ADVANCED THEORY AND SIMULATIONS
www.advtheorysimul.com

Probing the Stochastic Dynamics of Coronaviruses: Machine Learning Assisted Deep Computational Insights with Exploitable Dimensions

T. Mukhopadhyay, S. Naskar, K. K. Gupta, R. Kumar, S. Dey, and S. Adhikari**

A machine learning assisted efficient, yet comprehensive characterization of the dynamics of coronaviruses, in conjunction with finite element (FE) approach, is presented. Without affecting the accuracy of prediction in low-frequency vibration analysis, an equivalent model for the FE analysis is proposed, based on which the natural frequencies corresponding to first three non-rigid modes are analyzed. To quantify the inherent system-uncertainty efficiently, Monte Carlo simulation is proposed in conjunction with the machine learning based FE computational framework for obtaining complete probabilistic descriptions considering both individual and compound effect of stochasticity. A variance based sensitivity analysis is carried out to enumerate the relative significance of different material parameters corresponding to various constituting parts of the coronavirus structure. Using the modal characteristics like natural frequencies and mode shapes of the virus structure including their stochastic bounds, it is possible to readily identify coronaviruses by comparing the experimentally measured dynamic responses in terms of the peaks of frequency response function. Results from this first of its kind study on coronaviruses along with the proposed generic machine learning based approach will accelerate the detection of viruses and create efficient pathways toward future inventions leading to cure and containment in the field of virology.

1. Introduction

It is well known that a novel form of coronavirus (SARS-CoV-2) is responsible for the COVID-19 disease in a wide population across the globe in recent times.[1–3] Coronaviruses are a large family of viruses, most of which circulate among animals like pigs, camels, bats, and cats. Sometimes these viruses can jump to humans, which is referred as a spillover event. Though majority of the known coronaviruses cause only mild to moderate respiratory illnesses in humans, three of them (SARS coronavirus, MERS coronavirus, and the more recent novel coronavirus i.e. SARS-CoV-2) have been known to cause more serious, even fatal diseases. In this paper we propose a machine learning assisted generic approach for characterizing the probabilistic dynamics of coronaviruses leading to deep computational insights with exploitable dimensions of detection and containment.

The treatment of COVID-19 through direct medical interventions is still in its infancy.[4] As a result, several authors have proposed non-pharmaceutical interventions[5] as an indirect means of controlling and containing COVID-19. Here we provide a concise and representative review of such research investigations. To et al., conducted the viral culture for self-collected saliva of 12 patients and found 91.7% samples to be novel corona virus infected.[6] They also reported that the investigation of saliva is a promising approach of diagnosis, monitoring, and infection control in patients. Shen et al. in their review article reported various methods to detect the nucleic acid of the coronavirus based on polymerase chain reaction, isothermal nucleic acid amplification, and microarray based approaches.[7] Corman et al. made

T. Mukhopadhyay
Department of Aerospace Engineering
Indian Institute of Technology Kanpur
Kanpur, India
E-mail: tanmoy@iitk.ac.in

S. Naskar
Department of Aerospace Engineering
Indian Institute of Technology Bombay
Mumbai, India

K. K. Gupta, R. Kumar, S. Dey
Department of Mechanical Engineering
National Institute of Technology Silchar
Silchar, India

S. Adhikari
College of Engineering
Swansea University
Swansea, United Kingdom
E-mail: s.adhikari@swansea.ac.uk

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adts.202000291

DOI: 10.1002/adts.202000291

use of synthetic nucleic acid technology to develop a validated diagnostic workflow for novel coronavirus.[8] Sivasankarapillai et al. have presented a review on the use of nanomaterials and medicine concerning the prospective strategies for detection and cure of coronaviruses.[9] Besides these conventional testing procedures for the novel coronavirus, few groups have come up with machine learning and artificial intelligence (AI) based detection of coronavirus. Narin et al. used chest X-ray radiographs to model the convolutional neural network for prediction of novel coronavirus.[10] Gozes et al. developed AI assisted CT image analysis based on classification to predict the presence of coronavirus in the patients.[11] In this context, it may be noted that machine learning and AI based approaches have attracted a wide range of attention from the scientific community recently across different fields to explore physical systems in an extra-ordinarily detailed and insightful way.[12–14] However, machine learning assisted investigations in the filed of virology is rather limited, despite an immense potential in this field.

Most of the indirect approaches to identify and detect coronaviruses involve direct testing of samples. The purpose of this paper is proposing a novel methodology to obtain resonance frequencies of coronavirus by numerical methods in view of identifying it from its modal signature. Using vibrational approaches to characterize molecules is an established technique in nanoscience.[15] Traditionally molecular dynamics methods[16] have been used to compute resonance frequencies of complex molecules. It is well-known that such methods can be computationally expensive for large and complex systems like a coronavirus. Over the past two decades, simplified methods based on the theory of elasticity and structural dynamics have been developed for understanding the vibration of a wide range of molecules. Such approaches lead to achieving computational efficiency up to certain extent, suitable for carrying out a purely deterministic analysis involving only a few simulations. Examples include vibration of single and multilayer carbon nanotubes,[17] graphene sheets along with other 2D materials,[18–23] family of fullerene molecules,[24] to mention a few. Analysis of nano-scale biological systems including different viruses have also seen the application of elastic continuum and structural mechanics based approaches including finite element method (FEM). Mubeen et al. used FEM based numerical analysis to explore the cell dynamics and reported that the natural frequencies of the cell vary inversely with the size of the cell.[25] Molavi et al. presented FEM and scaled-up experimentation based modal analysis of spherical baker's yeast cells to determine the natural frequencies.[26] Varga et al. utilized atomic force microscopy to determine the elastic properties of endothelial and epithelial cells.[27] They also calculated the correlation between nucleus and periphery of the signals which resulted in a higher correlation factor for the endothelial cells. Ford[28] discussed two different analytical models based on liquid drop and elastic sphere to estimate the vibrational frequencies of virus particles, wherein it was concluded that the frequency appears in the range of a few GHz for the particles with a radius of around 50 nm. Zhang and Ru stated that viral capsids are the example of biopolymer spherical shells.[29] They studied the effect of high structural heterogeneity on the natural frequencies and vibration modes of biopolymer spherical shells. Carraso et al. reported that the mechanical strength of spherical viruses is a consequence of the interaction between crystallographically visible short DNA patches and the inner capsid wall.[30] Mateu presented a review article to elaborate on mechanics of the structural elements of viruses.[31] Fraldi et al. explored a frequency response-based hypothesis for mechanically targeting and selectively attacking cancer cells.[32] Jaganathan et al. proposed an unconventional mode of cancer treatment where cancer cells are damaged by utilizing their resonant frequencies.[33] A similar approach was utilized by Heckerman et al.[34] to treat the illness caused by pathogens. Yang et al. stated that the virus is used to resonate in the confined-acoustic dipolar mode when interacted with microwave at the same frequency.[35] Based on this understanding they demonstrated that the airborne virus could be inactivated with a reasonable microwave power density.

