

A Treecode Algorithm for 3D Stokeslets and Stresslets

Lei Wang^{1,*}, Svetlana Tlupova² and Robert Krasny³

¹ *Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA*

² *Department of Mathematics, Farmingdale State College, SUNY, Farmingdale, NY 11735, USA*

³ *Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA*

Received 27 August 2018; Accepted (in revised version) 13 February 2019

Abstract. The Stokeslet and stresslet kernels are commonly used in boundary element simulations and singularity methods for slow viscous flow. Evaluating the velocity induced by a collection of Stokeslets and stresslets by direct summation requires $\mathcal{O}(N^2)$ operations, where N is the system size. The present work develops a treecode algorithm for 3D Stokeslets and stresslets that reduces the cost to $\mathcal{O}(N \log N)$. The particles are divided into a hierarchy of clusters and well-separated particle-cluster interactions are computed by a far-field Cartesian Taylor approximation. The terms in the approximation are contracted to promote efficient computation. Serial and parallel results display the performance of the treecode for several test cases. In particular the method has relatively simple structure and low memory usage and this enhances parallel efficiency for large systems.

AMS subject classifications: 65D99, 76D07

Key words: Stokeslet, stresslet, fast summation, treecode, Taylor approximation.

1 Introduction

The slow steady flow of an incompressible viscous fluid is governed by the Stokes equations,

$$\nabla^2 \mathbf{u} - \nabla p = 0, \quad (1.1a)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (1.1b)$$

where \mathbf{u} is the fluid velocity, p is the pressure and the viscosity is taken to be unity. Many applications in fluid dynamics are modeled as particle interactions in Stokes flow,

*Corresponding author.

Emails: wang256@uwm.edu (Wang), tlupovs@farmingdale.edu (Tlupova), krasny@umich.edu (Krasny)

including for example particle-laden fluid jets [28], vibrations in microfluidic crystals [5], cilia- and flagella-driven flows [11, 33], free-surface flows of liquid drops [27] and vesicle flows [36], among others. The Stokeslet and stresslet kernels are fundamental solutions of the Stokes equations given in 3D (up to a numerical prefactor) by

$$S_{ij}(\mathbf{x}, \mathbf{y}) = \frac{\delta_{ij}}{|\mathbf{x} - \mathbf{y}|} + \frac{(x_i - y_i)(x_j - y_j)}{|\mathbf{x} - \mathbf{y}|^3}, \quad (1.2a)$$

$$T_{ijl}(\mathbf{x}, \mathbf{y}) = \frac{(x_i - y_i)(x_j - y_j)(x_l - y_l)}{|\mathbf{x} - \mathbf{y}|^5}, \quad (1.2b)$$

where δ_{ij} is the Kronecker delta, $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$ and indices $i, j, l = 1 : 3$ represent Cartesian coordinates. The Stokeslet and stresslet kernels are commonly used in boundary element simulations and singularity methods for slow viscous flow [29].

The i th component of the velocity induced by a set of Stokeslets and stresslets is

$$u_i(\mathbf{x}^m) = \sum_{\substack{n=1 \\ n \neq m}}^N S_{ij}(\mathbf{x}^m, \mathbf{x}^n) f_j^n + \sum_{\substack{n=1 \\ n \neq m}}^N T_{ijl}(\mathbf{x}^m, \mathbf{x}^n) h_j^n v_l^n, \quad i = 1, 2, 3, \quad (1.3)$$

where \mathbf{x}^m is a target particle, \mathbf{x}^n is a source particle, f_j^n is a force weight, h_j^n is a dipole weight and v_l^n are the components of a unit normal vector to a surface. Note that the Stokeslet term has an implicit sum over $j = 1 : 3$ and the stresslet term has an implicit sum over $j, l = 1 : 3$; for clarity in some places below these sums will be written out explicitly. Eq. (1.3) is written for the case in which the targets and sources coincide, but it is straightforward to handle problems where they are disjoint.

Evaluating the velocity (1.3) for $m = 1 : N$ by direct summation requires $\mathcal{O}(N^2)$ operations, which is prohibitively expensive when N is large. The same issue arises for interacting point masses, point charges and point vortices and many fast summation methods have been developed to reduce the cost, including particle-mesh methods [13, 20], the Fast Multipole Method (FMM) [19] and treecodes [4]. These methods reduce the operation count to $\mathcal{O}(N \log N)$ or $\mathcal{O}(N)$ in principle, while introducing approximations. The FMM and treecode use multipole expansions of particle clusters (near-field and far-field for the FMM, but only far-field for the treecode), while particle-mesh methods interpolate the particle strengths to a grid where often the fast Fourier transform (FFT) is used to compute the sum.

A number of these fast summation methods have been developed in the context of Stokes flow including a particle-mesh Ewald technique [31, 32] and a pre-corrected FFT method [38]. Several extensions of the FMM have also been developed for Stokes flow [7, 17, 24, 37, 39]. In one implementation [35], the Stokeslet and stresslet sums are decomposed into several terms which are computed by the FMM for Coulomb interactions [15]. The kernel-independent FMM [25, 26, 39] has been applied to simulate swimming microorganisms [30] using regularized Stokeslets [8, 9]. Recently the Spectral Ewald (SE) method was developed for Stokes flow using Gaussian spreading functions [1, 2].

These developments significantly improve the capability of fast summation methods for Stokes flow, but it is important to investigate different approaches and understand their properties, especially as new types of many-core computing platforms become available with new challenges to the goal of maintaining parallel efficiency. In particular, as the relative cost of memory access rises in comparison with arithmetic operations, it is worthwhile to investigate fast summation methods with different memory requirements and communication patterns.

In this context the present work contributes a treecode algorithm that reduces the cost of the Stokeslet and stresslet sums (1.3) to $\mathcal{O}(N \log N)$. In a treecode, the particles are divided into a hierarchy of clusters and well-separated particle-cluster interactions are computed by a far-field approximation, while nearby interactions are computed directly [4]. The original treecode used monopole approximations, but later work starting with the FMM showed the advantage of using higher-order multipole approximations [19]. Here we retain the structure of the treecode, but we employ higher-order Cartesian Taylor series for the far-field approximation [3, 10, 12, 21–23]. We derive novel expressions for the Taylor coefficients of the Stokeslet and stresslet kernels in terms of the Taylor coefficients of the Coulomb potential and the far-field approximation is contracted for efficient evaluation, following an approach used for direct summation in the FMMLIB3D code [14, 17]. A key feature of the proposed treecode is its relatively simple structure and low memory usage, which together can enhance parallel efficiency.

The paper is organized as follows. Section 2 explains the particle-cluster interaction on which the treecode is based. Section 3 derives expressions for the Taylor coefficients of the Stokeslet and stresslet kernels. Section 4 presents an efficient method for computing the particle-cluster approximations. Section 5 describes the treecode algorithm in detail. Section 6 presents serial and parallel numerical results showing the performance of the treecode in terms of accuracy, efficiency and memory usage. A summary is given in Section 7.

