4.3.3 Improvement of Models in the Physical Space

Statement of the Problem. Experience shows that the various models yield good results when they are applied to homogeneous turbulent flows and that the cutoff is placed sufficiently far into the inertial range of the spectrum, *i.e.* when a large part of the total kinetic energy is contained in the resolved scales[15].

In other cases, as in transitional flows, highly anisotropic flows, highly under-resolved flows, or those in high energetic disequilibrium, the subgrid models behave much less satisfactorily. Aside from the problem stemming from numerical errors, there are mainly two reasons for this:

1. The characteristics of these flows does not correspond to the hypotheses on which the models are derived, which means that the models are at fault. We then have two possibilities: deriving models from new physical hypotheses or adjusting existing ones, more or less empirically. The first choice is theoretically more exact, but there is a lack of descriptions of turbulence for frameworks other than that of isotropic homogeneous turbulence. Still, a few attempts have been made to consider the anisotropy appearing in this category. These are discussed in Chap. 5. The other solution, if the physics of the models is put to fault, consists in reducing their importance, *i.e.* increasing the cutoff frequency to capture a larger part of the flow physics directly. This means increasing the number of degrees of freedom and striking a compromise between the grid enrichment techniques and subgrid modeling efforts.
2. Deriving models based on the energy at cutoff or the subgrid scales (with no additional evolution equation) for simulations in the physical space runs up against Gabor-Heisenberg's generalized principle of uncertainty [91, 285], which stipulates that the precision of the information cannot be improved in space and in frequency at the same time. This is illustrated by Fig. 4.15. Very good frequency localization implies high non-localization in space, which reduces the possibilities of taking the intermittency[16] into account and precludes the treatment of highly inhomogeneous flows. Inversely, very good localization of the information in space prevents any good spectral resolution, which leads to high errors, e.g. in computing the energy at the cutoff. Yet this frequency localization is very important, since it alone can be used to detect the presence of the subgrid scales. It is important to recall here that large-eddy simulation is based on a selection in frequency of modes making up the exact solution. Problems arise here, induced by the localization of statistical average relations that are exact, as this localization may correspond to a statistical average. Two solutions may be considered: developing an acceptable compromise between the precision in space and frequency, or enriching the information contained in the simulation, which is done either by adding variables to it as in the case of models based on the kinetic energy of the subgrid modes, or by assuming further hypotheses when deriving models.

[15] Certain authors estimate this share to be between 80 % and 90 % [53]. Another criterion sometimes mentioned is that the cutoff scale should be of the order of Taylor's microscale. Bagget *et al.* [10] propose to define the cutoff length in such a way that the subgrid scales will be quasi-isotropic, leading to $\overline{\Delta} \approx L_\varepsilon/10$, where L_ε is the integral dissipation length.

[16] Direct numerical simulations and experimental data show that the true subgrid dissipation and its surrogates do not have the same scaling s [52, 232].

