IMA Journal of Applied Mathematics (2015) **80**, 1099–1123
doi:10.1093/imamat/hxu042
Advance Access publication on 18 September 2014

Optimizing performance of the deconvolution model reduction for large ODE systems

Lyudmyla L. Barannyk*
Department of Mathematics, University of Idaho, Moscow, ID 83844, USA
*Corresponding author: barannyk@uidaho.edu

and

Alexander Panchenko
Department of Mathematics, Washington State University, Pullman, WA 99164, USA

[Received on 27 February 2013; revised on 19 March 2014; accepted on 4 August 2014]

We investigate the numerical performance of the regularized deconvolution closure introduced recently by the authors. The purpose of the closure is to furnish constitutive equations for Irving–Kirkwood–Noll procedure, a well-known method for deriving continuum balance equations from the Newton's equations of particle dynamics. A version of this procedure used in the paper relies on spatial averaging developed by Hardy, and independently by Murdoch and Bedeaux. The constitutive equations for the stress are given as a sum of non-linear operator terms acting on the mesoscale average density and velocity. Each term is a 'convolution sandwich' containing the deconvolution operator, a composition or a product operator and the convolution (averaging) operator. Deconvolution is constructed using filtered regularization methods from the theory of ill-posed problems. The purpose of regularization is to ensure numerical stability. The particular technique used for numerical experiments is truncated singular value decomposition (SVD). The accuracy of the constitutive equations depends on several parameters: the choice of the averaging window function, the value of the mesoscale resolution parameter, scale separation, the level of truncation of singular values and the level of spectral filtering of the averages. We conduct numerical experiments to determine the effect of each parameter on the accuracy and efficiency of the method. Partial error estimates are also obtained.

Keywords: FPU chain; particle chain; Irving–Kirkwood–Noll procedure; Hardy–Murdoch averaging; upscaling; model reduction; dimension reduction; closure; regularized deconvolution; ill-posed problems; truncated SVD; spectral filtering.

1. Introduction

Many biological and advanced man-made materials are highly heterogeneous, with internal structure spanning a broad range of spatial scales. Micron-to-nanometre scales seem to be of particular importance for understanding the response of such materials. Traditional continuum models miss important effects at sub-micron scales: micro-instabilities, dispersive energy transport, radiative damping and presence of phonon gaps (see, e.g. Charlotte & Truskinovsky, 2012). To capture these features, one could use molecular dynamics (MD), but analysis of such models is difficult, and their computational efficiency is poor. Thus, there is a need for new mesoscopic continuum-style models capable of describing smaller-scale effects. A related problem in computational science is the design of fast multiscale algorithms for simulating coarse scale particle dynamics. Another relevant question, perhaps the most fundamental of the three, is to find a general method for generating continuum constitutive equations from the underlying atomistic models.

Continuum balance equations can be derived by averaging particle dynamics. This approach was originated by Irving & Kirkwood (1950), who used ensemble averages to derive hydrodynamics equations directly from Hamiltonian dynamics. Soon after, Noll (1955) improved the mathematical foundation of the method by introducing finite size 'window functions' instead of delta functions. Several decades later, Hardy (1982), Murdoch & Bedeaux (1994), Murdoch & Bedeaux (1996), Murdoch & Bedeaux (1997) and Murdoch (2007) independently developed a theory based entirely on space-time averaging. Recently, Hardy–Murdoch approach has attracted much attention in the applied mathematics, materials science and physics communities. It has found use both for solids by Tadmor & Miller (2011) and fluids far from equilibrium by Evans & Morriss (2008). These books also contain a large number of references. A few representative examples are E *et al.* (2009) (complex fluids), Lehoucq & Sears (2011) (peridynamics) and Zimmerman *et al.* (2010) (solids). In the latter work, the authors develop a material (Lagrangian) frame approach as opposed to spatial (Eulerian) formulation of the standard Hardy–Murdoch equations. The recent papers by Admal & Tadmor (2010, 2011) extend the averaging approach to multibody potentials.

Despite its many attractive features, Hardy–Murdoch averaging does not produce an effective continuum theory. The fluxes in their balance equations are functions of particle positions and velocities. This means that trajectories of all particles must be known before fluxes can be calculated. To recover the efficiency of continuum description, one needs to formulate constitutive equations for Hardy–Murdoch fluxes, that is, approximate them by operators acting on the averages.

Mesoscale continuum theories differ from classical continuum. The two features they exhibit most often are non-locality and scale dependence. For example, spatially non-local solid mechanics theories were proposed by Eringen (1976) and Kunin (1982). Charlotte & Truskinovsky (2012) introduced a temporally non-local (history dependent) model of 1D lattices with linear nearest neighbour interactions. Another non-local and scale dependent approach is peridynamics proposed by Silling (2000). In the past 12 years, peridynamics has gained popularity and has been the subject of many works in applied mathematics, notably by Du, Gunzburger, Lehoucq, Silling and many others (see the review papers by Silling & Lehoucq (2010), Lehoucq & Sears (2011), Du *et al.* (2012) and references therein). The original theory is phenomenological, so the constitutive equations are postulated or fitted from experiments. Scale dependence is introduced by using a fixed size window function (called horizon by Silling). The latter feature is similar to Hardy–Murdoch procedure, so one expects that peridynamic equations can be derived by averaging atomistic models. This was recently done by Lehoucq & Sears (2011). They provided an exact description of the peridynamical force state in terms of the particle positions and velocities.

A method for generating non-local and scale-dependent constitutive equations for Hardy–Murdoch averages was introduced in Panchenko *et al.* (2011) and tested numerically in Panchenko *et al.* (2014). Applications to discrete models of fluid flow were studied in Tartakovsky *et al.* (2011) and Panchenko & Tartakovsky (2014). Recently, similar methods have been successfully used in large eddy simulation of turbulence by Berselli *et al.* (2006), Kim *et al.* (2012) and Layton & Rebholz (2012).

In this paper, we study algorithmic realization of the regularized deconvolution closure. Since the problem at hand is non-linear, the error estimates that we provide are not likely to be tight, so we also study the error evolution numerically. The objective is to understand the dependence of the error on various parameters of the method. Specifically, we study the effects of a choice of a window function ψ, resolution parameter η, scale separation, truncation level in singular value decomposition (SVD) and filtering of spectral coefficients. The numerical experiments are conducted as follows. We solve the system of particle dynamics ordinary differential equations (ODEs) to have microscopic positions and velocities that we call 'exact' and compute all relevant averages. These averages can be approximated

as convolutions of a window function and certain functions of particle trajectories. Next, we apply a regularized deconvolution to computed average density and momentum to obtain approximations of the Jacobian J and a velocity interpolant $\tilde{\mathbf{v}}$. These approximations are used to calculate the approximate stress that is then compared with the directly computed ('exact') stress. We also compare 'exact' and approximated Jacobian and microscopic velocity.

A discretization of the integral convolution equation is a linear system. As a consequence of ill-posedness, the matrix of this system is ill-conditioned. To regularize the discrete problem, we use truncated SVD. In the continuum case, ill-posedness increases with the increase of the rate of decay of the singular values of the kernel (see Kirsch, 1996; Hansen, 1987). It is then reasonable to expect that the rate of decay of singular values plays a similar role in the discrete setting. However, the overall error, that is, the error in computing the approximate stress (1.1), does not follow this pattern. The Gaussian kernel, which has the fastest decay rate among the considered window functions, corresponds to a smaller overall error.

An important aspect of the method's performance is dependence of the error on scale separation characterized by, for example, the product of η and N. Numerical experiments conducted in this paper show that, as the scale separation increases, stresses are computed more accurately even though the recoverable functions are approximated more poorly. This is a non-trivial conclusion since the stress is a sum of several operators with the structure of a non-linear 'convolution sandwich'

$$R_\eta S Q_\eta^\alpha, \tag{1.1}$$

where R_η is the convolution (averaging) operator, Q_η^α is the deconvolution operator (this is an approximate regularized inverse of R_η that depends on a regularization parameter α) and S is a non-linear composition or a product-type operator.

The classical theory of ill-posed problems Kirsch (1996), Tikhonov & Arsenin (1987) and Morozov (1984) deals mostly with operators between Hilbert spaces, because for integral operators, the Hilbert space structure is required for existence of SVD. In the discrete setting, one can use SVD with any p-norm, including the ∞-norm. In the paper, we provide such error estimates for $p \in [1, \infty]$. The values of p for the right-hand side and the solution may be different. We also use the l^∞-norm for computational assessment of the quality of approximation.

We would like to make clear that this work is based on a spatial averaging version of Irving–Kirkwood–Noll procedure developed by Hardy (1982) and Murdoch & Bedeaux (1994). Ensemble averaging is not used and initial conditions for atomistic systems are supposed to be deterministic and known completely. This makes sense from a practical point of view since ensemble averages are difficult to compute and most of MD simulations are run one realization at a time. Thus, the constitutive theory in Panchenko *et al.* (2011), Tartakovsky *et al.* (2011), Panchenko *et al.* (2014) and present work applies to a single realization. Ensemble averaging can be employed in addition to spatial averaging, which would permit a comparison of the suggested constitutive theory to the classical models of statistical thermodynamics. This interesting question is not addressed here.

The organization of the paper is as follows. The background concerning the fine scale problem and spatial averaging is briefly described in Section 2. Section 3 contains the necessary facts about filtered regularization methods in a discrete setting. In Section 4, we study various window functions used to set-up averaging and the effect of a choice of a window function on the quality of the stress approximation. In Section 5, we analyse the effect of the resolution parameter on the reconstruction of microscopic quantities of interest and eventual stress approximation. In Section 6, we study the Fourier spectra of the reconstruction error and the overall error in the stress. The effect of the scale separation is considered

in Section 7. Error estimates for filtered regularization methods for $p \in [1, \infty]$ are provided in Section 8. Finally, in Section 9 we give conclusions. In appendices, we define various window functions used in the numerical experiments and provide the formula for the Lennard–Jones potential.

2. Governing equations

2.1 *Microscale equations*

The object of our study is a system of N particles that move according to Newton's law of motion, where $N \gg 1$. Particle positions and velocities are denoted by $\boldsymbol{q}_i(t)$ and $\boldsymbol{v}_i(t)$, $i = 1, \ldots, N$, respectively. For simplicity, we assume that all particles have the same mass M/N where M is the total mass of the system. The equations of motion are

$$\dot{\boldsymbol{q}}_i = \boldsymbol{v}_i, \tag{2.1}$$

$$\frac{M}{N}\dot{\boldsymbol{v}}_i = \boldsymbol{f}_i + \boldsymbol{f}_i^{(\mathrm{ext})}, \tag{2.2}$$

with initial conditions

$$\boldsymbol{q}_i(0) = \boldsymbol{q}_i^0, \quad \boldsymbol{v}_i(0) = \boldsymbol{v}_i^0. \tag{2.3}$$

Here $\boldsymbol{f}_i^{(\mathrm{ext})}$ is an external force (gravity) and $\boldsymbol{f}_i = \sum_j \boldsymbol{f}_{ij}$, where $\boldsymbol{f}_{ij}$ is an interaction force exerted on a particle i by a particle j. These forces are generated by a pair potential $U(|\boldsymbol{q}_i - \boldsymbol{q}_j|)$. In numerical experiments, we use the classical Lennard–Jones potential (see equation (B.1)).

To study the behaviour of the system for large N, it is convenient to introduce a small parameter

$$\varepsilon = N^{-1/d}, \tag{2.4}$$

where $d = 1, 2, 3$ is the dimension of the physical space, the microscopic length scale εL and a mesoscopic length scale η satisfying $\varepsilon L \ll \eta$, where $L = \mathrm{diam}(\Omega)$ and Ω is the computational domain. To make the total energy of the system bounded independent of N, we scale interparticle forces by ε as in Panchenko *et al.* (2014).

2.2 *Mesoscale averaging and dynamics*

The first step of model reduction in Panchenko *et al.* (2011, 2014) is to use space-time averaging pioneered by Noll (1955), Hardy (1982), Murdoch & Bedeaux (1994) and Murdoch (2007). In this section, we briefly describe the averaging approach to make the exposition self-contained. For simplicity, we consider only spatial averaging and refer to Noll (1955), Hardy (1982), Murdoch & Bedeaux (1994) and Murdoch (2007) for more details.