From the brief discussion above, it is evident that idealized computational methods on vibrational frequencies for both biological and non-biological molecules play a pivotal role in their understanding. Motivated by this, here we focus on the dynamic characteristics of coronaviruses using finite element method. The vibrational frequencies can also be relevant to nuclear magnetic resonance studies of coronaviruses[36,37] for further structural understandings. A key objective of this paper is to identify natural frequencies and mode shapes of the coronavirus structure using an equivalent mechanics-based model. We would mathematically model the virus as an elastic spherical shell structure with attached spikes and soft infill, and thereby solve the resulting equations using the finite element method. In **Figure 1** we depict the conceptual paradigm of the proposed analysis. Knowing the modal characteristics like natural frequencies and mode shapes of the virus structure, it would be possible to readily identify coronaviruses by comparing the experimentally measured dynamic response signals (note that the peaks in a frequency response function indicate the natural frequencies, as depicted in Figure 1E). In this context, it is important to note that the material properties of such viruses are expected to vary from sample to sample within a certain bound, leading to a random variation of the natural frequencies. Since the physics based finite element model is computationally intensive, it is practically impossible to carry out thousands of finite element simulations corresponding to every possible random combination of the system (i.e. the coronavirus structure) parameters for quantifying the natural frequencies in a probabilistic framework. Moreover, it is crucial to quantify the sensitivity of each system parameter to the natural frequencies for accessing their relative degree of importance. Such sensitivity analysis also normally requires multiple simulations ($\approx 10^4$) of the coronavirus structure corresponding to different sets of input parameters. To mitigate this lacuna, here we would develop an efficient machine learning assisted integrated approach coupled with the finite element simulation for achieving computational feasibility in prediction the natural frequencies. Exploitation of machine learning would allow us to carry out a rapid, efficient, yet comprehensive dynamic characterization of coronaviruses, leading to deep computational insights.

2. Results and Discussion

In this section, we present the numerical results for the dynamic behavior of coronaviruses focusing primarily on first three modes

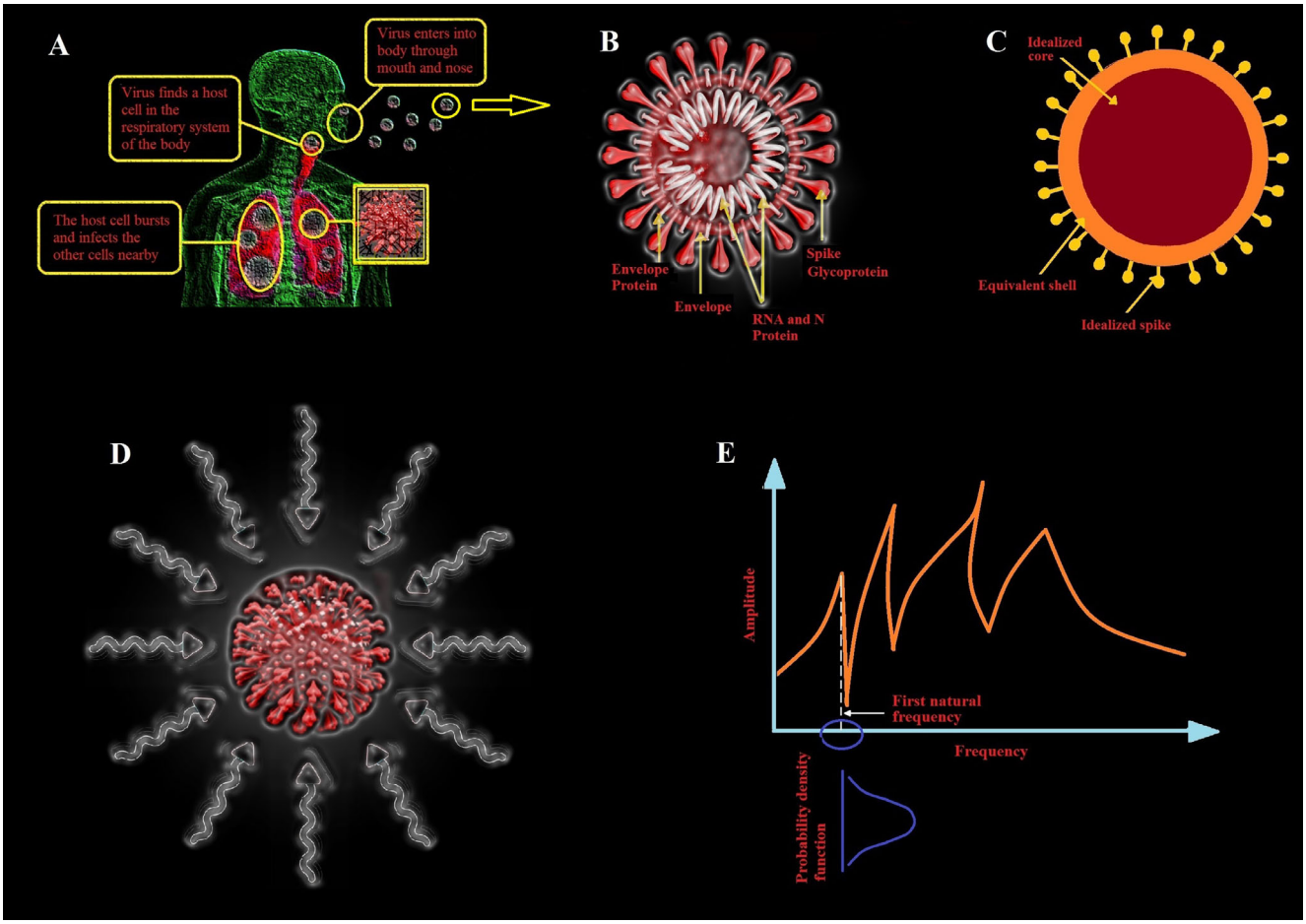

Figure 1. Structural idealization of coronavirus and overview of the analysis. A) A general overview of the stages of infection by coronaviruses. B) Detailed cross-sectional view of coronavirus showing the individual structural components. C) Idealization of coronavirus structure for low-frequency dynamic modeling. D) A typical representation of coronavirus under dynamic excitation. E) Dynamic response of a coronavirus structure in terms of amplitude and frequency, wherein the first peak response indicates the first natural frequency of the system. Noting that the material properties and geometric dimensions of a coronavirus structure may vary randomly within a bound, the probabilistic character of natural frequencies is indicated using a probability density function plot.

of non-rigid vibration. It may be noted here that first few modes of vibration in case of a structure with free boundary condition (such as the structure of coronaviruses considered in this investigation) consist of rigid body motion, which is not of our interest in the present context. We will start by a detailed description about the structure of coronaviruses including the scheme of idealization adopted for numerical modeling in this paper. Thereafter numerical results for vibration analysis would be presented following deterministic and probabilistic frameworks (including sensitivity analysis), assisted by a machine learning based finite element approach.