2 Particle-cluster interactions

We start by expressing the velocity (1.3) in terms of particle-cluster interactions. Assume the particles have been divided into a set of clusters $\{C\}$ (the procedure will be described below). Then write the Stokeslet part of the velocity as a sum of particle-cluster interactions,

$$\begin{aligned} u_i^{sto}(\mathbf{x}^m) &= \sum_{n=1}^N S_{ij}(\mathbf{x}^m, \mathbf{x}^n) f_j^n \\ &= \sum_C \sum_{\mathbf{y}^n \in C} S_{ij}(\mathbf{x}^m, \mathbf{y}^n) f_j^n = \sum_C u_i^{sto}(\mathbf{x}^m, C), \end{aligned} \quad (2.1)$$

where

$$u_i^{sto}(\mathbf{x}^m, C) = \sum_{\mathbf{y}^n \in C} S_{ij}(\mathbf{x}^m, \mathbf{y}^n) f_j^n \quad (2.2)$$

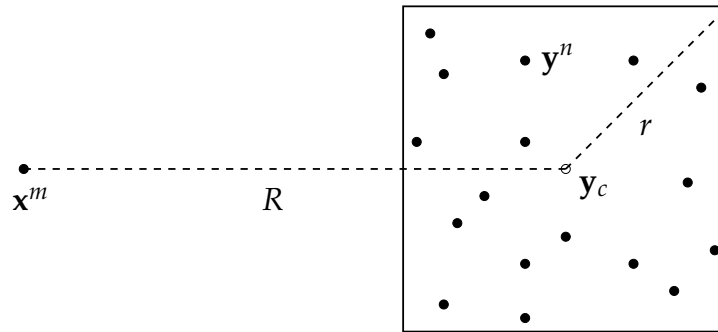


Figure 1: Particle-cluster interaction between a target particle \mathbf{x}^m and a cluster of source particles $C = \{\mathbf{y}^n\}$, with cluster center \mathbf{y}_c , cluster radius r , particle-cluster distance R .

is the interaction between a target particle \mathbf{x}^m and a cluster of source particles $C = \{\mathbf{y}^n\}$. Fig. 1 depicts the particle-cluster interaction, showing also the cluster center \mathbf{y}_c , cluster radius $r = \max_n |\mathbf{y}^n - \mathbf{y}_c|$ and particle-cluster distance $R = |\mathbf{x}^m - \mathbf{y}_c|$. The stresslet part of the velocity is treated similarly.

If the particle \mathbf{x}^m and cluster C are not well-separated (the criterion is given later), then direct summation is used in (2.2). If they are well-separated, then we use a far-field approximation given by Taylor expanding the Stokeslet $S_{ij}(\mathbf{x}, \mathbf{y})$ about the cluster center $\mathbf{y} = \mathbf{y}_c$,

$$\begin{aligned}
 u_i^{sto}(\mathbf{x}^m, C) &= \sum_{\mathbf{y}^n \in C} S_{ij}(\mathbf{x}^m, \mathbf{y}^n) f_j^n \\
 &= \sum_{\mathbf{y}^n \in C} \sum_{\|\mathbf{k}\|=0}^{\infty} \frac{1}{\mathbf{k}!} D_{\mathbf{y}}^{\mathbf{k}} S_{ij}(\mathbf{x}^m, \mathbf{y}_c) (\mathbf{y}^n - \mathbf{y}_c)^{\mathbf{k}} f_j^n \\
 &= \sum_{\|\mathbf{k}\|=0}^{\infty} \frac{1}{\mathbf{k}!} D_{\mathbf{y}}^{\mathbf{k}} S_{ij}(\mathbf{x}^m, \mathbf{y}_c) \sum_{\mathbf{y}^n \in C} (\mathbf{y}^n - \mathbf{y}_c)^{\mathbf{k}} f_j^n \\
 &\approx \sum_{\|\mathbf{k}\|=0}^p a_{ij}^{\mathbf{k}}(\mathbf{x}^m, \mathbf{y}_c) M_j^{\mathbf{k}}(C), \tag{2.3}
 \end{aligned}$$

where p is the order of approximation, $\mathbf{k} = (k_1, k_2, k_3)$ is an integer multi-index with all $k_i \geq 0$,

$$a_{ij}^{\mathbf{k}}(\mathbf{x}^m, \mathbf{y}_c) = \frac{1}{\mathbf{k}!} D_{\mathbf{y}}^{\mathbf{k}} S_{ij}(\mathbf{x}^m, \mathbf{y}_c) \tag{2.4}$$

is the \mathbf{k} th Taylor coefficient of the Stokeslet and

$$M_j^{\mathbf{k}}(C) = \sum_{\mathbf{y}^n \in C} (\mathbf{y}^n - \mathbf{y}_c)^{\mathbf{k}} f_j^n \tag{2.5}$$

is the \mathbf{k} th moment of cluster C . Note that the following conventions are used for multi-index notation, $\|\mathbf{k}\| = k_1 + k_2 + k_3$, $\mathbf{k}! = k_1! k_2! k_3!$, $D_{\mathbf{y}}^{\mathbf{k}} = D_{y_1}^{k_1} D_{y_2}^{k_2} D_{y_3}^{k_3}$, $\mathbf{y}^{\mathbf{k}} = y_1^{k_1} y_2^{k_2} y_3^{k_3}$.

Note that the particle-cluster approximation (2.3) separates the target and source particles and this promotes an efficient computation. The key points are (i) the Taylor coefficients a_{ij}^k depend on the target particle \mathbf{x}^m and cluster center \mathbf{y}_c , but not on the particles \mathbf{y}^n in the cluster and (ii) the cluster moments M_j^k depend only on the particles in the cluster C , so they can be stored and re-used for different target particles \mathbf{x}^m .

Similarly for the stresslet particle-cluster interaction,

$$u_i^{str}(\mathbf{x}^m, C) = \sum_{\mathbf{y}^n \in C} T_{ijl}(\mathbf{x}^m, \mathbf{y}^n) h_j^n v_l^n \approx \sum_{\|\mathbf{k}\|=0}^p \tilde{a}_{ijl}^k(\mathbf{x}^m, \mathbf{y}_c) \tilde{M}_{jl}^k(C), \tag{2.6}$$

with stresslet Taylor coefficients,

$$\tilde{a}_{ijl}^k(\mathbf{x}^m, \mathbf{y}_c) = \frac{1}{\mathbf{k}!} D_{\mathbf{y}}^k T_{ijl}(\mathbf{x}^m, \mathbf{y}_c), \tag{2.7}$$

and cluster moments,

$$\tilde{M}_{jl}^k(C) = \sum_{\mathbf{y}^n \in C} (\mathbf{y}^n - \mathbf{y}_c)^k h_j^n v_l^n. \tag{2.8}$$

Combining (2.3) and (2.6), when a particle \mathbf{x}^m and cluster C are well-separated, the far-field approximation for the induced velocity of the particle-cluster interaction is computed from

$$u_i(\mathbf{x}^m, C) \approx \sum_{\|\mathbf{k}\|=0}^p a_{ij}^k(\mathbf{x}^m, \mathbf{y}_c) M_j^k(C) + \sum_{\|\mathbf{k}\|=0}^p \tilde{a}_{ijl}^k(\mathbf{x}^m, \mathbf{y}_c) \tilde{M}_{jl}^k(C). \tag{2.9}$$

As explained in the next two subsections, the strategy for evaluating (2.9) has two parts. First we derive alternative expressions for the Stokeslet and stresslet Taylor coefficients, $a_{ij}^k, \tilde{a}_{ijl}^k$ and then using those expressions we derive a method for efficient computation of (2.9).

