1

Introduction and overview: Mathematical strategies for filtering turbulent systems

Filtering is the process of obtaining the best statistical estimate of a natural system from partial observations of the true signal from nature. In many contemporary applications in science and engineering, real-time filtering of a turbulent signal from nature involving many degrees of freedom is needed to make accurate predictions of the future state. This is obviously a problem with significant practical impact. Important contemporary examples involve the real-time filtering and prediction of weather and climate as well as the spread of hazardous plumes or pollutants. Thus, an important emerging scientific issue is the real-time filtering through observations of noisy signals for turbulent nonlinear dynamical systems as well as the statistical accuracy of spatio-temporal discretizations for filtering such systems. From the practical standpoint, the demand for operationally practical filtering methods escalates as the model resolution is significantly increased. In the coupled atmosphere–ocean system, the current practical models for prediction of both weather and climate involve general circulation models where the physical equations for these extremely complex flows are discretized in space and time and the effects of unresolved processes are parametrized according to various recipes; the result of this process involves a model for the prediction of weather and climate from partial observations of an extremely unstable, chaotic dynamical system with several billion degrees of freedom. These problems typically have many spatio-temporal scales, rough turbulent energy spectra in the solutions near the mesh scale, and a very large-dimensional state space, yet real-time predictions are needed.

Particle filtering of low-dimensional dynamical systems is an established discipline (Bain and Crisan, 2009). When the system is low dimensional or when it has a low-dimensional attractor, Monte Carlo approaches such as the particle filter (Chorin and Krause, 2004) with its various up-to-date resampling strategies (Del Moral, 1996; Del Moral and Jacod, 2001; Rossi and Vila, 2006) provide better estimates in the presence of strong nonlinearity and highly non-Gaussian distributions. However, with the above practical computational constraint in mind, these accurate nonlinear particle filtering strategies are not feasible since sampling a high-dimensional variable is computationally impossible for the foreseeable future. Recent mathematical theory strongly supports this curse of dimensionality for particle filters (Bengtsson *et al.*, 2008; Bickel *et al.*, 2008). Nevertheless some progress in developing particle filtering with small ensemble size for non-Gaussian turbulent dynamical systems is discussed in Chapter 15. These approaches, including the new maximum entropy particle filter (MEPF) due to the authors, all make judicious use of partial marginal distributions to avoid particle collapse. In the second direction, Bayesian hierarchical modeling (Berliner *et al.*, 2003) and reduced-order filtering strategies (Miller *et al.*, 1999; Ghil and Malanotte-Rizolli, 1991; Todling and Ghil, 1994; Anderson, 2001, 2003; Chorin and Krause, 2004; Farrell and Ioannou, 2001, 2005; Ott *et al.*, 2004; Hunt *et al.*, 2007; Harlim and Hunt, 2007b) based on the Kalman filter (Anderson and Moore, 1979; Chui and Chen, 1999; Kaipio and Somersalo, 2005) have been developed with some success in these extremely complex high-dimensional nonlinear systems. There is an inherently difficult practical issue of small ensemble size in filtering statistical solutions of these complex problems due to the large computational overload in generating individual ensemble members through the forward dynamical operator (Haven *et al.*, 2005). Numerous ensemble-based Kalman filters (Evensen, 2003; Bishop *et al.*, 2001; Anderson, 2001; Szunyogh *et al.*, 2005; Hunt *et al.*, 2007) show promising results in addressing this issue for synoptic-scale mid-latitude weather dynamics by imposing suitable spatial localization on the covariance updates; however, all these methods are very sensitive to model resolution, observation frequency and the nature of the turbulent signals when a practical limited ensemble size (typically less than 100) is used. They are also less skillful for more complex phenomena like gravity waves coupled with condensational heating from clouds which are important for the tropics and severe local weather.

Here is a list of fundamental new difficulties in the real-time filtering of turbulent signals that need to be addressed as mentioned briefly above.

1(a) **Turbulent dynamical systems to generate the true signal.** The true signal from nature arises from a turbulent nonlinear dynamical system with extremely complex noisy spatio-temporal signals which have significant amplitude over many spatial scales.

1(b) **Model errors.** A major difficulty in accurate filtering of noisy turbulent signals with many degrees of freedom is model error; the fact that the true signal from nature is processed for filtering and prediction through an imperfect model where by practical necessity, important physical processes are parametrized due to inadequate numerical resolution or incomplete physical understanding. The model errors of inadequate resolution often lead to rough turbulent energy spectra for the truth signal to be filtered on the order of the mesh scale for the dynamical system model used for filtering.