Figure 1B shows the cross-sectional structure of coronaviruses, the main components of which are different classes of protein and RNA.[38] From a structural mechanics view-point, an equivalent spherical shell-like structure can be idealized having attached spikes (beams with top mass) and a soft core, as depicted in Figure 1C. The effective homogenized material properties of the spherical shell structure could be calculated based on a volumetric homogenization of the properties of constituent proteins. The RNA and N protein inside the spherical shell could be modeled as an elastic soft core with equivalent material properties like Young's moduli and mass density, significantly lesser compared to that of the shell structure. For mechanical analysis like low-frequency dynamic behavior of a structure, it is quite common in literature to adopt such an equivalent continuum based modeling for achieving computational efficiency without compromising the accuracy of results.[39] However, for analyzing the high-frequency vibration modes (not within the scope and objective of this investigation), a more detailed modeling of the structure may be required. Most structural parts of a coronavirus are made of different kinds of protein. We have presented the natural frequencies in this article following a non-dimensional framework, wherein the elastic properties reported in literature (such as ref. [40]) or obtained using experimental characterization could be used to determine the exact numerical values (refer to the Section 4 for further details). It is worthy to note in this context that the main aim of this paper is to propose a machine learning assisted computational framework for characterizing the dynamic

behavior of coronaviruses including the effect of inevitable uncertainty and sensitivity analysis. The proposed framework would be helpful in identifying the viruses in a practically relevant uncertain environment along with quantification of the detection probability. We focus more here on the demonstration of methodological development rather than the exact material properties. For this reason, we have presented the results in a non-dimensional form to have wider impact in terms of their usability. Dimensions of different components of the coronavirus structure are defined relative to the diameter of the spherical shell (50–200 nm) as, thickness of the spherical shell wall $\approx D/30$, beam length of the spikes $\approx D/6 - D/4$, diameter of the cross-section of the beams $\approx D/40$ and diameter of beam top mass we have considered $D = 120$ nm (within the specified range) in this investigation.

We have used the finite element approach to establish a physics based model for vibration analysis of coronaviruses (for detailed description refer to Section SM1, Supporting Information). However, before investigating the dynamic behavior of coronaviruses, it is necessary to gain adequate confidence on the numerical modeling approach by means of carrying out validation studies. Since suitable literature concerning the vibration analysis of coronaviruses is not available, we resort to other spherical bio/nano structures for the purpose of gaining confidence on the adopted finite element modeling. In this context it may be noted that analysis of shell structures has been extensively reported in the literature of structural mechanics.[41–45] For validation, first we have modeled two different spherical breast cells (normal and cancerous) with exactly same geometric dimensions and material properties as the reference article following the current finite element approach and compared the results of computed natural frequencies with the numerical values reported in literature.[33] A good agreement between the results (refer to Table S2, Supporting Information) corroborates the validity of our modeling approach. However, the dimension of a breast cell is 100 times bigger than that of the virus under consideration. Thus, to investigate the validity of the current finite element model further, we consider another spherical system (fullerene C_{60}) with dimension 100 times lesser than that of the virus (since the corresponding results of vibration analysis for fullerene C_{60} are available in the literature). We noticed from the numerical results presented in Table S3, Supporting Information, that a finite element simulation carried out in ANSYS (with the idealization of a spherical structure using brick elements) can produce close results to the fundamental natural frequencies obtained using Raman spectroscopy and molecular mechanics based approaches.[46–49] Thus, in the absence of adequate numerical results for validating the natural frequencies, considering two different systems having diameters 100 times more and 100 times less than the diameter of coronaviruses, we show that a finite element approach can produce accurate results for a spherical bio/nano structure. It is expected that the finite element approach would produce accurate results for a spherical structure where the dimension lies in between these two extreme cases. Previous experiences of the authors in this research area also support the observation that the physics behind the low-frequency vibration of a spherical structure is not affected significantly by the change of scale. The two-fold validation presented here provides us adequate confidence to extend the finite element model further for predicting the dynamic behavior of the present coronavirus structure. The numerical results concerning natural frequencies of coronaviruses can be considered as the first of its kind to be reported in the literature.

Based on the preceding paragraphs, the assumptions in this work can be viewed separately in terms of the equivalent continuum based modeling of the virus structure and the subsequent analysis method. It is noted that a reduced-order modeling approach using equivalent continuum models (with the linear definitions of strain and material properties) is sufficient to capture the physics of low-frequency vibration of a spherical nano/bio structure. We have shown this assumption to be valid in a finite element based analysis for spherical structures considering dimensions in a range of 100 times more and 100 times less than that of the virus. Further, one of the most critical effects that needs to be accounted in nanoscale elastic analysis is the nonlocal effect. However, a well-documented fact may be noted that the non-local effect becomes negligible beyond the dimension of 30 nm.[50] Since the mean diameter of the current virus structure is considered as 120 nm, the nonlocal effect is neglected in the current analysis.

The numerical results for vibration analysis of the coronavirus structure are presented hereafter considering 60 spikes, unless otherwise explicitly mentioned. A finite element method based modal analysis is performed here to determine the mode shapes and natural frequencies of the coronavirus structure for the first three non-rigid modes, as depicted in **Figure 2**A–C. It can be noted here that the first mode shape shows a radial breathing mode, while the two higher modes represent peripheral deformation of the circular shell in two orthogonal directions. The corresponding three non-dimensional natural frequencies are reported as 9.97×10^8, 1.03×10^9, and 1.09×10^9. Here all the numerical results of natural frequencies are presented in a non-dimensional form for generality, following a scheme which is widely adopted in the area of shell structures[51] (refer to the Section 4 for further details). Deterministic variation of non-dimensional natural frequencies of the coronavirus structure is investigated for individual material properties in the range of $[0.98X, 1.02X]$, where X represents the nominal values of a material property (refer to Figure 2D–F). The relative slope of the plots gives an impression of the sensitivity of each material property to the natural frequencies, while their nature (positive or negative) shows whether the natural frequencies increase or decrease with the variation of the material properties. To demonstrate the effect of different number of spikes, the variation of natural frequencies is plotted against the number of spikes in Figure 2G, where a general trend of increasing frequencies can be noticed as the number of spikes increases.