3 Alternative expressions for Taylor coefficients

First consider the Coulomb potential,

$$G(\mathbf{x}, \mathbf{y}) = \frac{1}{|\mathbf{x} - \mathbf{y}|}, \tag{3.1}$$

with Taylor coefficients

$$b^k(\mathbf{x}, \mathbf{y}) = \frac{1}{\mathbf{k}!} D_{\mathbf{y}}^k G(\mathbf{x}, \mathbf{y}). \tag{3.2}$$

It was shown in [12,23] that these coefficients satisfy the recurrence relation

$$\|\mathbf{k}\| \cdot |\mathbf{x} - \mathbf{y}|^2 b^{\mathbf{k}} - (2\|\mathbf{k}\| - 1) \sum_{i=1}^3 (x_i - y_i) b^{\mathbf{k} - \mathbf{e}_i} + (\|\mathbf{k}\| - 1) \sum_{i=1}^3 b^{\mathbf{k} - 2\mathbf{e}_i} = 0, \tag{3.3}$$

with $b^0 = G(\mathbf{x}, \mathbf{y})$, $b^{\mathbf{k}} = 0$ when any $k_i < 0$ and \mathbf{e}_i is the i th Cartesian unit vector. First the Stokeslet (1.2a) is written as

$$S_{ij}(\mathbf{x}, \mathbf{y}) = \delta_{ij}G(\mathbf{x}, \mathbf{y}) + (x_j - y_j)D_{y_i}G(\mathbf{x}, \mathbf{y}), \quad (3.4)$$

and then applying the Leibniz rule for differentiating a product, we have

$$\begin{aligned} \frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k}}[(x_j - y_j)D_{y_i}G] &= \frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} - k_j\mathbf{e}_j}D_{y_j}^{k_j}[(x_j - y_j)D_{y_i}G] \\ &= \frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} - k_j\mathbf{e}_j}[(x_j - y_j)D_{y_j}^{k_j}D_{y_i}G - k_jD_{y_j}^{k_j-1}D_{y_i}G] \\ &= (x_j - y_j)\frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} + \mathbf{e}_i}G - \frac{k_j}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j}G \\ &= (x_j - y_j)(k_i + 1)b^{\mathbf{k} + \mathbf{e}_i} - (k_i + 1 - \delta_{ij})b^{\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j}. \end{aligned} \quad (3.5)$$

This yields the following expression for $a_{ij}^{\mathbf{k}}$ in terms of the $b^{\mathbf{k}}$,

$$a_{ij}^{\mathbf{k}} = \delta_{ij}b^{\mathbf{k}} + (x_j - y_j)(k_i + 1)b^{\mathbf{k} + \mathbf{e}_i} - (k_i + 1 - \delta_{ij})b^{\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j}. \quad (3.6)$$

Similarly the stresslet (1.2b) is written as

$$T_{ijl}(\mathbf{x}, \mathbf{y}) = \frac{1}{3}[(x_l - y_l)D_{y_i}D_{y_j}G(\mathbf{x}, \mathbf{y}) + \delta_{ij}D_{y_l}G(\mathbf{x}, \mathbf{y})], \quad (3.7)$$

so the Taylor coefficients are

$$\tilde{a}_{ijl}^{\mathbf{k}} = \frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k}}T_{ijl} = \frac{1}{3}\left[\frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k}}[(x_l - y_l)D_{y_i}D_{y_j}G] + \delta_{ij}\frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k}}D_{y_l}G\right]. \quad (3.8)$$

The first term on the right is

$$\begin{aligned} \frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k}}[(x_l - y_l)D_{y_i}D_{y_j}G] &= (x_l - y_l)\frac{1}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j}G - \frac{k_l}{\mathbf{k}!}D_{\mathbf{y}}^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_l}G \\ &= (x_l - y_l)(k_i + 1)(k_j + 1 + \delta_{ij})b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j} \\ &\quad - (k_i + 1 - \delta_{il})(k_j + 1 + \delta_{ij} - \delta_{jl})b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_l}. \end{aligned} \quad (3.9)$$

This yields the following expression for $\tilde{a}_{ijl}^{\mathbf{k}}$ in terms of the $b^{\mathbf{k}}$,

$$\begin{aligned} 3\tilde{a}_{ijl}^{\mathbf{k}} &= (x_l - y_l)(k_i + 1)(k_j + 1 + \delta_{ij})b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j} \\ &\quad - (k_i + 1 - \delta_{il})(k_j + 1 + \delta_{ij} - \delta_{jl})b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_l} + \delta_{ij}(k_l + 1)b^{\mathbf{k} + \mathbf{e}_l}. \end{aligned} \quad (3.10)$$

The Taylor coefficients $a_{ij}^{\mathbf{k}}$, $\tilde{a}_{ijl}^{\mathbf{k}}$ could be computed explicitly using (3.6) and (3.10), however instead we will employ these relations implicitly to obtain a more efficient evaluation of the particle-cluster approximation (2.9). The details are explained in the next section.

4 Efficient computation of particle-cluster approximations

In this section we explain how to rewrite various sums for improved computational efficiency, first for the direct sum (1.3) to illustrate the idea and then for the particle-cluster approximation (2.9).

4.1 Direct sum

The procedure for efficient evaluation of the direct sum (1.3) was previously used in the FMMLIB3D code [14, 17]. It is repeated here to illustrate the idea and then it is applied to the particle-cluster approximation in the treecode. The idea is to contract the sums and re-use quantities wherever possible. First note that the Stokeslet part of (1.3) can be expressed as

$$\begin{aligned}
 u_i^{sto}(\mathbf{x}^m) &= \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{j=1}^3 S_{ij}(\mathbf{x}^m, \mathbf{y}^n) f_j^n \\
 &= \sum_{\substack{n=1 \\ n \neq m}}^N \left(\sum_{j=1}^3 \frac{\delta_{ij}}{|\mathbf{x}^m - \mathbf{y}^n|} f_j^n + \sum_{j=1}^3 \frac{(x_i^m - y_i^n)(x_j^m - y_j^n)}{|\mathbf{x}^m - \mathbf{y}^n|^3} f_j^n \right) \\
 &= \sum_{\substack{n=1 \\ n \neq m}}^N \left(\frac{f_i^n}{|\mathbf{x}^m - \mathbf{y}^n|} + (x_i^m - y_i^n) s^{mn} \right), \tag{4.1}
 \end{aligned}$$

where

$$s^{mn} = |\mathbf{x}^m - \mathbf{y}^n|^{-3} \sum_{j=1}^3 (x_j^m - y_j^n) f_j^n$$

can be re-used for $i = 1:3$. Similarly, the stresslet part of (1.3) can be expressed as

$$\begin{aligned}
 u_i^{str}(\mathbf{x}^m) &= \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{j=1}^3 \sum_{l=1}^3 T_{ijl}(\mathbf{x}^m, \mathbf{y}^n) h_j^n v_l^n \\
 &= \sum_{\substack{n=1 \\ n \neq m}}^N \sum_{j=1}^3 \sum_{l=1}^3 \frac{(x_i^m - y_i^n)(x_j^m - y_j^n)(x_l^m - y_l^n)}{|\mathbf{x}^m - \mathbf{y}^n|^5} h_j^n v_l^n = \sum_{\substack{n=1 \\ n \neq m}}^N (x_i^m - y_i^n) t^{mn}, \tag{4.2}
 \end{aligned}$$

where

$$t^{mn} = |\mathbf{x}^m - \mathbf{y}^n|^{-5} \sum_{j=1}^3 (x_j^m - y_j^n) h_j^n \sum_{l=1}^3 (x_l^m - y_l^n) v_l^n$$

can be re-used for $i = 1:3$. Using (4.1) and (4.2), the operation count for direct summation is still $\mathcal{O}(N^2)$, but we observe empirically that the CPU run time is reduced since this procedure avoids explicitly forming the tensors S_{ij} , T_{ijl} .

