J. Fluid Mech. (1999), *vol.* 398, *pp.* 321–346. Printed in the United Kingdom

Optimal LES formulations for isotropic turbulence

By JACOB A. LANGFORD AND ROBERT D. MOSER

Department of Theoretical and Applied Mechanics, University of Illinois, Urbana, IL 61801, USA

(Received 17 July 1998 and in revised form 15 June 1999)

It is shown that there is an abstract subgrid model that is in all senses ideal. An LES using the ideal subgrid model will exactly reproduce all single-time, multi-point statistics, and at the same time will have minimum possible error in instantaneous dynamics. The ideal model is written as an average over the real turbulent fields whose large scales match the current LES field. But this conditional average cannot be computed directly. Rather, the ideal model is the target for approximation when developing practical models, though no new practical models are presented here. To construct such models, the conditional average can be formally approximated using stochastic estimation. These *optimal formulations* are presented, and it is shown that a relatively simple but general class of one-point estimates can be computed from two-point correlation data, and that the estimates retain some of the statistical properties of the ideal model.

To investigate the nature of these models, optimal formulations were applied to forced isotropic turbulence. A variety of optimal models of increasing complexity were computed. In all cases, it was found that the errors between the real and estimated subgrid force were nearly as large as the subgrid force itself. It is suggested that this may also be characteristic of the ideal model in isotropic turbulence. If this is the case, then it explains why subgrid models produce reasonable results in actual LES while performing poorly in *a priori* tests. Despite the large errors in the optimal models, one feature of the subgrid interaction that is exactly represented is the energy transfer to the subgrid scales by each wavenumber.

1. Introduction

1.1. *The current LES formulation*

Large-eddy simulation of turbulence (LES) is essentially an under-resolved turbulence simulation that uses a model to account for the lack of small-scale resolution (see Rogallo & Moin 1984). Such a simulation is motivated by the recognition that the large scales of turbulence often dominate heat transfer, mixing, and other quantities of engineering interest, while the small scales of turbulence are important in these cases only because they affect the large scales. There has been optimism that LES can be used as a robust predictive tool, partially because small scales of turbulence are believed to be more isotropic and more universal than the large scales, and partially because LES has already been successful in many flows. For reviews of LES see Rogallo & Moin (1984) and Lesieur & Métais (1996).

The standard formulation of LES generally begins by applying a spatial filter to

the Navier–Stokes and continuity equations:

$$\frac{\widetilde{\partial u_i}}{\partial t} = -\frac{\widetilde{\partial u_i u_j}}{\partial x_j} - \frac{\widetilde{\partial p}}{\partial x_i} + \frac{1}{Re}\frac{\widetilde{\partial^2 u_i}}{\partial x_j \partial x_j}. \tag{1.1}$$

The filtered Navier–Stokes equations can then be written as

$$\frac{\partial \widetilde{u}_i}{\partial t} = -\frac{\partial \widetilde{u}_i \widetilde{u}_j}{\partial x_j} - \frac{\partial \widetilde{p}}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 \widetilde{u}_i}{\partial x_j \partial x_j} + M_i, \tag{1.2}$$

where

$$M_i = -\frac{\partial \tau_{ij}}{\partial x_j} + C_i, \tag{1.3}$$

$$\tau_{ij} = \widetilde{u_i u_j} - \widetilde{u}_i \widetilde{u}_j, \tag{1.4}$$

and C_i is a term that arises when the filter does not commute with spatial differentiation. Notice that M_i must be modelled because it contains terms that cannot be written as functions of the filtered field $\widetilde{u}$. LES research has largely been directed towards modelling the subgrid stress τ_{ij}.

The most common class of subgrid models is the eddy viscosity type. These models are generally postulated by analogy to the stresses produced by molecular viscosity. However, molecular viscosity models work because there is a clear separation between molecular and continuum fluid dynamic scales, and no such separation of scales exists for the LES problem (see Chollet & Lesieur 1981). Nevertheless, an eddy viscosity model takes the form

$$\tau_{ij} = -2\nu_T S_{ij}, \tag{1.5}$$

$$S_{ij} = \frac{1}{2}\left(\frac{\partial \widetilde{u}_i}{\partial x_j} + \frac{\partial \widetilde{u}_j}{\partial x_i}\right), \tag{1.6}$$

where ν_T is the eddy viscosity and S_{ij} is the strain rate of the resolved scales. Smagorinsky (1963) proposed

$$\nu_T = (C_S \Delta)^2 |S|, \tag{1.7}$$

where C_S is a constant, and Δ is a length characteristic of the filter. The constant C_S can be deduced analytically by assuming a filter cutoff in the inertial range (see Lilly 1964), or it can be determined empirically.

Eddy-viscosity models have worked relatively well for homogeneous flows, but there are problems near walls, where subgrid-scale structures play an important role in momentum transfer to the wall (see Panton 1997). *A priori* studies comparing eddy-viscosity stresses to real subgrid stresses have found that for eddy-viscosity models, real and predicted stresses do not correlate well (see Clark, Ferziger & Reynolds 1979; Meneveau 1994).

The concept of a wavenumber-dependent eddy viscosity emerged from the direct interaction theory of Kraichnan (1959) and the EDQNM theory of Orszag (1970). Further work by Kraichnan (1976) showed that a nearly constant eddy viscosity with a sharp increase at the cutoff can be used to accurately predict the energy spectrum in homogeneous turbulence. Further, the value of the plateau scales with the cutoff wavenumber and energy at the cutoff, provided that the cutoff is in an inertial range. However, a wavenumber-dependent eddy viscosity is difficult to compute in physical space. It was found by Leslie & Quarini (1979) that with different filters the eddy viscosity should display less dependence on wavenumber. The structure function

model of Métais & Lesieur (1992) uses a physical space formulation and a spatially-varying eddy viscosity. The model provides a mechanism for evaluating the energy spectrum at the cutoff, which can then be used to compute the local eddy viscosity.

The class of dynamic models was introduced by Germano *et al.* (1991), and is intended to dynamically adjust the Smagorinsky constant (or other adjustable constant) to local features of the flow. The basic concept is to apply a second filter of larger width Δ' to the LES field. The second filter defines a second subgrid stress,

$$T_{ij} = \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u}}_i \widetilde{\overline{u}}_j.$$

Assuming that the second subgrid stress obeys a similar eddy viscosity relation, i.e. that

$$T_{ij} = 2(C_S \Delta')^2 |S'| S'_{ij}, \tag{1.8}$$

one can write an over-determined but closed system of equations for the eddy viscosity coefficient C_S:

$$\begin{aligned} \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u}}_i \widetilde{\overline{u}}_j &= T_{ij} - \overline{\tau_{ij}} \\ &= 2(C_S \Delta')^2 |S'| S'_{ij} - \overline{2(C_S \Delta)^2 |S| S_{ij}}. \end{aligned} \tag{1.9}$$

The over-determinacy can be removed by a contraction with S_{ij}, but it is more common to use the least-squares solution proposed by Lilly (1992).

There are some mathematical and practical problems that the dynamic method has had to overcome. To allow an explicit solution for C_S, it is generally assumed that $\overline{C_S^2 |S| S_{ij}} = C_S^2 \overline{|S| S_{ij}}$, but it has been found that the coefficient C_S varies strongly in space, so that the assumption is not justified (see Lund, Ghosal & Moin 1993). The large variance of C_S also presents practical issues of stability, particularly because large negative values of C_S cause an unphysical backscatter of energy that destabilizes the simulation. There are several ways to deal with the excessive variance of C_S, mostly involving filters or averages over homogeneous directions or fluid path lines (see Germano *et al.* 1991; Meneveau, Lund & Cabot 1996). But to some degree these techniques all destroy the spatial locality that was intended of the dynamic model. An alternative is the dynamic localization procedure of Ghosal *et al.* (1995), in which one solves an integral equation for C_S to avoid removing C_S^2 from the filter operator in (1.9). In the process, one can also constrain C_S^2 to be positive to avoid stability problems, or one can allow it to become locally negative to model backscatter.

The scale-similarity model of Bardina, Ferziger & Reynolds (1980) is based on a second application of the filter, or more generally the application of a different second filter. A new filter defines a second subgrid stress

$$T_{ij} = \widetilde{\overline{u}_i \overline{u}_j} - \widetilde{\overline{u}}_i \widetilde{\overline{u}}_j, \tag{1.10}$$

that can be assumed proportional to the real subgrid stress. Unlike eddy-viscosity models, scale-similarity models predict stresses that correlate relatively well with real stresses (see Liu, Meneveau & Katz 1994). But real simulations have found that the model does not properly dissipate energy, so the model is often coupled with an eddy-viscosity model to achieve the proper dissipation. The result is called a mixed model.

1.2. *Need for a new approach*

As discussed above, a broad array of LES models have been developed using a variety of different approximations. Such models have been successful in many flows,

yet they all have shortcomings. Several of the difficult issues that underlie current LES practice are listed below:

(*a*) *Filtering*: In LES practice, there is commonly little attention paid to the definition of the filter, and in many cases the filter is defined implicitly by the numerics of the simulation. Further the only dependence of the subgrid model on the filter is generally through the filter width; yet the interaction between subgrid and resolved scales depends strongly on the choice of filter (Leslie & Quarini 1979; Piomelli 1988).

(*b*) *Numerics*: As has been pointed out by Ghosal (1996), even when using high-order numerical methods, the numerical discretization errors can far exceed the subgrid model term. This can be circumvented by using grid resolution that is finer than the filter width, though this is rarely done due to the expense. The alternative is to account for numerical discretization in the formulation of the subgrid model. Generally, current modelling techniques do not account for discretization errors, though some dissipative numerical discretizations have been proposed as implicit subgrid models (Boris *et al.* 1992).

(*c*) *Modelling*: Current subgrid models are derived under restrictive assumptions regarding the nature of the subgrid turbulence. In particular, it is generally assumed that the subgrid scales are locally homogeneous and isotropic and that the filter scale is in an inertial range. However, these assumptions are violated in flow situations of common interest such as near walls. Development of wall models that will not require the refinement of the near-wall grid with increasing Reynolds number is considered by many to be the most pressing challenge in LES research.

(*d*) *Inhomogeneity*: For inhomogeneous flows, it is usually necessary to introduce inhomogeneous filters. As is well known, such filters do not in general commute with spatial derivative operators, and this introduces new terms that in principle should be modelled. These 'commutation errors' are not commonly modelled, largely because modelling techniques have not been available for them. However, Ghosal & Moin (1995) provide suggestions of how commutation error might be treated, and Vasilyev, Lund & Moin (1998) show that approximately commuting filters can be defined for one-dimensional filter inhomogeneities.

(*e*) *Diagnostics*: Efforts at LES development have been hindered by a lack of effective diagnostic techniques. Currently available diagnostics are *a priori* tests in which models are compared to actual subgrid terms computed from DNS, and *a posteriori* tests in which an LES is performed and statistics are compared with available data. The issue of interest, that is the performance of LES simulations, is measured directly by *a posteriori* tests, but they provide little diagnostic information regarding why a simulation fails or succeeds. On the other hand, *a priori* tests can provide detailed diagnostics of model accuracy. But such diagnostics are generally misleading since most models do poorly on such tests, while performing adequately in actual simulations.