1(c) **Curse of ensemble size.** For forward models for filtering, the state space dimension is typically large, of order 10^4–10^8, for these turbulent dynamical systems, so generating an ensemble size with such a direct approach of order 50–100 members is typically all that is available for real-time filtering.

1(d) **Sparse, noisy, spatio-temporal observations for only a partial subset of the variables.** In systems with multiple spatio-temporal scales, the sparse observations of the truth signal might automatically couple many spatial scales, as shown below in Chapter 7 or in Harlim and Majda (2008b), while the observation of a partial subset of variables might mix together temporal slow and fast components of the system (Gershgorin and Majda, 2008, 2010) as discussed in Chapter 10. For example, observations of pressure or temperature in the atmosphere mix slow vortical and fast gravity wave processes.

This book is an introduction to filtering with an emphasis on the central new issues in 1(a)–(d) for filtering turbulent dynamical systems through the "modus operandi" of the modern applied mathematics paradigm (Majda, 2000a) where rigorous mathematical theory, asymptotic and qualitative models, and novel numerical algorithms are all blended together interactively to give insight into central "cutting edge" practical science problems. In the last several years, the authors have utilized the synergy of modern applied mathematics to address the following:

2(a) How to develop simple off-line mathematical test criteria as guidelines for filtering extremely stiff multiple space–time scale problems that often arise in filtering turbulent signals through plentiful and sparse observations? (Majda and Grote, 2007; Castronovo *et al.*, 2008; Grote and Majda, 2006; Harlim and Majda, 2008b)

2(b) For turbulent signals from nature with many scales, even with mesh refinement, the model has inaccuracies from parametrization, under-resolution, etc. Can judicious model errors help filtering and simultaneously overcome the curse of dimensionality? (Castronovo *et al.*, 2008; Harlim and Majda, 2008a,b, 2010a)

2(c) Can new computational strategies based on stochastic parametrization algorithms be developed to overcome the curse of dimensionality, to reduce model error and improve the filtering as well as the prediction skill? (Gershgorin *et al.*, 2010a,b; Harlim and Majda, 2010b)

2(d) Can exactly solvable models be developed to elucidate the central issues in 1(d) for turbulent signals, to develop unambiguous insight into model errors and to lead to efficient new computational algorithms? (Gershgorin and Majda, 2008, 2010)

The main goals of this book are the following: first, to introduce the reader to filtering from this viewpoint in an elementary fashion where no prior background on these topics is assumed (Chapters 2–4); secondly, to describe in detail the recent and ongoing developments, emphasizing the remarkable new mathematical and physical phenomena that emerge from the modern applied mathematics modus operandi applied to filtering turbulent dynamical systems. Next, in this introductory chapter, we provide an overview of turbulent dynamical systems and basic filtering followed by an overview of the basic applied mathematics motivation which leads to the new developments and viewpoint emphasized in this book.

1.1 Turbulent dynamical systems and basic filtering

The large-dimensional turbulent dynamical systems which define the true signal from nature to be filtered in the class of problems studied here have a fundamentally different statistical character than in more familiar low-dimensional chaotic dynamical systems. The most well-known low-dimensional chaotic dynamical system is Lorenz's famous three-equation model (Lorenz, 1963) which is weakly mixing with one unstable direction on an attractor with high symmetry. In contrast, realistic turbulent dynamical systems have a large phase space dimension, a large dimensional unstable manifold on the attractor, and are strongly mixing with exponential decay of correlations. The simplest prototype example of a turbulent dynamical system is also due to Lorenz and is called the L-96 model (Lorenz, 1996; Lorenz and Emanuel, 1998). It is widely used as a test model for algorithms for prediction, filtering and low-frequency climate response (Majda *et al.*, 2005; Majda and Wang, 2006). The L-96 model is a discrete periodic model given by the following system

$$\frac{du_j}{dt} = (u_{j+1} - u_{j-2})u_{j-1} - u_j + F, \quad j = 0, \ldots, J-1, \tag{1.1}$$