Define the spatial average density and linear momentum

$$\bar{\rho}^{\eta}(t, \boldsymbol{x}) = \frac{M}{N} \sum_{i=1}^{N} \psi_{\eta}(\boldsymbol{x} - \boldsymbol{q}_i(t)), \tag{2.5}$$

$$\bar{\rho}^{\eta}\bar{\boldsymbol{v}}^{\eta}(t, \boldsymbol{x}) = \frac{M}{N} \sum_{i=1}^{N} \boldsymbol{v}_i(t)\psi_{\eta}(\boldsymbol{x} - \boldsymbol{q}_i(t)), \tag{2.6}$$

where $\psi(x) \geqslant 0$ is a fast decreasing window function such that $\int_{-\infty}^{\infty} \psi(x)dx = 1$. The purpose of scaling $\psi_\eta(x) = (1/\eta)\psi(x/\eta)$ is to take into account only contributions from particles such that the distance $|\boldsymbol{x} - \boldsymbol{q}_i|$ is on the order of η.

Differentiating equations (2.5) and (2.6) with respect to t, and using the ODEs (2.1) and (2.2) yields exact balance equations for $\bar{\rho}^\eta$ and $\bar{\rho}^\eta \bar{\boldsymbol{v}}^\eta$ (see Hardy, 1982; Murdoch & Bedeaux, 1994):

$$\partial_t \bar{\rho}^\eta + \operatorname{div}(\rho^\eta \bar{\boldsymbol{v}}^\eta) = 0, \tag{2.7}$$

$$\partial_t(\bar{\rho}^\eta \bar{\boldsymbol{v}}^\eta) + \operatorname{div}(\bar{\rho}^\eta \bar{\boldsymbol{v}}^\eta \otimes \bar{\boldsymbol{v}}^\eta) - \operatorname{div}\boldsymbol{T}^\eta = 0. \tag{2.8}$$

Here the system is assumed to be isolated ($\boldsymbol{f}_i^{(\mathrm{ext})} = 0$); $\boldsymbol{T}^\eta$ is the total stress written as $\boldsymbol{T}^\eta = \boldsymbol{T}^\eta_{(c)} + \boldsymbol{T}^\eta_{(\mathrm{int})}$ (see Murdoch, 2007), where

$$\boldsymbol{T}^\eta_{(c)}(t,\boldsymbol{x}) = -\sum_{i=1}^{N} m_i(\boldsymbol{v}_i - \bar{\boldsymbol{v}}^\eta(t,\boldsymbol{x},))) \otimes (\boldsymbol{v}_i - \bar{\boldsymbol{v}}^\eta(\boldsymbol{x},t))\psi_\eta(\boldsymbol{x} - \boldsymbol{q}_i) \tag{2.9}$$

is the *convective stress*, and

$$\boldsymbol{T}^\eta(t,\boldsymbol{x})_{(\mathrm{int})} = \sum_{(i,j)} \boldsymbol{f}_{ij} \otimes (\boldsymbol{q}_j - \boldsymbol{q}_i) \int_0^1 \psi_\eta(s(\boldsymbol{x} - \boldsymbol{q}_j) + (1-s)(\boldsymbol{x} - \boldsymbol{q}_i))\,\mathrm{d}s \tag{2.10}$$

is the *interaction stress*. The summation in (2.10) is over all pairs of particles (i,j) that interact with each other. The mesosystem (2.7), (2.8) is not closed because one must know the trajectories of all particles before $\boldsymbol{T}^\eta$ can be determined. Therefore, solving the exact mesosystem is no more efficient than solving the original ODEs (2.1) and (2.3).

2.3 *Integral approximations and closure*

To improve efficiency, we use the closure method introduced in Panchenko *et al.* (2011) and developed further in Panchenko *et al.* (2014). We outline it here for convenience of the reader.

Introducing interpolants $\tilde{\boldsymbol{q}}(t,\boldsymbol{X})$, $\tilde{\boldsymbol{v}}(t,\tilde{\boldsymbol{q}})$ of particle positions and velocities, one obtains the integral approximations of the average density and average momentum:

$$\begin{aligned}\bar{\rho}^\eta(t,\boldsymbol{x}) &= \frac{M}{N}\sum_{i=1}^{N}\psi_\eta(\boldsymbol{x} - \boldsymbol{q}_i(t)) \approx \frac{M}{|\Omega|}\int_\Omega \psi_\eta(\boldsymbol{x} - \tilde{\boldsymbol{q}}(t,\boldsymbol{X}))\,\mathrm{d}\boldsymbol{X} \\ &\approx \frac{M}{|\Omega|}\int_\Omega \psi_\eta(\boldsymbol{x} - \boldsymbol{y})J(t,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \frac{M}{|\Omega|}R_\eta[J](\boldsymbol{x}),\end{aligned} \tag{2.11}$$

$$\bar{\rho}^\eta\bar{\boldsymbol{v}}^\eta(t,\boldsymbol{x}) = \frac{M}{N}\sum_{i=1}^{N}\boldsymbol{v}_i(t)\psi_\eta(\boldsymbol{x} - \boldsymbol{q}_i(t)) \approx \frac{M}{|\Omega|}\int_\Omega \psi_\eta(\boldsymbol{x} - \boldsymbol{y})\tilde{\boldsymbol{v}}(t,\boldsymbol{y})J(t,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \frac{M}{|\Omega|}R_\eta[\tilde{\boldsymbol{v}}J](\boldsymbol{x}). \tag{2.12}$$

Here R_η is a convolution operator with the kernel $\psi_\eta(\boldsymbol{x})$ and the Jacobian

$$J = |\det \nabla \tilde{\boldsymbol{q}}^{-1}|, \tag{2.13}$$

describes local volume changes.

Integral approximations of $\boldsymbol{T}^{\eta}_{(\mathrm{int})}$ and $\boldsymbol{T}^{\eta}_{(c)}$ are obtained similarly (see Panchenko *et al.*, 2014):

$$\boldsymbol{T}^{\eta}_{(\mathrm{int})}(t,\boldsymbol{x}) = \frac{1}{|\Omega|^2}\int \psi_\eta(\boldsymbol{x}-\boldsymbol{R})\left(\int U'(|\boldsymbol{\rho}|)\frac{\boldsymbol{\rho}\otimes\boldsymbol{\rho}}{|\boldsymbol{\rho}|}J\left(t,\boldsymbol{R}+\frac{\varepsilon}{2}\boldsymbol{\rho}\right)J\left(t,\boldsymbol{R}-\frac{\varepsilon}{2}\boldsymbol{\rho}\right)\mathrm{d}\boldsymbol{\rho}\right)\mathrm{d}\boldsymbol{R}, \quad (2.14)$$

$$\boldsymbol{T}^{\eta}_{(c)}(t,\boldsymbol{x}) \approx -\frac{M}{|\Omega|}\int_{\Omega}\psi_\eta(\boldsymbol{x}-\boldsymbol{y})(\tilde{\boldsymbol{v}}(t,\boldsymbol{y})-\bar{\boldsymbol{v}}^{\eta}(t,\boldsymbol{x},))\otimes(\tilde{\boldsymbol{v}}(t,\boldsymbol{y})-\bar{\boldsymbol{v}}^{\eta}(\boldsymbol{x},t))J(t,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}. \quad (2.15)$$

Equations (2.14) and (2.15) can be compactly written as

$$\boldsymbol{T}^{\eta}_{(\mathrm{int})} = R_\eta S_1[J], \quad \boldsymbol{T}^{\eta}_{(c)} = R_\eta S_2[J,\tilde{\boldsymbol{v}}], \quad (2.16)$$

where S_1 is a composition of shifted multiplication $J(t,\boldsymbol{R})\to J(t,\boldsymbol{R}+(\varepsilon/2)\rho)J(t,\boldsymbol{R}-(\varepsilon/2)\boldsymbol{\rho})$ and $\boldsymbol{\rho}$-integration with the weight $|\Omega|^{-2}|\boldsymbol{\rho}|^{-1}\boldsymbol{\rho}\otimes\boldsymbol{\rho}U'(|\boldsymbol{\rho}|)$, and S_2 is a multiplication of $-M/|\Omega|J$ by the diadic product of $\tilde{\boldsymbol{v}}(t,\boldsymbol{y})-\bar{\boldsymbol{v}}^{\eta}(t,\boldsymbol{x})$ with itself. Equations (2.14) and (2.15) show that stress can be represented as an operator acting on the two fine scale functions J and $\tilde{\boldsymbol{v}}$. These functions can be approximately recovered by inverting the convolution operator R_η (see (2.11) and (2.12)). Since R_η is compact, it is not surjective. Thus, R_η^{-1} is unbounded, and straightforward inversion is ill-posed and unstable. Such problems can be approximately solved using regularization (see, e.g. Kirsch, 1996; Morozov, 1984; Tikhonov & Arsenin, 1987). Regularization consists of approximating the exact unstable problem by a parameter α dependent stable problem whose solution operator Q^{α}_{η} provides an approximation to R_η^{-1}. Computational implementation of regularized deconvolution is described in Section 3.

The closure is obtained by replacing exact J and $\tilde{\boldsymbol{v}}$ with their regularized approximations:

$$\boldsymbol{T}^{\eta}_{(\mathrm{int})} \approx \frac{|\Omega|}{M}R_\eta S_1 Q^{\alpha}_{\eta}[\bar{\rho}^{\eta}], \quad \boldsymbol{T}^{\eta}_{(c)} \approx \frac{|\Omega|}{M}R_\eta S_2\left(Q^{\alpha}_{\eta}[\bar{\rho}^{\eta}], \frac{Q^{\alpha}_{\eta}[\bar{\rho}^{\eta}\bar{\boldsymbol{v}}^{\eta}]}{Q^{\alpha}_{\eta}[\bar{\rho}^{\eta}]}\right).$$

2.4 *Test cases*

For computational testing, we use Lennard–Jones potential defined in (B.1) and two sets of the initial conditions. The problem is assumed to be periodic with period L. The initial positions in both cases are equally spaced with

$$q^0_j = \left(j-\frac{1}{2}\right)\Delta x \quad \text{where} \quad \Delta x = \frac{L}{N},\ j=1,\ldots,N. \quad (2.17)$$

First test case. The initial velocity is a one mode sine function

$$v^{(1)}(x,0) = 10^{-2}\sin\frac{2\pi x}{L}, \quad 0\leqslant x\leqslant L. \quad (2.18)$$

Second test case. The initial velocity is a continuous function that is a truncated fourth-order polynomial on $[L/3, 2L/3]$ with the double roots at $x=L/3$ and $x=2L/3$ and zero otherwise:

$$v^{(2)}(x,0) = \begin{cases} 25\left(x-\dfrac{L}{3}\right)^2\left(x-\dfrac{2L}{3}\right)^2 & \text{if } \dfrac{L}{3}\leqslant x\leqslant\dfrac{2L}{3}, \\ 0 & \text{otherwise.}\end{cases} \quad (2.19)$$

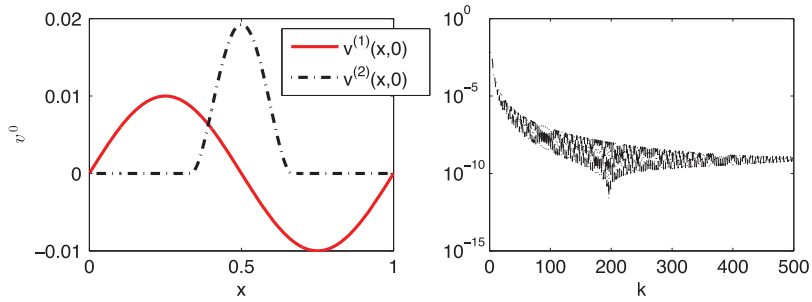

FIG. 1. Initial velocities and their Fourier spectra. Left panel shows initial velocities. Solid curve is for the first test initial velocity $v^{(1)}(x,0)$ defined in (2.18), dash-dot curve—second test initial velocity $v^{(2)}(x,0)$ given in (2.19). Right panel has discrete Fourier transform of $v^{(2)}(x,0)$.

The initial velocities are shown in the left panel of Fig. 1. The main difference between these initial velocities can be seen by looking at their Fourier spectra. The first initial velocity has only one low-frequency Fourier mode while the second initial velocity has all Fourier modes, shown in the right panel of Fig. 1, present that level off at 10^{-9} for $N = 1000$. As N increases, the tail of the Fourier spectrum levels off at a lower value. For example, for $N = 10,000$, the tail is at 10^{-12}. Because all Fourier modes are present, we expect to see more complicated non-linear dynamics and a stronger effect of high frequencies in the second case.