In an effort to address these issues, we introduce in this paper a new approach to the development of large-eddy simulation, which we call optimal LES formulation. In principle, by using this approach, one can account for all the complications discussed above, and diagnostics are an integral part of the procedure. This work starts from the trivial observation that the loss of information incurred during filtering places a limit on the achievable accuracy of LES. The unique best possible LES model is thus the one that achieves this limit, and we call it the ideal LES. The optimal LES formulation then involves systematic approximation of the ideal LES through formal optimization.

In the remainder of this paper, the details of the optimal LES formulation are presented and it is applied to LES of isotropic turbulence with Fourier cutoff filters as an example. Of course, LES of isotropic turbulence does not suffer from most of the difficulties cited above. The goal here is to determine how the optimal formulation relates to well established techniques for simulation of isotropic turbulence, rather than to develop new 'practical' models to address the above difficulties. Such new practical models will require further analysis in different flows and with different filters. Such analysis is being pursued.

The ideal LES is defined in § 2 and its statistical and dynamical accuracy are determined. A systematic technique for approximating the ideal LES is presented in § 3, and this is applied to isotropic turbulence in § 4. Concluding remarks appear in § 5.

2. Definition of ideal LES

The most generally accepted goal of LES is to accurately reproduce large-scale statistics. However, there is also interest in details of the large-scale flow dynamics, though it must be recognized that such details can only be predicted for short times due to the chaotic nature of turbulence. An example application for which this is an issue is the flow in turbomachinery. Here, the large-scale vortices that are shed from a stator (for example) interact with a downstream rotor, and the details of this interaction are of interest. It will be shown in §§ 2.3–2.4 that there is a unique LES evolution that generates accurate spatial statistics *and* minimizes error of the large-scale dynamics. These two properties are so important that this unique evolution will be called *ideal LES*. For an LES field w, a real turbulent field u, and a filter $\sim$, ideal LES is governed by the conditional average:

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \left\langle \widetilde{\frac{\mathrm{d}u}{\mathrm{d}t}} \middle| \tilde{u} = w \right\rangle. \tag{2.1}$$

Note that the symbols u and w denote entire fields, which is why the time-derivatives are not partial derivatives. When working with ideal LES, it is important to think of entire turbulent fields as single entities. Throughout this paper, u will be used to represent an entire field. Normal vector quantities at a particular spatial location will be written with direct notation as $\boldsymbol{u}$ or with indicial notation as u_i. The symbols w, $\boldsymbol{w}$, and w_i denote LES quantities, while $\tilde{u}$, $\tilde{\boldsymbol{u}}$, and $\tilde{u}_i$ denote an application of the filter to velocities from a particular real field.

2.1. *The role of filtering*

Before exploring the properties of this ideal LES, we must more carefully define the mathematical properties of filters. One of the most important features of the large-eddy simulation of turbulence is that only the large scales of turbulence are to be simulated. The large scales to be simulated must be defined, and of course this is the role of the filter. For the filter to be useful in this context, it cannot be invertible, that is, it must discard information. More precisely, the filter is a mapping from the large-dimensional space of Navier–Stokes solutions for the problem under consideration (formally infinite-dimensional), to a smaller dimensional space that can be practically represented on a computer. Unlike in a DNS (which also uses a finite-dimensional representation), the loss of information is severe in an LES, so severe that the error committed by ignoring the missing information cannot be tolerated. If one were to use an invertible filter, then the dynamics of the filtered system would be identical to

that of the unfiltered; only the dependent variables describing it would be different. A properly resolved simulation of such a filtered system would be a DNS.

A variety of different filters have been used in performing LES simulations, but two common filters demand discussion since their properties are subtle. They are the implicit filter and the Gaussian filter.

It is common in LES to allow the numerical discretization to define the large scales being simulated. Since the numerical representation is inherently computable and finite-dimensional, it certainly defines a low-dimensional space in which the LES can be performed. However, with such numerically defined filters, less attention has been paid to the definition of the mapping from the Navier–Stokes space to the LES space. When the numerical method is finite difference, the mapping is commonly described as the top-hat filter, which when acting on some function g in one dimension is written

$$\widetilde{g}(x) = \frac{1}{\delta^+ - \delta^-} \int_{x-\delta^-(x)}^{x+\delta^+(x)} g(x')\,\mathrm{d}x'. \tag{2.2}$$

However, this filter is nearly invertible. It is the sampling of $\widetilde{g}$ on a grid that actually reduces the dimension of the representation space. The combination of the top-hat filter (2.2) and the sampling on a grid produces a filter that is more correctly thought of as a volume average on discrete volumes (as in a finite volume method).

Another common LES filter is a Gaussian, in one dimension:

$$\widetilde{g}(x) = \int_{-\infty}^{\infty} \mathrm{e}^{(x-x')^2/\delta^2} g(x')\,\mathrm{d}x'. \tag{2.3}$$

But the Gaussian filter is also formally invertible, so in such simulations it is again the numerical truncation (commonly a Fourier spectral method) that defines the reduced dimensional space of the solution. This is subtle because the Gaussian filter so strongly damps the small-scale turbulence that representing the filtered velocity with a truncated Fourier series is extremely accurate, if all one needs to do is represent the structure of the filtered field. However, when attempting to compute the evolution of the filtered field as in LES, the truncation of the representation is critical because the exact evolution of the Gaussian-filtered velocity field is determined by inverse filtering the field, computing the right-hand side of the Navier–Stokes equations and filtering the results. The defiltering operation amplifies the small scales to their original magnitude, making them non-negligible. However, if the filtered field is represented as a truncated Fourier series, the information about the small scales has been eliminated. Of course there are other difficulties with Gaussian defiltering due to the finite precision of computer calculations, but this is an independent issue.

In the remainder of the paper, we will therefore consider only non-invertible filters, that is, filters that discard information. Thus, in some cases our notion of a filter includes the numerical discretization that causes the actual loss of information. We will also make one mild assumption regarding the filter, and that is that it is a linear operator. The authors know of no examples of a nonlinear filter being used or proposed for an LES.

2.2. *The important LES spaces*

Following the above discussion, we define the filter to be a non-invertible linear operator $\mathrm{F} : \mathscr{V} \to \mathscr{W}$ that maps elements from the space $\mathscr{V}$ of turbulent velocity fields to the space $\mathscr{W}$ of LES fields. (Note that $\mathscr{W}$ need not be a subspace of $\mathscr{V}$, for example $\mathscr{W}$ could be a space of equivalence classes.) It will be convenient to define

two more spaces related to F. One is the null space (or kernel) $\mathscr{V}'$ of F, which is simply the space of all elements of $\mathscr{V}$ that map to 0 in $\mathscr{W}$. The other space we will need is a complement to $\mathscr{V}'$ in $\mathscr{V}$. A complement space $\widetilde{\mathscr{V}}$ is defined by the properties that all the elements of $\widetilde{\mathscr{V}}$ are linearly independent of the null space $\mathscr{V}'$ and that $\widetilde{\mathscr{V}}$ and $\mathscr{V}'$ together span $\mathscr{V}$. Note that the choice of $\widetilde{\mathscr{V}}$ is not unique, but the results of our analysis will not depend on this choice.

Given $\mathscr{V}'$ and $\widetilde{\mathscr{V}}$, there is a unique pair of projection operators $\mathrm{P}_{\widetilde{\mathscr{V}}} + \mathrm{P}_{\mathscr{V}'} = \mathrm{I}_{\mathscr{V}}$ that decompose an element of $\mathscr{V}$ into elements of $\mathscr{V}'$ and $\widetilde{\mathscr{V}}$:

$$u = \tilde{u} + u', \tag{2.4}$$

$$\tilde{u} = \mathrm{P}_{\widetilde{\mathscr{V}}} u, \quad \mathrm{P}_{\widetilde{\mathscr{V}}} : \mathscr{V} \to \widetilde{\mathscr{V}},$$
$$u' = \mathrm{P}_{\mathscr{V}'} u. \quad \mathrm{P}_{\mathscr{V}'} : \mathscr{V} \to \mathscr{V}'.$$

The term projection is used here to describe a transformation that has the property $\mathrm{PP} = \mathrm{P}$.

For any linear filter, once the space $\widetilde{\mathscr{V}}$ is defined, the filter can be written as the projection $\mathrm{P}_{\widetilde{\mathscr{V}}}$ from $\mathscr{V}$ to $\widetilde{\mathscr{V}}$ followed by an invertible linear transformation L from space $\widetilde{\mathscr{V}}$ to $\mathscr{W}$, the range of the filter. Thus

$$\mathrm{F} = \mathrm{L}\mathrm{P}_{\widetilde{\mathscr{V}}} \tag{2.5}$$

$$\mathrm{F} : \mathscr{V} \to \mathscr{W}, \quad \mathrm{P} : \mathscr{V} \to \widetilde{\mathscr{V}}, \quad \mathrm{L} : \widetilde{\mathscr{V}} \to \mathscr{W}, \quad \mathrm{L}^{-1} : \mathscr{W} \to \widetilde{\mathscr{V}}.$$

This representation of a filter will be needed in § 2.3 below.

2.3. *Accurate LES statistics*

In this subsection we aim to find an LES model that is guaranteed to get correct spatial statistics, leading to results suggested by S. Pope. By spatial statistics we mean any single-time multi-point statistical quantity of the large-scale field. To understand the statistical properties of dynamical systems it is convenient to consider the evolution of probability densities in phase-space. The phase-space probability density function (p.d.f.) of a system characterizes an entire ensemble of flows. The p.d.f. specifies the probability that any particular realization might be chosen at random, at that instant in time. But because each realization has a deterministic evolution, the p.d.f. has a deterministic evolution.

Let u denote an element of $\mathscr{V}$, the space of turbulent velocity fields as defined in § 2.2. Note that since $\mathscr{V}$ is a phase-space, u denotes an *entire* turbulent field. Then $\rho_{\mathscr{V}}(u, t)$ is a p.d.f. denoting the probability that a field randomly chosen from the turbulent ensemble at time t is equal to u. Let w denote an element of $\mathscr{W}$, the phase space of the LES which is also the image of $\mathscr{V}$ through the filter F. Then $\rho_{\mathscr{W}}(w, t)$ is a p.d.f. denoting the probability that a field randomly chosen from the LES ensemble at time t is equal to w.

Any spatial statistic can be computed by integrating the statistic over the probability measure space, i.e.

$$\langle G(w) \rangle = \int_{\mathscr{W}} G(w) \rho_{\mathscr{W}}(w) \, \mathrm{d}\mathscr{W}(w). \tag{2.6}$$

So to find the conditions under which an LES will produce correct spatial statistics, one must find the conditions under which the LES p.d.f. will match the real p.d.f.

Note that in general it is impossible to require that $\rho_{\mathscr{V}}(u) = \rho_{\mathscr{W}}(\mathrm{F}u)$, since the

dimensions of $\mathscr{V}$ and $\mathscr{W}$ are not the same. It will thus be possible to vary u in a way such that Fu remains constant. The problem is handled by defining a marginal p.d.f., which is obtained by integrating $\rho_{\mathscr{V}}(u)$ over the space in which Fu is constant. The region of integration is $\mathscr{V}' = \ker(\mathrm{F})$, the kernel (nullspace) of the filter.