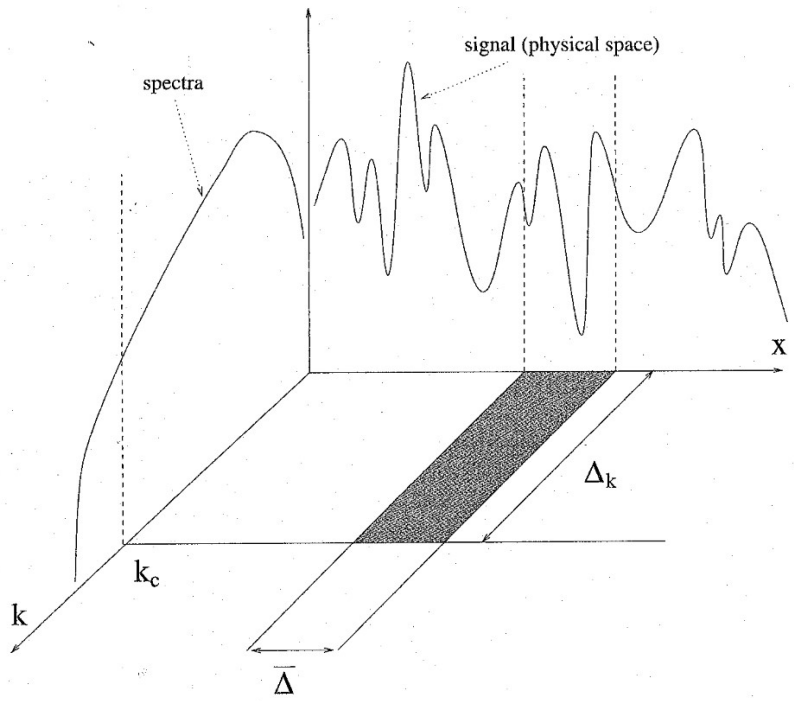

Fig. 4.15. Representation of the resolution in the space-frequency plane. The spatial resolution $\overline{\Delta}$ is associated with frequency resolution Δ_k. Gabor-Heisenberg's uncertainty principle stipulates that the product $\overline{\Delta} \times \Delta_k$ remains constant, *i.e.* that the area of the gray domain keeps the same value (from [91], with the permission of F. Ducros).

In the following, we present techniques developed to improve the simulation results, though without modifying the structure of the subgrid models deeply. The purpose of all these modifications is to adapt the subgrid model

better to the local state of the flow and remedy the lack of frequency localization of the information.

We will be describing:

1. Dynamic procedures for computing subgrid model constants (p.105). These constants are computed in such a way as to reduce an a priori estimate of the error committed with the model considered, locally in space and time, in the least squares sense. This estimation is made using the Germano identity, and requires the use of an analytical filter. It should be noted that the dynamic procedures do not change the model in the sense that its form (e.g. subgrid viscosity) remains the same. All that is done here is to minimize a norm of the error associated with the form of the model considered. The errors committed intrinsically[17] by adopting an a priori form of the subgrid tensor are not modified. These procedures, while theoretically very attractive, do pose problems of numerical stability and can induce non-negligible extra computational costs. This variation of the constant at each point and each time step makes it possible to minimize the error locally for each degree of freedom, while determining a constant value offers only the less efficient possibility of an overall minimization. This is illustrated by the above discussion of the constant in the Smagorinsky model.
2. Structural sensors (p.118), which condition the existence of the subgrid scales to the verification of certain constraints by the highest frequencies of the resolved scales. More precisely, we consider here that the subgrid scales exist if the highest resolved frequencies verify topological properties that are expected in the case of isotropic homogeneous turbulence. When these criteria are verified, we adopt the hypothesis that the highest resolved frequencies have a dynamics close to that of the scales contained in the inertial range. On the basis of energy spectrum continuity (see the note of page p.83), we then deduce that unresolved scales exist, and the subgrid model is then used, but is otherwise canceled.
3. The accentuation technique (p.120), which consists in artificially increasing the contribution of the highest resolved frequencies when evaluating the subgrid viscosity. This technique allows a better frequency localization of the information included in the model, and therefore a better treatment of the intermittence phenomena, as the model is sensitive only to the higher resolved frequencies. This result is obtained by applying a frequency high-pass filter to the resolved field.
4. The damping functions for the near-wall region (p.123), by which certain modifications in the turbulence dynamics and characteristic scales of the subgrid modes in the boundary layers can be taken into account. These functions are established in such a way as to cancel the subgrid viscosity models in the near-wall region so that they will not inhibit the driving mechanisms occurring in this area. These models are of limited generality as they presuppose a particular form of the flow dynamics in the region considered. They also require that the relative position of each point with respect to the solid wall be known, which can raise problems in practice such as when using multidomain techniques or when several surfaces exist. And lastly, they constitute only an amplitude correction of the subgrid viscosity models for the forward energy cascade: they are not able to include any changes in the form of this mechanism, or the emergence of new mechanisms.

[17] For example, the subgrid viscosity models described above all induce a linear dependency between the subgrid tensor and the resolved-scale tensor:

$$\tau_{ij}^{\mathrm{d}} = -\nu_{\mathrm{sgs}}\overline{S}_{ij} \quad .$$

The three "generalist" techniques (dynamic procedure, structural sensor, accentuation) for adapting the subgrid viscosity models are all based on extracting a test field from the resolved scales by applying a test filter to these scales. This field corresponds to the highest frequencies catured by the simulation, so we can see that all these techniques are based on a frequency localization of the information contained in the subgrid models. The loss of localness in space is reflected by the fact that the number of neighbors involved in the subgrid model computation is increased by using the test filter.

Dynamic Procedures for Computing the Constants. Dynamic Models

Germano–Lilly Dynamic Procedure. In order to adapt the models better to the local structure of the flow, Germano *et al.* [118] proposed an algorithm for adapting the Smagorinsky model by automatically adjusting the constant at each point in space and at each time step. This procedure, described below, is applicable to any model that makes explicit use of an arbitrary constant C_{d}, such that the constant now becomes time- and space-dependent: C_{d} becomes $C_{\mathrm{d}}(\mathbf{x}, t)$.