The 1D version of system of ODEs (2.1) and (2.2) with the initial conditions defined in (2.17)–(2.19) is solved using the Velocity Verlet method (see, for example, Griebel *et al.*, 2007) until $t = 1$ for various N. The solutions of this system, q and v, are considered to be 'exact' positions and velocities. They are used to compute the 'exact' Jacobian (2.13), the averages (2.5) and (2.6) and 'exact' stresses (2.9) and (2.10), that are later compared with respective approximations.

3. Filtered regularization methods

On a discrete level, equations (2.11) and (2.12) reduce to a linear system

$$A^\eta \boldsymbol{x} = \boldsymbol{b}, \tag{3.1}$$

where $\boldsymbol{b}$ is a known average quantity such as $(|\Omega|/M)\bar{\rho}^\eta$ and $(|\Omega|/M)\bar{\rho}^\eta\bar{\boldsymbol{v}}^\eta$ and $\boldsymbol{x}$ is either J or $\tilde{\boldsymbol{v}}J$. To produce $\boldsymbol{b}$, averages are discretized on a coarse mesh with B nodes, $B \sim 1/\eta$ and $B \ll N$, and the solutions are rendered on a finer mesh with $N' \geqslant B$ nodes. In practice, N' can vary between B and N. In our numerical experiments, we use $N' = N$.

The matrix A^η, obtained by discretizing the kernel ψ_η, has dimensions $B \times N'$ and $r = \operatorname{rank}(A^\eta) \leqslant B$. There are two difficulties associated with solving (3.1): (i) the condition number of A^η is large because of the ill-posedness and (ii) the system is underdetermined and has multiple solutions. Performing SVD of $A^\eta = \sum_{j=1}^{r} \boldsymbol{\xi}_j \sigma_j \hat{\boldsymbol{\xi}}_j^\top$ we obtain r non-zero singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$, and left and right singular vectors $\boldsymbol{\xi}_j$ and $\hat{\boldsymbol{\xi}}_j$ of length B and N', respectively, that are orthonormal in their respective spaces. In this paper, we work only with matrices A^η satisfying $\sigma_j \in (0, 1]$. This assumption holds for discretizations of all convolution kernels under consideration.

Because of ill-conditioning of A^η, a straightforward minimum l_2 norm solution of (3.1) may be highly inaccurate. To stabilize the computation, one can use regularization (see, e.g. Kirsch, 1996;

Hansen, 1987). Here we limit ourselves to filtered linear regularization methods that produce a regularized minimum l_2 norm solution

$$\boldsymbol{x}^{\alpha} = \sum_{j=1}^{r} b_j \frac{\phi(\sigma_j, \alpha)}{\sigma_j} \hat{\boldsymbol{\xi}}_j, \quad b_j = \boldsymbol{\xi}_j^{\top} \boldsymbol{b}, \tag{3.2}$$

where ϕ is a filter function. For properties of filter functions, see Kirsch (1996).

In this work, we use the truncated SVD method with the filter function

$$\phi^{\alpha} = \begin{cases} 1 & \text{if } \sigma_j \leqslant \sigma_{\alpha}, \\ 0 & \text{otherwise}, \end{cases} \tag{3.3}$$

with $\sigma_{\alpha} = 10^{-13}$. Another example of a filtered method is Landweber iteration with

$$\phi_L^n = 1 - (1 - \sigma_j^2)^{n+1}. \tag{3.4}$$

The algorithmic realizations of the above and other regularization techniques are discussed in the book by Hansen (1987).

In addition to truncating singular values below σ_{α}, we also use a spectral filtering technique on the coefficients b_j of the right-hand side $\boldsymbol{b}$, which is similar to the Fourier filtering used in Krasny (1986). The coefficients b_j below a cut-off value γ that is close to the machine precision are set to 0. We use $\gamma = 10^{-13}$. We find this filtering helpful when both σ_j and b_j are very small since the smallest singular values and their corresponding singular vectors may not be computed accurately, especially when the singular values are densely spaced as in our case (see Anderson *et al.*, 1999). By discarding the small coefficients b_j, we discard potentially inaccurate coefficients b_j/σ_j in a reconstructed solution $\boldsymbol{x}^{\alpha}$, which is turn produces more accurate approximations. Varying γ between 10^{-14} and 10^{-11} does not essentially change approximations to stresses, while setting the filter level at a higher value increases the error as expected.

4. Choice of a window function

We restrict our attention to a window function ψ satisfying the conditions:

$$\psi \text{ is non-negative, continuous and differentiable almost everywhere on the interior of its support;} \tag{4.1}$$

$$\int_{-\infty}^{\infty} \psi(x) = 1, \tag{4.2}$$

$$\psi(x) \to 0 \quad \text{as } |x| \to \infty. \tag{4.3}$$

The function $\psi(x)$ can be compactly supported or fast decreasing as Gaussian. We consider several window functions: piecewise constant (characteristic) $\psi^{(1)}(x)$, piecewise linear trapezoidal shape $\psi^{(2)}(x)$, piecewise linear triangular $\psi^{(3)}(x)$, truncated quadratic $\psi^{(4)}(x)$, truncated fourth-order polynomial $\psi^{(5)}(x)$ and Gaussian $\psi^{(6)}(x)$. The functions are chosen to have the compact support on $[-L/2, L/2]$. For the Gaussian function $\psi^{(6)}(x) = (1/\sqrt{2\pi\sigma}) \exp(-x^2/2\sigma^2)$, the standard deviation is $\sigma = L/6$, i.e. only about 0.2% of the area under $\psi^{(6)}$ is outside $[-L/2, L/2]$. Window functions are depicted in Fig. 2. The shifted and rescaled window function $\psi_{\eta}(\boldsymbol{x} - \boldsymbol{q}_i(t))$ defines the average properties of the particles

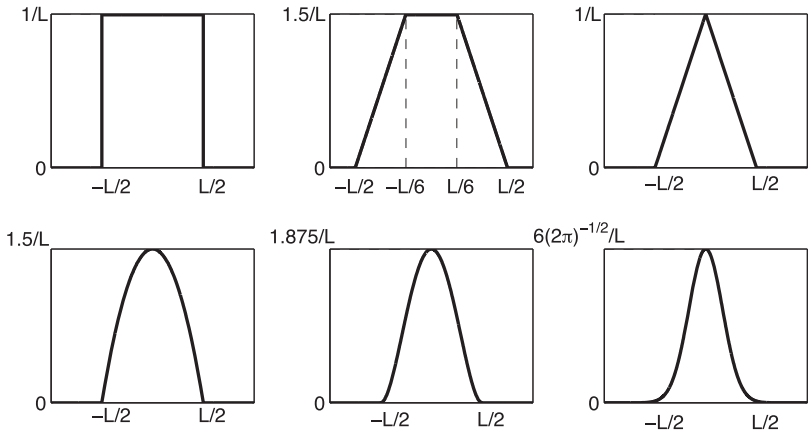

FIG. 2. Window functions. Top left: piecewise constant (characteristic) $\psi^{(1)}(x)$; top middle: piecewise linear trapezoidal shape $\psi^{(2)}(x)$; top right: triangular function $\psi^{(3)}(x)$; bottom left: truncated quadratic $\psi^{(4)}(x)$; bottom middle: truncated fourth-order polynomial $\psi^{(5)}(x)$; bottom right: Gaussian $\psi^{(6)}(x)$.

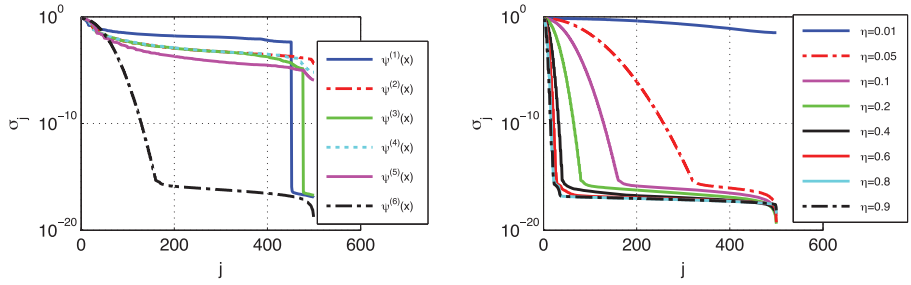

FIG. 3. Left panel: singular values σ_j for window functions $\psi^{(i)}$, $i = 1, \ldots, 6$, with $\eta = 0.1$, $N = 1000$ and $B = 500$. Right panel: singular values for various η, $0.01 \leqslant \eta \leqslant 0.9$, $\psi^{(6)}$, $N = 1000$, $B = 500$. In both cases, the graphs of singular values with $N = 100,00$ and $B = 500$ are similar.

in the vicinity of $\boldsymbol{x}$. In the case of the characteristic function $\psi^{(1)}$, the contributions of all particles within a fixed distance of $\boldsymbol{x}$ are weighted equally. A better choice of a window function is obtained when the weight decreases to zero as the distance between the particle and the observation point increases. This adds another 'desired' characteristics of a window function (see Root *et al.* (2003)):

$$\psi(x) \text{ has a maximum at } x = 0. \tag{4.4}$$

Clearly, functions $\psi^{(1)}(x)$ and $\psi^{(2)}(x)$ do not satisfy this property, whereas the rest of the functions do. A window function can have a different degree of differentiability, ranging from piecewise constant (as characteristic function) to infinitely many times differentiable as Gaussian. The order of differentiability affects the speed of decay of singular values to zero. In the left panel of Fig. 3, we plot singular values of $\psi^{(i)}(x)$, $i = 1, \ldots, 6$, for $\eta = 0.1$, $N = 1000$ and $B = 500$. Results are similar for $N = 10,000$. Since we have to invert an operator with kernel ψ_η, increasing smoothness of the kernel increases ill-posedness of the inverse problem. On a discrete level, the situation is somewhat different. First, discretization itself regularizes the inverse problem (see Kirsch, 1996). Second, truncating SVD provides additional

regularization (see, for example, Hansen, 1987). It is natural to ask how smoothness of the window function effects the reconstruction quality. The least smooth function is a characteristic function $\psi^{(1)}(x)$, followed by piecewise linear trapezoidal $\psi^{(2)}(x)$ and triangular function $\psi^{(3)}(x)$. Then we have a truncated quadratic $\psi^{(4)}(x)$, truncated fourth-order polynomial $\psi^{(5)}(x)$ and Gaussian $\psi^{(6)}(x)$. The singular values of the corresponding matrices A^{η} are given in the left panel of Fig. 3. As expected, the decay of spectral coefficients is fastest for the Gaussian and slowest for the characteristic function. A sharp drop of singular values for the characteristic function $\psi^{(1)}(x)$ and triangular function $\psi^{(3)}(x)$ indicates that the discrete problem is numerically rank-deficient. The same can be said about the Gaussian that has only about a third of singular values above the machine zero.

If the convolution kernel ψ_{η} and function f are periodic or extended periodically, then R_{η} is a circular convolution operator (see Mallat, 2009). If both f and $R_{\eta}[f]$ are discretized on the same grid, then the eigenvectors of the circular convolution operator are the discrete complex exponentials and the eigenvalues are Fourier modes of the window function ψ_{η} (see, e.g. Mallat, 2009). Therefore, eigenvectors corresponding to the smallest eigenvalues of the R_{η} carry information about high-frequency component of a solution. If $R_{\eta}[f]$ and f are sampled at different scales, we have to deal with singular value decomposition instead of eigenvalue decomposition. In this case, singular vectors corresponding to the largest singular values would have contribution not only from low frequencies, but also from some high frequencies. Nevertheless, singular vectors corresponding to smallest singular values will be the most oscillatory and we can still think that singular vectors corresponding to the smallest singular values represent the oscillatory part of a solution. In this regard, the fact that singular values for all window functions but Gaussian decay slowly indicates that solutions corresponding to these window functions will have high-frequency component present. Hence, if there is a numerical error in computation of singular values and singular vectors, especially in those corresponding to the smallest singular values, there may be a significant error in computation of a solution. On the other hand, singular values of a scaled Gaussian window function decay very fast. This means that a fewer singular values can be used to represent a solution. Thus, using a Gaussian is more efficient for large systems of ODEs if the associated error is comparable with other choices of window functions.