4.2 Stokeslet particle-cluster interaction

We take a similar approach in computing the particle-cluster approximations in the treecode; the sums are contracted to reduce the number of operations, terms are re-used wherever possible and we avoid explicitly forming the Taylor coefficient tensors $a_{ij}^{\mathbf{k}}, \tilde{a}_{ijl}^{\mathbf{k}}$. Consider the Stokeslet particle-cluster approximation (2.9); using the expression for the Taylor coefficients (3.6), we have

$$\begin{aligned} u_i^{sto}(\mathbf{x}^m, C) &\approx \sum_{\|\mathbf{k}\|=0}^p \sum_{j=1}^3 a_{ij}^{\mathbf{k}}(\mathbf{x}^m, \mathbf{y}_c) M_j^{\mathbf{k}}(C) \\ &= \sum_{\|\mathbf{k}\|=0}^p \sum_{j=1}^3 \left[\delta_{ij} b^{\mathbf{k}} + (x_j - y_j)(k_i + 1) b^{\mathbf{k} + \mathbf{e}_i} - (k_i + 1 - \delta_{ij}) b^{\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j} \right] M_j^{\mathbf{k}}(C). \end{aligned} \quad (4.3)$$

This simplifies to

$$u_i^{sto}(\mathbf{x}^m, C) \approx \sum_{\|\mathbf{k}\|=0}^p \left[2b^{\mathbf{k}} M_i^{\mathbf{k}}(C) + (k_i + 1) \left[b^{\mathbf{k} + \mathbf{e}_i} \sigma^{\mathbf{k}}(C) - \sum_{j=1}^3 b^{\mathbf{k} + \mathbf{e}_i - \mathbf{e}_j} M_j^{\mathbf{k}}(C) \right] \right], \quad (4.4)$$

where

$$\sigma^{\mathbf{k}}(C) = \sum_{j=1}^3 (x_j - y_j) M_j^{\mathbf{k}}(C)$$

can be re-used for $i = 1 : 3$.

4.3 Stresslet particle-cluster interaction

Next consider the stresslet particle-cluster approximation (2.9); using the expression for the Taylor coefficients (3.10), we have

$$\begin{aligned} u_i^{str}(\mathbf{x}^m, C) &\approx \sum_{\|\mathbf{k}\|=0}^p \sum_{j=1}^3 \sum_{l=1}^3 \tilde{a}_{ijl}^{\mathbf{k}}(\mathbf{x}^m, \mathbf{y}_c) \tilde{M}_{jl}^{\mathbf{k}}(C) \\ &= \frac{1}{3} \sum_{\|\mathbf{k}\|=0}^p \sum_{j=1}^3 \sum_{l=1}^3 \left[(x_l - y_l)(k_i + 1)(k_j + 1 + \delta_{ij}) b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j} \right. \\ &\quad \left. - (k_i + 1 - \delta_{il})(k_j + 1 + \delta_{ij} - \delta_{jl}) b^{\mathbf{k} + \mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_l} + \delta_{ij}(k_l + 1) b^{\mathbf{k} + \mathbf{e}_l} \right] \tilde{M}_{jl}^{\mathbf{k}}(C). \end{aligned} \quad (4.5)$$

To keep the next few intermediate formulas more concise, in the remainder of this section we drop the arguments \mathbf{x}^m, C . Then moving the sums over indices j, l as far as possible to

the right and splitting (4.5) into two terms, we have

$$u_i^{str} \approx \frac{1}{3} \sum_{\|\mathbf{k}\|=0}^p \left[(k_i+1) \sum_{j=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j} \sum_{l=1}^3 (x_l-y_l) \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.6a}$$

$$\left. - (k_i+1) \sum_{j=1}^3 \sum_{l=1}^3 (k_j+1+\delta_{ij}-\delta_{jl}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.6b}$$

$$\left. + \sum_{j=1}^3 \sum_{l=1}^3 \delta_{il} (k_j+1+\delta_{ij}-\delta_{jl}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.6c}$$

$$\left. + \sum_{l=1}^3 (k_l+1) b^{\mathbf{k}+\mathbf{e}_l} \sum_{j=1}^3 \delta_{ij} \tilde{M}_{jl}^{\mathbf{k}} \right]. \tag{4.6d}$$

Then split (4.6b) into two terms and simplify (4.6c) and (4.6d) to obtain

$$u_i^{str} \approx \frac{1}{3} \sum_{\|\mathbf{k}\|=0}^p \left[(k_i+1) \sum_{j=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j} \sum_{l=1}^3 (x_l-y_l) \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.7a}$$

$$\left. - (k_i+1) \sum_{j=1}^3 \sum_{l=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.7b}$$

$$\left. + (k_i+1) \sum_{j=1}^3 \sum_{l=1}^3 \delta_{jl} b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.7c}$$

$$\left. + \sum_{j=1}^3 (k_j+1) b^{\mathbf{k}+\mathbf{e}_j} \tilde{M}_{ji}^{\mathbf{k}} + \sum_{l=1}^3 (k_l+1) b^{\mathbf{k}+\mathbf{e}_l} \tilde{M}_{il}^{\mathbf{k}} \right]. \tag{4.7d}$$

Defining

$$\tau_j = \sum_{l=1}^3 (x_l-y_l) \tilde{M}_{jl}^{\mathbf{k}}$$

in (4.7a), simplifying (4.7c) and combining the terms in (4.7d), we have

$$u_i^{str} \approx \frac{1}{3} \sum_{\|\mathbf{k}\|=0}^p \left[(k_i+1) \sum_{j=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j} \tau_j \right. \tag{4.8}$$

$$\left. - (k_i+1) \sum_{j=1}^3 \sum_{l=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}} \right. \tag{4.8}$$

$$\left. + (k_i+1) b^{\mathbf{k}+\mathbf{e}_i} \sum_{j=1}^3 \tilde{M}_{jj}^{\mathbf{k}} + \sum_{j=1}^3 (k_j+1) b^{\mathbf{k}+\mathbf{e}_j} (\tilde{M}_{ji}^{\mathbf{k}} + \tilde{M}_{ij}^{\mathbf{k}}) \right]. \tag{4.8}$$

Next let

$$m^{\mathbf{k}} = \sum_{j=1}^3 \tilde{M}_{jj}^{\mathbf{k}}, \quad m_{ij}^{\mathbf{k}} = \tilde{M}_{ij}^{\mathbf{k}} + \tilde{M}_{ji}^{\mathbf{k}},$$

and restore the arguments \mathbf{x}^m, C , so that the stresslet particle-cluster approximation is

$$\begin{aligned}
 u_i^{str}(\mathbf{x}^m, C) \approx & \frac{1}{3} \sum_{|\mathbf{k}|=0}^p \left[(k_i+1) \sum_{j=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j} \tau_j(C) \right. \\
 & - (k_i+1) \sum_{j=1}^3 \sum_{l=1}^3 (k_j+1+\delta_{ij}) b^{\mathbf{k}+\mathbf{e}_i+\mathbf{e}_j-\mathbf{e}_l} \tilde{M}_{jl}^{\mathbf{k}}(C) \\
 & \left. + (k_i+1) b^{\mathbf{k}+\mathbf{e}_i} m^{\mathbf{k}}(C) + \sum_{j=1}^3 (k_j+1) b^{\mathbf{k}+\mathbf{e}_j} m_{ij}^{\mathbf{k}}(C) \right]. \quad (4.9)
 \end{aligned}$$

In summary, the particle-cluster approximation (2.9) is computed using (4.4) for the Stokeslet part and (4.9) for the stresslet part, where the coefficients $b^{\mathbf{k}}$ are computed using the recurrence relation (3.3). The introduction of the quantities $\sigma^{\mathbf{k}}, \tau^{\mathbf{k}}, m^{\mathbf{k}}, m_{ij}^{\mathbf{k}}$ enables a more efficient computation. The operation count for the treecode is still $\mathcal{O}(N \log N)$, but we observe empirically that the CPU run time is reduced since this procedure avoids explicitly forming the Taylor coefficient tensors $a_{ij}^{\mathbf{k}}, \tilde{a}_{ijl}^{\mathbf{k}}$.

5 Description of treecode

The treecode implementation starts by inputting the particle positions and weights and building a hierarchical tree of particle clusters [4]. The root cluster is a cube containing all the source particles. The root is bisected along the Cartesian axes and the eight children become subclusters of the root. The child clusters are similarly bisected and the process continues until a cluster contains fewer than N_0 particles, where N_0 is a user-specified parameter. Each cluster has a data structure containing necessary information, e.g., pointers to the particles belonging to the cluster, coordinates of the cluster center, pointers to the children of the cluster and so on.