The dynamic procedure is based on the Germano identity (3.78), now written in the form:

$$L_{ij} = T_{ij} - \tilde{\tau}_{ij} \quad , \tag{4.126}$$

in which

$$\tau_{ij} \equiv L_{ij} + C_{ij} + R_{ij} = \overline{u_i u_j} - \overline{u}_i \overline{u}_j \quad , \tag{4.127}$$

$$T_{ij} \equiv \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u}}_i \widetilde{\overline{u}}_j \quad , \tag{4.128}$$

$$L_{ij} \equiv \widetilde{\overline{u}_i \overline{u}_j} - \widetilde{\overline{u}}_i \widetilde{\overline{u}}_j \quad , \tag{4.129}$$

in which the tilde symbol *tilde* designates the test filter. The tensors τ and T are the subgrid tensors corresponding, respectively, to the first and second

filtering levels. The latter filtering level is associated with the characteristic length $\widetilde{\overline{\Delta}}$, with $\widetilde{\overline{\Delta}} > \overline{\Delta}$. Numerical tests show that an optimal value is $\widetilde{\overline{\Delta}} = 2\overline{\Delta}$. The tensor L can be computed directly from the resolved field.

We then assume that the two subgrid tensors τ and T can be modeled by the same constant C_d for both filtering levels. Formally, this is expressed:

$$\tau_{ij} - \frac{1}{3}\tau_{kk}\delta_{ij} = C_\mathrm{d}\beta_{ij} \quad , \tag{4.130}$$

$$T_{ij} - \frac{1}{3}T_{kk}\delta_{ij} = C_\mathrm{d}\alpha_{ij} \quad , \tag{4.131}$$

in which the tensors α and β designate the deviators of the subgrid tensors obtained using the subgrid model deprived of its constant. It is important noting that the use of the same subgrid model with the same constant is equivalent to a scale-invariance assumption on both the subgrid fluxes and the filter, to be discussed in the following.

Some examples of subgrid model kernels for α_{ij} and β_{ij} are given in Table 4.1.

Table 4.1. Examples of subgrid model kernels for the dynamic procedure.

Model	β_{ij}	α_{ij}
(4.90)	$-2\overline{\Delta}^2\lvert\overline{S}\rvert\overline{S}_{ij}$	$-2\widetilde{\overline{\Delta}}^2\lvert\widetilde{\overline{S}}\rvert\widetilde{\overline{S}}_{ij}$
(4.102)	$-2\overline{\Delta}\sqrt{\overline{F}(\overline{\Delta})}\overline{S}_{ij}$	$-2\widetilde{\overline{\Delta}}\sqrt{\widetilde{\overline{F}}(\widetilde{\overline{\Delta}})}\widetilde{\overline{S}}_{ij}$
(4.116)	$-2\overline{\Delta}^{1+\alpha}\lvert\mathcal{F}(\overline{\mathbf{u}})\rvert^\alpha(q_c^2)^{\frac{1-\alpha}{2}}\overline{S}_{ij}$	$-2\widetilde{\overline{\Delta}}^{1+\alpha}\lvert\mathcal{F}(\widetilde{\overline{\mathbf{u}}})\rvert^\alpha(\tilde{q}_c^2)^{\frac{1-\alpha}{2}}\widetilde{\overline{S}}_{ij}$

Introducing the above two formulas in the relation (4.126), we get[18]:

$$L_{ij} - \frac{1}{3}L_{kk}\delta_{ij} \equiv L_{ij}^\mathrm{d} = C_\mathrm{d}\alpha_{ij} - \widetilde{C_\mathrm{d}\beta_{ij}} \quad . \tag{4.132}$$

We cannot use this equation directly for determining the constant C_d because the second term uses the constant only through a filtered product [281]. In order to continue modeling, we need to make the approximation:

$$\widetilde{C_\mathrm{d}\beta_{ij}} = C_\mathrm{d}\widetilde{\beta}_{ij} \quad , \tag{4.133}$$

which is equivalent to considering that C_d is constant over an interval at least equal to the test filter cutoff length. The parameter C_d will thus be computed