In order to approximate the stress, one needs to approximate first the exact microscopic positions and velocities (solutions of microscopic initial value problem (2.1)–(2.3)). Their approximations are obtained by generating deconvolution approximations of J and $\tilde{v}J$, recovering microscopic velocities $\tilde{v}$ by dividing $\tilde{v}J$ by J and reconstructing microscopic positions $\tilde{q}$ from J using (2.13). Then these approximations of microscopic positions and velocities are used to compute approximate stresses, thus, closing the system. Each of the above steps carries some error, and smaller deconvolution error does not necessarily yield to smaller overall error in approximating stresses $\boldsymbol{T}^{\eta}_{(\mathrm{int})}$ and $\boldsymbol{T}^{\eta}_{(c)}$.

Next we investigate the effect of a choice of a window function on the quality of the stress approximation. Solutions of the underlying microscopic initial value problem (2.1)–(2.3) serve a role of 'exact' solutions: microscopic positions and velocities are generated at different times by solving this system numerically and used to compute average density, average velocity and exact stresses. These averages are then employed to recover microscopic positions and velocities, and approximate the stresses. In Fig. 4, we compare the l^{∞}-relative errors in approximation of the convective $T^{\eta}_{(c)}$ and interaction $T^{\eta}_{(\mathrm{int})}$ stresses in the first test case. Clearly, the characteristic function $\psi^{(1)}$ has the worst performance. The absolute error (not shown here) in $T^{\eta}_{(c)}$ is at most 10^{-7} or 13% and it is by two orders of magnitude larger than with other window functions. The smallest error (2–3%) is achieved with the fourth-order truncated polynomial $\psi^{(5)}$ and Gaussian $\psi^{(6)}$. The other functions produce errors in between. The results in the right panel of Fig. 4 suggest that the approximation of $T^{\eta}_{(\mathrm{int})}$ also has the largest error with $\psi^{(1)}$. The error with the piecewise linear $\psi^{(2)}$ function reaches 75–100% at some times. The smallest within 0.5% error

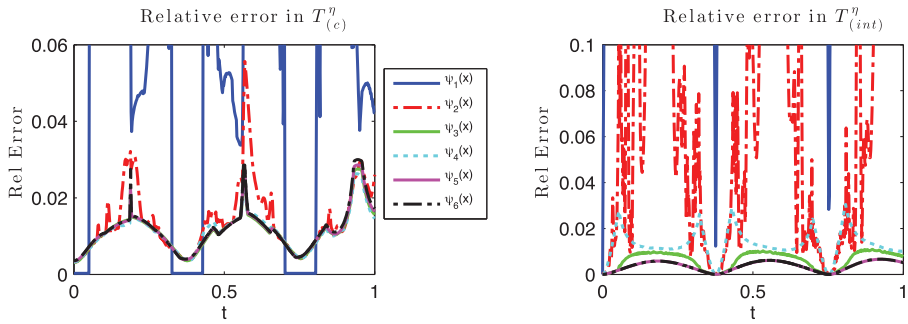

FIG. 4. Effect of a choice of ψ on the error in approximation of stresses in the first test case with $N = 1000$, $B = 500$ and $\eta = 0.1$. Left panel: $T^{\eta}_{(c)}$. Right panel: $T^{\eta}_{(\text{int})}$.

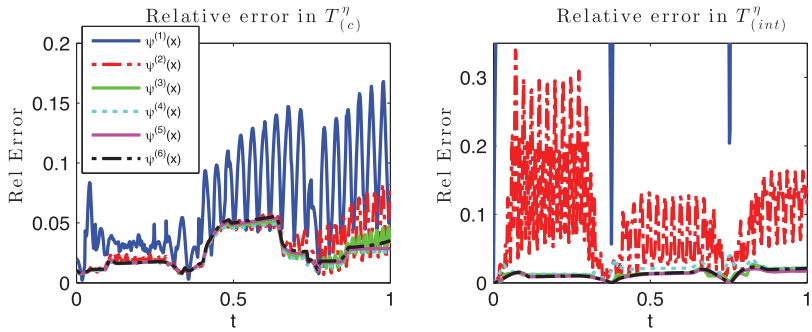

FIG. 5. The effect of a choice of ψ on the error in approximation of stresses in the second test case. The relative error is shown for the case with $N = 1000$, $B = 500$ and $\eta = 0.1$. Left panel: $T^{\eta}_{(c)}$. Right panel: $T^{\eta}_{(\text{int})}$.

is again produced by $\psi^{(5)}$ and $\psi^{(6)}$. Functions $\psi^{(3)}$ and $\psi^{(4)}$ give slightly bigger 1–3% error. The results for the second test case are shown in Fig. 5. The findings are similar as in the first test case: the worst performance is by $\psi^{(1)}$, followed by $\psi^{(2)}$; the smallest error is obtained with $\psi^{(5)}$ and $\psi^{(6)}$, the rest of the functions give intermediate results. The overall errors in the second (high-frequency) case are slightly larger than in the first case (low frequency), which is expected since we filter out high-frequency content of solutions.

The worst behaviour of $\psi^{(1)}$, followed by $\psi^{(2)}$, could be explained by the fact that function $\psi^{(1)}$ is not even continuous everywhere that violates condition (4.1), while $\psi^{(2)}$ is only piecewise continuous. Moreover, both functions do not have a strict maximum at $x = 0$, thus violating the condition (4.4). For the rest of the window functions considered here, all conditions (4.1)–(4.4) are satisfied. The results suggest that the smoother window function is, the better stress approximation can be achieved, despite the fact that smoothness increases ill-conditioning of the problem.

The experiments in this section show that the most accurate approximation of the stresses is obtained using either truncated fourth-order polynomial function $\psi^{(5)}$ or Gaussian $\psi^{(6)}$. The absolute error is typically smaller with $\psi^{(6)}$ but the relative error is sometimes slightly smaller with $\psi^{(5)}$, so the difference in performance of these functions is not significant. However, with $\psi^{(5)}$ and, for example, with $N = 10,000$ and $B = 500$, one has to use all 500 singular values and singular vectors since all singular values are above the threshold $\sigma_\alpha = 10^{-13}$ (they decay to at most 10^{-6}, see the left panel of Fig. 3),

and only 147 singular values with $\psi^{(6)}$, which is more efficient for large systems of particles. For this reason, in what follows we fix Gaussian window function $\psi^{(6)}$ and study the effect of other parameters such as averaging width η and scale separation on stress approximation.

5. Choice of mesoscale resolution parameter η

The parameter η in (2.5) and (2.6) determines the size of the averaging region and the amount of high-frequency filtering. For smaller η, there is less damping of high frequency content of the solution, while using a larger η produces smoother and smaller averages. The latter is discussed in Section 6. As η increases, the singular values of the corresponding matrix A^η decay at a higher rate. This can be seen in the right panel of Fig. 3 where we show the singular values for $N = 1000$ and $B = 500$ and $\eta = 0.01, 0.05, \ldots, 0.9$. (The results with $N = 10,000$ and $B = 500$ are similar since they are more sensitive to the choice of B and not of N.)

On the one hand, for larger η, the average contains less high-frequency information, and it may be more difficult to reconstruct the same microscopic solution from increasingly smoothed averages. This may increase the error in computation of Q^α_η. On the other hand, computation of the stress involves another averaging (the outer layer R_η in (1.1)) that may decrease the overall error even if J and $\tilde{\boldsymbol{v}}$ are recovered more poorly. In this section, we study the cumulative effect of these two competing tendencies on the overall relative error in a stress approximation. We analyse the error in reconstruction of the microscopic Jacobian and velocity, as well as the error in the subsequent stress approximation. We fix $N = 1000$, $B = 500$ and the window function $\psi^{(6)}(x)$, and vary η between 10^{-2} and 0.9.

5.1 *First test case*

In this case, the errors in the Jacobian and velocity approximation, shown in Fig. 6, are essentially independent of η, though there is a slight dependence on η for times after $t = 0.8$. Both absolute and relative errors in the Jacobian oscillate in a quasiperiodic manner. The amplitude of these oscillations increases slightly (by less than 0.01%) during the simulation time. The behaviour of the error can be connected to the evolution of the total computed energy of the system shown in left panel of Fig. 7. As can be seen from the graph, the energy is not completely conserved. Instead it oscillates periodically and deviates from its value at $t = 0$ (the true energy of the system) by at most 0.05%. In the regions where the energy starts deviating from its initial value, the error in Jacobian reconstruction increases. When the energy comes back to its initial value, the error in Jacobian also decreases. The error in velocity

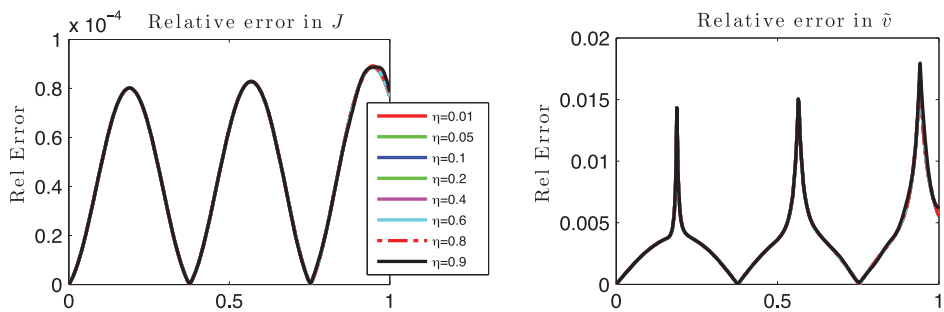

FIG. 6. The effect of a choice of the resolution parameter η on the error in approximation of J (left panel) and $\tilde{v}$ (right panel) in the first test case with $N = 1000$ and $B = 500$.

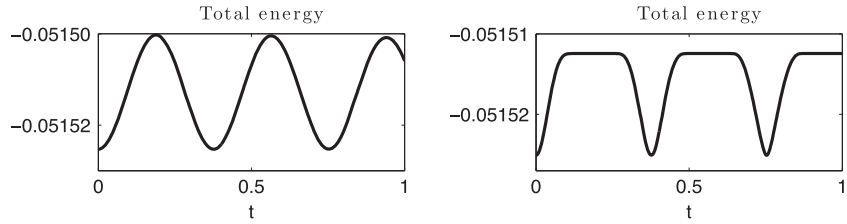

FIG. 7. Evolution of the total energy for $N = 1000$, $B = 500$ and $\eta = 0.1$. Left panel: the first test case. Right panel: the second test case.

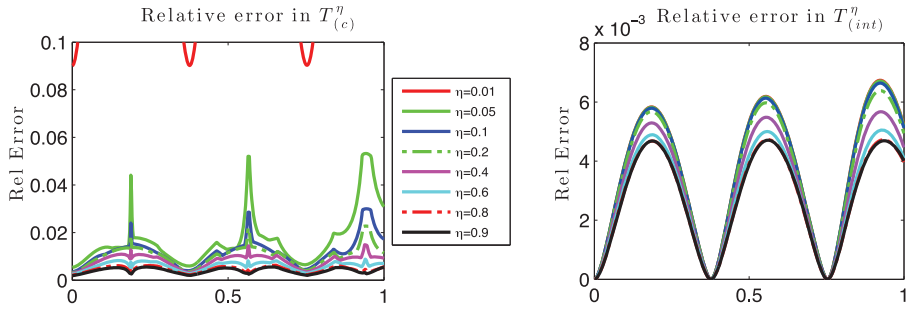

FIG. 8. The effect of the choice of the resolution parameter η on the error in approximation of stresses in the first test case with $N = 1000$ and $B = 500$. Left panel: $T^{\eta}_{(c)}$. Right panel: $T^{\eta}_{(\text{int})}$.

reconstruction, shown in the right panel of Fig. 6, has a more non-linear dynamics but similarly to the error in the Jacobian, the error in velocity increases or decreases in phase with oscillations of the total computed energy. During the simulation time, the error in velocity increases by at most 1.8%.