The marginal p.d.f. $\rho_{\tilde{\mathscr{V}}}(\tilde{u}, t)$ is thus defined as

$$\rho_{\tilde{\mathscr{V}}}(\tilde{u}) = \int_{\mathscr{V}'} \rho_{\mathscr{V}}(\tilde{u} + u')\, \mathrm{d}\mathscr{V}'(u'). \tag{2.7}$$

With this definition, the LES p.d.f. $\rho_{\mathscr{W}}(w)$ should be compared to $\rho_{\tilde{\mathscr{V}}}(\mathrm{L}^{-1}w)$. Thus, for the LES to produce the correct spatial statistics, $\rho_{\mathscr{W}}(w, t)$ and $\rho_{\tilde{\mathscr{V}}}(\mathrm{L}^{-1}w, t)$ must be equal for all times.

To find what will make this true, one needs evolution equations for $\rho_{\mathscr{W}}(w, t)$ and $\rho_{\tilde{\mathscr{V}}}(\mathrm{L}^{-1}w)$. The turbulent system has been defined as a deterministic dynamical system with a random ensemble of initial conditions. Suppose one defines an arbitrary region in the phase-space of the dynamical system. The region then flows through phase-space as each point in the region evolves according to the dynamical system, but the probability associated with the region remains constant. If the probability density function is differentiable, one can then express this conservation of probability as a 'conservation law' in phase space:

$$\frac{\partial}{\partial t}\rho_{\mathscr{V}}(u) + \frac{\partial}{\partial u} \cdot \left(\rho_{\mathscr{V}}(u) \frac{\mathrm{d}u}{\mathrm{d}t} \right) = 0, \tag{2.8}$$

where $\partial/\partial u\cdot$ is a generalized divergence in phase space. One can also regard the p.d.f. as the expected value of a Dirac delta function, and arrive at the same evolution equation by carefully applying the chain rule (Lundgren 1967; Pope 1985). Conservation of probability also holds for the LES p.d.f.,

$$\frac{\partial}{\partial t}\rho_{\mathscr{W}}(w) + \frac{\partial}{\partial w} \cdot \left(\rho_{\mathscr{W}}(w) \frac{\mathrm{d}w}{\mathrm{d}t} \right) = 0, \tag{2.9}$$

but it cannot be applied to the marginal p.d.f., because phase trajectories cross when the real turbulence system is projected into $\tilde{\mathscr{V}}$, so there is no well-defined flow on which to apply the conservation law. The evolution of the marginal p.d.f. $\rho_{\tilde{\mathscr{V}}}(\tilde{u})$ must be obtained by integrating (2.8) over the small scales, that is, $\mathscr{V}'$, the null space of F:

$$\int_{\mathscr{V}'} \frac{\partial}{\partial t}\rho_{\mathscr{V}}(u) + \frac{\partial}{\partial u} \cdot \left(\rho_{\mathscr{V}}(u) \frac{\mathrm{d}u}{\mathrm{d}t} \right) \mathrm{d}\mathscr{V}'(u') = 0. \tag{2.10}$$

Applying the definition of the marginal p.d.f., and decomposing the divergence,

$$\frac{\partial}{\partial t}\rho_{\tilde{\mathscr{V}}}(\tilde{u}) + \int_{\mathscr{V}'} \frac{\partial}{\partial \tilde{u}} \cdot \left(\rho_{\mathscr{V}}(u) \mathrm{P}_{\tilde{\mathscr{V}}} \frac{\mathrm{d}u}{\mathrm{d}t} \right) + \frac{\partial}{\partial u'} \cdot \left(\rho_{\mathscr{V}}(u) \mathrm{P}_{\mathscr{V}'} \frac{\mathrm{d}u}{\mathrm{d}t} \right) \mathrm{d}\mathscr{V}'(u') = 0. \tag{2.11}$$

The second term in the integrand can be eliminated by applying the divergence theorem and requiring that the p.d.f. is zero outside a finite region:

$$\frac{\partial}{\partial t}\rho_{\tilde{\mathscr{V}}}(\tilde{u}) + \int_{\mathscr{V}'} \frac{\partial}{\partial \tilde{u}} \cdot \left(\rho_{\mathscr{V}}(u) \mathrm{P}_{\tilde{\mathscr{V}}} \frac{\mathrm{d}u}{\mathrm{d}t} \right) \mathrm{d}\mathscr{V}'(u') = 0. \tag{2.12}$$

Exchanging the order of the divergence and integration,

$$\frac{\partial}{\partial t}\rho_{\tilde{\mathscr{V}}}(\tilde{u}) + \frac{\partial}{\partial \tilde{u}} \cdot \int_{\mathscr{V}'} \rho_{\mathscr{V}}(u) \mathrm{P}_{\tilde{\mathscr{V}}} \frac{\mathrm{d}u}{\mathrm{d}t}\, \mathrm{d}\mathscr{V}'(u') = 0. \tag{2.13}$$

The remaining integral can be rewritten as a conditional average:

$$\frac{\partial}{\partial t}\rho_{\widetilde{\mathscr{V}}}(\widetilde{u}) + \frac{\partial}{\partial \widetilde{u}} \cdot \left(\rho_{\widetilde{\mathscr{V}}}(\widetilde{u}) \left\langle \mathrm{P}_{\widetilde{\mathscr{V}}} \frac{\mathrm{d}u}{\mathrm{d}t} \middle| \mathrm{P}_{\widetilde{\mathscr{V}}} u = \widetilde{u} \right\rangle \right) = 0. \tag{2.14}$$

Using L to express the result in the LES space,

$$\frac{\partial}{\partial t}\rho_{\widetilde{\mathscr{V}}}(\mathrm{L}^{-1}w) + \frac{\partial}{\partial w} \cdot \left(\rho_{\widetilde{\mathscr{V}}}(\mathrm{L}^{-1}w) \left\langle \mathrm{F} \frac{\mathrm{d}u}{\mathrm{d}t} \middle| \mathrm{F}u = w \right\rangle \right) = 0. \tag{2.15}$$

Comparing (2.15) to (2.9), it is clear that if $\rho_{\mathscr{W}}(w)$ equals $\rho_{\widetilde{\mathscr{V}}}(\mathrm{L}^{-1}w)$ at some initial time, then $\rho_{\mathscr{W}}(w,t)$ equals $\rho_{\widetilde{\mathscr{V}}}(\mathrm{L}^{-1}w,t)$ for all time *if and only if*

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \left\langle \mathrm{F} \frac{\mathrm{d}u}{\mathrm{d}t} \middle| \mathrm{F}u = w \right\rangle. \tag{2.16}$$

Note that this result is independent of the choice of complement space $\widetilde{\mathscr{V}}$. The corollary, which follows from (2.6), is that all ensemble-averaged spatial statistics will be correct if and only if (2.16) holds. To obtain correct space–time correlations, one would need to include past time information in the condition of (2.16). We will consider only single-time conditions here.

There are some subtleties to this result. For example, it may be possible for an individual trajectory to be driven into a region where the density function is zero and the conditional average is undefined. Also, in the usual application of LES, one evolves a *single* field w through time. For the formal result to hold, the conditional average must be over an ensemble in the real system whose p.d.f. matches that of the LES. For a single LES field, the p.d.f. could be matched by a single DNS field, and evolution by the conditional average would result in the single LES following the single DNS.

An important question is what will happen to a single LES that is evolved according to the conditional average over a larger ensemble. In statistically stationary turbulence, if the LES system is ergodic, then we expect statistics gathered from the time average of a single LES will match the LES ensemble averages, and thus statistics of the real system. The question of whether a single LES has these properties needs further exploration. Berkooz (1994) explored these ideas with a simulation of the filtered Lorenz equations and concluded that there were problems recreating the statistics. But his reduction of the Lorenz equations to two dimensions guaranteed a periodic, non-chaotic solution that clearly could not recreate the proper statistics. The procedure needs to be tested on a different model system, or better, on turbulence. There has also been work with evolving an ensemble of LES systems simultaneously (see Carati, Wray & Cabot 1996). The analysis of this subsection suggests that such ensemble techniques may extend the ability of LES to predict spatial statistics.

2.4. *Error-minimizing dynamics*

One can define a mean-square error between the evolution of an LES field $w(t)$ and an instance of the real system $u(t)$ as an instantaneous pointwise measurement on $\partial\widetilde{u}_i/\partial t$,

$$e_i(\boldsymbol{x}) = \frac{\partial w_i}{\partial t}(\boldsymbol{x}) - \frac{\partial \widetilde{u}_i}{\partial t}(\boldsymbol{x}). \tag{2.17}$$

To see how this definition of error addresses the dynamics of LES, write the value of the filtered field at a later time as a Taylor series expansion of the filtered field at

the current time. Then the leading term of the expansion is $\partial\widetilde{u}_i/\partial t$, which suggests that the error of (2.17) is appropriate for measuring accuracy of short-time large-scale dynamics (see Adrian 1990).

The error defined in (2.17) is minimized in an RMS sense when the LES field w evolves as the conditional average of the real field u:

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \left\langle \widetilde{\frac{\mathrm{d}u}{\mathrm{d}t}} \middle| \widetilde{u} = w \right\rangle. \tag{2.18}$$

The relationship between conditional averages, stochastic estimation, and RMS error is outlined in Papoulis (1965) and will not be repeated here.

As discussed above, the most fundamental property of LES is that state information about the turbulent field is lost, so it is reasonable to expect that the lost information limits the accuracy of LES in some way. It was shown in § 2.3 that accurate large-scale statistics can still be obtained. The loss that occurs is in the small-scale statistics that cannot be computed. In the case of dynamical error, the results are not so promising. Although ideal LES may minimize the error, there is nothing to guarantee that the error will be small. This error is fundamental to the LES representation.

Note that the ideal LES may be interpreted as the deterministic component of the evolution of the large scales. Then the mean-square error in the time derivative, which is minimized by the ideal LES, measures the magnitude of the stochastic component of the small-scale contribution to large-scale dynamics.

Measuring error of the large-scale dynamics also provides an *a priori* framework with which to measure the performance of any subgrid model. Even though the limiting error of the ideal model may be impractical to compute, the *a priori* errors of practical models may still be computed so that models may be compared to each other.

2.5. *An ideal model term*

It is customary to write the LES equations as the Navier–Stokes terms operating on the filtered field, plus a subgrid force, as in (1.2). This has been done with the expectation that Navier–Stokes terms capture a significant fraction of the LES behaviour. It has already been shown that ideal LES evolves as the conditional average of $\partial\widetilde{u}_i/\partial t$. However, it is useful to continue the practice of decomposing the equations into Navier–Stokes and model terms, so that an ideal subgrid model may be defined.

The discussion of § 2.2, in which the evolution space for the LES is defined, illuminates a subtlety to using Navier–Stokes terms in the LES equations. In particular, it is clear that the LES evolution equations must be a mapping from $\mathscr{W}$ to $\mathscr{W}$, i.e. mappings within the LES solution space. For non-invertible filters, it is likely that Navier–Stokes terms cannot be exactly represented as such mappings. For example, if the LES solution space is a piecewise-linear continuous function space, then terms with spatial derivatives are not in the space. If the LES solution space is a spherically truncated Fourier space, then quadratic products are not represented in the space. In general, the best one can do is to approximate the Navier–Stokes operator by a mapping within the LES solution space. This observation is almost trivial, except that the approximation is not unique. Details of the terms will depend on the choice of filter, on the choice of numerical method, and on the choice of mapping back into $\mathscr{W}$. This implies that the subgrid force is not unique, and that the model should depend on everything just mentioned.