[18] It is important to note that, for the present case, the tensor L_{ij} is replaced by its deviatoric part L_{ij}^d, because we are dealing with a zero-trace subgrid viscosity modeling.

in such a way as to minimize the error committed[19], which is evaluated using the residual E_{ij}:

$$E_{ij} = L_{ij} - \frac{1}{3}L_{kk}\delta_{ij} - C_\mathrm{d}\alpha_{ij} + C_\mathrm{d}\widetilde{\beta}_{ij} \quad . \tag{4.134}$$

This definition consists of six independent relations, which in theory makes it possible to determine six values of the constant[20]. In order to conserve a single relation and thereby determine a single value of the constant, Germano *et al.* propose contracting the relation (4.134) with the resolved strain rate tensor. The value sought for the constant is a solution of the problem:

$$\frac{\partial E_{ij}\overline{S}_{ij}}{\partial C_\mathrm{d}} = 0 \quad . \tag{4.135}$$

This method can be efficient, but does raise the problem of indetermination when the tensor $\overline{S}_{ij}$ cancels out. To remedy this problem, Lilly [195] proposes calculating the constant C_d by a least-squares method, by which the constant C_d now becomes a solution of the problem:

$$\frac{\partial E_{ij}E_{ij}}{\partial C_\mathrm{d}} = 0 \quad , \tag{4.136}$$

or

$$C_\mathrm{d} = \frac{m_{ij}L_{ij}^\mathrm{d}}{m_{kl}m_{kl}} \quad , \tag{4.137}$$

in which

$$m_{ij} = \alpha_{ij} - \widetilde{\beta}_{ij} \quad . \tag{4.138}$$

The constant C_d thus computed has the following properties:

- It can take negative values, so the model can have an anti-dissipative effect locally. This is a characteristic that is often interpreted as a modeling of the backward energy cascade mechanism. This point is detailed in Sect. 4.4.
- It is not bounded, since it appears in the form of a fraction whose denominator can cancel out[21].

[19] Meneveau and Katz [227] propose to use the dynamic procedure to rank the subgrid models, the best one being associated with the lowest value of the residual.

[20] Which would lead to the definition of a tensorial subgrid viscosity model.

[21] This problem is linked to the implementation of the model in the simulation. In the continuous case, if the denominator tends toward zero, then the numerator cancels out too. These are calculation errors that lead to a problem of division by zero.

These two properties have important practical consequences on the numerical solution because they are both potentially destructive of the stability of the simulation. Numerical tests have shown that the constant can remain negative over long time intervals, causing an exponential growth in the high frequency fluctuations of the resolved field. The constant therefore needs an ad hoc process to ensure the model's good numerical properties. There are a number of different ways of performing this process on the constant: statistical average in the directions of statistical homogeneity [118, 354], in time or local in space [366]; limitation using arbitrary bounds [366] (clipping); or by a combination of these methods [354, 366]. Let us note that the averaging procedures can be defined in two non-equivalent ways [367]: by averaging the denominator and numerator separately, which is denoted symbolically:

$$C_{\mathrm{d}} = \frac{\langle m_{ij} L_{ij}^{\mathrm{d}} \rangle}{\langle m_{kl} m_{kl} \rangle} \quad , \tag{4.139}$$

or by averaging the quotient, *i.e.* on the constant itself:

$$C_{\mathrm{d}} = \langle C_{\mathrm{d}} \rangle = \langle \frac{m_{ij} L_{ij}^{\mathrm{d}}}{m_{kl} m_{kl}} \rangle \quad . \tag{4.140}$$

The ensemble average can be performed over homogeneous directions of the simulation (if they exist) or over different realizations, *i.e.* over several statistically equivalent simulations carried out in parallel [48, 51].

The time average process is expressed:

$$C_{\mathrm{d}}(\mathbf{x}, (n+1)\Delta t) = a_1 C_{\mathrm{d}}(\mathbf{x}, (n+1)\Delta t) + (1 - a_1) C_{\mathrm{d}}(\mathbf{x}, n\Delta t) \quad , \tag{4.141}$$

in which Δt is the time step used for the simulation and $a_1 \leq 1$ a constant. Lastly, the constant clipping process is intended to ensure that the following two conditions are verified:

$$\nu + \nu_{\mathrm{sgs}} \geq 0 \quad , \tag{4.142}$$

$$C_{\mathrm{d}} \leq C_{\max} \quad . \tag{4.143}$$

The first condition ensures that the total resolved dissipation $\varepsilon = \nu \overline{S}_{ij} \overline{S}_{ij} - \tau_{ij} \overline{S}_{ij}$ remains positive or zero. The second establishes an upper bound. In practice, $C_{\max}$ is of the order of the theoretical value of the Smagorinsky constant, *i.e.* $C_{\max} \simeq (0.2)^2$.