Unlike approximation of the Jacobian and velocity, the errors in approximation of both convective and interaction stresses depend on η. While the absolute error in $T^{\eta}_{(c)}$ is not monotonic in η at all times, it is typically larger for larger η. However, the relative error, shown in the left panel of Fig. 8, decreases monotonically as η increases. More specifically, the error with $\eta = 0.01$ is the largest and varies between 9 and 100%, whereas it is the smallest with $\eta = 0.9$ and it does not exceed 0.6%. Both absolute and relative errors in approximation of $T^{\eta}_{(\text{int})}$ oscillate in time and decrease monotonically as η increases. The latter is shown in the right panel of Fig. 8. We also note that the dependence on η is weaker for the interaction stress and as η varies from $\eta = 0.01$ to $\eta = 0.9$, the error changes slightly within 0.7%.

5.2 *Second test case*

The simulation results with various η are presented in Figs 9 and 10. Figure 9 indicates that the reconstruction of both Jacobian and velocity gets worse as η and time t increase. However, the situation with stress approximation is different. The error in approximating the interaction stress is the smallest with the largest $\eta = 0.9$ used. It is below 1% during the entire simulation time. As for the convective stress, the error does not depend on η monotonically. It also varies non-linearly in time. Overall, this error does not exceed 9% for all considered values of η except of the smallest $\eta = 0.01$. It should be noted that the convective stress is much smaller than the interaction stress (by at least two orders of magnitude), so relatively large errors in the convecting stress do not contribute much to the errors in the entire stress.

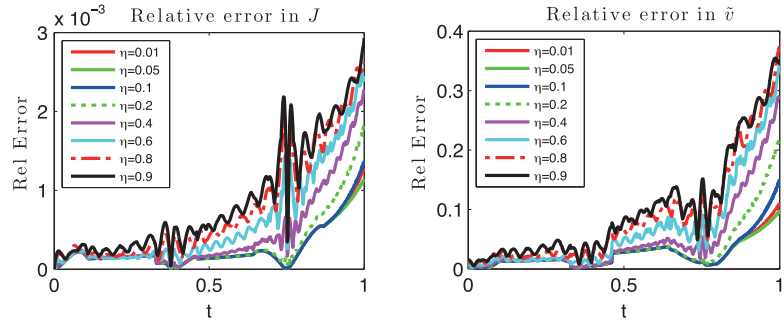

FIG. 9. Effect of a choice of η on the error in reconstruction of J (left panel) and $\tilde{v}$ (right panel) in the second test case with $N = 1000$, $B = 500$, $\eta = 0.01,\ 0.05,\ 0.1, \ldots, 0.9$.

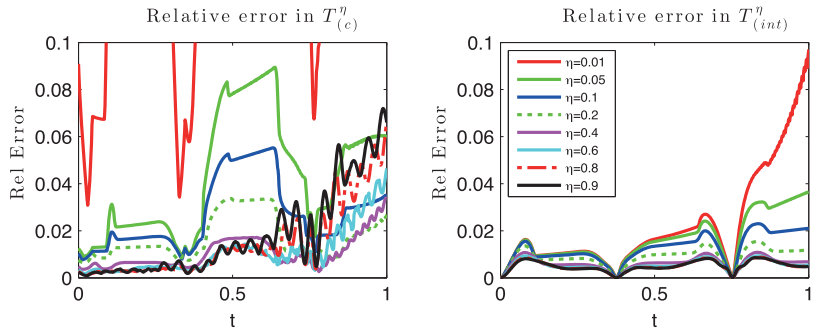

FIG. 10. Effect of a choice of η on the error in stress approximation in the second test case with $N = 1000$, $B = 500$, $\eta = 0.01,\ 0.05,\ 0.1, \ldots, 0.9$. Left panel: $T^{\eta}_{(c)}$. Right panel: $T^{\eta}_{(\text{int})}$.

Hence, for all considered values of η, the error in approximation of J stays very small during the simulation period, though it grows slowly with time (at most linearly). Despite of larger error in approximation of $\tilde{v}$, we still get very good stress approximation and overall its quality increases with η.

6. Spectral evolution of averages and stresses

As was mentioned in Section 5, parameter η determines the size of the averaging window and the amount of high-frequency filtering. With larger η, the averages are smoother and smaller. This can be seen by considering the Fourier transform of an average as follows.

Recall a typical one-particle dynamical function used in statistical mechanics

$$g_{sm}(t, \boldsymbol{x}) = \sum_{i=1}^{N} g(\boldsymbol{q}_i(t), \boldsymbol{v}_i(t)) \delta(\boldsymbol{x} - \boldsymbol{q}_i(t)), \tag{6.1}$$

where δ is delta-distribution. The Fourier transform of g_{sm} with respect to $\boldsymbol{x}$ is

$$\hat{g}_{sm}(t, \boldsymbol{\xi}) = \sum_{i=1}^{N} g(\boldsymbol{q}_i(t), \boldsymbol{v}_i(t))\, \mathrm{e}^{\mathrm{i}\boldsymbol{\xi}\cdot\boldsymbol{q}_i}. \tag{6.2}$$

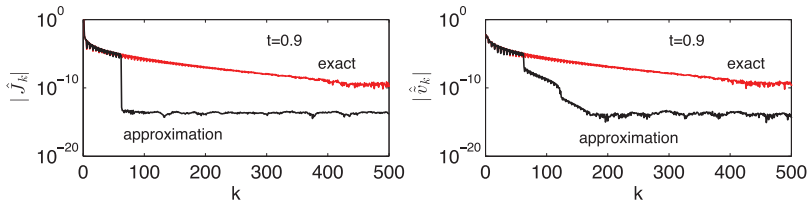

FIG. 11. Discrete Fourier coefficients of J (left panel) and $\tilde{v}$ (right panel) and their approximations in the second test case with $N = 10,000$, $B = 500$, $\eta = 0.1$. The logarithm of coefficients' amplitudes is plotted against wavenumber k at $t = 0.9$.

Now compare this with a windowed spatial average

$$\bar{g}(t,\boldsymbol{x}) = \sum_{i=1}^{N} g(\boldsymbol{q}_i(t),\boldsymbol{v}_i(t))\psi_\eta(\boldsymbol{x}-\boldsymbol{q}_i(t)), \tag{6.3}$$

and the corresponding Fourier transform

$$\hat{\bar{g}}(t,\boldsymbol{x}) = \sum_{i=1}^{N} g(\boldsymbol{q}_i(t),\boldsymbol{v}_i(t))\hat{\psi}(\eta\boldsymbol{\xi})\,\mathrm{e}^{\mathrm{i}\boldsymbol{\xi}\cdot\boldsymbol{q}_i} = \hat{\psi}(\eta\boldsymbol{\xi})\hat{g}_{sm}, \tag{6.4}$$

where $\hat{\psi}$ denotes the Fourier transform of ψ. Thus, the Fourier transform of $\bar{g}$ is obtained from $\hat{g}_{sm}$ by low-pass filtering (multiplication by $\hat{\psi}(\eta\boldsymbol{\xi})$). If $\lim_{|\boldsymbol{k}|\to\infty}|\hat{\psi}(\boldsymbol{k})| = 0$, which is true for any L_1 function by Riemann–Lebesgue Lemma, then $\hat{\psi}(\eta\boldsymbol{\xi})$ converges to 0 for each $\boldsymbol{\xi} \neq 0$ as $\eta \to \infty$. Since $\hat{g}_{sm}$ does not depend on η, equation (6.4) implies $\hat{\bar{g}}$ approaches 0. The rate of decay of $\hat{\psi}$ increases with smoothness as ψ. Thus, increasing η produces progressively more filtered versions of $\hat{g}_{sm}$.

Flexibility afforded by varying η is convenient for studying large scale behaviour of averages. For example, in statistical physics, the derivation of hydrodynamical equations and computation of fluid viscosity by Green–Kubo formulas (see, e.g. Berne, 1977) employs truncated Taylor expansions of $\hat{g}_{sm}$ at $\boldsymbol{\xi} = 0$. Estimating the error of these approximations for large $\boldsymbol{\xi}$ may be difficult, while multiplication by $\hat{\psi}(\eta\boldsymbol{\xi})$ makes analysis easier. Another useful feature of (6.4) is the possibility to adjust the size of the low-frequency neighbourhood of interest by changing η.

To investigate quality of our approximations in the Fourier space, we analyse Fourier spectra of the exact Jacobian, velocity, stresses and their approximations. We consider only the second test case because the initial velocity in this case has full spectrum unlike the first test case. Since the approximation tends to become worse at later times, we show the relevant spectra at the 'worst case scenario' time $t = 0.9$. Figure 11 depicts the spectra of J and $\tilde{v}$ (gradually decaying to 10^{-10} curves) and their approximations (curves that drop to about the machine precision 10^{-15}). The left panel of Fig. 11 indicates that we only capture about 70 first low-frequency modes of the Jacobian. Similarly, the first 70 modes of the velocity are well reconstructed, while modes between 70 and 180 have much smaller amplitudes than in the exact velocity. While spectra of both Jacobian and velocity are not very accurate, spectral approximations of both convective and interaction stresses are very good (see Fig. 12). Both exact stresses have only low-frequency components (110 for the convective stress and only 70 for the interaction) and all these modes are captured perfectly! This demonstrates that it is not necessary to recover higher frequency modes of the Jacobian and velocity in order to approximate stress accurately.

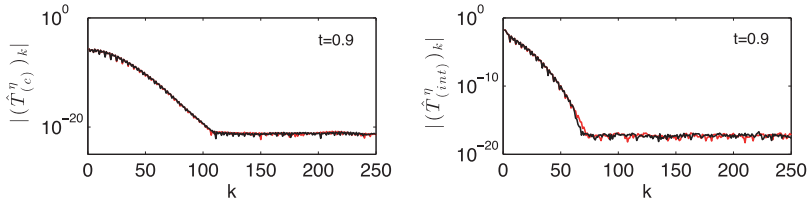

FIG. 12. Discrete Fourier coefficients of $T^{\eta}_{(c)}$ (left panel) and $T^{\eta}_{(\text{int})}$ (right panel) in the second test case with $N = 10,000$, $B = 500$, $\eta = 0.1$ and $t = 0.9$. The red/lighter curves are exact solutions, black/darker—approximations.

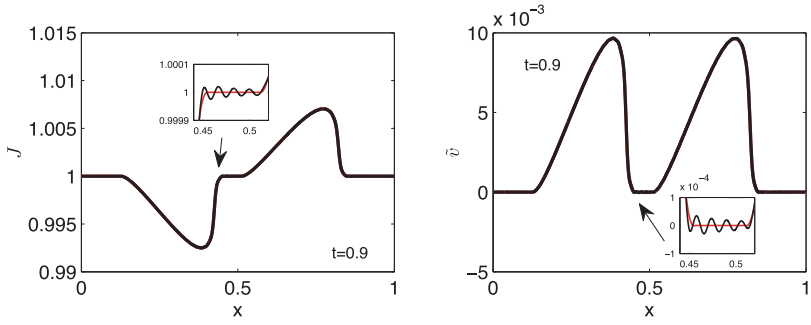

FIG. 13. Reconstruction of J (left panel) and $\tilde{v}$ (right panel) in the second test case with $N = 10,000$, $B = 500$, $\eta = 0.1$. Exact (red/lighter curves) and approximate (black/darker, slightly oscillatory in some regions, curves) solutions are shown at $t = 0.9$.

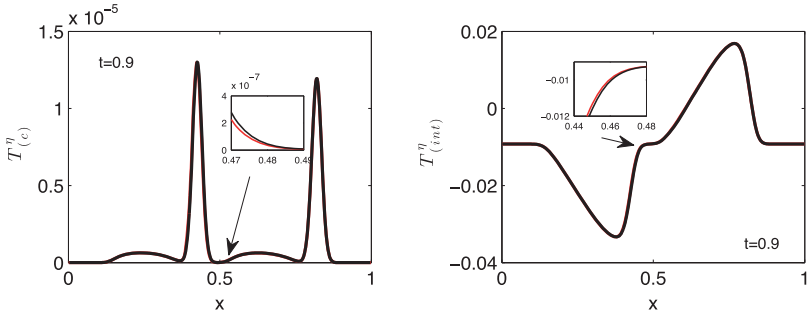

FIG. 14. Reconstruction of $T^{\eta}_{(c)}$ (left panel) and $T^{\eta}_{(\text{int})}$ (right panel) in the second test case with $N = 10,000$, $B = 500$ and $\eta = 0.1$. Exact (red/lighter curves) and approximate (black/darker curves) solutions are shown at $t = 0.9$.