Note that this problem does not occur in direct simulations because there are no

important subspaces that arise from the filter. In DNS, one has the possibility of converging to a solution through grid refinement. With large-eddy simulation, the filter defines an LES solution space that is in general not closed under the Navier–Stokes operator. Even explicit filtering and grid refinement cannot circumvent the issue.

Thus, we write the LES evolution equations as

$$\frac{\partial \boldsymbol{w}}{\partial t} = \mathbf{NS}(w) + \boldsymbol{m}(w), \tag{2.19}$$

where **NS** and $\boldsymbol{m}$ are both mappings from the LES function space to itself, and **NS** is some non-unique approximation to the Navier–Stokes time-derivative. Note that depending on how **NS** is formulated, it may or may not be Galilean invariant. However, provided only that the filter operator is Galilean invariant, $\boldsymbol{m}$ as defined below will be such that the right-hand side of (2.19) will be Galilean invariant.

The remainder of this paper will refer to estimation of $\boldsymbol{M}(u)$, not $\partial\widetilde{\boldsymbol{u}}/\partial t$, and LES subgrid models $\boldsymbol{m}(w)$ are regarded as approximations to the ideal subgrid model

$$\boldsymbol{m}(w) = \langle M(u) | \widetilde{u} = w \rangle, \tag{2.20}$$

where the real subgrid force $\boldsymbol{M}(u)$ must be consistent with the LES equations (2.19), i.e.

$$\boldsymbol{M}(u) = \widetilde{\frac{\partial \boldsymbol{u}}{\partial t}} - \mathbf{NS}(\widetilde{u}). \tag{2.21}$$

The dynamical error defined in § 2.4 can be written as

$$\boldsymbol{e}(\boldsymbol{x}) = \boldsymbol{M}(\boldsymbol{x}) - \boldsymbol{m}(\boldsymbol{x}). \tag{2.22}$$

3. Optimal formulations

The ideal LES defined above has all the features one could wish for in an LES model, but unfortunately it cannot be computed directly. The statistical information embodied in the conditional average of ideal LES is so tremendous that it is unlikely that the ideal model for a given flow or filter could ever be found exactly. Still, one may use approximation techniques to investigate the nature of the ideal model. The term *optimal LES* will be used to describe formulations that most closely approximate ideal LES within some class. By systematically developing approximations to the ideal model, one may determine bounds on the minimum error, which is associated with the stochastic nature of the subgrid force. Also, one can regard the task of subgrid modelling as a problem of approximating the conditional average that defines the ideal model.

One strategy for constructing such approximations is to use a stochastic estimate of the LES time derivative or the LES model term, as suggested Adrian (1990). Given a functional form for the estimate, the stochastic estimate minimizes the mean-square error between the quantity being estimated (the subgrid force in this case) and the estimate. Stochastic estimation is a well-established method of approximating conditional averages (see Adrian 1977; Adrian & Moin 1988; Adrian *et al.* 1989), and the technique has already been applied to optimal subgrid models by Berkooz (1994), who tested the the procedure on the Lorenz equations.

As discussed in § 2.5, the real subgrid force $M_i(\boldsymbol{x})$ is the term to be estimated. The approximation is based on the convolution of an estimation kernel K_{ij} with velocity

event data at N points:

$$m_i(\boldsymbol{x}) = \int K_{ij}(\boldsymbol{x}, \xi_1, \ldots, \xi_N) E_j(w; \xi_1, \ldots, \xi_N)\, \mathrm{d}\xi_1 \ldots \mathrm{d}\xi_N, \tag{3.1}$$

where E_j is an event vector that contains (possibly nonlinear) functions of $\boldsymbol{w}$ at N different points. Note that one can choose

$$\vec{E}(\xi_1, \xi_2, \ldots) = (1,\ w_i(\xi_1),\ w_j(\xi_1) w_k(\xi_2), \ldots) \tag{3.2}$$

to recover the expansion

$$m_i(\boldsymbol{x}) = A_i(\boldsymbol{x}) + \int B_{ij}(\boldsymbol{x}, \xi_1) w_j(\xi_1)\, \mathrm{d}\xi_1 + \int C_{ijk}(\boldsymbol{x}, \xi_1, \xi_2) w_j(\xi_1) w_k(\xi_2)\, \mathrm{d}\xi_1 \mathrm{d}\xi_2 + \cdots, \tag{3.3}$$

which is complete in the sense that any functional $m_i(\boldsymbol{x}, w)$ has such a representation, although the expansion may be infinite. One can also choose event data that include higher-order terms at a single point, such as

$$\vec{E}(\xi_1) = (1,\ w_i(\xi_1),\ w_j(\xi_1) w_k(\xi_1), \ldots), \tag{3.4}$$

to form a higher-order expansion using single-point event data. Such an expansion is not complete, but one may still improve the accuracy of an N-point estimate by including higher-order terms. This framework enables one to speak of linear one-point estimates, quadratic one-point estimates, quadratic two-point estimates, etc.

By the statistical orthogonality principle (see Papoulis 1965), the RMS error between M_i and the estimate m_i is minimized when the event data are uncorrelated with the error $e_i = M_i - m_i$:

$$\langle (M_i(\boldsymbol{x}) - m_i(\boldsymbol{x})) E_k(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_N) \rangle = 0. \tag{3.5}$$

Substituting (3.1), the optimal kernel is governed by

$$\langle M_i(\boldsymbol{x}) E_k(\eta_1, \ldots, \eta_N) \rangle = \int K_{ij}(\boldsymbol{x}, \xi_1, \ldots, \xi_N) \langle E_j(\xi_1, \ldots, \xi_N) E_k(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_N) \rangle\, \mathrm{d}\xi_1 \cdots \mathrm{d}\xi_N. \tag{3.6}$$

Note that in minimizing the mean-square error between the true subgrid force M_i and its estimate m_i, one also minimizes the mean-square difference between $\langle M_i | \widetilde{u} = w \rangle$ and m_i. This allows us to interpret the estimate as an approximation of the conditional average.

To solve for the kernel K_{ij} of the optimal N-point estimate, one must know the $(N+1)$-point correlation of the real subgrid term M_i and the event data, and one must know the $(2N)$-point correlation of the event data with itself. The solution will be unique provided the event data are linearly independent. Further, a unique solution must correspond to a minimum error, not a maximum, because the domain from which K_{ij} is chosen is unbounded while the error associated with K_{ij} has a lower bound of zero.

The mean quadratic error of the optimal N-point estimate is

$$\langle e_i(\boldsymbol{x}) e_j(\boldsymbol{x}') \rangle = \langle M_i(\boldsymbol{x}) M_j(\boldsymbol{x}') \rangle - \int K_{jk}(\boldsymbol{x}', \xi_1, \ldots, \xi_N) \langle M_i(\boldsymbol{x}) E_k(\xi_1, \ldots, \xi_N) \rangle\, \mathrm{d}\xi_1 \ldots\, \mathrm{d}\xi_N. \tag{3.7}$$

Computing the error of the optimal estimate thus requires an additional two-point correlation of the real subgrid term with itself.

The subgrid models based on these optimal N-point estimates have remarkable *a priori* properties. One can show that the correlation of the model term m_i with any event data is the same as the correlation of the exact subgrid term M_i with the event data:

$$\langle m_i(\boldsymbol{x}) E_j(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N) \rangle = \langle M_i(\boldsymbol{x}) E_j(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N) \rangle. \tag{3.8}$$

Since the only contribution of the small scales to $\partial \widetilde{u}_i / \partial t$ is contained in the subgrid term m_i, an optimal subgrid model also has the *a priori* property that the correlation of $\partial \widetilde{u}_i / \partial t$ with the event data is predicted exactly. The event data of a linear one-point estimate includes $\widetilde{u}_i$, so that the subgrid model will accurately predict all kinetic energy transfer in an *a priori* test. The emphasis is that this is an *a priori* result because the ensemble averages of (3.8) are averages over the real system. For a discussion of *a priori* testing, see Meneveau (1994).

Another interesting property of these estimates is that they preserve Galilean invariance. If **NS** in (2.19) is Galilean invariant, then an estimate that includes a linear dependence on $\widetilde{u}_i$ is sufficient to assure invariance of m_i. If on the other hand **NS** is not invariant, then a one-point quadratic estimate is sufficient to assure invariance of the LES equations.

4. Application to isotropic turbulence

We applied the quadratic one-point estimation to the case of isotropic turbulence with spherical Fourier truncation filters, using DNS data. The DNS data are from a well-resolved 256^3 simulation of forced isotropic turbulence. Except when noted otherwise, numerical results are quoted using the arbitrary normalization of the DNS. In these units, the minimum and maximum resolved wavenumbers have magnitude $k = 1$ and $k = 121$, and the viscosity is $\nu = 1/150$. Forcing was achieved with a negative viscosity on modes with wavenumber magnitude $k \leqslant 3$. The simulation was pseudo-spectral: nonlinear terms were dealiased with a combination of masks and random phase shifts. Nonlinear terms were integrated with a second-order Runga–Kutta scheme, and viscous terms were advanced with an integrating factor. For further details see Rogallo (1981). In our arbitrary units, the resulting statistics of the simulation are kinetic energy $q^2/2 = 41.1$, dissipation $\epsilon = 62.9$, Taylor microscale $\lambda_g = 0.209$, and Reynolds number $R_\lambda = 164$. The three-dimensional energy spectrum is shown in figure 1.

The subgrid term to be estimated is written

$$M_i(\boldsymbol{x}) = \mathscr{P}\left(\frac{\partial}{\partial x_j} \tau_{ij}(\boldsymbol{x}) \right), \tag{4.1}$$

where $\mathscr{P}$ denotes the divergence-free projection into the band-limited Fourier space. Need for the projection arises from the nonlinear filtered product in τ_{ij}, which would otherwise cause **M** to be outside the LES space (see § 2.5). The Fourier coefficient of **M** is

$$\hat{M}_i(\boldsymbol{k}) = \mathrm{i} k_j \hat{\tau}_{ij}(\boldsymbol{k}) - \mathrm{i} \frac{k_i k_j k_l}{k^2} \hat{\tau}_{jl}(\boldsymbol{k}). \tag{4.2}$$

The subgrid term (4.1) was approximated with a *linear one-point estimate,* a *quadratic one-point estimate*, and a composite estimate of up to 46 terms, including up to cubics.

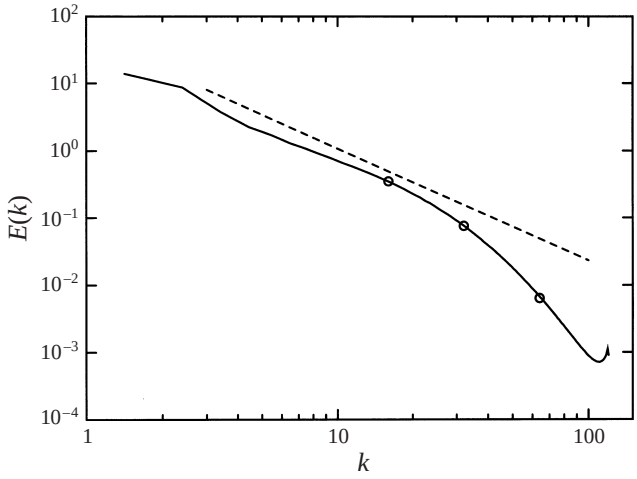

FIGURE 1. The three-dimensional energy spectrum E(k) (———), a $-\frac{5}{3}$ power law (– – – –), and locations (○) of the three sharp Fourier cutoff filters.

The event data for the linear and quadratic estimates are

$$\vec{E}(w,\xi_1) = \begin{pmatrix} w_1(\xi_1) \\ w_2(\xi_1) \\ w_3(\xi_1) \end{pmatrix} \quad \text{and} \quad \vec{E}(w,\xi_1) = \begin{pmatrix} w_1(\xi_1) \\ w_2(\xi_1) \\ w_3(\xi_1) \\ w_1(\xi_1)w_1(\xi_1) \\ w_1(\xi_1)w_2(\xi_1) \\ w_1(\xi_1)w_3(\xi_1) \\ w_2(\xi_1)w_2(\xi_1) \\ w_2(\xi_1)w_3(\xi_1) \\ w_3(\xi_1)w_3(\xi_1) \end{pmatrix}, \tag{4.3}$$

respectively, and the 46-term model is described in § 4.4. Note that the quadratic one-point estimate has a form general enough to subsume the incompressible Navier–Stokes equations. This estimation was done for three sharp Fourier cutoff filters with cutoff wavenumbers $k_c = 64$, $k_c = 32$, and $k_c = 16$, which are marked on the energy spectrum in figure 1.

While the DNS used to gather the estimation data is a model for isotropic turbulence, it is not exactly isotropic. In particular, it is computed in a cubical domain with periodic boundary conditions. Therefore, the estimates were performed in two ways, first by assuming the DNS data are not isotropic and then by assuming they are isotropic. By admitting any anisotropies that may have arisen from the periodic representation of the simulation, we were able to ensure that details of the estimation were not consequences of an isotropic assumption. Analysis of the two estimates and their associated errors led us to conclude that the assumption of strict isotropy was, in fact, a good one to make. In this paper only the isotropic results are presented, although details of both estimation procedures are provided.

4.1. *Data requirements*

To compute the optimal linear one-point estimate and its associated error, the following statistical quantities were needed:

$$\langle \widetilde{u}_i(\boldsymbol{\xi}_1)\widetilde{u}_j(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle,\ \langle \tau_{ij}(\boldsymbol{\xi}_1)\widetilde{u}_k(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle,\ \langle \tau_{ij}(\boldsymbol{\xi}_1)\tau_{kl}(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle.$$

These are the correlations of the event data with itself, of the event data with the subgrid stress, and of the subgrid stress with itself. The optimal quadratic one-point estimate required the following additional correlations:

$$\langle \widetilde{u}_i(\boldsymbol{\xi}_1)\widetilde{u}_j(\boldsymbol{\xi}_1)\widetilde{u}_k(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle,\quad \langle \widetilde{u}_i(\boldsymbol{\xi}_1)\widetilde{u}_j(\boldsymbol{\xi}_1)\widetilde{u}_k(\boldsymbol{\xi}_1+\boldsymbol{x})\widetilde{u}_l(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle,$$

$$\langle \widetilde{u}_i(\boldsymbol{\xi}_1)\widetilde{u}_j(\boldsymbol{\xi}_1)\tau_{kl}(\boldsymbol{\xi}_1+\boldsymbol{x})\rangle.$$

In general, correlations of the event vectors with the model term M_i (or equivalently τ_{ij}), and with themselves are required. A set of these correlations was required for each filter.

These statistics were gathered from 15 DNS fields, each separated in time by approximately one-half of an eddy-turnover time, where the eddy-turnover timescale is defined as ϵ/q^2. To increase the statistical sample base, there was an average over directions of spatial homogeneity. There was also an average over the 48 reflections and 90° rotations of each field, all of which were assumed to occur with equal probability. The statistics were not gathered with a full set of isotropic symmetries so that anisotropic estimates could be computed. However, statistics for the isotropic estimates were computed using standard shell-averaging techniques.

Convergence of the statistics was checked by comparing quantities averaged over 7 and 14 fields. The most sensitive check for convergence was to compare individual Fourier modes of the correlation tensors, a test that is relevant for the anisotropic estimate. By this method, several of the correlation tensors appeared not to have converged. For example, the third-order correlation of filtered velocity with subgrid stress had not converged at each Fourier mode. A less sensitive test for convergence was to make comparisons of quantities that had been fit to an isotropic form through a shell-averaging procedure. In this sense the isotropic estimates were well-converged, and the anisotropic estimates appeared to be converging to the isotropic ones.

4.2. *Exact properties of the subgrid scales*

Next, several exact features of the subgrid term M_i are presented. The spectra of M_i and $\partial\widetilde{u}_i/\partial t$ are shown in figure 2. The spectra of the subgrid terms are characterized by a quadratic growth with k followed by a sharp increase near the cutoff. It is evident that the subgrid term is only significant near the cutoff.

Because M_i is a term contributing directly to $\partial\widetilde{u}_i/\partial t$, the correlation $\langle \widetilde{u}_i(\boldsymbol{\xi})M_i(\boldsymbol{\xi}+\boldsymbol{x})\rangle$ represents the average contribution of M_i to the evolution of large-scale kinetic energy. This correlation integrated in wavespace over wavenumbers from magnitude 0 to magnitude k is shown in figure 3. This represents the average transfer of kinetic energy from modes with wavenumber magnitude less than k to the subgrid. The Fourier transform of the correlation, normalized by the magnitudes of M_i and $\widetilde{u}_i$ is shown in figure 4. This shows the average alignment of $\boldsymbol{M}$ and $\widetilde{\boldsymbol{u}}$, which indicates the direction of energy exchange between a particular resolved scale and the combined subgrid scales. It is clear from these figures that there is a small net backscatter of energy to low wavenumbers with the two narrowest filters, but no backscatter of energy with the coarsest filter. Recall from the previous section that these quantities will be *a priori* predicted exactly with the optimal estimates described in this section.

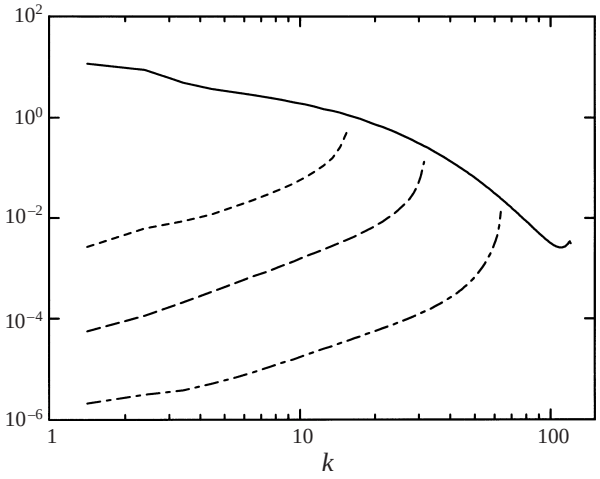

FIGURE 2. Spectra of $\partial\widetilde{u}_i/\partial t$ (———) and the subgrid force terms $\mathscr{P}((\partial/\partial x_j)(\widetilde{u_i u_j} - \widetilde{u}_i\widetilde{u}_j))$ for sharp Fourier cutoff filters at cutoffs $k_c = 16$ (– – – –), $k_c = 32$ (— — —), and $k_c = 64$ (– · – · –).

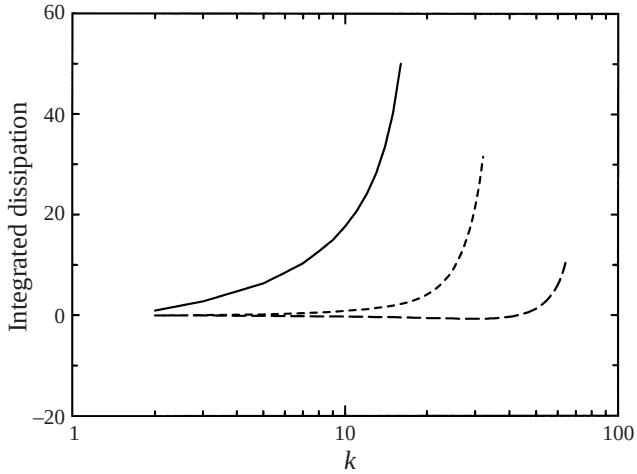

FIGURE 3. Average transfer of energy $\int_0^k \langle\hat{\widetilde{u}}_i(\boldsymbol{k}')\hat{M}_i^*(\boldsymbol{k}')\rangle\,\mathrm{d}k'$ from the largest scales to the subgrid for cutoff filters at widths $k_c = 16$ (———), $k_c = 32$ (– – – –), and $k_c = 64$ (— — —).

4.3. *Homogeneous one-point estimation*

For a one-point estimate in a homogeneous system, the estimation is local in Fourier space, i.e.

$$\hat{m}_i(\boldsymbol{k}) = \hat{K}_{ij}(\boldsymbol{k})\hat{E}_j(\boldsymbol{k}). \tag{4.4}$$

The optimal kernel $\hat{K}_{ij}$ can be determined from

$$\langle\hat{M}_i(\boldsymbol{k})\hat{E}_j^*(\boldsymbol{k})\rangle = \hat{K}_{il}(\boldsymbol{k})\langle\hat{E}_l(\boldsymbol{k})\hat{E}_j^*(\boldsymbol{k})\rangle. \tag{4.5}$$

This is a simple linear system. For the linear estimate the matrices are all 3×3, and for the quadratic estimate the matrices are 3×9 and 9×9. Because $\hat{w}_1$, $\hat{w}_2$, and $\hat{w}_3$ are linearly related through continuity, the correlation of the event data with itself is a singular matrix and the optimal kernel $\hat{K}_{ij}$ is not unique, although the estimate

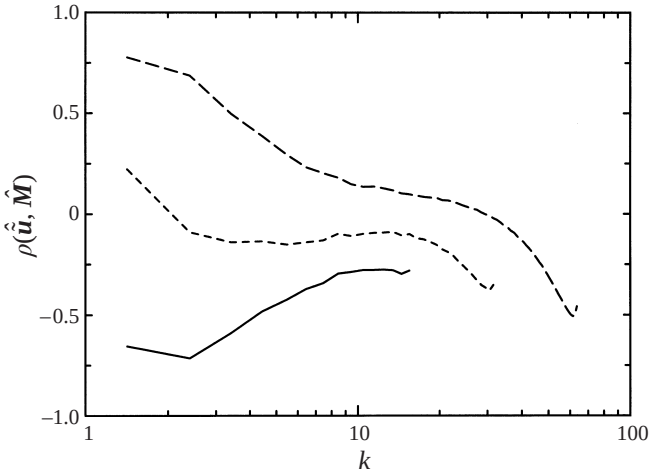

FIGURE 4. Correlation $\rho(\hat{\bar{u}}, \hat{M}) = Re\,[\langle \hat{\bar{u}}_i \hat{M}_i^* \rangle]/\langle \hat{M}_i \hat{M}_i^* \rangle^{1/2} / \langle \hat{\bar{u}}_i \hat{u}_i^* \rangle^{1/2}$ of the filtered velocity with the real subgrid term for cutoff filters at widths $k_c = 16$ (———), $k_c = 32$ (– – – –), and $k_c = 64$ (— — —). Note that the value $\rho = 1$ corresponds to a complete alignment between $\hat{M}_i$ and $\hat{\bar{u}}_i$.

(4.5) and the error (4.6) are. The error between the one-point optimal estimate and the real subgrid term is

$$\langle \hat{e}_i(\boldsymbol{k}) \hat{e}_j^*(\boldsymbol{k}) \rangle = \langle \hat{M}_i(\boldsymbol{k}) \hat{M}_j^*(\boldsymbol{k}) \rangle - \hat{K}_{il}(\boldsymbol{k}) \langle \hat{M}_j^*(\boldsymbol{k}) \hat{E}_l(\boldsymbol{k}) \rangle. \tag{4.6}$$

4.4. *Isotropic estimation equations*

For isotropic turbulence, the homogeneous one-point estimate simplifies considerably. The most general one-point estimate can be expressed as

$$\hat{m}_i(\boldsymbol{k}) = C^1(k) \hat{E}_i^1(\boldsymbol{k}) + C^2(k) \hat{E}_i^2(\boldsymbol{k}) + \cdots + C^N(k) \hat{E}_i^N(\boldsymbol{k}), \tag{4.7}$$

where the event vectors $\boldsymbol{E}^j$ must be in the space $\mathscr{W}$ in which the LES is performed. Thus, when a Fourier cutoff filter is used, the event vectors $\boldsymbol{E}^j$ are divergence-free and band-limited. The scalar coefficients C^j are simple functions of k, and for each k they are governed by the linear system

$$\begin{pmatrix} \langle \hat{E}_i^{1*} \hat{E}_i^1 \rangle & \langle \hat{E}_i^{1*} \hat{E}_i^2 \rangle & \cdots & \langle \hat{E}_i^{1*} \hat{E}_i^N \rangle \\ \langle \hat{E}_i^{2*} \hat{E}_i^1 \rangle & \langle \hat{E}_i^{2*} \hat{E}_i^2 \rangle & \cdots & \langle \hat{E}_i^{2*} \hat{E}_i^N \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \hat{E}_i^{N*} \hat{E}_i^1 \rangle & \langle \hat{E}_i^{N*} \hat{E}_i^2 \rangle & \cdots & \langle \hat{E}_i^{N*} \hat{E}_i^N \rangle \end{pmatrix} \begin{pmatrix} C^1 \\ C^2 \\ \vdots \\ C^N \end{pmatrix} = \begin{pmatrix} \langle \hat{E}_i^{1*} \hat{M}_i \rangle \\ \langle \hat{E}_i^{2*} \hat{M}_i \rangle \\ \vdots \\ \langle \hat{E}_i^{N*} \hat{M}_i \rangle \end{pmatrix}. \tag{4.8}$$

Furthermore, the error is given by

$$\langle \hat{e}_i^*(\boldsymbol{k}) \hat{e}_i(\boldsymbol{k}) \rangle = \langle \hat{M}_i^*(\boldsymbol{k}) \hat{M}_i(\boldsymbol{k}) \rangle - \sum_{j=1}^{N} C^j \langle \hat{E}_i^{j*}(\boldsymbol{k}) \hat{M}_i(\boldsymbol{k}) \rangle. \tag{4.9}$$

For the linear estimate, one uses $\hat{E}_i^1 = \hat{w}_i$, and for the quadratic estimate, one also includes $\hat{E}_i^2 = \mathscr{P}\big(\mathrm{i}k_j \widehat{w_i w_j}\big)$, where $\mathscr{P}$ is a solenoidal projection into the LES space. Note that the homogeneous anisotropic estimates outlined in § 4.3 appear to converge to the isotropic estimates shown above. Consequently, all results in this paper will be shown only for the isotropic estimates.

Event vectors	Event scalars
w_i	1
$\mathscr{P}(\partial/\partial x_j)(w_i^{16}w_j^{16})$	$(S_{ij}S_{ij})^{1/2}$
$\mathscr{P}(\partial/\partial x_j)(w_i^{14}w_j^{14})$	$S_{ij}S_{ij}$
$\mathscr{P}(\partial/\partial x_j)(w_i^{12}w_j^{12})$	$(w_iw_i)^{1/2}$
$\mathscr{P}(\partial/\partial x_j)(w_i^{10}w_j^{10})$	w_iw_i
$\mathscr{P}(\partial/\partial x_j)(w_i^{8}w_j^{8})$	$(w_iw_i)^2$
$\mathscr{P}(\partial/\partial x_j)(w_i\mathscr{P}(\partial/\partial x_k)w_jw_k)$	
$\mathscr{P}(\partial/\partial x_j)(w_i\mathscr{P}(\partial/\partial x_j)w_iw_k)$	

TABLE 1. Table of vectors and scalars used to construct estimation events. An estimation event is formed by multiplying a vector by a scalar. Here, $\boldsymbol{w}^N$ denotes the turbulent velocity field that has been filtered with a cutoff $k_c = N$, and S_{ij} is the filtered strain-rate magnitude defined in (1.5).

Most of the results in this paper will be provided only for the linear and quadratic estimates. However, more complicated estimates were also explored for the case of $k_c = 16$. The most comprehensive estimate was a model with 46 terms, which were generated by taking a basic set of 8 event vectors, scaling each event vector by one of 6 scalar fields (including the constant 1), and removing terms that were not independent, as determined by the fact that their inclusion made the correlation matrix in (4.8) singular.

The event vectors and scaling vectors are listed in table 1. Included are quadratic terms computed from velocities that have been filtered with various Fourier cutoffs including $k_c = 16$, the LES cutoff under consideration, and coarser filters with cutoffs as small as $k_c = 8$. These coarse-filtered quadratic terms are included because they subsume the generalized scale-similarity model as proposed by Liu *et al.* (1994). Also, by using several filter widths, we include some of the information in the general two-point quadratic estimate (3.3). Also included are two cubic terms that generalize the cubic term suggested by the RNG analysis of Zhou & Vahala (1993). Of the scalars, the most notable is the magnitude of the strain-rate tensor, which is included to allow the Smagorinsky model to be subsumed. Of the 48 terms that can be formed from a product of one of the vector events and one of the scalars, two were found to be statistically dependent on the other 46. They were the two cubic terms scaled by $(w_iw_i)^2$.

4.5. *Properties of the optimal estimates*

The optimal linear isotropic estimate can be written as a k-dependent eddy viscosity:

$$\hat{m}_i(\boldsymbol{k}) = -k^2\hat{\nu}_T(k)\hat{w}_j(\boldsymbol{k}), \tag{4.10}$$

where

$$\hat{\nu}_T(k) = \frac{-\langle\hat{\widetilde{u}}_i^*(\boldsymbol{k})\hat{M}_i(\boldsymbol{k})\rangle}{k^2\langle\hat{\widetilde{u}}_j^*(\boldsymbol{k})\hat{\widetilde{u}}_j(\boldsymbol{k})\rangle}. \tag{4.11}$$

The optimal eddy viscosity for the three cutoffs is normalized by $[\mathrm{E}(k_c)/k_c]^{1/2}$ and shown in figure 5.

The optimal eddy viscosities all turn up sharply near the cutoff as predicted by Kraichnan (1976) and observed by Domaradzki *et al.* (1987) and Lesieur & Rogallo (1989). Also note that the eddy viscosities are nearly zero, and in fact go negative far

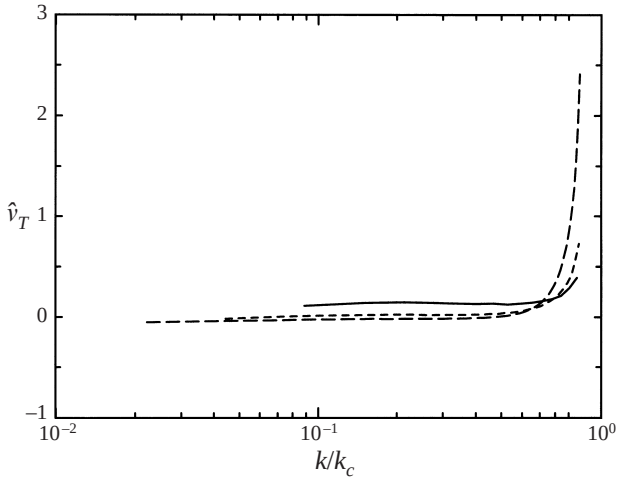

FIGURE 5. Optimal eddy viscosities $\hat{\nu}_T$ normalized by $[\mathrm{E}(k_c)/k_c]^{1/2}$ for cutoff filters at widths $k_c = 16$ (———), $k_c = 32$ (– – – –), and $k_c = 64$ (— — —).

from the cutoff for the two highest wavenumber cutoffs. These cases are consistent with the results of Domaradzki *et al.* (1987), who also obtained slightly negative plateau values when the cutoffs were in a dissipative range. It has been suggested by Lesieur (1998) that when the cutoff is located where the spectrum has a significantly steeper slope than the $k^{-5/3}$ of the Kolmogorov inertial range, then the plateau level of the eddy viscosity should go to zero (see figure 1). The $k_c = 16$ case has behaviour that is more representative of a high Reynolds number LES with the cutoff in the inertial range. Results of EDQNM suggest that for cutoff in the inertial range, the eddy viscosity has a plateau value of 0.267 (see Chollet & Lesieur 1981), which is the basis of the structure function model (Métais & Lesieur 1992). A plateau value of 0.09 was observed by Lesieur & Rogallo (1989), and in our case of $k_c = 16$ the plateau value is approximately 0.13, which is consistent with the other results.

The optimal quadratic isotropic estimate includes one additional term, and can be written so that the k-dependent coefficients multiply terms of order one. The normalization is achieved by multiplying $C^j(k)$ by $-\langle \hat{E}_i^j(\boldsymbol{k})\hat{E}_i^{j*}(\boldsymbol{k})\rangle^{1/2}$ (see § 4.4). These coefficients are shown in figure 6. The quadratic coefficients are characterized by a plateau that is nearly zero, with the level tending to zero as k_c increases into the dissipative range (see figure 1). For the $k_c = 16$ case, which is most representative of a high Reynolds number LES, the quadratic coefficient is at least a factor of 6 smaller than the linear term in the plateau region. The quadratic coefficients increase sharply near the cutoff, but are still 2 or 3 times smaller than the linear ones, suggesting that the linear coefficient is most significant.

Another measure of the significance of the quadratic and other higher-order estimates is the mean-square difference between the linear and higher-order estimates, which is shown in figure 7. It has been normalized by the magnitude of the linear estimate at the cutoff, which is where model performance is most important (see figure 8). The difference between the estimates is insignificant, except near the cutoff, and there the mean-square difference is at most 15% of the linear estimate for the quadratic, but is as much as 70% of the linear estimate for the 46-term estimate.

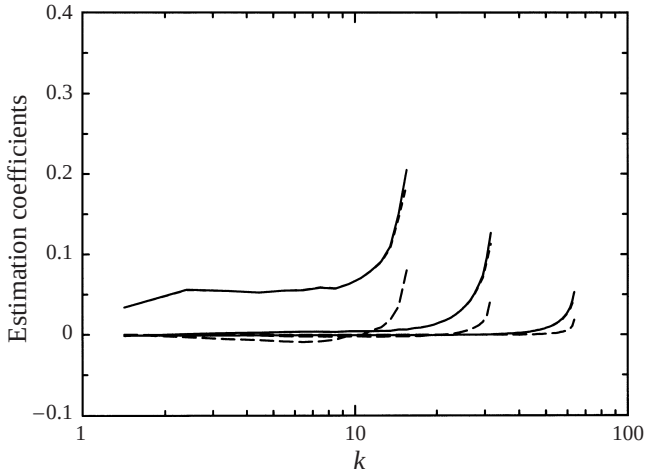

FIGURE 6. Optimal estimation coefficients for three cutoff filters, normalized to multiply quantities of order one. Shown are the linear coefficient from the linear estimate (———), the linear coefficient from the quadratic estimate (– – – –), and the quadratic coefficient from the quadratic estimate (— — —).

