A machine learning based Bayesian optimization solution to non-linear responses in dusty plasmas

IOP Publishing Mach. Learn.: Sci. Technol. 2 (2021) 035017 https://doi.org/10.1088/2632-2153/abe7b7

MACHINE LEARNING
Science and Technology

OPEN ACCESS

RECEIVED
2 November 2020

REVISED
25 January 2021

ACCEPTED FOR PUBLICATION
18 February 2021

PUBLISHED
14 June 2021

PAPER

A machine learning based Bayesian optimization solution to non-linear responses in dusty plasmas

Zhiyue Ding, Lorin S Matthews and Truell W Hyde

Department of Physics, Baylor University, One Bear Place 97316, Waco, TX, United States of America

E-mail: dingzhiyue@gmail.com

Keywords: dusty plasma, machine learning, non-linear dynamics

Abstract

Nonlinear frequency response analysis is a widely used method for determining system dynamics in the presence of nonlinearities. In dusty plasmas, the plasma–grain interaction (e.g. grain charging fluctuations) can be characterized by a single-particle non-linear response analysis, while grain–grain non-linear interactions can be determined by a multi-particle non-linear response analysis. Here a machine learning-based method to determine the equation of motion in the non-linear response analysis for dust particles in plasmas is presented. Searching the parameter space in a Bayesian manner allows an efficient optimization of the parameters needed to match simulated non-linear response curves to experimentally measured non-linear response curves.

1. Introduction

Machine learning (or deep learning) has recently become one of the hottest analysis techniques in the scientific world as application of this powerful numerical method has proven useful in solving problems across a wide range of fields. For example, convolutional neural networks (LeNet, AlexNet) [1, 2] now fulfill object recognition tasks to a high degree of accuracy; recurrent neural networks (LSTM) [3] are improving computational understanding of natural language; reinforcement learning agents are out-performing human experts in strategic decision making (AlphaGo) [4, 5]; and generative adversarial networks (GANs) [6] are showing the ability to create music, paintings and dialogue in a human manner. In addition to industrial applications, machine learning techniques are now also being applied to solve physics problems. Examples include the prediction of molecular atomization energies by employing regression models [7], the application of a neural network to solve quantum many-body problems [8], and crystallization recognition through the use of a shallow neural network [9].

In this paper, a machine learning-based method is applied to non-linear response problems in dusty plasmas. Dusty plasmas [10–12] are systems containing both weakly ionized gas and charged micron-sized dust particles. Due to the higher thermal velocities of electrons compared to ions, dust particles in a dusty plasma become negatively charged [13] in response to the frequent collisions between the plasma particles and a dust grain's surface. Dust particle behavior in plasmas is determined by many factors, with the restoring confinement caused by the balance between the electrostatic force and gravity, the neutral gas drag, and particle-particle interactions between dust particles among the primary of these. Understanding the physics behind dust particle behavior (i.e. investigating these factors) is one of the most important tasks in dusty plasmas. One of the ways in which this can be fulfilled is by studying the response of the particles to external excitations [14, 15]. This is known as the non-linear frequency response analysis, which has a wide application in mechanics, material science and nano science [16–18].

Here, a Bayesian optimization framework [19] is used to resolve a non-linear response analysis [20–23] in a numerical manner in a dusty plasma. The undetermined coefficients in the equation of motion for a dust grain are derived by optimizing the simulated motion to match that obtained from experimental results. These are compared to the analytic results from a multiple-scale perturbation method. The non-linearity is measured to the 4th order in displacement, which helps correctly characterize the potential energy of particle in the plasma sheath to an accuracy not obtained before.

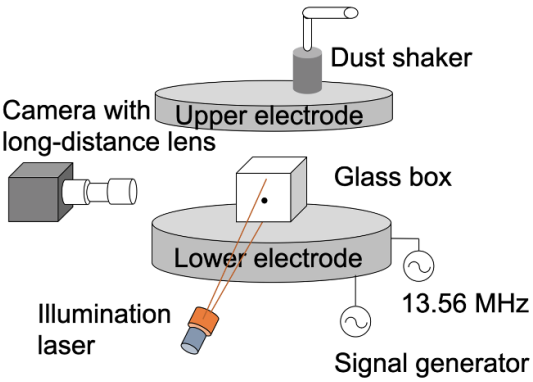

Figure 1. Sketch of the modified GEC RF reference cell.

It is important to note that this framework is not limited to non-linear response analysis, but can also be applied to the more general case of physics problems where the experimental results can be reproduced by simulations. In these cases, undetermined physical quantities can be revealed efficiently (especially when the simulation is very computational expensive) by optimizing the simulations to experimental results in this Bayesian manner.

2. Experiment and Bayesian optimization

The experiment which will be discussed in this paper was conducted in a modified Gaseous Electronics Conference (GEC) RF reference cell (see figure 1) filled with argon gas. A single melamine formaldehyde (MF) particle having a diameter of $8.89 \pm 0.09\ \mu$m was inserted into a glass box (height: 20 mm, length: 18 mm, width: 18 mm) placed on the lower electrode which was powered at 13.56 MHz. The plasma power and pressure were fixed at 1.68 W and 40 mTorr, respectively. The MF particle was levitated in the plasma sheath region due to the balance between gravity and the electrostatic force produced by the negatively charged lower electrode. The dust particle was illuminated by a laser sheet (wavelength of 660 nm) with the resulting motion recorded at a rate of 500 fps by a high speed camera mounted at the side port of the cell.

A primary amplitude-frequency response curve was measured by applying a sinusoidal excitation signal to the lower electrode with a fixed amplitude at various frequencies. Particle motion was recorded and then transformed into the frequency domain (FFT spectrum) using a Fourier transform for each value of the excitation frequency. The peak height of the FFT spectrum at the excitation frequency was measured, providing the primary response at this excitation. The secondary (super-harmonic) response to the excitation (a non-linear response), can also be measured from the peak height of the FFT spectrum at twice the excitation frequency. (Sample response curves are shown in figure 2.)

The motion of a single such particle levitating inside the plasma sheath under a vertical sinusoidal excitation can be modeled as a forced oscillator [20],

$$\ddot{x} + \mu\dot{x} + \omega^2 x + \alpha x^2 + \beta x^3 = F\exp(i\Omega t) + \text{c.c.}, \tag{1}$$

where μ is the neutral drag coefficient, ω is the restoring constant, α and β are the second and third order derivatives of the restoring field, Ω is the frequency of the sinusoidal excitation, F is the amplitude of the excitation (in units of acceleration) and c.c. stands for the complex conjugate. The neutral drag coefficient can be theoretically identified as [24]:

$$\mu = \delta\frac{4\pi}{3}Nm_n\bar{c}_n r_p^2, \tag{2}$$

where N, m_n, $\bar{c}_n$ and r_p are the neutral gas number density, mass of a neutral gas atom, the thermal speed of the gas and the radius of the dust particle, respectively. The coefficient δ accounts for the type of reflection or absorption of the neutral gas particle. An estimate of the drag coefficient for the given experimental conditions yields $\mu = 7.71$–$11.09\ \text{s}^{-1}$ [25], with δ in a range from 1.26 to 1.44 (with uncertainty considered) [26]. Often, the effective restoring force experienced by the particle at equilibrium can be approximated as a linear function in displacement (i.e. $-\omega^2 x$ where ω is the natural resonance frequency) under the assumption that the particle is levitating in a region where the sheath can be considered to exhibit a perfect parabolic sheath potential [27, 28]. For the given experimental conditions, ω is estimated to be 11.3–12.6 Hz

by taking the spectral power density of the particle's thermal motion, as described in [29]. Unfortunately, this approximation becomes invalid in most realistic situations, such as where charge fluctuations are considered [21, 22], or when the oscillation of the dust particle is so large that the sheath potential can no longer be considered parabolic. In this case, it is necessary to extend the restoring force to the non-linear regime as $-\omega^2 x - \alpha x^2 - \beta x^3$ with terms higher than $O(x^3)$ ignored for simplicity.

Based on equation (1), the particle motion as a function of time $x(t)$ under an excitation with frequency Ω can be simulated (or numerically solved) given a set of known parameters $\theta = \{\mu, \omega, \alpha, \beta, F\}$. In this case, the particle motion was simulated employing the velocity Verlet algorithm, which updates the position and velocity in each iteration as

$$\begin{aligned} x(t+dt) &= x(t) + v(t)dt + \frac{a(t)}{2}(dt)^2, \\ v(t+dt) &= v(t) + \frac{a(t+dt)+a(t)}{2}dt, \end{aligned} \tag{3}$$

where dt is the time step of the simulation, $v(t)$ is the velocity at time t and $a(t)$ is the acceleration normalized by the particle mass at time t as determined by equation (1):

$$a(t) = -\mu v(t) - \omega^2 x(t) - \alpha x^2(t) - \beta x^3(t) + 2F\cos(\Omega t). \tag{4}$$

Following the same approach described above for the experiment, the primary and secondary amplitude-frequency response curves can also be measured from the simulated particle motion $x(t)$ by varying the excitation frequency Ω over the range of excitation frequencies used in the experiment.

This allows a parameter set $\theta^* = \{\mu^*, \omega^*, a^*, b^*, F^*\}$ characterizing the properties of the dust motion (which depend on properties of the nearby plasma environment) to be determined by searching the parameter space for the optimal set of parameters that generates a simulated amplitude-frequency response curve which most closely matches the experimentally measured amplitude-frequency response curve. This process can be quantified as

$$\theta^* = \underset{\theta}{\operatorname{argmin}}(L(R_e, R_s(\theta))), \tag{5}$$

where R_e represents the experimentally measured response curve and $R_s(\theta)$ represents the simulated response curve for a given set of parameters $\{\mu, \omega, \alpha, \beta, F\}$. $L(R_e, R_s)$ is a measure of the difference between the experimentally measured and simulated response curves. In order to quantify this difference, we define L as a function $L : \theta = \{\mu, \omega, \alpha, \beta, F\} \mapsto \mathbb{R}$ that maps a set of parameters to a real value which measures the 'distance' between the experimentally measured and the simulated response curve as

$$L(\theta) = \sum_{i=1}^{N} \left(\frac{r_e(\Omega_i) - r_s(\Omega_i, \theta)}{r_e(\Omega_i)} \right)^2, \tag{6}$$

where $r_e(\Omega_i)$ and $r_s(\Omega_i)$ are the experimentally measured and simulated response amplitudes at the excitation frequency Ω_i, respectively. The summation is carried out over the span of the excitation frequencies employed in the experiment. The difference are squared to ensure that $L(\theta)$ does not yield negative values, which guarantees the existence of minimal points. It is important to mention that this type of loss function $L(\theta)$ may not be well-defined everywhere. For unrealistic parameters sets, i.e. sets that either have no physical meaning or are not suitable for describing the condition of the plasma sheath, the simulated non-linear response curves diverge, resulting in an undefined distance function. In these cases, a large value is assigned to the distance function (e.g. $L = 10^5$) in order to ensure optimization success.

As can be seen from equation (6), calculation of the loss function $L(\theta)$ for even one set of parameters requires multiple simulations of the particle's motion, i.e. one for each excitation frequency used for measuring the response curve in the experiment. For example, the response curve shown in figure 2 requires 71 independent simulations to calculate the distance function for just one set of parameters. This is computationally expensive and, as such, a minimization of the distance function $L(\theta)$ based on a random search of the parameter space θ is infeasible.

Therefore, this loss function is minimized employing a Bayesian optimization. This technique has shown great promise in machine learning, especially for the fine tuning of neural networks for model selection. Instead of randomly searching the parameter space and then conducting simulations for each set, only those parameter sets selected in a Bayesian manner are simulated. A surrogate function f is introduced to model

Algorithm: Bayesian optimization (TPE).

Define a loss function $L(\theta)$;
Initialing a data set $\mathcal{D}_{1:t} = \{\theta_{1:t}, L(\theta_{1:t})\}$ by randomly sampling parameters $\theta_{1:t}$;
while NOT converged **do**
  Calculate f^* as the lower 25% quantile of the loss functions for the up-to-date data set $\mathcal{D}_{1:t}$;
  Estimate $l(\theta)$ using data with loss function less than f^*;
  Estimate $g(\theta)$ using data with loss function greater than f^*;
  Calculate θ_{t+1} that maximizes $l(\theta)/g(\theta)$;
  Run simulation on θ_{t+1}, and update the data set $\mathcal{D}_{1:t}$ with $\{\theta_{t+1}, L(\theta_{t+1})\}$;
end

the distribution of the value of the loss function $L(\theta)$. The posterior distribution of this surrogate function at the parameter θ given the data observed $\mathcal{D}_{1:t} = \{\theta_{1:t}, L(\theta_{1:t})\}$ can be derived using Bayes' law:

$$p(f|\theta;\mathcal{D}_{1:t}) = \frac{p(\theta|f;\mathcal{D}_{1:t})p(f;\mathcal{D}_{1:t})}{p(\theta;\mathcal{D}_{1:t})}, \tag{7}$$

where tree-structured Parzen density estimators [30] (a generative model) are used to model the likelihood function $p(\theta|f;\mathcal{D}_{1:t})$ defined as

$$p(\theta|f;\mathcal{D}_{1:t}) = \begin{cases} l(\theta), \text{if } f < f^* \\ g(\theta), \text{if } f \geq f^*. \end{cases} \tag{8}$$

In this likelihood function, $l(\theta)$ and $g(\theta)$ are non-parametric Parzen density estimators. To estimate $l(\theta)$ and $g(\theta)$, the data set needs to be split into two subsets $\mathcal{D}_<$ and $\mathcal{D}_>$, where $\mathcal{D}_<$ contains the data with loss functions less than a threshold f^* (the lower 25% quantile of the loss function) while $\mathcal{D}_>$ contains the rest of the data. In this case, $l(\theta)$ and $g(\theta)$ can be evaluated as $l(\theta) = \frac{1}{n_<}\sum_{\theta_i \in \mathcal{D}_<} K(\theta, \theta_i)$ and $g(\theta) = \frac{1}{n_>}\sum_{\theta_i \in \mathcal{D}_>} K(\theta, \theta_i)$, where $n_<$, $n_>$ are the size of $\mathcal{D}_<$ and $\mathcal{D}_>$ respectively, and $K(\theta, \theta_i)$ is a kernel function (e.g. Gaussian kernel). As such, the marginal distribution of the parameter set given the observed data set $\mathcal{D}_{1:t}$ (the denominator of equation (7)) can in turn be calculated as

$$\begin{aligned} p(\theta;\mathcal{D}_{1:t}) &= \int_{-\infty}^{\infty} p(\theta|f;\mathcal{D}_{1:t})p(f;\mathcal{D}_{1:t})df \\ &= (l(\theta) - g(\theta)) \int_{-\infty}^{f^*} p(f;\mathcal{D}_{1:t})df + g(\theta). \end{aligned} \tag{9}$$

The criteria for exploring the overall parameter space is to choose the next set of simulation parameters that maximizes the expected improvement $\mathbb{E}[\max(f^* - f, 0)]$ [31] as

$$\begin{aligned} \theta_{t+1} &= \underset{\theta}{\text{argmax}} \int_{-\infty}^{\infty} \max(f^* - f, 0)p(f|\theta;\mathcal{D}_{1:t})df \\ &= \underset{\theta}{\text{argmax}} \frac{\int_{-\infty}^{f^*}(f^* - f)p(f;\mathcal{D}_{1:t})df}{\frac{g(\theta)}{l(\theta)}(1 - \int_{-\infty}^{f^*} p(f;\mathcal{D}_{1:t})df) + \int_{-\infty}^{f^*} p(f;\mathcal{D}_{1:t})df} \\ &= \underset{\theta}{\text{argmax}} \frac{l(\theta)}{g(\theta)}, \end{aligned} \tag{10}$$

where the last equation holds since the cumulative distribution $\int_{-\infty}^{f^*} p(f;\mathcal{D}_{1:t})df$ is strictly less than 1, and this result is not affected by the exact form of the prior $p(f;\mathcal{D}_{1:t})$. The next parameter set whose loss function will be simulated is chosen to maximize the quotient of the Parzen density estimators $l(\theta)/g(\theta)$. As each new simulation is conducted, the data set $\mathcal{D}$ is updated with the new simulated data points. The algorithm is generalized as the following.

Since secondary responses (as non-linear responses) are very sensitive to non-linear terms, i.e. αx^2 and βx^3, while the primary responses are more sensitive to the linear terms, it is necessary to minimize the loss functions for both primary and secondary responses simultaneously. One simple way of achieving this is to minimize a weighted sum of these two loss functions rather than minimizing them individually (e.g. $L = L_p + 0.05L_s$).

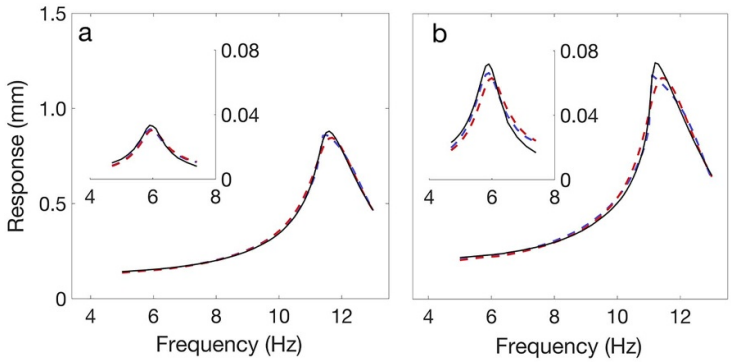

Figure 2. The primary Bayesian optimized (based on equation (1)) and experimentally measured response curves plotted in dashed red and solid black, respectively, under (a) 1.0 V excitation and (b) 1.5 V excitation. The corresponding secondary (super-harmonic) response curves are shown in the subplots with the same color distribution. Dashed blue curves are the Bayesian optimized response curves based on equation (14).

Table 1. The parameter space measured for Model 1 (equation (1)) from the Bayesian optimization method and the multiple-scale perturbation method, and for Model 2 (equation (14)) from the Bayesian optimization method. For the Bayesian optimization method, the measurements are averages of five independent experimental trials, with the corresponding CV shown in parentheses.

Methods	μ (s^{-1})	ω (Hz)	$\lvert\alpha\rvert$ ($\mu m^{-1} \times s^{-2}$)	β ($\mu m^{-2} \times s^{-2}$)	F ($\mu m^{-1} \times s^{-2}$)	$\lvert\gamma\rvert$ ($\mu m^{-3} \times s^{-2}$)
Excitation 1.0 V						
Model 1	10.1 (0.6%)	11.9 (0.1%)	2.1 (0.9%)	3.9×10^{-4} (5.7%)	6.3×10^{5} (0.4%)	
Multiple-scale	9.6	11.9	2.0	3.5×10^{-4}	6.2×10^{5}	
Model 2	9.7 (1.5%)	11.8 (0.3%)	2.0 (1.2%)	14.2×10^{-4} (7.3%)	6.2×10^{5} (0.4%)	1.8×10^{-6} (6.6%)
Excitation 1.5 V						
Model 1	10.9 (0.1%)	11.9 (0.0%)	2.1 (0.8%)	3.8×10^{-4} (4.0%)	9.4×10^{5} (0.2%)	
Multiple-scale	10.2	11.9	1.9	2.1×10^{-4}	9.1×10^{5}	
Model 2	10.3 (1.3%)	11.7 (0.2%)	2.0 (1.2%)	15.5×10^{-4} (3.5%)	9.0×10^{5} (0.6%)	1.2×10^{-6} (5.6%)

3. Results

Figure 2 shows the Bayesian-optimized simulated primary response curves (dashed red curves) of a single dust particle levitated in the plasma sheath in the GEC RF reference cell at a plasma power of 1.68 W and pressure of 40 mTorr. The corresponding secondary response curves are shown in the subplots. Particles excited under excitation amplitudes of 1.0 and 1.5 V are plotted in figures 2(a) and (b), respectively. As shown, the optimized response curves (dashed red curves simulated according to equation (1)) resemble the experimentally measured responses curves (solid black curves) in both the primary and secondary regions. Also, note that the spring softening phenomenon (i.e. the non-linear phenomenon that results in the primary resonance peak being 'bent' in the low frequency direction) becomes more obvious as the excitation amplitude increases (figure 2(a)).