Loss of accuracy in using truncated spectra for deconvolution often leads to Gibbs phenomenon. It is indeed present in both Jacobian and velocity reconstructions shown in Fig. 13. The amplitude of Gibbs ripples seem to increase with η. In contrast, the approximations to stresses do not suffer from Gibbs oscillations as can be seen from Fig. 14. Gibbs phenomenon is typical for solutions of linear systems using a truncated SVD approach (see Boyd, 2002; Bruno, 2003; Bruno *et al.*, 2007; Boyd & Ong, 2009). In Lyon (2012), Gibbs phenomenon is controlled by using Sobolev smoothing. In our case, smoothing is done naturally by averaging present in the stress approximation (see (1.1)).

7. Scale separation with fixed η, B and varying N

In this section, we investigate how the scale separation, i.e. ratio B to N, affects the accuracy of reconstruction of the Jacobian, velocity and stress. We use two test initial conditions as before with Gaussian window function $\psi^{(6)}(x)$, $\eta = 0.1$, $B = 500$ and $N = 1000$, 2000, 5000 and 10,000.

In the first test case, results of which are presented in Figs 15 and 16, we observe that the error in reconstruction of the Jacobian (shown in the left panel of Fig. 15) decreases as the scale separation increases (N in this case), while the error in velocity reconstruction (right panel of Fig. 15) does not depend on N and it is less accurate than the Jacobian. Both errors oscillate in time and their oscillatory dynamics is related to the total computed energy oscillations depicted in Fig. 7 in the left panel, i.e. they oscillate in phase with energy oscillations. Similarly to the approximation of the Jacobian, the error in approximation of the convective stress (left panel of Fig. 16) decreases as the scale separation increases. Reconstruction of the interaction stress, depicted in the right panel of Fig. 16, does not depend of N like velocity and it is by about one order accurate than the convective stress approximation. During the entire simulation time, the error in the convective stress approximation does not exceed 3%, while the error in the interaction stress (recall that the interaction stress values are dominant over the convective stress) is less than 0.6%.

In the second test case, reconstruction of both Jacobian and velocity depends on N in a non-monotonic manner (see Fig. 17). The intervals where the error with a larger N is larger or smaller

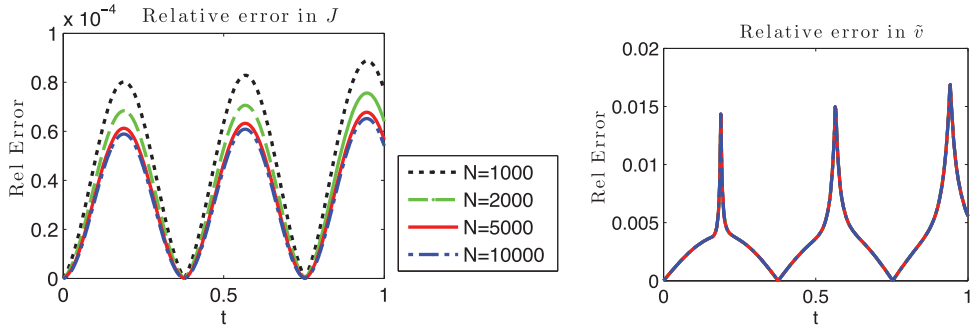

FIG. 15. Effect of the scale separation on the reconstruction of J (left panel) and $\tilde{v}$ (right panel) in the first test case with $\eta = 0.1$, $\psi^{(6)}(x)$, $B = 500$ and $N = 1000$, 2000, 5000 and 10,000.

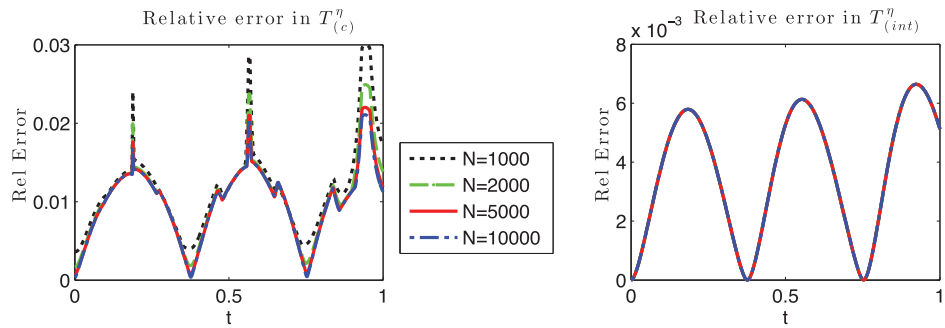

FIG. 16. Effect of the scale separation on the error in approximation of $T^{\eta}_{(c)}$ (left panel) and $T^{\eta}_{(\mathrm{int})}$ (right panel) in the first test case with $\eta = 0.1$, $\psi^{(6)}(x)$, $B = 500$ and $N = 1000$, 2000, 5000 and 10,000.

alternate in phase with oscillations of the total computed energy. Similarly to the first test case, the Jacobian is approximated more accurately than velocity. The error in the approximation of the convective stress is shown in the left panel of Fig. 18. It is not monotonic in N, oscillates quasiperiodically, and does not exceed 6% during the simulation time. The error in the more dominant interaction stress, shown in the right panel of Fig. 18, does not depend on N and stays below 2.5%.

Overall, we see that the approximation of the stress becomes more accurate as the scale separation increases, especially when the total computed energy is close to its exact value. The approach performs well for both low-frequency (first test case) and high-frequency (second test case) problems with the error only slightly greater in the second test case.

8. Error estimates

8.1 *Estimates for filtered regularization methods*

In practice, the right-hand side of (3.1) is known imprecisely, so instead of the exact $\boldsymbol{b}$, one has an approximate vector $\boldsymbol{b}^{\delta}$. The computed regularized solution is thus

$$\boldsymbol{x}^{\alpha,\delta} = \sum_{j=1}^{D} b_j^{\delta} \frac{\phi(\sigma_j,\alpha)}{\sigma_j} \hat{\boldsymbol{\xi}}_j. \tag{8.1}$$

Our goal is to estimate the error $\| \boldsymbol{x} - \boldsymbol{x}^{\alpha,\delta} \|_p$, in some vector p-norm. Usually $p = 2$, but here we assume $p \in [1, \infty)$. By triangle inequality,

$$\begin{aligned} \| \boldsymbol{x} - \boldsymbol{x}^{\alpha,\delta} \|_p &\leqslant \left\| \sum_{j=1}^{D} b_j \frac{1-\phi(\sigma_j,\alpha)}{\sigma_j} \hat{\boldsymbol{\xi}}_j \right\|_p + \left\| \sum_{j=1}^{D} (b_j - b_j^{\delta}) \frac{\phi(\sigma_j,\alpha)}{\sigma_j} \hat{\boldsymbol{\xi}}_j \right\|_p \\ &\leqslant C(\hat{\boldsymbol{\xi}}, p) \left(\left(\sum_{j=1}^{D} |b_j|^p \frac{|1-\phi(\sigma_j,\alpha)|^p}{\sigma_j^p} \right)^{1/p} + \left(\sum_{j=1}^{D} |b_j - b_j^{\delta}|^p \frac{|\phi(\sigma_j,\alpha|^p}{\sigma_j^p} \right)^{1/p} \right). \end{aligned} \tag{8.2}$$

The constant $C(\hat{\boldsymbol{\xi}}, p)$ depends only on p and the components of the singular vectors $\hat{\boldsymbol{\xi}}_j$.

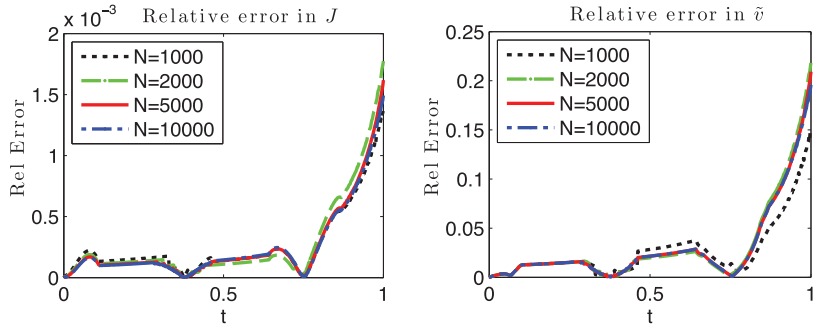

FIG. 17. Effect of the scale separation on the reconstruction of J (left panel) and $\tilde{v}$ (right panel) in the second test case with $\eta = 0.1$, $\psi^{(6)}(x)$, $B = 500$ and $N = 1000$, 2000, 5000 and 10,000.

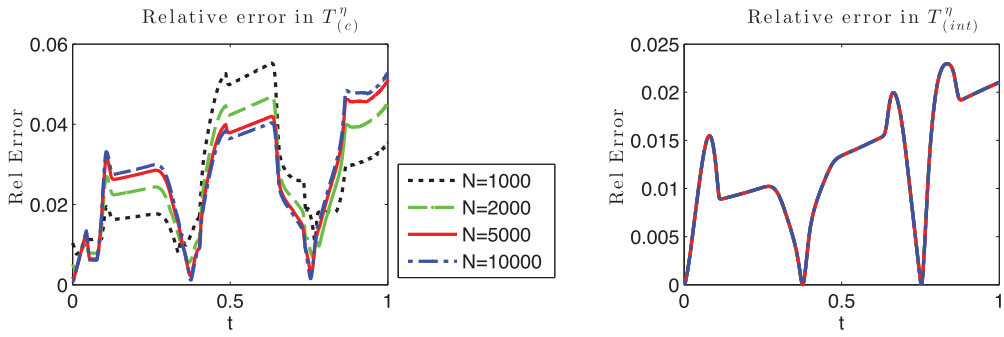

FIG. 18. Effect of the scale separation on the error in approximation of $T^{\eta}_{(c)}$ (left panel) and $T^{\eta}_{(\text{int})}$ (right panel) in the second test case with $\eta = 0.1$, $\psi^{(6)}(x)$, $B = 500$ and $N = 1000$, 2000, 5000 and 10,000.

Writing $|b_j|/\sigma_j = |x_j|$, we see that the first term on the very right of (8.2) is bounded by

$$C(\hat{\boldsymbol{\xi}}, p) \max_j |x_j| \left(\sum_{j=1}^{D} |1 - \phi(\sigma_j, \alpha)|^p \right)^{1/p} \leqslant C_1(\hat{\boldsymbol{\xi}}, p) \, \| \boldsymbol{x} \|_\infty \left(\int_0^{D+1} |1 - \phi(f(t), \alpha)|^p \, \mathrm{d}t \right)^{1/p}, \quad (8.3)$$

where we introduced a function $f(t) : [0, \infty) \to (0, 1]$ that interpolates between singular values:

$$f(j) = \sigma_j, \quad f(0) = 1, \quad \lim_{t \to \infty} f(t) = 0.$$

The function f is chosen to be continuous, non-negative and strictly decreasing. To see that the integral is larger than the corresponding sum, note that the sum is the left-endpoint Riemann sum for the integral, and the function under the integral is increasing.

Similarly,

$$C(\hat{\boldsymbol{\xi}}, p) \left(\sum_{j=1}^{D} |b_j - b_j^\delta|^p \frac{|\phi(\sigma_j, \alpha)|^p}{\sigma_j^p} \right)^{1/p} \leqslant C_1(\hat{\boldsymbol{\xi}}, p) \, \| \boldsymbol{b} - \boldsymbol{b}^\delta \|_\infty \left(\int_0^{D+1} \frac{|\phi(f(t), \alpha)|^p}{f(t)^p} \, \mathrm{d}t \right)^{1/p}. \quad (8.4)$$

Combining (8.2)–(8.4), we have

$$\begin{aligned} \| \boldsymbol{x} - \boldsymbol{x}^{\alpha,\delta} \|_p \leqslant\ & C_1(\hat{\boldsymbol{\xi}}, p) \, \| \boldsymbol{x} \|_\infty \| \, 1 - \phi(f(t), \alpha) \, \|_{L^p(0, D+1)} \\ & + C_1(\hat{\boldsymbol{\xi}}, p) \, \| \boldsymbol{b} - \boldsymbol{b}^\delta \|_\infty \| \, \phi(f(t), \alpha) f^{-1}(t) \, \|_{L^p(0, D+1)} \, . \end{aligned} \quad (8.5)$$

By definition of ϕ, the first term can be made arbitrarily small by choosing α small enough. In the second term, as $\alpha \to 0$, the norm of $\phi(f(t), \alpha) f^{-1}(t)$ typically increases. To control the second term, we need $\boldsymbol{b} - \boldsymbol{b}^\delta$ to be small. This is typical of the error estimates available in the literature. Our inequalities differ from the standard ones because we use a p-norm for the error, and ∞-norms for $\boldsymbol{x}$ and $\boldsymbol{b} - \boldsymbol{b}^\delta$. The standard estimates use 2-norms of $\boldsymbol{x} - \boldsymbol{x}^{\alpha,\delta}$, $\boldsymbol{x}$, $\boldsymbol{b} - \boldsymbol{b}^\delta$, and what is essentially the ∞-norm for the ϕ- and f-dependent terms. Depending on the actual ϕ and $f(t)$, our approach can yield tighter bounds. Improvement occurs if, loosely speaking, the integrals involving ϕ, f in (8.3) and (8.4) are smaller than maximal values of the integrands.