The procedure is outlined in Algorithm 1. The code uses a multipole acceptance criterion (MAC) to determine whether a given target particle \mathbf{x}^m and source cluster C are well-separated. The criterion for being well-separated is

$$\frac{r}{R} \leq \theta, \quad (5.1)$$

where as shown in Fig. 1, r is the cluster radius, R is the distance between the particle and the cluster center \mathbf{y}_c and θ is a user-specified parameter which together with the order p controls the approximation error. The code cycles through the target particles and each particle interacts with source clusters starting at the root. If the MAC (5.1) is satisfied, the particle-cluster approximation is computed as explained above. If the MAC is not satisfied, the code checks the children of the cluster, or if the cluster is a leaf (no children), direct summation is performed, again using the efficient formulation explained above. This structure follows the original Barnes-Hut treecode algorithm [4], modified to

accommodate higher-order particle-cluster approximations of the Stokeslet and stresslet kernels.

Algorithm 1 treecode.

```

1: program main
2:   input particle positions  $\mathbf{x}^n$  and weights  $f_j^n, h_j^n, v_j^n$ , treecode parameters  $p, \theta, N_0$ 
3:   build tree, compute cluster moments  $M_j^k(C), \tilde{M}_{jl}^k(C)$ 
4:   for  $n = 1 : N$ 
5:     compute-velocity ( $\mathbf{x}^n$ , root-cluster)
6: subroutine compute-velocity ( $\mathbf{x}, C$ )
7:   if MAC is satisfied
8:     compute particle-cluster interaction by far-field Taylor approximation
9:   else
10:    if  $C$  is a leaf, compute particle-cluster interaction by direct sum
11:    else
12:      for each child  $C'$  of  $C$ 
13:        compute-velocity ( $\mathbf{x}, C'$ )
  
```

The treecode algorithm was programmed in C++ and compiled using the Intel icpc compiler with -O2 optimization flag. The source code is available for download [18]. The computations were performed on the University of Wisconsin-Milwaukee Mortimer Faculty Research Cluster which has 55 standard compute nodes and each node is a Dell PowerEdge R430 server with two 12-core Intel Xeon E5-2680 v3 processors at 2.50GHz and 64 GB RAM. Serial computations were done on one core and parallel computations were done using MPI with each process running on one core.

6 Numerical results

We present results for two test cases. The first test case has Stokeslet and stresslet particles on the surface of a unit sphere, as in boundary element simulations of exterior Stokes flow. The particle distribution is given by triangulating the sphere as follows. Starting from an icosahedron with 20 triangular faces, the faces are refined by connecting the centers of the edges, resulting in triangulations with $N = 20 \cdot 4^L$ faces for L levels of refinement. The particles are obtained by projecting the triangle centroids onto the sphere. In this test case each cluster in the tree is shrunk to the bounding rectangular box containing its particles. The second test case follows [2] and considers Stokeslet particles randomly distributed in cubes of different sizes, with a fixed particle number density. In both test cases the particle weights f_j^n, h_j^n are random numbers in $[-1, 1]$.

There are three user-specified parameters required for the treecode. The parameter N_0 is the maximum number of particles in a leaf of the tree; the value $N_0 = 2000$ is used throughout this work. The other parameters are the order of approximation p and MAC

parameter θ ; we will vary these parameters to investigate their effect on the code's performance. The relative error in velocity (1.3) is defined by

$$E = \left(\frac{\sum_{n=1}^N |\mathbf{u}^d(\mathbf{x}^n) - \mathbf{u}^t(\mathbf{x}^n)|^2}{\sum_{n=1}^N |\mathbf{u}^d(\mathbf{x}^n)|^2} \right)^{1/2}, \quad (6.1)$$

where \mathbf{u}^d is the velocity obtained by direct summation and \mathbf{u}^t is the treecode approximation. The CPU time is given in units of seconds (s). The following subsection presents serial results for each test case, followed by a subsection with parallel results for the second test case.

6.1 Serial computations

6.1.1 Test case 1: particles on the surface of a sphere

The first test case has Stokeslet and stresslet particles on a unit sphere at locations determined by the icosahedral triangulation as explained above. Fig. 2 plots the treecode CPU time versus error E for systems of size $N = 82920, 317680, 1310720$, with Taylor approximation order $p = 0:2:10$ (increasing from right to left) and MAC parameter $\theta = 0.8, 0.5, 0.2$ (decreasing from right to left). As expected, smaller error E is attained by increasing the order p and decreasing the MAC parameter θ , but this increases the CPU time. Note that as p increases, the error decreases more rapidly for smaller θ . Also note that for a given order p and MAC parameter θ , the error is relatively insensitive to the system size N . We can distinguish three regimes; for low accuracy $\theta = 0.8$ is most efficient, for medium accuracy $\theta = 0.5$ is most efficient and for high accuracy $\theta = 0.2$ is most efficient. Fig. 2 also shows the direct sum CPU time; as the system size increases, the treecode becomes more efficient in comparison with direct summation.

The dependence of the error and CPU time on the MAC parameter θ can be explained as follows. Choosing a smaller θ has two effects, (i) the code descends deeper into the tree, so the CPU time increases, but the clusters have smaller radius, so the Taylor approximation is more accurate and (ii) there is higher likelihood the code will reach the leaf clusters of the tree, which again increases the CPU time, but in that case direct summation is performed, which incurs no error.

Table 1 gives the speedup (ratio d/t of CPU times for direct sum and treecode) and the error E . The treecode achieves higher speedup as the system size increases; for example, a medium accuracy $E \approx 5e-05$ is attained using MAC parameter $\theta = 0.5$ and order $p = 6$; in this case the treecode is 3.5 times faster than direct summation for $N = 81920$, 11 times faster for $N = 327680$ and 38 times faster for $N = 1310720$.

Fig. 3 shows results graphically for MAC parameter $\theta = 0.5$, order $p = 0:2:10$ and system size between $N = 20480$ and $N = 1310720$. Fig. 3(a) plots the treecode error E

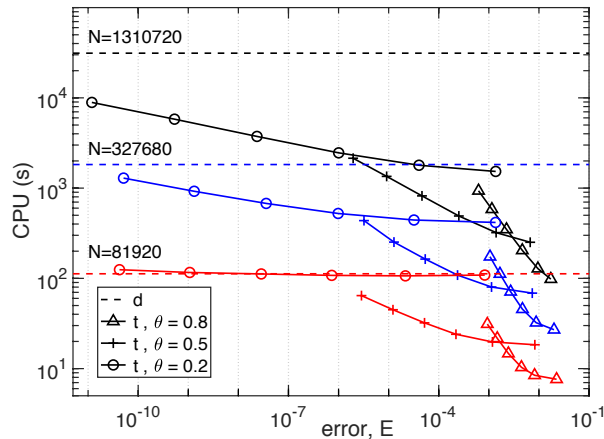


Figure 2: Test case 1, Stokeslets and stresslets on a sphere, CPU time versus error E , system size $N=1310720$ (upper, black), $N=327680$ (middle, blue), $N=81920$ (lower, red), direct sum (d, dashed), treecode (t, solid), MAC parameter $\theta=0.8(\Delta), 0.5(+), 0.2(\circ)$ (decreasing right to left), order $p=0:2:10$ (increasing right to left).

Table 1: Test case 1, Stokeslets and stresslets on a sphere, treecode MAC parameter θ , order p , speedup d/t =ratio of direct sum and treecode CPU times, error E , (a) $N=81920$, $d=112s$, (b) $N=327680$, $d=1824s$, (c) $N=1310720$, $d=31267s$.