The corresponding optimized parameters obtained from model 1 (equation (1)) are calculated as the average over five independent trials of the optimizing experiment and their values are listed in table 1 (method 'Model 1'), with the corresponding coefficient of variations (CV) shown in parentheses. Notice that the sign of the coefficient of the quadratic non-linearity α is irrelevant in this response analysis since it only changes the direction of the asymmetric motion of the dust particle. Even though there is a large variation in randomness in this Bayesian search of the parameter space, the optimizing experiment converges to yield consistent results as evidenced by the low CV. The relatively high CV for the parameter β (the coefficient of the cubic non-linearity) is due to the fact that the response curves are robust in responding to variation of nonlinearities of higher order. As such, a small variation in β will not significantly perturb the entire response curve.

Figure 3 shows the loss (equation (6)) as a function of the number of iterations for the dust particle excited at both 1.0 V (a) and 1.5 V (b), each of which has five independent experimental trials. As shown, the loss values decrease rapidly after the first several iterations, reaching convergence after a few hundred iterations. (This again indicates the efficiency of the presented Bayesian optimization method in exploitation of the parameter space.) However, in order to boost overall accuracy and ensure wider exploration of the parameter space, we conducted a large number of iterations. The observed higher convergency loss value for the dust particle under 1.5 V excitation ($\approx 4.3 \times 10^{-3}$) as compared to that under 1.0 V excitation

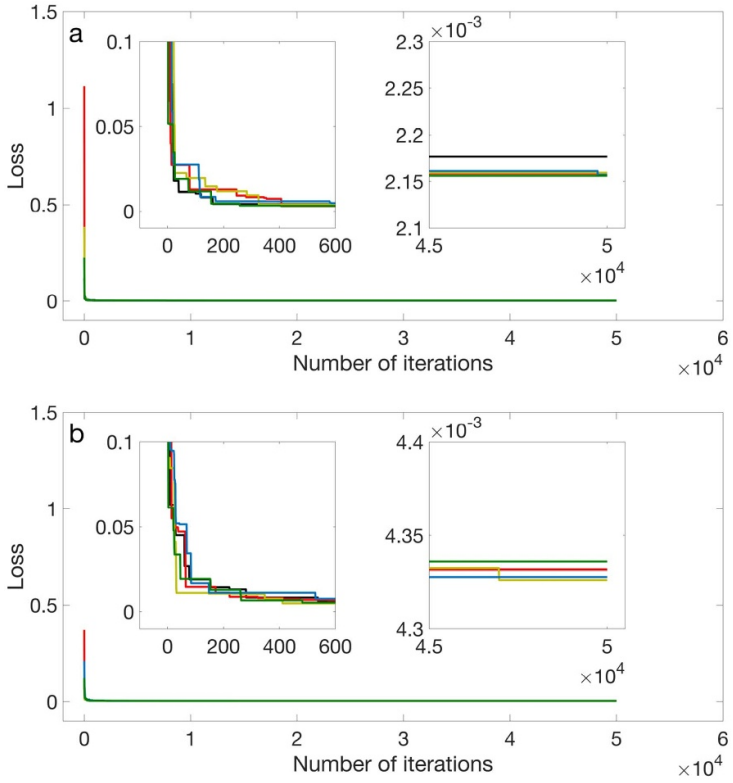

Figure 3. The loss (the 'distance' between the experimentally measured and the simulated response curve) as a function of the number of iterations for a dust particle excited under an excitation amplitude of (a) 1.0 V and (b) 1.5 V. Colors denote the five independent trials.

($\approx 2.2 \times 10^{-3}$) can be attributed to the increased difficulty of capturing the spring softening phenomenon as the excitation amplitude becomes larger (compare figures 2(a) and (b)).

4. Multiple-scale perturbation method

The parameters can also be derived analytically by solving the equation of motion (equation (1)) employing the multiple-scale perturbation method. The details of this method are given in [23, 32], with the main results needed for the analysis given below.

Assuming an external excitation at a frequency of approximately half that of the oscillator resonance frequency ω, i.e. $\Omega \approx \frac{1}{2}\omega$, the solution to equation (1), to first order, yields

$$x(t) = \frac{F}{\omega^2 - \Omega^2}\cos(\Omega t) - \frac{\alpha F^2}{4\omega(\omega^2 - \Omega^2)^2\sqrt{\frac{\mu^2}{4} + (2\Omega - \omega)^2}}\sin(2\Omega t - \phi), \tag{11}$$

where ϕ is the shifted phase which is dependent on the excitation frequency as $\phi = \arctan(\frac{4\Omega - 2\omega}{\mu})$. The parameters ω, μ and α are determined using the experimentally measured secondary response curve fitted to the steady state theoretical secondary response $\alpha F^2/4\omega(\omega^2 - \Omega^2)^2\sqrt{\frac{\mu^2}{4} + (2\Omega - \omega)^2}$. Considering an external excitation having a frequency approximately equal to the oscillator resonant frequency, i.e. $\Omega \approx \omega$, the solution to equation (1), to first order of approximation yields

$$x(t) = A(\Omega)\cos(\Omega t - \phi'), \tag{12}$$

where the shifted phase is $\phi' = \Omega - \omega - \beta$. By eliminating the secular term appearing in the equation of motion to second order of approximation, the steady state theoretical primary response $A(\Omega)$ can now be derived as

$$\frac{F^2}{4\omega^2} = \left(\frac{A\mu}{2}\right)^2 + \left[\left(\frac{9\beta\omega^2 - 10\alpha^2}{24\omega^3}\right)A^3 - (\Omega - \omega)A\right]^2. \tag{13}$$

By fitting the corresponding experimentally measured primary response curve to equation (13), the parameter space for the cubic nonlinearities β and F can now be investigated. The parameters obtained in this way are shown in table 1 (method 'Multiple-scale'). As shown, the parameters measured from the Bayesian optimization are consistent with those measured from the multiple-scale perturbation, except for the value of β under 1.5 V excitation (with approximately 57.6% difference). Notice that the measurements of the parameters from the multiple-scale perturbation serve as a benchmark and should not be considered as true values since they are derived and are precise only to the first order of approximation.

Due to limitations of the multiple-scale perturbation method, any extension of the model (equation (1)) (e.g. including higher order nonlinearities of either displacement x or velocity $\dot{x}$) and derivation of the corresponding approximate solutions would be tediously complicated. However, the Bayesian optimization scheme described here allows this process to be simplified greatly. As an example, the model is extended to include an additional non-linearity of higher order in displacement x:

$$\ddot{x} + \mu\dot{x} + \omega^2 x + \alpha x^2 + \beta x^3 + \gamma x^4 = F\exp(i\Omega t) + \text{c.c.} \tag{14}$$

Applying the Bayesian method, the optimized parameters are measured, with the results shown in table 1 (with method 'Model 2') and the corresponding simulated response curves shown in figure 2 (dashed blue curves). Again, the corresponding CV for the five independent trials are shown in parentheses.

5. Discussion and conclusion

By considering nonlinearities to fourth order, the primary response curves (dashed blue line) in figure 2(a) more closely resemble the spring softening behavior than do those considering nonlinearities to third order (dashed red line). Also, the loss is further reduced, reaching 1.6×10^{-3} for a 1.0-V excitation and 2.3×10^{-3} for a 1.5 V excitation, as compared to values based on the model with third order nonlinearities (equation (1)) which results in loss values of 2.2×10^{-3} and 4.3×10^{-3}, respectively. This indicates a closer match of the simulated response curves to the experimentally measured ones. After introducing nonlinearities to the fourth order, the measured drag coefficient μ, excitation amplitude F, and the coefficient of the quadratic non-linearity α more closely approach the values measured from the multiple-scale perturbation. However, it is noted that the coefficient for the cubic nonlinearities, β exhibits a large deviation. Considering the conditions for the existence of the spring softening effect (seen from equation (13)):

$$9\beta\omega^2 - 10\alpha^2 < 0, \tag{15}$$

the critical value of β for the existence of the spring softening phenomenon can be derived as $\beta < \beta_c \approx 8.1 \times 10^{-4}$ (by taking into consideration the fact that the measured values for the coefficient of the quadratic non-linearity α are consistent in both models given by equations (1) and (14)). In this case, a large value of β as measured for both a 1.0 V excitation and a 1.5 V excitation, based on the model represented by equation (14), seems to violate the condition of the existence of the spring softening phenomenon (equation (15)). However, equation (15) is derived from only the first order of approximation in the multiple-scale perturbation. Thus, although a response curve simulated with a parameter β violating the condition given by equation (15) still reveals the spring softening phenomenon, this indicates a limited ability and accuracy of the multiple-scale perturbation method to explain non-linear responses since it ignores higher order nonlinearities. In order to accurately determine the coefficient of the cubic nonlinearities β (an important factor characterizing the non-linearity of the plasma sheath [20, 21]), the effects from higher order nonlinearities are important and should not be ignored. Figure 4(a) shows the effective restoring potential energy Φ of the particle (divided by the particle mass) in the vicinity of its equilibrium position for both Model 1 (red) and Model 2 (blue). The difference in restoring potential energy is observable for these two models.

With the coefficient for the fourth order non-linearity available, the change in the electric field and the grain charge at varying levitation positions can be further investigated. By considering an expansion in the electric field E and grain charge Q:

$$\begin{aligned} E &= E_0 + E_1 x + E_2 x^2 + E_3 x^3, \\ Q &= Q_0 + Q_1 x + Q_2 x^2 + Q_3 x^3, \end{aligned} \tag{16}$$

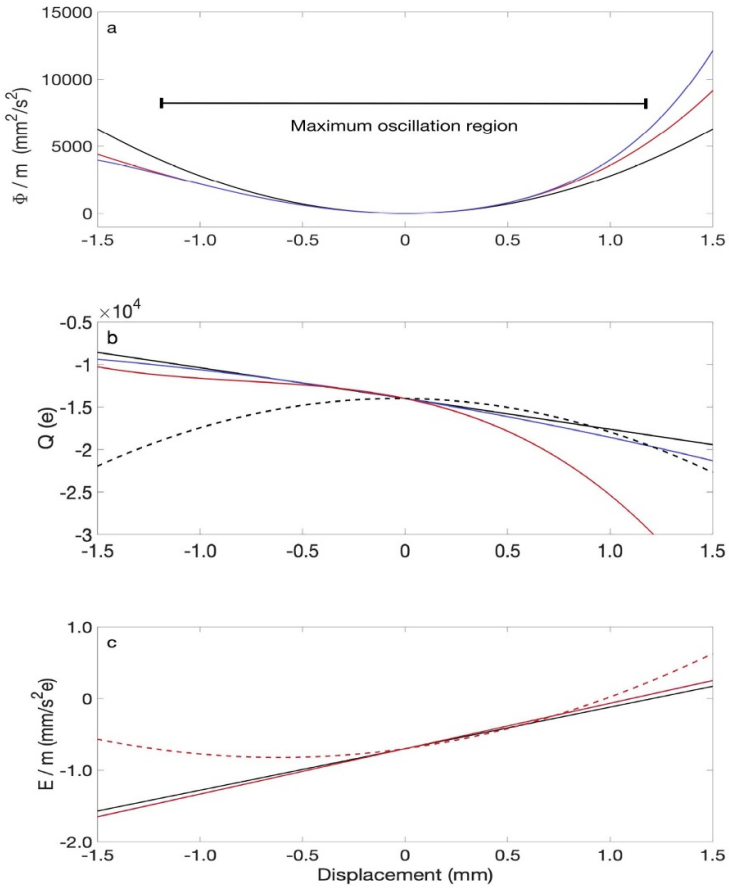

Figure 4. (a) Restoring potential energy around the equilibrium levitation position (divided by the mass of the particle) for Model 1 (red) and Model 2 (blue). The perfect parabolic approximation is shown by the black line as a benchmark. The region over which a particle oscillates in this experiment (with maximum amplitude of 1.2 mm) is also indicated. Grain charge (b) and the electric field (c) in the vicinity of the equilibrium position. Red, blue and black solid lines correspond to the third order, second order and linear polynomial expansion of the grain charge with a linear electric field, respectively. The dashed lines correspond to the second order polynomial expansion for both grain charge and the electric field.

the electrostatic force can be written as

$$\begin{aligned} F_{stat} = &(E_0Q_0) + (E_0Q_1 + E_1Q_0)x \\ &+ (E_0Q_2 + E_1Q_1 + E_2Q_0)x^2 \\ &+ (E_0Q_3 + E_1Q2 + E_2Q_1 + E_3Q_0)x^3 \\ &+ (E_1Q_3 + E_2Q_2 + E_3Q_1)x^4. \end{aligned} \tag{17}$$

The coefficients of the polynomial in this expansion are related to the corresponding coefficients in equation (14). By assuming a linear electric field (i.e. $E_2 = E_3 = 0$), the nonlinearities in charge Q can be explored to the third order (Q_3) with γ provided by the Bayesian optimization approach (Model 2), while without γ, the nonlinearities in charge can only be explored up to the second order (Q_2). Figure 4(b) shows the grain charge in the vicinity of the equilibrium position for the third order polynomial ($Q = Q_0 + Q_1x + Q_2x^2 + Q_3x^3$) and second order polynomial ($Q = Q_0 + Q_1x + Q_2x^2$) expansion in red and blue, respectively. As a reference, a linear charge model ($Q = Q_0 + Q_1x$) is also shown in black. The corresponding linear electric field (divided by the particle mass) for each charge expansion are shown in figure 4(c). The equilibrium charge $Q_0 \approx 1.4 \times 10^4 e$ was estimated by using the levitation position comparison method [33]. In this method, a vertically aligned two-particle pair is formed, and the difference in the levitation position for the upstream particle with and without the presence of the downstream particle (where the downstream particle is knocked out of the system using a laser pulse), is measured. As shown, the third order polynomial charge model predicts a weaker charge reduction in the downstream region, but a stronger charge accumulation in the region above the equilibrium position.

Beyond assuming a linear electric field, we can also investigate the non-linear expansions for the E-field and the grain charge simultaneously (i.e. $E = E_0 + E_1x + E_2x^2$ and $Q = Q_0 + Q_1x + Q_2x^2$). However, due to a lack of constraints, this investigation can only be explored to the second order for both electric field and

grain charge even with γ known. Figure 4(b) and (c) show the result of this charge model and the corresponding non-linear electric field as dashed lines. As shown, both the grain charge and the electric field can be very different from the cases where the electric field is assumed to be linear in the downstream region. This indicates a reasonable assumption of a linear electric field in the close vicinity of the equilibrium position. It is instructive to compare these results against the usual linear models for both the particle charge and electric field. With the assumption of a linear electric field, the charge varies considerably from the linear charge model in the upstream direction. Conversely, with the assumption of a quadratic electric field, it is seen that the charge varies significantly from the linear charge model in the downstream direction. Future experiments may be designed to determine which of these models is correct.

In conclusion, a non-linear response analysis for dust particles in plasma was provided employing a machine learning based method. An efficient technique for optimizing the comparison between numerically simulated and experimentally measured response curves by searching the parameter space in a Bayesian manner was described. Using this approach, the physical parameters characterizing the plasma conditions can be derived. The non-linearity of the response was determined to the fourth order, which is necessary in order to accurately determine the coefficients for lower-order nonlinearities, as well as to correctly characterize the potential energy of the particle in the sheath. Beyond the field of dusty plasmas, the proposed framework provides a general method for measuring physical quantities by optimizing simulation parameters to match experimental observations in an efficient manner, especially when the simulation is computationally expensive.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

Support from NASA Grant No. 1571701, NSF Grant No. 1740203 and NSF Grant No. 1707215 is gratefully acknowledged.

ORCID iDs

Zhiyue Ding https://orcid.org/0000-0003-1554-5911
Lorin S Matthews https://orcid.org/0000-0003-1781-2868
Truell W Hyde https://orcid.org/0000-0002-8603-2737

References

[1] Lecun Y, Bottou L, Bengio Y and Haffner P 1998 *Proc. IEEE* **86** 2278–324
[2] Krizhevsky A, Sutskever I and Hinton G E 2017 *Commun. ACM* **60** 84?90
[3] Hochreiter S and Schmidhuber J 1997 *Neural Comput.* **9** 1735–80
[4] Silver D *et al* 2016 *Nature* **529** 484–9
[5] Silver D *et al* 2017 *Nature* **550** 354–9
[6] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets *Advances in Neural Information Processing Systems 27 2014* 2672–80
[7] Rupp M, Tkatchenko A, Müller K R and von Lilienfeld O A 2012 *Phys. Rev. Lett.* **108** 058301
[8] Carleo G and Troyer M 2017 *Science* **355** 602–6
[9] Dietz C, Kretz T and Thoma M H 2017 *Phys. Rev.* E **96** 011301
[10] Melzer A, Schweigert V A, Schweigert I V, Homann A, Peters S and Piel A 1996 *Phys. Rev.* E **54** R46
[11] Fortov V E, Ivlev A V, Khrapak S A, Khrapak A G and Morfill G E 2005 *Phys. Rep.* **421** 1–103
[12] Morfill G E and Ivlev A V 2009 *Rev. Mod. Phys.* **81** 1353–404
[13] Goree J 1994 *Plasma Sources Sci. Technol.* **3** 400–6
[14] Carstensen J, Jung H, Greiner F and Piel A 2011 *Phys. Plasmas* **18** 033701
[15] Ding Z, Qiao K, Ernst N, Kong J, Chen M, Matthews L S and Hyde T W 2019 *New J. Phys.* **21** 103051
[16] Peng Z, Lang Z and Billings S 2007 *J. Sound Vib.* **301** 777–88
[17] Cottone F, Vocca H and Gammaitoni L 2009 *Phys. Rev. Lett.* **102** 080601
[18] Samanta C, Yasasvi Gangavarapu P R and Naik A K 2015 *Appl. Phys. Lett.* **107** 173110
[19] Močkus J 1975 On Bayesian methods for seeking the extremum *Optimization Techniques IFIP Technical Conference Novosibirsk 1974* 400–4
[20] Ivlev A V, Sütterlin R, Steinberg V, Zuzic M and Morfill G 2000 *Phys. Rev. Lett.* **85** 4060–3
[21] Zafiu C, Melzer A and Piel A 2001 *Phys. Rev.* E **63** 066403
[22] Wang Y N, Hou L J and Wang X 2002 *Phys. Rev. Lett.* **89** 155001
[23] Ding Z, Qiao K, Kong J, Matthews L S and Hyde T W 2019 *Plasma Phys. Control. Fusion* **61** 055004
[24] Epstein P S 1924 *Phys. Rev.* **23** 710–33

[25] Chen M, Dropmann M, Zhang B, Matthews L S and Hyde T W 2016 *Phys. Rev.* E **94** 033201
[26] Liu B, Goree J, Nosenko V and Boufendi L 2003 *Phys. Plasmas* **10** 9–20
[27] Tomme E B, Law D A, Annaratone B M and Allen J E 2000 *Phys. Rev. Lett.* **85** 2518–21
[28] Carstensen J, Greiner F, Block D, Schablinski J, Miloch W J and Piel A 2012 *Phys. Plasmas* **19** 033702
[29] Melzer A 2003 *Phys. Rev.* E **67** 016411
[30] Bergstra J, Bardenet R, Bengio Y and Kégl B 2011 Algorithms for hyper-parameter optimization *Proc. 24th Int. Conf. Neural Information Processing Systems* NIPS-11, pp 2546–54
[31] Jones D R, Schonlau M and Welch W J 1998 *J. Glob. Optim.* **13** 455–92
[32] Nayfeh A H and Mook D T 1979 *Nonlinear Oscillations* (New York: Wiley)
[33] Ding Z, Qiao K, Matthews L and Hyde T 2016 A quick method to determine the charge on dust particles in a complex plasma *APS Division of Plasma Physics Meeting Abstracts* vol 2016 p CO7.008