Figure 3 presents the dynamic response of two coronavirus structural configurations (with 20 and 60 spikes) in terms of frequency response function (FRF). The results are shown considering two damping ratios and two different measuring points (as indicated in inset of the figures). Here the frequencies corresponding to peak amplitudes represent the natural frequencies of the system. It is interesting to note that the frequencies corresponding to the structural configurations with 20 and 60 spikes exactly match with the respective natural frequencies shown in Figure 2G, corroborating an indirect validation of the current

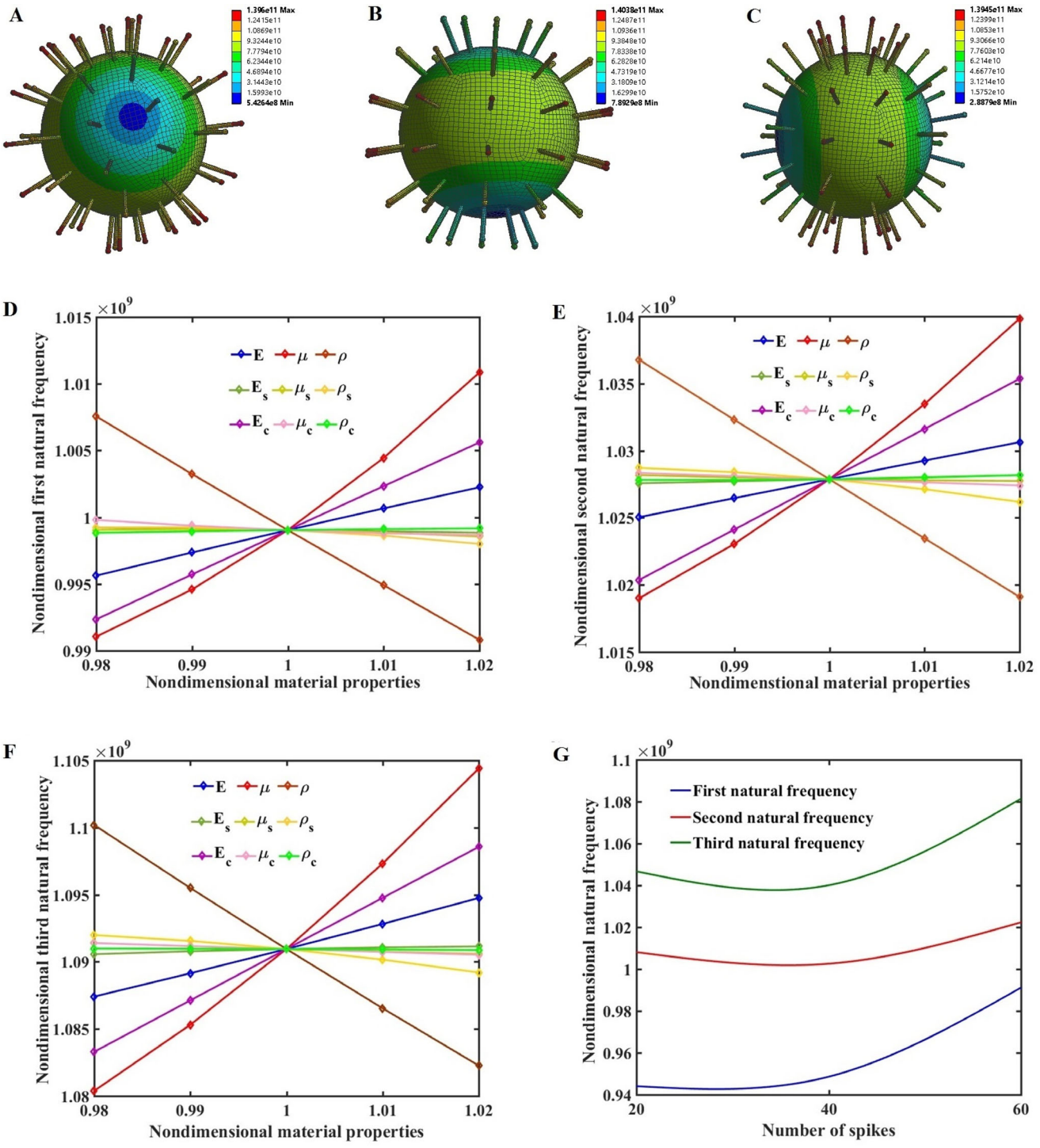

Figure 2. Deterministic characterization of the dynamics of coronavirus structure. A–C) Vibration mode shapes of the coronavirus structure for first three non-rigid modes considering 60 number of spikes in the structure (The corresponding three non-dimensional natural frequencies are reported as 9.97×10^8, 1.03×10^9, and 1.09×10^9). D–F) Variation of non-dimensional natural frequencies of coronavirus structure with individual material properties. The material properties are investigated in the range of [0.98*X*, 1.02*X*], where *X* represents the nominal values of a material property. Here the material properties are non-dimensionalized using their respective nominal values (*X*). G) Variation of deterministic non-dimensional natural frequencies with respect to different number of spikes.

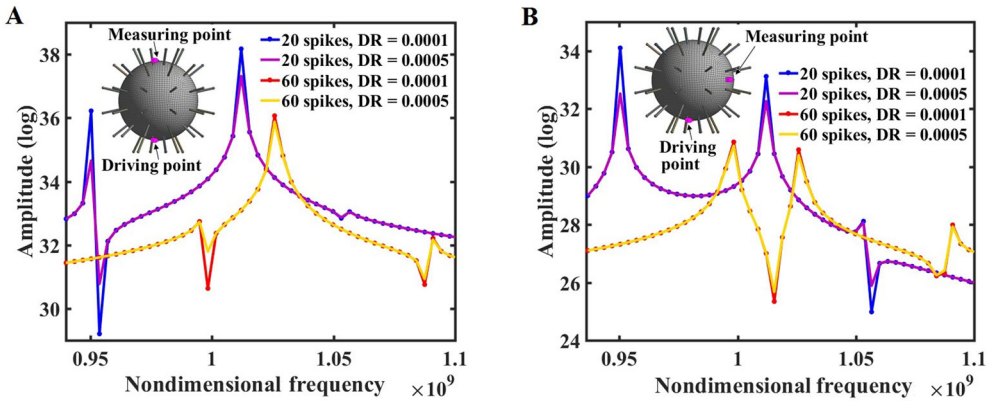

Figure 3. Frequency response function (FRF) of coronaviruses. A) FRF plots for a location of the measuring point when it is considered in a radially opposite direction. B) FRF plots for a location of the measuring point when it is considered in a radially orthogonal direction. Locations of the driving points and measuring points are clearly indicated in the insets of each figure. The FRFs are shown here for two different structural configurations with 20 and 60 spikes, wherein two damping ratios (DR) are considered in each case. Here the amplitude of vibration is plotted using logarithmic scale. It can be noted that the frequencies corresponding to peak amplitudes represent natural frequencies of the system.