θ	(a) $N=81920$			(b) $N=327680$			(c) $N=1310720$		
	p	d/t	E	p	d/t	E	p	d/t	E
0.2	0	1.03	8.3e-04	0	4.38	1.4e-03	0	20.43	1.4e-03
	2	1.05	2.2e-05	2	4.11	3.2e-05	2	17.44	4.0e-05
	4	1.04	7.3e-07	4	3.48	9.8e-07	4	12.71	1.0e-06
	6	1.00	2.9e-08	6	2.69	3.6e-08	6	8.36	2.4e-08
	8	0.97	1.1e-09	8	1.97	1.3e-09	8	5.38	5.3e-10
	10	0.90	4.2e-11	10	1.41	5.1e-11	10	3.51	1.2e-11
0.5	0	6.12	8.9e-03	0	26.59	7.4e-03	0	124.49	6.7e-03
	2	5.69	1.2e-03	2	22.75	1.1e-03	2	97.36	1.4e-03
	4	4.68	2.2e-04	4	16.79	2.4e-04	4	63.67	2.5e-04
	6	3.50	5.2e-05	6	11.15	5.4e-05	6	38.12	4.6e-05
	8	2.50	1.2e-05	8	7.23	1.3e-05	8	23.18	9.1e-06
	10	1.74	2.9e-06	10	4.18	3.2e-06	10	14.59	2.0e-06
0.8	0	14.75	2.3e-02	0	67.63	2.0e-02	0	318.73	1.7e-02
	2	13.32	8.2e-03	2	56.52	8.8e-03	2	243.78	9.5e-03
	4	10.87	4.5e-03	4	40.07	4.7e-03	4	153.01	4.7e-03
	6	7.65	2.5e-03	6	25.66	2.7e-03	6	89.71	2.3e-03
	8	5.19	1.5e-03	8	16.29	1.7e-03	8	53.60	1.2e-03
	10	3.61	9.3e-04	10	10.49	1.1e-03	10	33.48	6.2e-04

versus N . The error is relatively insensitive to the system size and for a given N , the error decreases with increasing order p . Fig. 3(b) plots the CPU time versus N . The direct sum CPU time scales like $\mathcal{O}(N^2)$, while the treecode CPU time is consistent with $\mathcal{O}(N \log N)$

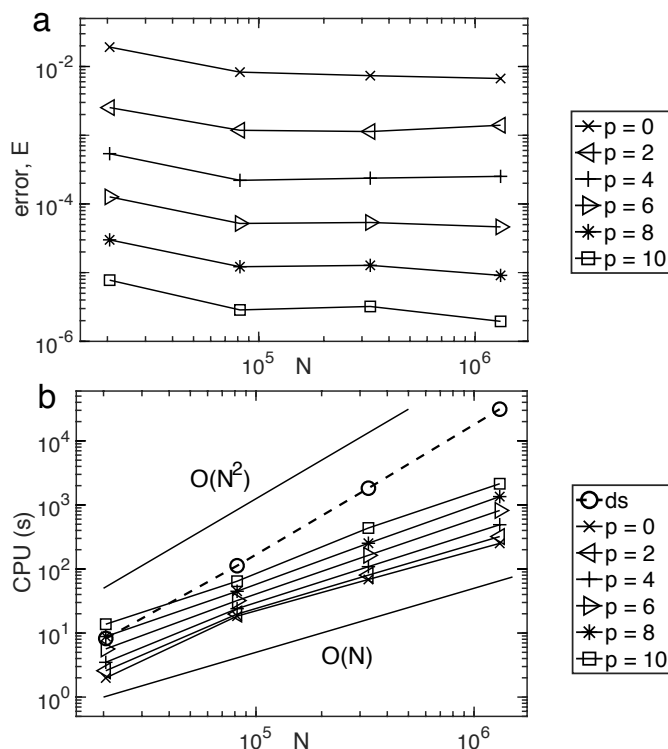


Figure 3: Test case 1, Stokeslets and stresslets on a sphere, treecode MAC parameter $\theta=0.5$, order $p=0:2:10$, system size $N=20480:1310720$, (a) error, E , (b) CPU time for direct sum (dashed), treecode (solid).

and hence the treecode is faster than direct summation except for small system size N and large order p . In the remainder of this section we use MAC parameter $\theta=0.5$.

Table 2 displays the peak memory used by the treecode as a function of system size N for order $p=0:2:10$. The memory used by direct summation is also given. The memory usage statistics were obtained using the Valgrind massif analysis tool (valgrind.org). The treecode and direct summation both store the particles in arrays of size $\mathcal{O}(N)$, but the treecode uses additional memory of size $\mathcal{O}((N/N_0)p^3)$ for the cluster moments, where the factor N/N_0 represents the number of clusters in the tree and the factor p^3 is the memory associated with the moments of a cluster. Recall that the number of particles in a leaf cluster is set here at $N_0=2000$.

In the current treecode implementation, the additional memory used for the moments was reduced as follows. Note that for a given cluster C and indices $j, l=1:3$, the moments $M_j^k(C)$, $\tilde{M}_{jl}^k(C)$ could be stored in square three-dimensional arrays of size p^3 , corresponding to the index $\mathbf{k}=(k_1, k_2, k_3)$. However since we only need the indices with $\|\mathbf{k}\|=k_1+k_2+k_3=0:p$, a large portion of these arrays would be empty. So instead the moments are stored in one-dimensional arrays with no empty space, by accessing the indices (k_1, k_2, k_3) in a fixed order; we refer to this as flattening the moment arrays. While

Table 2: Test case 1, Stokeslets and stresslets on a sphere, memory usage (MB) for system size N , direct sum, treecode with order p .

	$N = 20480$	$N = 81920$	$N = 327680$	$N = 1310720$
direct sum	3.3	13.1	52.4	209.8
treecode order p				
0	3.3	13.1	52.6	210.7
2	3.4	13.2	53.0	214.3
4	3.7	13.6	54.2	224.1
6	4.3	14.2	56.6	243.4
8	5.4	15.2	60.6	275.2
10	7.6	16.6	66.6	322.8

the treecode memory usage still has a term scaling like $\mathcal{O}((N/N_0)p^3)$, the prefactor is greatly reduced. Table 2 shows that in this range of system size N and order $p \leq 10$, except for one case with the smallest N and largest p , the treecode uses less than twice as much memory as direct summation.

6.1.2 Test case 2: particles in a cube

The second test case has Stokeslet particles located randomly in a cube of side length L , with number density $N/L^3 = 2500$ [2]. Table 3 shows the speedup (ratio d/t of CPU times for direct sum and treecode) and the error E , for system size $N = 125K$ and $N = 1000K$. The trends in error and CPU time with respect to MAC parameter θ and order p are similar to test case 1. The speedup in this case is somewhat less than in the previous case, because here the system sizes are smaller and the stresslet part of the sum is omitted; nonetheless the treecode is faster than direct summation except for one case with small θ and large p .

Table 4 presents the memory usage in test case 2. As before, the treecode uses less than twice as much memory as direct summation. The relatively low memory usage of the treecode is an advantage in parallel simulations, where it enables a simple replicated data approach as shown below.

6.2 Parallel computations

The parallel treecode implementation used here relies on the observation that the target particle computations in Algorithm 1 are independent of each other, enabling a replicated data approach [16]. The procedure is sketched in Algorithm 2 and was implemented using MPI. The idea is that the particle array of length N is divided into np segments of length N/np , where np is the chosen number of processes and each segment is assigned to one process (recall that each process is running on a single core). Each process receives a copy of the entire particle array, builds a local copy of the tree and computes the cluster moments. The processes run concurrently and each one computes the induced velocity at its assigned target particles. The scheme assumes that the entire particle array and tree structure fit into the memory of each core, which is facilitated by the treecode's relatively

Table 3: Test case 2, random Stokeslets in a cube with number density $N/L^3 = 2500$ [2], treecode MAC parameter θ , order p , speedup d/t =ratio of direct sum and treecode CPU times, error E , (a) $N=125K$, $d=150s$, (b) $N=1000K$, $d=9873s$.

