COMMUN. MATH. SCI.
Vol. 16, No. 2, pp. 425–439

DATA-DRIVEN MODELING FOR THE MOTION OF A SPHERE FALLING THROUGH A NON-NEWTONIAN FLUID*

ZONGMIN WU† AND RAN ZHANG‡

Abstract. In this paper, we will introduce a mathematical model of the jerk equation to simulate the unstable oscillations of the motion of a falling sphere in the wormlike micellar solution. This differential/algebraic equation (DAE) is established only by learning the experimental data of time vs velocity with the sparse optimization method. To simulate the solutions of the DAE, four discretization schemes are proposed and compared. Periodic and damped harmonic motion, and nonuniform transient and sustaining oscillations can be observed for the sedimentation of a sphere through the non-Newtonian fluid in the numerical experiments. It successfully presents chaos consistent with the physical behavior, which are highly sensitive to the initial values and experimentally nonreproducible. We can conclude that our model has the ability to capture the primary pattens of the dynamics, which is more meaningful than predict an individual trajectory for the chaotic systems.

Keywords. falling sphere; non-Newtonian fluid; data-driven modeling; sparsity; differential/algebraic equation (DAE); jerk.

AMS subject classifications. 34K28; 34B60; 41A30; 65P20; 65Z05.

1. Introduction

A data-driven mathematical model and its numerical simulations will be present by learning the experimental result of the sphere falling in a wormlike micellar fluid [14]. It has been reported that in this complex non-Newtonian solution, the motion of a falling sphere experiences unsteady oscillations and never reaches a steady terminal velocity. The probable main reason for the instabilities is suggested as the breakup of the flow-induced structure (FIS) formed in the shear region around the sphere and a critical velocity gradient is required for the onset of the oscillations. For more observations about the phenomena of instabilities, see [3, 8, 9].

Actually, for Newtonian fluid, Stokes has already derived that the drag force exerted on a very small size sphere with small Reynolds number (i.e., no turbulence) is proportional to the velocity. This actually describes the motion of the falling sphere with a first-order linear ordinary differential equation (ODE) of velocity by Newton's second law of motion. However, unlike Newtonian fluid whose viscosity is constant and in which the shear stress is linearly related to strain rate, the rheology of non-Newtonian liquids includes shear banding [7], the phenomena of shear-thinning [1] and shear-thickening [18].

In fact, it is difficult to choose a proper constitutive equation characterizing the relation between the shear stress and the strain rate to simulate the sustaining oscillatory motion of the sphere falling in non-Newtonian fluid. Rajagopalan et al. [28] considered the model and numerical method of the transient motion of a sphere accelerating from

*Received: October 21, 2017; accepted (in revised form): November 22, 2017. Communicated by Pingwen Zhang.
This work is supported by NSFC Key Project (11631015), NSFC (91330201), NSFC(11571078), and Joint Research Fund by National Science Foundation of China and Research Grants Council of Hong Kong (11461161006).

†Shanghai Center for Mathematical Sciences, Shanghai Key Laboratory for Contemporary Applied Mathematics, School of Mathematical Sciences, Fudan University, Shanghai, 200433, P.R. China (zmwu@fudan.edu.cn).

‡Corresponding author, Shanghai Key Laboratory for Contemporary Applied Mathematics, School of Mathematical Sciences, Fudan University, Shanghai, 200433, P.R. China (ranzhang12@fudan.edu.cn).

rest with a multimode nonlinear Phan–Thien–Tanner model. In the Ph.D thesis of Lee [15], he attempts to simulate the continual oscillation of a falling sphere in a wormlike micellar solution and obtains decaying oscillations using the Johnson–Segalman model exhibiting the non-monotonic shear stress-strain rate relation which is also conjectured by Jayaraman and Belmonte in [14]. And Lee and Xu discuss more about the properties of the numerical schemes to deal with non-Newtonian fluid flows in [16].

Here instead of discovering the unknown physical regulation simply based on the former laws, by learning the time vs velocity data $\{(t_j, v_j^*)\}_{j=1}^n$ shown in Figure 1.1, we will establish a nonlinear second-order quadratic differential/algebraic equation (DAE) of velocity v with respect to time (i.e., the third-order DAE of z, the z-axial component of position) to model the motion of a falling sphere in the wormlike micellar solution. This model is introduced, since the third-order nonlinear system $J(z''', z'', z', z) = 0$, i.e., jerk equation [24, 33], is regarded as the simplest equation in a single variable possessing chaos, which is consistent with the randomly unstable and experimentally nonreproducible motion of a sphere in polymeric non-Newtonian fluids.

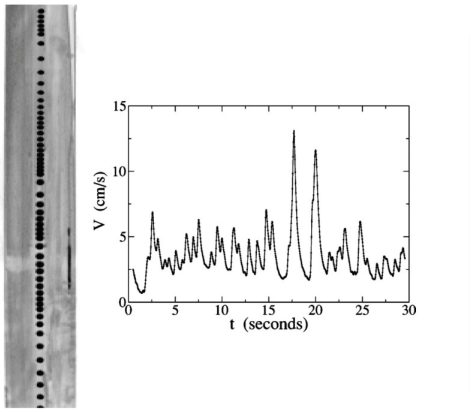


Fig. 1.1: *(Left) Collage of video images showing the sedimentation of a 3/16-inch-diameter Teflon sphere in an aqueous solution of* 6.0 *mM CTAB/NaSal (image shown is* 50 *cm in height, with* $\Delta t = 0.13$ *s). (Right) Velocity vs time for a 1/4-inch-diameter Teflon sphere falling through 9.0 mM CTAB/NaSal. This experiment was originally published by Jayaraman and Belmonte [14].*

One may be more familiar with the laws of physics described by the second-order differential equations of position, like classical Newton's equation of motion, Maxwell equations and Einstein equations, etc. However, at least three phase-space variables are needed for the unpredictable chaotic long-term solutions in continuous-time systems. Except the standard examples of chaotic autonomous systems, such as Lorenz [20] and Rössler [29] attractors, many nonlinear systems have been applied to simulate the oscillatory behavior in geophysical fluid dynamics, biological process, chemical reactions, and psychology [4].

