



# Physics-informed neural networks for gravity field modeling of small bodies

John Martin<sup>1</sup> · Hanspeter Schaub<sup>2</sup>

Received: 14 January 2022 / Revised: 29 August 2022 / Accepted: 30 August 2022 /

Published online: 5 October 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

## Abstract

The physics-informed neural network (PINN) gravity model offers a novel and efficient way to represent high-fidelity gravity fields. PINNs leverage modern deep learning strategies to generate custom basis functions capable of modeling idiosyncratic features of a celestial body's gravity field, bypassing the inefficiencies encumbered by gravity models which prescribe geometries like spherical harmonics. Prior research on the PINN gravity model focuses on its ability to represent the gravity fields of large, near-spherical celestial bodies with highly discontinuous surface features. This research extends the investigation of the PINN gravity model to the small-body regime. Specifically, the results demonstrate that the PINN gravity model is capable of solving the additional challenges associated with modeling small-body gravity fields such as divergence within the Brillouin sphere, cumbersome computational requirements, and sample-inefficient regression. Further, this paper investigates strategies to improve network gravity modeling performance—demonstrating how additional physics constraints in network cost function can increase robustness to noise in the training data and how introducing transformer-inspired changes to the network architecture can offer order-of-magnitude improvements in modeling accuracy.

**Keywords** Gravity field modeling · Machine learning · Physics-informed neural network · Spherical harmonics · Small-bodies · Astrodynamics

---

This article is part of the topical collection on Dynamics of Space Debris and NEO.

---

Guest Editors: Xiyun Hou, Massimiliano Vasile and Alessandra Celletti.

---

This article is part of the topical collection on Machine Learning in Celestial Mechanics and Dynamical Astronomy.

---

Guest Editors: Massimiliano Vasile, Xiyun Hou, Roberto Furfaro and Alessandra Celletti.

---

✉ John Martin  
john.r.martin@colorado.edu

<sup>1</sup> Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, USA

<sup>2</sup> Professor and Department Chair, Schaden Leadership Chair, Aerospace Engineering Sciences, University of Colorado Boulder, Boulder, USA

## 1 Introduction

The rise of small-body exploration within the past decade (Lauretta et al. 2017; Kawaguchi et al. 2008; Oh et al. 2016; Russell et al. 2007; Watanabe et al. 2017) has spurred an increasing need for robust gravity models that are able to represent the exotic gravitational environments about asteroids and comets. Historically, the community has relied extensively on the spherical harmonic representation (Lemoine et al. 2014; Tapley 2008; Han et al. 2002) as it provides an exact, analytic solution to Laplace's equation, while conveniently expressing planetary oblateness with a single coefficient in the expansion ( $J_2$  or  $C_{2,0}$ ). Spherical harmonics' ability to capture such oblateness makes it particularly popular for use around near-spherical bodies like the Earth, Moon, and other large planetary objects (Jekeli 2015).

The spherical harmonics gravity model, however, is not a universally convenient representation. Within the sphere that encloses all mass elements of the body (the Brillouin sphere), the representation begins to diverge (Werner and Scheeres 1997). At altitudes just below the Brillouin sphere, the numerical effects are generally negligible—however, the divergence grows increasingly large at lower altitudes and higher degrees due to the  $(R/r)^l$  term within the expansion— $R$  being the Brillouin radius,  $r$ , the radius to the field point, and  $l$ , the degree of the expansion. This poses a problem for irregularly shaped objects, like asteroids, where spacecraft may operate at substantially lower radii than the Brillouin sphere.

There exist multiple alternative representations to combat this divergence. Among the most prominent are the mascon gravity model, the interior spherical harmonic gravity model, and the polyhedral gravity model. The mascon gravity model distributes a layer of point mass elements under the surface of the body and sums their potentials and resulting accelerations to form a global solution (Wittick and Russell 2018). While this representation does not diverge within the Brillouin sphere, it can struggle toward the surface of the body because the discrete nature of the distributed elements grows increasingly apparent (Tardivel 2016).

The interior spherical harmonic gravity model reformulates the original spherical harmonics model by expanding the harmonics about an origin that exists outside of the body. This representation can extend its effective radius down to a single point at the surface of the body (Takahashi et al. 2013). In this manner, the expansion is characterized by a  $(r/R)^l$  term such that the field can be expanded to arbitrarily high degree without risk of divergence at lower altitudes. The disadvantage of this approach, however, is that the representation is only stable within that sphere and, by consequence, only down to a single point on the body. As such, this approach is only viable when attempting a touch-and-go or landing operation on a localized point on the surface.

Takahashi and Scheeres (2014) propose the interior and exterior spherical Bessel gravity field representation which produces the complement of the exterior and interior spherical harmonic gravity fields, respectively, by solving Poisson's equation within the Brillouin sphere. This approach offers a more efficient global mapping of the gravity field than solving for multiple interior spherical harmonic representations, but it comes with considerable analytic complexity and sensitivity to the sourcing data. In addition, it is not free of user heuristics (an arbitrarily prescribed weight function is used in the conversion between exterior and interior Bessel coefficients) and may require optimization for best performance.

The polyhedral gravity model resolves the challenges associated with both the spherical harmonic and the mascon gravity models by generating a gravity field directly from the shape model of the object (Werner and Scheeres 1997). This representation maintains stability both within the Brillouin sphere and down to the surface of the body, making it a popular choice for use in recent asteroid missions (Chesley et al. 2020; Leonard et al. 2020; McMahon

et al. 2018). However, the polyhedral gravity model does require that an accurate shape model of the object exists. This is a non-trivial requirement which can involve processing radiometric and optical science products with precise orbit determination that can take many months (Miller et al. 2002; Lauretta et al. 2017). Moreover, the polyhedral representation requires scientists and engineers to assume a density profile of the body, which is challenging to uniquely estimate (Takahashi and Scheeres 2014). Finally, the polyhedral model can be quite computationally expensive to evaluate when left unparallelized (Wittick and Russell 2019)—often leaving the modeling efforts to ground based simulation rather than onboard feed-forward control purposes.

In Martin and Schaub (2022), the authors introduce the physics-informed neural network (PINN) gravity model which aims to resolve all limitations of the prior models while using fewer computational resources and no assumptions about the body in question. The PINN gravity model uses an artificial neural network to generate a convenient mapping from position to acceleration, and it “physics-informs” the network by requiring that the accelerations are computed by taking the gradient of a learned gravitational potential using automatic differentiation (Baydin et al. 2018). The network therefore learns to simultaneously penalize modeling error of the accelerations and violations of the underlying differential equations that govern the system. Such an approach allows the PINN to efficiently learn custom and physically motivated basis functions which represent the natural features in the gravitational potential of a planetary body (e.g., craters on the Moon or mountain ranges on the Earth) rather than imposing basis functions like spherical harmonics which are inefficient at capturing these idiosyncratic and often discontinuous features. Moreover, these networks have demonstrated order-of-magnitude improvements in computationally efficiency over past analytic representations.

Despite encouraging findings presented in the original PINN gravity model paper, there remains open questions regarding the capabilities and limits of the PINN gravity model particularly in the context of small-body exploration. For example, are these network representations equally robust when modeling the gravity field of non-spherical celestial bodies? Do the networks continue to generate accurate potentials and accelerations once inside of the Brillouin sphere? What training data conditions are necessary to achieve a robust model, and how sensitive are these networks to noise in the training data? Can these networks expedite the recovery of a gravity field model as compared to more traditional filtering methods?

In addition, little work has been done discussing how these networks can be designed and tuned for increased performance. In Martin and Schaub (2022), only a single physics constraint was used in the cost function of the original PINN gravity model ( $\mathbf{a} = -\nabla U$ ); however, other constraints can be introduced. Do such constraints assist in generalization or does the multi-objective nature of the additional constraints impede the learning process? Does the densely connected architecture of the network preclude the ability of the networks to learn abstract symmetries and conservation laws as suggested in Wang et al. (2020)?

This paper aims to address these questions and extend the discussion of the PINN gravity model to small-body scenarios looking specifically at performance of these models around the asteroids Eros and Bennu. Eros offers a unique gravitational environment due to its highly irregular shape and relatively smooth surface, whereas Bennu provides a more spherical shape but substantially less smooth topography.

The paper is organized as follows: Section 2 discusses the three gravity models that will be used throughout this paper (spherical harmonics, polyhedral, and PINN). Section 3 investigates the accuracy of the PINN gravity model as a function of altitude and training data distribution. Section 4 investigates how the PINN representation can be modified to improve modeling accuracy and robustness in the presence of noise. Section 5 investigates

how efficiently the PINN gravity model can estimate a gravity field using sequential data following the trajectory of NEAR-Shoemaker about Eros and the trajectory of OSIRIS-REx about Bennu. Section 6 summarizes the findings and offers insights into how performance may improve with increased training data.

## 2 Methods

### 2.1 Spherical harmonic representation

The spherical harmonic expansion of the gravitational potential,  $U$ , is defined as:

$$U_l(r) = \frac{\mu}{r} \sum_{l=0}^L \sum_{m=0}^L \left(\frac{R}{r}\right)^l P_{l,m}[\sin(\phi)][C_{l,m} \cos(m\lambda) + S_{l,m} \sin(m\lambda)], \quad (1)$$

where  $r$  is the field point,  $\mu$  is the gravitational parameter of the body,  $R$  is the circumscribing radius of the celestial body,  $l$  is the degree of the model,  $m$  is the order of the model,  $C_{l,m}$  and  $S_{l,m}$  are the spherical harmonic coefficients,  $\lambda$  is the longitude,  $\phi$  is the latitude, and  $P_{l,m}$  are the associated Legendre polynomials (Kaula 1966).

Spherical harmonics are a popular choice of gravity model as they are one of the analytic solutions to Laplace's equation:

$$\nabla^2 U = 0. \quad (2)$$

However, spherical harmonics grow unstable when evaluating field points that exist inside of the Brillouin sphere due to the  $(\frac{R}{r})^l$  term, particularly with expansions that extend to high degree,  $l$ . At radii just below  $R$ , this divergence is generally small—however, for bodies which have surfaces that exist far within the Brillouin radius, this divergence can lead to numerical error and consequently unstable dynamics within a simulation.