with $J = 40$ and with F the forcing parameter. The model is designed to mimic baroclinic turbulence in the mid-latitude atmosphere with the effects of energy-conserving nonlinear advection and dissipation represented by the first two terms in (1.1). For sufficiently strong forcing values such as $F = 6, 8, 16$, the L-96 model is a prototype turbulent dynamical system which exhibits features of weakly chaotic turbulence ($F = 6$), strongly chaotic turbulence ($F = 8$), and strong turbulence ($F = 16$) (Majda *et al.*, 2005). In order to quantify and compare the different types of turbulent chaotic dynamics in the L-96 model as F is varied, it is convenient to rescale the system to have unit energy for statistical fluctuations around the constant mean statistical state, $\bar{u}$ (Majda *et al.*, 2005); thus, the transformation $u_j = \bar{u} + E_p^{1/2}\tilde{u}_j$, $t = \tilde{t}E_p^{-1/2}$ is utilized where E_p represents the energy fluctuations (Majda *et al.*, 2005). After this normalization, the mean state becomes zero and the energy fluctuations are unity for all values of F. The dynamical equation in terms of the new variables, $\tilde{u}_j$, becomes

$$\frac{d\tilde{u}_j}{d\tilde{t}} = (\tilde{u}_{j+1} - \tilde{u}_{j-2})\tilde{u}_{j-1} + E_p^{-1/2}((\tilde{u}_{j+1} - \tilde{u}_{j-2})\bar{u} - \tilde{u}_j) + E_p^{-1}(F - \bar{u}). \tag{1.2}$$

Table 1.1 lists, in the non-dimensional coordinates, the leading Lyapunov exponent, λ_1, the dimension of the unstable manifold, N^+, the sum of the positive Lyapunov exponents (the KS entropy) and the correlation time, T_{corr}, of any $\tilde{u}_j$ variable with itself as F is varied

Table 1.1 Dynamical properties of the L-96 model for regimes with $F = 6, 8, 16$. λ_1 denotes the largest Lyapunov exponent, N^+ denotes the dimension of the expanding subspace of the attractor, KS denotes the Kolmogorov–Sinai entropy and T_{corr} denotes the decorrelation time of the energy-rescaled time correlation function.

	F	λ_1	N^+	KS	T_{corr}
Weakly chaotic	6	1.02	12	5.547	8.23
Strongly chaotic	8	1.74	13	10.94	6.704
Fully turbulent	16	3.945	16	27.94	5.594

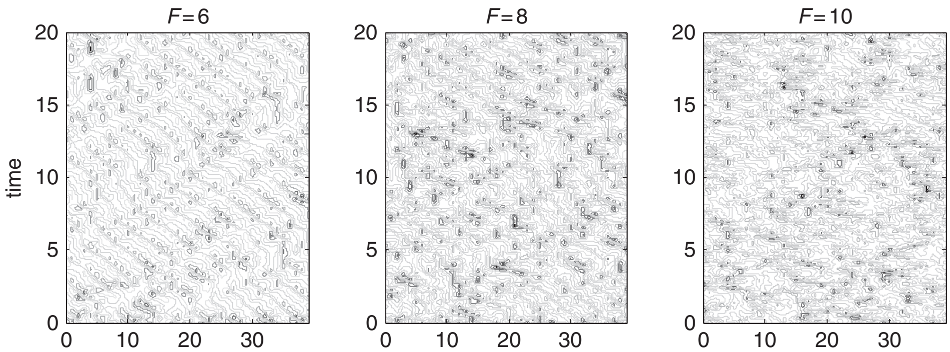

Figure 1.1 Space–time diagrams of numerical solutions of the L-96 model for the weakly chaotic ($F = 6$), strongly chaotic ($F = 8$) and fully turbulent ($F = 16$) regimes.

through $F = 6, 8, 16$. Note that λ_1, N^+ and KS increase significantly as F increases while T_{corr} decreases in these non-dimensional units; furthermore, the weakly turbulent case with $F = 6$ already has a 12-dimensional unstable manifold in the 40-dimensional phase space. Snapshots of the time series for (1.1) with $F = 6, 8, 16$, as depicted in Fig. 1.1, qualitatively confirm the above quantitative intuition with weakly turbulent patterns for $F = 6$, strongly chaotic wave turbulence for $F = 8$, and fully developed wave turbulence for $F = 16$. It is worth remarking here that smaller values of F around $F = 4$ exhibit the more familiar low-dimensional weakly chaotic behavior associated with the transition to turbulence.