Finally, we note that using Hölder inequality with exponents $q, q', q^{-1} + (q')^{-1} = 1$ in the right-hand side of (8.2) results in the estimates

$$\begin{aligned} \| \boldsymbol{x} - \boldsymbol{x}^{\alpha,\delta} \|_p \leqslant C_2(\hat{\boldsymbol{\xi}}, p, q) \, \| \boldsymbol{x} \|_{pq} \| \, 1 - \phi(f(t), \alpha) \, \|_{L^{pq'}(0,D+1)} \\ + C_2(\hat{\boldsymbol{\xi}}, p, q) \, \| \boldsymbol{b} - \boldsymbol{b}^{\delta} \|_{pq} \| \, \phi(f(t), \alpha) f^{-1}(t) \, \|_{L^{pq'}(0,D+1)} \, . \end{aligned} \tag{8.6}$$

8.2 *Error in the interaction stress approximation*

The purpose of this section is to estimate the difference between the exact integral representation of the interaction stress $\boldsymbol{T}^{\eta}_{(\mathrm{int})}$ in (2.14), and its closed form approximation

$$\bar{\boldsymbol{T}}^{\eta}_{(\mathrm{int})}(t, \boldsymbol{x}) = \frac{1}{|\Omega|^2} \int \psi_\eta(\boldsymbol{x} - \boldsymbol{R}) \left(\int U'(|\boldsymbol{\rho}|) \frac{\boldsymbol{\rho} \otimes \boldsymbol{\rho}}{|\boldsymbol{\rho}|} Q_\eta[\bar{\rho}^\eta] \left(t, \boldsymbol{R} + \frac{\varepsilon}{2}\boldsymbol{\rho} \right) Q_\eta[\bar{\rho}^\eta] \left(t, \boldsymbol{R} - \frac{\varepsilon}{2}\boldsymbol{\rho} \right) \mathrm{d}\boldsymbol{\rho} \right) \mathrm{d}\boldsymbol{R}. \tag{8.7}$$

Since estimates will be local in time, we will suppress the dependence on t in the remainder of this section. Define the error

$$E(\boldsymbol{x}) = \boldsymbol{T}^{\eta}_{(\mathrm{int})}(\boldsymbol{x}) - \bar{\boldsymbol{T}}^{\eta}_{(\mathrm{int})}(\boldsymbol{x}).$$

Next, introduce the abbreviated notation $J^+ = J(\boldsymbol{R} + (\varepsilon/2)\boldsymbol{\rho})$, $J^- = J(\boldsymbol{R} - (\varepsilon/2)\boldsymbol{\rho})$, and

$$Q_\eta[\bar{\rho}^\eta]^+ = Q_\eta[\bar{\rho}^\eta] \left(\boldsymbol{R} + \frac{\varepsilon}{2}\boldsymbol{\rho} \right), \quad Q_\eta[\bar{\rho}^\eta]^- = Q_\eta[\bar{\rho}^\eta] \left(\boldsymbol{R} - \frac{\varepsilon}{2}\boldsymbol{\rho} \right),$$

and denote

$$\Phi(\boldsymbol{\rho}) = U'(|\boldsymbol{\rho}|) \frac{\boldsymbol{\rho} \otimes \boldsymbol{\rho}}{|\boldsymbol{\rho}|}. \tag{8.8}$$

This function is smooth and can be assumed compactly supported on a shell $D = \{\boldsymbol{\rho} : c_1 \leqslant |\boldsymbol{\rho}| \leqslant c_2\}$ where $c_1 > 0$. With these notation, using an elementary identity

$$a_1 a_2 - b_1 b_2 = a_1(a_2 - b_2) + a_2(a_1 - b_1) - (a_1 - b_1)(a_2 - b_2),$$

we have

$$\begin{aligned} J^+ J^- - Q_\eta[\bar{\rho}^\eta]^+ Q_\eta[\bar{\rho}^\eta]^- &= J^+(J^- - Q_\eta[\bar{\rho}^\eta]^-) + (J^+ - Q_\eta[\bar{\rho}^\eta]^+) Q_\eta[\bar{\rho}^\eta]^- \\ &= J^+(J^- - Q_\eta[\bar{\rho}^\eta]^-) + J^-(J^+ - Q_\eta[\bar{\rho}^\eta]^+) \\ &\quad - (J^- - Q_\eta[\bar{\rho}^\eta]^-)(J^+ - Q_\eta[\bar{\rho}^\eta]^+). \end{aligned} \tag{8.9}$$

Now

$$\begin{aligned} |E(\boldsymbol{x})| &\leqslant |\Omega|^{-2} \sup |\psi_\eta| \iint |\Phi(\boldsymbol{\rho})| |J^+ J^- - Q_\eta[\bar{\rho}^\eta]^+ Q_\eta[\bar{\rho}^\eta]^-|(\boldsymbol{R}, \boldsymbol{\rho}) \, \mathrm{d}\boldsymbol{\rho} \, \mathrm{d}\boldsymbol{R} \\ &\leqslant |\Omega|^{-2} \sup |\psi_\eta| \sup_D |\Phi| \iint (|J^+||J^- - Q_\eta[\bar{\rho}^\eta]^-| + |J^-||J^+ - Q_\eta[\bar{\rho}^\eta]^+| \\ &\quad + |J^+ - Q_\eta[\bar{\rho}^\eta]^+||J^- - Q_\eta[\bar{\rho}^\eta]^-|) \, \mathrm{d}\boldsymbol{\rho} \, \mathrm{d}\boldsymbol{R}. \end{aligned}$$

Changing variables in the last integral to $\boldsymbol{y}_1 = \boldsymbol{R} + (\varepsilon/2)\boldsymbol{\rho}$, $\boldsymbol{y}_2 = \boldsymbol{R} - (\varepsilon/2)\boldsymbol{\rho}$, and observing that the Jacobian of this transformation is ε^{-d} and that the quantities marked by $^+$ (respectively, by $^-$) depend

only on $\boldsymbol{y}_1$ (respectively, on $\boldsymbol{y}_2$), we find

$$|E(\boldsymbol{x})| \leqslant \varepsilon^{-d}|\Omega|^{-2}C(\psi_\eta, \Phi)\left[2\int_\Omega |J(\mathbf{y}_1)|\,\mathrm{d}\mathbf{y}_1 \int_\Omega |J - Q_\eta[\bar{\rho}^\eta]|(\mathbf{y}_2)\,\mathrm{d}\mathbf{y}_2 + \left(\int_\Omega |J - Q_\eta[\bar{\rho}^\eta]|(\mathbf{y})\,\mathrm{d}\mathbf{y}\right)^2\right].$$

Suppose now that J and $Q_\eta[\bar{\rho}^\eta]$ are given by their discretizations on the fine mesh. Thus, we can assume that they are piecewise constant functions having values J_j, $Q_\eta[\bar{\rho}^\eta]_j$ on the sets $S_j \subset \Omega, j = 1, 2, \ldots, N$ of measure $|\Omega|/N$. In this way, J and $Q_\eta[\bar{\rho}^\eta]$ can be identified with, respectively, the vectors $\boldsymbol{J} = (J_1, J_2, \ldots, J_N)^\top$ and $\boldsymbol{Q} = (Q_1, Q_2, \ldots, Q_N)^\top$.

Suppose that there exists a constant M such that

$$J \leqslant M. \tag{8.10}$$

With this,

$$\begin{aligned}|E(\boldsymbol{x})| &\leqslant \varepsilon^{-d}|\Omega|^{-2}C(\psi_\eta, \Phi)\left(\frac{|\Omega|}{N}\right)^2 \left(2M\,\|\boldsymbol{J} - \boldsymbol{Q}\|_1 + \|\boldsymbol{J} - \boldsymbol{Q}\|_1^2\right)\\ &= C(\psi_\eta, \Phi)\left(2M\,\|\boldsymbol{J} - \boldsymbol{Q}\|_1 + \|\boldsymbol{J} - \boldsymbol{Q}\|_1^2\right).\end{aligned} \tag{8.11}$$

The last equality holds since $N = \varepsilon^{-1/d}$. The norms are vector 1-norms that can be estimated using (8.6) with $\boldsymbol{x} = \boldsymbol{J}$, $\boldsymbol{x}^{\alpha,\delta} = \boldsymbol{Q}$ and $\boldsymbol{b}$ representing a discretization of $\bar{\rho}^\eta$.

The results of this section can be summarized in the following theorem.

THEOREM 8.1 Suppose that

(i) $J(t, \boldsymbol{x})$ satisfies (8.10) uniformly in t;

(ii) Φ defined in (8.8) is bounded;

(iii) $J(t, \boldsymbol{x}) = \sum_{j=1}^N J_j(t, \boldsymbol{x})\chi_j(\boldsymbol{x})$, $Q_\eta[\bar{\rho}^\eta] = \sum_{j=1}^N Q_j(t, \boldsymbol{x})\chi_j(\boldsymbol{x})$, where χ_j are characteristic functions of sets S_j such that $\bigcup_{j=1}^N = \Omega$, $S_j \cap S_k = \emptyset$ if $j \neq k$, and $|S_j| = N^{-1}|\Omega|$. Then the error

$$E = \boldsymbol{T}^\eta_{(\mathrm{int})} - \bar{\boldsymbol{T}}^\eta_{(\mathrm{int})},$$

satisfies

$$|E(t, \boldsymbol{x})| \leqslant \sup|\psi_\eta| \sup|\Phi|(2M\,\|\boldsymbol{J} - \boldsymbol{Q}\|_1 + \|\boldsymbol{J} - \boldsymbol{Q}\|_1^2),$$

where $\boldsymbol{J} = (J_1, J_2, \ldots, J_N)^\top$, $\boldsymbol{Q} = (Q_1, Q_2, \ldots, Q_N)^\top$.

The estimates for the error in $\boldsymbol{T}^\eta_{(c)}$ can be derived similarly, but would require more technical work because of the triple product structure of the integrand. It is also worth noting that, while a pointwise bound on J can be reasonably expected, similar bounds on the velocity $\tilde{v}$ would blow up as $\varepsilon \to 0$. Consequently, estimating of the error $\boldsymbol{T}^\eta_{(c)} - \bar{\boldsymbol{T}}^\eta_{(c)}$ is left to future work.

9. Conclusions

We study the numerical performance of the regularized deconvolution closure introduced in Panchenko *et al.* (2011, 2014). The closure method consists of the following. The average density and linear momentum are written as convolutions acting on respective fine scale functions: J and $J\tilde{\boldsymbol{v}}$, where J

is the Jacobian of the inverse deformation map, and $\tilde{\mathbf{v}}$ is a particle velocity interpolant. These functions are approximately recovered by applying a regularized deconvolution to the averages. To construct the deconvolution operator, we use the theory of ill-posed problems. Closure is obtained by using these deconvolution approximations in the exact flux equations. This gives constitutive equations that express stress in terms of the average density and velocity. The exact stress is thus approximated by a sum of terms that have the 'convolution sandwich' structure: they combine the convolution operator, a non-linear composition or a product type operator and the deconvolution operator. The resulting constitutive equations are non-linear and non-local.

We show that the closure has the 'convolution sandwich' structure and involves the convolution operator, a non-linear composition or a product-type operator and the deconvolution operator, which makes the problem non-linear. The approximation quality depends on a choice of the window function ψ used to define averages, magnitude of scale separation and values of the resolution and regularization parameters. Because of the non-linearity of the problem, the error estimates tend to be too pessimistic. Therefore, we conduct numerical experiments to determine the dependence of the error on the above parameters. Since the Fourier spectrum of velocity seems to have a strong effect on the error, we consider two sets of initial conditions. In the first test case, the initial velocity is a low-frequency mode sine function, while in the second test the initial velocity has full Fourier spectrum. The initial positions in both cases are equally spaced.