### 2.2 Polyhedral gravity model

At low altitudes, the gravity fields of asteroids are often generated from a shape model using the polyhedral gravity representation. In contrast to spherical harmonics, the polyhedral model maintains stable numerics within the Brillouin sphere making it a valuable alternative for missions that attempt touch-and-go (TAG) or landing maneuvers. The polyhedral gravity model takes a shape model of a body comprised of facets and vertices and predict accelerations using the following equation:

$$\nabla U = -G\sigma \sum_{e \in \text{edges}} \mathbf{E}_e \cdot \mathbf{r}_e \cdot L_e + G\sigma \sum_{f \in \text{facets}} \mathbf{F}_f \cdot \mathbf{r}_f \cdot \omega_f, \quad (3)$$

where  $G$  is the gravitational constant,  $\sigma$  is the density of the body,  $\mathbf{E}_e$  is an edge dyad,  $\mathbf{r}_e$  is the position vector between the center of the edge and the field point,  $L_e$  is an analog to the potential contribution by the edge,  $\mathbf{F}_f$  is the face normal dyad,  $\mathbf{r}_f$  is the distance between the face normal and the field point, and  $\omega_f$  is analog to the potential contribution by the face. More details can be found in Werner and Scheeres (1997).

While the polyhedral gravity model remains stable within the circumscribing sphere, it suffers from two disadvantages. First, there is an assumption of uniform density within the body as indicated by the lack of subscript for  $\sigma$  in Eq. (3). While a unique density profile can be forced onto the model, such profiles are challenging to resolve in practice (Park et al. 2010;

Takahashi and Scheeres (2014). The second disadvantage is the computational overhead of the model. Depending on the fidelity of the shape model, evaluating Eq. (3) can include looping over 100,000s of edges and facets to evaluate the acceleration at a single field point. This quickly becomes computationally expensive and poses a challenge for both offline simulation and onboard control unless properly parallelized.

### 2.3 Physics-informed neural network gravity model

Raissi et al. (2019) introduced the physics-informed neural network (PINN)—a neural network which learns to approximate a physical process or system by using traditional deep learning stochastic gradient descent with a cost function that penalizes both modeling inaccuracies and systems that violate their underlying differential equations. In this way, not only does the network learn a custom and efficient basis to represent the natural features in the problem, it also imposes that the learned basis respects relevant physics.

For example, in the case of the gravity field modeling problem, a traditional neural network would learn the mapping between the position vector,  $\mathbf{x}$ , and the corresponding acceleration vector,  $\mathbf{a}$  by minimizing the following cost function

$$J_{00}(\Theta) = \frac{1}{N_f} \sum_{i=1}^{N_f} |\mathbf{a}_i - f(\mathbf{x}_i | \Theta)|^2, \quad (4)$$

where  $f(\mathbf{x} | \Theta)$  is the learned mapping of the acceleration vector parameterized by the weights and biases of the network  $\Theta$  (i.e.,  $f(\mathbf{x} | \Theta) = \hat{\mathbf{a}}$ ). While this cost function will yield a mapping between these two quantities, the underlying dynamics of the system may not be satisfied.

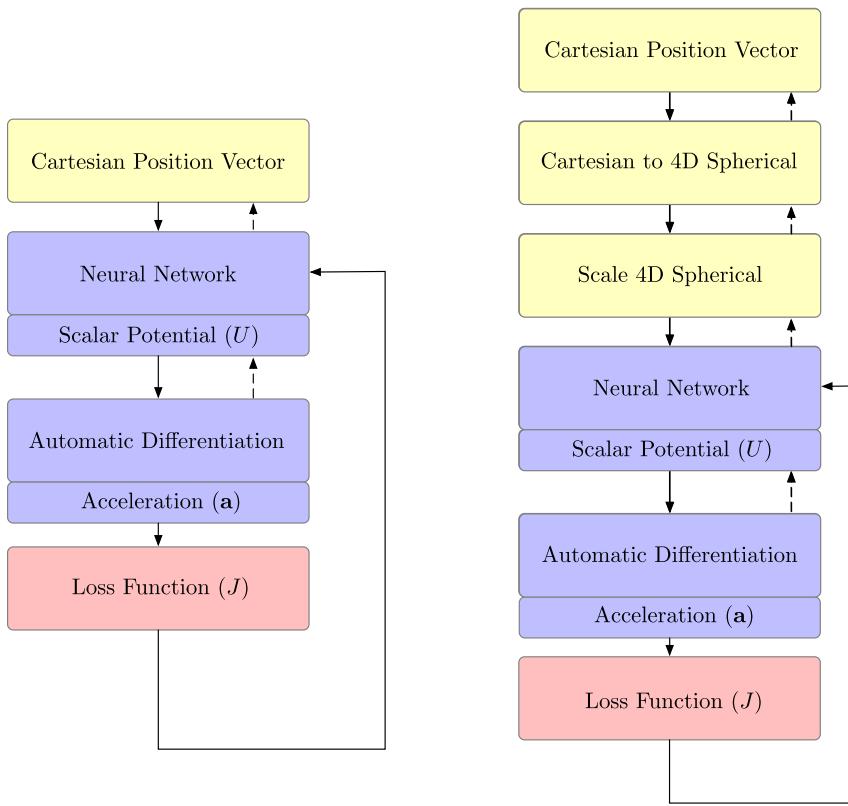
Given that the network is attempting to represent the force of gravity, it is also true that the acceleration is a by-product of a more fundamental scalar potential  $U$  through:

$$\mathbf{a} = -\nabla U. \quad (5)$$

By training with only the cost listed in Eq. (4), the network will be agnostic to this constraint and potentially generate a gravity field representation which does not satisfy this property. In contrast to this traditional approach, a physics-informed neural network specifically includes these dynamical constraints in the cost function—forcing that the network solution be consistent with the physics. To accomplish this, the network must instead learn a representation of the scalar potential,  $f(\mathbf{x} | \Theta) = \hat{U}$ , and then use automatic differentiation (AD) to take the corresponding gradient of the potential to enforce that the dynamics of Eq. (5) are satisfied through the following cost function:

$$J_A(\Theta) = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| \mathbf{a}_i - \left( -\nabla f(\mathbf{x}_i | \Theta) \right) \right|^2. \quad (6)$$

Despite the simple change, Eq. (6) provides a number of valuable consequences. Most obvious is the fact that the network is now trained with the knowledge that there is a relationship between the accelerations it produces via automatic differentiation and the more fundamental scalar potential. Second, because the network is learning a representation of the potential rather than the accelerations, all three acceleration vector components of the training data are now being used to constrain a single scalar function. This is a much more efficient regression for the network, learning a single potential rather than being forced to learn three separate acceleration features. Finally, the network now provides a differentiable function of



(a) Original PINN Gravity Model

(b) Modified PINN Gravity Model

**Fig. 1** Network structure

the scalar potential which can be used in a variety of other astrodynamics applications which will be explored in future work.

## 2.4 Modifications to original PINN gravity model

The PINN gravity models proposed in this paper deviate slightly from the original conception in Martin and Schaub (2022). This section highlights the primary changes in the implementation.

### 2.4.1 Normalization

In many machine learning applications, the data used to train neural networks are normalized to ensure that the weights and biases learned by the network are not of exceedingly large magnitude, with the goal of avoiding numerical issues (Jayalakshmi and Santhakumaran 2011). In the original PINN paper, the training data were normalized via a min-max transform on each acceleration component such that all training data existed between –1 and 1. While effective,

this approach fails to account for the natural scaling evident in the system (i.e., the accelerations should be scaled by the two independent variables responsible for its creation—the scalar potential and the position). In this paper, a physically relevant non-dimensionalization is introduced. Specifically, the physics equations are non-dimensionalized using the following scheme:

$$\mathbf{x}^* = \frac{\mathbf{x} - \mathbf{x}_0}{x_s} \quad (7)$$

$$U^* = \frac{U - U_0}{U_s} \quad (8)$$

$$\mathbf{a}^* = \frac{x_s}{U_s} \frac{dU}{d\mathbf{x}}, \quad (9)$$

where  $x_0$  corresponds to the minimum value of all position components and  $x_s$  corresponds to the range of the position vector.  $U_0$  and  $U_s$  are the minimum and range of the potential.

## 2.4.2 Feature engineering

In the original PINN gravity model, the PINNs are trained using Cartesian position and acceleration training data. While this approach is not inherently problematic, the domain of each unscaled feature ( $x, y, z, a_x, a_y, a_z$ ) extends from  $(-\infty, \infty)$ , forcing the network to learn a substantially wider feature space than if represented in spherical coordinates. This paper changes this construct by instead training PINNs on a 4D spherical coordinate set.

The 4D spherical coordinate set is chosen over the more traditional 3D set because in the traditional 3D spherical coordinate space of radius, longitude, and latitude or  $(r, \theta, \phi)$ , there exists a singularity in the gradient at the poles ( $\phi = -90^\circ$  and  $90^\circ$ ):

$$\nabla U = \frac{\partial U}{\partial r} + \frac{1}{r} \frac{\partial U}{\partial \phi} + \frac{1}{r \cos \phi} \frac{\partial U}{\partial \theta}. \quad (10)$$

This singularity poses a number of numerical challenges during PINN training. As such, the PINNs trained for this paper pass Cartesian position vectors into the Tensorflow graph, but these vectors are immediately preprocessed to a non-singular spherical coordinate representation introduced in Pines (1973). This representation uses a four-dimensional vector  $(r, s, t, u)$  where

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ s &= \frac{x}{r} \\ t &= \frac{y}{r} \\ u &= \frac{z}{r}, \end{aligned}$$

such that  $s, t$ , and  $u$  represent the tangents of angles between the field points and each of the Cartesian unit vectors  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$ . This new coordinate system maintains the advantage of substantially reducing the domain of the training data to  $r \in [0, \infty]$  and  $s, t, u \in [-1, 1]$ . Moreover, because the conversion between Cartesian coordinates and the 4D Pines' spherical coordinate set is written within the Tensorflow graph (see Fig. 1b), the potential can still be automatically differentiated with respect to the Cartesian inputs. This graph construction

produces the best of both coordinate sets: the network trains on a significantly reduced feature space (increasing network convergence and accuracy), whereas the derivatives are kept in a convenient Cartesian basis.