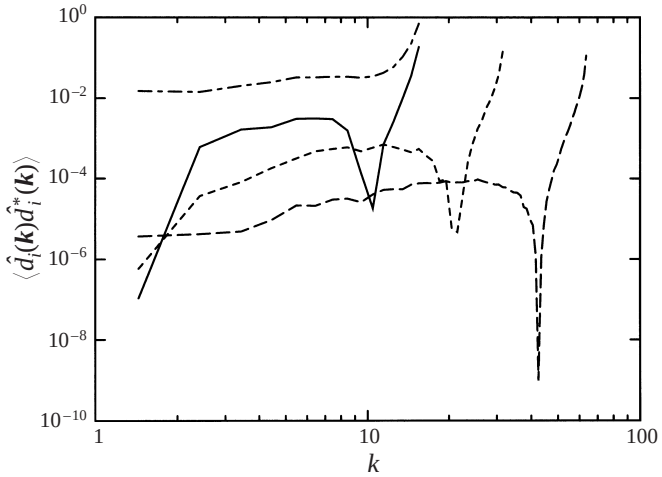

FIGURE 7. Spectral density of the difference d_i between the optimal linear and quadratic estimates, normalized by the mean-square value of the linear estimate at the cutoff. Sharp Fourier cutoffs are located at $k_c = 16$ (———), $k_c = 32$ (– – – –), and $k_c = 64$ (— — —). Also shown is the difference between the linear and 46-term estimate at $k_c = 16$ (– · – · –).

4.6. *Error of the optimal estimates*

Mean-square error between the real and estimated subgrid force is shown in figure 8. Note that the error has been normalized by the spectrum of the real subgrid force M_i. This normalization is used to clearly show how well a particular model can represent the real subgrid term. It would also be sensible to normalize errors by the magnitude of $\partial\widetilde{u}_i/\partial t$, which would show the expected impact of the errors on the large-scale dynamics. But throughout much of the spectrum, the subgrid term is much smaller than $\partial\widetilde{u}_i/\partial t$, which would make it difficult to differentiate between the errors

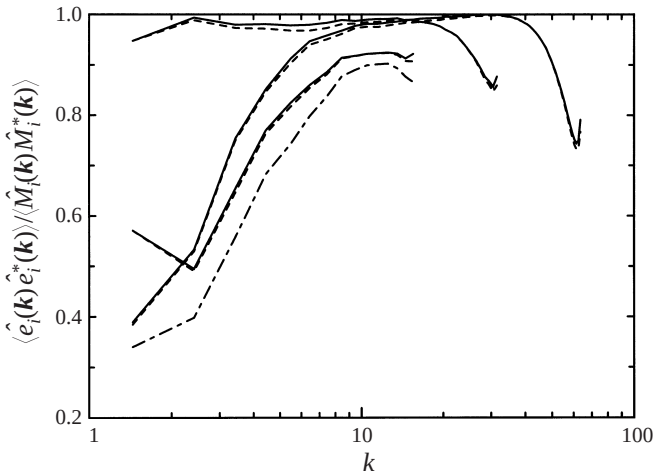

FIGURE 8. Spectral densities of the estimation errors for the linear (———), quadratic (– – – –), and 46-term (– · – · –) estimates of the subgrid term at three different sharp cutoff widths, normalized by the spectral density of the real subgrid term.

of different models. The error is thus normalized by the spectrum of the real subgrid term and the reader is referred to figure 2 to understand its contribution to $\partial\widetilde{u}_i/\partial t$.

The most striking feature of the error is that it is large. This means that each of the optimal estimates that we computed is only able to represent a small fraction of the true subgrid force. Although the error is uniformly large, the magnitude of the subgrid term is much smaller than the magnitude of $\partial\widetilde{u}_i/\partial t$ at low wavenumbers (see figure 2), so that the impact of the error on the accuracy of the LES dynamics is most important near the filter cutoff.

Recall from § 2.4 that the limiting error associated with the ideal model is a measure of the stochastic nature of the subgrid, and quantifies the fundamental limits on accuracy that are inherent to LES modelling. Obviously one would like to know the magnitude of this limiting error. Such knowledge, for example, would allow one to sensibly assess whether a particular model can be improved. Unfortunately it may be impossible to directly compute the fundamental error of the ideal model, which leaves only the possibility of deducing its magnitude from the errors of optimal estimates. It is known that the mean-square errors of the optimal estimates (shown in figure 8) are equal to the fundamental error of the ideal model, plus the error committed in approximating the conditional average, which again is not known. The error of an optimal estimate also provides an upper bound on the fundamental error. As the accuracy of the stochastic estimates is improved (e.g. by including more terms), this bound will be refined, and the error we are measuring will approach the fundamental error of the ideal model.

It has been previously established that stochastic estimates can closely approximate conditional averages of turbulent quantities in isotropic turbulence (see Adrian *et al.* 1989). Further, the error of the linear model is only slightly higher than the error of the quadratic model. In the case of filter cutoff $k_c = 16$, a variety of higher-order models formed from the event vectors and scalars listed in table 1 were used. The error from the best of these, which is the 46-term model discussed in table 1 is also plotted in figure 8. This 46-term model yields a few percent smaller error than the linear model, but there are no terms that contribute particularly strongly to this

improvement. Each nonlinear term on its own yields error improvements on the order of that provided by the quadratic model. Thus it is the accumulated contribution of all 46 terms that yields the reduced error.

These observations lead us to hypothesize that in isotropic turbulence with Fourier cutoff filters, the estimation errors of the ideal model are nearly as large as the errors shown in figure 8. The implication is that the subgrid force is mostly stochastic, and further, that the linear model possesses the dominant characteristics of the ideal model. Of course the validity of this hypothesis depends on how well our estimates approximate the conditional average, and we cannot preclude the possibility that a broader class of multi-point estimation terms would have a smaller error and establish a different bound on the error of the ideal model. What is clear is that model terms of the form used in many practical subgrid models (e.g. Smagorinsky, scale similarity, RNG-based, dynamic Smagorinsky and structure function) do not significantly improve on the simple linear model.

There are several other interesting features of the error shown in figure 8. The error appears to decrease near the cutoff, except for the case of cutoff $k_c = 16$, where it is uniformly smaller. The error also appears to decrease at the lowest wavenumbers, except for the case of cutoff $k_c = 32$. To explain these features, recall that the optimal one-point estimates can exactly represent average energy transfer between the resolved scales and the subgrid scales (the average energy transfer appears in the left-hand side of (4.5)). This suggests that scales exchanging energy with the subgrid will have smaller error than scales not exchanging energy with the subgrid. Indeed, figure 4 shows that there is a point of no average alignment (no average energy exchange) between wavenumbers $k = 2$ and $k = 3$ for cutoff $k_c = 32$, and a point of no average alignment between wavenumbers $k = 30$ and $k = 40$ for cutoff $k_c = 64$. These points correspond to peaks in the error shown in figure 8. This suggests that the error should be normalized by the part of the subgrid that is statistically orthogonal to the velocity, a quantity that is exactly the error of the linear estimate.

The quantity

$$\frac{\left\langle \hat{e}_i(\boldsymbol{k})\hat{e}_i^*(\boldsymbol{k})\right\rangle}{\left\langle \hat{M}_i(\boldsymbol{k})\hat{M}_i^*(\boldsymbol{k})\right\rangle - \left\langle \hat{M}_i(\boldsymbol{k})\hat{\bar{u}}_i^*(\boldsymbol{k})\right\rangle^2 / \left\langle \hat{\bar{u}}_j(\boldsymbol{k})\hat{\bar{u}}_j^*(\boldsymbol{k})\right\rangle} \tag{4.12}$$

is shown in figure 9 for the optimal quadratic and 46-term estimate. It is the estimation error normalized by the part of $\hat{M}$ that is statistically orthogonal to $\hat{\bar{\boldsymbol{u}}}$. It is also the ratio of the higher-order to linear errors of figure 8. Note that the low-wavenumber data shown in figure 9 are characterized by some statistical noise, although the data from just before the peak through the cutoffs are well-converged. For the linear estimate the quantity is identically 1, and for the quadratic estimate it is nearly 1. The renormalized error of the 46-term model ranges from 94% to 97% in the high-wavenumber range where the subgrid term is of any importance. This again suggests how difficult it is to significantly improve on the linear, purely dissipative linear model.

It has already been shown that the optimal estimates can correctly model average energy exchanges. We have also hypothesized that the errors in the higher-order estimates are representative of those for the conditional average. If this is so, then this new normalization of the error suggests that, except for the energy exchange, the dynamics of the subgrid term are mostly stochastic, that is, the subgrid term may be essentially a random fluctuation about a mean energy exchange. If so, the best subgrid model may be able to do little more than estimate average dissipation. Note

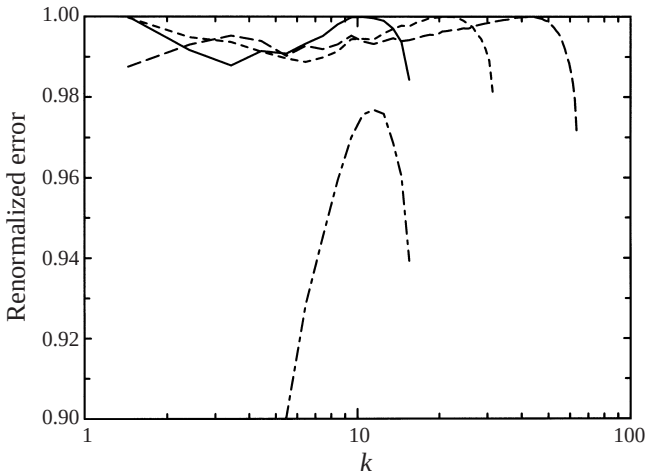

FIGURE 9. Ratio of the quadratic estimation error to the linear estimation error (spectral density) for the isotropic estimates of the subgrid term at sharp Fourier cutoffs of width $k_c = 16$ (———), $k_c = 32$ (– – – –), and $k_c = 64$ (— — —), and of the 46-term estimation error to the linear estimation error at $k_c = 16$ (– · – · –). Interpreted also as error normalized by the part of the subgrid that is statistically orthogonal to velocity (see discussion).

that the discussion of this section is limited to isotropic turbulence and sharp cutoff filters; it is expected that the role of the model is more significant in complex flows.

5. Conclusions

We have shown that there is an LES model that can be written as the conditional average (2.1) and is in all senses ideal. The ideal model yields accurate spatial statistics and minimizes error of the short-time large-scale dynamics. Unfortunately, the conditional average is not directly computable. But this result establishes a well-defined target for practical models, and suggests that formal approximation of the conditional average is a good strategy for developing subgrid models. Furthermore, it was demonstrated that minimizing dynamical error is necessary for accurate statistics, but that the minimum possible dynamical error is not expected to be zero. An obvious consequence is that in performing an LES one must tolerate a level of dynamical error.

Measuring error of the large-scale dynamics also provides an *a priori* framework in which to assess the performance of any proposed subgrid model. Because the limiting error of the ideal model is not known, one does not know how closely a practical model approximates the ideal model. However, the errors of practical models can still be compared to each other to determine which are closest to the ideal. This can be meaningful even when the error is large, since the intrinsic error of the ideal model may itself be large.

Stochastic estimation can be used to develop classes of optimal subgrid models that formally approximate the ideal model by minimizing error in the large-scale dynamics. Expressions are derived for a general class of estimates of the conditional average and it is shown that the simplest (one-point) estimates are able to identically represent the average energy exchanges between the resolved and subgrid scales at each wavenumber, as well as a variety of other one-point statistics.

The formulation was applied to a forced isotropic turbulence, with the large scales

defined using a Fourier cutoff filter. For each model tested, it was found that the mean-square error between the real and estimated subgrid force was large. It is hypothesized that the error of the ideal model is also large; if so, then the subgrid term is mostly stochastic. This would be consistent with previous experience with stochastic estimation of conditional averages in isotropic turbulence (Adrian *et al.* 1989), in which the error in estimating the conditional average was small, while the stochastic component was large. However, without knowing the error committed in approximating the conditional average, this hypothesis cannot be proved directly. There appears to be no good way to evaluate the proposition that the errors of our optimal estimates are comparable to those of the ideal model, except to continue identifying and testing potential dependences. It is important that such modelling activities continue.

These observations have important implications. It has been observed since Clark *et al.* (1979) that LES models do not provide a good instantaneous representation of the subgrid terms, and yet these models appear to work well in actual simulations. The current results suggest that this is inherent to LES modelling; that it is neither necessary nor possible for a deterministic subgrid model to accurately represent the real subgrid terms. Rather, the subgrid model needs to represent the conditional average defined in (2.20).

Even if the exact subgrid term is mostly stochastic, the local dynamics of the largest scales may still be predicted for short times because the subgrid contribution to $\partial\widetilde{u}_i/\partial t$ is small. Furthermore, as discussed in § 2.3, there is also reason to believe that even with a mostly stochastic subgrid term, the turbulent statistics can still be computed correctly; indeed it has been seen that the simplest one-point estimates are capable of recovering the average dissipation of the real system.

If it is indeed the case that the estimates and errors presented in § 4.6 are representative of the conditional average, then, for isotropic turbulence with Fourier cutoff filters, the best subgrid model may be able to do little more than account for the average dissipative effect of the subgrid on each large-scale mode. In any case, it is clear that among the many estimation terms evaluated in § 4, the only identifiable subgrid effect is the dissipative energy transfer represented by the linear estimate. The modest, but discernible improvement of the 46-term estimate over the linear estimate is the cumulative effect of all 46 terms; we were unable to attribute it to any single term or small group of terms.

The dominance of dissipative energy transfer in the deterministic part of the model term M_i (to the extent we can determine it) is clearly specific to isotropic turbulence. In a strongly inhomogeneous flow (e.g. near a wall) a similar analysis is not expected to yield this result. However, the apparent dominance of the dissipation is likely to apply when anisotropy or inhomogeneity are not too strong, provided the filter is sufficiently narrow so that subgrid turbulence is isotropic. This is probably the reason that simple eddy-viscosity models such as Smagorinsky and its dynamic variant have been relatively successful. It is also consistent with the observation by Baggett, Jiménez & Kravchenko (1997) that for the dynamic Smagorinsky model to work well in the core region of a channel flow, the subgrid contribution to the mean Reynolds stress must be negligible. This is a necessary condition for the subgrid turbulence to be isotropic.

Finally, several details of the current results are also specific to the Fourier cutoff filter used here. In particular, the fact that the model term M_i is so small compared to $\partial\widetilde{u}_i/\partial t$ (see figure 1), and the apparent dominance of the dissipation in the deterministic part of M_i is a testament to how well the Navier–Stokes terms in the LES formulation

(1.2) represent the deterministic part of $\partial\widetilde{u}_i/\partial t$. This property almost certainly depends on the filter, so with other filters we should expect the Navier–Stokes terms to represent less of the deterministic $\partial\widetilde{u}_i/\partial t$, leaving more for the optimal LES model term to represent. This is consistent with observations by Piomelli (1988) and Liu *et al.* (1994) that the scale-similarity model yields significant improvements in subgrid model error over Smagorinsky for a Gaussian filter, but not for a Fourier cutoff filter.

Clearly, the optimal LES analysis discussed in this paper needs to be applied to other filters and to inhomogeneous flows so that the necessary characteristics of subgrid models for these cases can be determined. Such analyses are being pursued.

This research was jointly supported by the National Science Foundation and the Air Force Office of Scientific Research under NSF grant CTS-9616219, and by NASA grant NGT 2-52229. We would also like to thank R. Adrian and S. Balachandar for their insightful comments.

REFERENCES

Adrian, R. J. 1977 On the role of conditional averages in turbulence theory. In *Turbulence in Liquids* (ed. S. Zahin & G. Patterson). Princeton, NJ: Science Press.

Adrian, R. J. 1990 Stochastic estimation of sub-grid scale motions. *Appl. Mech. Rev.* **43**, 214–218.

Adrian, R. J., Jones, B. G., Chung, M. K., Hassan, Y., Nithianandan, C. K. & Tung, A. 1989 Approximation of turbulent conditional averages by stochastic estimation. *Phys. Fluids* **1**, 992–998.

Adrian, R. J. & Moin, P. 1988 Stochastic estimation of organized turbulent structure: homogeneous shear flow. *J. Fluid Mech.* **190**, 531–559.

Baggett, J., Jiménez, J. & Kravchenko, A. 1997 Resolution requirements in large-eddy simulations of shear flows. In *Annual Research Briefs*, pp. 51–66. Center for Turbulence Research, Stanford University.

Bardina, J., Ferziger, J. H. & Reynolds, W. C. 1980 Improved subgrid-scale models for large-eddy simulation. *AIAA Paper* 80-1357.

Berkooz, G. 1994 An observation on probability density equations, or, when do simulations reproduce statistics? *Nonlinearity* **7**, 313–328.

Boris, J. P., Grinstein, F. F., Oran, E. S. & Kolbe, R. L. 1992 New insights into large eddy simulation. *Fluid Dyn. Res.* **10**, 199–228.

Carati, D., Wray, A. & Cabot, W. 1996 Ensemble averaged dynamic modeling. In *Proceedings of the Summer Program.* Center for Turbulence Research, Stanford University.

Chollet, J. & Lesieur, M. 1981 Parameterization of small scales of three-dimensional isotropic turbulence using spectral closures. *J. Atmos. Sci.* **38**, 2747–2757.

Clark, R., Ferziger, J. & Reynolds, W. 1979 Evaluation of subgrid-scale models using an accurately simulated turbulent flow. *J. Fluid Mech.* **91**, 1–16.

Domaradzki, J., Metcalfe, R., Rogallo, R. & Riley, J. 1987 Analysis of subgrid-scale eddy viscosity with use of results from direct numerical simulations. *Phys. Rev. Lett.* **68**, 547–550.

Germano, M., Piomelli, U., Moin, P. & Cabot, W. H. 1991 A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids* A **3**, 1760–1765.

Ghosal, S. 1996 An analysis of numerical errors in large-eddy simulations of turbulence. *J. Comput. Phys.* **125**, 187–206.

Ghosal, S., Lund, T., Moin, P. & Akselvoll, K. 1995 A dynamic localization model for large-eddy simulation of turbulent flows. *J. Fluid Mech.* **286**, 229–255.

Ghosal, S. & Moin, P. 1995 The basic equations for the large-eddy simulation of turbulent flows in complex geometries. *J. Comput. Phys.* **118**, 24–37.

Kraichnan, R. 1959 The structure of isotropic turbulence at very high Reynolds numbers. *J. Fluid Mech.* **5**, 497–543.

Kraichnan, R. 1976 Eddy viscosity in two and three dimensions. *J. Atmos. Sci.* **33**, 1521–1536.

Lesieur, M. 1998 Spectral eddy-viscosity based LES of incompressible and compressible shear flows. *AIAA Paper* 98-2894.

Lesieur, M. & Métais, O. 1996 New trends in large-eddy simulations of turbulence. *Ann. Rev. Fluid Mech.* **28**, 45–82.

Lesieur, M. & Rogallo, R. 1989 Large-eddy simulation of passive scalar diffusion in isotropic turbulence. *Phys. Fluids* **1**, 718–722.

Leslie, D. & Quarini, G. 1979 The application of turbulence theory to the formulation of subgrid modelling procedures. *J. Fluid Mech.* **91**, 65–91.

Lilly, D. K. 1964 On the application of the eddy viscosity concept in the inertial subrange of turbulence. *NCAR Manuscr.* 123.

Lilly, D. K. 1992 A proposed modification of the Germano subgrid-scale closure method. *Phys. Fluids* **4**, 633–635.

Liu, S., Meneveau, C. & Katz, J. 1994 On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet. *J. Fluid Mech.* **275**, 83–119.

Lund, T., Ghosal, S. & Moin, P. 1993 Numerical experiments with highly-variable eddy-viscosity models. In *Engineering Applications to Large Eddy Simulation* (ed. U. Piomelli & S. Ragab), pp. 7–11. ASME.

Lundgren, T. S. 1967 Distribution functions in the statistical theory of turbulence. *Phys. Fluids* **10**, 969–975.

Meneveau, C. 1994 Statistics of turbulence subgrid-scale stresses: Necessary conditions and experimental tests. *Phys. Fluids* **6**, 815–833.

Meneveau, C., Lund, T. & Cabot, W. H. 1996 A Lagrangian dynamic subgrid-scale model of turbulence. *J Fluid Mech.* **319**, 353–385.

Métais, O. & Lesieur, M. 1992 Spectral large-eddy simulation of isotropic and stably stratified turbulence. *J. Fluid Mech.* **239**, 157.

Orszag, S. 1970 Analytical theories of turbulence. *J. Fluid Mech.* **41**, 363–386.

Panton, R. 1997 *Self-Substaining Mechanisms of Wall Turbulence*. Computational Mechanics.

Papoulis, A. 1965 *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill.

Piomelli, U. 1988 Model consistency in large eddy simulation of turbulent channel flows. *Phys. Fluids* **31**, 1884–1891.

Pope, S. B. 1985 P.d.f. methods for turbulent reactive flows. *Prog. Energy Combust. Sci.* **11**, 119–192.

Rogallo, R. S. 1981 Numerical experiments in homogeneous turbulence. *NASA TM*-81315.

Rogallo, R. S. & Moin, P. 1984 Numerical simulation of turbulent flows. *Ann. Rev. Fluid Mech.* **16**, 99–137.

Smagorinsky, J. 1963 General circulation experiments with the primitive equations. *Mon. Weath. Rev.* **91**, 99–164.

Vasilyev, O. V., Lund, T. S. & Moin, P. 1998 A general class of commutative filters for les in complex geometries. *J. Comput. Phys.* **146**, 82–104.

Zhou, Y. & Vahala, G. 1993 Reformulation of recursive-renormalization-group-based subgrid modelling of turbulence. *Phys. Rev.* E **47**, 2503–2519.