In regimes to realistically mimic properties of nature, virtually all atmosphere, ocean and climate models with sufficiently high resolution are turbulent dynamical systems with features as described above. The simplest paradigm model of this type is the two-layer quasi-geostrophic (QG) model in doubly periodic geometry that is externally forced by a mean vertical shear (Smith *et al.*, 2002), which has baroclinic instability (Salmon, 1998); the properties of the turbulent cascade have been extensively discussed in this setting, e.g. see Salmon (1998) and citations in Smith *et al.* (2002). The governing equations for the two-layer QG model with a flat bottom, rigid lid and equal-depth layers H can be written as

$$\begin{aligned} \frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U\frac{\partial q_1}{\partial x} + (\beta + k_d^2 U)\frac{\partial \psi_1}{\partial x} + \nu\nabla^8 q_1 = 0, \\ \frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U\frac{\partial q_2}{\partial x} + (\beta - k_d^2 U)\frac{\partial \psi_2}{\partial x} + \kappa\nabla^2\psi_2 + \nu\nabla^8 q_2 = 0, \end{aligned} \tag{1.3}$$

where subscript 1 denotes the top layer and 2 the bottom layer; ψ is the perturbed stream function; $J(\psi, q) = \psi_x q_y - \psi_y q_x$ is the Jacobian term representing nonlinear advection; U is the zonal mean shear; β is the meridional gradient of the Coriolis parameter; q is the perturbed quasi-geostropic potential vorticity, defined as follows

$$q_i = \beta y + \nabla^2\psi_i + \frac{k_d^2}{2}(\psi_{3-i} - \psi_i), \quad i = 1, 2, \tag{1.4}$$

where $k_d = \sqrt{8}/L_d$ is the wavenumber corresponding to the Rossby radius L_d; κ is the Ekman bottom drag coefficient; and ν is the hyperviscosity coefficient. Note that Eqns (1.3) are the prognostic equations for perturbations around a uniform shear with stream function $\Psi_1 = -Uy, \Psi_2 = Uy$ as the background state, and the hyperviscosity term, $\nu\nabla^8 q$, is added to filter out the energy buildup on the smaller scales.

This is the simplest climate model for the poleward transport of heat in the atmosphere or ocean and with a modest resolution of $128 \times 128 \times 2$ grid points has a phase space of more than 30,000 variables. Again for modeling the atmosphere and ocean, this model in the appropriate parameter regimes is a strongly turbulent dynamical system with strong cascades of energy (Salmon, 1998; Smith *et al.*, 2002; Kleeman and Majda, 2005); it has been utilized recently as a test model for algorithms for filtering sparsely observed turbulent signals in the atmosphere and ocean (Harlim and Majda, 2010b).

1.1.1 Basic filtering

We assume that observations are made at uniform discrete times, $m\Delta t$, with $m = 1, 2, 3, \ldots$ For example, in global weather prediction models, the observations are given as inputs in the model every six hours and for large-dimensional turbulent dynamical systems, it is a challenge to implement continuous observations, practically. As depicted in Fig. 1.2, filtering is a two-step process involving statistical prediction of a probability distribution for the state variable u through a forward operator on the time interval between observations followed by an analysis step at the next observation time which corrects this probability distribution on the basis of the statistical input of noisy observations of the system. In the present applications, the forward operator is a large-dimensional dynamical system perhaps with noise written in the Itô sense as

$$\frac{du}{dt} = F(u,t) + \sigma(u,t)\dot{W}(t) \tag{1.5}$$

for $u \in \mathbb{R}^N$, where σ is an $N \times K$ noise matrix and $\dot{W} \in \mathbb{R}^K$ is K-dimensional white noise. The Fokker–Planck equation for the probability density, $p(u,t)$, associated with (1.5) is

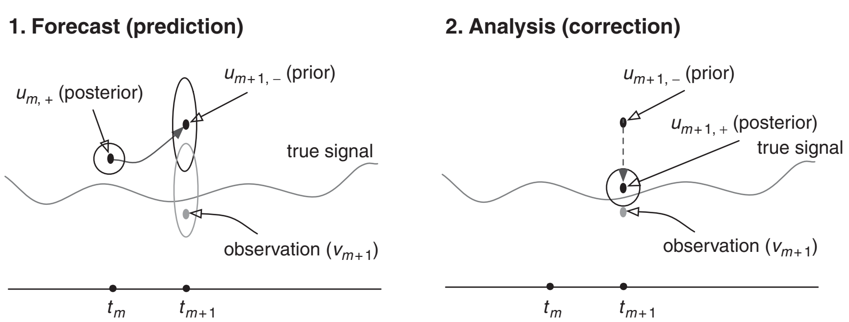

Figure 1.2 Filtering: Two-step predictor–corrector method.

