure approximations. These model errors affect GCF, DMF, SDMF and TEKF in our filter suite (see §3).

The two sources of model error listed above are, in general, interdependent. In this study, however, we use the system (1) with exactly solvable statistics for both generating the synthetic truth and as the forecast model in all filters. This approach allows us to disentangle the effects of the two sources of model error by:

- *Filtering with correctly specified parameters in the model.* In this case the effects of model error on filtering which arise solely from the use of incorrect statistics can be examined.

- *Filtering with incorrectly specified filter parameters.* This configuration can be used to study the effects of model error due to inadequate parameterization of unresolved physical processes in the model by comparing solutions filtered with correct and incorrect parameters.

The above procedure for analyzing the filter skill is carried out next in the three regimes discussed in §2.1, mimicking different modes in the turbulent spectrum. This comprehensive study also provides insight into the robustness of different filters in a wide range of physical conditions; filter robustness is a critical feature in more complex cases when filtering high-dimensional, spatially-extended turbulent systems.

7. Filter performance and parameter estimation in regime with plentiful short-lasting, transient instabilities (Regime I)

In this and the subsequent two sections we study in detail the skill of the filters derived in §3 for reconstructing the observed signal $u(t)$ and estimating the stochastic parameters $\gamma(t)$ and $b(t)$. All tests are carried out for signals generated in different regimes of *mean-stable dynamics* discussed in §2.1. Detailed tests and discussion of filtering skill as a function of the observation time step, observation noise variance, and incorrect filter parameters are preceded by a summary of the most important findings.

7.1. Characteristics of the filtered signal

This configuration corresponds to regime I (cf. §2.1) of the mean-stable dynamics of the system (1) and it is characterized by very frequent but short and intermittent phases of transient instability in the dynamics of $u(t)$. The ability of the filters to detect these intermittent instabilities will largely determine the filter skill. The decorrelation time of $\gamma(t)$, given by $1/d_\gamma$, is here very small. The decorrelation time of $u(t)$, given approximately by $1/\hat{\gamma}$, is controlled by the mean damping $\hat{\gamma}$ which can vary widely in this regime; this is because the mean-stability condition (2) does not impose an upper bound on $\hat{\gamma}$ in this regime. Both the the mean damping strength of $u(t)$ and the separation between decorrelation times of $u(t)$ and $\gamma(t)$ have important consequences on the filtering skill.

We distinguish the following three subcategories of dynamical behavior in this regime:

- *Strongly damped dynamics* of $u(t)$ with small-amplitude fluctuations around low-frequency deterministic mean: In this case $\hat{\gamma} \gg 1$, the mean-stability parameter $\chi \ll -1$, and the decorellation times of both $\gamma(t)$ and of $u(t)$ are very short.

- *Weakly damped dynamics of* $u(t)$ with very large fluctuations: In this case $\hat{\gamma}$ is small and $0 < \chi \ll 1$ so that the system evolves near the boundary of mean-stable dynamics. The decorrelation time of $u(t)$ is long compared to the decorrelation time of $\gamma(t)$.

- *Moderately damped dynamics of* $u(t)$ with large-amplitude, intermittent fluctuations: Here $\hat{\gamma} \sim \mathcal{O}(1)$ and the mean-stability parameter, $-\chi \sim \mathcal{O}(1)$. The decorrelation time of $u(t)$ is well separated from the very short decorrelation time of $\gamma(t)$.

In the numerical tests we have chosen the truth signals generated from (1) with moderately damped dynamics so that the low frequency modulation is significantly affected, but not completely dominated, by the transient instabilities. The truth signal in the simulations was generated with

$$\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \qquad \omega_u = 1.78, \sigma_u = 0.5, \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5; \tag{29}$$

(see Figure 1 for an example of a path-wise solution in this configuration). For parameters (29) the decorrelation time of u is approximately $1/\hat{\gamma} \approx 0.833$, decorrelation time of γ is $1/d_\gamma = 0.05$ and the mean stability parameter is $\chi = -0.7$. In this configuration the decorrelation times of u and γ are well separated and the interplay between the low-frequency modulation and the intermittent instabilities provides a good test bed for analyzing the effects of various sources of model error on the filtering skill.

7.2. Filtering with imperfect models

Here we summarize the most important facts about filtering in regime I. Properties of filtering with a perfect filter in regime I, which is used as a benchmark here, were discussed in §5 (see Cases 1, 2, 4). A more detailed discussion of the filter skill as a function of the observation time step, noise variance and filter parameters, as well as a description of the figures referred to below, are presented in §7.2.1.

Filter performance in regime I:

- For $\{d_\gamma^M \geqslant d_\gamma,\ \sigma_\gamma^M \leqslant \sigma_\gamma, \sigma_u^M \leqslant \sigma_u\}$, i.e., possibly underestimated decorrelation time of γ and possibly underestimated noise levels in the dynamics of u and γ, the skill of the analyzed filters satisfies

$$SPEKF > GCF > DMF > \ SDMF > TEKF$$

  The differences in the filter skill increase monotonically with Δt^{obs} and r^o but the ordering in this hierarchy remains unchanged. For other parameters the skill of SPEKF is somewhat worse than that of GCF and DMF and in extreme cases it may become comparable with TEKF (figures 7-9 and 5); however, in such cases all filters tend to have a good skill.

- Effects of model error due to various moment closures are more pronounced here than in other regimes (compare especially figures 7-9 and 16-18).

- TEKF and SDMF diverge for large observation times $\Delta t^{obs} \sim 1/\hat{\gamma}$ and small observation noise levels (figure 5); they are also sensitive to errors in filter parameters (figures 7-9).

- SPEKF in regime I has the closest skill to the perfect filter and it is far superior to other filters for significantly underestimated decorrelation time, i.e., when $d_\gamma^M \gg d_\gamma$ (see figure 8). SPEKF is the best at capturing the extrema of instability (i.e., dominant minima of γ) which improves its skill (see, e.g., figure 4). Its performance deteriorates for overestimated decorrelation time, i.e., when $d_\gamma^M < d_\gamma$.

- GCF is the least sensitive filter to parameter errors in regime I but it has lower skill than SPEKF for $\{d_\gamma^M \geqslant d_\gamma,\ \sigma_\gamma^M \leqslant \sigma_\gamma, \sigma_u^M \leqslant \sigma_u\}$ (figures 7-9).

- DMF has a similar skill to GCF except for largely overestimated decorrelation times of γ, i.e., when $d_\gamma^M \ll d_\gamma$ (figure 8).

- TEKF is consistently the worst performer in regime I. The largest errors occur for $\Delta t^{obs} \sim 1/\hat{\gamma}$. This behavior is to be expected since this linearized filter fails to resolve the short, large-amplitude unstable events in the dynamics of $u(t)$ (figures 3, 4). For noise dominated signal and $\Delta t^{obs} \gg 1/\hat{\gamma}$, TEKF struggles even with recovering the low frequency mean (not shown).

- SDMF performs well only for small observation times, i.e., when $\Delta t^{obs} \ll 1/\hat{\gamma}$ (figures 5, 7-9). For small noise levels its skill is comparable with SPEKF, GCF and DMF. For increasing observation noise variance and incorrect filter parameters its skill quickly deteriorates and becomes comparable with TEKF (figures 7-9).

Parameter estimation in regime I:

- As expected, the parameter estimation in this regime is generally poor but the mean of $\gamma(t)$ and $b(t)$ are recovered well by SPEKF, GCF and DMF (figures 1-4).

- SPEKF, GCF and DMF have a similar skill for estimating $\gamma(t)$. For sufficiently small observation times they detect the extrema of instability (figure 4). In regions of transient instability the signal-to-noise ratio is large and these filters trust the observations, learning sufficiently quickly about the onset of instability. The estimation of $\gamma(t)$ in other regions is unreliable.
- The estimation of $\gamma(t)$ by TEKF and SDMF is similar to the other filters when filtering with correct parameters and small observation times (figure 3). For incorrect filter parameters, these filters are completely unreliable at estimating $\gamma(t)$ (see figure 4).
- For sufficiently short observation times, i.e., $\Delta t^{obs} < 1/\gamma_b$, the estimation of $b(t)$ is similar for all filters and characterized by a lag of the minima in the filtered signal with respect to the true dynamics (figure 3, 4).

7.2.1. Specific examples

We discuss here the results of specific tests where we examine the filter skill of our suite of filters in regime I as a function of the observation time step, observation noise variance, and incorrect filter parameters.

Path-wise examples of filtering in regime I

In figures 1, 2 we show two examples of filtering with the perfect filter which, as discussed earlier, is given by SPEKF with correct parameter values. In both examples the observed signal is filtered with the observation time step much shorter than the decorrelation time of $u(t)$. Figure 1 shows results of filtering and parameter estimation for the observation noise variance $r^o = 0.1$ which corresponds to a moderate signal-to-noise ratio for the system parameters (29). The observation time here is too long to resolve the dynamics of $\gamma(t)$ but its mean value is recovered correctly. Accurate estimation of γ is not essential for skillful filtering of u as long as the extrema of instability are recovered. In figure 2 the observed signal is dominated by noise and the filter trusts the model, failing to detect many unstable events; however, it still outperforms the observations. The estimation of $\gamma(t)$ is unsurprisingly poor but some extrema of instability are still recovered.

Figures 3 and 4 show path-wise examples of filtering with imperfect filters. Figure 3 shows examples of filtering with correct parameters for three different observation time steps. Figure 4 shows three examples of filtering with incorrect noise amplitude in the dynamics of $u(t)$ and a relatively large observation time ($\Delta t^{obs} = 0.6$, decorrelation time $1/\hat{\gamma} = 0.83$). Similarly, to the perfect filter scenario, accurate estimation of γ is not essential for skillful filtering of u as long as the extrema of instability are recovered. SPEKF, GCF and DMF clearly outperform TEKF and SDMF here. When filtering with correct parameters and large observation time steps (figure 3) the effects of particular sampling of the signal during the short-lasting unstable bursts affect all filters but SPEKF, GCF and DMF remain more skillful than SDMF and TEKF. When filtering with incorrect noise amplitudes σ_u^M, SPEKF has the best skill for recovering the extrema of instability, especially for $\sigma_u^M < \sigma_u$, but it is somewhat worse than GCF and DMF in other intervals.

Filtering skill in regime I as a function of observation time step

In figure 5 we show the average RMS errors of the filtered solutions $u(t)$ which are obtained by sampling the same truth signal with different observation time steps, Δt^{obs}, for fixed values of the observation noise variance r^o. The particular values of r^o we have chosen are such that the smallest considered value, $r^o = 0.05$, corresponds to high overall signal-to-noise ratio, while for the largest considered value, $r^o = 1$, the signal is dominated by noise. The first column in figure 5 shows the RMS errors for filtering with correct parameters. In such a case SPEKF is the perfect filter and the other filters are affected only by model errors due to incorrect statistics (see §6). Columns 2 and 3 in figure 5 show the RMS errors of the estimated solution $u(t)$ when filtering with incorrect parameters. The use of incorrect parameter values in the filters introduces additional model error, mimicking the effects of inadequate parameterization of unresolved processes. In column 2 of figure 5 all filters (except for the perfect filter) overestimate the decorrelation time of γ in the truth signal ($d_\gamma^M < d_\gamma$); in column 3 of figure 5 all filters underestimate the decorrelation time of γ in the truth signal ($d_\gamma^M < d_\gamma$).

Based on extensive numerical simulations summarized in figure 5, we draw the following conclusions:

- TEKF and SDMF diverge for sufficiently large observation times and sufficiently small observation noise levels. The skill of these filters improves for strongly damped dynamics of $u(t)$ (i.e., when $\hat{\gamma} \gg 1$, not shown).
- The effects of model error become increasingly important with increasing observation time step Δt^{obs}; the RMS error differences between filters increase with Δt^{obs} (see figure 3).
- In extreme cases the skill of SPEKF can become worse than TEKF when filtering with significantly overestimated decorrelation time of $\gamma(t)$, i.e., $d_\gamma^M < d_\gamma$, and for large observation time step (column 2 of figure 5).
- For small observation time steps ($\Delta t^{obs} \ll 1/\hat{\gamma}$) the skill of all filters is good and comparable.
- For all filters the RMS errors of the filtered signal increase monotonically with increasing observation time step Δt^{obs}. (The oscillations seen in figure 5 for $\Delta t^{obs} \sim 1/\hat{\gamma}$ are due to the particular sampling of the truth signal during large-amplitude unstable phases; see figures 3 and 4 for path-wise examples.)
- The RMS errors approach some finite value as Δt^{obs} increases for a fixed value of r^o. For weakly damped dynamics of u, long observation times $\Delta t^{obs} \sim 1/\hat{\gamma}$ and small enough r^o, the skill of all filters is comparable with the observation error.
- For noise dominated signals all filters have comparable skill. This is to be expected since in such a case all filters trust the models and the model errors due to different closures are negligible compared to the observation error.

Filtering skill in regime I as a function of observation noise variance

Here we study how the filter performance depends on the observation noise levels for fixed values of the observation time step Δt^{obs}. We summarize this analysis in figure 6 which shows the average RMS errors of the filtered signal $u(t)$ as a function of the observation variance r^o for fixed values of the observation time step and various combinations of the filter parameters. We chose four different observation time steps to illustrate the dependence of the RMS errors on the observation noise variance. The same truth signal as in the previous tests (see figure 1) was used here. Column 1 of figure 6 shows results of filtering with correct parameters in all filters; columns 2 and 3 show results of analogous computations with, respectively, overestimated and underestimated decorrelation time of $\gamma(t)$.

Based on our analysis summarized in figure 6, we make the following conclusions:

- The effects model error due to the Gaussian closure are small but become noticeable at large observation times ($\Delta t^{obs} \sim 1/\hat{\gamma}$) for increasing observation noise variance r^o.
- TEKF and SDMF diverge for small noise levels and sufficiently large observation times.
- The RMS errors grow monotonically with r^o for all filters. They all beat observations for sufficiently large observation noise levels.
- The effects of model error grow with increasing r^o; this can be inferred from the increasing differences in the RMS errors between different filters for increasing r^o.

Filtering skill in regime I as a function of filter parameters

Here we study the performance of imperfect filters when the noise amplitudes σ_γ, σ_u in the dynamics of γ and u are varied. We also analyze the filtering skill as a function of incorrect decorrelation time of γ assumed by the filters. Departures of these parameters from their true values have the most interesting effects on the filter skill. Each test is carried out for four different pairs of fixed values of the observation time and observation noise variance. The decorrelation time of the truth signal $u(t)$ used in the tests is approximately $1/\hat{\gamma} = 0.8$. Here we have chosen two different observation time steps: $\Delta t^{obs} = 0.1$ which is much shorter than the decorrelation time of u and $\Delta t^{obs} = 0.6$ which is comparable with the decorrelation time of u. The two values of the observation noise variance are chosen such that the smaller one, $r^o = 0.05$, corresponds to small overall signal-to-noise levels and $r^o = 0.5$ corresponds to moderate noise values. The four pairs of the observation time and observation noise variance are listed in table 1.

Figure 7 shows the average RMS errors for the filtered signal $u(t)$ as a function of the noise amplitude, σ_γ^M, assumed in the filter for the dynamics of $\gamma(t)$. Figure 8 shows the RMS error for the filtered signal $u(t)$ for varying decorrelation time of $\gamma(t)$, i.e., the filter

$(\Delta t^{obs}, r^o)$=(0.1, 0.05)	$(\Delta t^{obs}, r^o)$=(0.1, 0.5)
$(\Delta t^{obs}, r^o)$=(0.6, 0.05)	$(\Delta t^{obs}, r^o)$=(0.6, 0.5)

Table 1: Pairs of observation times Δt^{obs} and observation noise variance r^o used in figures 7-9.

parameter d_γ^M is varied. Finally, figure 9 shows the RMS errors of $u(t)$ as a function of the noise amplitude, σ_u^M, assumed in the filter for the dynamics of $u(t)$.

Based on the results illustrated in figures 7-9 and also figure 5 discussed earlier, we make the following points:

- For $\{d_\gamma^M \geqslant d_\gamma,\ \sigma_\gamma^M \leqslant \sigma_\gamma,\ \sigma_u^M \leqslant \sigma_u\}$, i.e., possibly underestimated decorrelation time of $\gamma(t)$ and possibly underestimated noise levels in the dynamics of $u(t)$ and $\gamma(t)$, the skill of the filters satisfies

  $$SPEKF > GCF > DMF > SDMF > TEKF$$

  For other parameter values the filter skill hierarchy is

  $$GCF \sim DMF\ > SDMF \sim TEKF$$

  and the skill of SPEKF is somewhat worse than that of GCF and DMF and it may become comparable with TEKF (figures 7-9); however, in such cases all filters have a good skill.

- The skill of SPEKF is most dramatically reduced by underestimating the decorrelation time of $\gamma(t)$ in the filter (i.e., $d_\gamma^M \gg d_\gamma$) at sufficiently large observation noise levels (figure 8).

- TEKF and SDMF are the most sensitive filters to parameter variations. These filters diverge for observation times comparable with the decorrelation time of $u(t)$ when filtering with overestimated decorrelation time of $\gamma(t)$, i.e., when $\Delta t^{obs} \sim 1/\hat{\gamma}$ and $d_\gamma^M \ll d_\gamma$ (figure 8), or when filtering with underestimated noise variances $\sigma_u^M, \sigma_\gamma^M$ (figures 7 and 9).

- GCF is less sensitive to variations of d_γ^M and σ_γ^M than SPEKF. However, SPEKF has significantly better skill than GCF for $\{d_\gamma^M \geqslant d_\gamma,\ \sigma_\gamma^M \leqslant \sigma_\gamma,\ \sigma_u^M \leqslant \sigma_u\}$ and large observation times.

8. Filter performance and parameter estimation skill in regime of intermittent large bursts of instability (Regime II)

8.1. Characteristics of the filtered signal

This configuration corresponds to regime II of mean-stable dynamics of the system (1) discussed in §2.1. Here, the damping d_γ in the dynamics of $\gamma(t)$ is much weaker than in regime I. Consequently, the intermittent phases of transient instabilities are less frequent here but they are, on average, longer lasting than in regime I. Path-wise solutions in this regime are characterized by intermittent, large amplitude bursts of instability separated by quiescent periods dominated by a low-frequency modulation due to the deterministic forcing (see figure 10).

In this regime the decorrelation time of $\gamma(t)$, given by $1/d_\gamma$, is much longer than in regime I. However, similarly to regime I, the mean damping, $\hat{\gamma}$, and decorrelation time, $1/\hat{\gamma}$, of $u(t)$ can vary widely, as long as $\chi < 0$. Consequently, the dynamics in this regime can be roughly divided into three subcategories according to the strength of the mean damping $\hat{\gamma}$:

- Strongly damped dynamics and short decorrelation times of $u(t)$, i.e., $\hat{\gamma} \gg 1$,
- Weakly damped dynamics of and long decorrelation times of $u(t)$, i.e., $0 < \hat{\gamma} < 1$,
- Moderately damped dynamics and moderate decorrelation times of $u(t)$, i.e., $\hat{\gamma} \sim 1$.

In the filtering tests carried out in this section we have chosen true signals obtained from (1) with weakly damped dynamics so that the low frequency modulation is interleaved with intermittent, large-amplitude unstable bursts in $u(t)$. For moderate and strong damping strengths in $u(t)$ the signals are too similar to those discussed in regime I.

In all simulations carried out in this section the true signal is generated from (1) with parameters

$$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \qquad \omega_u = 1.78, \sigma_u = 0.4, \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4, \tag{30}$$

so that the mean stability parameter is $\chi = -0.05$.

Path-wise solutions of (1) in regime II are characterized by drastically different dynamical phases. Consequently, the RMS errors averaged over the whole observation interval do not provide a sufficiently detailed diagnostics of filter performance. Therefore, we focus here on three distinct intervals of our truth signal (see figure 10):

Interval 1 Phase of a large-amplitude transient instability in $u(t)$ which dominates the low frequency modulation due to the deterministic forcing.

Interval 2 Phase of two successive instabilities in $u(t)$. Here, the two intervals where $\gamma(t) < 0$ occur in a short succession so that the first unstable burst is not completely damped before the next one occurs. This is a tough test for any filter since there is not enough observations between the two events to predict the next transient instability in the signal.

Interval 3 Quiescent phase. The low frequency modulation of the signal is accompanied by short, isolated, small-amplitude instabilities.

8.2. Filtering with imperfect models

Here we summarize the most important facts about filtering in regime II (cf. §2.1). Properties of filtering with the perfect filter in this regime, which is used as a benchmark here, were discussed in §5 (see Cases 1-4). A more detailed discussion of the filter skill as a function of the observation time, noise variance and filter parameters, as well as a description of the figures referred to below, are presented in §8.2.1.

Filtering skill in regime II:

- In regime II the skill of the analyzed filters satisfies

$$SPEKF \sim GCF \sim DMF \ \gg \ SDMF \sim TEKF$$

  The differences in filter skill increase monotonically with the observation time, Δt^{obs}, and observation noise variance, r^o, but this hierarchy remains unchanged (figures 13-15). The above hierarchy is also largely insensitive to errors in the filter parameters.

- TEKF and SDMF experience a catastrophic divergence which is absent only for sufficiently small observation time step (figure 13). The main cause of divergence of these two filters stems from their failure to predict the intermittent, large-amplitude unstable events in the dynamics of $u(t)$ (figures 11, 14).

- The skill of SPEKF, GCF and DMF remains good for overestimated decorrelation time of γ (i.e., when $d_\gamma^M \leqslant d_\gamma$) and it somewhat deteriorates for underestimated decorrelation time of γ; however, even in extreme cases these filters outperform TEKF and SDMF (figure 17).

- The largest skill differences between the examined filters occur when filtering with large observation times ($\Delta t^{obs} \sim 1/\hat{\gamma}$) within intervals associated with large-amplitude intermittent instabilities (figure 14).

- Effects of model error due to incorrect statistics are insignificant in this regime (compare figures 16-18 and 7-9).

- The dominant source of model error stems here from incorrect filter parameters. The RMS errors can also be significantly affected by the particular sampling of the truth signal (figure 11).

Parameter estimation in regime II:

- The estimation of the damping, $\gamma(t)$, from SPEKF, GCF and DMF is similar and is the best in regions of intermittent instabilities when the signal-to-noise ratio is high and the filters trust the observations (figures 19, 20 and 11,12). This allows SPEKF, GCF and DMF to recover the extrema of instability (i.e., dominant minima of $\gamma(t)$), resulting in a good overall skill of these filters in this regime. In quiescent intervals the estimation skill of γ deteriorates.

- The estimation of γ and b by TEKF and SDMF is completely unreliable during the large-amplitude, intermittent instabilities. For sufficiently large observation times ($\Delta t^{obs} \sim 1/\hat{\gamma}$) these filters predict erroneous phases of strongly stable dynamics (i.e., $\gamma^M(t) \gg 1$) during unstable phases in the truth signal $\gamma(t) < 0$ (figure 11)

- The estimation of $b(t)$ by SPEKF, GCF and DMF is comparable in all intervals (figures 21, 22 and 11, 12) and it is good provided that the observation time is shorter that the decorrelation time of $b(t)$.

- The estimation of $b(t)$ by TEKF and SDMF is completely unreliable within intervals associated with large-amplitude instabilities (figures 11, 12 and 21, 22).