The models in which the constant is computed by this procedure are called "dynamic" because they automatically adapt to the local state of the flow. When the Smagorinsky model is coupled with this procedure, it is habitually called the dynamic model, because this combination was the first to be put to the test and is still the one most extensively used among the dynamic models.

The use of the same value of the constant for the subgrid model at the two filtering levels appearing in equation (4.126) implicitely relies on the two following self-similarity assumptions:

- The two cutoff wave numbers are located in the inertial range of the kinetic energy spectrum;
- The filter kernels associated to the two filtering levels are themselves self-similar.

These two constraints are not automatically satisfied, and the validity of the dynamic procedure for computing the constant requires a careful analysis.

Meneveau and Lund [229] propose an extension of the dynamic procedure for a cutoff located in the viscous range of the spectrum. Writing the constant of the subgrid-scale model C as an explicit function of the filter characteristic length, the Germano–Lilly procedure leads to

$$C(\overline{\Delta}) = C(\widetilde{\overline{\Delta}}) = C_{\mathrm{d}} \quad . \tag{4.144}$$

Let η be the Kolmogorov lengthscale. It was said in the introduction that the flow is fully resolved if $\overline{\Delta} = \eta$. Therefore, the dynamic procedure is consistent if, and only if

$$\lim_{\overline{\Delta} \to \eta} C_{\mathrm{d}} = C(\eta) = 0 \quad . \tag{4.145}$$

Numerical experiments carried out by the two authors show that the Germano–Lilly procedure is not consistent, because it returns the value of the constant associated to the test filter level

$$C_{\mathrm{d}} = C(\widetilde{\overline{\Delta}}) \quad , \tag{4.146}$$

yielding

$$\lim_{\overline{\Delta} \to \eta} C_{\mathrm{d}} = C(r\eta) \neq 0, \quad r = \widetilde{\overline{\Delta}} / \overline{\Delta} \quad . \tag{4.147}$$

Numerical tests also showed that taking the limit $r \to 1$ or computing the two values $C(\overline{\Delta})$ and $C(r\overline{\Delta})$ using least-square-error minimization without assuming them to be equal yield inconsistent or ill-behaved solutions. A solution is to use prior knowledge to compute the dynamic constant. A robust algorithm is obtained by rewriting equation (4.134) as follows:

$$E_{ij} = L_{ij}^{\mathrm{d}} - C(\overline{\Delta}) \left(f(\overline{\Delta}, r) \alpha_{ij} - \widetilde{\beta}_{ij} \right) \quad , \tag{4.148}$$

where $f(\overline{\Delta}, r) = C(r\overline{\Delta}) / C(\overline{\Delta})$ is evaluated by calculations similar to those of Voke (see page 89). A simple analytical fitting is obtained in the case $r = 2$:

$$f(\overline{\Delta}, 2) \approx \max(100, 10^{-x}), \; x = 3.23 (Re_{2\overline{\Delta}}^{-0.92} - Re_{\overline{\Delta}}^{-0.92}) \quad , \tag{4.149}$$

where the mesh-Reynolds numbers are evaluated as (see equation (4.70)):

$$Re_{\overline{\Delta}} = \frac{\overline{\Delta}^2|\overline{S}|}{\nu}, \quad Re_{2\overline{\Delta}} = \frac{4\overline{\Delta}^2|\widetilde{\overline{S}}|}{\nu} \quad .$$

We now consider the problem of the filter self-similarity. Let G_1 and G_2 be the filter kernels associated to the first and second filtering level. For sake of simplicity, we use the notations $\overline{\Delta} = \Delta_1$ and $\widetilde{\overline{\Delta}} = \Delta_2$. We assume that the filter kernel are re-written in a form such that:

$$\overline{u}(x) = G_1 \star u(x) = \int G_1\left(\frac{|x-\xi|}{\Delta_1}\right) u(\xi)d\xi \quad , \tag{4.150}$$

$$\widetilde{\overline{u}}(x) = G_2 \star u(x) = \int G_2\left(\frac{|x-\xi|}{\Delta_2}\right) u(\xi)d\xi \quad . \tag{4.151}$$