$$p_t = -\nabla_u \cdot (F(u,t)p) + \frac{1}{2}\nabla_u \cdot \nabla_u (Qp) \equiv L_{\mathrm{FP}}\, p \tag{1.6}$$
$$p_t|_{t=t_0} = p_0(u)$$

with $Q(t) = \sigma\sigma^T$. For simplicity in exposition, here and throughout the remainder of the book we assume M linear observations, $\vec{v}_m \in \mathbb{R}^M$, of the true signal from nature given by

$$\vec{v}_m = Gu(m\Delta t) + \vec{\sigma}_m^o, \; m = 1, 2, \ldots \tag{1.7}$$

where G maps $\mathbb{R}^N$ into $\mathbb{R}^M$ while the observational noise, $\vec{\sigma}_m^o \in \mathbb{R}^M$, is assumed to be a zero-mean Gaussian random variable with $M \times M$ covariance matrix,

$$R^o = \langle \vec{\sigma}_m^o \otimes (\vec{\sigma}_m^o)^T \rangle. \tag{1.8}$$

Gaussian random variables are uniquely determined by their mean and covariance; here and below, we utilize the standard notation $\mathcal{N}(\vec{X}, R)$ to denote a vector Gaussian random variable with mean $\vec{X}$ and covariance matrix R. With these preliminaries, we describe the two-step filtering algorithm with the dynamics in (1.5), (1.6) and the noisy observations in (1.7), (1.8). Start at time step $m\Delta t$ with a posterior probability distribution, $p_{m,+}(u)$, which takes into account the observations in (1.7) at time $m\Delta t$. Calculate a prediction or forecast probability distribution, $p_{m+1,-}(u)$, by using (1.6), in other words, let p be the solution of the Fokker–Planck equation,

$$p_t = L_{\mathrm{FP}}\, p, \quad m\Delta t < t \leq (m+1)\Delta t \tag{1.9}$$
$$p|_{t=m\Delta t} = p_{m,+}(u).$$

Define $p_{m+1,-}(u)$, the prior probability distribution before taking observations at time $m+1$ into account, by

$$p_{m+1,-}(u) \equiv p(u, (m+1)\Delta t) \tag{1.10}$$

with p determined by the forward dynamics in (1.9). Next, the analysis step at time $(m+1)$ Δt which corrects this forecast and takes the observations into account is implemented by using Bayes' theorem

$$\begin{aligned} p_{m+1,+}(u)p(v_{m+1}) &= p_{m+1}(u|v_{m+1})p(v_{m+1}) \\ &= p_{m+1}(u,v) = p_{m+1}(v_{m+1}|u)p_{m+1,-}(u). \end{aligned} \tag{1.11}$$

With Bayes' formula in (1.11), we calculate the posterior distribution

$$p_{m+1,+}(u) \equiv p_{m+1}(u|v_{m+1}) = \frac{p_{m+1}(v_{m+1}|u)p_{m+1,-}(u)}{\int p_{m+1}(v_{m+1}|u)p_{m+1,-}(u)du}. \tag{1.12}$$

The two steps described in (1.9), (1.10), (1.12) define the basic nonlinear filtering algorithm which forms the theoretical basis for practical design of algorithms for filtering turbulent dynamical systems (Jazwinski, 1970; Bain and Crisan, 2009). While this is conceptually clear, practical implementation of (1.9), (1.10), (1.12), directly in turbulent dynamical systems, is impossible due to large state space, $N \gg 1$, as well as the fundamental difficulties elucidated in 1(a)–(d) in the introduction.