8.2.1. Specific examples

Here we analyze the performance of different filters as a function of the observation time step, the observation noise variance and incorrect parameter values used in the filters from our suite (cf. §3). Properties of filtering with the perfect filter in this regime, which is used as a benchmark here, were discussed in §5 (see Cases 1-4).

Path-wise examples of filtering in regime II

In figures 10-12 we show three path-wise examples of filtering in this regime; all of them for correct filter parameters. Figure 10 illustrates filtering and parameter estimation using the perfect filter which is given by SPEKF with correct parameter values. The observed signal is filtered for moderate observation noise variance, r^o, and the observation time step, Δt^{obs}, much shorter than the decorrelation times of u, γ and b in (1). The outcome of the filtering in the three intervals discussed in §8.1 is shown in separate insets. The filtering skill of the observed component, $u(t)$, is good in all three intervals. The dynamics of $\gamma(t)$ is estimated much better in the intervals 1 and 2 than in the quiescent interval 3. In particular, the extrema of instability, corresponding to the dominant minima of $\gamma(t)$, are well detected in interval 1; this fact is important for a skillful filtering of $u(t)$.

Figures 11 and 12 show path-wise examples of filtering with imperfect filters in intervals 1 and 3 (cf. §8.1). Figure 3 shows examples of filtering with correct parameters for three different observation time steps. SPEKF, GCF and DMF clearly outperform TEKF and SDMF in interval 1 which contains a large-amplitude burst of transient instability. The dominant minima of γ are well detected by SPEKF, GCF and DMF. When filtering with correct parameters and large observation time steps, the effects of specific sampling of the truth signal during the bursts of instability affect all filters but SPEKF, GCF and DMF remain more skillful than SDMF and TEKF; in fact, TEKF and SDMF erroneously predict phases of strong stability in regions of transient instability in the truth signal. The skill differences in quiescent regions are much less pronounced, as illustrated in figure 12.

Filtering skill in regime II as a function of the observation time step

In figures 13-14 we show the average RMS errors of the filtered solutions $u(t)$ as a function of the observation time step, Δt^{obs}, for different, fixed values of the observation variance, r^o, and various combinations of the filter parameters. The filtering procedure is performed by sampling the same truth signal, generated from (1) with parameters (30), with different observation times. The particular values of the observation noise variance, r^o, we have chosen are such that the smallest considered value, $r^o = 0.01$, corresponds to high overall signal-to-noise ratio, while for the largest considered value, $r^o = 2$, the signal is dominated by observation noise. In the first column of figure 13 correct parameter values are used in all filters so that the model errors arise only from the incorrect statistics. Columns 2 and 3 show results of analogous computations but with incorrect parameter values used in the filters, introducing an additional model error. In column 2 the of figure 13 all filters overestimate the decorrelation time of γ (i.e., $d_\gamma^M < d_\gamma$); in column 3 of figure 13 all filters underestimate the decorrelation time of γ (i.e., $d_\gamma^M > d_\gamma$).

Based on the results summarized in figures 13-14, we make the following points:

- For increasing observation time step, Δt^{obs}, the skill of the analyzed filters satisfies

$$SPEKF > GCF > DMF \gg SDMF \sim TEKF$$

  The RMS errors grow monotonically with Δt^{obs} for all filters and the RMS error differences between the filters increase with Δt^{obs} and with the noise variance r^o.

- The largest RMS errors between the filtered solution and the truth occur within intervals associated with large amplitude instabilities (see intervals 1 and 2 in figure 14). However, for SPEKF, GCF and DMF these errors are small relative to the signal amplitude.

- The smallest RMS errors between the filtered solution and the truth, and the smallest differences between various filters, occur within the quiescent intervals with no large-amplitude transient instabilities (see *interval 3* in figure 14)

- TEKF and SDMF diverge for sufficiently large observation times (i.e., $\Delta t^{obs} \sim 1/\hat{\gamma}$). These filters fail to detect the large amplitude instabilities in the dynamics of u and even produce erroneous super stability ($\gamma^M(t) \gg 1$) in regions of transient instability in the true signal (figure 11). SDMF and TEKF perform well only in the quiescent intervals in which all filters have a good and comparable skill (see figures 14 and 12).

- The skill of SPEKF is generally good and it is somewhat better than GCF and DMF, except for small observation times and small r^o. The skill of all filters deteriorates for underestimated noise amplitude σ_γ^M in γ (figure 16), or for severely underestimated decorrelation time of γ (see figure 17 for $d_\gamma^M \gg d_\gamma$). Note that this trend is opposite to that observed in regime I (figure 8).

- Differences in skill between GCF and DMF due to different moment closures are negligible in this regime.

Filtering skill in regime II as a function of the observation noise variance

Here we study how the filter performance depends on the observation noise levels for fixed values of the observation time Δt^{obs}. We summarize this analysis in figure 15 which shows the average RMS errors of the filtered signal, $u(t)$, as a function of the observation noise variance, r^o, for fixed values of the observation time step and various combinations of the filter parameters. We chose four different observation time steps to illustrate the dependence of the RMS errors on r^o. The same truth signal as in the previous tests was used here. Column 1 of figure 15 shows results of filtering with correct parameters in all filters; columns 2 and 3 show results of analogous computations with, respectively, overestimated and underestimated decorrelation time of γ.

We make the following points based on inspection of figure 15:

- The RMS errors of all filters increase with the observation noise variance r^o. Both TEKF and SDMF diverge for sufficiently large Δt^{obs} and r^o. These filters can also diverge for small Δt^{obs} and small r^o (e.g., figure 15 with $\Delta t^{obs} = 0.2$).

- The differences between SPEKF, GCF and DMF are negligible for increasing r^o.

Filtering skill in regime II as a function of filter parameters

Here we study the performance of imperfect filters as a function of the filter parameters which are different than those used for generating the truth.

We test the filter performance for varying noise amplitudes σ_γ^M, σ_u^M assumed by the filters for the dynamics of γ and u. We also analyze the filtering skill as a function of incorrect decorrelation time of, γ assumed by the filters. Each test is carried out for four different pairs of fixed values of the observation time step and observation noise variance. The decorrelation time of the truth signal $u(t)$ used in the tests is approximately $1/\hat{\gamma} \approx 1.8$. Here, we chose two values of the observation times: $\Delta t^{obs} = 0.1$ which is much shorter than the decorrelation time of u and $\Delta t^{obs} = 1$ which is comparable with the decorrelation time of u. The two values of the observation noise variance are chosen such that $r^o = 0.05$ corresponds to small signal-to-noise levels in the quiescent interval 3 and $r^o = 0.7$ corresponds to moderate signal-to-noise values. The four pairs of the observation time steps and observation noise variance are listed in the table 2.

$(\Delta t^{obs}, r^o)$=(0.1, 0.05)	$(\Delta t^{obs}, r^o)$=(0.1, 0.7)
$(\Delta t^{obs}, r^o)$=(1, 0.05)	$(\Delta t^{obs}, r^o)$=(1, 0.7)

Table 2: Pairs of observation times Δt^{obs} and observation noise variance r^o used in figures 16-18.

Figure 16 shows the average RMS errors for the filtered signal $u(t)$ as a function of the noise amplitude, σ_γ^M, assumed in the filters for the dynamics of $\gamma(t)$. The truth signal is the same in all simulations and the true value of σ_γ remains unchanged. Figure 17 shows the average RMS errors for the filtered signal $u(t)$ for different values of decorrelation time of $\gamma(t)$ assumed in the filters, i.e., the filter parameter d_γ^M is varied. Finally, figure 18 shows the RMS errors of $u(t)$ as a function of the noise amplitude, σ_u^M, assumed in the filters for the the dynamics of $u(t)$.

Based on the results of numerical tests summarized in figures 16-18, we make the following points:

- According to the RMS errors of the filtered signal $u(t)$ the skill of different filters satisfies

$$SPEKF \sim GCF \sim DMF \gg SDMF \sim TEKF$$

  The skill of SPEKF is somewhat worse than that of GCF and DMF for small observation times and overestimated decorrelation times (i.e., $d_\gamma^M < d_\gamma$) or for overestimated noise amplitudes $\sigma_u^M, \sigma_\gamma^M$ in the dynamics of u and γ. However, in such cases all filters are significantly better than the observation error.

- TEKF and SDMF are completely unreliable in this regime. They diverge even for correct parameter values.

- The effects of model error due to moment closures in GCF and DMF are negligible here compared to model errors introduced by incorrect parameters in the filters.

Estimation of $\gamma(t)$ and $b(t)$ in regime II

We discuss here the estimation of the stochastic parameters γ and b whose augmented dynamics in the test model (1) is hidden from observations.

Figure 19 shows the RMS errors of $\gamma(t)$ within the three distinct intervals discussed earlier (see figure 10) as a function of the observation time step Δt^{obs} and for different fixed values of the observation noise variance r^o; the range of the noise variances examined was chosen such that the smallest value, $r^o = 0.05$, corresponds to large signal-to-noise ratio in the quiescent interval 3 and for largest value, $r^o = 1$, the signal was dominated by noise in interval 3. It is important to note here that due to the large-amplitude of the unstable burst in interval 1, the signal-to-noise ratio in this interval remains large for both noise levels. Figure 20 shows the RMS errors of $\gamma(t)$ within three distinct intervals as a function of incorrect noise amplitude, σ_γ^M, assumed in the filters for the dynamics of $\gamma(t)$.

Figure 21 shows the RMS errors of $b(t)$ within the three distinct intervals discussed earlier (see figure 10) as a function of the observation time step and for different fixed values of the observation noise variance r^o. Figure 22 shows the RMS errors of $b(t)$ within the three distinct intervals as a function of incorrect noise amplitude σ_γ^M in the dynamics of γ.

We make the following points based on the results illustrated in figures 19-22:

- The estimation of $\gamma(t)$ from SPEKF, GCF and DMF is similar and is the best in intervals containing bursts of transient instabilities when the signal-to-noise ratio is high and the filters trust the observations.

- Parameter estimation by SPEKF is the least sensitive to errors in filter parameters.
- The estimation of γ and b from TEKF and SDMF is completely unreliable during the large-amplitude transient instabilities. For $\Delta t^{obs} \sim 1/\hat{\gamma}$, these filters predict erroneous phases of strongly stable dynamics $\gamma^M(t) \gg 1$.
- The estimation of $b(t)$ by SPEKF, GCF and DMF is comparable in all intervals provided that the decorrelation time of $b(t)$ is longer that the observation time step.

9. Parameter estimation in the laminar regime (Regime III)

Here, we mostly focus on the estimation of the stochastic parameters γ and b whose augmented dynamics in the test model (1) is hidden from observations.

9.1. Characteristics of the filtered signal

This configuration corresponds to regime III of mean-stable dynamics of the system (1) which was identified in §2.1. Signals in this regime are characterized by long decorrelation times of γ (in the sense that $d_\gamma \ll \sigma_\gamma$, $\sigma_\gamma \sim \mathcal{O}(1)$) and a rapidly decorrelating observed component $u(t)$ (i.e., $1/\hat{\gamma} \ll 1$). Thus, the path-wise solutions of (1) in this regime are dominated by low frequency modulation due to deterministic forcing with superimposed small-amplitude fluctuations. Transient instabilities are very rare in this configuration but they can have very large amplitudes.

In the numerical tests we have chosen the truth signals generated from (1) with parameters

$$\hat{\gamma} = 8.1,\ d_\gamma = 0.25,\ \sigma_\gamma = 1, \qquad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5; \quad (31)$$

see figure 23 for a path-wise example of solution generated with these parameters. In this configuration the observed component $u(t)$ decorrelates much faster than $\gamma(t)$. For parameters (31) the decorrelation time of u is approximately $1/\hat{\gamma} \approx 0.12$, the decorrelation time of γ is $1/d_\gamma = 4$ and the mean stability parameter is $\chi = -0.1$.

9.2. Filtering with imperfect models

Here, we summarize the most important facts about filtering and parameter estimation in regime III. Properties of filtering with a perfect filter in regime III, which is used as a benchmark here, were discussed in §5 (see Cases 2-4). A more detailed discussion of filter performance as a function of the observation time, and filter parameters, as well as a description of the figures referred to below, are presented in §9.2.1.

Filter performance in regime III:

- The filtering skill for all filters is comparable and good except for large observation times, $\Delta t^{obs} \sim 1/\hat{\gamma}$, and small observation noise levels (figure 24).
- The effects of model error due to different moment closure approximations are insignificant in this regime.

Parameter estimation in regime III: (figures 25-27)

- The estimation of $\gamma(t)$ from SPEKF, GCF, DMF and SDMF is similar. The mean value of $\gamma(t)$ and its dominant minima are recovered well, except for large observation times, $\Delta t^{obs} \sim 1/\hat{\gamma}$.
- TEKF is completely unreliable at estimating γ and b.
- The parameter estimation by SPEKF, GCF, DMF and SDMF is largely insensitive to variations in filter parameters.
- $b(t)$ is recovered well by SPEKF, GCF, DMF and SDMF provided that it decorrelates sufficiently slowly, i.e., $1/\gamma_b > \Delta t^{obs}$.

9.2.1. Specific examples

We only briefly discuss the results of specific tests where we examine the filter skill as a function of the observation time step, observation noise variance, and incorrect filter parameters.

Filtering skill in regime III as a function of observation time step

In figure 24 we show the average RMS errors of the filtered solutions $u(t)$ which are sampled with different observation time steps Δt^{obs} with correct filter parameters. In such a case SPEKF represents the perfect model (see 5) and other filters are affected only by model errors due to incorrect statistics (see §6). We show results of filtering of the same truth signal for different fixed values of the observation noise variance r^o. The particular values of r^o we have chosen are such that the smallest considered value, $r^o = 5\times10^{-5}$, corresponds to high overall signal-to-noise ratio, while for the largest considered value, $r^o = 0.5$, the signal is dominated by noise. Analogous tests for incorrect parameter values do not reveal new phenomena except for uniform changes in RMS errors for all filters.

Based on numerical tests summarized in figure 24, we make the following points:

- The RMS errors of the filtered solution $u(t)$ increase with increasing observation time step Δt^{obs}.
- All filters behave similarly in this regime for a range of observation times and observation noise variances.
- All filters can diverge for very small observation noise levels and large observation times (i.e., when $\Delta t^{obs} \sim 1/\hat{\gamma}$.)

Parameter estimation in regime III

Here, we test the skill of various filters for estimating the stochastic parameters as a function of incorrect noise amplitudes σ_γ^M, σ_u^M assumed by the filters for the dynamics of γ and u. We also analyze the filtering skill as a function of incorrect decorrelation time of γ controlled by d_γ^M. Incorrect values of these parameters have the most pronounced effects on the filter performance.

Each test is carried out for four different pairs of fixed values of the observation time and observation noise variance. The decorrelation time of the truth signal $u(t)$ used in the tests is approximately $1/\hat{\gamma} = 0.12$. We chose two values of observation time: $\Delta t^{obs} = 0.02$ which is much shorter than the decorrelation time of u and $\Delta t^{obs} = 0.08$ which is comparable with the decorrelation time of u. The two values of the observation noise variance are chosen such that $r^o = 8 \times 10^{-5}$ corresponds to small noise levels and $r^o = 8 \times 10^{-4}$ corresponds to moderate noise values. The four pairs of the observation time steps and observation noise variance are listed in the table 3 below.

$(\Delta t^{obs}, r^o) = (0.02,\ 8 \cdot 10^{-5})$	$(\Delta t^{obs}, r^o) = (0.02,\ 8 \cdot 10^{-4})$
$(\Delta t^{obs}, r^o) = (0.08,\ 8 \cdot 10^{-5})$	$(\Delta t^{obs}, r^o) = (0.08,\ 8 \cdot 10^{-4})$

Table 3: Pairs of observation times Δt^{obs} and observation noise variance r^o used in figures 25-27.

Figure 25 shows the average RMS errors for γ and b estimated from the forecast model (1) and the observations of $u(t)$ as a function of the noise amplitude σ_γ^M assumed in the filters for the dynamics of γ. The truth signal is the same in all simulations and the true value of σ_γ remains unchanged. Figure 26 shows the average RMS errors for γ and b as a function of the decorrelation time of $\gamma(t)$ assumed by the filters, i.e., the filter parameter d_γ^M is varied. Finally, figure 27 shows the RMS errors for γ and b as a function of the filter noise amplitude σ_u^M in the dynamics of u.

Based on the results of numerical tests summarized in figures 25-27, we make the following points:

- SPEKF, GCF and DMF and SDMF recover the mean value of γ and its dominant minima. There are essentially no transient instabilities in this regime and a good estimate of the mean leads to a high skill of the filtered solution $u(t)$.

- TEKF is completely unreliable at estimating $\gamma(t)$ and $b(t)$. TEKF is also very sensitive to incorrect filter parameters.

- The parameter estimation by SPEKF, GCF and DMF and SDMF is largely insensitive to variations in the filter parameters.

- The model error due to various moment closure approximations has no effect on parameter estimation in this regime. All filters, except for TEKF, are comparable with the perfect filter when filtering with correct parameters. For incorrect filter parameters the estimation skill deteriorates with SPEKF, GCF, DMF and SDMF shadowing each other.

- The additive stochastic parameter, $b(t)$, is estimated well by SPEKF, GCF and DMF and SDMF provided that its decorrelation time is sufficiently long $\Delta t^{obs} < 1/\gamma_b$.

10. Concluding discussion and future directions

We have tested the performance of a suite of nonlinear algorithms employing stochastic parameter estimation for filtering multiscale turbulent signals. Updating the parameters associated with unresolved or unknown processes in the imperfect forecast model through stochastic parameter estimation is an efficient way to increase filtering skill and model performance. The examined filtering algorithms were based on the same test model but they implemented different moment closure approximations when propagating the second-order statistics in the filtering procedure. We used as a benchmark the Stochastic Parameterization Extended Kalman Filter (SPEKF) which involves exact formulas for updating the mean and covariance of the augmented system and it systematically corrects both multiplicative and additive biases in the observed dynamics. The remaining filters introduced additional model error through the use of incorrect statistics. A comprehensive study was presented of the filter performance in the presence of model error for various combinations of the observation time step and observation noise levels. For filters employing the stochastic parameter estimation we identified the following two main sources of model error:

1) *Imperfect forecast models used for filtering.* These errors arise when the signal from nature is processed through a forecast model where important physical processes are parameterized due to inadequate numerical resolution or incomplete physical understanding.

2) *Incorrect statistics* used in the filters due to particular moment closure approximations.

These two sources of model error are, in general, interdependent. Moreover, it is not immediately obvious when the use of incorrect statistics introduces a nonnegligible error compared to the model error due to the imperfect forecast model. In order to disentangle these two effects, the synthetic "truth" signal was generated using the same test model as the one used for filtering but with different parameters. We showed that the exactly solvable test model used here can mimic various modes in the turbulent spectrum, ranging from signals with intermittent bursts of instability to laminar behavior, and its mathematical tractability allows for an unambiguous analysis of the effects of various model errors on the filtering skill. Moreover, this approach allowed for analyzing the filter skill in the presence of different model errors by:

1) *Filtering with correctly specified parameters in the model.* In this case the effects of model error on filtering which arise solely from the incorrect statistics could be examined.

2) *Filtering with incorrectly specified filter parameters.* This configuration was used to study the effects of model error due to inadequate parameterization of unresolved

physical processes in the model by comparing solutions filtered with correct and incorrect parameters.

The true dynamics of the stochastic parameters, $\gamma(t)$ and $b(t)$, in the test model was known but hidden from observations. Consequently, the skill of different filters for estimating the hidden stochastic parameters from the observations of the resolved component $u(t)$ was studied by comparing the estimated solutions with the synthetic truth.

The analysis carried out here enabled a comprehensive understanding of the filtering skill and parameter estimation for various filters implementing stochastic parameter estimation for turbulent modes in different regimes of the turbulent spectrum. The main findings of this study are:

Filtering skill:

1) For the three examined regimes characteristic of turbulent modes in the energy transfer range (regime I), dissipation range (regime II) and the laminar modes (regime III), the hierarchy of filters (from best to worst) is

$$SPEKF \gtrsim GCF \gtrsim DMF \gg SDMF \gtrsim TEKF$$

2) The effects of model error due to the moment closure approximations are most pronounced in regime I where they can be comparable to errors due to imperfect model; this occurs for moderate and large observation noise levels and large observation times. In regimes II and III the effects of incorrect statistics used in the filters are insignificant.

3) The effects of model error due to imperfect forecast model are dominant in regimes II and III; in these regimes SPEKF, GCF and DFM shadow each other.

4) TEKF and SDMF are generally unreliable. They diverge in regimes I and II for a wide range of filter parameters and observation time steps. This is mainly due to the failure of these filters to detect the intermittent, large-amplitude instabilities in the signal.

5) SPEKF is the best at capturing the extrema of instability (associated with dominant minima of $\gamma(t)$) which improves its skill. The performance of SPEKF is most sensitive to overestimated decorrelation time (i.e., $d_\gamma^M \ll d_\gamma$) in regime I. In regime II, SPEKF, GCF and DMF shadow each other.

6) GCF is generally comparable to DMF but it performs slightly better in regime I. GCF is also the most robust filter in regime I. It is the least sensitive to incorrect parameters, although SPEKF can have a better skill this regime.

Parameter estimation:

1) SPEKF is generally the best at predicting the extrema of instability of $\gamma(t)$ (i.e., the dominant minima of γ) even in cases where the unresolved processes $\gamma(t)$ and $b(t)$ decorrelate too quickly compared to the observation time step. This feature helps SPEKF improve the overall filtering skill. GCF and DMF are generally comparable to SPEKF but they are somewhat less skillful at large observation time steps.

2) TEKF and SDMF are completely unreliable at estimating the stochastic parameters in signals containing intermittent bursts of transient instability. They often predict phases of strong stability in the filtered signal $u(t)$ during phases of transient instability in the truth, i.e., when $\gamma(t) < 0$.

The results presented here should provide useful guidelines for developing cheap, skillful and robust techniques for filtering spatially extended systems with multiple spatio-temporal scales in the presence of significant model errors and sparse observations. The robustness of SPEKF and GCF for real-time stochastic parameterization of the unresolved scales across multiple modes in the turbulent spectrum might prove valuable, for example, for developing more skillful eddy parameterization schemes for ocean climate models [27].

Appendix A. Derivation of the exact criterion for mean-stable dynamics of the test model

We recapitulate here some exact analytical formulas for the path-wise solutions and the first moments of the system (1) which are necessary in deriving the mean-stability criterion (see (2) in Proposition 1). The complete set of analytical formulas for the second order statistics of (1) can be found in [13]. The proof of Proposition 1 is outlined at the end.