### 3 PINN gravity model accuracy

This section investigates how well the PINN gravity model is able to learn the gravity field of asteroid 433-Eros from position and acceleration estimates. These experiments are designed assuming that a preexisting high-fidelity gravity model exists (be it spherical harmonics, polyhedral, mascon, or other) from which perfect or noisy measurements can be drawn and used as training data to generate the PINN representation. For this section, the polyhedral representation is assumed to be ground truth and makes use of the Glaskell shape model<sup>1</sup>. Later sections will investigate how the PINN gravity model can be trained if a preexisting gravity model does not exist and instead the PINN model must be constructed using estimates made in situ (Fig. 2).

#### 3.1 Training data distributions

Generally, network modeling accuracy hinges on the quantity and distribution of its training data. As such, the PINN gravity model is trained on three unique datasets each drawn from a different distributions detailed below.

##### 3.1.1 Uniform distribution $0 - 3R(r)$

The first training distribution samples data uniformly from the surface of the asteroid to a radius equal to three times the maximum radius of the asteroid,  $R$  (Fig. 3a). This sampling represents the best case scenario—where the network is trained on data that span the entire problem domain (within, at, and above the Brillouin sphere). These conditions require that a high-fidelity, a priori gravity model exist such that low-altitude samples can be generated, or a particularly risk tolerant mission which allows the spacecraft to regularly enter these low-altitude regimes.

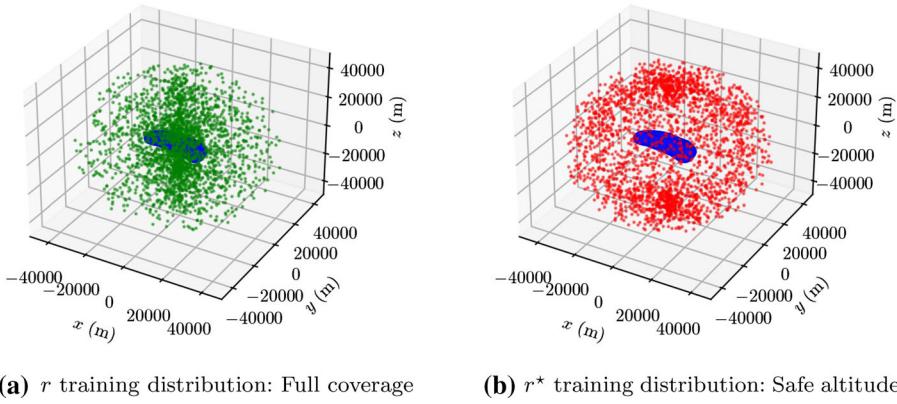
##### 3.1.2 Uniform distribution $2R - 3R(r^*)$

The second training distribution reflects more realistic circumstances—where samples are collected uniformly, but only from a range of safe altitudes above the asteroid (2–3 times the maximum radius). Such distribution can be seen in Fig. 3b. Such altitudes are generally favorable for estimating low-degree spherical harmonic models or collect images that can be used to resolve a shape model for the polyhedral representation.

##### 3.1.3 Hybrid distribution $2R - 3R$ plus sparse low-altitude samples ( $\bar{r}$ )

The final training distribution forms a hybrid set from the prior two distributions. The majority of the samples are collected from the  $2R - 3R(r^*)$  distribution, but 10% are drawn between the surface and  $2R$ . These sparser samples could represent two potential phenomena: active

<sup>1</sup> <https://arcnav.psi.edu/urn:nasa:pds:gaskell.ast-eros.shape-model>



**Fig. 2** Training distributions asteroid 433-Eros

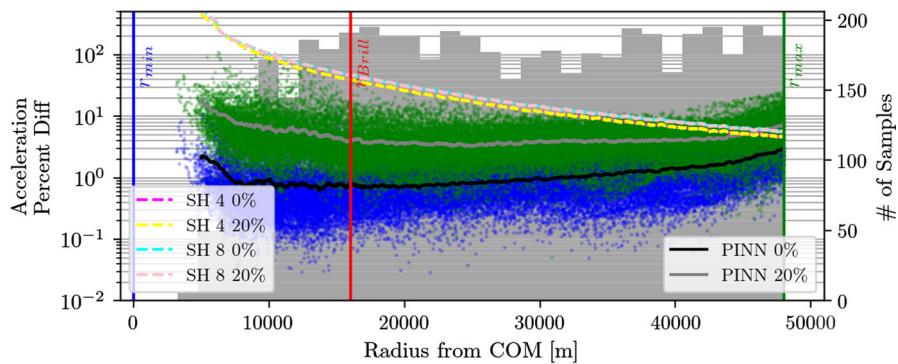
asteroid ejecta or advance mission concepts like gravity poppers. As recently found in the OSIRIS-REx radio science experiment, some asteroids eject particles which can be tracked using optical measurements and their trajectories estimated to be used in gravity field recovery (Chesley et al. 2020; McMahon et al. 2020). These particles greatly enhanced the resolution of the original radio science returns of OSIRIS-REx and has consequently motivated a new mission technology called gravity poppers (Villa et al. 2021). Gravity poppers are a technology demonstration which deploy artificial, uncontrolled probes from a mothercraft onto the surface of a small-body. Those probes then hop off of the surface of the asteroid, and the resulting dynamics are used to assist in estimating the gravity field of the body.

### 3.2 Experimental setup

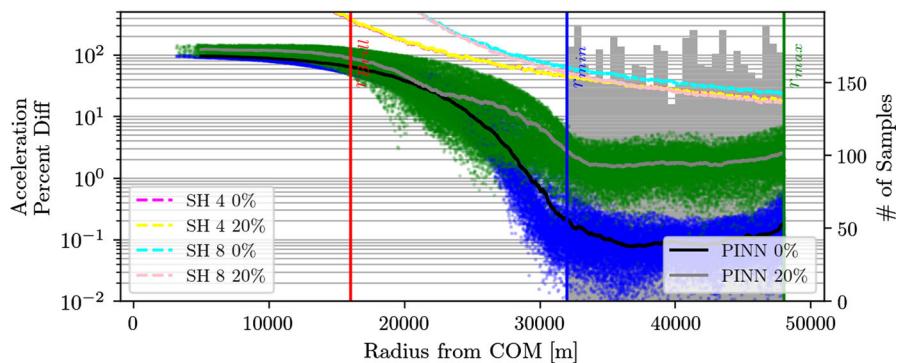
5000 training data are sampled from each of the aforementioned distributions and used to train a corresponding PINN gravity model for 7500 epochs. An epoch corresponds to a full iteration through the data set over which one or many stochastic gradient descent updates are applied depending on the mini-batch size. This process is then repeated with artificial error added to the acceleration measurements to generate a more representative, noisy data set produced by an orbit determination process using dynamic model compensation. These noisy estimates are generated by adding an error vector with a random orientation and magnitude equal to 20% of the true acceleration vector magnitude. The error is therefore non-Gaussian, and the larger the true acceleration, the larger the corresponding error.

The hyperparameters used to train these networks are listed in Table 1.

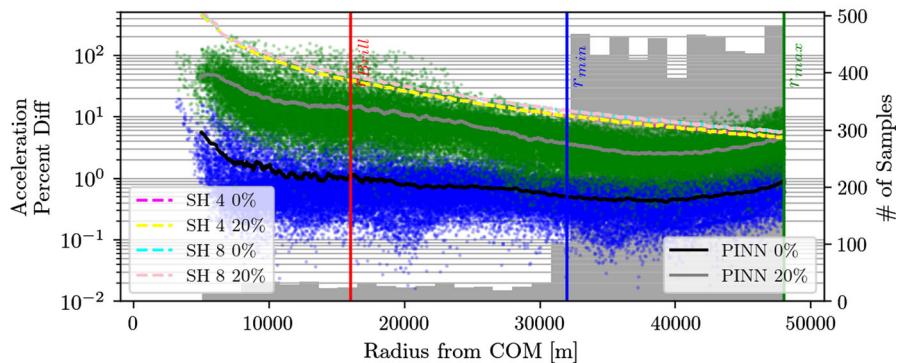
After training each model, 20,000 test samples are selected randomly from the surface of the body to an altitude corresponding to a  $3R$  radius. The true accelerations at the test locations are then differenced with the acceleration estimates produced by the PINN and the corresponding residuals and moving average of the percent error are plotted. In addition, two spherical harmonic models (one of degree and order 4, the other of degree and order 8) are fit to the same 5000 training data using least squares regression and are tested on the same 20,000 test samples. The spherical harmonic model residuals and average error are also plotted to serve as a comparison point. These curves are presented over the corresponding training data distribution histograms in Fig. 3.



**(a)** Training distribution  $r$  and corresponding model errors.



**(b)** Training distribution  $r^*$  and corresponding model errors.



**(c)** Training distribution  $\bar{r}$  and corresponding model errors.

**Fig. 3** Acceleration residuals,  $|\mathbf{a} - \hat{\mathbf{a}}|/|\mathbf{a}| \times 100$ , their moving averages, and training data distribution as a function of radius. Blue scatter plots correspond to error of the PINN trained on perfect measurements, and green corresponds to the PINN trained on noisy measurements. Dashed lines represent the error of the spherical harmonic models fit on the same data. The gray histogram represents the radial distribution of the training data

**Table 1** Nominal Hyperparameters

Parameter	Value
Activation	GELU
Batch size	5000
Optimizer	adam
Epochs	7500
Hidden layers	8
Nodes per layer	20
Weight initialization	Glorot normal
Learning rate	0.002

### 3.3 Results

In the best case scenario ( $r$  without noise), the PINN models accelerations with less than 3% average error all the way down to the surface—an encouraging feat given that the network was exposed to only 5000 data points. When the 20% acceleration error is introduced into the estimates, the PINN's error also increases, but remains below an average of 10%. The regressed spherical harmonics models, in contrast, never achieve better than 5% average error—*independent* of amount of error in the estimates—and can be seen diverging as they approach lower altitudes.

In the case of training PINNs on  $r^*$ , the learned representations achieve  $< 0.1\%$  average error with perfect measurements and  $< 3\%$  when error is included in the training data domain; however, the PINNs suffer from extrapolation error beyond these bounds. While this behavior is undesirable, it is not unique to the PINN models. The regressed spherical harmonic models also suffer from greater extrapolation error. Moreover, even within the training domain, spherical harmonics models generate higher error than the PINNs. For both PINNs and spherical harmonics, these results suggest that generalizable models require training samples from lower altitudes if they are to be used in close proximity to the body.