The most important and famous example of filtering is the Kalman filter where the analysis step in (1.5) is associated with linear dynamics which can be integrated between observation time steps $m\Delta t$ and $(m+1)\Delta t$ to yield the forward operator

$$u_{m+1} = Fu_m + \bar{f}_{m+1} + \sigma_{m+1}. \tag{1.13}$$

Here F is the $N \times N$ system operator matrix and σ_m is the system noise assumed to be zero-mean and Gaussian with $N \times N$ covariance matrix

$$R = \langle \sigma_m \otimes \sigma_m^T \rangle, \ \forall m, \tag{1.14}$$

while $\bar{f}_m$ is a deterministic forcing. Next, we present the simplified Kalman filter equations for the linear case. First assume the initial probability density $p_0(u)$ is Gaussian, i.e. $p_0(u) = \mathcal{N}(\bar{u}_0, R_o)$ and assume by recursion that the posterior probability distribution, $p_{m,+}(u) = \mathcal{N}(\bar{u}_{m,+}, R_{m,+})$, is also Gaussian. By using the linear dynamics in (1.13), the forecast or prediction distribution at time $(m+1)\Delta t$ is also Gaussian,

$$\begin{aligned} p_{m+1,-}(u) &= \mathcal{N}(\bar{u}_{m+1,-}, R_{m+1,-}) \\ \bar{u}_{m+1,-} &= F\bar{u}_{m,+} + \bar{f}_{m+1} \\ R_{m+1,-} &= FR_{m,+}F^T + R. \end{aligned} \tag{1.15}$$

With the assumptions in (1.7), (1.8) and (1.13), (1.15), the analysis step in (1.12) becomes an explicit regression procedure for Gaussian random variables (Chui and Chen, 1999; Anderson and Moore, 1979) so that the posterior distribution, $p_{m+1,+}(u)$, is also Gaussian yielding the **Kalman filter**

$$\begin{aligned} p_{m+1,+}(u) &= \mathcal{N}(\bar{u}_{m+1,+}, R_{m+1,+}) \\ \bar{u}_{m+1,+} &= (\mathcal{I} - K_{m+1}G)\bar{u}_{m+1,-} + K_{m+1}v_{m+1} \\ R_{m+1,+} &= (\mathcal{I} - K_{m+1}G)R_{m+1,-} \\ K_{m+1} &= R_{m+1,-}G^T(GR_{m+1,-}G^T + R^o)^{-1}. \end{aligned} \tag{1.16}$$

The $N \times M$ matrix, K_{m+1}, is the Kalman gain matrix. Note that the posterior mean after processing the observations is a weighted sum of the forecast and analysis contributions through the Kalman gain matrix and also that the observations reduce the covariance, $R_{m+1,+} \leq R_{m+1,-}$. In this Gaussian case with linear observations, the analysis step going from (1.15) to (1.16) is a standard linear least-squares regression. An excellent treatment of this can be found in chapter 3 of Kaipio and Somersalo (2005). There is a huge literature on Kalman filtering; two excellent basic texts are Chui and Chen (1999) and Anderson and Moore (1979) where more details and references can be found. Our intention in the introductory parts in this book in Chapters 2 and 3 is not to repeat the well-known material in (1.15), (1.16) in detail; instead we introduce this elementary material in a fashion to set the stage for the mathematical guidelines developed in Part II (Chapters 5–8) and the applications to filtering turbulent nonlinear dynamical systems presented in Part III (Chapters 9–15).

Naively, the reader might expect that everything is known about filtering linear systems; however, when the linear system is high dimensional, i.e. $N \gg 1$, the same issues elucidated in 1(a)–(d) occur for linear systems in a more transparent fashion. This is the viewpoint emphasized and developed in Part II of the book (Chapters 5–8) which is motivated next. For linear systems without model errors, the recursive Kalman filter is an optimal estimator but the recursive nonlinear filter in (1.7)–(1.12) may not be an optimal estimator for the nonlinear stochastic dynamical system without model error in (1.5).

1.2 Mathematical guidelines for filtering turbulent dynamical systems

How can useful mathematical guidelines be developed in order to elucidate and ameliorate the central new issues in 1(a)–(d) from the introduction for turbulent dynamical systems? This is the topic of this section. Of course, to be useful, such mathematical guidelines have to be general yet still involve simplified models with analytical tractability. Such criteria have been developed recently by Majda and Grote (2007); Castronovo *et al.* (2008) and Harlim and Majda (2008b) through the modern applied mathematics paradigm and the goal here is to outline this development and discuss some of the remarkable phenomena which occur. The starting point for this development for filtering turbulent dynamical systems involves the symbiotic interaction of three different disciplines in applied mathematics/physics, as depicted in Fig. 1.3: stochastic modeling of turbulent signals, numerical analysis of PDEs and classical filtering theory outlined in (1.13)–(1.16) of Section 1.1. Here is the motivation from the three legs of the triangle.