Path-wise solutions of the nonlinear model (1) are given by

$$b(t) = \hat{b} + (b_0 - \hat{b})e^{-\lambda_b(t-t_0)} + \sigma_b \int_{t_0}^{t} e^{-\lambda_b(t-s)}\mathrm{d}W_b(s), \tag{A.1}$$

$$\gamma(t) = \hat{\gamma} + (\gamma_0 - \hat{\gamma})e^{-d_\gamma(t-t_0)} + \sigma_\gamma \int_{t_0}^{t} e^{-d_\gamma(t-s)}\mathrm{d}W_\gamma(s), \tag{A.2}$$

$$u(t) = e^{-J(t_0,t)+\hat{\lambda}(t-t_0)}u_0 + \int_{t_0}^{t}(b(s)+f(s))e^{-J(s,t)+\hat{\lambda}(t-s)}\mathrm{d}s + \sigma_u \int_{t_0}^{t} e^{-J(s,t)+\hat{\lambda}(t-s)}\mathrm{d}W(s), \tag{A.3}$$

where b_0, γ_0, u_0 are the initial conditions at t_0 and

$$\lambda_b = -\gamma_b + i\omega_b, \quad \hat{\lambda} = -\hat{\gamma} + i\omega, \quad J(s,t) = \int_s^t (\gamma(s) - \hat{\gamma})\mathrm{d}s. \tag{A.4}$$

The mean of $u(t)$ is

$$\begin{aligned} \langle u(t)\rangle = &\big(\langle u_0\rangle - Cov(u_0, J(t_0,t))\big)e^{\hat{\lambda}(t-t_0)-\langle J(t_0,t)\rangle+\frac{1}{2}Var(J(t_0,t))} \\ &+ \int_{t_0}^{t}\big(\hat{b} + e^{\lambda_b(s-t_0)}(\langle b_0\rangle - \hat{b} - Cov(b_0, J(s,t)))\big)e^{\hat{\lambda}(t-s)-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}\mathrm{d}s \\ &+ \int_{t_0}^{t} f(s)e^{\hat{\lambda}(t-s)-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}\mathrm{d}s, \end{aligned} \tag{A.5}$$

where

$$\langle J(s,t)\rangle = \frac{1}{d_\gamma^2}\left(e^{-d_\gamma(s-t_0)} - e^{-d_\gamma(t-t_0)}\right)(\langle\gamma_0\rangle - \hat{\gamma}), \tag{A.6}$$

$$\begin{aligned} Var(J(s,t)) = &\frac{1}{d_\gamma^2}\left(e^{-d_\gamma(s-t_0)} - e^{-d_\gamma(t-t_0)}\right)Var(\gamma_0) \\ &- \frac{\sigma_\gamma^2}{d_\gamma^3}\left[1 + d_\gamma(s-t) + e^{-d_\gamma(s+t-2t_0)}\left(\cosh(d_\gamma(s-t)) - 1 - e^{2d_\gamma(s-t_0)}\right)\right], \end{aligned} \tag{A.7}$$

and

$$Cov(u_0, J(s,t)) = \frac{1}{d_\gamma}\left(e^{-d_\gamma(s-t_0)} - e^{-d_\gamma(t-t_0)}\right) Cov(u_0, \gamma_0), \tag{A.8}$$

$$Cov(b_0, J(s,t)) = \frac{1}{d_\gamma}\left(e^{-d_\gamma(s-t_0)} - e^{-d_\gamma(t-t_0)}\right) Cov(b_0, \gamma_0). \tag{A.9}$$

The variance of $u(t)$ and all components of the covariance matrix are given in [13]. It is important to note that, due to the nonlinearity of (1), solutions with Gaussian initial statistics at $t = t_0$ will not remain Gaussian for $t > t_0$.

Proof of Proposition 1

We prove here Proposition 1 (cf. §2), which determines a condition for the mean-stability of the dynamics of the system (1). We first show that (2) is a necessary condition for the mean stability of the system (1). We then show that it is also sufficient.

Consider the first term in A.5) given by

$$\big(\langle u_0\rangle - Cov(u_0, J(t_0,t))\big)e^{\hat{\lambda}(t-t_0)-\langle J(t_0,t)\rangle+\frac{1}{2}Var(J(t_0,t))}. \tag{A.10}$$

Assuming that the statistics at t_0 is bounded and $d_\gamma > 0$, the terms in the bracket in (A.10) are bounded on $[t_0, \infty)$ since $\langle u_0\rangle < \infty$ and, using (A.8), we have

$$\max_{t_0 \leqslant t < \infty} Cov(u_0, J(t_0,t)) = \frac{1}{d_\gamma} Cov(u_0, \gamma_0) < \infty. \tag{A.11}$$

It can be easily seen from (A.6) and (A.7) that the exponential term in (A.10) is bounded for any finite $t \geqslant t_0$ and it is also bounded on $[t_0, \infty)$ provided that

$$\lim_{t\to\infty} \Re e\left[\hat{\lambda}(t-t_0) - \langle J(t_0,t)\rangle + \frac{1}{2}Var(J(t_0,t))\right] < 0. \tag{A.12}$$

The condition (A.12) can be rewritten, using (A.6) and (A.7), as

$$\begin{aligned} \mathop{lim}_{t\to\infty} \Bigg\{ &-\hat{\gamma}(t-t_0) - \frac{1}{d_\gamma^2}\left(1 - e^{-d_\gamma(t-t_0)}\right)\left(\langle\gamma_0\rangle - \hat{\gamma} + \frac{1}{2}Var(\gamma_0)\right) \\ &- \frac{\sigma_\gamma^2}{2d_\gamma^3}\left[1 - d_\gamma(t-t_0) + e^{-d_\gamma(t-t_0)}\left(\cosh(d_\gamma(t-t_0)) - 2\right)\right]\Bigg\} < 0. \end{aligned} \tag{A.13}$$

For $t \gg 1$, the terms in the brackets can be approximated as

$$\Re e[\hat{\lambda}(t-t_0) - \langle J(t_0,t)\rangle + \frac{1}{2}Var(J(t_0,t))] = \left(-\hat{\gamma} + \frac{\sigma_\gamma^2}{2d_\gamma^2}\right)t + \mathcal{O}(1). \tag{A.14}$$

Clearly, the condition (A.12) is satisfied when

$$-\hat{\gamma} + \frac{\sigma_\gamma^2}{2d_\gamma^2} < 0, \tag{A.15}$$

as claimed. Note that this condition cannot be satisfied when $\hat{\gamma} < 0$.

We now show that the condition (A.15) is also sufficient for mean stability of the system (1). Consider the second part of (A.5) given by the integral

$$\int_{t_0}^{t} \big(\,\hat{b} + e^{\lambda_b(s-t_0)}(\langle b_0\rangle - \hat{b} - Cov(b_0, J(s,t)))\,\big)e^{\hat{\lambda}(t-s)-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}\mathrm{d}s, \tag{A.16}$$

and two functions $\mathcal{A}_1(t,t_0)$, $\mathcal{A}_2(t,t_0)$ which are bounded on $[t_0,\ \infty)$ and given by

$$\mathcal{A}_1(t,t_0) = \max_{s\in[t_0,t]}\big|\,\hat{b} + e^{\lambda_b(s-t_0)}(\langle b_0\rangle - \hat{b} - Cov(b_0, J(s,t)))\big|, \tag{A.17}$$

and

$$\mathcal{A}_2(t,t_0) = \max_{s\in[t_0,t]} \Re e\left[-\langle J(s,t)\rangle + \frac{1}{2}Var(J(s,t)) - \frac{\sigma^2}{2d_\gamma^2}(t-s)\right]. \tag{A.18}$$

Boundedness of these functions can be easily deduced from (A.6)-(A.9). In particular, we can use the following estimates (for $d_\gamma > 0$)

$$\mathcal{A}_1(t,t_0) \leqslant \frac{1}{d_\gamma^2}\big|Var(\gamma_0)/2 - \langle\gamma_0\rangle + \hat{\gamma}\big| + \frac{\sigma_\gamma^2}{2d_\gamma^3}, \tag{A.19}$$

and

$$\mathcal{A}_2(t,t_0) \leqslant |\hat{b}| + |\langle b_0\rangle| + \frac{1}{d_\gamma}\,|Cov(b_0,\gamma_0)|. \tag{A.20}$$

Using the functions $\mathcal{A}_1, \mathcal{A}_2$ and assuming that (A.15) holds, we obtain the following estimate

$$\begin{aligned}\max_{t\in[t_0,\infty)}\left|\int_{t_0}^{t} \big(\,\hat{b} + e^{\lambda_b(s-t_0)}(\langle b_0\rangle - \hat{b} - Cov(b_0, J(s,t)))\,\big)e^{\hat{\lambda}(t-s)-\langle J(s,t)\rangle+\frac{1}{2}Var(J(s,t))}\mathrm{d}s\right| \leqslant \\ \max_{t\in[t_0,\infty)}\left(\mathcal{A}_1(t,t_0)\,e^{\mathcal{A}_2(t,t_0)}\int_{t_0}^{t} e^{\left(-\hat{\gamma}+\frac{\sigma_\gamma^2}{2d_\gamma^2}\right)(t-s)}\mathrm{d}s\right) < \infty,\end{aligned} \tag{A.21}$$

where we use the fact that $|\int h(t,s)\mathrm{d}s| \leqslant \int |h(t,s)|\mathrm{d}s$. Analogous argument holds for the second integral in (A.5) as long as the (scalar) forcing remains bounded, i.e., there exists a constant C such that $\max_{t\in[t_0,\ \infty)}|f(t)| < C$. This completes the proof.

Appendix B. Details of filter implementations

Appendix B.1. Tangent EKF (TEKF)

This is the classical procedure of deriving the EKF in which the system (1) is linearized about the posterior mean at the previous observation time $(u_{m|m}, b_{m|m}, \gamma_{m|m})^T$, leading to a system for the expected value $\mathbf{U}^T = (\bar{u}, \bar{b}, \bar{\gamma})^T$ in the form

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} = L_{m|m}\mathbf{U} + \mathbf{\Phi}(t), \tag{B.1}$$

where

$$L_{m|m} = \begin{pmatrix} \lambda_{m|m} & 1 & -u_{m|m} \\ 0 & \lambda_b & 0 \\ 0 & 0 & -d_\gamma^M \end{pmatrix}, \quad \lambda_{m|m} = -\gamma_{m|m} + i\,\omega_u^M, \lambda_b = -\gamma_b^M + i\,\omega_b^M, \tag{B.2}$$

and the inhomogeneity $\mathbf{\Phi}(t) = (f(t), 0, 0)^T$ represents the deterministic forcing.

The prior mean at the next time step $(u_{m+1|m}, b_{m+1|m}, \gamma_{m+1|m})^T$ is obtained by integrating (B.1) between successive observation times with initial condition

$$(\bar{u}(t_m), \bar{b}(t_m), \bar{\gamma}(t_m))^T = (\bar{u}_{m|m}, \bar{b}_{m|m}, \bar{\gamma}_{m|m})^T,$$

which leads to

$$\begin{pmatrix} \bar{u}_{m+1|m} \\ \bar{b}_{m+1|m} \\ \bar{\gamma}_{m+1|m} \end{pmatrix} = \begin{pmatrix} e^{\lambda_{m|m}\Delta t} & \frac{e^{\lambda_b \Delta t} - e^{\lambda \Delta t}}{\lambda_b - \lambda_{m|m}} & -u_{m|m}\frac{e^{\lambda\Delta t} - e^{-d_\gamma \Delta t}}{d_\gamma - \lambda_{m|m}} \\ 0 & e^{\lambda_b \Delta t} & 0 \\ 0 & 0 & e^{-d_\gamma \Delta t} \end{pmatrix} \begin{pmatrix} \bar{u}_{m|m} \\ \bar{b}_{m|m} \\ \bar{\gamma}_{m|m} \end{pmatrix} + \begin{pmatrix} \tilde{f} \\ 0 \\ 0 \end{pmatrix} \tag{B.3}$$

where

$$\tilde{f} = \int_{t_m}^{t_{m+1}} f(s)\, e^{\lambda_{m|m}(t_{m+1}-s)} \mathrm{d}s. \tag{B.4}$$

The covariance is updated in the same way as in SDMF using the linear equation (B.1) with the Jacobian $L_{m|m}$ rewritten in real variables and integrated between successive observation times, leading to

$$R_{m+1|m} = e^{A_{m|m}\Delta t} R_{m|m} e^{A_{m|m}^T \Delta t} + \int_{t_m}^{t_{m+1}} e^{A_{m|m}(t_{m+1}-s)} Q\, e^{A_{m|m}^T (t_{m+1}-s)} \mathrm{d}s. \tag{B.5}$$

Appendix B.2. Deterministic Mean Filter (DMF)

Following the derivation in §3, the mean $(\bar{u}, \bar{b}, \bar{\gamma})$ and covariance matrix

$$R(t) = \begin{pmatrix} u' \\ b' \\ \gamma' \end{pmatrix} \cdot \big(u', b', \gamma'\big) \tag{B.6}$$

in DMF is propagated using

$$\begin{cases} (a) \quad \dot{\bar{u}}(t) = (-\bar{\gamma}(t) + i\,\omega_u^M)\bar{u}(t) + \bar{b}(t) + f(t), \\ (b) \quad \dot{\bar{b}}(t) = (-\gamma_b^M + i\omega_b^M)(\bar{b}(t) - \hat{b}^M), \\ (c) \quad \dot{\bar{\gamma}}(t) = -d_\gamma^M(\bar{\gamma}(t) - \hat{\gamma}^M), \\ (d) \quad \dot{R}(t) = AR(t) + R(t)A^T + Q, \end{cases} \tag{B.7}$$

where A is the Jacobian of (B.7) evaluated at the current mean $(\bar{u}, \bar{b}, \bar{\gamma})^T$, which is given by

$$A\big(\bar{u}(t), \bar{b}(t), \bar{\gamma}(t)\big) = \begin{pmatrix} -\bar{\gamma}(t) & -\omega^M & 1 & 0 & -\Re e[\bar{u}(t)] \\ \omega^M & -\bar{\gamma}(t) & 0 & 1 & -\Im m[\bar{u}(t)] \\ 0 & 0 & -\gamma_b^M & -\omega_b^M & 0 \\ 0 & 0 & \omega_b^M & -\gamma_b^M & 0 \\ 0 & 0 & 0 & 0 & -d_\gamma^M \end{pmatrix}, \tag{B.8}$$

and $Q = diag\big[(\frac{1}{2}((\sigma_u^M)^2, \frac{1}{2}(\sigma_u^M)^2, \frac{1}{2}(\sigma_b^M)^2, \frac{1}{2}(\sigma_b^M)^2, (\sigma_\gamma^M)^2)\big]$ is the covariance matrix.

Appendix B.3. Split Deterministic Mean Filter (SDMF)

In this filter the prior mean and covariance are updated separately. Here the mean is updated first by solving (B.7a-c) on a time interval $[t_m,\ \ t_{m+1}]$ with initial condition given by the posterior mean at t_m, i.e., $\big(\bar{u}(t_m), \bar{b}(t_m), \bar{\gamma}(t_m)\big)^T = \big(\bar{u}_{m|m}, \bar{b}_{m|m}, \bar{\gamma}_{m|m}\big)^T$.

The covariance is updated using the linear equation (B.7) with the Jacobian A evaluated at the posterior mean, i.e., $A(t_m) = A_{m|m}$, and integrated between successive observation times, leading to

$$R_{m+1|m} = e^{A_{m|m}\Delta t} R_{m|m} e^{A^T_{m|m}\Delta t} + \int_{t_m}^{t_{m+1}} e^{A_{m|m}(t_{m+1}-s)} Q e^{A^T_{m|m}(t_{m+1}-s)} \mathrm{d}s. \tag{B.9}$$

Appendix B.4. Gaussian Closure Filter (GCF)

In a Gaussian Closure Filter the prior statistics is propagated using a nonlinear dynamical system which is obtained by neglecting third and higher moments in the probability distribution associated with the system (1). The posterior update is carried out using the steps (8)-(10) as in the other filters. The second-order statistics can be obtained directly from the system (16) by substituting (1) for $\boldsymbol{f}(\boldsymbol{x}, t)$. Here, we sketch an alternative derivation specific to the system (1).

Given the system (1), we use the average Reynolds decomposition to represent all variables in (1) as the sum of a mean and fluctuations around the mean, i.e.,

$$u = \bar{u} + u', \quad b = \bar{b} + b', \quad \gamma = \bar{\gamma} + \gamma', \tag{B.10}$$

$$\langle u \rangle = \bar{u},\ \langle b \rangle = \bar{b},\ \langle \gamma \rangle = \bar{\gamma},\ \langle u' \rangle = 0,\ \langle b' \rangle = 0,\ \langle \gamma' \rangle = 0. \tag{B.11}$$

The equations for the mean and fluctuations can be easily obtained from (1) and (B.10)-(B.11) in the form

$$\mathrm{d}\bar{u} = \big[(-\bar{\gamma} + i\omega)\bar{u} - \overline{u'\gamma'} + \bar{b} + f\,\big]\mathrm{d}t, \tag{B.12}$$

$$\mathrm{d}u' = \big[(-\bar{\gamma} + i\omega)u' + b' - \bar{u}\gamma' - u'\gamma' + \overline{u'\gamma'}\,\big]\mathrm{d}t + \sigma_u \mathrm{d}W_u, \tag{B.13}$$

$$\mathrm{d}\bar{b} = (-\gamma_b + i\omega_b)\left(\bar{b} - \hat{b}\right)\mathrm{d}t \qquad , \tag{B.14}$$

$$\mathrm{d}b' = (-\gamma_b + i\omega_b)\, b'\mathrm{d}t + \sigma_b \mathrm{d}W_b, \tag{B.15}$$

$$\mathrm{d}\bar{\gamma} = -d_\gamma\left(\bar{\gamma} - \hat{\gamma}\right)\mathrm{d}t, \tag{B.16}$$

$$\mathrm{d}\gamma' = -d_\gamma \gamma' \mathrm{d}t + \sigma_\gamma \mathrm{d}W_\gamma. \tag{B.17}$$

A closed twelve-dimensional dynamical system for the first and second moments,

$$\bar{u},\ \bar{b},\ \bar{\gamma},\quad \overline{|u'|^2},\ \overline{u'^2},\ \overline{|b'|^2},\ \overline{b'^2}, \overline{\gamma'^2},\quad \overline{u'b'},\ \overline{u'b'^*},\ \overline{u'\gamma'},\ \overline{b'\gamma'}, \tag{B.18}$$

which are needed in the Kalman filter can be found using (B.12)-(B.17) and the multivariate Ito formula (e.g., [12]) and assuming the Gaussian closure (i.e., $E[(X - E(X))^p] = 0$, where p odd).

As an example, consider the evolution of $\overline{|u'|^2}$ which can be derived using the Ito formula and (B.13) as

$$\mathrm{d}\overline{|u'|^2} = \left[-2\,\bar{\gamma}\overline{|u'|^2} + \overline{u'b'^*} + \overline{u'^*b'} - \bar{u}\,\overline{u'^*\gamma'} - \bar{u}^*\,\overline{u'\gamma'} - 2\overline{|u'|^2\gamma'} + \sigma_u^2\right]\mathrm{d}t. \tag{B.19}$$

Following the Gaussian closure approximation, we assume that the third moment, $\overline{|u'|^2\gamma'}$, in (B.19) vanishes. Consequently, the resulting approximation leads to

$$\mathrm{d}\overline{|u'|^2} \approx \left[-2\,\bar{\gamma}\overline{|u'|^2} + \overline{u'b'^*} + \overline{u'^*b'} - \bar{u}\,\overline{u'^*\gamma'} - \bar{u}^*\,\overline{u'\gamma'} + \sigma_u^2\right]\mathrm{d}t. \tag{B.20}$$

Similar procedure leads to dynamical equations for all the variables (B.18).

References

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

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

[3] Bain, A., and D. Crisan, *Fundamentals of Stochastic Filtering*, Springer, New York, 2009.

[4] Bickel, P., B. Li and T. Bengtsson, Sharp failure rates for the bootstrap filter in high dimensions, in *IMS Lecture Notes - Monograph Series: Essays in Honor of J. K. Gosh*, Institute of Mathematical Sciences (2008), 318–329.

[5] Bishop, C., B. Etherton, S. Majumdar, Adaptive sampling with the ensemble transform Kalman filter part I: The theoretical aspects, *Monthly Weather Review*, 129 (2001), 420–436.

[6] Chorin, A., P. Krause, Dimensional reduction for a Bayesian filter, *PNAS*, 101 (2004), 15013–15017.

[7] Crassidis J.L., Markley F.L., Unscented filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 26 (2003), 536–542.

[8] Delsole, T., Stochastic model of quasigeostrophic turbulence, *Surveys in Geophysics*, 25(2) (2004), 107–149.

[9] Evensen, G., Using the Extended Kalman Filter With a Multilayer Quasi-Geostrophic Ocean Model, *J. Geophys. Res.*, 97(C11) (1992), 17,905–17,924.

[10] Evensen, G.,The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynamics*, 53 (2003), 343–367.

[11] Farrell, B., and P. Ioannou, State estimation using a reduced-order Kalman filter, *Journal of the Atmospheric Sciences*, 58 (2001), 3666–3680.

[12] Gardiner, C., *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Springer-Verlag, 1997.

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

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

[15] Gershgorin, B., A.J. Majda, A nonlinear test model for filtering slow-fast systems, *Comm. Math. Sci.*, 6 (2008), 611–649.

[16] Gershgorin, B., A.J. Majda, Filtering a nonlinear slow-fast system with strong fast forcing, *Comm. Math. Sci.*, 8 (2009), 67–92.

[17] Gershgorin, B., A.J. Majda, Filtering a Statistically Exactly Solvable Test Model for Turbulent Tracers from Partial Observations, *J. Comp. Phys*, 230(4) (2011), 1602–1638.

[18] Ghil, M., P. Malanotte-Rizolli, Data assimilation in meteorology and oceanography, *Advances in Geophysics*, 33 (1991), 141–266.

[19] Haven, K., Majda, A.J., Abramov, R., Quantifying predictability through information theory: Small sample estimation in a non-gaussian framework, *J. Comp. Phys.*, 206 (2005), 334–362.

[20] Haykin, S. (Editor), *Kalman Filtering and Neural Networks*, John Wiley & Sons, Inc., 2001

[21] Harlim, J., and A. Majda, Filtering turbulent sparsely observed geophysical flows, *Monthly Weather Review*, 138(4) (2010), 1050–1083.

[22] Hunt, B., E. Kostelich and I. Szunyogh, Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D*, 230 (2007), 112–126.

[23] Jazwinski, A. H., *Stochastic processes and filtering theory*, Academic Press, New York, 1970

[24] Judd, K., T. Stemler, Failures of sequential Bayesian filters and the successes of shadowing filters in tracking of nonlinear deterministic and stochastic systems, *Phys. Rev. E*, 79 (2009), 066206

[25] Kalman, R.E., A new approach to linear filtering and prediction problems. *Journal of Basic Engineering (Series D)*, 82 (1960), 35–45.

[26] Kalman, R. E., Bucy, R.S., New results in linear filtering and prediction problems, *Journal of Basic Engineering (ASME)*, (1961) 83D.

[27] Keating, S.R., Majda A. J., Smith. K. S., New methods for estimating poleward eddy heat transport using satellite altimetry, submitted to *J. Phys. Oceanogr.* (2011).

[28] Lefferts E.J., Markley F.L., Shuster M.D., Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5 (1982), 417–429.

[29] Majda, A.J., R.V. Abramov, M.J. Grote, Information theory and stochastics for multiscale nonlinear systems. CRM Monograph Series, vol. 25, American Mathematical Society, Providence, Rhode Island, USA, 2005.

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

[31] Majda, A.J., J. Harlim, B. Gershgorin, Mathematical strategies for filtering turbulent dynamical systems, DC DS 27(2) (2010), 441–486.

[32] Majda, A.J., I. Timofeyev and E. Vanden-Eijnden, Models for stochastic climate prediction, *Proc. Nat. Acad. Sci.*, 96 (1999), 15687–15691.

[33] Majda, A.J., I. Timofeyev and E. Vanden-Eijnden, Systematic strategies for stochastic mode reduction in climate, *Journal of the Atmospheric Sciences*, 60 (2003), 1705–1722.

[34] Majda, A.J., I. Timofeyev, Low dimensional chaotic dynamics versus intrinsic stochastic chaos: A paradigm model, *Physica D*, 199 (2004), 339–368.

[35] Majda, A.J., C. Franzke and B. Khouider, An applied mathematics perspective on stochastic model ling for climate, *Phil. Trans. A*, 366 (2008), 2429–2455.

[36] Moral, P.D., Nonlinear filtering: Interacting particle solutions, Markov processes and related fields (1996), 2, 555–580.

[37] Moral, P.D., J. Jacod, Interacting particle filtering with discrete observation, in *Sequential Monte-Carlo Methods in Practice* (eds. A. Doucet, N. de Freitas and N. Gordon), vol. Statistics for engineering and information science, Springer, (2001), 43–75.

[38] Pedlosky, J.,Geophysical Fluid Dynamics, Springer, New York, 1979.

[39] Rye, B.J., R.M. Hardesty, Nonlinear Kalman filtering techniques for incoherent backscatter lidar: return power and log power estimation, *Applied optics*, 28(18) (1989), 3908–3917.

[40] Salmon, R., *Lectures on Geophysical Fluid Dynamics*, vol. 378, Oxford University Press, 1998.

[41] Sun X., Jin L., Xiong M., Extended Kalman Filter for Estimation of Parameters in Nonlinear State-Space Models of Biochemical Networks. *PLoS ONE* (2008), 3(11): e3758, doi:10.1371/journal.pone.0003758

[42] Slight, T.J., Ironside, C.N., Stanley, C.R., Hopkinson, M., Farmer, C.D. Integration of a resonant tunneling diode and optical communications laser. *IEEE Photonics Tech. Letters*, 18(14) 2006, 1518–1520.

[43] Snyder, Ch., T. Bengtsson, P. Bickel, J. Anderson, Obstacles to High-dimensional Particle Filtering, *Monthly Weather Review*, 136 (2008), 4629–4640.

[44] Szunyogh, I., Kostelich, E.G., Gyarmati, D. Patil, B. Hunt, E. Kalnay, E. Ott and J. Yorke, Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model, *Tellus A*, 57 (2005), 528–545.

[45] Todling, R., M. Ghil, Tracking atmospheric instabilities with the Kalman filter. Part I: Methodology and one-layer results, *Monthly Weather Review*, 122 (1994), 183–204.

[46] Tribbia, J.J., D.P. Baumhefner, The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies, *Monthly Weather Review* 116 (1988) 2276–2288.

[47] Xiong Kai, Liu Liangdong and Liu Yiwu, Robust extended Kalman filtering for nonlinear systems with multiplicative noises, *Optim. Control Appl. Meth.*, 32 (2011), 47–63

[48] Valverde, G., Terzija, V., Unscented kalman filter for power system dynamic state estimation, *Generation, Transmission & Distribution, IET* , 5(1) (2011), 29–37, doi: 10.1049/iet-gtd.2010.0210

[49] Zerouali, S., A. Allag, S. M. Mimoune, K. Srairi, A. H. Hamida, M. Féliachi, Extended Kalman filter for uninterruptible power supplies applied to non linear loads, *Int. J. Appl. Electromagnetics and Mechanics*, 25 (2007), 565–569.

Regime I. Filtering with perfect model

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.
Observation time: $\Delta t^{obs} = 0.2$
Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$
True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$
Observation error: $\sqrt{r^o} = 0.316$

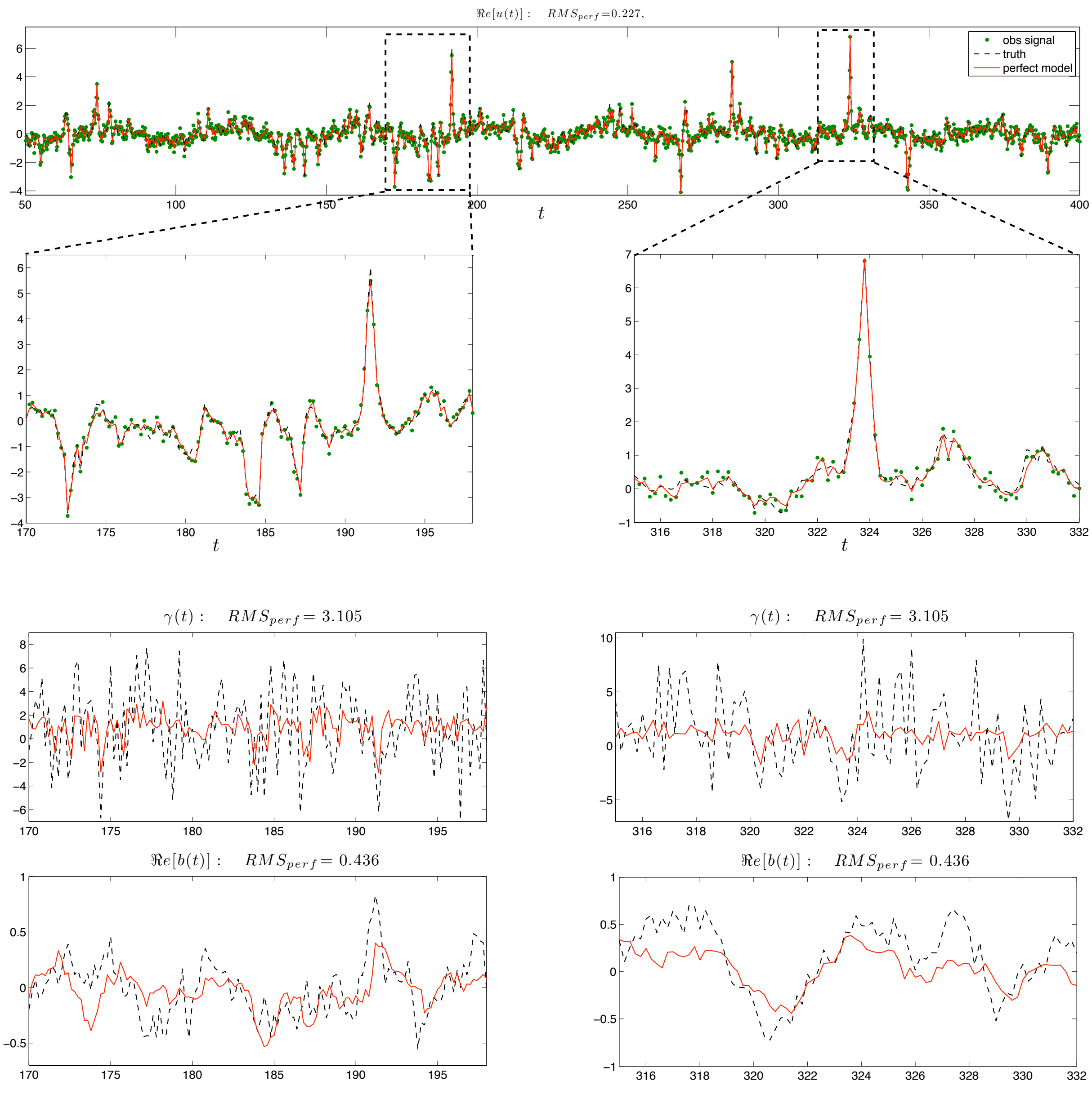

Figure 1: (Regime I) Path-wise example of filtering with perfect model for small observation noise. The perfect filter is given here by SPEKF with correct system parameters (see §5).

Regime I. Filtering with perfect model.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Observation time: $\Delta t^{obs} = 0.2$

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Observation error: $\sqrt{r^o} = 1.41$

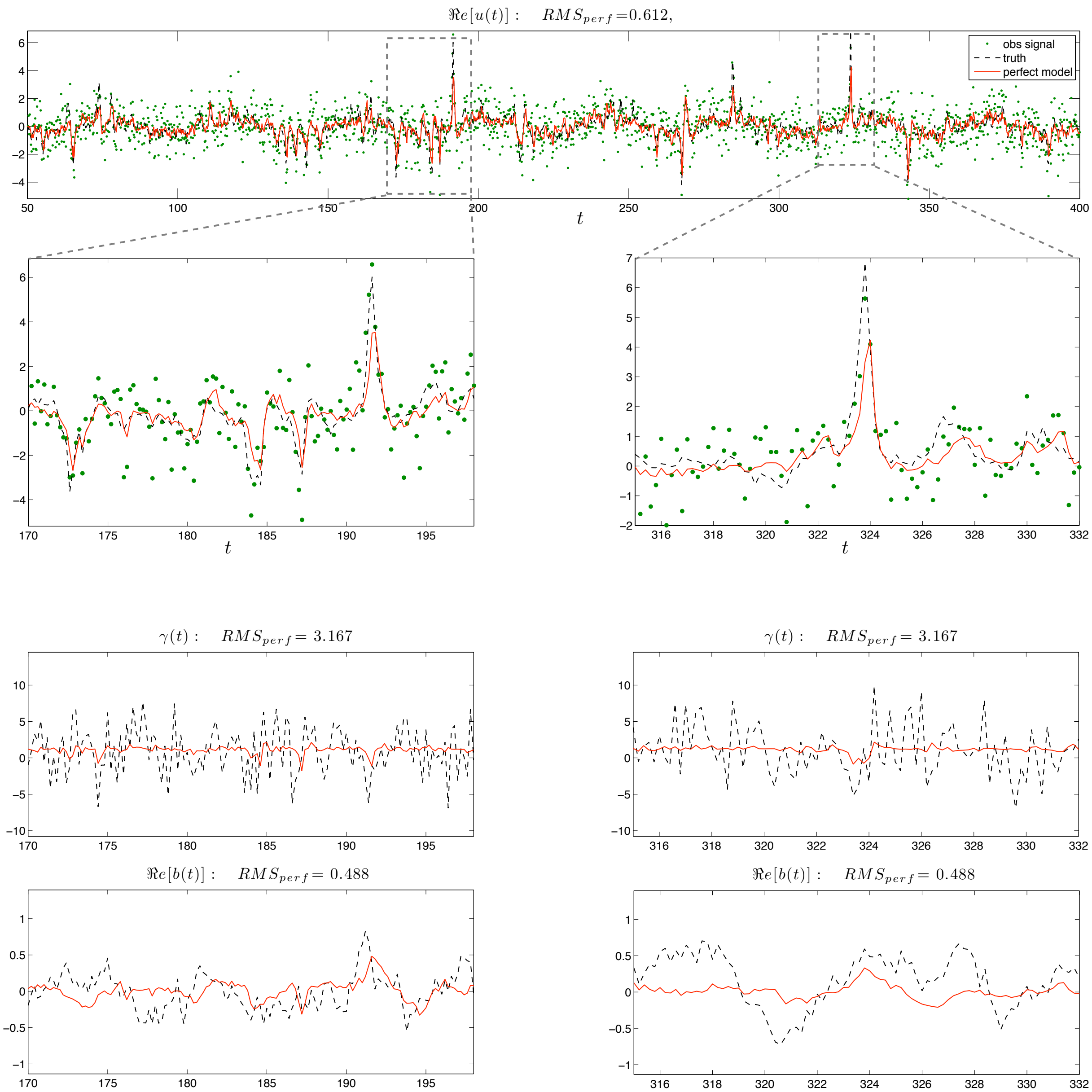

Figure 2: (Regime I) Path-wise example of filtering with perfect model when the signal is dominated by noise but the observation time is much shorter than the decorrelation time of $u(t)$. The perfect filter is given here by SPEKF with correct system parameters (see §5).

Regime I. Filtering with correct parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

Observation time: $\Delta t^{obs} = 0.6$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Observation error: $\sqrt{r^o} = 0.316.$

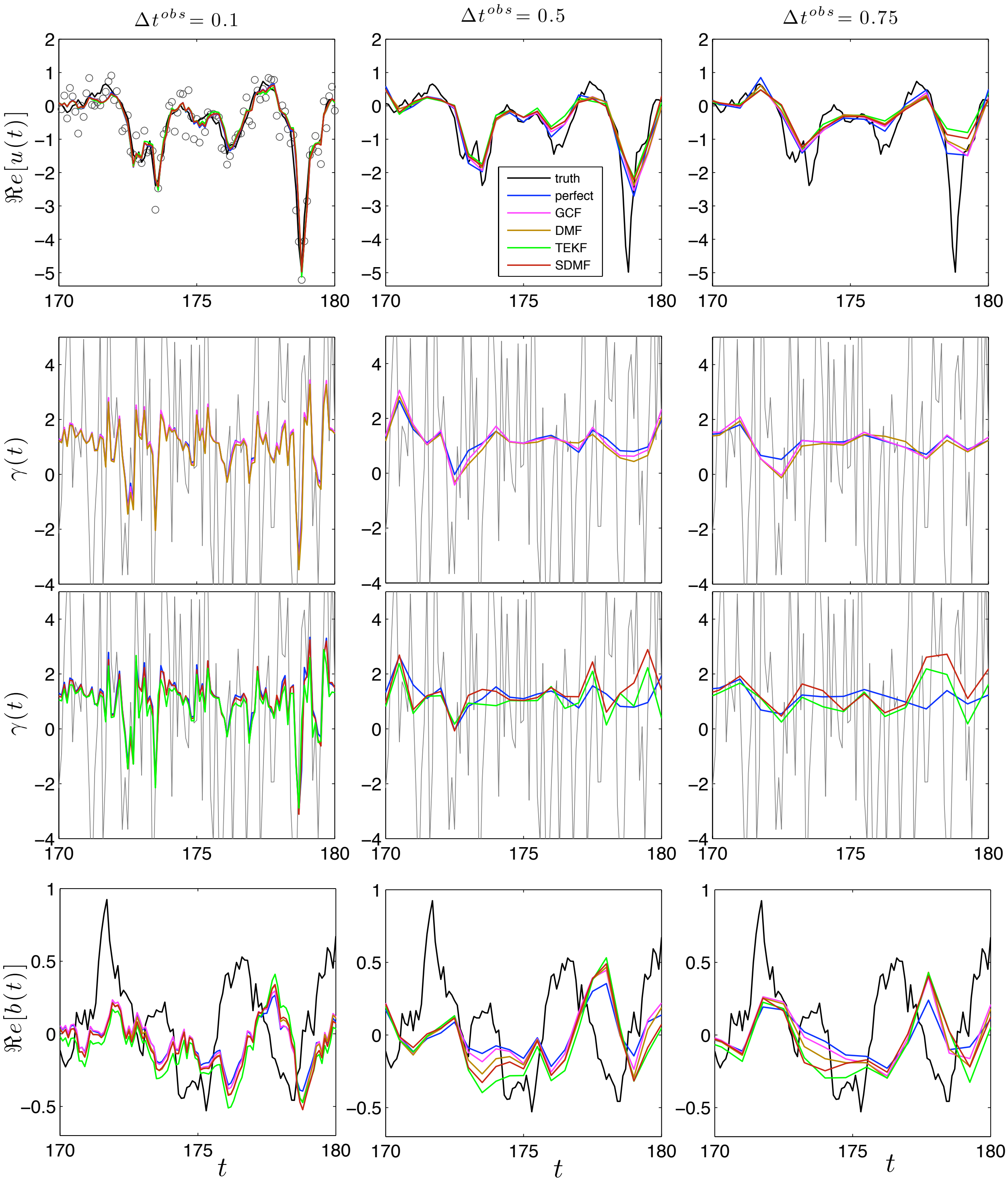

Figure 3: (Regime I) Path-wise example of filtering with correct system parameters for three different values of the observation time Δt^{obs}. The suite of filters used here is described in §3.

Regime I. Filtering with incorrect parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

Observation time: $\Delta t^{obs} = 0.6$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Observation error: $\sqrt{r^o} = 0.316$.

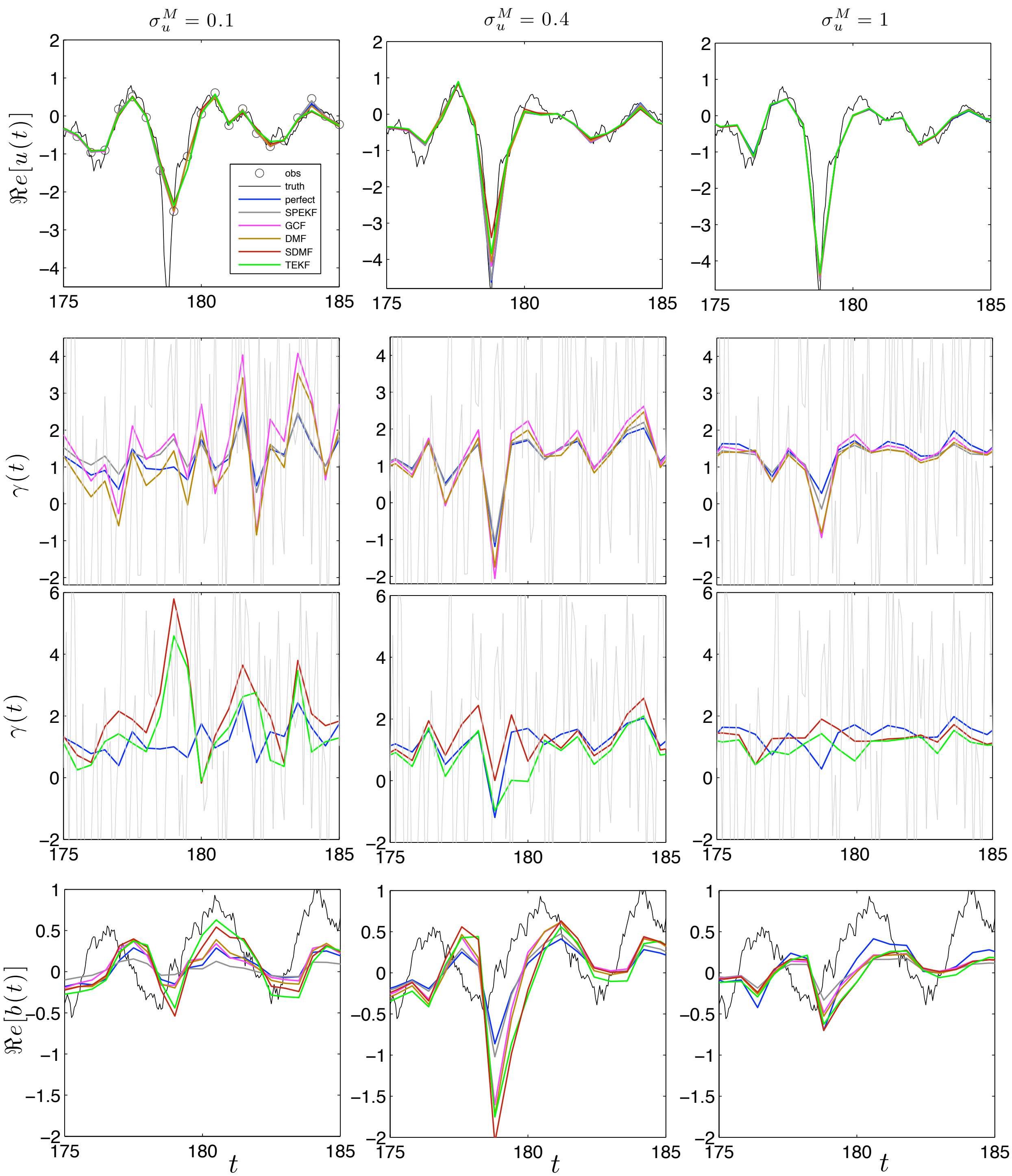

Figure 4: (Regime I) Example of path-wise filtering with incorrect parameter values for three different values of the noise amplitude σ_u^M assumed in the filters for the dynamics of $u(t)$. The suite of filters used here is described in §3.

Regime I. Filtering skill as a function of observation time step.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: (Column 2) $d_\gamma^M = 15,\ \sigma_\gamma^M = 23,$ (Column 3) $d_\gamma^M = 23,\ \sigma_\gamma^M = 21.$

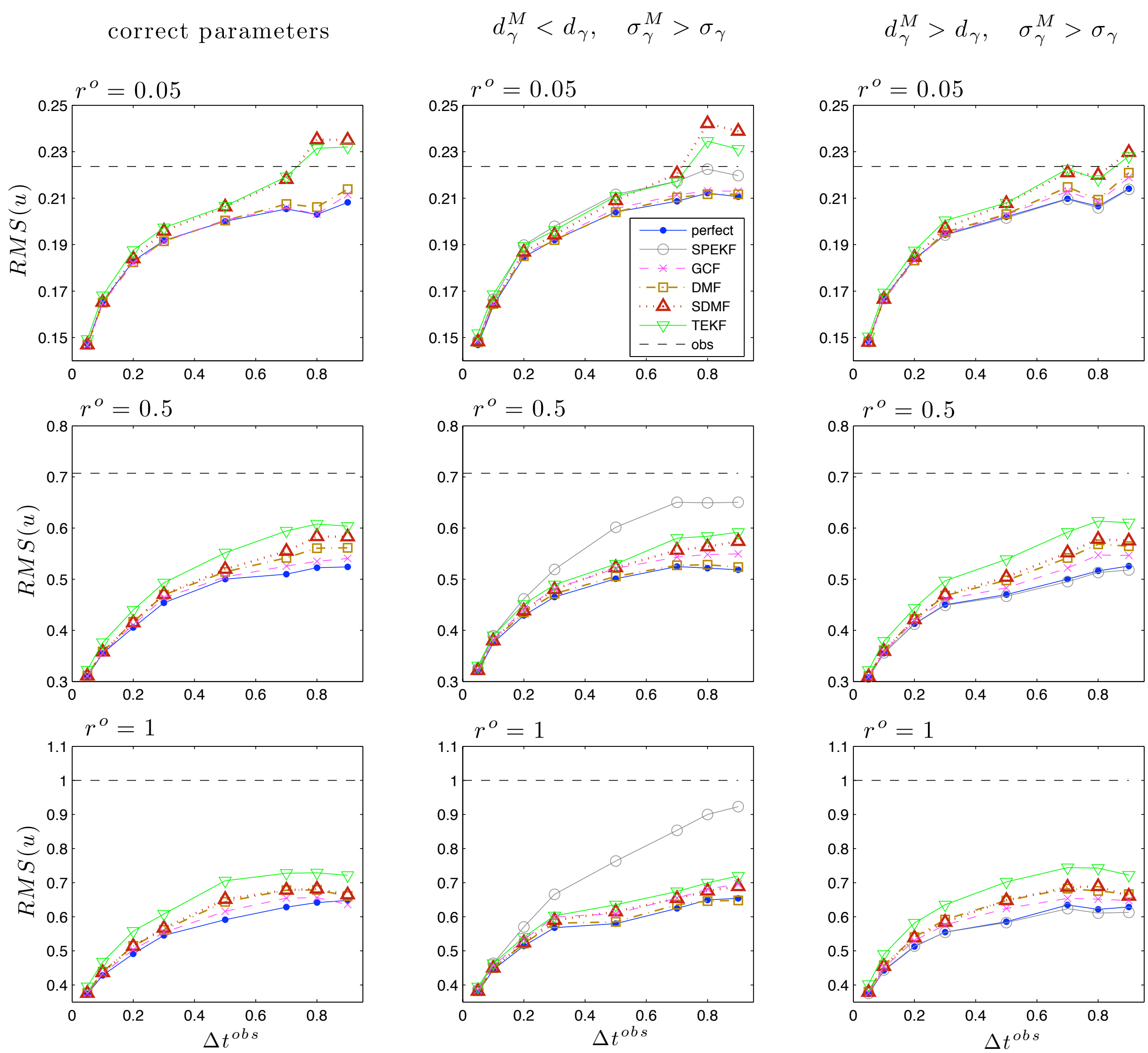

Figure 5: (Regime I) Average RMS errors of the filtered solution $u(t)$ as a function of the observation time Δt^{obs} for fixed values of the observation noise variance r^o and different filter parameters. For $r^o \gtrsim 1$ the signal is dominated by noise. The filters used here are described in §3.