study. Figure 3 shows that the amplitude of vibration increases with decreasing damping ratio and the nature of FRF changes with different measuring points. However, the natural frequencies obtained using the peak amplitudes remain constant for any particular structural configuration since it is a function of the stiffness and mass matrix of the system (not dependent on the value of damping ratio, or location of driving and measuring points). Using experimental methods, it is possible to obtain the dynamic response signals (FRF) of a virus, as presented in Figure 3 (, albeit numerically obtained here) and the illustration in Figure 1E. By comparing the frequencies corresponding to peak of the FRFs with the computationally obtained known natural frequencies of the virus, it would be possible to detect coronaviruses readily. It may be noted in this context that under certain circumstances the virus would be suspended in a liquid droplet. The impact of such liquid (with various physical parameters) surrounding the virus structure on the dynamic behavior will be primarily in terms of a damping effect. However, the natural frequencies of the virus, being mainly a function of the stiffness and mass matrices of the structure, is expected to be minimally affected by such damping effects of the surrounding liquids. The observations of Figure 3 further corroborates the above discussion.

So far we have presented physics-based deterministic results of the vibration analysis of coronaviruses. However, it may be noted that the material properties of such viruses are expected to vary from sample to sample within a certain bound, leading to a random variation of the natural frequencies. Since the physics based finite element model is computationally intensive, it is practically impossible to carry out thousands of finite element simulations corresponding to every possible random combination of the system (i.e. the coronavirus structure) parameters for quantifying the natural frequencies in a probabilistic framework. Moreover, it is important to quantify the sensitivity of each system parameter to the natural frequencies for accessing their relative degree of importance while characterizing them mechanically. Such sensitivity analysis also normally requires multiple simulations ($\approx 10^4$) of the coronavirus structure corresponding to different sets of input parameters. We have developed an efficient machine learning assisted integrated approach coupled with the finite element simulation for achieving computational efficiency in prediction of the natural frequencies (refer to Section SM2, Supporting Information). In this approach, first a machine learning model of the computationally intensive coronavirus structure is developed based on a few optimally chosen samples. Once the machine learning model is formed, it can be regarded as a computationally efficient digital substitute of the actual simulation model. Based on the machine learning model it is possible to predict the natural frequencies corresponding to any combination of the system input parameters efficiently, which in turn can be used for several computationally intensive investigations such as probabilistic quantification and sensitivity analysis (refer to Section SM3, Supporting Information).

For stochastic characterization and high dimensional data assisted sensitivity analysis, we have developed a support vector regression based machine learning model in conjunction with finite element simulation for vibration analysis of coronaviruses.[52–54] Before using the machine learning model for any further analysis, its prediction capability is checked extensively using statistical measures (such as R^2, refer to Table S3, Supporting Information) and scatter plots. From the scatter plots presented in **Figure 4**A–C, it can be noticed that a sample size of $N = 32$ (drawn from the quasi-random Sobol sequence with faster convergence rate compared to other sampling methods[55]) provides adequate prediction capability (noting that closer proximity of the data points from the diagonal line indicates a better formation of the machine learning model), the corresponding R^2 of which is also found to be close to 1. We have presented the results of sensitivity analysis and uncertainty quantification of the natural frequencies (first three non-rigid vibration modes) in the following paragraph using the machine learning models formed based on 32 data samples.

To understand the relative importance of the material properties quantitatively on the first three natural frequencies of a coronavirus structure, a sensitivity analysis is carried out as an integral part of this investigation. A variance based sensitivity