To investigate how many low-altitude samples are necessary, the  $\bar{r}$  distribution supplements the  $r^*$  distribution with 500 additional samples collected between the asteroid surface and  $2R$  radius. By including this relatively small set of measurements, the PINN modeling accuracy stabilizes considerably. When no error is included within the training data, the average error of the PINN remains  $< 1\%$  for the majority of altitudes. Only at the surface of the body does the PINN model achieve an average modeling error of 5%. When error is added to the acceleration training data, the modeling accuracy decreases, but the total error remains below 10% for the majority of altitudes and only reaches a maximum average modeling error of 40% when evaluated at the surface—both sizable improvements in accuracy compared to the networks trained only on  $\bar{r}$ . This experiment is particularly encouraging as it implies that PINNs are quite efficient at extracting meaningful information about the gravity field when given access to sparse low-altitude samples. Such findings provide strong evidence in favor of concepts like gravity poppers. Given how little data are needed to achieve generalizable models, there may not be a need for many gravity poppers or even poppers that can produce diverse trajectories. A formal analysis of this claim is left for future work.

## 4 Improving PINN performance

The results thus far have demonstrated that the PINN gravity model, in its simplest form, is capable of achieving superior modeling accuracy than its spherical harmonic counterpart given sparse training sets. This section investigates various strategies for improving PINN gravity model accuracy; specifically looking at how modifying the cost function to include additional physics-informed constraints affects modeling accuracy, as well as determining whether transitioning from a densely connected network to a transformer-inspired architecture provides modeling advantages.

### 4.1 Additional PINN constraints

The cost function proposed for the original PINN gravity model only penalized errors in the predicted acceleration (as denoted by the subscript A in Eq. 6). While effective, Eq. (6) fails to account for other dynamical properties that must be observed by the system. For example, the force of gravity is a conservative force—therefore the scalar potential learned by the network must also obey these additional physics properties:

$$\nabla^2 U = 0 \quad \text{Laplacian (L)} \quad (11)$$

$$\nabla \times \mathbf{a} = 0 \quad \text{Curl (C)} \quad (12)$$

These conservative vector field properties can therefore be embedded into the original PINN cost function through:

$$J_{\text{ALC}}(\Theta) = \frac{1}{N_f} \sum_{i=1}^{N_f} \left| \mathbf{a}_i + \overset{\text{AD}}{\nabla} f(\mathbf{x}_i | \Theta) \right|^2 + \left| \overset{\text{AD}}{\nabla^2} f(\mathbf{x}_i | \Theta) \right|^2 + \left| \overset{\text{AD}}{\nabla} \times \overset{\text{AD}}{\nabla} f(\mathbf{x}_i | \Theta) \right|^2. \quad (13)$$

In addition, if a model of the gravitational potential exists, the cost function can be extended to penalize those modeling errors (subscript P):

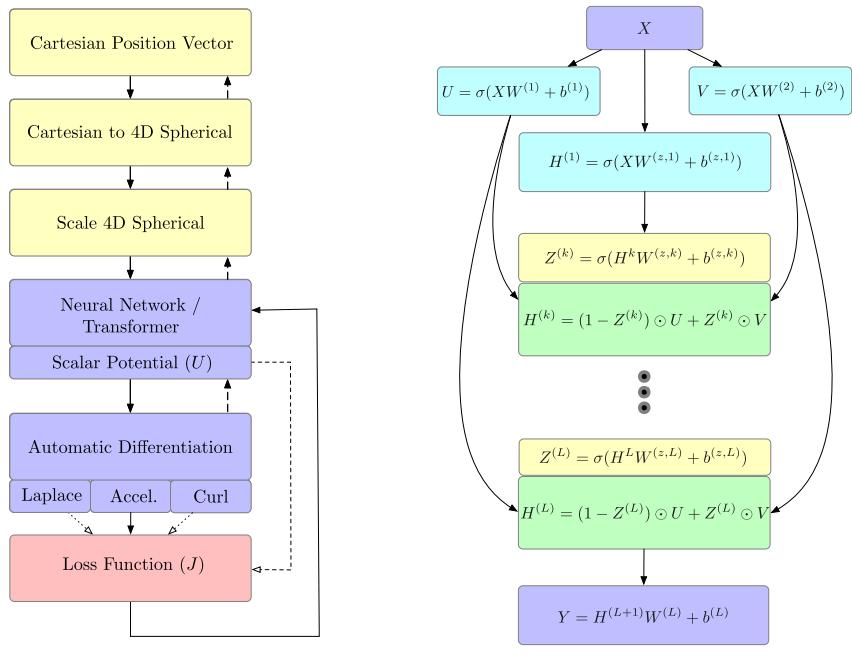
$$J_{\text{APLC}}(\Theta) = J_{\text{ALC}}(\Theta) + \frac{1}{N_f} \sum_{i=1}^{N_f} |U_i - f(\mathbf{x}_i | \Theta)|^2. \quad (14)$$

These additions to the cost function act as a form of regularization during training which is hypothesized to help improve robustness, particularly in the presence of noise.

### 4.2 Network architectures

In addition to investigating the effect of additional physics constraints on network model accuracy, this work also looks at the effect of changing the network architecture. In Wang et al. (2020), the authors observe that the success of many modern machine learning models relies on embedding symmetry groups or invariances into the network architectures (be it translational invariances in convolutional networks or temporal invariances in recurrent neural networks). Embedding these properties into the network architectures for physical processes can be challenging as it assumes preexisting knowledge of these symmetry groups which may not be present. For this reason, it is common to rely only on fully connected networks in modeling physical processes as opposed to more advanced architectures.

Wang et al. (2020) imply that this choice may not be sufficient, as there does not yet exist a proof of the convergence of fully connected Physics-Informed Neural Networks. By conse-

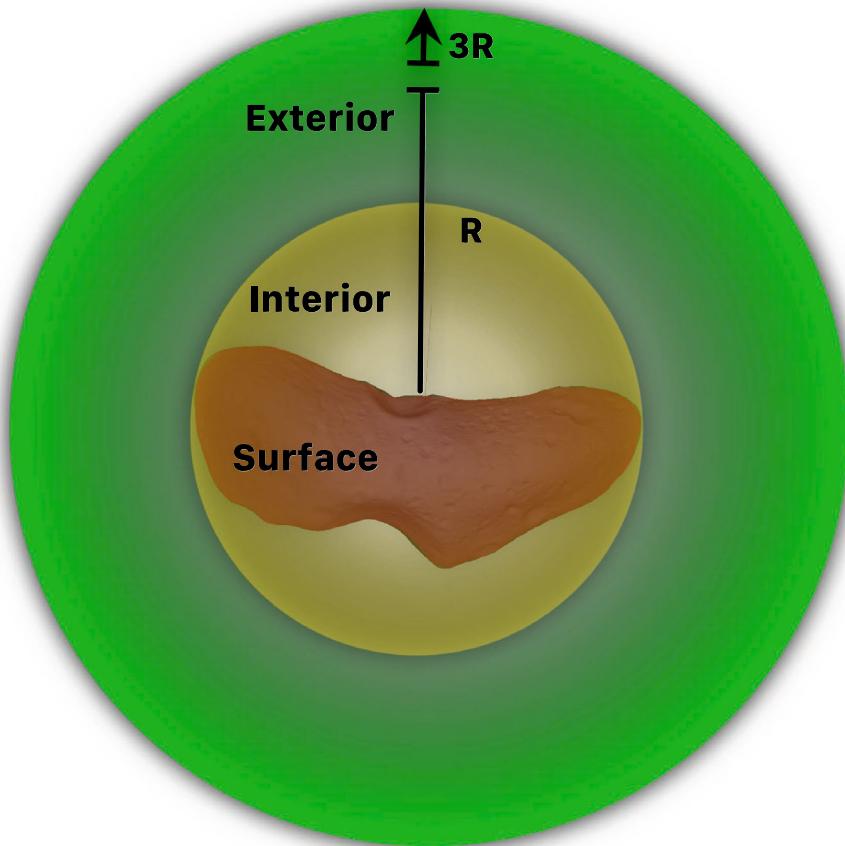


**Fig. 4** Modified network structure

quence, researchers can only hope that fully connected networks provide sufficient modeling capacity and flexibility to represent the solution to the corresponding differential equation. The authors supplement this claim by introducing how a relatively simple change in network architecture can significantly improve the accuracy of the learned model in a collection of archetypal PINN problems. This paper adopts the modified architecture presented in Wang et al. (2020) which makes use of advances in neural attention mechanisms. Specifically, the PINNs in this paper extend the fully connected network with two transformers, which project the inputs into a high-dimensional space. In addition, the hidden states are enhanced with residual connections from prior layers through multiplicative, element-wise interactions as shown in Fig. 4b. These modifications introduce relatively minor changes to the memory and computational footprint of the fully connected network but offer appreciable gains in model accuracy. This architecture will be referred to as the physics-informed transformer (PIT) for the remainder of this study.

### 4.3 Experimental setup

To evaluate the degree to which these modifications improve model accuracy, this experiment trains the modified PINNs and PITs on increasingly sparse datasets ( $N = \{2500, 1250, 625\}$ ) comprised of position/acceleration pairs that are sampled randomly from  $0 - 3R$ . Once trained, the modeling error of these PINN and PIT representations is then evaluated on separate test datasets comprised of 20,000 randomly distributed position/acceleration pairs. The first test set contains 20,000 positions/acceleration pairs sampled outside of the Brillouin



**Fig. 5** Eros data distributions

sphere, and within  $3R$  of the body (exterior), the second test set is sampled between the surface of the asteroid and the Brillouin sphere (interior), and finally the third set samples positions/accelerations from 20,000 randomly selected facets of the polyhedral shape model (surface). These three test dataset regions are shown in Fig. 5. Each network model is trained using only one of the cost functions ( $J_A$ ,  $J_{AP}$ ,  $J_{ALC}$ ,  $J_{APLC}$  and a non-physics-informed cost function, denoted  $J_{00}$ —see Eq. 4). In addition, each of the representations is trained on datasets with increasing levels of error added to the acceleration data. The erroneous acceleration estimates are generated in the same manner as presented in Sect. 3. The percent errors tested are 0%, 10%, and 20% of the true acceleration magnitude. Three spherical harmonic gravity models (corresponding to degree  $l = \{4, 8, 16\}$ ) are also regressed on each training dataset using least squares and evaluated using the same test datasets.