Regime I. Filtering skill as a function of observation noise variance.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: (Column 2) $d_\gamma^M = 15,\ \sigma_\gamma^M = 23$, (Column 3) $d_\gamma^M = 23,\ \sigma_\gamma^M = 21$.

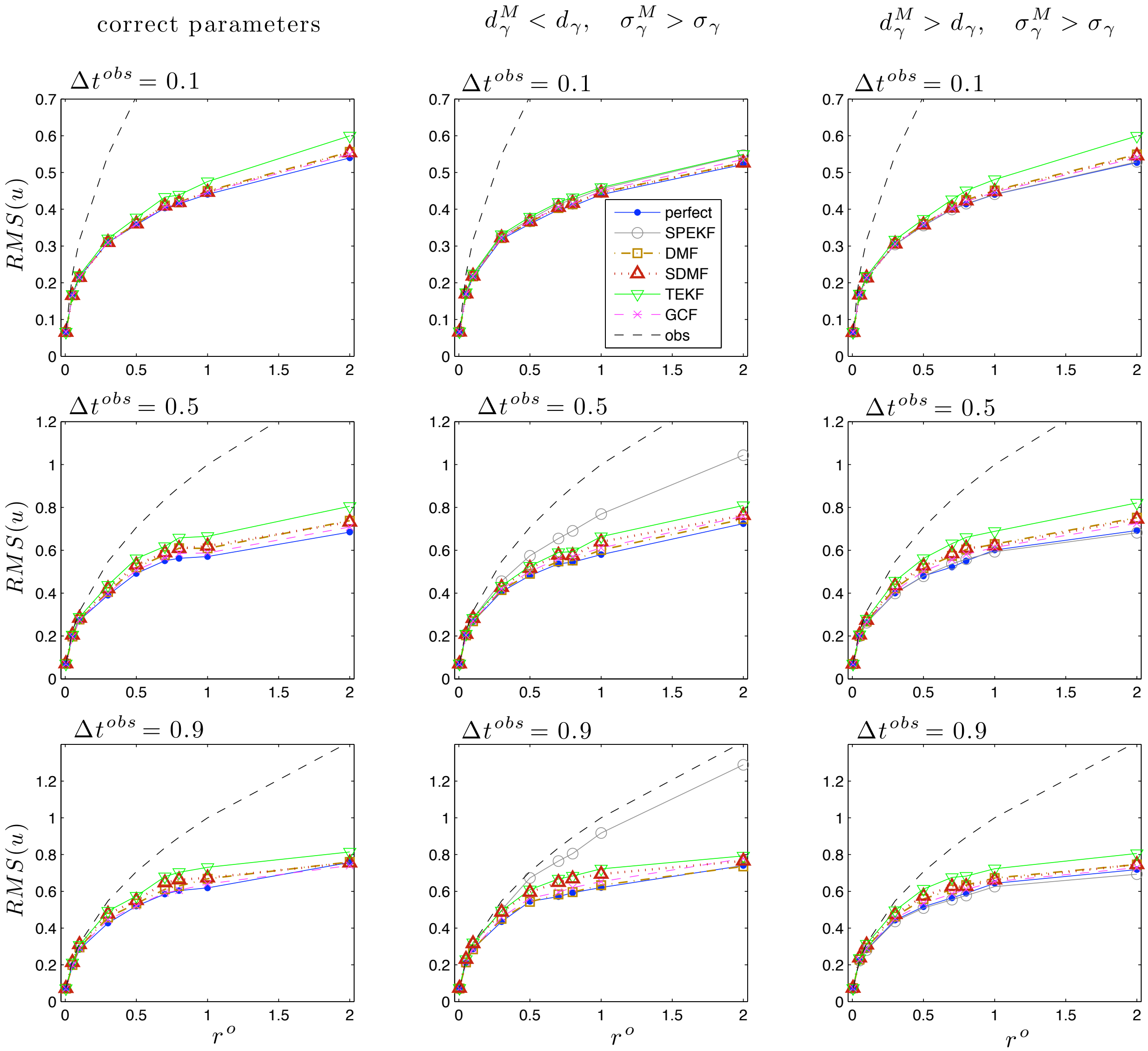

Figure 6: (Regime I) Average RMS errors of the filtered solution $u(t)$ as a function of the observation noise variance r^o for fixed values of the observation time step Δt^{obs} and different filter parameters. For $r^o \gtrsim 1$ the signal is dominated by noise. The filters used here are described in §3.

Regime I. Filtering with incorrect parameter values.

The mean stability parameter:	$\boldsymbol{\chi = -0.7}$.
Decorrelation time of u:	$1/\hat{\gamma} \approx 0.833$
True signal parameters:	$\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$
Incorrect filter parameters:	σ_γ^M varied.

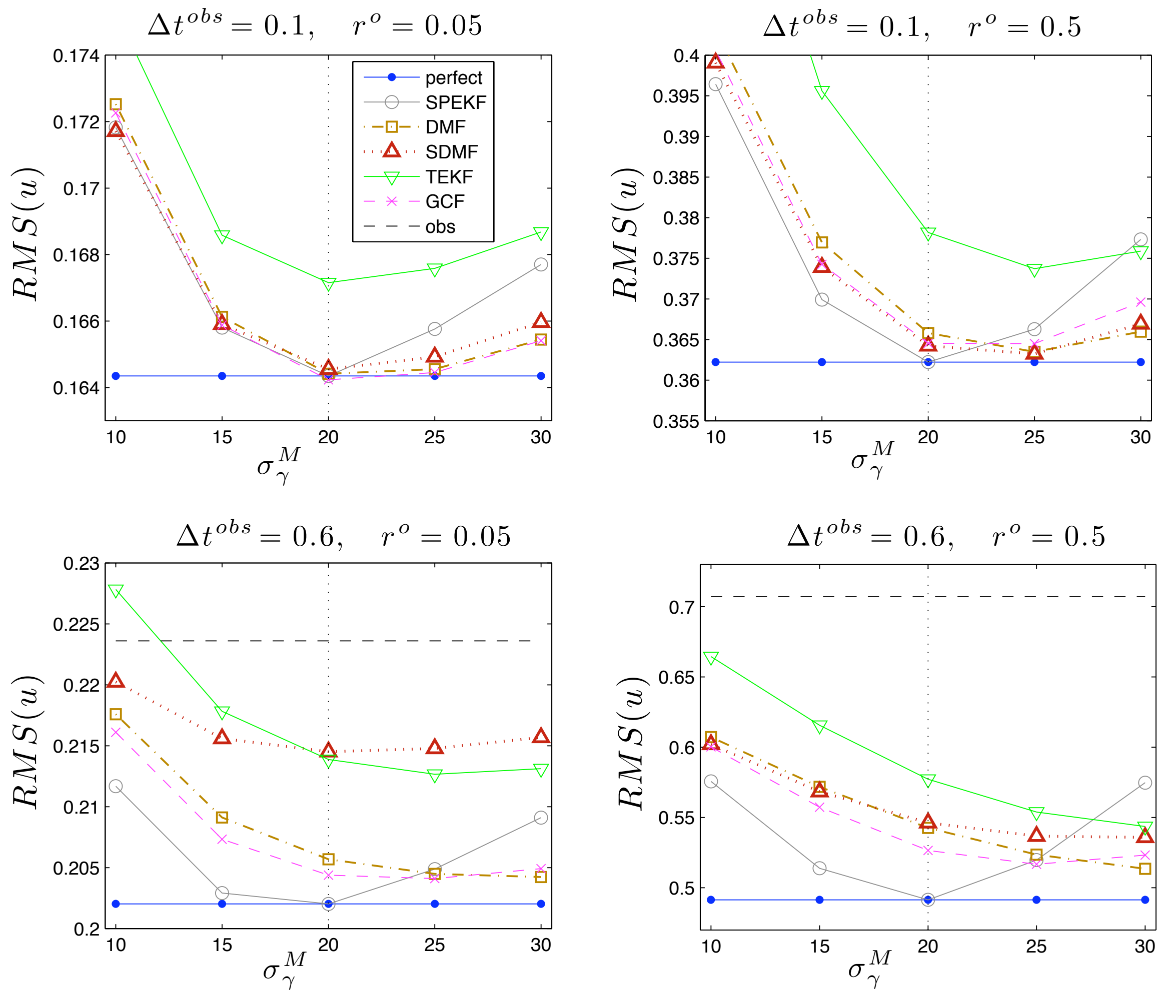

Figure 7: (Regime I) Filtering with imperfect models (see §6). Average RMS errors of the filtered solution $u(t)$ as a function of the filter parameter σ_γ^M (i.e., incorrect noise amplitude assumed in γ) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime I. Filtering with incorrect parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: d_γ^M varied.

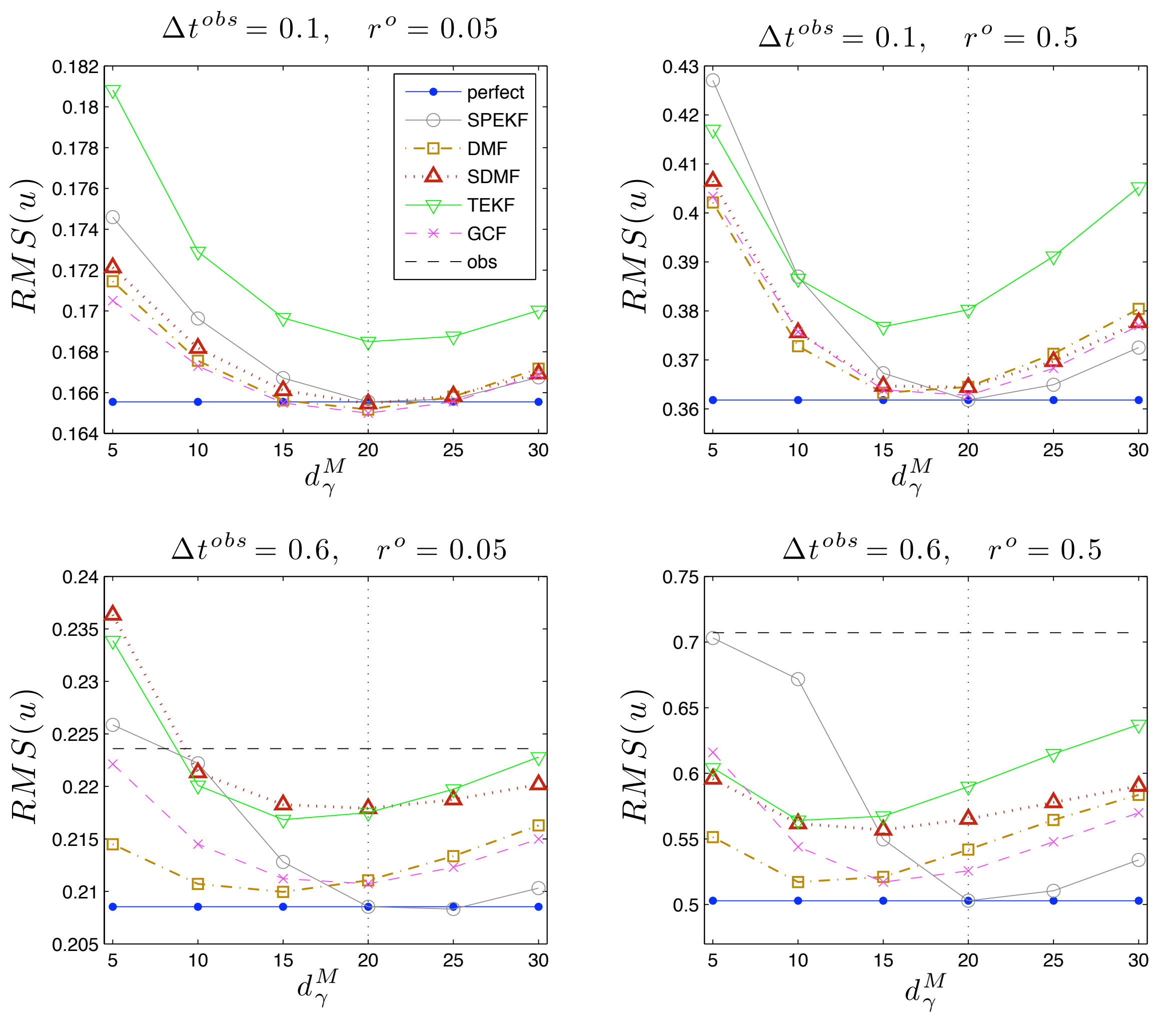

Figure 8: (Regime I) Filtering with imperfect models (see §6). Average RMS errors of the filtered solution $u(t)$ as a function of the filter parameter d_γ^M (i.e., incorrect decorrelation time assumed for γ) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime I. Filtering with incorrect parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.7}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.833$

True signal parameters: $\hat{\gamma} = 1.2, d_\gamma = 20, \sigma_\gamma = 20, \quad \omega_u = 1.78, \sigma_u = 0.5, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: σ_u^M varied.

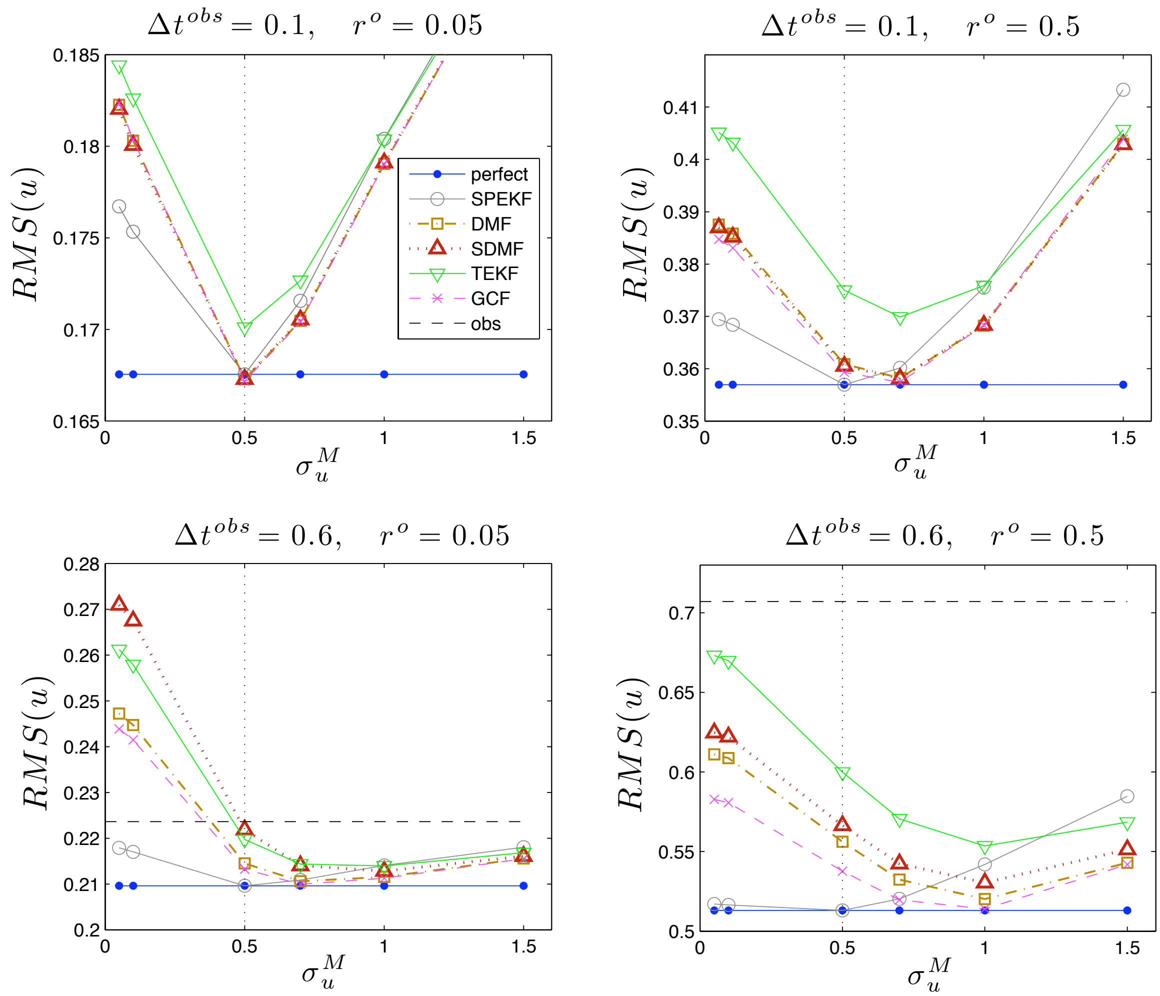

Figure 9: (Regime I) Filtering with imperfect models (see §6). Average RMS errors of the filtered solution $u(t)$ as a function of the filter parameter σ_u^M (i.e., incorrect noise amplitude assumed in u) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime II. Filtering with correct parameters within distinct intervals.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$.
Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$
Observation time: $\Delta t^{obs} = 0.2$
True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$
Observation error: $\sqrt{r^o} = 0.316$.

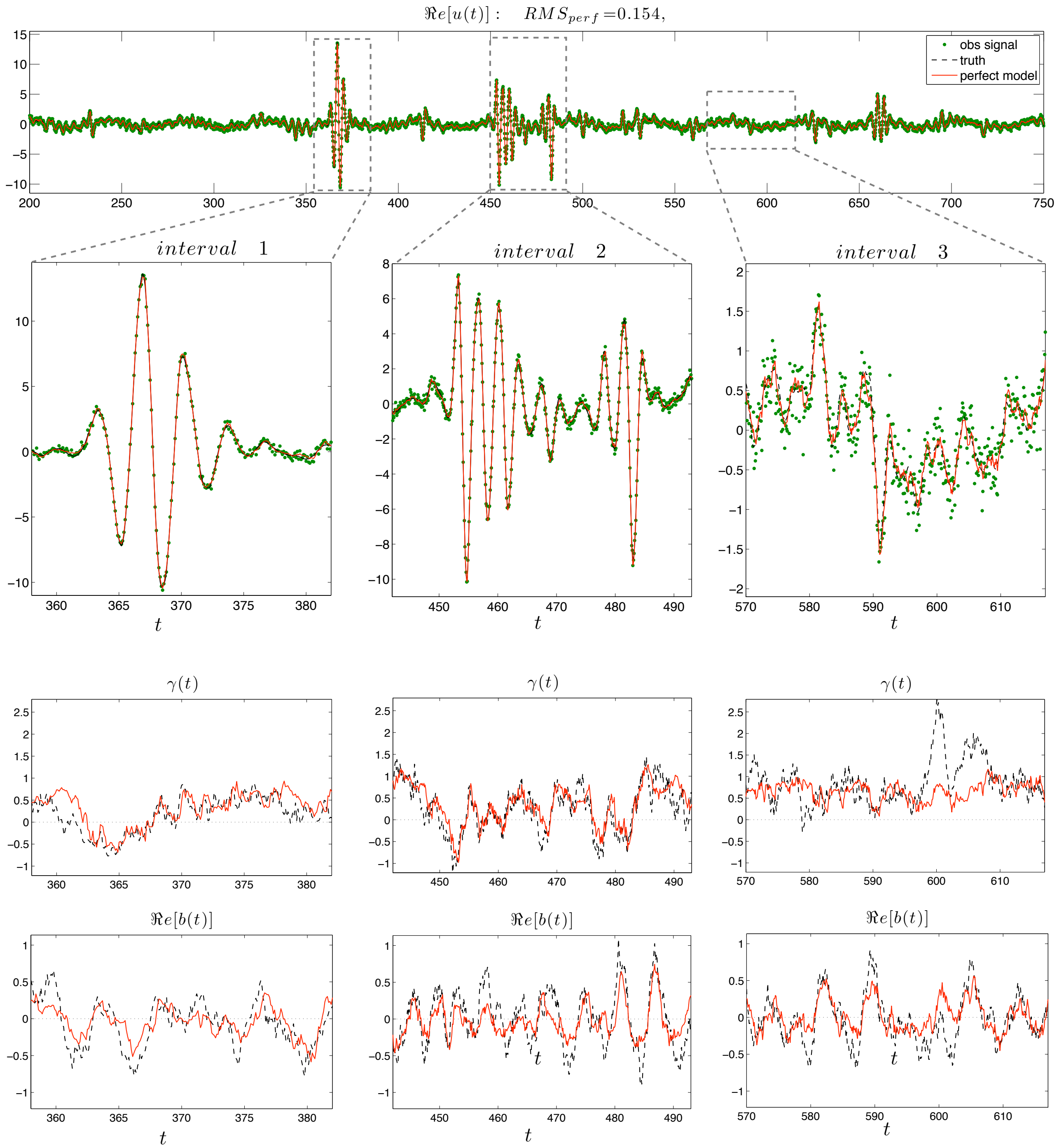

Figure 10: (Regime II) Filtering with the perfect model within three dynamically distinct intervals: ***Interval 1***: large-amplitude transient instability in $u(t)$; **Interval 2**: two subsequent instabilities dominating the low-frequency modulation; ***Interval 3***: quiescent phase with one short unstable episode. The perfect filter is given here by SPEKF with correct parameters (see §5)

Regime II. Filtering with correct parameter values in **INTERVAL 1**.

The mean stability parameter:	$\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u:	$1/\hat{\gamma} \approx 1.81$
Observation time:	Δt^{obs} varied
True signal parameters:	$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$
Observation error:	$\sqrt{r^o} = 0.44.$

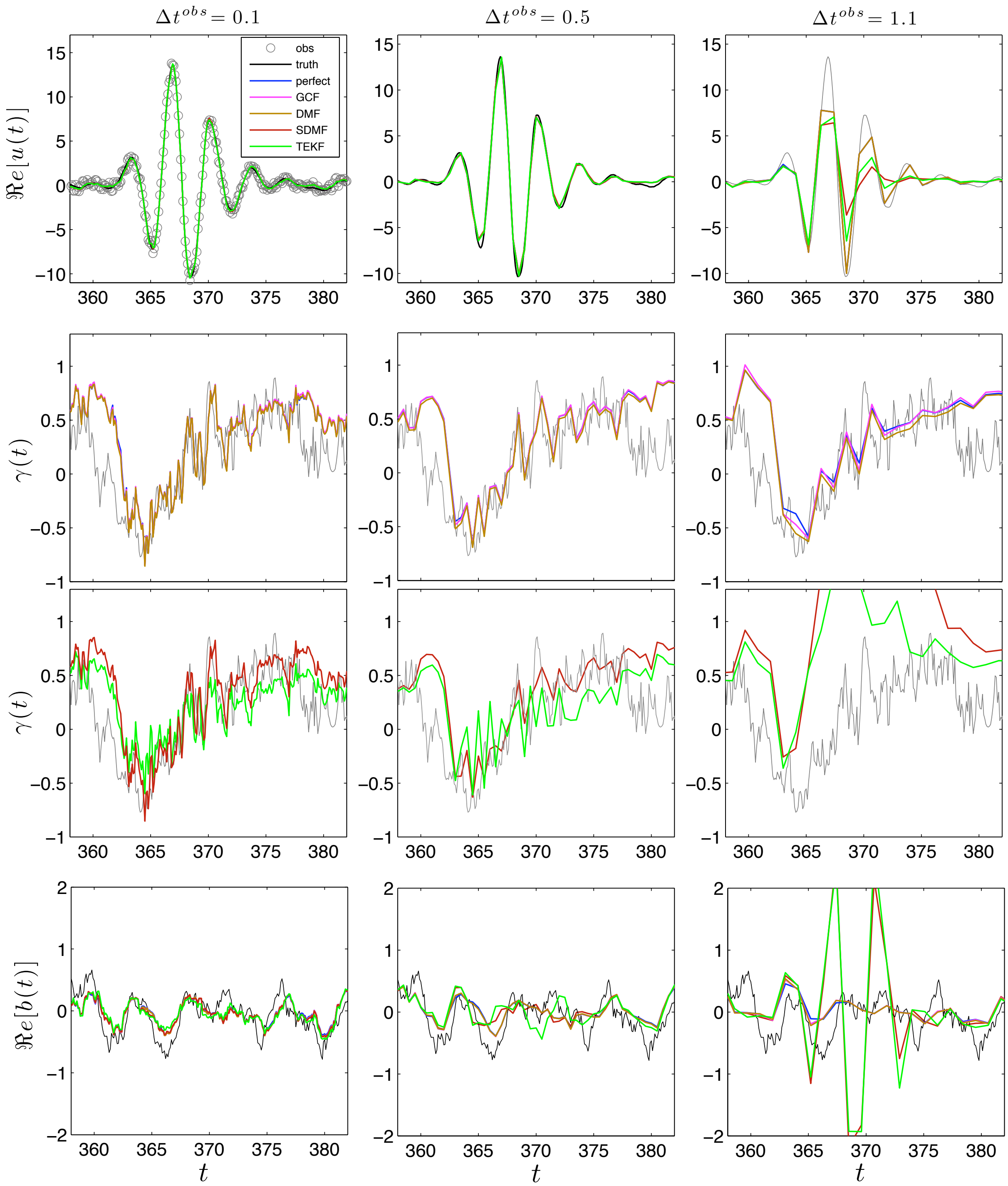

Figure 11: (Regime II). Path-wise example of filtering within interval 1 (see Figure 10) for three different observation times, Δt^{obs}, and a fixed value of the observation noise variance r^o. The filter parameters are here the same as those used for generating the truth signal; the filter suite is described in §3.

Regime II. Filtering with correct parameter values in **INTERVAL 3**.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$
Observation time: Δt^{obs} varied
True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5,\quad \omega_u = 1.78, \sigma_u = 0.1,\quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$
Observation error: $\sqrt{r^o} = 0.44.$

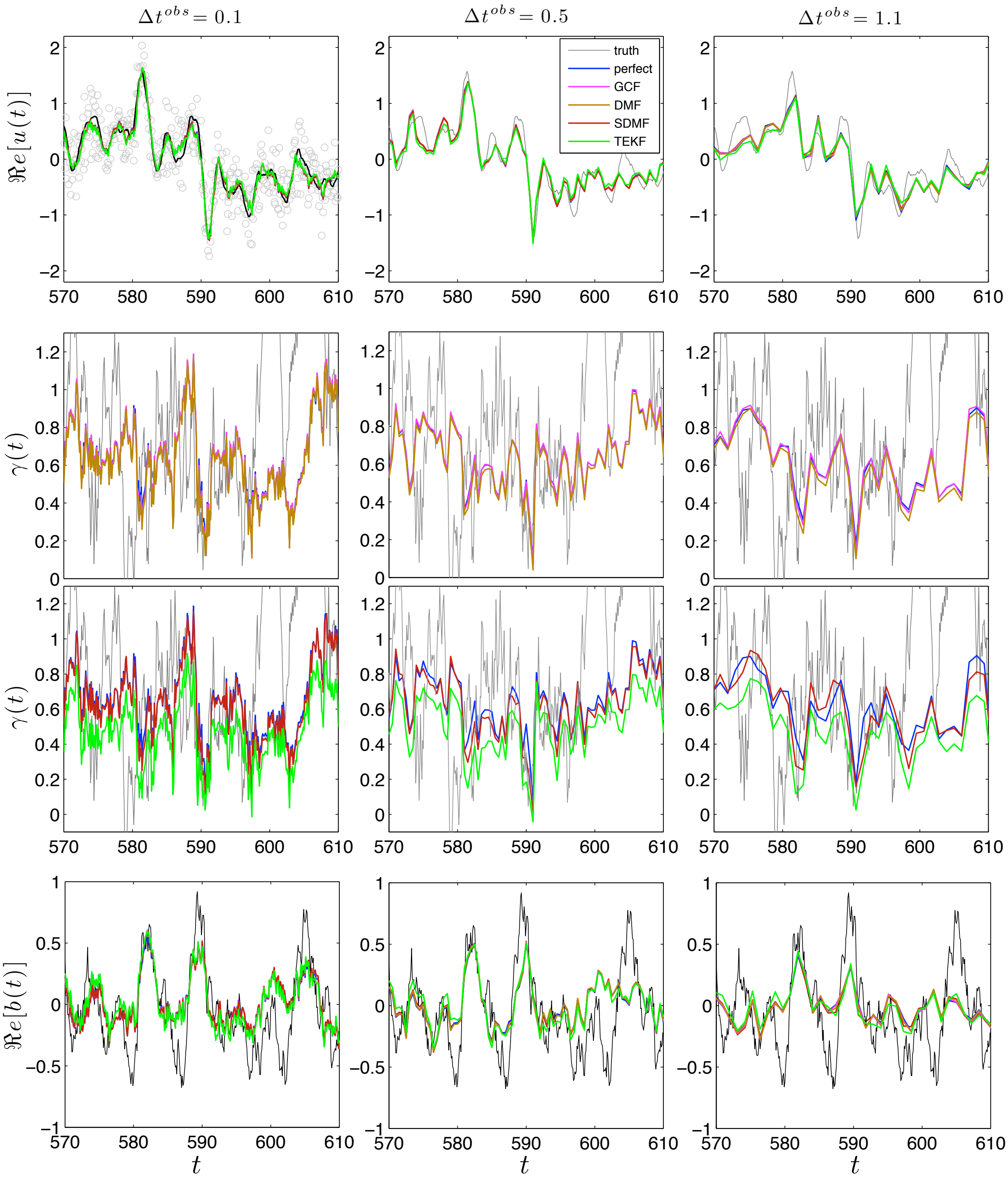

Figure 12: (Regime II). Path-wise example of filtering within interval 2 (see Figure 10) for three different observation times, Δt^{obs}, and a fixed value of the observation noise variance r^o. The filter parameters are here the same as those used for generating the truth signal; the filter suite is described in §3.

Regime I. Filtering skill as a function of the observation noise variance.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$

True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$

Incorrect filter parameters: $\sigma_u^M = 0.2$ and (Column 2) $d_\gamma^M = 0.35,\ \sigma_\gamma^M = 0.55$, (Column 3) $d_\gamma^M = 0.6,\ \sigma_\gamma^M = 0.4,$

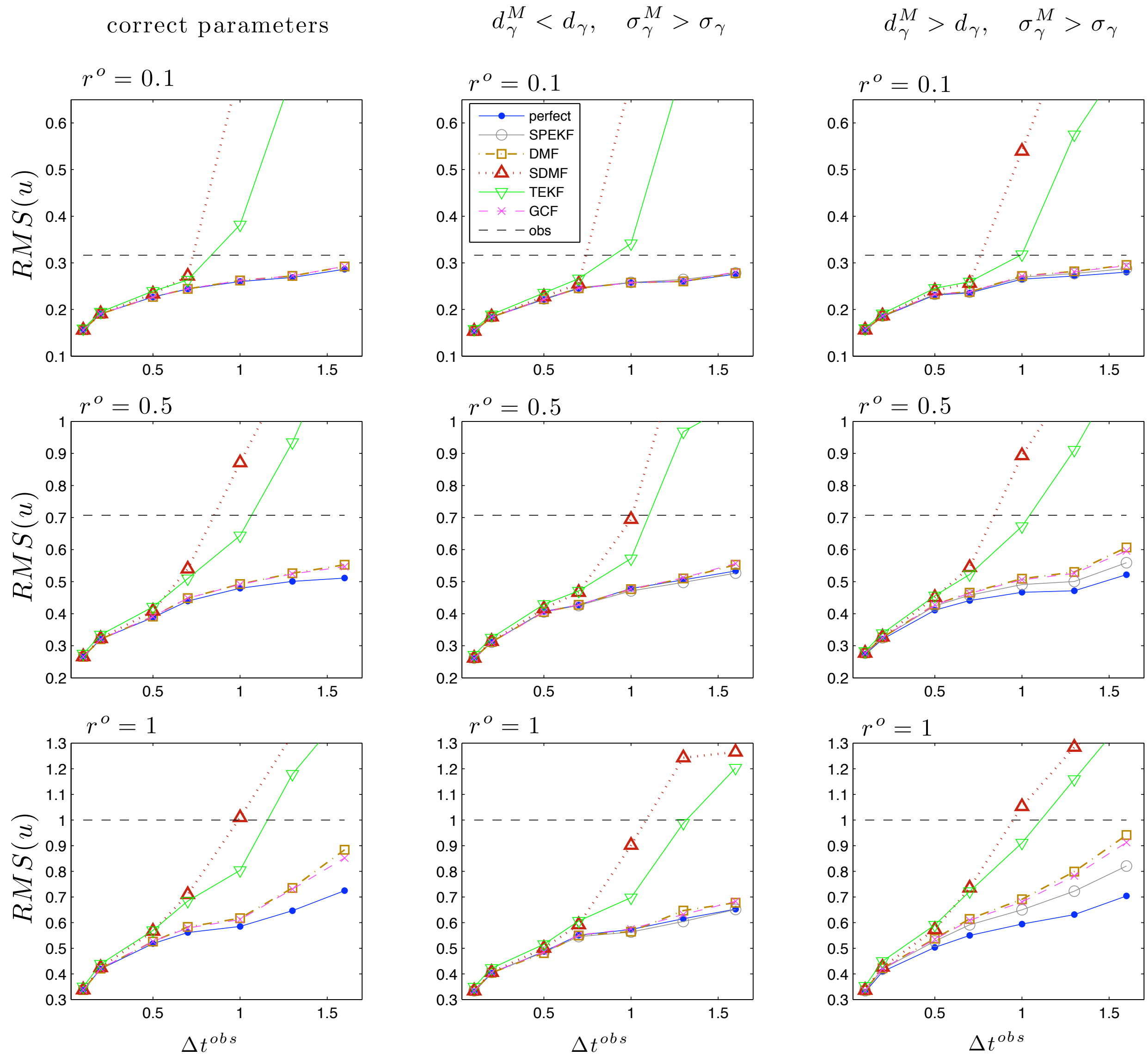

Figure 13: (Regime II) Average RMS errors of the filtered solution $u(t)$ as a function of the observation time step Δt^{obs} for fixed values of the observation noise variance r^o and different filter parameters. For $r^o \gtrsim 1$ the signal is dominated by noise. The filter suite is described in §3.

Regime II. Filtering with correct parameter values within different intervals.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).

Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$

Observation time: Δt^{obs} varied

True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.1.$

Observation error: $\sqrt{r^o}$ varied.

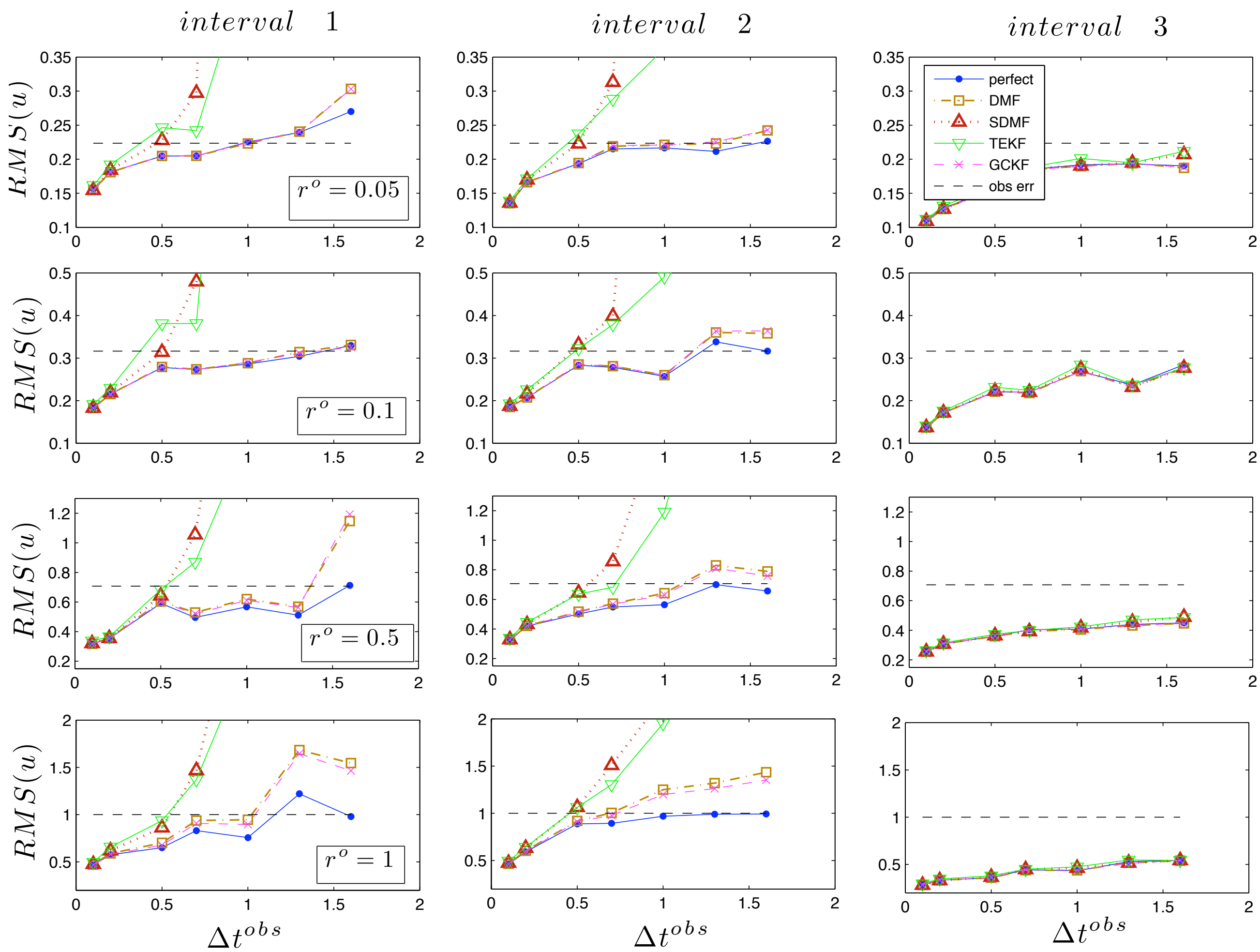

Figure 14: (Regime II) Filtering with correct parameter values. Average RMS errors of $u(t)$ as a function of the observation time Δt^{obs} for fixed true signal and fixed values of the observation noise variance r^o. Columns show the RMS errors in the three intervals shown in Figure 10. The filter suite is described in §3.

Regime I. Filtering skill as a function of observation noise variance.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$

True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$

Incorrect filter parameters: $\sigma_u^M = 0.2$, (Column 2) $d_\gamma^M = 0.35$, $\sigma_\gamma^M = 0.55$, (Column 3) $d_\gamma^M = 0.6$, $\sigma_\gamma^M = 0.4$,

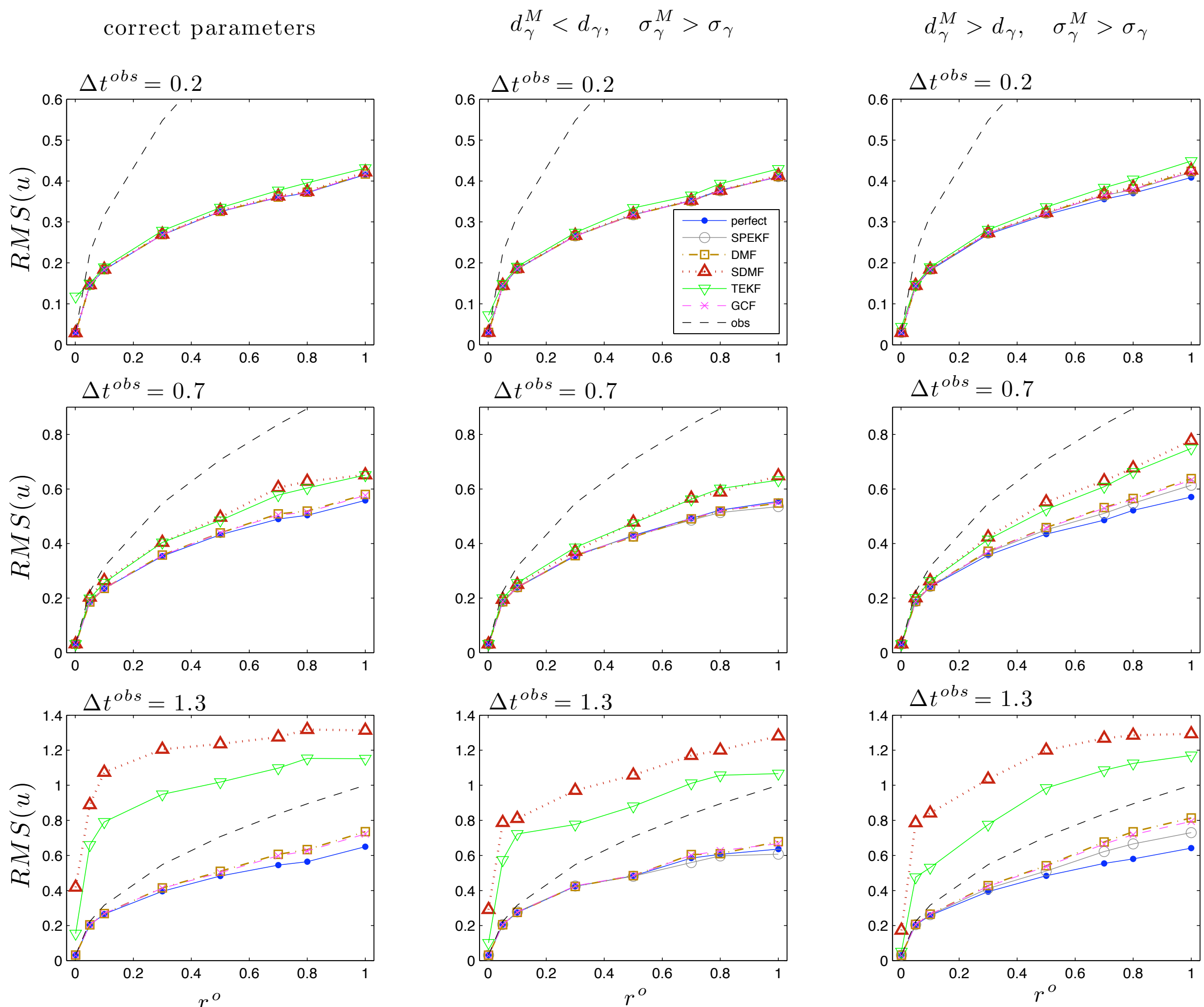

Figure 15: (Regime II) Average RMS errors of the filtered solution $u(t)$ as a function of the observation noise variance r^o for fixed values of the observation time step Δt^{obs} and different filter parameters. For $r^o \gtrsim 1$ the signal is dominated by noise. The filter suite is described in §3.

Regime II. Filtering with imperfect models.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$

True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$

Incorrect filter parameters: σ_γ^M varied.

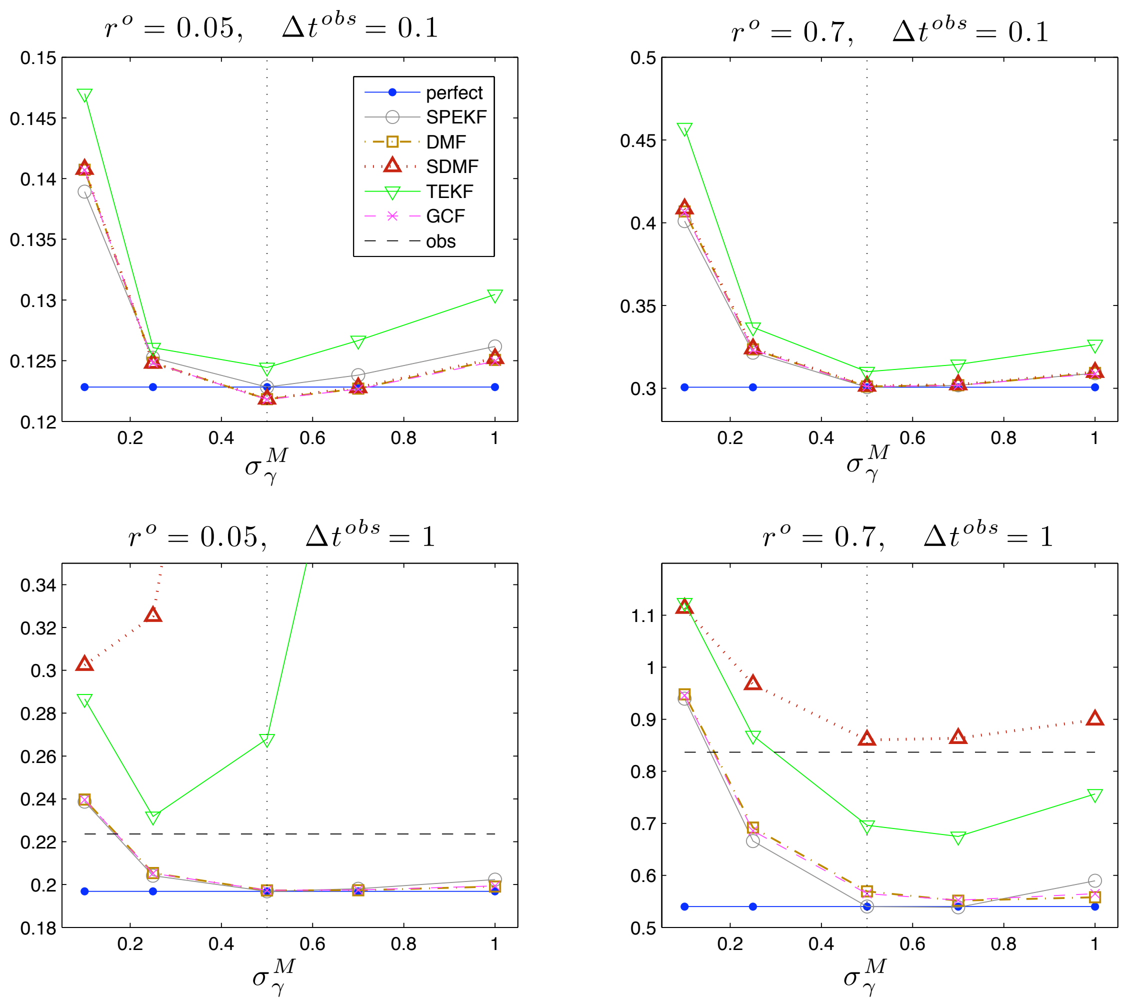

Figure 16: (Regime II) Filtering with imperfect models (cf §6). Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter σ_γ^M (incorrect noise amplitude assumed for $\gamma(t)$) for fixed values of the observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime II. Filtering with imperfect models.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$

True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$

Incorrect filter parameters: d_γ^M varied.

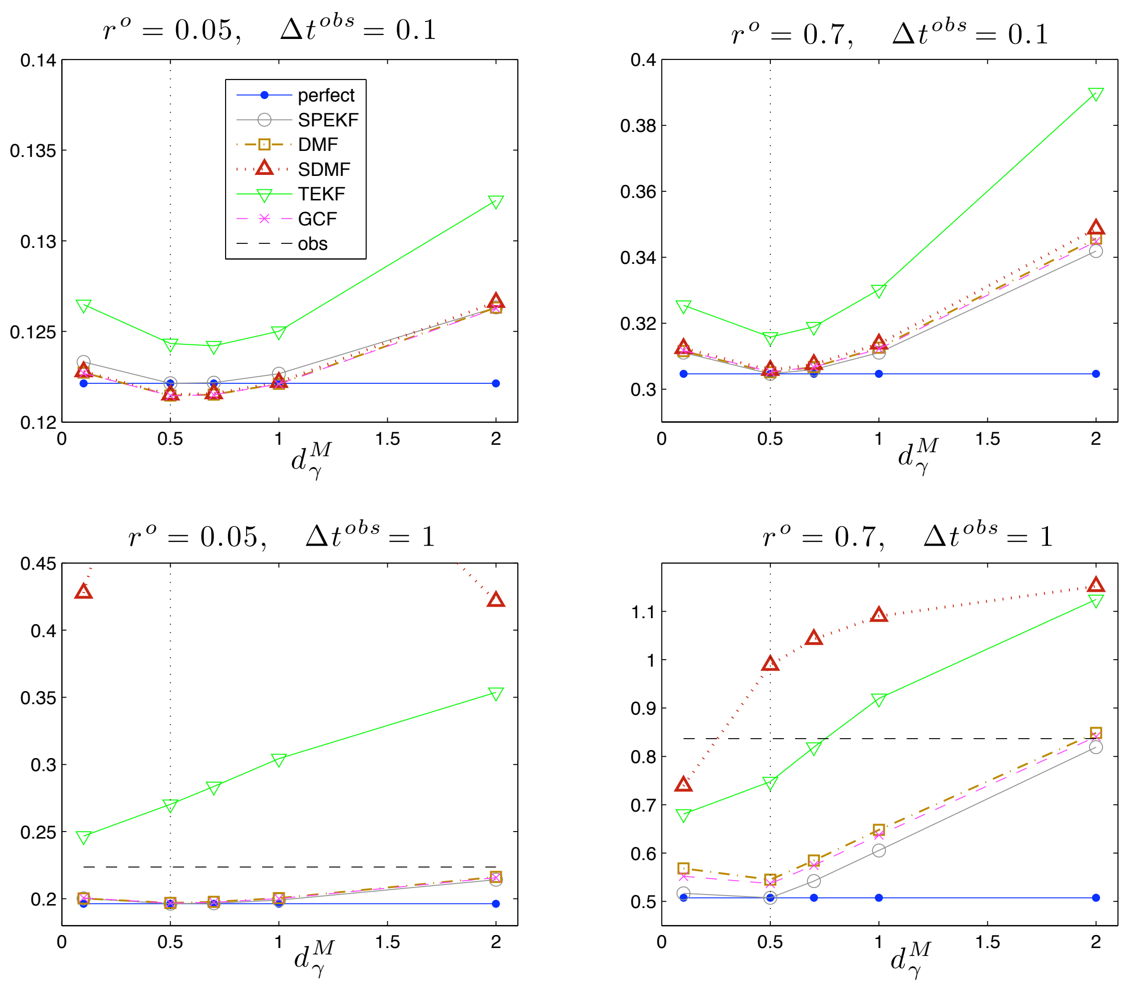

Figure 17: (Regime II) filtering with imperfect models (cf. §6). Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter d_γ^F (incorrect decorrelation time assumed for $\gamma(t)$) for fixed values of the observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime II. Filtering with imperfect models.

The mean stability parameter:	$\boldsymbol{\chi = -0.05}$.
Decorrelation time of u:	$1/\hat{\gamma} \approx 1.81$
True signal parameters:	$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.4.$
Incorrect filter parameters:	σ_u^M varied.

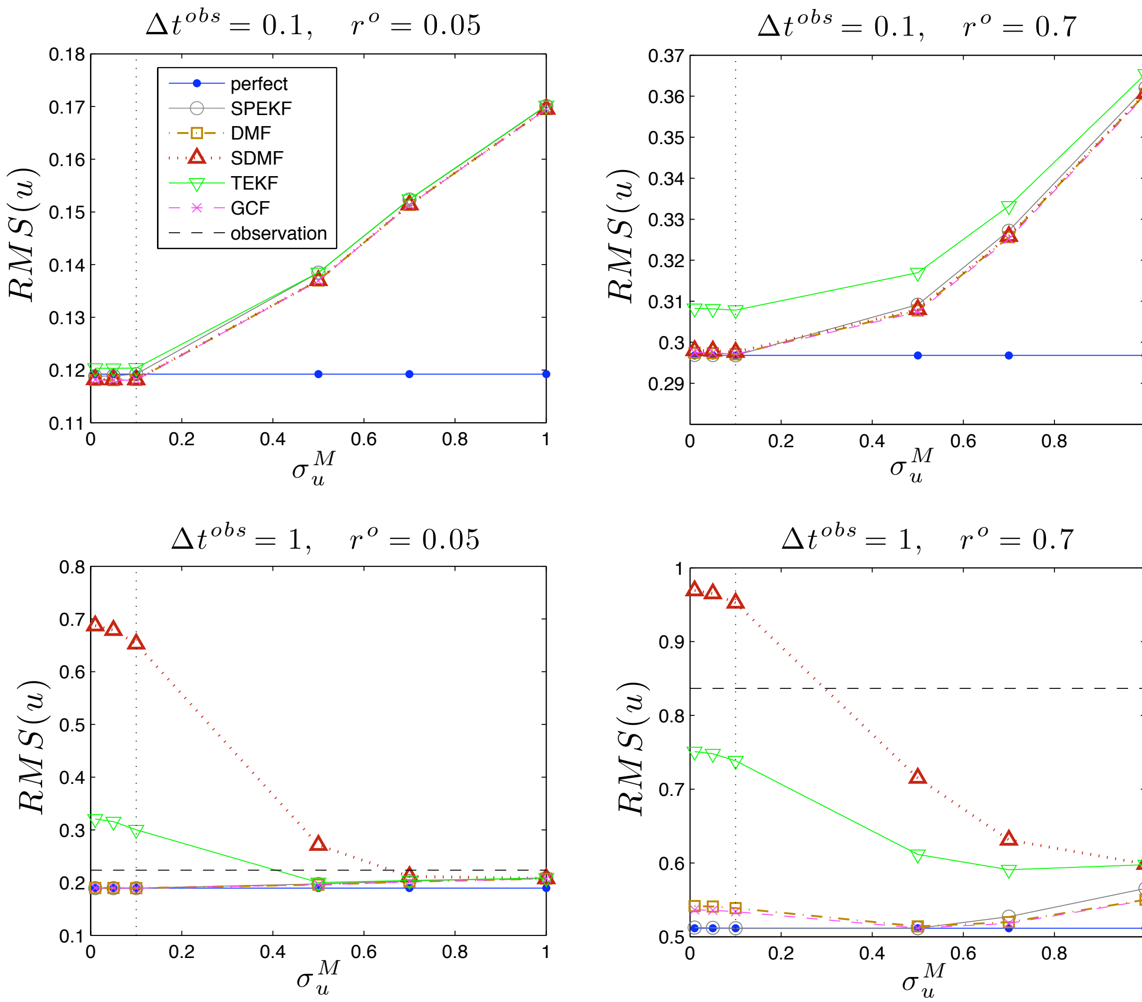

Figure 18: (Regime II) Filtering with imperfect models (cf. §6). Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter σ_u^M (incorrect noise amplitude assumed in $u(t)$) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime II. Parameter estimation.

The mean stability parameter:	$\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u:	$1/\hat{\gamma} \approx 1.81$
Observation time:	Δt^{obs} varied
True signal parameters:	$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.1.$
Filter parameters:	all correct
Observation error:	$\sqrt{r^o}$ varied.

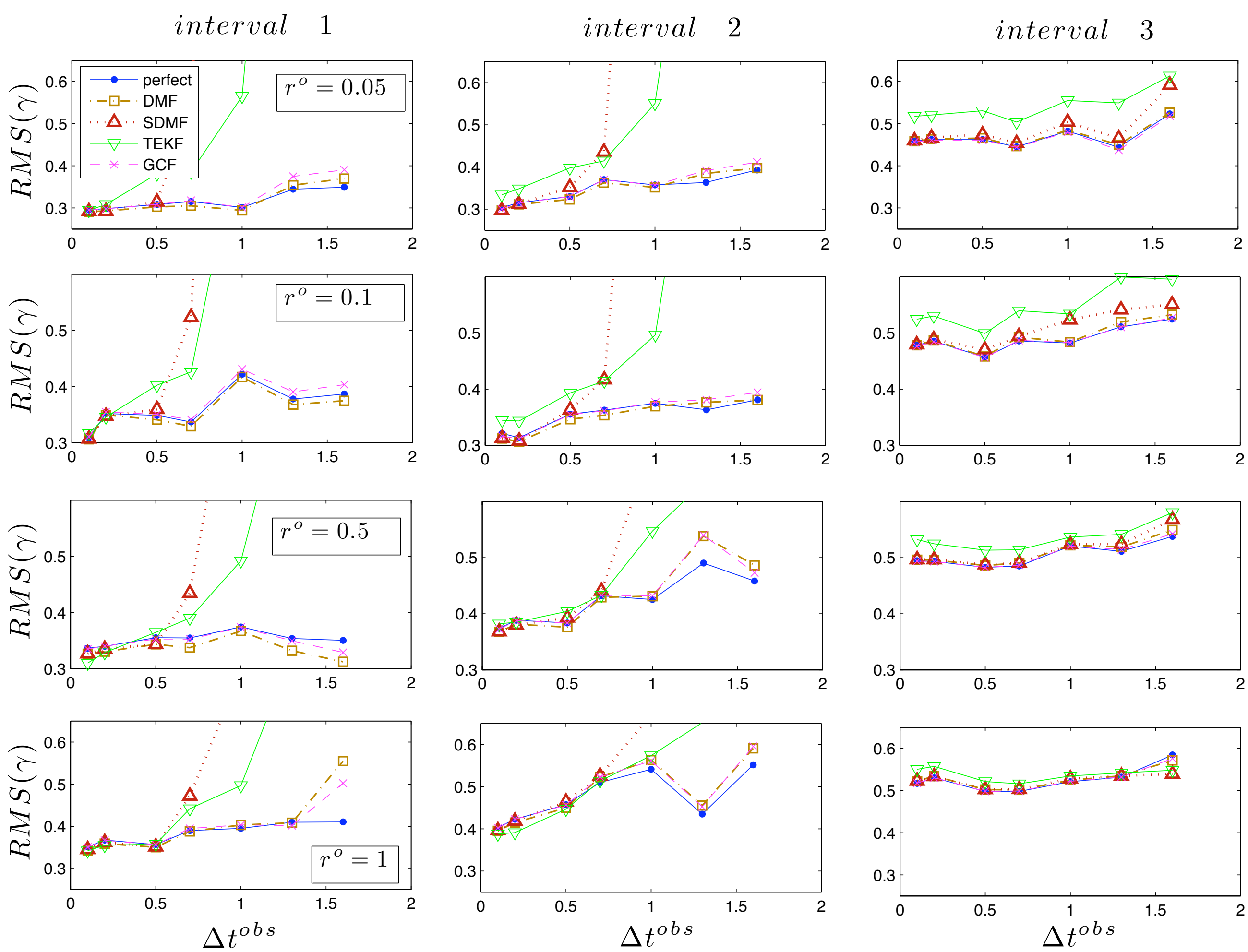

Figure 19: (Regime II) Parameter estimation. Average RMS errors of $\gamma(t)$ as a function of the observation time Δt^{obs} for fixed true signal and fixed values of the observation noise variance r^o. Columns show the RMS errors in the three intervals shown in Figure 10. The filter suite is described in §3.

Regime II. Parameter estimation with imperfect models within distinct intervals.

The mean stability parameter:	$\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u:	$1/\hat{\gamma} \approx 1.81$
Observation time:	Δt^{obs} varied
Observation error:	$\sqrt{r^o}$ varied.
True signal parameters:	$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.1.$
Incorrect filter parameters:	σ_γ^M varied.

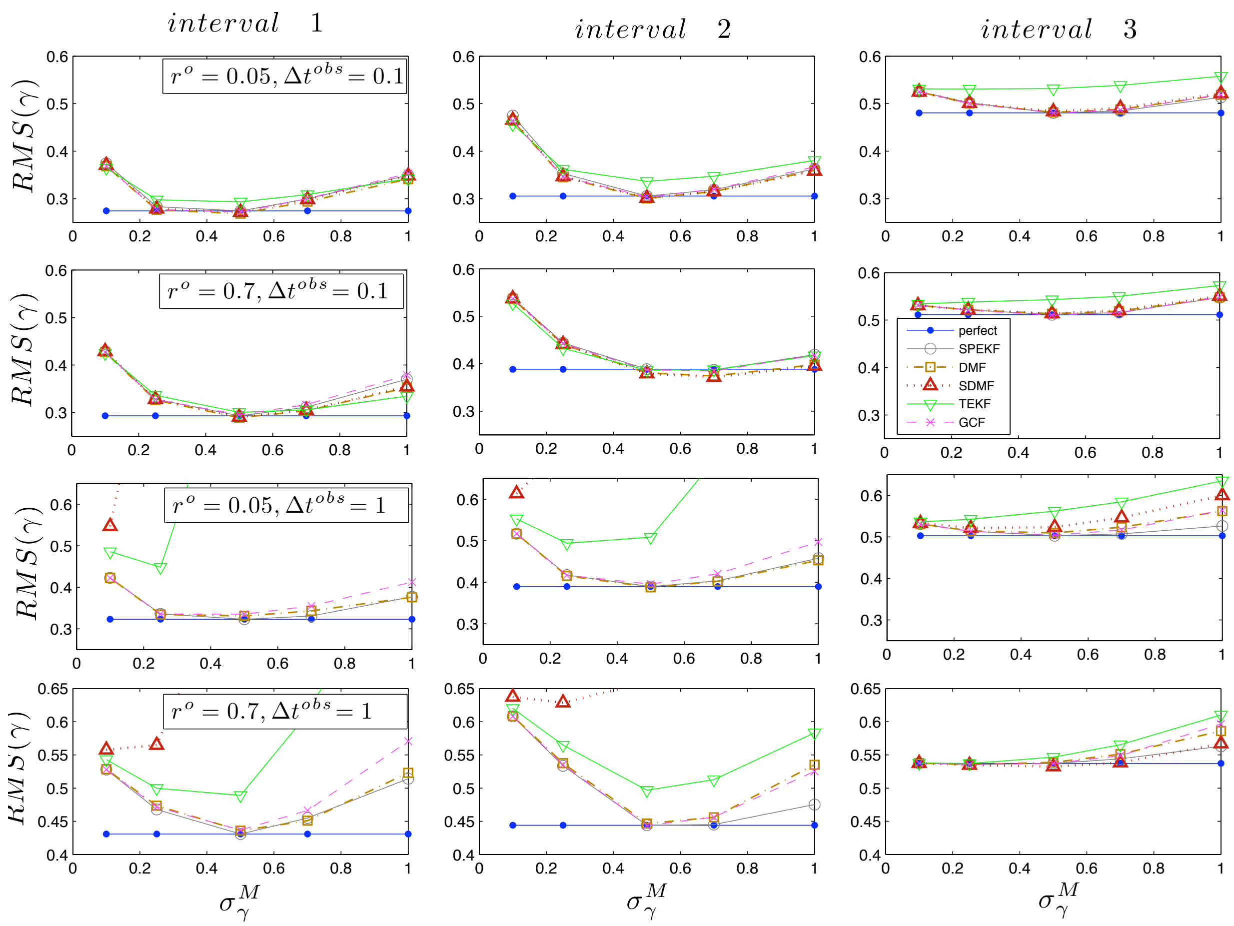

Figure 20: (Regime II) Parameter estimation. Average RMS errors of $\gamma(t)$ as a function of the filter parameter σ_γ^M (incorrect noise amplitude assumed in $\gamma(t)$) for fixed values of the observation noise variance r^o and the observation time Δt^{obs}. The filter suite is described in §3.

Regime II. Filtering with correct parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u: $1/\hat{\gamma} \approx 1.81$
Observation time: Δt^{obs} varied
True signal parameters: $\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5,\quad \omega_u = 1.78, \sigma_u = 0.1,\quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.1.$
Filter parameters: all correct
Observation error: $\sqrt{r^o}$ varied.

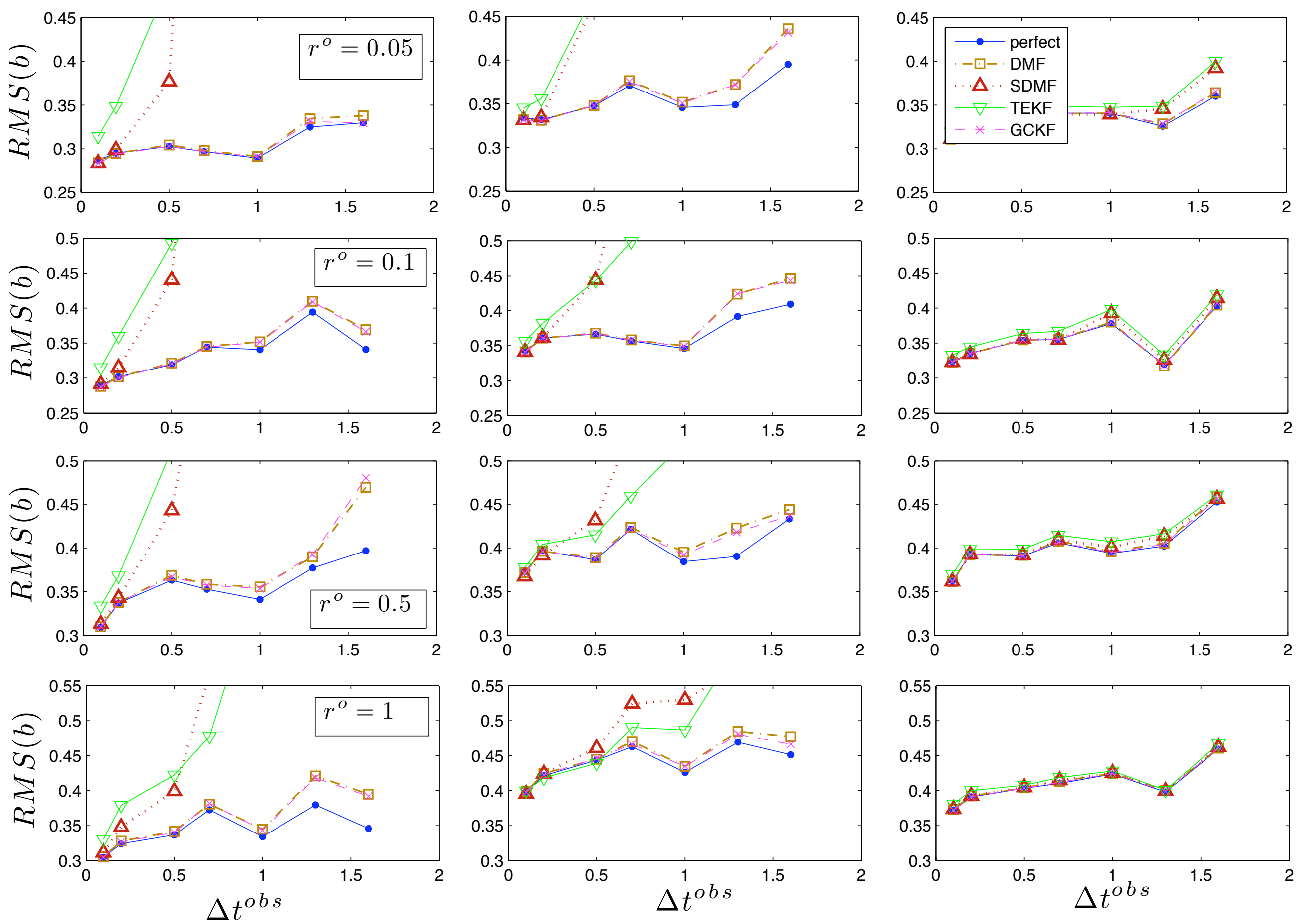

Figure 21: (Regime II) Parameter estimation. Average RMS errors of $b(t)$ as a function of the observation time Δt^{obs} for fixed true signal and fixed values of the observation noise variance r^o. Columns show the RMS errors in the three intervals shown in Figure 10. The filter suite is described in §3.

Regime II. Parameter estimation with imperfect models within distinct intervals.

The mean stability parameter:	$\boldsymbol{\chi = -0.05}$ (weakly damped dynamics of u).
Decorrelation time of u:	$1/\hat{\gamma} \approx 1.81$
Observation time:	Δt^{obs} varied
Observation error:	$\sqrt{r^o}$ varied.
True signal parameters:	$\hat{\gamma} = 0.55, d_\gamma = 0.5, \sigma_\gamma = 0.5, \quad \omega_u = 1.78, \sigma_u = 0.1, \quad \gamma_b = 0.4, \omega_b = 1, \sigma_b = 0.1.$
Incorrect filter parameters:	σ_γ^M varied.

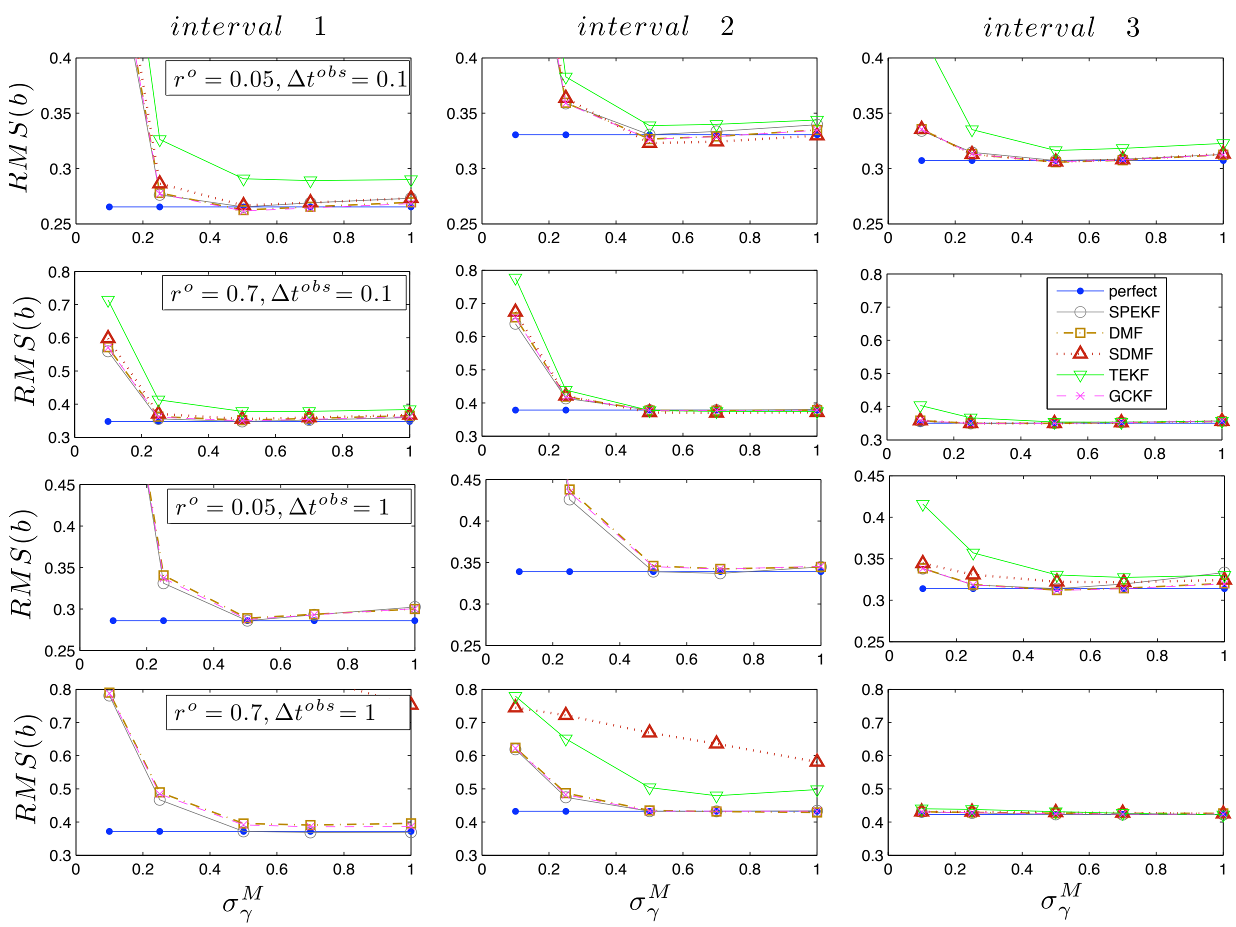

Figure 22: (Regime II) Parameter estimation. Average RMS errors of $\gamma(t)$ as a function of the filter parameter σ_γ^M (incorrect noise amplitude assumed for $\gamma(t)$) for fixed values of observation noise variance r^o and observation time Δt^{obs}. The filter suite is described in §3.

Regime III. Filtering and parameter estimation.

The mean stability parameter: $\boldsymbol{\chi = -0.1}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.12$

True signal parameters: $\hat{\gamma} = 8.1, d_\gamma = 0.25, \sigma_\gamma = 1, \quad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: $d_\gamma^M = 0.3$.

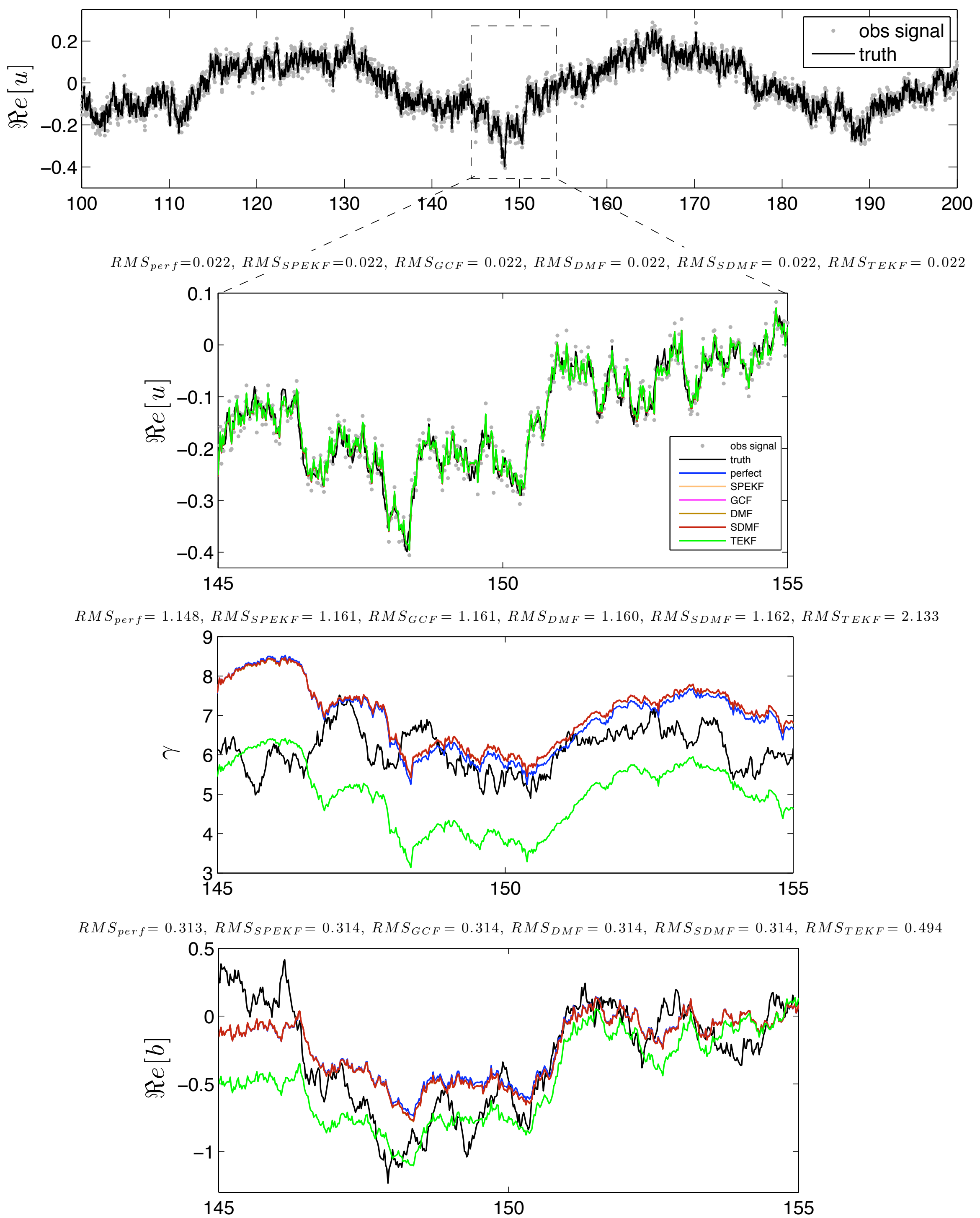

Figure 23: (Regime III) Path-wise filtering with incorrect parameters (cf. §6) when the decorrelation time of $\gamma(t)$ assumed in the filters is underestimated (i.e., $d_\gamma^M > d_\gamma$). The filter suite is described in §3.

Regime III. Filtering with correct parameter values.

The mean stability parameter: $\boldsymbol{\chi = -0.1}$.

Observation time: Δt^{obs} varied

Decorrelation time of u: $1/\hat{\gamma} \approx 0.12$

True signal parameters: $\hat{\gamma} = 8.1, d_\gamma = 0.25, \sigma_\gamma = 1, \quad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 0.3, \sigma_b = 0.5.$

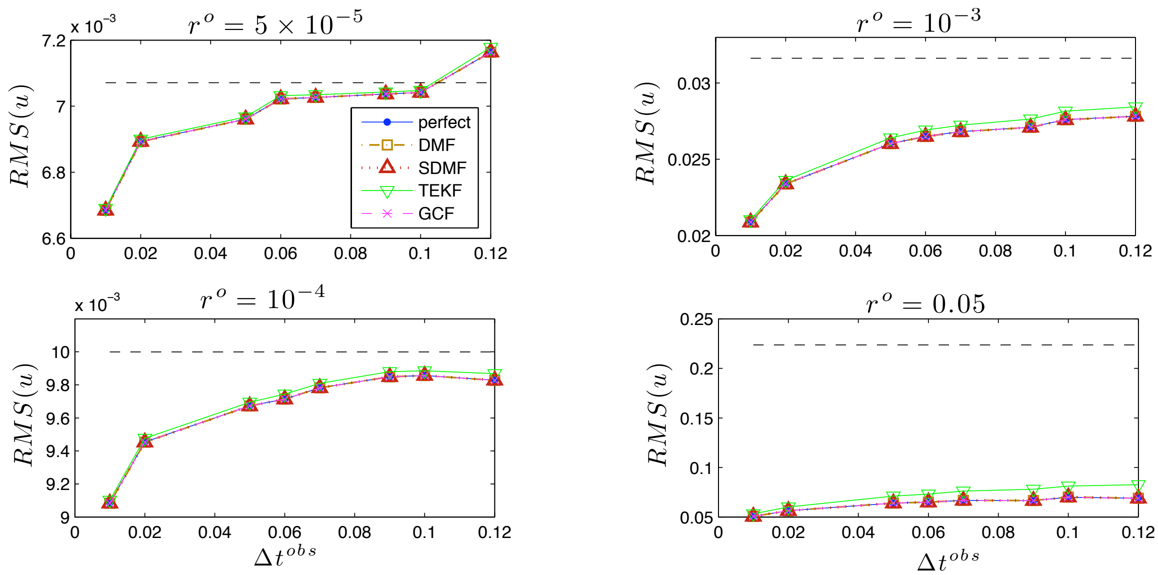

Figure 24: (Regime III) Filtering with correct parameter values. Average RMS errors of the filtered signal $u(t)$ as a function of the observation time Δt^{obs} for fixed values of the observation noise variance r^o. For $r^o \gtrsim 0.03$ the signal is dominated by observation noise. The filter suite is described in §3.

Regime III. Parameter estimation with imperfect models.

The mean stability parameter: $\boldsymbol{\chi = -0.1}$.

Decorrelation time of u: $1/\hat{\gamma} \approx 0.12$

True signal parameters: $\hat{\gamma} = 8.1, d_\gamma = 0.25, \sigma_\gamma = 1, \quad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$

Incorrect filter parameters: σ_γ^M varied.

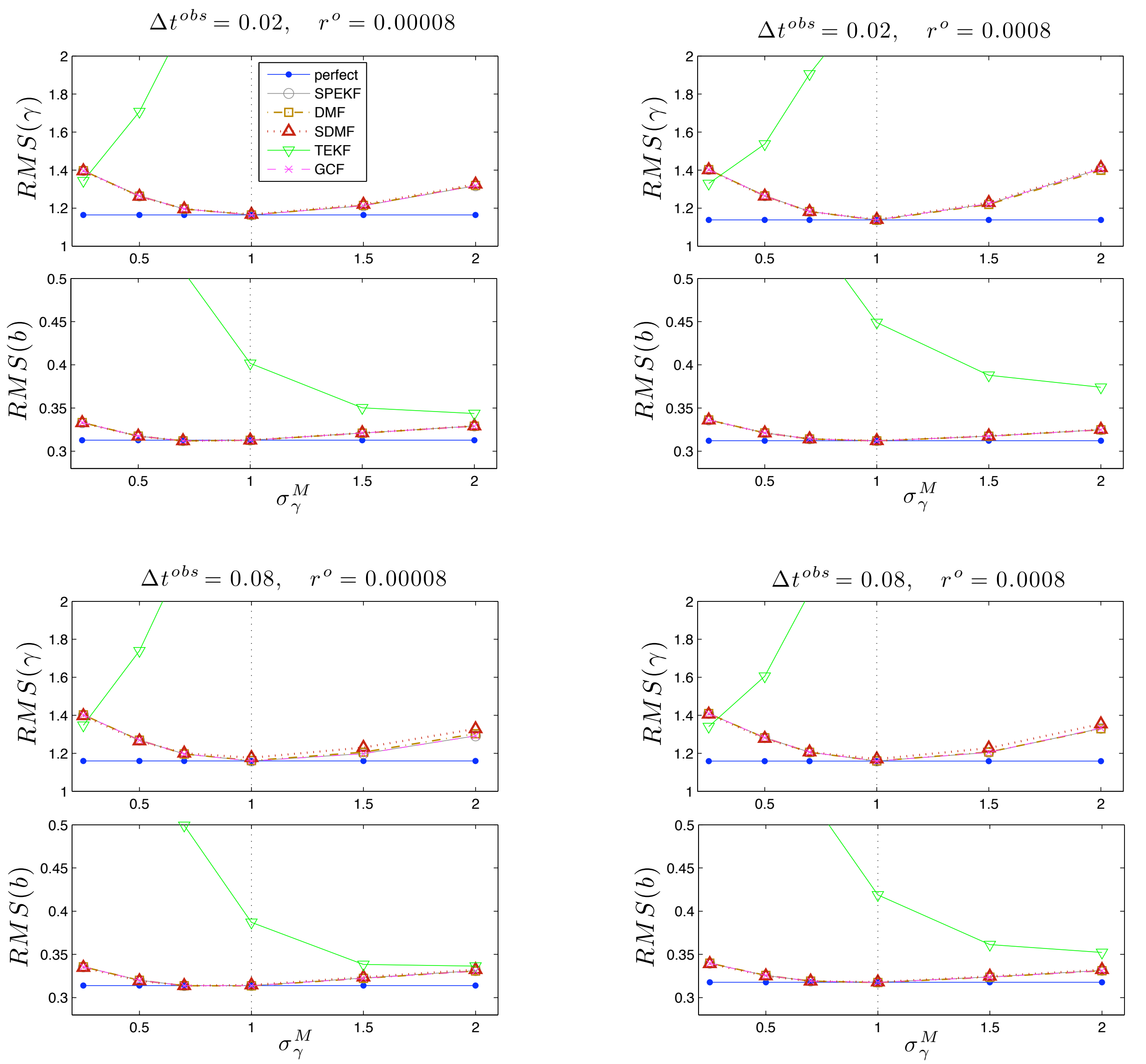

Figure 25: (Regime III) Parameter estimation with imperfect models. Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter σ_γ^M (incorrect noise amplitude in the dynamics of γ) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime III. Parameter estimation with imperfect models.

The mean stability parameter: $\boldsymbol{\chi = -0.1}$.
Decorrelation time of u: $1/\hat{\gamma} \approx 0.12$
True signal parameters: $\hat{\gamma} = 8.1, d_\gamma = 0.25, \sigma_\gamma = 1, \quad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$
Incorrect filter parameters: d_γ^M varied.

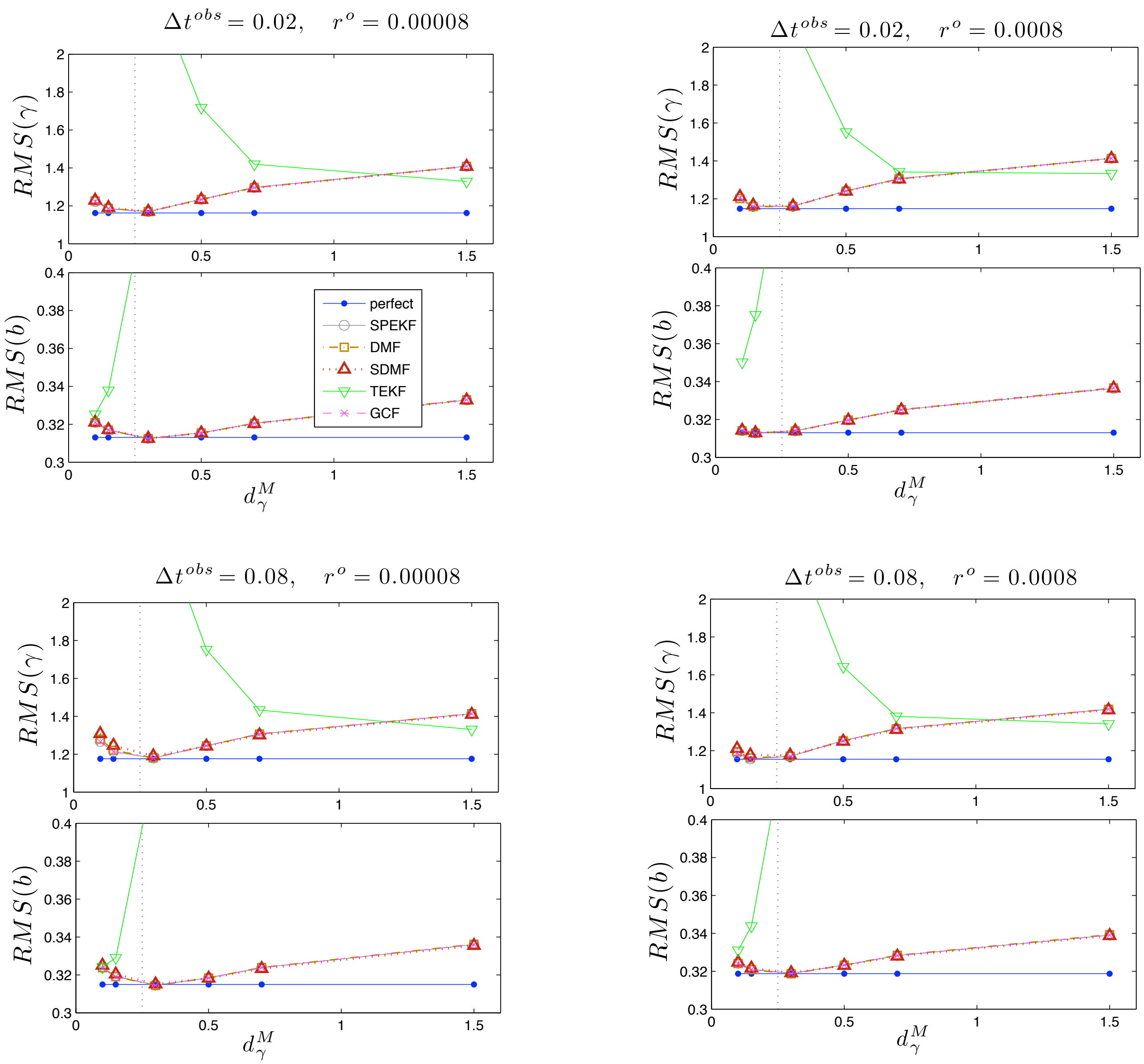

Figure 26: (Regime III) Parameter estimation with imperfect models. Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter d_γ^M (incorrect damping assumed for γ) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.

Regime III. Parameter estimation with imperfect models.

The mean stability parameter: $\boldsymbol{\chi = -0.1}$.
Decorrelation time of u: $1/\hat{\gamma} \approx 0.12$
True signal parameters: $\hat{\gamma} = 8.1, d_\gamma = 0.25, \sigma_\gamma = 1, \quad \omega_u = 1.78, \sigma_u = 0.25, \quad \gamma_b = 0.5, \omega_b = 1, \sigma_b = 0.5.$
Incorrect filter parameters: σ_u^M varied.

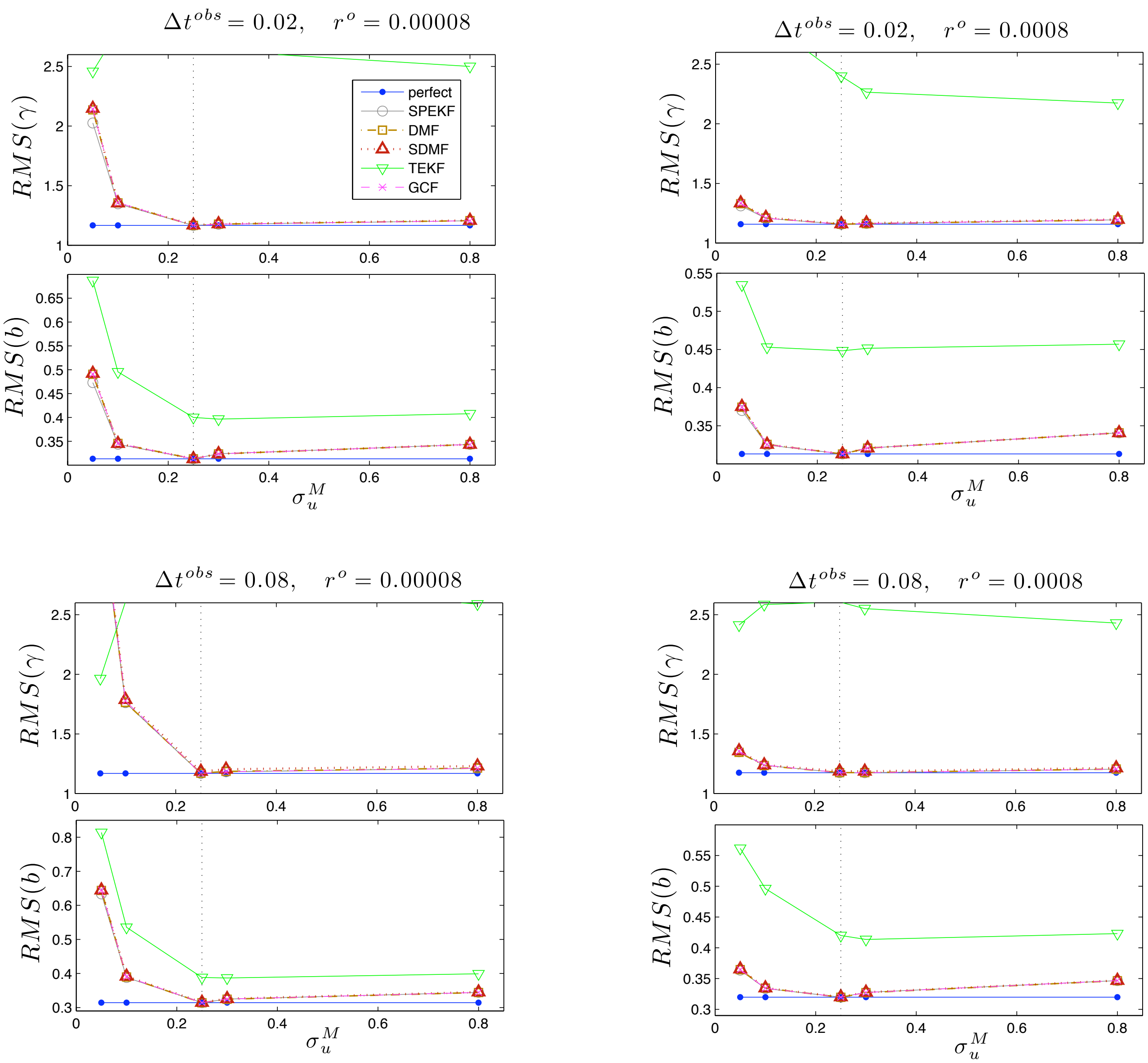

Figure 27: (Regime III) Parameter estimation with imperfect models. Average RMS errors of the filtered signal $u(t)$ as a function of the filter parameter σ_u^M (incorrect noise amplitude assumed in u) for fixed values of observation time Δt^{obs} and fixed observation noise variance r^o. The filter suite is described in §3.