We study window functions of different smoothness, varying from piecewise continuous to infinitely smooth. Among these functions, the Gaussian, an infinitely smooth function, provides the best overall performance despite the fact that the corresponding integral deconvolution problem has the highest degree of ill-posedness. Numerical deconvolution amounts to solving an ill-conditioned linear system. We use a truncated SVD method with an additional spectral filtering of the right-hand side. Filtering helps to reduce the effect of error that is present in every standard numerical SVD routine.

The choice of the resolution parameter η affects the size of the averaging window and the amount of high-frequency filtering in the computed averages. Larger values of η produce smoother and smaller averages, thus causing the reconstruction to deteriorate. This tendency is counteracted by the presence of the convolution operator in the stress equations. We find that the overall error in the stress approximation tends to decrease with increasing η. Therefore, it is not necessary to have very good reconstruction of the Jacobian and velocity to have good approximations of the stresses. This 'self-correcting' property is a noteworthy feature of the deconvolution closure. Another method to increase scale separation is to vary the number of particles while keeping η fixed. The results in this case are less clear-cut compared with the case of increasing η. However, at times when the computed total energy is close to its exact value, the error in the stress decreases with increasing scale separation.

The deconvolution error estimates derived in the paper are applicable to general SVD-based filtered regularization methods (see, e.g. Kirsch, 1996). We also obtain error estimates for the interaction stress (the part of the total stress induced by interparticle forces). We believe that similar estimates can be also obtained for the remaining convective stress, but such estimates will be developed elsewhere.

Acknowledgements

The author L.L.B. would like to acknowledge the availability of computational resources at Idaho National Laboratory (INL).

References

Admal, N. C. & Tadmor, E. B. (2010) A unified interpretation of stress in molecular systems. *J. Elasticity*, **100**, 63–143.

Admal, N. C. & Tadmor, E. B. (2011) Stress and heat flux for arbitrary multibody potentials: a unified framework. *J. Chem. Phys.*, **134**,184106.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A. & Sorensen, D. (1999) *LAPACK Users' Guide*, 3rd edn. Philadelphia, PA: SIAM.

Berne, B. J. (1977) Projection operator techniques. *Modern Theoretical Chemistry: Statistical Mechanics of Time Dependent Processes*, New York: Plenum, pp. 233–257.

Berselli, L. C., Iliescu, T. & Layton, W. J. (2006) *Mathematics of Large Eddy Simulation of Turbulent Flows*. New York: Springer.

Boyd, J. P. (2002) A comparison of numerical algorithms for Fourier extension of the first, second, and third kinds. *J. Comput. Phys.*, **178**, 118–160.

Boyd, J. P. & Ong, J. R. (2009) Exponentially-convergent strategies for defeating the Runge phenomenon for the approximation of non-periodic functions. I. Single-interval schemes. *Commun. Comput. Phys.*, **5**, 484–497.

Bruno, O. P. (2003) Fast, high-order, high-frequency integral methods for computational acoustics and electromagnetics. *Topics in Computational Wave Propagation*. Lecture Notes in Computational Science and Engineering 31. Berlin: Springer, pp. 43–82.

Bruno, O. P., Han, Y. & Pohlman, M. M. (2007) Accurate, high-order representation of complex three-dimensional surfaces via Fourier continuation analysis. *J. Comput. Phys.*, **227**, 1094–1125.

Charlotte, M. & Truskinovsky, L. (2012) Lattice dynamics from a continuum viewpoint. *J. Mech. Phys. Solids*, **60**, 1508–1544.

Du, Q., Gunzburger, M., Lehoucq, R. B. & Zhou, K. (2012) Analysis of the volume-constrained peridynamic Navier equation of linear elasticity. *J. Elasticity*, doi:10.1007/s10659–012–9418–x.

E, W., Ren, W. & Vanden-Eijnden, E. (2009) A general strategy for designing seamless multiscale methods. *J. Comput. Phys.*, **228**, 5437–5453.

Eringen, A. C. (1976) Nonlocal field theories. *Continuum Physics*, vol. 4. New York: Academic Press.

Evans, D. J. & Morriss, G. (2008) *Statistical Mechanics of Non-equilibrium Liquids*, 3rd edn. Cambridge: Cambridge University Press.

Griebel, M., Knapek, S. & Zumbusch, G. (2007) *Numerical Simulation in Molecular Dynamics: Numerics, Algorithms, Parallelization, Applications*. Texts in Computational Science and Engineering. Berlin: Springer.

Hansen, P. C. (1987) *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA: SIAM.

Hardy, R. J. (1982) Formulas for determining local properties in molecular-dynamics simulations: shock waves. *J. Chem. Phys.*, **76**, 622–628.

Irving, J. H. & Kirkwood, J. G. (1950) The statistical theory of transport processes IV. The equations of hydrodynamics. *J. Chem. Phys.*, **18**, 817–829.

Kim, T.-Y., Rebholz, L. & Fried, E. (2012) A deconvolution enhancement of the Navier–Stokes-$\alpha\beta$-model. *J. Comp. Phys.*, **231**, 4015–4027.

Kirsch, A. (1996) *An Introduction to the Mathematical Theory of Inverse Problems*. New York: Springer.

Krasny, R. (1986) A study of singularity formation in a vortex sheet by the point-vortex approximation. *J. Fluid Mech.*, **167**, 65–93.

Kunin, I. (1982) *Elastic Media with Microstructure, V.I (One dimensional models)*. Berlin: Springer.

Layton, W. J. & Rebholz, L. G. (2012) *Approximate Deconvolution Models of Turbulence*. Berlin: Springer.

Lehoucq, R. B. & Sears, M. P. (2011) Statistical mechanical foundation of the peridynamic nonlocal continuum theory: energy and momentum conservation laws. *Phys. Rev. E*, **84**, 031112.

Lyon, M. (2012) Sobolev smoothing of SVD-based Fourier continuations. *Appl. Math. Lett.*, **25**, 2227–2231.

Mallat, S. (2009) *A Wavelet Tour of Signal Processing: The Sparse Way*. Burlington, MA: Academic Press.

Morozov, V. A. (1984) *Methods for Solving Incorrectly Posed Problems*. New York: Springer.

Murdoch, A. I. (2007) A critique of atomistic definitions of the stress tensor. *J. Elasticity*, **88**, 113–140.

Murdoch, A. I. & Bedeaux, D. (1994) Continuum equations of balance via weighted averages of microscopic quantities. *Proc. R. Soc. Lond. A*, **445**, 157–179.

Murdoch, A. I. & Bedeaux, D. (1996) A microscopic perspective on the physical foundations of continuum mechanics—part I: macroscopic states, reproducibility, and macroscopic statistics, at prescribed scales of length and time. *Int. J. Eng. Sci.*, **34**, 1111–1129.

Murdoch, A. I. & Bedeaux, D. (1997) A microscopic perspective on the physical foundations of continuum mechanics—part II: a projection operator approach to the separation of reversible and irreversible contributions to macroscopic behaviour. *Int. J. Eng. Sci.*, **35**, 921–949.

Noll, W. (1955) Der Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik. *J. Ration. Mech. Anal.*, **4**, 627–646.

Panchenko, A., Barannyk, L. L. & Cooper, K. (2014) Deconvolution closure for mesoscopic continuum models of particle systems, *J. Mech. Phys. Solids*, *Preprint*, arXiv:1109.5984.

Panchenko, A., Barannyk, L. L. & Gilbert, R. P. (2011) Closure method for spatially averaged dynamics of particle chains. *Nonlinear Anal. Real World Appl.*, **12**, 1681–1697.

Panchenko, A. & Tartakovsky, A. (2014) Discrete models of fluids: spatial averaging, closure, and model reduction. *SIAM J. Appl. Math.*, **74**, 477–515.

Root, S., Hardy, R. J. & Swanson, D. R. (2003) Continuum predictions from molecular dynamics simulations: shock waves. *J. Chem. Phys.*, **118**, 3161–3165.

Silling, S. A. (2000) Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids*, **48**, 175–209.

Silling, S. & Lehoucq, R. B. (2010) Peridynamic theory of solid mechanics. *Adv. Appl. Mech.*, **44**, 73–168.

Tadmor, E. B. & Miller, R. E. (2011) *Modeling Materials. Continuum, Atomistic and Multiscale Techniques*. Cambridge: Cambridge University Press.

Tartakovsky, A., Panchenko, A. & Ferris, K. (2011) Dimension reduction method for ODE fluid models. *J. Comput. Phys.*, **230**, 8554–8572.

Tikhonov, A. N. & Arsenin, V. Y. (1987) *Solutions of Ill-Posed Problems*. New York: Wiley.

Zimmerman, J. A., Jones, R. E. & Templeton, J. A. (2010) A material frame approach for evaluating continuum variables in atomistic simulations. *J. Comput. Phys.*, **229**, 2364–2389.

Appendix A. Window functions

A window function ψ is chosen to define a mesoscale average. This function has to satisfy several conditions: be non-negative, fast decreasing, compactly supported (we also consider non-compactly supported functions like Gaussian), continuous and differentiable almost everywhere in the interior of its domain and $\int_{\infty}^{\infty} \psi(x) = 1$. We use functions $\psi^{(i)}(x)$, $i = 1, \ldots, 6$ of different order of smoothness starting from the characteristic function $\psi^{(1)}(x)$ that is discontinuous at $x = \pm L/2$ up to infinitely many times differentiable Gaussian $\psi^{(6)}(x)$. The window functions are defined in (A.1)–(A.6) and plotted in Fig. 2.

$$\psi^{(1)}(x) = \begin{cases} 1/L & \text{if } |x| \leqslant L/2, \\ 0 & \text{otherwise}, \end{cases} \tag{A.1}$$

$$\psi^{(2)}(x) = \begin{cases} 1/(2L) & \text{if } |x| \leqslant L/2, \\ -2/L(x - 3L/2) & \text{if } L/2 < x \leqslant 3L/2, \\ -2/L(x + 3L/2) & \text{if } -3L/2 \leqslant x < -L/2, \\ 0 & \text{otherwise}, \end{cases} \tag{A.2}$$

$$\psi^{(3)}(x) = \begin{cases} -4/L^2(x - L/2) & \text{if } 0 \leqslant x \leqslant L/2, \\ 4/L^2(x + L/2) & \text{if } -L/2 \leqslant x < 0, \\ 0 & \text{otherwise}, \end{cases} \tag{A.3}$$

$$\psi^{(4)}(x) = \begin{cases} -6/L^3(x^2 - L^2/4) & \text{if } |x| < L/2, \\ 0 & \text{otherwise,} \end{cases} \tag{A.4}$$

$$\psi^{(5)}(x) = \begin{cases} 30/L^5(x^2 - L^2/4)^2 & \text{if } |x| \leqslant L/2, \\ 0 & \text{otherwise,} \end{cases} \tag{A.5}$$

$$\psi^{(6)}(x) = \frac{6}{L\sqrt{2\pi}} \exp\left(-\frac{18x^2}{L^2}\right). \tag{A.6}$$

Appendix B. Lennard–Jones potential

The dynamics of particles considered in this paper is governed by Lennard–Jones potential defined as

$$U(\xi) = 4\epsilon \left[\left(\frac{\sigma}{\xi}\right)^{12} - \left(\frac{\sigma}{\xi}\right)^{6} \right], \tag{B.1}$$

with the potential well depth $\epsilon = 0.025$. The same potential but with $\epsilon = 0.25$ was used in Panchenko *et al.* (2014). The magnitude of ϵ defines how strong interaction between particles is. The potential is zero at the distance given by σ and reaches its minimum at the distance $h = 2^{1/6}\sigma$ at which particles are in equilibrium. For smaller distances $\xi < h$, the potential is repulsive whereas for $\xi > h$ it is mildly attractive. When the distance $\xi > 2.5h$, the force is very small and we set it to zero to speed up computations. This truncation of the potential tail typically takes into account three particles on each side from a current particle. Truncating at larger distances slightly decreases deviations of the total computed energy from its exact value (at $t = 0$) and as a result slightly decreases the error in the approximation of the stresses, more so in the interaction stress.