#### 4.4 Results

The performance of the PINNs, PITs, and spherical harmonics gravity models are plotted as a function of amount of training data,  $N$ , and error in the training data (0%, 10%, and 20%).

The corresponding results for each test dataset (exterior, interior, and surface) are plotted in Figs. 6, 7, and 8, respectively. Note that the model errors are quantified as acceleration percent error distributions and they are represented as box-and-whisker plots. The box represents the 1st and 3rd quartiles (25% to 75%) of the percent error distribution, i.e., 50% of data lies within the box. The line within the box corresponds to the median or 2nd quartile. The lines extending beyond the box correspond to  $1.5 \times IQR$  where  $IQR$  is the interquartile range, and the points beyond those lines correspond qualitatively with outliers. Also, Figs. 6, 7, and 8 cluster the PINN, spherical harmonics (SH), and PIT models into different color clusters. As such, the rows within each cluster of the legend correspond to the models in the figure plotted from left to right.

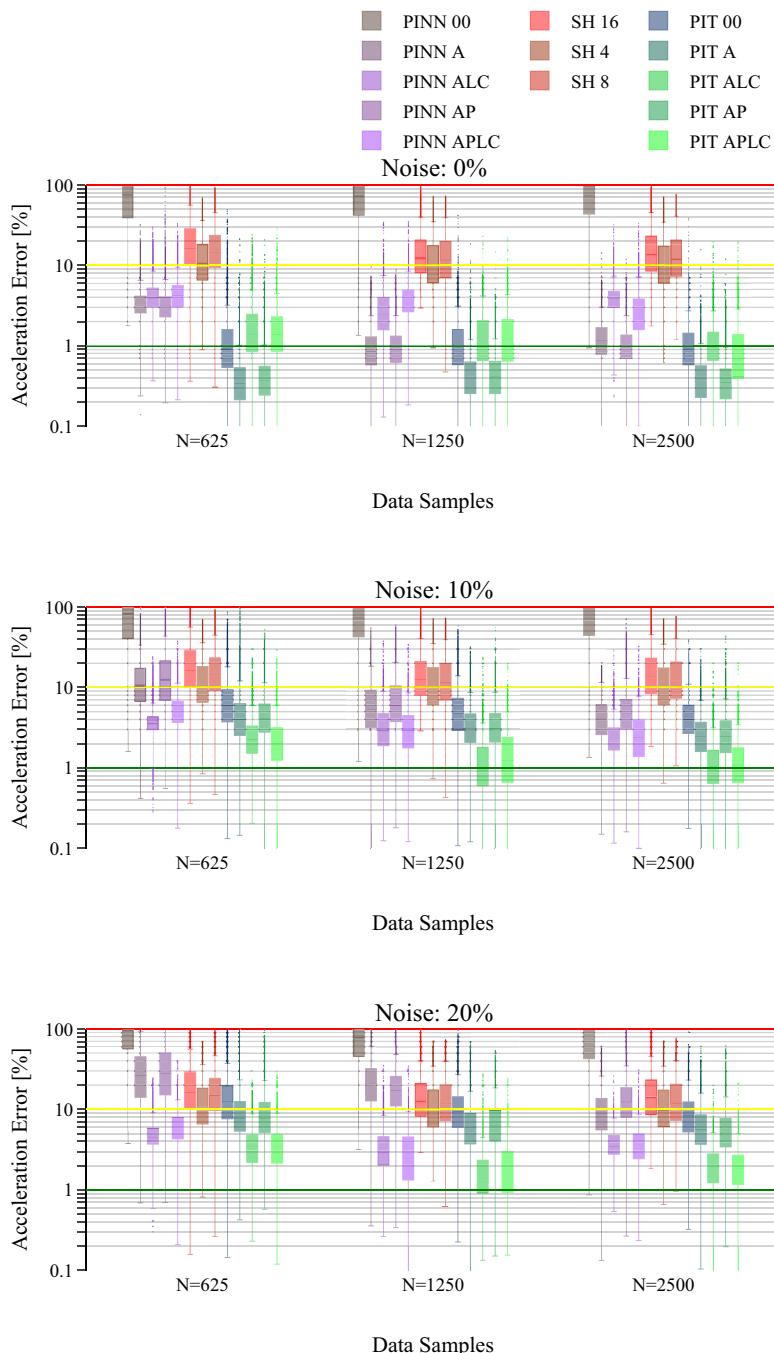
Figures 6, 7, and 8 demonstrate a number of interesting behaviors of the PINN and PIT gravity models. First, the additional constraints of the cost function act as a form of regularization to the regression. When there exists sufficient data ( $N = 2500$ ) with little-to-no noise, the multiple-constraint PINNs (ALC, APLC) tend to perform worse than their simpler counterparts (00, A, AP). However, when the noise in the dataset increases, or the number of samples decrease, the multiple-constraint networks are more robust. These networks verify that the additional constraints act as a form of regularization, minimizing the risk of overfitting to noise or sparse data and making the multiple-constraint PINN a more trustworthy representation in uncertain or novel environments. This can also be seen by comparing Figs. 3 and 9 and noting how the PINN ALC network trained on the noisy  $\bar{r}$  distribution is functionally desensitized to the noise.

The second interesting finding from Figs. 6, 7, and 8 is the consistent performance improvement of the PITs over the PINNs. In both the low- and high-sample regimes, the transformers can offer as high as an order-of-magnitude reduction in error despite being trained with the same data as best illustrated in the low-to-no noise cases of Figs. 6 and 7. It is worth noting, however, that the transformers are also more prone to overfitting than the PINNs due to their greater modeling capacity—as seen when using the simpler physics constraints (00, A, AP) on the 10% and 20% noise cases in Figs. 7 and 8. In the 0% noise case, this modeling flexibility could be considered a strength as the fewer physics constraints allow the model to converge more quickly than their more regularized variants. However, in more practical applications like when the field is not known a priori, it is recommended to use the additional constraints to minimize the risk of overfitting.

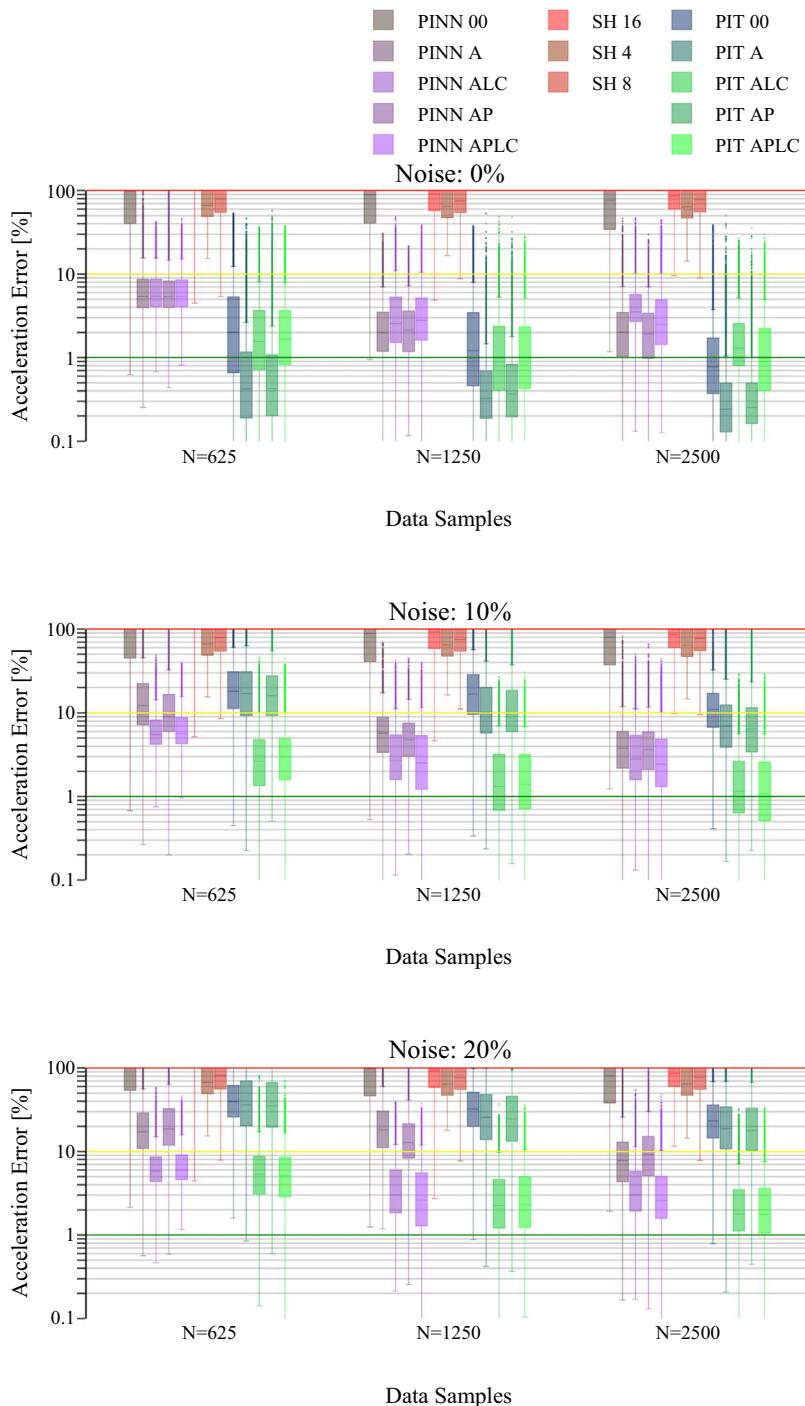
Finally, both the PINNs and PITs offer sizable performance gains over their spherical harmonic predecessor not only in the interior and surface distributions for which spherical harmonics will diverge, but also the exterior distribution where spherical harmonics has often had the analytical high-ground. This consistently worse performance of the spherical harmonic representation is a function of the small sample size, the random distribution of these samples, and the low-altitude data—each capable of posing numerical challenges in the least squares estimate of the spherical harmonic coefficients.

## 5 PINN gravity field recovery

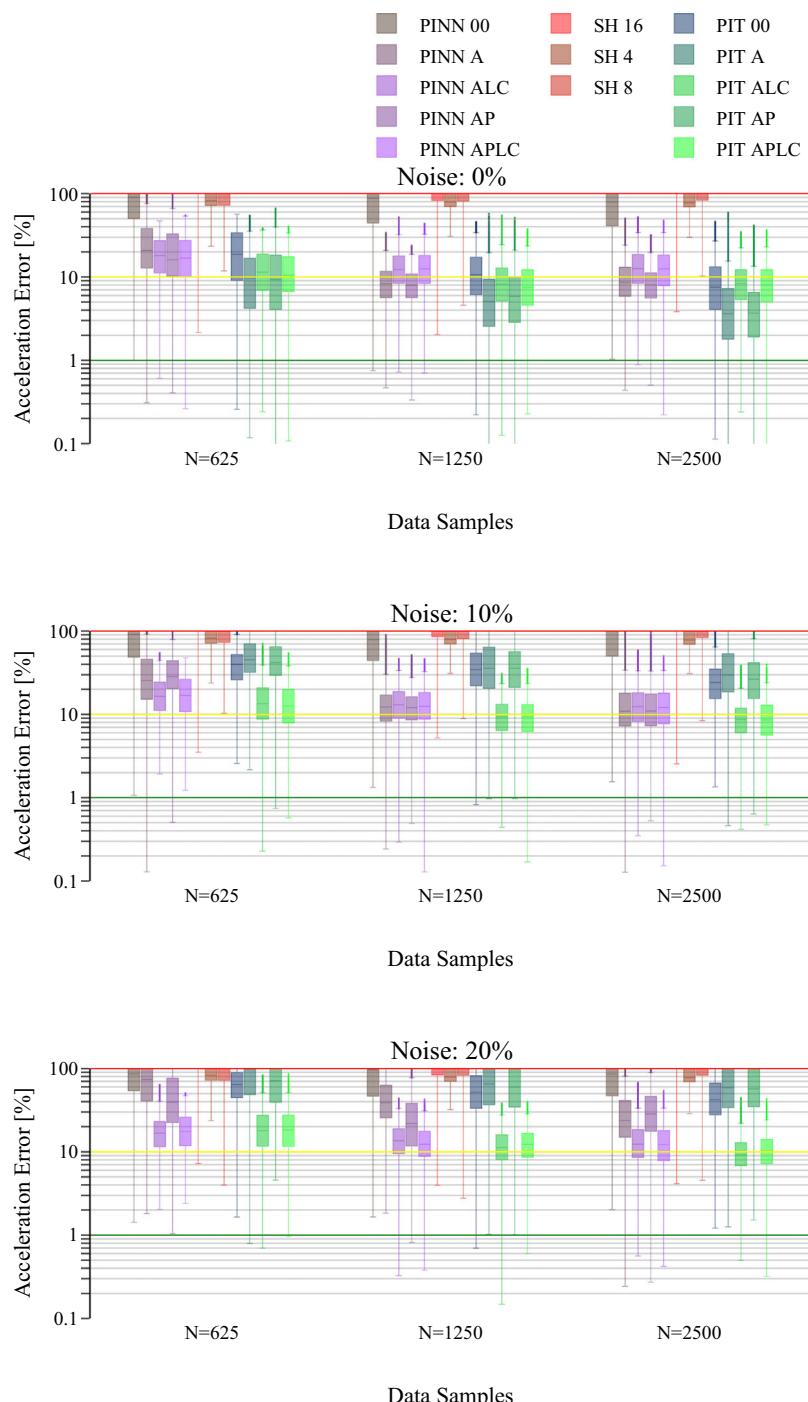
The physics-informed gravity models trained thus far assume that a prior gravity model exists from which training data can be generated and used to explore model sensitivity to data distribution, quantity, and measurement error. While the networks demonstrate robustness to each of these conditions, the experiments fail to investigate the more representative operational circumstance: when a spacecraft enters orbit around an asteroid and has no high-



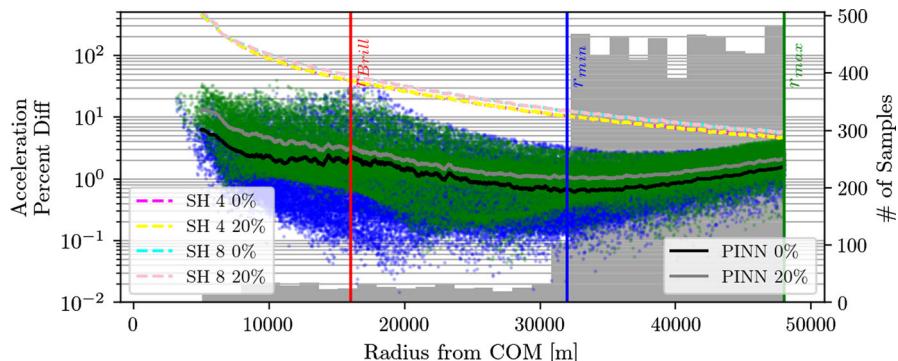
**Fig. 6** Model error *outside* the Brillouin sphere (exterior) as a function of amount of training data and noise added to the training data



**Fig. 7** Model error *inside* the Brillouin sphere (interior) as a function of amount of training data and noise added to the training data



**Fig. 8** Model error at the surface of the asteroid (surface) as a function of amount of training data and noise added to the training data



**Fig. 9** Acceleration residuals as a function of training distribution with  $J_{ALC}$  demonstrate how additional physics constraints help desensitize to noise in the training data

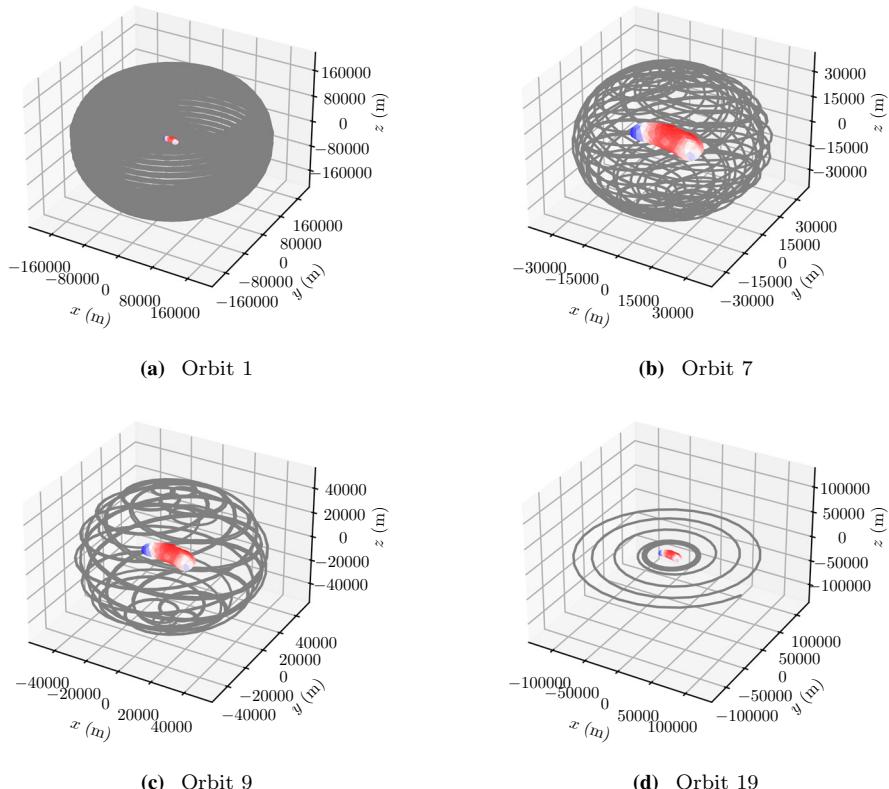
fidelity model from which training data can be generated. In this setting, the physics-informed models must act as part of the navigation filter—continuously updating its model based on incoming estimates of the spacecraft position and gravitational accelerations. This section aims to investigate the in situ construction of a PINN gravity model for the asteroids Eros and Bennu by sampling training data only along the ephemeris of their corresponding orbiters, NEAR-Shoemaker and OSIRIS-REx. By following the real trajectories of asteroid missions, this section demonstrates whether a PIT gravity model trained with the ALC cost offers advantages in gravity field recovery over spherical harmonics.

## 5.1 NEAR-shoemaker trajectory

The NEAR-Shoemaker spacecraft entered orbit around 433-Eros on February 14, 2000 beginning a year-long campaign comprised of 24 increasingly close orbits used to study the asteroid (Miller et al. 2002). A sample of these trajectories are plotted in Fig. 10.

This experiment assumes that an erroneous acceleration estimate is calculated and stored once every 10 minutes during the mission lifetime. These assumptions are supported from two articles, Leonard et al. (2013) and Villa et al. (2021). The former article demonstrates that high-order gravitational accelerations can be directly estimated in a onboard filter as part of the state using high-order forms of dynamic model compensation. The latter article provides justification for the 10-minutes cadence between state estimates in asteroids settings. These noisy acceleration estimates are generated by first calculating the true acceleration at the current spacecraft position, and then adding an error vector with a magnitude equal to 10% of the true acceleration magnitude in a random direction. This is consistent with the methodology presented in Sects. 3 and 4. It is worth emphasizing that this choice results in higher acceleration errors when the spacecraft is closer to the asteroid than when it is far away. While this choice is not necessarily reflective of true filter estimates and evolving covariances, it remains useful as it provides a challenging dataset that can be used to highlight the strengths and weaknesses of different gravity field recovery strategies.

The experiment progresses by generating the corresponding position and noisy acceleration data along the ephemeris of NEAR-Shoemaker. Each time the spacecraft transitions to a new orbital configuration, of which there are 24, all previously stored acceleration estimates are then used in a batched least squares algorithm to estimate three different spherical har-

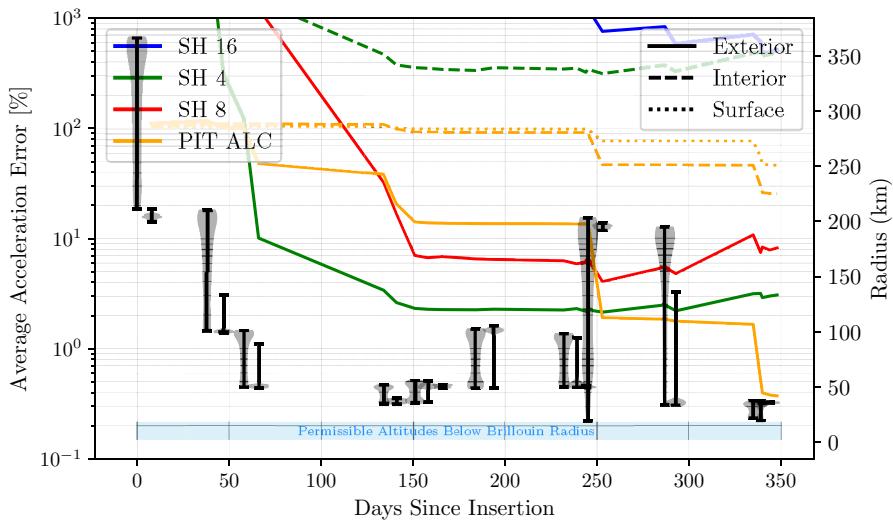


**Fig. 10** A random selection of NEAR-Shoemaker science orbits around 433-Eros

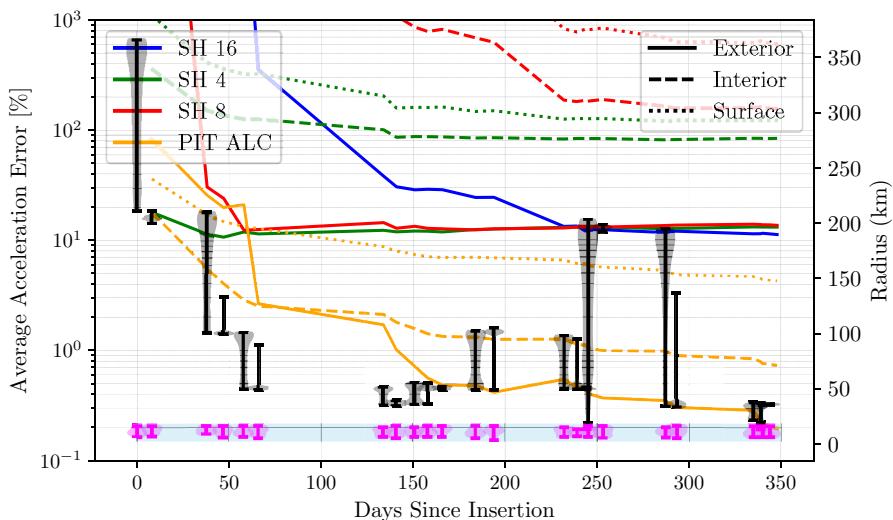
monic models of degree and order 4, 8, and 16 respectively. Likewise, a PIT model with the ALC constraint cost is also trained on this data and is given 7500 epochs to converge.

There exists a rich body of literature discussing more advanced algorithms to estimate high-degree spherical harmonic models (Xu and Rummel 1994; Xu 1992, 1998; Xu et al. 2006). The choice to use a traditional least squares approach rather than these alternative algorithms is purposeful. Many of these more sophisticated algorithms rely on embedding heuristic insights about the structure of the spherical harmonic coefficients into the regression (e.g., using Kaula's rule for ridge regression), or they rely on the assumption that the expected measurement error has a mean of zero. Because the zero mean error requirement is not satisfied in this problem, and because the PITs are not given any additional user heuristics, traditional linear least squares regression is chosen to yield a more balanced comparison between these two model types.

Figure 11 plots the acceleration percent error produced by each intermediate model as the spacecraft progresses through the mission. The model error is calculated by evaluating each model on 20,000 randomly distributed test data within each of the three major distributions shown in Fig. 5 (exterior, interior, and surface). The lines corresponding to each test dataset distribution are plotted as solid, dashed, and dotted linestyles in the figure, respectively. In addition, the altitude distribution of the spacecraft in each of the 24 orbits is shown as black violin plots to demonstrate at what point in time the spacecraft is close to the body and when



**Fig. 11** Error of regressed PIT gravity model of Eros without Ejecta



**Fig. 12** Error of regressed PIT gravity model of Eros with Ejecta measurements (Magenta)

it is far away. Note that some of the resulting curves/linestyles exceed the y-bounds of the figure due to their high error.

Interestingly, the PIT performs slightly worse than the low-degree spherical harmonic model on the exterior distribution for the majority of the mission. Only after 250 days of orbiting does the PIT offer a more accurate model. This reflects one of the conclusions reached in Sect. 3: PINNs are particularly sensitive to the amount of low-altitude training data. Nearly all major improvements in the PIT model accuracy came from orbits which had close passes with the asteroid. This is best demonstrated in the near constant modeling error from day 170 to 240 which included orbits with the same minimum altitude. As soon as the spacecraft had

a closer approach starting on day 240, the modeling error for the exterior distribution drops by nearly an order-of-magnitude from a brief low-altitude pass.

Another interesting feature of Fig. 11 is how the low degree spherical harmonic models outperform the high-degree spherical harmonic models. Multiple factors contribute to this result. Foremost, the least squares algorithm does not have a way to filter out the error included within the acceleration estimates. A weighted least squares algorithm could be used instead, but such choice assumes that the error in the measurement is of zero mean for the solution to be optimal which is not true in this experiment. Second, because the high-degree models have greater modeling capacity than their low-degree counterparts, they are more susceptible to capturing these errors in their solution, thereby producing a more erroneous fit when run on the test set. Encouragingly, the PIT gravity model is not sensitive to these acceleration errors. By including the additional physics constraints in the cost function, the model remains desensitized to the error similar to what is seen in Sect. 4.

Given the PITs sensitivity to the low-altitude training data, a second experiment is proposed which seeks to exploit this phenomenon. The second experiment supplements the original training data set with infrequent acceleration estimates produced by particles tracked between the surface of the asteroid and the Brillouin radius (via ejecta events or gravity poppers). A total of 504 additional acceleration estimates are sampled from this range and are evenly distributed across the entire mission duration. This choice attempts to produce similar conditions to those found during the OSIRIS-REx mission in which a minimum of 600 particle observations were used to update the high-degree spherical harmonic coefficients (Chesley et al. 2020).

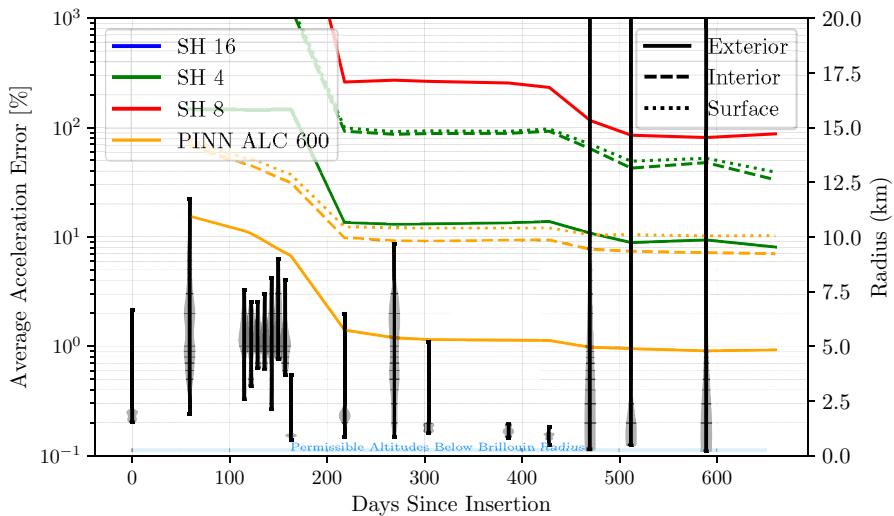
By including these additional 504 measurements over a year-long mission, the PIT ALC gravity model achieves remarkably better performance than the traditional spherical harmonics—converging to  $< 10\%$  error within 10 days, and  $< 1\%$  error after only 100 days (see Fig. 12)—feats never accomplished by spherical harmonics over the entire mission lifetime. This rapid convergence suggests that if a mission could actively search for additional ejecta events near the surface of the body or artificially generate them via gravity poppers, the time necessary to characterize an asteroid’s gravity field could be reduced by an order of magnitude, saving valuable mission time and resources to be repurposed for closer approaches and additional science opportunities. Together, these results demonstrate how the PIT gravity model is able to learn a representation that is substantially more accurate and faster to converge than both its low- and high-degree spherical harmonic counterparts.

## 5.2 OSIRIS-REx trajectory

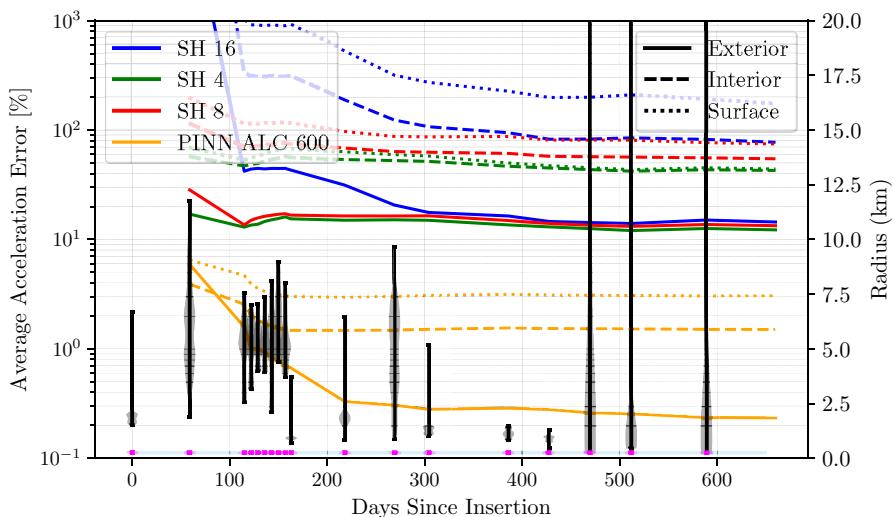
To supplement the findings of Sect. 5.1, the same experiment is repeated following the trajectory of OSIRIS-REx mission about the asteroid Bennu. OSIRIS-REx spent approximately two years in orbit during which it transitioned from preliminary survey orbits into increasingly close approaches in anticipation for the Touch-and-Go (TAG) maneuver to collect a sample of regolith that is ultimately to be returned to Earth (Lauretta et al. 2017).

Like the last experiment, noisy acceleration estimates are generated every 10 minutes for the duration of the mission. Spherical harmonic and PIT models are regressed on all prior data after each mission phase/orbital configuration, and their respective error in the exterior, interior, and surface test distributions is plotted in Fig. 13.

Unlike the 433-Eros field recovery case, the PIT does not require additional low-altitude samples from ejecta or poppers for the PIT model to achieve better model accuracy than spherical harmonics. Given the closer orbits OSIRIS-REx was able to enter around the aster-



**Fig. 13** Error of regressed PIT gravity model of Bennu without Ejecta



**Fig. 14** Error of regressed PIT gravity model of Bennu with Ejecta measurements (Magenta)

oid, the PITs had an advantage over spherical harmonics, achieving order-of-magnitude better modeling accuracy than the corresponding spherical harmonic representation in each respective testing domain. Moreover, if the original acceleration estimates of the spacecraft are supplemented with ejecta or popper measurements as before (one measurement per every 1000 minutes), the transformer is able to achieve  $< 3\%$  error in all three test distributions after a mere 150 days (see Fig. 14).

## 6 Conclusions

This paper characterizes the modeling accuracy and robustness of the PINN gravity model in the small-body environment. The findings demonstrate that PINNs and their evolved counterpart, PITs, provide substantial improvements in modeling accuracy and convergence as compared to their spherical harmonic and polyhedral predecessors about small bodies. Moreover, the results presented here are generally conservative, having trained the PINNs and PITs on datasets that are artificially small. Given additional data and more confident state estimates, it is reasonable to assume that all models presented in this paper will see appreciable modeling improvements.

In addition, this paper presents an analysis which investigates how the PINN gravity model performs when an a priori gravity model does not exist. The results demonstrate that given the same amount of noisy data, the PINN gravity model is capable of producing a more reliable estimate of the gravity field both outside and inside the Brillouin sphere in considerably less time than would be required if using spherical harmonics. Future work will build on these findings, discussing how the PINN gravity model can be directly embedded into the orbit determination process and comparing its performance to past filtering methods for gravity field estimation.

**Acknowledgements** This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 2040434.

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## References

- Baydin, A.G., Pearlmutter, B.A., Siskind, J.M.: Automatic differentiation in machine learning: a survey. *J. Mach. Learn. Res.* **18**, 1–43 (2018)
- Chesley, S.R., et al.: Trajectory estimation for particles observed in the vicinity of 101955 Bennu. *J. Geophys. Res. Planets* (2010). <https://doi.org/10.1029/2019JE006363>
- Han, S.-C., Jekeli, C., Shum, C.K.: Efficient gravity field recovery using in situ disturbing potential observables from CHAMP. *Geophys. Res. Lett.* **29**(16), 1–4 (2002). <https://doi.org/10.1029/2002GL015180>
- Jayalakshmi, T., Santhakumaran, A.: Statistical normalization and back Propagationfor classification. *Int. J. Comput. Theory Eng.* **3**(1), 89–93 (2011). <https://doi.org/10.7763/ijcte.2011.v3.288>
- Jekeli, C.: Potential theory and the static gravity field of the earth. In: *Treatise on Geophysics: Second Edition* 3, pp. 9–35. issn: 9780444538031. (2015) <https://doi.org/10.1016/B978-0-444-53802-4.00056-7>.
- Kaula, W.M.: *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*. Blaisdell Publishing Co, Waltham, Mass. (1966)
- Kawaguchi, J., Fujiwara, A., Uesugi, T.: Hayabusa-its technology and science accomplishment summary and Hayabusa-2. *Acta Astron.* **62**(10–11), 639–647 (2007). <https://doi.org/10.1016/j.actaastro.2008.01.028>
- Lauretta, D.S., et al.: OSIRIS-REx: sample return from asteroid (101955) Bennu. *Space Sci. Rev.* **212**(1–2), 925–984 (2017). <https://doi.org/10.1007/s11214-017-0405-1>
- Lemoine, F.G., et al.: GRGM900C: a degree 900 lunar gravity model from GRAIL primary and extended mission data. *Geophys. Res. Lett.* **41**(10), 3382–3389 (2014). <https://doi.org/10.1002/2014GL060027>
- Leonard, J.M., Geeraert, J.L., et al.: Osiris-Rex orbit determination performance during the Navigation Campaign. *Adv. Astron. Sci.* **171**, 3031–3050 (2020)
- Leonard, J.M., Nievinski, F.G., Born, G.H.: Gravity error compensation using second-order gauss-markov processes. *J. Spacecr. Rockets* **50**(1), 217–229 (2013). <https://doi.org/10.2514/1.A32262>
- Martin, J., Schaub, H.: Physics-informed neural networks for gravity field modeling of the Earth and Moon. *Celest. Mech. Dyn. Astron.* (2022). <https://doi.org/10.1007/s10569-022-10069-5>
- McMahon, J.W., Scheeres, D.J., Hesar, S.G., et al.: The OSIRIS-REx Radio science experiment at Bennu. *Space Sci. Rev.* (2018). <https://doi.org/10.1007/s11214-018-0480-y>

- McMahon, J.W., Scheeres, D.J., Chesley, S.R., et al.: Dynamical evolution of simulated particles ejected from Asteroid Bennu. *J. Geophys. Res. Planets* **125**(8), 1–18 (2020). <https://doi.org/10.1029/2019JE006229>
- Miller, J.K., et al.: Determination of shape, gravity, and rotational state of asteroid 433 Eros. *Icarus* **155**(1), 3–17 (2002). <https://doi.org/10.1006/icar.2001>
- Oh, D. Y. et al.: psyche: journey to a MetalWorld. In: 52nd AIAA/SAE/ASEE Joint Propulsion Conference, 2016 2014, pp. 1–11. issn: 9781624104060. (2016) <https://doi.org/10.2514/6.2016-4541>.
- Park, R.S., Werner, R.A., Bhaskaran, S.: Estimating small-body gravity field from shape model and navigation data. *J. Guid. Control Dyn.* **33**(1), 212–221 (2010). <https://doi.org/10.2514/1.41585>
- Pines, S.: Uniform representation of the gravitational potential and its derivatives. *AIAA J.* **11**(11), 1508–1511 (1973). <https://doi.org/10.2514/3.50619>
- Raiassi, M., Perdikaris, P., Karniadakis, G.: Physics-Informed Neural Networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019). <https://doi.org/10.1016/j.jcp.2018.10.045>
- Russell, C.T., et al.: Dawn Mission to Vesta and Ceres. *Earth Moon Planets* **101**(1–2), 65–91 (2007). <https://doi.org/10.1007/s11038-007-9151-9>
- Takahashi, Y., Scheeres, D.J.: Small body surface gravity fields via spherical harmonic expansions. *Celest. Mech. Dyn. Astron.* **119**(2), 169–206 (2014). <https://doi.org/10.1007/s10569-014-9552-9>
- Takahashi, Y., Scheeres, D.J., Werner, R.A.: Surface gravity fields for asteroids and comets. *J. Guid. Control Dyn.* **36**(2), 362–374 (2013). <https://doi.org/10.2514/1.59144>
- Takahashi, Y., Scheeres, D.: Morphology driven density distribution estimation for small bodies. *Icarus* **233**, 179–193 (2014). <https://doi.org/10.1016/j.icarus.2014.02.004>
- Tapley, B.D.: Gravity model determination from the GRACE Mission. *J. Astron. Sci.* **56**(3), 273–285 (2008). <https://doi.org/10.1007/BF03256553>
- Tardivel, S.: The Limits of the Mascons Approximation of the Homogeneous Polyhedron. In: AIAA/AAS Astrodynamics Specialist Conference. September. Reston, Virginia: American Institute of Aeronautics and Astronautics, pp. 1–13. isbn: 978-1-62410-445-9, (2016) <https://doi.org/10.2514/6.2016-5261>.
- Villa, J. et al.: Gravity estimation of small bodies via optical tracking of hopping artificial probes. In: AAS/AIAA Astrodynamics Specialist Conference. Big Sky, Montana, pp. 1–21. (2021)
- Wang, S., Teng, Y., Perdikaris, P.: Understanding and mitigating gradient pathologies in physics-informed neural networks. In: arXiv, pp. 1–28. (2020)
- Watanabe, S., et al.: Hayabusa2 mission overview. *Space Sci. Rev.* **208**(1–4), 3–16 (2017). <https://doi.org/10.1007/s11214-017-0377-1>
- Werner, R., Scheeres, D.: Exterior gravitation of a polyhedron derived and compared with harmonic and mascon gravitation representations of Asteroid 4769 castalia. *Celest. Mech. Dyn. Astron.* **65**(3), 313–344 (1997). <https://doi.org/10.1007/BF00053511>
- Wittick, P.T., Russell, R.P.: Mascon models for small body gravity fields. *Adv. Astron. Sci.* **162**, 2003–2020 (2018)
- Wittick, P.T., Russell, R.P.: Mixed-model gravity representations for small celestial bodies using mascons and spherical harmonics. *Celest. Mech. Dyn. Astron.* **131**(7), 31–317 (2019). <https://doi.org/10.1007/s10569-019-9904-6>
- Xu, P.: The value of minimum norm estimation of geopotential fields. *Geophys. J. Int.* **111**(1), 170–178 (1992). <https://doi.org/10.1111/j.1365-246X.1992.tb00563.x>
- Xu, P.: Truncated SVD methods for discrete linear Ill-Posed Prob- lems. *Geophys. J. Int.* **135**(2), 505–514 (1998). <https://doi.org/10.1046/j.1365-246X.1998.00652.x>
- Xu, P., Fukuda, Y., Liu, Y.: Multiple parameter regularization: numerical solutions and applications to the determination of geopotential from precise satellite orbits. *J. Geodesy* **80**(1), 17–27 (2006). <https://doi.org/10.1007/s00190-006-0025-0>
- Xu, P., Rummel, R.: Generalized ridge regression with applications in determination of geopotential fields. *Manuscr. Geoda* pp. 8–20 (1994) issn: 0340-8825

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.