First, the simplest stochastic models for modeling turbulent fluctuations consist of replacing the nonlinear interaction at these modes by additional dissipation and white noise

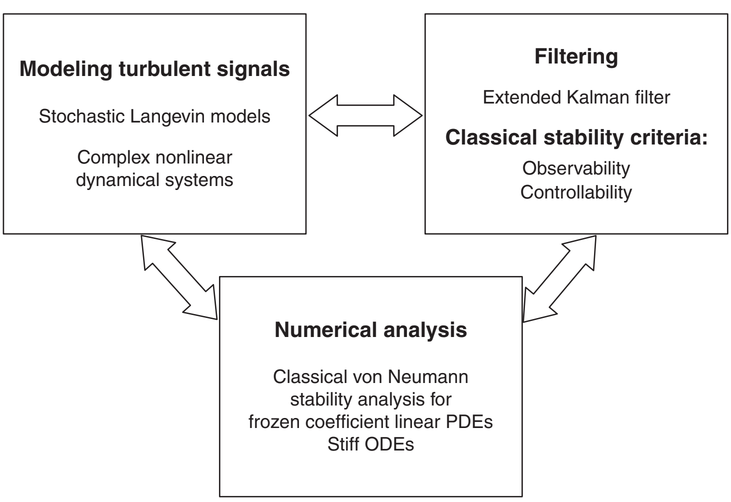

Figure 1.3 Modern applied mathematics paradigm for filtering.

forcing to mimic rapid energy transfer (Salmon, 1998; Majda *et al.*, 1999, 2003; Majda and Timofeyev, 2004; Delsole, 2004; Majda *et al.*, 2008). Conceptually, we view this stochastic model for a given turbulent (Fourier) mode as given by the linear Langevin SDE or Ornstein–Uhlenbeck process for the complex scalar

$$du(t) = \lambda u(t)dt + \sigma dW(t), \qquad (1.17)$$
$$\lambda = -\gamma + \mathrm{i}\omega, \gamma > 0,$$

with $W(t)$ a complex Wiener process, and σ its noise strength. Of course, the amplitude and strength of these coefficients, γ, σ, vary widely for different Fourier modes and depend empirically on the nonlinear nature of the turbulent cascade, the energy spectrum, etc. These simplest turbulence models are developed in detail in Chapter 5 and an important extension with intermittent instability at large scales is developed in Chapter 8. Quantitative illustrations of this modeling process for the L-96 model in (1.1) in a variety of regimes and the two-layer model in (1.3) are developed in Part III in Chapters 12 and 13, together with cheap stochastic filters with judicious model errors based on these linear stochastic models.

Secondly, the most successful mathematical guideline for numerical methods for deterministic nonlinear systems of PDEs is von Neumann stability analysis (Richtmeyer and Morton, 1967): The nonlinear problem is linearized at a constant background state, and Fourier analysis is utilized for this constant-coefficient PDE, resulting in discrete approximations for a complex scalar test model for each Fourier mode,

$$\frac{du(t)}{dt} = \lambda u(t), \lambda = -\gamma + \mathrm{i}\omega, \gamma > 0. \qquad (1.18)$$

All the classical mathematical phenomena such as, for example, the CFL stability condition on the time step Δt and spatial mesh h, $|c|\Delta t/h < 1$, for various explicit schemes for the advection equation $u_t + cu_x = -du$, occur because, at high spatial wavenumbers, the scalar test problem in (1.18) is a stiff ODE, i.e.

$$|\lambda| \gg 1. \qquad (1.19)$$

For completeness, Chapter 4 provides a brief introduction to this analysis.

The third leg of the triangle involves classical linear Kalman filtering as outlined in (1.13)–(1.16). In conventional mathematical theory for filtering linear systems, one checks algebraic observability and controllability conditions (Chui and Chen, 1999; Anderson and Moore, 1979) and is automatically guaranteed asymptotic stability for the filter; this theory applies for a fixed state dimension and is a very useful mathematical guideline for linear systems that are not stiff in low-dimensional state space. Grote and Majda (2006) developed striking examples involving unstable differencing of the stochastic heat equation where the state space dimension is $N = 42$ with 10 unstable modes where the classical observability (Cohn and Dee, 1988) and controllability conditions were satisfied yet the filter covariance matrix had condition number 10^{13} so there is no practical filtering skill!