θ	(a) $N=125K$			(b) $N=1000K$		
	p	d/t	E	p	d/t	E
0.2	0	1.14	1.1e-02	0	3.46	4.3e-02
	2	1.13	2.2e-04	2	3.37	7.1e-04
	4	1.11	4.4e-06	4	3.18	1.5e-05
	6	1.08	1.1e-07	6	2.89	3.9e-07
	8	1.02	3.2e-09	8	2.50	1.1e-08
	10	0.95	9.6e-11	10	2.08	3.2e-10
0.5	0	5.50	1.1e-01	0	33.89	1.5e-01
	2	5.28	8.2e-03	2	31.39	1.6e-02
	4	4.79	1.1e-03	4	26.86	2.0e-03
	6	4.17	1.7e-04	6	21.13	3.3e-04
	8	3.40	3.0e-05	8	13.77	5.7e-05
	10	2.64	5.5e-06	10	11.28	1.0e-05
0.8	0	16.94	2.1e-01	0	123.78	2.7e-01
	2	15.85	4.3e-02	2	111.61	7.0e-02
	4	13.81	1.4e-02	4	90.79	2.3e-02
	6	11.32	5.4e-03	6	67.34	9.0e-03
	8	8.77	2.4e-03	8	47.13	4.1e-03
	10	6.40	1.1e-03	10	32.54	1.9e-03

Table 4: Test case 2, random Stokeslets in a cube with number density $N/L^3 = 2500$ [2], memory usage (MB) for system size N , direct sum, treecode with order p .

	$N=125K$	$N=1000K$
direct sum	11.0	88.0
treecode order p		
0	14.0	112.4
2	14.1	112.8
4	14.3	114.0
6	14.6	116.4
8	15.0	120.4
10	15.8	126.3

low memory usage; this is not an issue for the examples considered here, but in case of a much larger system size where this assumption does not hold, a distributed memory approach would be required.

The parallel treecode performance is demonstrated below for test case 2. First we consider strong scaling (fixed N) and then weak scaling (increasing N), as the number of processes np increases. We also implemented a parallel direct sum using the same replicated data approach for comparison with the parallel treecode.

Algorithm 2 parallel treecode.

-
- 1: in main process
 - 2: input particle positions \mathbf{x}^n , weights f_j^n, h_j^n , treecode parameters p, θ, N_0
 - 3: broadcast particle array to each process
 - 4: in each process
 - 5: build local copy of tree, compute cluster moments
 - 6: use treecode to compute induced velocity at assigned target particles
 - 7: send result to main process
-

Table 5: Test case 2, random Stokeslets in a cube, parallel strong scaling, system size $N = 1000K$, treecode parameters $\theta = 0.5$, $p = 6$, error $E = 3.1e-04$, number of MPI processes (np), CPU time (d =direct sum, t =treecode), ratio of CPU time for one process and np processes ($d_1/d_{np}, t_1/t_{np}$), parallel efficiency (PE, ratio/ np), treecode speedup (d/t).

np	d CPU (s)	d_1/d_{np}	d PE (%)	t CPU (s)	t_1/t_{np}	t PE (%)	d/t
1	9707.0	1.00	100.0	336.6	1.00	100.0	28.8
2	4911.4	1.98	98.8	176.3	1.91	95.4	27.9
4	2451.2	3.96	99.0	86.6	3.89	97.2	28.3
8	1345.6	7.21	90.2	48.1	7.00	87.6	28.0
16	702.0	13.83	86.4	25.6	13.16	82.2	27.4
32	350.7	27.68	86.5	13.6	24.75	77.3	25.8

6.2.1 Strong scaling

Table 5 shows results for test case 2 with $N = 1000K$ random Stokeslets in a cube. The treecode parameters are $\theta = 0.5$, $p = 6$, yielding error $E = 3.1e-04$. The Table displays the CPU time, ratio of CPU time for 1 process and np processes and parallel efficiency, for the direct sum (d) and treecode (t) and finally the speedup due to using the treecode (d/t), up to $np = 32$ processes. The parallel direct sum reduces the CPU time from 9707 s on 1 process to 350.7 s on 32 processes, for parallel efficiency 86.5%. As expected the treecode CPU times are smaller and the parallel performance is almost as good; the CPU time is reduced from 336.6 s on 1 process to 13.6 s on 32 processes, for parallel efficiency 77.3%. The treecode is 28 times faster than direct summation on 1 process and 25 times faster on 32 processes.

6.2.2 Weak scaling

Table 6 shows results for test case 2 starting with $N = 125K$ particles on 1 process, then doubling the number of particles and processes until reaching $N = 4000K$ particles on 32 processes. The box size increases so that the number density is $N/L^3 = 2500$ [2]. The treecode parameters are $\theta = 0.5$, $p = 6$, yielding error $E \leq 3.7e-04$. The results show that with each doubling of N and np , the direct sum CPU time approximately doubles, while the treecode CPU time increases more slowly. Hence the treecode performance improves as the system size increases; with $N = 125K$ on 1 process, the treecode is 6 times faster than

Table 6: Test case 2, random Stokeslets in a cube, parallel weak scaling, system size (N), number of MPI processes (np), treecode parameters $\theta=0.5$, $p=6$, error $E \leq 3.7e-04$, CPU time (d = direct sum, t = treecode), treecode speedup d/t .

N	np	d CPU (s)	t CPU (s)	d/t
125K	1	145.5	23.7	6.1
250K	2	291.7	27.2	10.7
500K	4	610.3	32.9	18.6
1000K	8	1345.7	48.1	28.0
2000K	16	2804.9	56.7	49.5
4000K	32	5565.9	67.7	82.3

direct summation, but with $N=4000K$ on 32 processes, the treecode is 82 times faster.

7 Summary

We presented a treecode algorithm for computing the velocity induced by a collection of Stokeslets and stresslets in 3D flow. The method uses a far-field Cartesian Taylor approximation to compute well-separated particle-cluster interactions. Expressions were derived for the Taylor coefficients of the Stokeslet and stresslet kernels in terms of the Taylor coefficients of the Coulomb potential and these expressions enable an efficient computation of higher-order approximations. Numerical results were presented for icosahedral particles on the surface of a sphere and random particles in a cube [2]. For a given level of accuracy, the treecode CPU time scales like $\mathcal{O}(N \log N)$, where N is the number of particles and a substantial speedup over direct summation is achieved for large systems. The memory usage increases with the system size N and Taylor approximation order p , but for the range of parameters considered here, in most cases the treecode used less than twice as much memory as direct summation. A relatively straightforward parallel treecode implementation was demonstrated.

It is beyond the scope of the present work to make a detailed performance comparison with other methods for fast summation of Stokeslets and stresslets such as the Fast Multipole Method (FMM) [17,25,26] and the Spectral Ewald (SE) method [2]. These methods have demonstrated excellent performance in terms of accuracy and efficiency. Yet the treecode may be an attractive option in some cases due to its relatively simple algorithmic structure and low memory usage, which together can enhance parallel efficiency.

Our simulations used representative values of the treecode parameters (order p , MAC parameter θ , maximum number of particles in a leaf N_0) and one future goal is to gain efficiency by tuning their values. There are several other directions for future work. The present approach can be extended to treat regularized Stokeslets and stresslets, which may help accelerate biofluid applications using those kernels [8, 9, 30]. Another goal is to apply the treecode in boundary element simulations of Stokes-Darcy porous medium flow [6, 34] and Stokes flow around solid bodies [29].

Acknowledgements

We thank the reviewers for suggestions that helped improve the article and Leighton Wilson for suggesting the method used in reducing the memory associated with particle clusters.