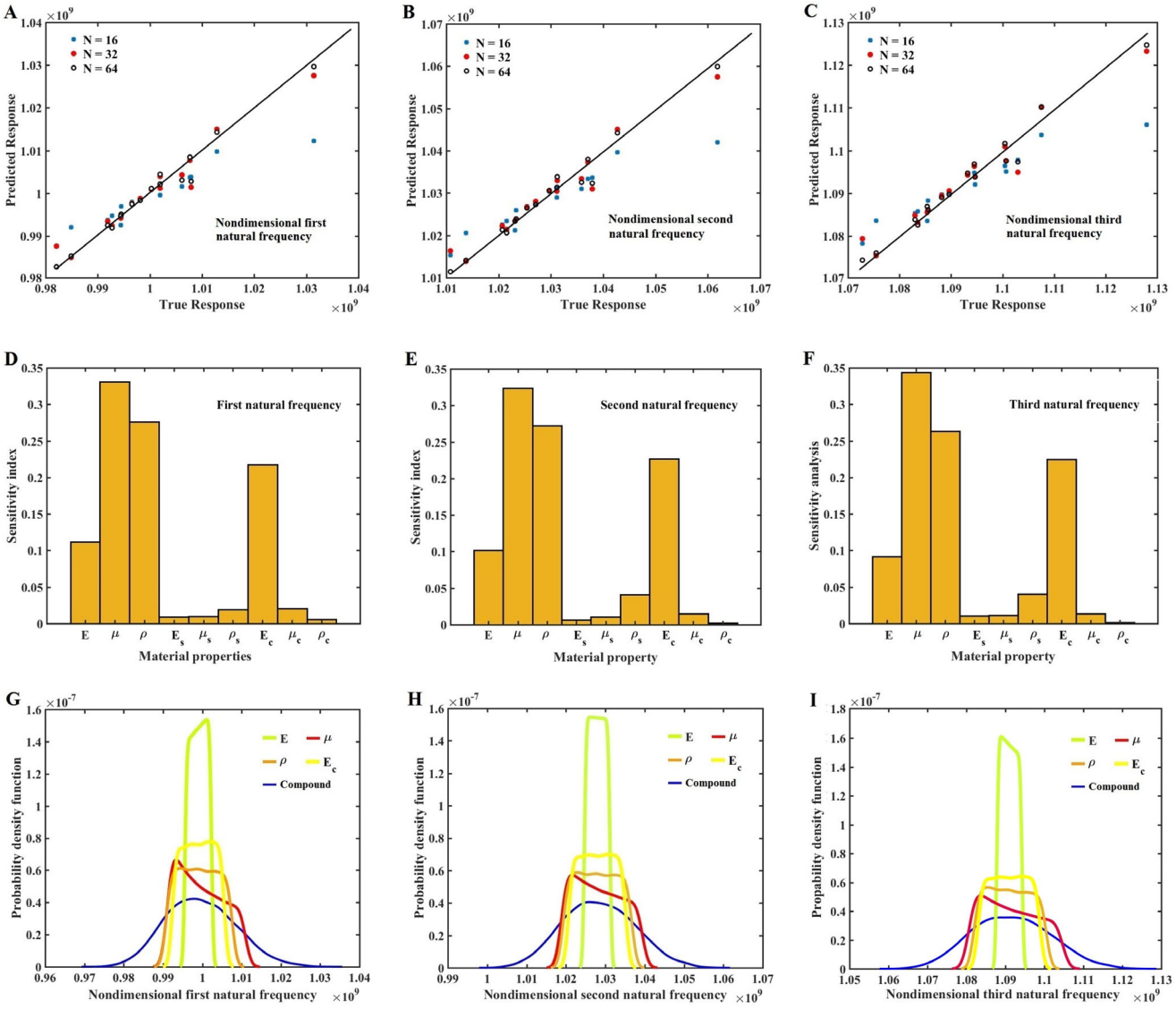

Figure 4. Machine learning assisted rapid dynamic analysis of coronavirus. A–C) Scatter plots for the first three non-dimensionalized natural frequencies considering three different numbers of training samples (N). D–F) Sensitivity analysis for the first three natural frequencies to establish the relative importance of individual material properties. Here E, μ, ρ, E_s, μ_s, ρ_s, E_c, μ_c, ρ_c represent the Young's modulus, Poisson's ratio, and mass density of the spherical shell (denoted without any subscript), spikes (denoted by subscript s) and the equivalent soft core (denoted by subscript c), respectively. G–I) Probability density function plots for first three natural frequencies corresponding to the individual stochastic effect of most sensitive material properties (E, μ, ρ, and E_c) and the compound stochastic variation in all the material properties.

analysis method is used here exploiting the computational advantage of the developed machine learning models.[56,57] It can be observed from the results presented in Figure 4D–F that the material properties of the spherical shell (E, μ, and ρ) and the Young's modulus of the equivalent soft core material (E_c) are the most sensitive parameters compared to the other material properties. It is interesting to note that the observation on relative sensitivity of the material properties are in good agreement with the results presented in Figure 2D–F, considering the relative slope of the curves corresponding to each individual material property. Having quantitative knowledge on the relative sensitivity of each material parameters, we have explored the individual stochastic influence of the most sensitive material properties first, followed by an investigation accounting the compound random variation of all material properties. Figure 4G–I presents Monte Carlo simulation (machine learning assisted) based probabilistic description of the three natural frequencies considering individual and compound effect of stochasticity in the material parameters. The results show that the compound effect of stochasticity leads to the highest stochastic response bounds, followed by the four individual material parameters in the order of their respective sensitivities. It is worthy to note that in a realistic situation, the natural frequencies are expected to have a response bound as shown here, instead of a single deterministic value. Thus the machine learning assisted computationally efficient approach for stochastic characterization of the natural frequencies would lead

to an inclusive paradigm for vibration based rapid detection and prospective containment of the coronaviruses in a practically relevant probabilistic framework. Interestingly, the probabilistic descriptions of the natural frequencies could also inevitably result in a computational framework for quantification of the crucial detection probability of coronaviruses following the proposed approach.

3. Conclusions and Perspective

We have explored the dynamic behavior of coronaviruses following an efficient machine learning assisted framework coupled with finite element approach. A physics-based modeling paradigm of the virus structure concerning low-frequency non-rigid vibration modes is developed using the concept of equivalent homogenized mechanical properties. To gain adequate confidence, the finite element approach is validated with existing literature before using it in the present analysis. It is demonstrated that such a reduced-order model could accurately capture the dynamics corresponding to low-frequency vibration modes of spherical bio/nano structures. Based on finite element model of the equivalent coronavirus structure, an insightful deterministic vibration analysis is performed first considering different number of spikes. It is noticed that the first non-rigid vibration mode corresponds to a radial breathing mode, while the subsequent modes involve peripheral deformations. Using the finite element model, dynamic response of two different coronavirus structural configurations are presented in terms of frequency response function (FRF) considering different sets of driving and measuring points. It is numerically demonstrated that the frequencies corresponding to peak amplitudes represent the natural frequencies of the system. The results reveal that the nature and amplitude of FRF changes with different measuring points and damping ratios, while the natural frequencies obtained using the peak amplitudes remain constant for any particular structural configuration since it is a function of the stiffness and mass matrix of the system. Using experimental methods it is possible to obtain the dynamic responses (FRF) of a virus, and thereby comparing the frequencies corresponding to peak of the FRFs with computationally obtained known natural frequencies of the virus, it would be possible to detect coronaviruses readily.

Normally the effective material properties of coronaviruses are expected to vary from sample to sample randomly, leading to a probabilistic variation of the natural frequencies. To quantify this random variation, Monte Carlo simulation is carried out in conjunction with machine learning based finite element computational framework for obtaining the complete probabilistic description of the natural frequencies. Since the physics based finite element model is computationally intensive, it is practically impossible to carry out thousands of finite element simulations corresponding to every possible random combination of the system (i.e. the coronavirus structure) parameters using a direct finite element approach. The coupled machine learning based finite element model reduces the computational time and cost significantly without compromising the accuracy of results. In the probabilistic analysis, both individual and compound effect of stochasticity in the material properties of the coronavirus structure are considered, wherein the compound variation leads to the highest level of response bound followed by the other material properties in an order of their respective sensitivity. The efficient machine learning based computational model is further exploited to quantify the sensitivity of material properties corresponding to different constituting components of the coronavirus structure. The sensitivity analysis would allow the experimentalists to decide on the degree of quality control needed while determining the physical properties of different components of the virus accurately. It is noticed that the material properties of the spherical shell and the Young's modulus of the equivalent soft core material are the most sensitive parameters to the first three natural frequencies. Realization of such deep computational insights involving large-scale data-driven analyses has only been possible here due to the seamless coupling of finite element modeling and machine learning. In general, the novelty and impact of this article lie in the proposed efficient machine learning assisted equivalent finite element approach for comprehensively characterizing the stochastic dynamics of coronaviruses (including insightful numerical results of the first of its kind study) and their potential rapid detection under a practically-relevant uncertain environment using the dynamic response signals. It is worthy to mention here that, based on the scope of this paper, we have restricted the study to computational simulations. Experimental investigations could be carried out in the future for a more accurate data-driven quantification of uncertainty in the dynamic behavior of coronaviruses.

In summary, we have presented the dynamic behavior of coronaviruses including the practically-relevant effect of stochasticity in the system parameters. Based on an efficient machine learning assisted equivalent finite element framework we develop coupled computational paradigm for obtaining the dynamic response bounds of natural frequencies under a realistic stochastic environment. It is demonstrated that a coronavirus structure could potentially be identified (including quantification of the detection probability) by comparing the stochastic response bound of natural frequencies with experimentally measured dynamic signals. It may be noted that the proposed coupled computational framework is generic in nature and it can be extended to other viruses and critical biological systems for efficient, yet comprehensive characterization, leading to the pathways for their rapid detection along with potential cure and containment.