Based on equation (2.2), we provided a parsimony four-term DAE model (2.7) of velocity by the least absolute shrinkage and selection operator (Lasso). It penalizes the unknown coefficients with L1 norm to obtain sparse result. Actually, with the advent of the theory of compressed sensing, Lasso and its generalizations are widely used as a model selection method to build a parsimony model, since they are computably

feasible numerical scheme and can efficiently promote sparse solutions. And for more information about the sparsity and optimization algorithms of Lasso and its applications in statistics and machine learning, see [10, 11, 35].

Recently, it has also been applied to physics and differential equations [6, 23, 31], and some of the researches especially focus on chaotic systems. Tran etc. [36] show that the governing equation of a sufficiently ergodic chaotic system can be exactly recovered from corrupted data in high probability by solving a partial L1-minimization problem. More discussions and numerical simulations including Lorentz system (ODE) and fluid flow (PDE) are presented in [5] by thresholded least square method to promote sparsity. However, one should notice that almost all the present literature studies the cases with simulated data, and utilize a dictionary of basis functions containing the true system to search for the 'unknown' features. In practice, one may not have any prior knowledge about the space of the dictionary. We are now going further to learn real data and intend to find the simplest model possessing chaotic behavior which can describe the motion of the falling sphere in the non-Newtonian fluid.

Four numerical schemes including backward differentiation formula (BDF) and equations (3.6), (3.7) and (3.8) are considered to simulate our DAE model (2.7), and the phenomena of the harmonic motion, transient oscillations and continual oscillations are observed which are coincident with the physical evidences.

The structure of the paper is as follows. In Section 2, we will present the main result of establishing the differential equation model. In Section 3, different numerical schemes of our model will be studied and compared. In the last Section 4, we make conclusions and more discussion.

2. Modeling

The main process of how to derive the differential equation which describes the motion of the falling sphere in the vicious fluid is investigated in this part.

Since in real life and experiments, the data are always perturbed, then we suppose

$$v_j^* \triangleq v^*(t_j) = v(t_j) + \epsilon_j, j = 1, 2, \ldots, n.$$

where $v(t_j)$ are the exact velocity values at time t_j, and ϵ_j are the independent, identically distributed noise with zero mean and finite variance σ^2.

For experimental data from the right of Figure 1.1, the size of the samples is $n = 160$ and the time step is equal with $\Delta t = 0.1875$. The samples noises usually come from the experimental setup, the process of the data visualization and other factors. Then we use the sample variance to approximate the true variance σ^2. And more about its relation to the parameter selection in our method, see Appendix A.

For the purpose of simplifying the differential equation model, firstly we will also use the nondimensionalization physical technique to preprocess the data. Based on the Stokes' terminal velocity (when Stokes' drag combined with the buoyant force balances the gravitational force)

$$v_s = \frac{2}{9}\frac{(\rho_s - \rho_f)}{\mu} g r^2,$$

we have the non-dimensional variables

$$\tilde{v} = \frac{v}{v_s}, \qquad \tilde{t} = \frac{v_s}{r} t, \tag{2.1}$$

where r and ρ_s are the radius and density of the sphere, and ρ_f and μ are the density and viscosity of the fluid, respectively.

For simplicity of the notations, we still use t and v to denote time and velocity respectively if there is no confusion.

Then we suppose that the second-order quadratic DAE of velocity with the unknown coefficients $\boldsymbol{\xi} = (\xi_0, \cdots, \xi_8)$ has the form of

$$J_2(v'', v', v) = \xi_0 v + \xi_1 v' + \xi_2 v'' + \xi_3 v^2 + \xi_4 v v' + \xi_5 v v'' + \xi_6 v'^2 + \xi_7 v' v'' + \xi_8 v''^2 = 0. \quad (2.2)$$

It is definitely tremendous possible to further consider a higher-order nonlinear DAE and higher polynomials or even trigonometric terms, such as v^3 or $\sin(v)$, to simulate the motion of the falling sphere. As aforementioned, the Stokes' law considers the equation of the motion as $mv' = F_d + F_g$ with m being the mass of the sphere, $F_d = Sv$, S being a constant associated with the fluid viscosity μ and sphere radius r, and F_g being the excess force due to the difference between the weight and buoyancy of the sphere. This is simply a first order linear ODE. And the decaying oscillations obtained by Lee [15] can also been regarded as the solution of a linear ODE. Considering the parsimony of a physical law, we propose to find a simple enough nonlinear system with the form of the Jerk Equation (2.2) which has been shown as the minimal setting for the solutions showing chaotic behavior.

To obtain the estimations of the unknown coefficients in the differential equations, various numerical approximation methods for function itself and its derivatives satisfied the differential equation have been considered. For example, Prony method employed divided difference by solving the least square problem to deal with the linear ordinary differential equations (ODE's) [25]. And then based on this idea, spline smoothing [27] and kernel smoothing [17] functions are introduced into the general nonlinear ODE's.

Multiquadric (MQ) quasi-interpolation is taken into account to provide the approximation of the velocity function $v(t)$ and its first and second derivatives with the discrete perturbation samples in this paper. Because it is a well-known meshless method suitable for scatted data while Prony related methods require equi-spaced samples, and no linear systems or least square formulations in overdetermined systems need to be computed. And it has been proved that MQ quasi-interpolation is much more stable than divided difference for noisy data [22].

As it is also mentioned in [5, 31], the high noise level will lead to the inaccurate approximation of the derivatives, and then the inaccurate recovery, so it may be necessary to filter the data. Actually, here MQ quasi-interpolation scheme is already a data smoothing approach.

More knowledge about it utilized in our model will be presented in the Appendix A.

Thus, after the preprocessing of the samples $\{(t_j, v_j^*)\}_{j=1}^n$ in Figure 1.1 by equation (2.1), with the estimators from the MQ quasi-interpolation scheme in equation (A.2) and equation (A.3), we have

$$\widehat{v}(t) = \mathcal{L}_{\mathcal{D}} v^*(t), \widehat{v}'(t) = (\mathcal{L}_{\mathcal{D}} v^*)'(t) \text{ and } \widehat{v}'' = (\mathcal{L}_D v^{1,*})'(t). \quad (2.3)$$

2.1. Lasso model. To get the unknown coefficients $\boldsymbol{\xi}$, substituting the time nodes $\{t_j\}_{j=1}^n$ and the above estimators equation (2.3) for the function itself, first and second derivatives of $v(t)$ accordingly, into DAE (2.2), we have the 'residuals',

$$P(D)\widehat{v}(t_j) = r_j, \quad j = 1, ..., n. \quad (2.4)$$

Rewrite it as the matrix form,

$$Q\boldsymbol{\xi} = \boldsymbol{r}, \quad (2.5)$$

where

$$Q=\begin{pmatrix} \widehat{v}(t_1) & \widehat{v}'(t_1) & \cdots & \widehat{v}''^2(t_1) \\ \widehat{v}(t_2) & \widehat{v}'(t_2) & \cdots & \widehat{v}''^2(t_2) \\ \vdots & \vdots & & \vdots \\ \widehat{v}(t_n) & \widehat{v}'(t_n) & \cdots & \widehat{v}''^2(t_n) \end{pmatrix}_{n\times 9}.$$

But a differential equation model with nine terms will be possibly not easy to have the stable numerical solution, not to mention the theoretical one. Of course, models with more terms lead to small 'residual', but simpler ones can reveal and interpret the law of the motion which are preferred in mathematics and physics.

Thus, based on Prony's idea which minimizes the 'residuals' in the principle of least square, we further consider the famous Lasso scheme in the following step to make a tradeoff between the 'best' and 'simple'. Then by adding the L1 regularization term $||\cdot||_1$, the estimator $\widehat{\boldsymbol{\xi}}$ of the coefficients $\boldsymbol{\xi}$ can be obtained by

$$\widehat{\overline{\boldsymbol{\xi}}}=\underset{\overline{\boldsymbol{\xi}}}{\operatorname{argmin}}||\boldsymbol{a}-B\overline{\boldsymbol{\xi}}||_2^2+\lambda||\overline{\boldsymbol{\xi}}||_1, \tag{2.6}$$

where $\boldsymbol{a}=(\widehat{v}''(t_1),...,\widehat{v}''(t_n))'$, $\overline{\boldsymbol{\xi}}=(\xi_0,\xi_1,\xi_3,...,\xi_8)'$,

$$B=-\begin{pmatrix} \widehat{v}(t_1) & \widehat{v}'(t_1) & \widehat{v}^2(t_1) & \cdots & \widehat{v}''^2(t_1) \\ \widehat{v}(t_2) & \widehat{v}'(t_2) & \widehat{v}^2(t_2) & \cdots & \widehat{v}''^2(t_2) \\ \vdots & \vdots & \vdots & & \vdots \\ \widehat{v}(t_n) & \widehat{v}'(t_n) & \widehat{v}^2(t_n) & \cdots & \widehat{v}''^2(t_n) \end{pmatrix}_{n\times 8},$$

and $\lambda>0$ is the regularization parameter to control the amount of shrinkage. Notice that we choose the linear scaling by setting the coefficient of the second order derivative term ξ_2 to be one (If there is no scale constraints, $\boldsymbol{\xi}=\mathbf{0}$ would be the solution), since the second-order derivative term v'' has the strongest correlation with the 'residual' in the full variate model and it can be guaranteed that our model is the second-order differential equation. Then the algorithm to compute equation (2.6) is implemented by solving the quadratic optimization of the corresponding constrained formulation [32].

By increasing the regularization parameter λ from zero corresponding to the least square method, to $\lambda=10$ in increments of 0.1 in our experiment, the terms v', vv', v'^2, $v'v''$ and at last v are deleted from the full model successively and no terms reenter the model. Of course, one can continue to enlarge the value of λ to obtain more sparse model. But considering the balance between small 'residual' and parsimony, the potential model can be the 5-term model with the square of the L_2-norm of the residual $||\boldsymbol{r}||_2^2=2.7730$ when $\lambda=3.2$. If λ goes up to 8 (the smallest value for the term v leaving the model), we obtain the 4-term Lasso model with $||\boldsymbol{r}||_2^2$ reaching 4.0996, which is basically twice of $||\boldsymbol{r}||_2^2$ $(=2.0460)$ of the full model.

Then we propose the model to be

$$J_2(v,v'')=\xi_0 v+v''+\xi_3 v^2+\xi_5 vv''+\xi_8 v''^2=0, \tag{2.7}$$

with $\xi_0=0.0338$, $\xi_3=-0.0105$, $\xi_5=-0.2964$, $\xi_8=-0.4555$.

And combined with appropriate initial conditions, like $v(0)$ and $v'(0)$, it becomes a solvable system. This 5-term model has the same order of error with the full model

by keeping almost half of the number of the variates. And even in the time interval [75,115] where the falling sphere experiences strong instability, the errors of the model are controlled, as shown in Figure 2.1.

Recall that at the beginning of this section, we use the nondimensionalization physical technique in equation (2.1). If we go back to the formula with physical dimension, the coefficients of equation (2.7) should depend on the properties of the sphere (radius r and density ρ_s), the fluid (viscosity μ and density ρ_f) and other physical quantities. In [38], we have further discussed the relationship between the coefficients of the model and different experimental observations to make it a complete physical law.

REMARK 2.1. From the numerical solutions implemented in the next section, as shown in Figure 3.3, the system with consistent initial conditions can successfully simulate the nonuniform oscillations of the motion of the falling sphere. However, one should notice that, because the instabilities show randomness and are experimentally nonreproducible, we are not intending to search for a model to exactly recover any individual experiment, but to extract a parsimony physical expression that can capture the main pattern of the motion. Actually, if one focus on improving the numerical accuracy of the model, it is proper to learn the discrete scheme corresponding to the continuous ODE model with the data. This has been proved numerically to enable more robust approximation, see [19].

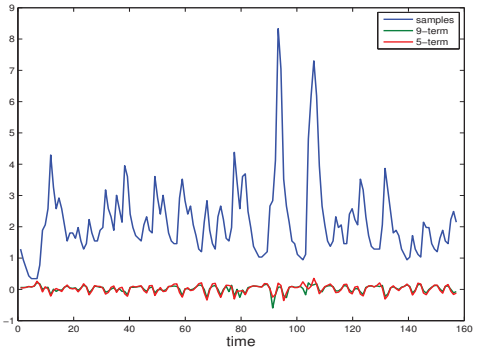


Fig. 2.1: *The 'residuals' of 9-term model with full variates (green lines) and 5-term model of Lasso (red lines) combined with the time vs velocity samples (blue lines).*

3. Numerical results

In this section, different numerical schemes will be discussed and compared to deal with the fully nonlinear differential/algebraic equation (DAE) of the form of equation (2.7). It is possible to rewrite it in the implicit ODE form. However, it is obviously an undesirable form since it has two branches and usually keeping only some branch can destroy the original physical relationship between the variable velocity and its derivatives. And this relationship is generated automatically by the modeling method. Thus numerical solution for the implicit form is studied and the typical idea is to approximate the derivatives of the variable of interest (here is the second derivative of $v(t)$) by a linear combination of the solution $v(t)$ at the current step and at several previous

steps. Thus the popular $k-$step backward differentiation formula (BDF) is

$$J_2(v^{(j)}, \frac{\pi(v^{(j)})}{(\delta t)^2}) = 0, \tag{3.1}$$

where $\pi(v^{(j)}) = \sum_{i=0}^{k} \beta_i v^{(j-i)}$ and β_i, $i = 0,1,...k$ are the coefficients of the approximation scheme for the second order derivative. In our simulations shown in Table 3.1, Houbolt's fourth order finite difference method in [30] is utilized. Then the nonlinear equations for $v^{(j)}$ at each time step is solved by Newton iteration. But difficulties are confronted in implementing the algorithm, for example, in our case equation (3.1) is a quadratic equation, then how can one decide the initial guess to make sure Newton iteration converges to the right solution from the two roots? And more related practical problems have been raised and investigated in [26].

Considering about these and the special case of our model, other three schemes are also proposed by rewriting equation (2.7) as the forms of

Scheme 1

$$v = \frac{-v'' + 0.4555v''^2 - 0.0105v^2}{0.0338 - 0.2964v''} \tag{3.2}$$

Scheme 2

$$v = \frac{0.4555v'' - 1}{0.0338 - 0.0105v - 0.2964v''} v'' \tag{3.3}$$

Scheme 3

$$v'' = \frac{-0.0338v + 0.0105v^2 + 0.4555v''^2}{1 - 0.2964v} \tag{3.4}$$

And approximating $v''^{(j)}$ by the backward differentiation formula with second order accuracy

$$v''^{(j)} = (v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})/(\delta t)^2, \tag{3.5}$$

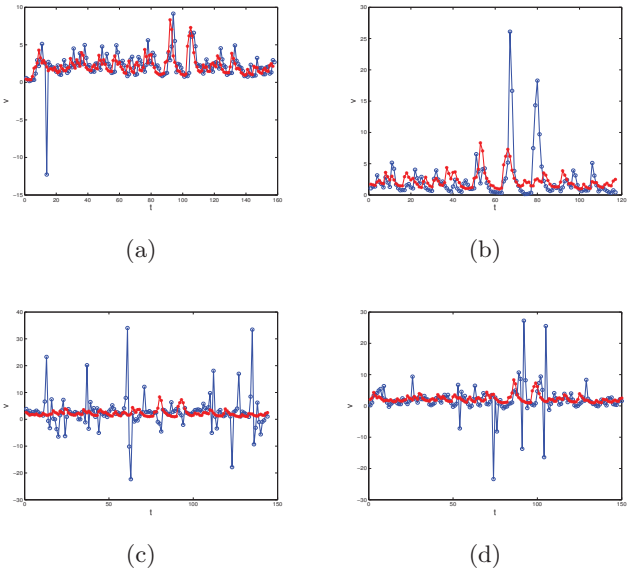
then at the each step $j = 3,4,\ldots,N$, we have the explicit iteration schemes accordingly,

Numerical Scheme 1

$$v^{(j)} = \frac{-(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})(\delta t)^2 + 0.4555(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})^2 - 0.0105(v^{(j-1)})^2(\delta t)^4}{0.0338(\delta t)^4 - 0.2964(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})(\delta t)^2} \tag{3.6}$$

Numerical Scheme 2

$$v^{(j)} = \frac{0.4555(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)}) - (\delta t)^2}{0.0338(\delta t)^2 - 0.0105v^{(j-1)}(\delta t)^2 - 0.2964(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})} \frac{v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)}}{(\delta t)^2}. \tag{3.7}$$



(a) (b)

(c) (d)

Fig. 3.1: *The comparisons between the samples* $\{v_j^*\}$ *(red lines) and estimated values* $\{\widetilde{v}_{j,i}\}$ *(blue lines) with different discrete schemes. (a) BDF; (b) Scheme 1; (c) Scheme 2; (d) Scheme 3.*

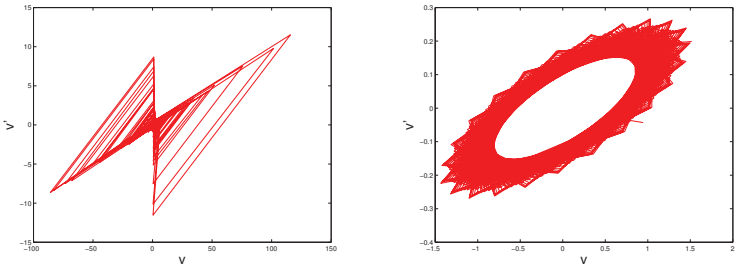
Numerical Scheme 3

$$v^{(j)} = 2v^{(j-1)} - v^{(j-2)} + \frac{-0.0338v^{(j-1)}(\delta t)^4 + 0.0105(v^{(j-1)})^2(\delta t)^4 + 0.4555(v^{(j-1)} - 2v^{(j-2)} + v^{(j-3)})^2}{(\delta t)^2(1 - 0.2964v^{(j-1)})}. \tag{3.8}$$

And the initial conditions $\{v^{(j)}\}_{j=0}^2$ have to be consistency, which means that $v^{(2)}$ is determined by $v^{(0)}$ and $v^{(1)}$ with the relation of equation (2.7).

Based on these four numerical schemes, we first calculate the local errors between the samples and the results of the model, i.e., $e_j = v_j^* - \widetilde{v}_{j,i}$ where $\widetilde{v}_{j,i} = F_i(v_{j-1}^*, v_{j-2}^*, v_{j-3}^*)$ is the estimated value computed with BDF $(i=0)$ and Scheme i $(i=1,2,3)$, respectively. Since the sample time step δt is 1 (after nondimensionalization), then it is directly utilized in the simulations initially. However, as the following analysis, the Scheme 1, 2 and 3 do not show unsteady oscillations when $\delta t = 1$. Thus, we choose $\delta t = 14$, 5 and 3 for Scheme 1, 2 and 3, respectively, which promote smaller mean values of the errors. From Figure 3.1, different choice of δt with the samples still makes the approximations less accurate. BDF almost captures all the oscillations except one big error, and the first scheme enlarges the two peaks and has some delay of time as well, while the scheme 2 and 3 provide totally different results compared with the samples. This again inspires one to learn the discrete model as aforementioned in Remark 2.1 if one focuses on the computation accuracy.

We now will consider to explore the main patterns of the dynamical system with varying time steps and initial values. In the following simulations, the samples of velocity values $\{v_p^*\}_{p=1}^3$ will be directly used to estimate the initial conditions $\{v^{(j)}\}_{j=0}^2$. By applying the aforementioned iteration schemes, more than 100,000 experiments have been done with different time steps δt and varying perturbations e to the initial values



(a) (b)

Fig. 3.2: *(a) plot of trajectory v–v' for $\delta = 10$ of Scheme 2; (b) plot of trajectory v–v' for $\delta = 8$ of Scheme 3.*

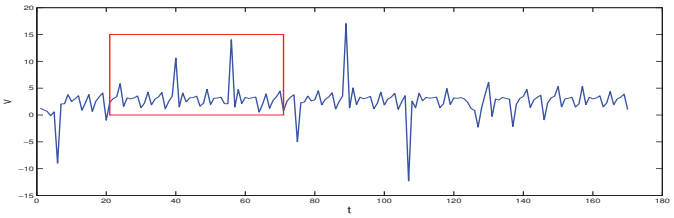


Fig. 3.3: *Numerical result of Scheme 2 with $\delta t = 4.8$, initial values being $\{v_p^*\}_{p=1}^3$ and $N = 169$.*

$\{v^{(j)}\}_{j=0}^2 = \{v_p^*\}_{p=1}^3$ (i.e., different initial values). We always keep the initial velocity (i.e., $v^{(0)}$) unchanged, but vary the acceleration and jerk (i.e., $v^{(1)}$ or $v^{(2)}$). We choose the time steps δt ranging from 0.1 to 10 in increments of 0.1 and the number of the iterations being 500. Some of the typical numerical examples are exhibited in Table 3.1-3.2.

We have the following observations:

- For smaller δt, the unstable sustaining oscillations are observed in BDF, while divergent and even blow up solutions are obtained for Scheme 1, 2 and 3. Then as δt increases, the unsteady oscillations will end up with harmonic motions after some critical point for $0.7 \le \delta t \le 1$ in BDF and $\delta t = 2.5$ and 2.6 in Scheme 2. If δt is larger than 1, the numerical results for BDF experience gradual decay and reach some terminal velocity near zero which is consistent with the classical Stokes' law. And the larger the δt is, the faster the decay is. However, sustaining unstable oscillations appear when $\delta t > 2.5$ and $\delta t > 2.6$ for Scheme 1 and Scheme 2, respectively. For Scheme 3, no such large unsteady fluctuations have been detected and their behavior are periodic ($0.5 \le \delta t \le 4$) or nearly periodic ($4 < \delta t \le 8.2$).
- In Figure 3.2, we show the plots of trajectory (v–v') for $\delta t = 10$ of Scheme 2 and $\delta t = 8$ of Scheme 3. The former has the similar structure with Lorenz attractor which resembles a butterfly and the later shows the evidence that the behavior is not periodic as aforementioned.
- Since the instability of the motion of the falling sphere appears randomly and

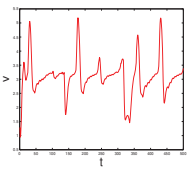
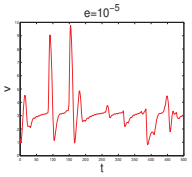
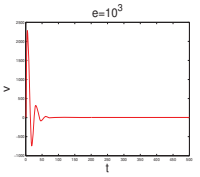
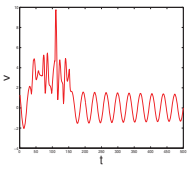
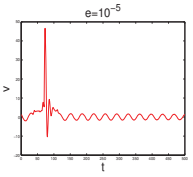
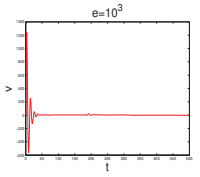
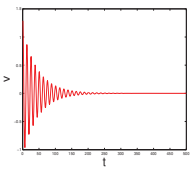
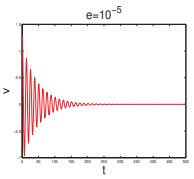
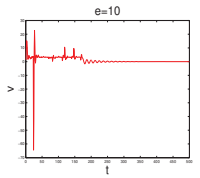
	$e=0$	small e	large e
$\delta t < 0.7$ unsteady oscillations (shown with $\delta t = 0.5$)			
$0.7 \le \delta t \le 1$ unsteady oscillations then harmonic motion (shown with $\delta t = 1$)			
$1 < \delta t \le 10$ gradual decay (shown with $\delta t = 3$)			

Table 3.1: *Numerical results for BDF for different δt and varying perturbations e to the initial values $\{v_p^*\}_{p=1}^3$ with selected examples are shown.*

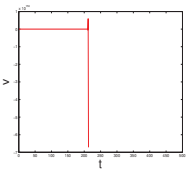
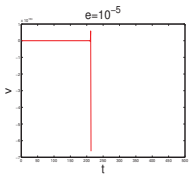
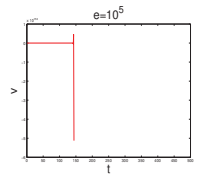
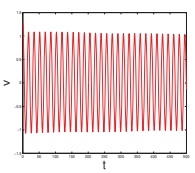
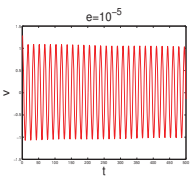
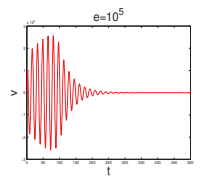
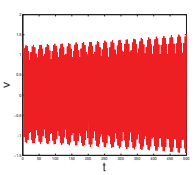
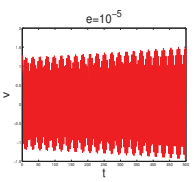
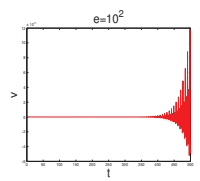
	$e=0$	small e	large e
$\delta t \le 0.4$ blow up (shown with $\delta t = 0.4$)			
$0.5 \le \delta t \le 4$ harmonic motion (shown with $\delta t = 2$)			
$4 < \delta t \le 8.2$ nearly harmonic motion (but not) (shown with $\delta t = 8$)			

Table 3.2: *Numerical results for Scheme 3 for different δt and varying perturbations e to the initial values $\{v_p^*\}_{p=1}^3$ with selected examples are shown.*

is experimentally non-repeatable, the harmonic motion, transient oscillations and continual oscillations observed in the numerical simulations are consistent with the physical evidences in a way. Actually, even though it is not possible to simulate the exactly same pattern, in the longer time domain of our simulation than that of the sample data, one can find some segment of the simulations showing the similar motion with the experiment in Figure 1.1, for example, the part marked in the rectangle of Figure 3.3 implemented by Scheme 2 with initial values being $\{v_p^*\}_{p=1}^3$, $\delta t = 4.8$ and $N = 169$.

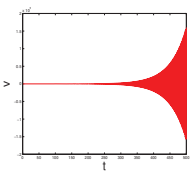
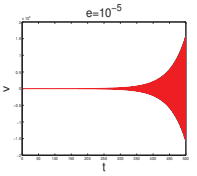
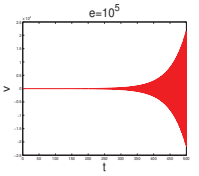
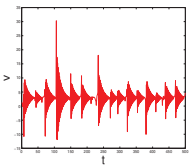
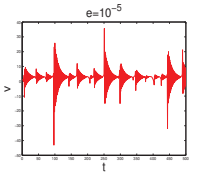
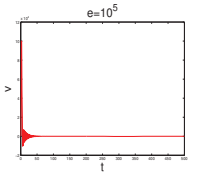
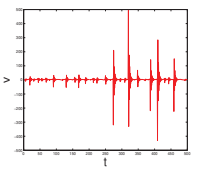
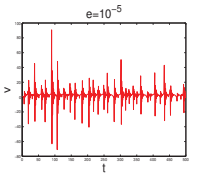
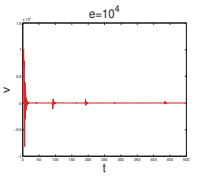
	$e = 0$	small e	large e
$\delta t \le 2.5$ divergent solution or blow up (shown with $\delta t = 2.5$)			
$2.5 < \delta t \le 10$ sustaining unsteady oscillations (shown with $\delta t = 2.8$ and $\delta t = 4$, respectively)			
			

Table 3.3: *Numerical results for Scheme 1 for different δt and varying perturbations e to the initial values $\{v_p^*\}_{p=1}^3$ with selected examples are shown.*

Remark 3.1. For the cases showing continuous unsteady oscillations, the velocity series are highly sensitive to initial conditions even for quite small perturbations $e = 10^{-5}$, which actually just demonstrates that the motion of the falling sphere is not experimentally reproducible. While for large enough e (different lower bound for each scheme and step size), as shown in the last column of Table 3.1 for $\delta t \le 1$, Table 3.3 for $2.5 < \delta t \le 10$, Table 3.4 for $2.5 < \delta t \le 10$ and Table 3.2 for $0.5 \le \delta t \le 4$, such motions can be observed to end up with some constant terminal velocity after the unsteady oscillations. We should also mention that the divergent or blow up solutions will always stay at the same state no matter how the initial values change, but a divergent solution appears at e larger than 10^2 for Scheme 3 of $4 < \delta t \le 8.2$ while there is nearly harmonic motion for smaller e (i.e., smaller speed). These phenomena can be related to the control of high speed motion.

Remark 3.2. Smaller step sizes tend to exhibit more unsteady behavior for the simulations of the falling sphere. In Table 3.1 of BDF, the nonuniform and uniform unstably oscillatory motions are observed when $\delta t \le 1$, and the damped harmonic motion

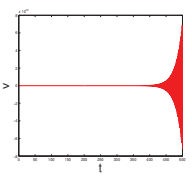
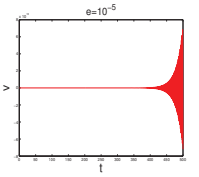
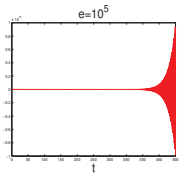
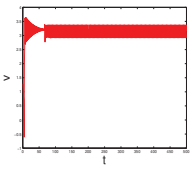
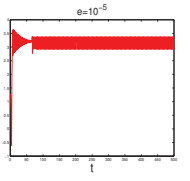
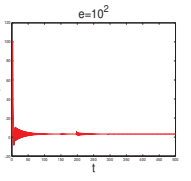
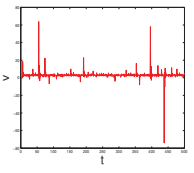
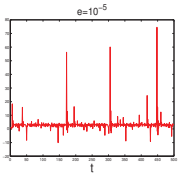
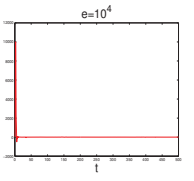
	$e=0$	small e	large e
$\delta t \le 2.4$ divergent solution or blow up (shown with $\delta t = 2.3$)			
$\delta t = 2.5, 2.6$ unsteady oscillations then nearly harmonic motion (shown with $\delta t = 2.5$)			
$2.7 \le \delta t \le 10$ sustaining unsteady oscillations (shown with $\delta t = 4.2$)			

Table 3.4: *Numerical results for Scheme 2 for different δt and varying perturbations e to the initial values $\{v_p^*\}_{p=1}^3$ with selected examples are shown.*

decay faster to reach the 'terminal velocity' as the increase of δt. And for Scheme 1, 2 and 3, when the time step is small, the solutions are even blow up. It can also be observed that after the sudden large increase in the velocity, the motion will undergo a series of gradually small waves until it reaches a temporary stable state. Then another instability appears and the period repeats again. However, with the growing of the time step, rapid descending in the velocity is in the wake of the fast ascending. It means that the smaller the amount of the fluid parcel (the smaller time step) is simulated, the more fluctuations and unstable the motion is.

REMARK 3.3. Like the Stokes terminal velocity of a falling sphere in the Newtonian fluid after the balance of the drag force with buoyant force and the gravity, one also expect to explore the 'terminal (or settling) velocity' in the non-Newtonian polymer solutions. And actually, we found that the constant velocity, periodic and nonuniform motion around some constant velocity can all be regarded as the settling states. And a larger step size δt always indicates a more stable terminal state, i.e., constant velocity, since the non-Newtonian part of the fluid is not kept, which is consistent with Remark 3.2.

REMARK 3.4. One should notice that the velocity can reach negative value while all the samples are all positive. But it is also observed in other physical experiments that the falling sphere moves in the reverse direction of gravity [8] and it is possible to happen when too much drag force is exerted to the falling sphere at some point. This phenomenon differs from the classical Newton's idea.

4. Conclusion and discussion

By simply learning the experimental data, we build the jerk equation (2.7) and different numerical schemes to study the motion of the falling sphere in the wormlike micellar solution. And to approximate the coefficients of the DAE, we introduce the feature selection method, L1 regularization with MQ quasi-interpolation scheme to obtain the parsimony result equation(2.7). More physical interpretations and investigations about the chaotic behavior of the model will become the further works. As the era of big data, this data-driven methodology of how to explore the laws of physics used in this paper can be generalized to chemistry, biology and economy etc., where large amount of data samples are available in the experiments and financial market, respectively.

Appendix A. MQ quasi-interpolation. Actually, quasi-interpolation formulations based on the multiquadric kernel function, or more generally radial basis functions for solving differential equations, and other approximation problems have been widely investigated and applied. Hardy [13] firstly introduced MQ function into quasi-interpolation for its well performance in both theoretical basis and numerical examples. Then more MQ quasi-interpolation schemes and their properties were presented, see [2,21,37]. The stability and asymptotic behavior of MQ quasi-interpolation schemes for estimating high-order derivatives fitting observations with independent, identically distributed white noises have also been investigated in [12, 22].

The information about quasi-interpolation method used in our model will be presented here. For a function $f:[a,b]\to\mathbb{R}$ and exact data value set $\{(t_j,f(t_j))\}_{j=1}^n$, $a=t_1<t_2<\cdots<t_n=b$ and $h:=\max_j(t_{j+1}-t_j)$, define the original MQ quasi-interpolation scheme as

$$\widehat{f}(t)\triangleq\mathcal{L}f(t)=\sum_j f(t_j)\psi_j(t), \tag{A.1}$$

where $\psi_j(t)$ is the linear combinations of the multiquadric functions

$$\phi_j(t)=\sqrt{c^2+(t-t_j)^2}$$

and c is shape parameter, which is usually chosen to be a small number. Since as c tends to zero, $\phi(t)$ will tend to $|t|$, then MQ actually tends to piecewise linear interpolation but possesses smoothness.

And we introduce the $\mathcal{L}_\mathcal{D}$ quasi-interpolation scheme to deal with the boundary, since no derivatives of f at endpoints are required and the shape preserving and convergent properties are proved in [37]. Then for noisy data $\{f_j^*\}_{j=1}^n$, the estimators of the function and the first derivative ($k=0,1$) are

$$\widehat{f}^{(k)}\approx(\mathcal{L}_D f^*)^{(k)}(t)\triangleq f_1^*\alpha_1^{(k)}(t)+f_2^*\alpha_2^{(k)}(t)+\sum_{j=3}^{n-2}f_j^*\psi_j^{(k)}(t)+f_{n-1}^*\alpha_{n-1}^{(k)}(t)+f_n^*\alpha_n^{(k)}(t), \tag{A.2}$$

where

$$\alpha_1(t)=\frac{1}{2}+\frac{\phi_2(t)-(t-t_1)}{2(t_2-t_1)},$$
$$\alpha_2(t)=\frac{\phi_3(t)-\phi_2(t)}{2(t_3-t_2)}-\frac{\phi_2(t)-(t-t_1)}{2(t_2-t_1)},$$

$$\alpha_{n-1}(t) = \frac{(t_n - t) - \phi_{n-1}(t)}{2(t_n - t_{n-1})} - \frac{\phi_{n-1}(t) - \phi_{n-2}(t)}{2(t_{n-1} - t_{n-2})},$$
$$\alpha_n(t) = \frac{1}{2} - \frac{(t_n - t) - \phi_{n-1}(t)}{2(t_n - t_{n-1})},$$

with $\phi_j(t)$ defined as aforementioned and

$$\psi_j(t) = \frac{\phi_{j+1}(t) - \phi_j(t)}{2(t_{j+1} - t_j)} - \frac{\phi_j(t) - \phi_{j-1}(t)}{2(t_j - t_{j-1})}.$$

And to achieve theoretical convergence and better numerical stabilization, for approximating second-order derivative, instead of directly using $(\mathcal{L}_D f^*)''(t)$, we use the technique of quasi-interpolation and then differentiation introduced in [34]. Specifically, denote $\{f^{1,*}(t_j)\}_{j=1}^n \triangleq \{(\mathcal{L}_D f^*)'(t_j)\}_{j=1}^n$ as the approximated value of the first derivative of f, then we apply the quasi-interpolation operator $\mathcal{L}_D$ again to this 'samples' as $\mathcal{L}_D f^{1,*}(t)$, and finally differential operator is taken to have $(\mathcal{L}_D f^{1,*})'(t)$ as the estimator of the second derivative of f,

$$\widehat{f}''(t) \approx (\mathcal{L}_D f^{1,*})'(t). \tag{A.3}$$

And by successively following this step, one can obtain the approximation of higher-order derivatives of the given function.

REMARK A.1. The choice of the bandwidth parameter c follows the Remark 3.3 in [12], where the convergence of the estimation of the kth-order derivative (in our case, $k=0,1,2$) is valid in the sense of MSE (Mean-Squared Error) if $\sigma^2 \le h^{\frac{k+4}{k+1}}$, $c = \mathcal{O}(h^{\frac{1}{k+1}})$; otherwise, $c = \mathcal{O}((\sigma^2 h)^{\frac{1}{2k+5}})$. Although originally it is proved to hold for MQ trigonometric quasi-interpolation, it works well numerically in the quasi-interpolation schemes we proposed. And notice that the true variance σ^2 is not known, in practice it will be approximated by the unbiased sample variance $s^2 = \frac{1}{n-1}\sum_{i=1}^n (v_j^* - \frac{1}{n}\sum_{i=1}^n v_j^*)^2$.

REFERENCES

[1] M.T. Arigo and G.H. McKinley, *An experimental investigation of negative wakes behind spheres settling in a shear thinning viscoelastic fluid*, Rheologica Acta, 37(4):307–327, 1998.

[2] R.K. Beatson and M.J.D. Powell, *Univariate multiquadric approximation: Quasi-interpolation to scattered data*, Constructive Approximation, 8(3):275C288, 1992.

[3] R.B. Bird, R.C. Armstrong, and O. Hassager, *Dynamics of polymetric liquids, Vol. 1, Fluid Mechanics*, Wiley Interscience, New York, 1987.

[4] M. Brøns, T.J. Kaper, and H.G. Rotstein, *Introduction to focus issue: mixed mode oscillations: experiment, computation, and analysis*, Chaos, 18, 015101, 2008.

[5] S.L. Brunton, J.L. Proctor, and J.N. Kutz, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proceedings of the National Academy Sciences, 113(15):3932–3937, 2016.

[6] R.E. Caflish, S.J. Osher, H. Schaeffer, and G. Tran, *PDEs with compressed solutions*, Commun. Math. Sci., 13:2155–2176, 2015.

[7] M.A. Calabrese, S.A. Rogers L Porcar, and N.J. Wagner, *Understanding steady and dynamic shear banding in a model wormlike micellar solution*, Journal of Rheology, 60, 2016.

[8] S. Chen and J.P. Rothstein, *Flow of wormlike micelle solutions past a confined circular sphere*, Journal of Non-Newtonian Fluid Mechanics, 116:205–234, 2004.

[9] R.P. Chhabra, *Bubbles, Drops and Particles in Non-Newtonian Fluids*, CRC Press, Boca Raton, FL, 1993.

[10] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, *Least angle regression*, Annals of Statistics, 32(2):407–451, 2004.

[11] J. Friedman, T. Hastie, and R. Tibshirani, *Regularization paths for generalized linear models via coordinate decent*, Journal of Statistical Software, 33(1):1–22, 2010.

[12] W.W. Gao and R. Zhang, *Multiquatric trigonometric spline quasi-interpolation for numerical differentiation of noisy data: a stochastic perspective*, Numerical Algorithms, 2017.
[13] R. Handy, *Multiquadric equations of topography and other irregular surfaces*, Journal of Geophysical Research, 76:1905–1915, 1971.
[14] A. Jayaraman and Andrew Belmonte, *Oscillations of a solid sphere falling through a wormlike micellar fluid*, Physical Review E67, 065301(R), 2003.
[15] Y.J. Lee, *Modelling and simulations of non-Newtonian fluid flows*, PhD Thesis, The Pennsylvania State University, University Park, PA, 2004.
[16] Y.J. Lee and J. Xu, *New formulations, positivity preserving discretizations and stability analysis for non-Newtonian flow models*, Computer Methods in Applied Mechanics and Engineering, 195:1180–1206, 2006.
[17] H. Liang and H. Wu, *Parameter Estimation for differential equation models using a framework of measurement error in regression models*, Journal of the American Statistical Association, 103(484):1570–1583, 2008.
[18] C. Liu and D.J. Pine, *Shear-induced gelation and fracture in micellar solutions*, Physical Review Letters, 77:2121–2124, 1996.
[19] Z.C. Long, Y.P. Lu, X.Z. Ma, and B. Dong, *PDE-Net: Learning PDEs from data*, preprint, 2017.
[20] E.N. Lorenz, *Deterministic nonperiodic flow*, Journal of the Atmospheric Sciences, 20:130–141, 1963.
[21] L.M. Ma and Z.M. Wu, *Approximation to the k-th derivatives by multiquadric quasi-interpolation method*, J. Comput. Appl. Math., 231:925–932, 2009.
[22] L.M. Ma and Z.M. Wu, *Stability of Multiquadric quasi-interpolation to approximate high order derivatives*, Science China Mathematics, 53:925-932, 2010.
[23] L.J. Nelson, G.L.W. Hart, F. Zhou, and V. Ozolins, *Compressive sensing as a paradigm for building physics models*, Physical Review B, 87, 035125, 2013.
[24] W. Marszalek and Z. Trzaska, *Mixed-mode oscillations and chaotic solutions of jerk (Newtonian) equations*, J. Comput. Appl. Math., 262:373–383, 2014.
[25] M.R. Osborne, *Some special nonlinear least squares problems*, SIAM Journal on Numerical Analysis, 12:571-592, 1975.
[26] L. Petzold, *Differential/Algebraic equations are not ODE's*, SIAM Journal on Scientific and Statistical Computing, 3(3):367-384, 1982.
[27] J.O. Ramsay, G. Hoocker, D. Campbell and J. Cao, *Parameter estiamtion for differential equations: a generalized smoothing approach*, Journal of the Royal Statistical Society, Ser.B, 69:741–796, 2007.
[28] D. Rajagopalan, M.T. Arigo, and G.H. McKinley, *The sedimentation of a sphere through an elastic fluid Part2. Transient motion*, Journal of Non-Newtonian Fluid Mechanics, 65:17–46, 1996.
[29] O.E. Rösslor, *An equation for continuous chaos*, Physics Letters A, 57, 397–398, 1976.
[30] C. Semler, W.C. Gentleman, and M.P. Païdoussis, *Numerical solutions of second order implicit non-linear ordinary differential equations*, Journal of Sound and Vibration, 195(4):553–574, 1996.
[31] H. Schaeffer, *Learning partial differential equations via data discovery and sparse optimization*, Proceedings: Mathematical, Physical and Engineering Sciences, 473(2197), 20160446, 2017.
[32] M. Schmidt, *Least squares optimization with L1-norm regularization*, CS542B Project Report, 98:230–238, 2005.
[33] J.C. Sprott, *Some simple chaotic flows*, Physical Review E, 50(2):647–650, 1994.
[34] Z.J. Sun, Z.M. Wu, and W.W. Gao, *Order preserving derivative approximation with quasi-interpolation*, preprint, 2017.
[35] R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society, Series B, 58:267–288, 1996.
[36] G. Tran and R. Ward, *Exact recovery of chaotic systems from highly corrupted data*, Multiscale Model. Simul., 15(3):1108–1129, 2017.
[37] Z.M. Wu and R. Schaback, *Shape preserving properties and convergence of univariate multiquadric quasi-interpolation*, Acta Mathematicae Applicatae Sinica, 10:441-446, 1994.
[38] Z.M. Wu and R. Zhang, *Learning physics by data for the motion of a sphere falling in a non-Newtonian fluid*, preprint, 2017.