References

- [1] L. AF KLINTEBERG AND A.-K. TORNBORG, *A fast integral equation method for solid particles in viscous flow using quadrature by expansion*, J. Comput. Phys., 326 (2016), pp. 420–445.
- [2] L. AF KLINTEBERG, D.S. SHAMSHIRGAR AND A.-K. TORNBORG, *Fast Ewald summation for free-space Stokes potentials*, Res. Math. Sci., 4 (2017), DOI 10.1186/s40687-016-0092-7.
- [3] D. M. AMBROSE, M. SIEGEL AND S. TLUPOVA, *A small-scale decomposition for 3D boundary integral computations with surface tension*, J. Comput. Phys., 247 (2013), pp. 168–191.
- [4] J. E. BARNES AND P. HUT, *A hierarchical $\mathcal{O}(N \log N)$ force-calculation algorithm*, Nature, 324 (1986), pp. 446–449.
- [5] T. BEATUS, T. TLUSTY AND R. BAR-ZIV, *Phonons in a one-dimensional microfluidic crystal*, Nature Phys., 2 (2006), pp. 743–748.
- [6] Y. BOUBENDIR AND S. TLUPOVA, *Domain decomposition methods for solving Stokes-Darcy problems with boundary integrals*, SIAM J. Sci. Comput., 35 (2013), pp. B82–B106.
- [7] E. CORONA, L. GREENGARD, M. RACHH AND S. VEERAPANENI, *An integral equation formulation for rigid bodies in Stokes flow in three dimensions*, J. Comput. Phys., 332 (2017), pp. 504–519.
- [8] R. CORTEZ, *The method of regularized Stokeslets*, SIAM J. Sci. Comput., 23 (2001), pp. 1204–1225.
- [9] R. CORTEZ, L. FAUCI AND A. MEDOVIKOV, *The method of regularized Stokeslets in three dimensions: Analysis, validation and application to helical swimming*, Phys. Fluids, 17 (2005), 031504.
- [10] C. I. DRAGHICESCU AND M. DRAGHICESCU, *A fast algorithm for vortex blob interactions*, J. Comput. Phys., 116 (1995), pp. 69–78.
- [11] K. DRESCHER, R. E. GOLDSTEIN, N. MICHEL, M. POLIN AND I. TUVAL, *Direct measurement of the flow field around swimming microorganisms*, Phys. Rev. Lett., 105 (2010), 168101.
- [12] Z.-H. DUAN AND R. KRASNY, *An adaptive treecode for computing nonbonded potential energy in classical molecular systems*, J. Comput. Chem., 22 (2001), pp. 184–195.
- [13] U. ESSMANN, L. PERERA, M. L. BERKOWITZ, T. DARDEN, H. LEE AND L. G. PEDERSEN, *A smooth particle mesh Ewald method*, J. Chem. Phys., 103 (1995), pp. 8577–8593.
- [14] FMMLIB3D 1.2, www.cims.nyu.edu/cmcl/fmm3dlib/fmm3dlib.html, (2012)
- [15] Y. FU AND G. J. RODIN, *Fast solution methods for three-dimensional Stokesian many-particle problems*, Commun. Numer. Meth. Eng., 16 (2000), pp. 145–149.
- [16] W.-H. GENG AND R. KRASNY, *A treecode-accelerated boundary integral Poisson-Boltzmann solver for solvated biomolecules*, J. Comput. Phys., 247 (2013), pp. 62–78.
- [17] Z. GIMBUTAS AND L. GREENGARD, *Simple FMM libraries for electrostatics, slow viscous flow and frequency-domain wave propagation*, Commun. Comput. Phys., 18 (2015), pp. 516–528.
- [18] github.com/Treecodes/stokes-treecode.
- [19] L. GREENGARD AND V. ROKHLIN, *A fast algorithm for particle simulations*, J. Comput. Phys., 73 (1987), pp. 325–348.

- [20] R. W. HOCKNEY AND J. H. EASTWOOD, *Computer Simulation Using Particles*, Taylor & Francis, 1988.
- [21] R. KRASNY AND L. WANG, *Fast evaluation of multiquadric RBF sums by a Cartesian treecode*, SIAM J. Sci. Comput., 33 (2011), pp. 2341–2355.
- [22] P. LI, H. JOHNSTON AND R. KRASNY, *A Cartesian treecode for screened Coulomb interactions*, J. Comput. Phys., 228 (2009), pp. 3858–3868.
- [23] K. LINDSAY AND R. KRASNY, *A particle method and adaptive treecode for vortex sheet motion in three-dimensional flow*, J. Comput. Phys., 172 (2001), pp. 879–907.
- [24] Y. J. LIU, *A new fast multipole boundary element method for solving 2D Stokes flow problems based on a dual BIE formulation*, Eng. Anal. Bound. Elem., 32 (2008), pp. 139–151.
- [25] D. MALHOTRA AND G. BIROS, *PVFMM: A parallel kernel independent FMM for particle and volume potentials*, Commun. Comput. Phys., 18 (2015), pp. 808–830.
- [26] D. MALHOTRA AND G. BIROS, *Algorithm 967: A distributed-memory Fast Multipole Method for volume potentials*, ACM Trans. Math. Soft., 43(2) (2016), 17.
- [27] L. C. NITSCHE AND P. PARTHASARATHI, *Cubically regularized Stokeslets for fast particle simulations of low-Reynolds-number drop flows*, Chem. Eng. Commun., 197 (2010), pp. 18–38.
- [28] F. PIGNATEL, M. NICOLAS, E. GUAZZELLI AND D. SAINTILLAN, *Falling jets of particles in viscous fluids*, Phys. Fluids, 21 (2009), 123303.
- [29] C. POZRIKIDIS, *Boundary Integral and Singularity Methods for Linearized Viscous Flow*, Cambridge University Press, 1992.
- [30] M. W. ROSTAMI AND S. D. OLSON, *Kernel-independent fast multipole method within the framework of regularized Stokeslets*, J. Fluid. Struct., 67 (2016), pp. 60–84.
- [31] D. SAINTILLAN, E. DARVE AND E. SHAQFEH, *A smooth particle-mesh Ewald algorithm for Stokes suspension simulations: The sedimentation of fibers*, Phys. Fluids, 17 (2005), 033301/21.
- [32] A. SIEROU AND J. F. BRADY, *Accelerated Stokesian Dynamics simulations*, J. Fluid Mech., 448 (2001), pp. 115–146.
- [33] D. J. SMITH, *A boundary element regularized Stokeslet method applied to cilia- and flagella-driven flow*, Proc. R. Soc. A, 465 (2009), pp. 3605–3626.
- [34] S. TLUPOVA AND R. CORTEZ, *Boundary integral solutions of coupled Stokes and Darcy flows*, J. Comput. Phys., 228 (2009), pp. 158–179.
- [35] A.-K. TORNBERG AND L. GREENGARD, *A fast multipole method for the three-dimensional Stokes equations*, J. Comput. Phys., 227 (2008), pp. 1613–1619.
- [36] S. K. VEERAPANENI, A. RAHIMIAN, G. BIROS AND D. ZORIN, *A fast algorithm for simulating vesicle flows in three dimensions*, J. Comput. Phys., 230 (2011), pp. 5610–5634.
- [37] H. WANG, T. LEI, J. LI, J. HUANG AND Z. YAO, *A parallel fast multipole accelerated integral equation scheme for 3D Stokes equations*, Int. J. Numer. Meth. Eng., 70 (2007), pp. 812–839.
- [38] X. WANG, J. KANPKA, W. YE, N. R. ALURU AND J. WHITE, *Algorithms in Fast Stokes and its application to micromachined device simulation*, IEEE Trans. Comput. Aided Des. Integra. Circ. Syst., 25 (2006), pp. 248–257.
- [39] L. YING, G. BIROS AND D. ZORIN, *A kernel independent adaptive fast multipole algorithm in two and three-dimensions*, J. Comput. Phys., 196 (2004), pp. 591–626.