We also introduce the test filter G_t, which is defined such that

$$\widetilde{\overline{u}} = G_2 \star u = G_t \star \overline{u} = G_t \star G_1 \star u \quad . \tag{4.152}$$

The filters G_1 and G_2 are self-similar if and only if

$$G_1(y) = \frac{1}{r^d} G_2\left(\frac{y}{r}\right), \quad r = \Delta_2/\Delta_1 \quad . \tag{4.153}$$

Hence, the two filters must have identical shapes and may only differ by their associated characteristic length. The problem is that in practice only G_t is known, and the self-similarity property might not be *a priori* verified. Carati and Vanden Eijnden [49] show that the interpretation of the resolved field is fully determined by the choice of the test filter G_t, and that the use of the same model for the two levels of filtering is fully justified. This is demonstrated by re-interpreting previous filters in the following way. Let us consider an infinite set of self-similar filters $\{F_n \equiv F(l_n)\}$ defined as

$$F_n(x) = \frac{1}{r^n}\mathcal{F}\left(\frac{x}{l_n}\right), l_n = r^n l_0 \quad , \tag{4.154}$$

where $\mathcal{F}$, $r > 1$ and l_0 are the filter kernel, an arbitrary parameter and a reference length, respectively. Let us introduce a second set $\{F_n^* \equiv F^*(l_n^*)\}$ defined by

$$F_n^* \equiv F_n \star F_{n-1} \star ... \star F_{-\infty} \quad . \tag{4.155}$$

For positive kernel $\mathcal{F}$, we get the following properties:

- The length l_n^* obeys the same geometrical law as l_n :

$$l_n^* = r l_{n-1}^*, \quad \text{and} \quad l_n^* = \frac{r}{\sqrt{r^2-1}} l_n \quad . \tag{4.156}$$

- $\{F_n^*\}$ constitute a set of self-similar filters.

Using these two set of filters, the classical filters involved in the dynamic procedure can be defined as self-similar filters:

$$G_t(\Delta_t) = F_n(l_n) \quad , \tag{4.157}$$

$$G_1(\Delta_1) = F_{n-1}^*(l_{n-1}^*) \quad , \tag{4.158}$$

$$G_2(\Delta_2) = F_n^*(l_n^*) \quad . \tag{4.159}$$

For any test-filter G_t and any value of r, the first filter operator can be constructed explicitly :

$$G_1 = G_t(\Delta_t/r) \star G_t(\Delta_t/r^2) \star ... \star G_t(\Delta_t/r^\infty) \quad . \tag{4.160}$$

This relation shows that for any test filter of the form (4.157), the two filtering operators can be rewritten as self-similar ones, justifying the use of the same model at all the filtering levels.

Lagrangian Dynamic Procedure. The constant regularization procedures based on averages in the homogeneous directions have the drawback of not being usable in complex configurations, which are totally inhomogeneous. One technique for remedying this problem is to take this average along the fluid particle trajectories. This new procedure [230], called the dynamic Lagrangian procedure, has the advantage of being applicable in all configurations.

The trajectory of a fluid particle located at position $\mathbf{x}$ at time t is, for times t' previous to t, denoted as:

$$\mathbf{z}(t') = \mathbf{x} - \int_{t'}^{t} \overline{\mathbf{u}}[\mathbf{z}(t''), t'']dt'' \quad . \tag{4.161}$$

The residual (4.134) is written in the following Lagrangian form:

$$E_{ij}(\mathbf{z}, t') = L_{ij}(\mathbf{z}, t') - C_{\mathrm{d}}(\mathbf{x}, t) m_{ij}(\mathbf{z}, t') \quad . \tag{4.162}$$

We see that the value of the constant is fixed at point $\mathbf{x}$ at time t, which is equivalent to the same linearization operation as for the Germano–Lilly procedure. The value of the constant that should be used for computing the subgrid model at $\mathbf{x}$ at time t is determined by minimizing the error along the fluid particle trajectories. Here too, we reduce to a well-posed problem by defining a scalar residual E_{lag}, which is defined as the weighted integral along the trajectories of the residual proposed by Lilly:

$$E_{\mathrm{lag}} = \int_{-\infty}^{t} E_{ij}(\mathbf{z}(t'), t') E_{ij}(\mathbf{z}(t'), t') W(t-t')dt' \quad , \tag{4.163}$$

in which the weighting function $W(t-t')$ is introduced to control the memory effect. The constant is a solution of the problem: