

# Machine learning for prediction of dynamical clustering in granular gases

Sai Preetham Sata<sup>1,2,4,\*</sup>, Ralf Stannarius<sup>3,4</sup>, and Dmitry Puzyrev<sup>2,4,\*\*</sup>

<sup>1</sup>IKS, Faculty of Informatics, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany

<sup>2</sup>MTRM, Medical Faculty, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany

<sup>3</sup>Dept. of Engineering, Brandenburg University of Appl. Sciences, Magdeburger Str. 50, 14770 Brandenburg an der Havel, Germany

<sup>4</sup>MARS, Otto von Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany

**Abstract.** Continuously excited dense granular gases in microgravity can develop spatial inhomogeneities of the particle distribution. Dynamical clustering is a phenomenon where a significant share of particles concentrate in strongly overpopulated regions. It is caused by a complex interplay between the energy influx and dissipation in collisions. The overall packing fraction, container geometry, and excitation parameters influence the gas-cluster transition. We perform Discrete Element Method (DEM) simulations for frictional spheres in a cuboid container and apply statistical criteria to the packing fraction profiles. Machine learning (ML) methods are used to study the dependence of the gas-cluster transition on system parameters. It is a promising alternative to predict the state of the system without the need for the time-consuming DEM simulations. We identify the best models for predicting the dynamical clustering of frictional spheres in a specific experimental geometry.

## 1 Introduction

Granular gases exhibit fascinating phenomena when subjected to mechanical excitation in a confined volume. Continuous energy supply is needed to maintain dynamic states due to the dissipative character of the particle collisions [1]. Transitions from a homogeneous gas to a condensed cluster state have been described [2–5]. In clusters, particles have frequent contacts and the dynamics resembles that of a liquid. Three-dimensional (3D) granular gases can be realized experimentally in microgravity. Our reference system is the VIP-Gran experiment [1] of the SPACE GRAINS [6] topical team, funded by ESA. It is designed for a planned ISS flight and was used in parabolic flights, as reported by Noirhomme et al. [4]. The authors described the gas-cluster transition, the main focus of the present study. They established a full state diagram for a range of system parameters. Another parametric study by Wu et al. [7] was based on numerical simulations.

The transition from the gaseous to a clustered state depends on the particle number density and on the excitation parameters. Often, a simple visual inspection gives evidence of clustering, but statistical criteria are desirable. The Kolmogorov–Smirnov (KS) test and a caging criterion were proposed [3]. The purpose of our study is to create datasets which allow one to train and test various ML algorithms to predict the gas-cluster transition without the need of time-consuming and computationally intensive simulations. With a limited data set, ML algorithms are trained and validated to identify the most appropriate one. Esti-

ated transition thresholds help assess the applicability of each test criterion when system parameters are varied.

## 2 Methods

The VIP-Gran setup consists of a cuboid container of  $L_y \times L_z \times L_x = 30 \times 30 \times 40 \text{ mm}^3$  filled with particles that can be excited by two opposing vibrating walls. Clustering occurs near the mid-plane of the box. We use this setup filled with an ensemble of frictional spheres in our simulations. The overall packing fraction is defined as  $\Phi = (N\pi d^3)/(6L_x L_y L_z)$ , where  $N$  is the number of particles inside the container and  $d$  is their diameter.

**Table 1.** Constant parameters in simulation software

Symbol	Parameter Name	Value
$\rho$	Particle density	8730 kg/m <sup>3</sup>
$d$	Particle diameter	1 mm
$m$	Particle mass	4.57 mg
$\nu$	Poisson's ratio	0.3
$\omega$	Frequency of the oscillation	15 Hz
$\mu$	Coefficient of friction	0.4
$\varepsilon$	Coefficient of restitution	0.85

The left and right wall positions  $x_1(t)$  and  $x_2(t)$  are

$$x_1(t) = -\frac{L_x}{2} + A \sin(\psi(t)), x_2(t) = \frac{L_x}{2} + A \sin(\psi(t) + \theta), \quad (1)$$

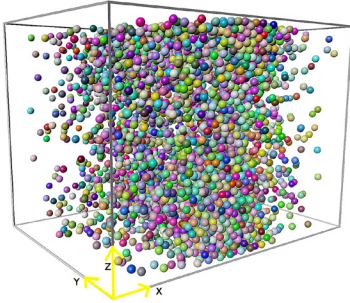
where  $A$  is the excitation amplitude, and  $\theta$  is a phase-shift. The vibration phase is  $\psi(t) = 2\pi Ft$ , with frequency  $F$  and time  $t$ . The parameters that are kept constant for all simulations in our study are listed in Tab. 1. The ranges of parameters that were varied are listed in Tab. 2.

\*e-mail: [sai.sata@ovgu.de](mailto:sai.sata@ovgu.de)

\*\*e-mail: [dmitry.puzyrev@ovgu.de](mailto:dmitry.puzyrev@ovgu.de)

**Table 2.** Variable parameters in simulation software

Symbol	Parameter Name	Range
$\Phi$	Filling fraction	2.6%, 8.2%
$A$	Amplitude	[2 mm, 5 mm]
$\theta$	Phase-shift	$[0, \pi]$



**Figure 1.** Snapshot of a CPU-GPU DEM simulation [8] of a dense granular gas of frictional spheres, rendered with Blender 3D software. The fill fraction is  $\Phi = 5.9\%$ . The pistons move along  $x$ . The other walls are fixed. The origin of the coordinate system is the center of the container. All the particles are identical, colors were assigned for visualization purpose.

The numerical simulation is performed with a hybrid CPU-GPU DEM code, kindly provided by Raúl Cruz Hidalgo [8]. We used the Hertz-Mindlin contact model [13] to describe inter-particle forces during collisions. A typical snapshot of the output is shown in Fig. 1. The normal restitution coefficient was set to 0.85 and friction coefficient to 0.4 [13]. Normal and tangential damping constants were set to 0.0943.

The KS-test criterion quantifies the distance between the measured cumulative distribution  $f(x)$  of the spheres to the ideal homogeneous distribution  $u(x)$ . The supremum distance between these distributions is

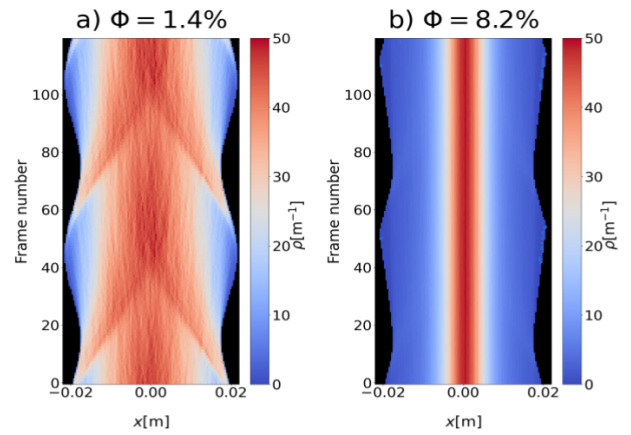
$$D_{\max} = \sup_{x_1 < x < x_2} |f(x) - u(x)|. \quad (2)$$

The cumulative distribution function  $f(x)$  is the integral of the probability density function  $P(x)$

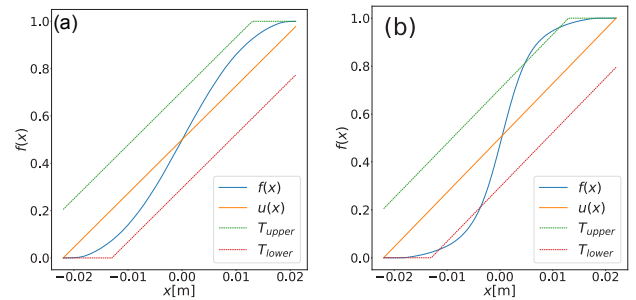
$$f(x) = \int_{x_1}^x P(x) dx.$$

$P(x)$  is obtained from the normalized actual particle number density integrated over  $y$  and  $z$ . We use the minimum difference  $\delta_{\min}$  between  $D_{\max}$  and a threshold value for a given set of system parameters  $\{N, A, \theta\}$  as one criterion for the gas-cluster transition.

With the caging criterion, clusters are identified by a caging effect that is predominant for large  $N$  [3]. A particle is considered caged when the local packing fraction defined as the ratio of the grain volume to the volume of its Voronoi cell (with the particle positions as generators) is greater than a given threshold  $T_{\text{caged}}$ , which we set to 0.285 following Ref. [3]. Based on this condition, the maximum fraction of caged particles  $\varphi_{\text{caged}}$  is evaluated to estimate the number of caged particles with respect to  $N$ .



**Figure 2.** Spatio-temporal profiles of the normalized number density for 120 frames (2 periods). The black region corresponds to regions with no particles, including the restricted area behind the piston walls. Red and blue colors symbolize higher and lower particle number densities integrated over  $y-z$  plane. The plots (a) and (b) corresponds to low filling fraction  $\Phi = 1.4\%$  ( $N = 1000$  particles) and  $8.2\%$  ( $N = 5643$  particles), respectively. Excitation parameters are  $A = 2$  mm and  $\theta = \pi$ .



**Figure 3.** Momentary cumulative distribution at phase  $\psi = 3\pi/2$  and the upper and lower limits of the KS criterion for  $\Phi = 1.4\%$  (a) and  $\Phi = 8.2\%$  (b),  $\theta = \pi$  and  $A = 2$  mm.

ML approaches have been used in various applications related to granular matter, including particle tracking in granular gases [9]. We employ ML to predict the gas-cluster transition boundary (shown as black dashed line in Fig 4 (a) and (b)) which depends on the system parameters  $N, A$ , and  $\theta$ . This can be achieved with the help of the estimates of  $\delta_{\min}$  and  $\varphi_{\text{caged}}$ . Datasets are prepared with fixed ranges of system parameters  $N, A, \theta$  using both criteria and ML algorithms are trained and evaluated on these datasets with the help of standard regression metrics such as the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE) and the coefficient of determination ( $R^2$ ). See [11] for the corresponding details.

### 3 Results

#### 3.1 Spacetime plots for number density profiles and cumulative distribution profiles

At antiphase excitation ( $\theta = \pi$ ), the distances between the walls are symmetric with respect to the center of the container for all  $\psi(t)$ . The piston walls move according to Eq. (1), thus the bins vary dynamically with the piston positions. Minimum and maximum limits of the KS-test vary accordingly. The frame numbers  $i_F$  are related to  $\psi(t)$  by  $i_F = (30/\pi) \cdot \psi(t)$ . The particle number density profiles for excitation parameters  $A = 2$  mm and  $\theta = \pi$  are shown in Fig. 2. These profiles were obtained by collecting the particles based on their  $x$  coordinate into specified bins along the  $X$  direction in every frame. The normalized density profiles with the same phases  $\psi(t)$  modulo  $2\pi$  are combined into space-time data arrays. These normalized profiles are used to evaluate  $\delta_{\min}$  in the KS-test.

Figure 3 shows the cumulative distribution profile of the particle positions at phase  $\psi(t) = 3\pi/2$  ( $i_F = 45$ ) as the blue curve. The uniform distribution  $u(x)$  is shown as orange line. The red and green lines show the lower and upper thresholds relative to  $u(x)$ . At a filling fraction  $\Phi = 1.4\%$  ( $N = 1000$  particles) (a), the cumulative distribution is within the threshold bounds and hence the system is in the gaseous state. As  $\Phi$  is increased to  $8.2\%$  ( $N = 5643$  particles) (b), the cumulative distribution profile crosses the threshold bounds and hence reach cluster state according to KS-test criterion.

#### 3.2 Effect of system parameters on the gas-cluster transition using the KS-test and caging criteria

Here, details of the dependence of the minimum difference  $\delta_{\min} = \min(\delta)$  and maximum fraction of caged particles  $\varphi_{\text{caged}}$  on the system parameters ( $A$ ,  $N$  and  $\theta$ ) are discussed. The ensemble is considered in the gaseous state by the KS test if the estimated  $\delta_{\min} > 0$ , and in the cluster state if  $\delta_{\min} \leq 0$ . Similarly, if  $0 \leq \varphi_{\text{caged}} < 0.05$ , the system is considered in the gaseous state by the caging test, and if  $\varphi_{\text{caged}} \geq 0.05$ , it is considered in the cluster state. The gas-cluster transition is defined as the transition boundary from the gaseous to the cluster regime as determined by  $\delta_{\min}$  or  $\varphi_{\text{caged}}$  [11]. Figure 4(a) shows the gaseous and cluster regimes for a given set of system parameters at phase-shift  $\theta = \pi$  using the KS-test criterion. The red region in Fig. 4(a) indicates that the system is permanently in the gaseous state, whereas the blue region indicates that the system becomes clustered at least for some time interval. The black dashed line sketches the gas-cluster transition.

At antiphase excitation ( $\theta = \pi$ ), when  $N$  is between 1834 ( $\Phi = 2.6\%$ ) and 3700 ( $\Phi = 5.3\%$ ) particles and amplitudes  $A$  are between 2 mm and 5 mm,  $\delta_{\min}$  becomes positive in some region, which indicates that the system is in the gaseous state there. When  $N$  is above 3700 ( $\Phi = 5.3\%$ ) particles,  $\delta_{\min}$  is always negative for a similar range of  $A$ . The system reaches the cluster state at least temporarily. Figure 4(b) shows the effect of the parameters  $A$ ,  $N$  for  $\theta = \pi$  on the maximum fraction of caged particles ( $\varphi_{\text{caged}}$ ).

The black line in Fig. 4(b) indicates where the gas-cluster transition occurs with 5 % significance level. The region above the line is the clustered regime, at least temporarily, during the cycles. There are practically no caged particles for  $\Phi$  less than approximately 4 %, indicating a gaseous state for all excitation parameters. As  $N$  increases above 3000 ( $\Phi = 4.3\%$ ), the number of caged particles increases and the clustered state is entered. In order to obtain reliable ML predictions, 385 simulations were performed for different parameter configurations within the given range as stated in Tab. 2. A striking difference to the results of the KS-test is that the amplitude  $A$  does hardly influence the clustering threshold. There is a slight tendency of lowering thresholds with increasing amplitude, but compared to the KS-test results, it is rather marginal.

Another substantial difference between the results of KS and caging tests is the dependence of the transition threshold on the phase-shift  $\theta$ . The oscillating boundaries influence the outcome of the KS-test and led to remarkable differences of thresholds predicted for inphase and antiphase excitation. At the same time, the caging test yields a much more stable and reliable results because the density and size of the particle cluster itself remain largely similar in the entire range of  $\theta$ . More detailed information on the difference between the two clustering criteria is provided elsewhere [11].

#### 3.3 Prediction of transition parameters using machine learning

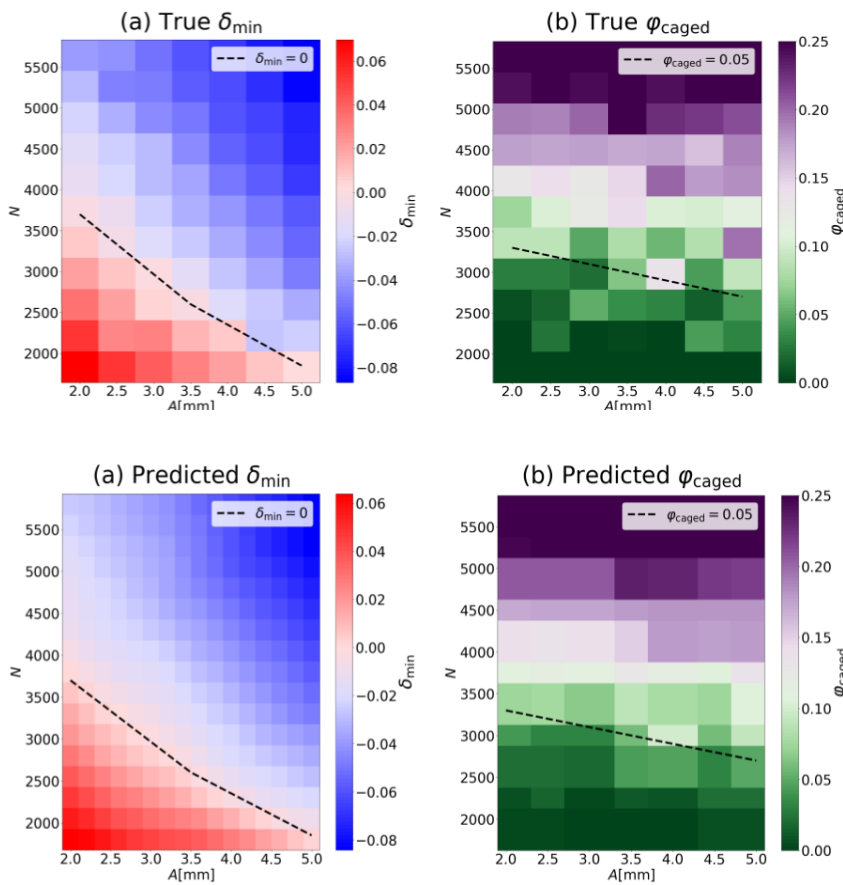
In this section, we show the predictions of two ML models that were trained and that showed the best evaluation results on the regression metrics mentioned in Sec. 2.

The plot in Fig. 5(a) shows exemplarily  $\delta_{\min}$  predicted from Neural Network (NN) regression when  $\theta = \pi$  (cf. Fig. 4(a) for the numerically calculated “true”  $\delta_{\min}$ ). The transition from the gaseous to the cluster regime at low amplitude,  $A = 2$  mm, is found near  $N \approx 3700$  particles ( $\Phi = 4.3\%$ ), and it lowers with increasing excitation strength to  $N \approx 1800$  ( $\Phi = 2.6\%$ ) at  $A = 5$  mm. The neural network regression model is able to predict the results of the KS-test with relatively good accuracy as the transition limits closely resemble the numerically computed data.

Similarly, the plots in Figs. 5(b) and 4(b) compare the predicted (by XGBoost regression) and “true”  $\varphi_{\text{caged}}$ . The particle number density for the gas-cluster transition predicted by the caging test is slightly decreasing with larger  $A$ , but much less pronounced than predicted by the KS-test. The XGBoost regression model thus accurately reproduces the results of the caging test for  $\theta = \pi$ . Additional details regarding ML based predictions of clustering, such as types of datasets or hyperparameters used in the algorithms, will be explored in greater detail elsewhere [12].

### 4 Conclusion

We have simulated permanently excited granular ensembles in the geometry of the VIP-Gran setup to obtain a



**Figure 4.** Minimum difference  $\delta_{\min}$  and  $\phi_{\text{caged}}$  evaluated with KS and caging tests. The plots corresponds to  $\delta_{\min}$  (a) and  $\phi_{\text{caged}}$  (b) at  $\theta = \pi$  respectively. The blue and purple regions in (a) and (b) indicate parameter ranges where clustering is detected and dashed line visualizes the onset.

**Figure 5.** Predicted  $\delta_{\min}$  and  $\phi_{\text{caged}}$  by NN and XGBoost regressions respectively for the same range of system parameters  $N, A$  as in Fig. 4. NN architecture contains three hidden layers with 16, 32, 16 neurons in each layer. ReLU activation is used. NN is trained for 100 epochs with batch size and learning rate set to 8 and 0.0001 respectively.

data basis for the evaluation of the gas-cluster transition and its dependence on system parameters. Then, we employed two established methods to determine the transition thresholds for a set of system parameters. Furthermore, we investigate ML approaches to predict the transition parameters of the original tests in order to reduce the complex and time-consuming numerical simulation and evaluation steps. The ML algorithms help to find the thresholds in a more efficient way but of course they cannot confirm or disprove the validity of the two evaluation criteria.

## Acknowledgments

The authors acknowledge funding by DLR under grants 50WM2252 and 50WK2348. We cordially thank Prof. Raúl Cruz Hidalgo for providing the simulation software.

## References

[1] Aumaitre, Sébastien, R. P. Behringer, A. Cazaubiel, E. Clément, et al., *Rev. Sci. Instr.* **89**, 075103 (2018).  
 [2] Falcon, Éric, Wunenburger, Régis, Évesque Pierre, Fauve, Stéphan, Chabot, Carole, Garrabos Yves, Beyens Daniel, *Phys. Rev. Lett.* **83**, 440–443 (1999).  
 [3] Opsomer, Eric, F. Ludewig, N. Vandewalle, *Europhys. Lett.* **99**, 4, 40001 (2012).

[4] Martial Noirhomme, Annette Cazaubiel, Alexis Daras, et al. *Europhys. Lett.* **123**, 14003 (2018).  
 [5] Puzyrev, Dmitry, Raúl Cruz Hidalgo, David Fischer, Kirsten Harth, Torsten Trittel, Ralf Stannarius, *EPJ Web of Conf.* **249**, 04004, (2021),  
 [6] Space Grains Team <https://spacegrains.org/>  
 [7] Qi-Lin Wu, Mei-Ying Hou, Lei Yang, Wei Wang, Guang-Hui Yang, Ke-Wei Tao, Liang-Wen Chen, and Sheng Zhang, *Chin. Phys. B*, **29**, 054502 (2020)  
 [8] R.C. Hidalgo, T. Kanzaki, F. Alonso-Marroquin and S. Luding, *AIP Conference Proceedings* **1542**, 169-172 (2013).  
 [9] Puzyrev, Dmitry, Kirsten Harth, Torsten Trittel, Ralf Stannarius, *Microgravity Sci. Technol.*, **32**, 897–906 (2020)  
 [10] Statistical significance: p value, 0.05 threshold, and applications to radiomics - reasons for a conservative approach. <https://rdcu.be/dv7LL>  
 [11] Sai Preetham Sata , Ralf Stannarius, Dmitry Puzyrev, Preprint available at (2025) <https://doi.org/10.21203/rs.3.rs-6243611/v1>  
 [12] Sai Preetham Sata , et al., Machine learning based prediction of dynamical clustering in granular gases (Unpublished)  
 [13] Wang, Jing, Statistical dynamics of soft low-friction grains, PhD Thesis, Otto von Guericke University Magdeburg, 2023. <http://dx.doi.org/10.25673/111395>