4. Methods

4.1. Finite Element Modeling of Equivalent Coronavirus Structure

Here, we briefly discuss the finite element modeling approach of the equivalent coronavirus structure (refer to Figure 1C) including the geometrical details. Positioning of the spikes in the corona structure is accomplished based on latitudes and longitudes of the spherical shell structure as depicted in Section SM1 of the Supporting Information. In the finite element analysis code, meshing of the coronavirus structure is carried out by utilizing a 3D hexahedron mesh element with a mesh size of 3.5 nm for the spherical shell and spikes, whereas for the inner soft core, a 2D quadrilateral mesh with mesh size of 1.2 nm is used. Following the standard procedure of finite element approach, the element-level mass and stiffness matrices are assembled to obtain the global mass and stiffness matrices of the entire coronavirus structure. The free vibration analysis is carried

out by solving an eigenvalue problem between the global mass and stiffness matrices, where the natural frequencies and mode-shapes are obtained from the eigenvalues and eigenvectors, respectively. Here, all the numerical results of natural frequencies are presented in a non-dimensional form for generality, following a scheme that is widely adopted in the area of shell structures:[51] $\omega_n = \frac{\omega}{D^2}[\sqrt{\frac{12\rho(1-\mu^2)}{Et^2}}]^{-1}$. The geometric parameters D and t denote the diameter and wall-thickness of the spherical shell, while ρ, μ, and E are the mass density, Poisson's ratio, and Young's modulus of the spherical shell. In order to present the natural frequencies of coronaviruses in non-dimensional terms of the material properties of the spherical shell structure, the material properties of the spikes and soft core are expressed as $\rho_s = \alpha_s\rho$, $\mu_s = \beta_s\mu$, $E_s = \gamma_s E$, $\rho_c = \alpha_c\rho$, $\mu_c = \beta_c\mu$, and $E_c = \gamma_c E$, where the non-dimensional constants α_s, β_s, γ_s, α_c, β_c, and γ_c depend on the mechanical properties of the respective structural components. One can find out the actual values of the natural frequencies readily from the material properties of the coronavirus utilizing the non-dimensionalization scheme described above.

4.2. Machine Learning Based Integrated Numerical Analysis Scheme in Virology

We have developed a machine learning assisted computational analysis scheme coupled with finite element modeling for investigating the computationally intensive probabilistic descriptions and sensitivity of natural frequencies. A support vector regression based machine learning algorithm is adopted here, the brief description of which is provided in Section SM2 of the Supporting Information. The fundamental stages involved in a machine learning based analysis are described in Figure S3 of the Supporting Information. In the first stage, we have used the Sobol sequence for generating the set of input parameters (within a closed bound of ±2% with respect to nominal values of the material properties), the natural frequencies corresponding to which are obtained using the physics based finite element model. Based on optimally designed training samples, the machine learning model is trained following a support vector regression algorithm. This stage also involves multiple checks on the prediction capability of the machine learning model, as described in Section SM2 of the Supporting Information. Once the machine learning model is formed with adequate accuracy of prediction, it can be regarded as a digital computationally efficient substitute of the original expensive finite element simulation model. In the third stage, the machine learning model could be exploited to predict the responses corresponding to any random set of input parameters, paving the way for carrying out computationally efficient probabilistic quantification (using Monte Carlo simulation) and sensitivity analysis (refer to Section SM3 of the Supporting Information). If θ is the considered input parameter then we define the bound for Monte Carlo simulation as $\theta_{\min} = \theta(1-\Delta)$ and $\theta_{\max} = \theta(1+\Delta)$, where $\Delta = 0.02$. The i^{th} perturbed sample of the Monte Carlo simulation is drawn as $\theta_i = \theta_{\min} + (\theta_{\max} - \theta_{\min})R_i$, $i \in 1, 2, 3, \ldots, N_{\text{MCS}}$. Here R_i is a random number generator following a particular probability distribution in the range of 0 to 1. The probability density function plots in Figure 4G–I are obtained on the basis of the natural frequencies evaluated corresponding to $N_{\text{MCS}}(\approx 10^4)$ dataset. The probabilistic descriptions of four most important individual effects (identified on the basis of variance based sensitivity analysis) and the compound effect of stochasticity are presented here (refer to Section SM3, Supporting Information).

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

T.M. and S.N. acknowledge the Initiation grants received from IIT Kanpur and IIT Bombay, respectively. S.A. acknowledges the support of UK-India Education and Research Initiative through grant number UKIERI/P1212.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

finite element modeling of viruses, machine learning assisted analysis in virology, natural frequency of viruses, rapid detection, vibration analysis of coronavirus

Received: November 12, 2020
Published online:

[1] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, *N. Engl. J. Med.* **2020**, *382*, 1199.
[2] R. O. J. H. Stutt, R. Retkute, M. Bradley, C. A. Gilligan, J. Colvin, *Proc. R. Soc. A* **2020**, *476*, 20200376.
[3] S. J. Fong, G. Li, N. Dey, R. G. Crespo, E. Herrera-Viedma, *Appl. Soft Comput.* **2020**, *93*, 106282.
[4] C.-C. Lai, T.-P. Shih, W.-C. Ko, H.-J. Tang, P.-R. Hsueh, *Int. J. Antimicrob. Agents* **2020**, *55*, 105924.
[5] N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. G. T. Walker, C. Waters, P. Winskill, C. Whittaker, C. A. Donnelly, S. Riley, A. C. Ghani, *Report 9: Impact of Nonpharmaceutical Interventions (npis) to Reduce Covid19 Mortality and Healthcare Demand*. **2020**.
[6] K. K. W. To, O. T. Y. Tsang, C. C. Y. Yip, K. H. Chan, T. C. Wu, J. M. C. Chan, D. C. Lung, *Clin. Infect. Dis.* **2020**, *71*, 1841.
[7] M. Shen, Y. Zhou, J. Ye, A. A. A. Al-Maskri, Y. Kang, S. Zeng, S. Cai, *J. Pharm. Anal.* **2020**, *10*, 97.
[8] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, D. G. Mulders, *Eurosurveillance* **2020**, *25*, 2000045.
[9] V. S. Sivasankarapillai, A. M. Pillai, A. Rahdar, A. P. Sobha, S. S. Das, A. C. Mitropoulos, M. H. Mokarrar, G. Z. Kyzas, *Nanomaterials* **2020**, *10*, 852.

[10] A. Narin, C. Kaya, Z. Pamuk, arXiv:2003.10849 **2020**.
[11] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, E. Siegel, *arXiv preprint arXiv* arXiv:2003.05037 **2020**.
[12] E. Samaniego, C. Anitescu, S. Goswami, V. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, *Comput. Methods Appl. Mech. Eng.* **2020**, *362*, 112790.
[13] Vaishali, T. Mukhopadhyay, P. Karsh, B. Basu, S. Dey, *Compos. Struct.* **2020**, *237*, 111870.
[14] A. Mahata, T. Mukhopadhyay, S. Adhikari, *Mater. Res. Express* **2016**, *3*, 036501.
[15] J. M. Hollas, *Modern Spectroscopy*, 3rd ed., John Wiley, UK **1996**.
[16] D. C. Rapaport, *The Art of Molecular Dynamics Simulation*, 2nd ed., Cambridge University Press, UK **2004**.
[17] R. Chowdhury, C. Y. Wang, S. Adhikari, *J. Phys. D: Appl. Phys.* **2010**, *43*, 1.
[18] Y. Chandra, R. Chowdhury, F. Scarpa, S. Adhikari, *Thin Solid Films* **2011**, *519*, 6026.
[19] Y. Chandra, S. Adhikari, E. I. Saavedra Flores, L. Figiel, *Mater. Sci. Eng., R* **2020**, *140*, 100544.
[20] T. Mukhopadhyay, A. Mahata, S. Adhikari, M. A. Zaeem, *2D Mater.* **2017**, *4*, 025006.
[21] T. Mukhopadhyay, A. Mahata, S. Adhikari, M. Asle Zaeem, *Nanoscale* **2018**, *10*, 5280.
[22] Y. Chandra, T. Mukhopadhyay, S. Adhikari, L. Figiel, *Nanotechnology* **2020**, *31*, 145705.
[23] T. Mukhopadhyay, A. Mahata, S. Adhikari, M. Asle Zaeem, *Sci. Rep.* **2017**, *7*, 15818.
[24] S. Adhikari, R. Chowdhury, *Phys. Lett. A* **2011**, *375*, 1276.
[25] A. A. Mubeen, R. Bharathwaj, A. Basavaraju, M. C. Kateel, *Vibroeng. Procedia* **2020**, *27*, 73.
[26] M. Molavi, A. Bonakdar, I. Stiharu, in *ASME Summer Bioengineering Conf.*, ASME, New York **2008**, pp. 493–494.
[27] B. Varga, C. Fazakas, I. Wilhelm, I. A. Krizbai, Z. Szegletes, G. Varo, A. G. Vegh, *Biochem. Biophys. Rep.* **2016**, *7*, 303.
[28] L. H. Ford, *Phys. Rev. E* **2003**, *67*, 051924.
[29] L. Zhang, C. Q. Ru, *AIP Adv.* **2018**, *8*, 075006.
[30] C. Carrasco, M. Castellanos, P. J. de Pablo, M. G. Mateu, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4150.
[31] M. G. Mateu, *Virus Res.* **2012**, *168*, 1.
[32] M. Fraldi, A. Cugno, L. Deseri, K. Dayal, N. M. Pugno, *J. R. Soc., Interface* **2015**, *12*, 20150656.
[33] S. K. Jaganathan, A. P. Subramanian, M. V. Vellayappan, A. Balaji, A. A. John, A. K. Jaganathan, E. Supriyanto, *Curr. Sci.* **2016**, *110*, 1828.
[34] D. E. Heckerman, S. J. Mercer, C. D. Karkanias, E. J. Horvitz, *U.S. 12/262,881*, **2010**.
[35] S. C. Yang, H. C. Lin, M. T. Liu, J. T. Lu, W. T. Hung, Y. R. Huang, C. K. Sun, *Sci. Rep.* **2015**, *5*, 18030.
[36] W. Peti, M. A. Johnson, T. Herrmann, B. W. Neuman, M. J. Buchmeier, M. Nelson, J. Joseph, R. Page, R. C. Stevens, P. Kuhn, K. Wüthrich, *J. Virol.* **2005**, *79*, 12905.
[37] A. Chatterjee, M. A. Johnson, P. Serrano, B. Pedrini, J. S. Joseph, B. W. Neuman, K. Saikatendu, M. J. Buchmeier, P. Kuhn, K. Wüthrich, *J. Virol.* **2009**, *83*, 1823.
[38] I. Seahand, X. Su, G. Lingam, *Eye* **2020**, *34*, 1155.
[39] T. Mukhopadhyay, S. Adhikari, *Int. J. Eng. Sci.* **2017**, *119*, 142.
[40] B. Stephanidis, S. Adichtchev, P. Gouet, A. McPherson, A. Mermet, *Biophys. J.* **2007**, *93*, 1354.
[41] S.-R. Li, X.-H. Fu, R. Batra, *Mech. Res. Commun.* **2010**, *37*, 577.
[42] M. Arefi, T. Rabczuk, *Composites, Part B* **2019**, *168*, 496.
[43] P. H. Shah, R. C. Batra, *AIAA J.* **2019**, *57*, 4942.
[44] A. Kumar, P. Bhargava, A. Chakrabarti, *Thin-Walled Struct.* **2013**, *63*, 82.
[45] S. Faroughi, E. Shafei, T. Rabczuk, *Comput. Methods Appl. Mech. Eng.* **2020**, *359*, 112668.
[46] M. Nasiri Sarvi, M. Ahmadian, *Sci. Iran.* **2012**, *19*, 1316.
[47] A. Sakhaee-Pour, A. Vafai, *Appl. Phys. Lett.* **2010**, *96*, 021903.
[48] A. C. Ferrari, J. Robertson, H. Kuzmany, R. Pfeiffer, M. Hulman, C. Kramberger, *Philos. Trans. R. Soc., A* **2004**, *362*, 2375.
[49] S. Adhikari, R. Chowdhury, *Phys. Lett. A* **2011**, *375*, 2166.
[50] M. Eltaher, M. Khater, S. A. Emam, *Appl. Math. Modell.* **2016**, *40*, 4109.
[51] S. Dey, T. Mukhopadhyay, S. Adhikari, *Composites, Part B* **2015**, *70*, 99.
[52] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, Berlin **2013**.
[53] T. B. Trafalis, H. Ince, in *Proc. of the IEEE-INNS-ENNS Int. Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, IEEE, Piscataway, NJ **2000**, pp. 348–353.
[54] S. Dey, T. Mukhopadhyay, S. Sahu, S. Adhikari, *Composites, Part B* **2016**, *105*, 188.
[55] T. Mukhopadhyay, *J. Sandwich Struct. Mater.* **2018**, *20*, 885.
[56] S. Dey, T. Mukhopadhyay, S. Adhikari, *Compos. Struct.* **2015**, *122*, 526.
[57] N. Vu-Bac, T. Lahmer, X. Zhuang, T. Nguyen-Thoi, T. Rabczuk, *Adv. Eng. Software* **2016**, *100*, 19.