Fast 3D Diffusion for Scalable Granular Media Synthesis and Homogenization Pipeline

M.-M. Hassan[1,4] , R. Cottereau[2], F. Gatti[3], P. Dec[4]

[1] *LMA (Université Aix-Marseille), Marseille, hassan@lma.cnrs-mrs.fr*
[2] *LMA (Université Aix-Marseille), Marseille, cottereau@lma.cnrs-mrs.fr*
[3] *LMPS (Université Paris Saclay), Gif sur Yvette, filippo.gatti@centralesupelec.fr*
[4] *Direction Innovation et Recherche, (SNCF), La Plaine Saint Denis, patryk.dec@sncf.fr*

Résumé — Discrete Element Method (DEM) simulations of granular media are very expensive, in particular during the sample initialization phase, which limits their use to short granular specimens. This paper proposes a data-driven 3D diffusion pipeline that learns packed voxel blocks from DEM databases and then stitches them into arbitrarily long assemblies. The method is demonstrated on railway ballast and lunar regolith simulants and yields speed-ups of more than two orders of magnitude with respect to DEM initialization. The generated assemblies are further exploited to run a large sleeper-track campaign, from which homogenized stress and strain fields are extracted and used to train a neural operator. The resulting surrogate provides fast predictions of continuum responses over long ballasted tracks while remaining statistically consistent with grain-scale simulations.
Mots clés — granular media, discrete element method, generative models.

1 Introduction

Granular materials such as railway ballast, sands or lunar regolith appear in many civil and geotechnical applications. Their mechanical behaviour is often studied with the Discrete Element Method (DEM) [1], which resolves grain kinematics and contact forces at very small time steps. This gives access to rich microstructural information, but the associated computational cost quickly becomes prohibitive when one wishes to simulate large domains and long loading histories.

In practice, the main bottleneck is the sample preparation phase. Starting from a random placement of grains, DEM simulations must let particles fall under gravity and dissipate kinetic energy until realistic densities are reached. For a single ballast sleeper setup this may already require hours of CPU time, even before the first wheel load is applied. For high-speed lines this limits simulations to short track segments and a few seconds of train passage. Similar constraints exist for lunar regolith simulants, where both Computed Tomography (CT) and DEM studies are typically restricted to millimetric samples [4, 5].

At the same time, generative diffusion models have emerged as powerful tools for learning complex high-dimensional distributions [2]. Rather than re-simulating the entire pouring process for each new sample, one can aim to learn directly the distribution of equilibrated packed states. The present work follows that idea and introduces a fast 3D diffusion pipeline tailored to granular media. Starting from DEM databases of equilibrated configurations, we learn to generate small voxel blocks and then assemble them into long volumes that can be segmented back into individual grains. On railway ballast and lunar regolith the method reproduces key statistics while dramatically reducing preparation time. The generated assemblies are further exploited in a discrete-to-continuum framework mentioned further down in the paper.

2 Diffusion-based generation of granular assemblies

2.1 Datasets and voxel blocks

The reference ballast dataset consists of DEM simulations of ballasted tracks performed with LMGC90. Each simulation models distance between two sleepers. Around 500 simulations with grains falling under gravity and compacting until a realistic density is reached provide a database of equilibrated final states.

For training, the DEM meshes are rasterised onto binary voxel grids after point in polygon operations, from which Representative Volume Elements (RVEs) are extracted. A similar procedure is applied to lunar regolith using synthetic grain shapes from the NU LHT 4M CT database [4, 5], packed into DEM cells with a target packing fraction of about 0.5.

2.2 Diffusion models and generation of long assemblies

The pipeline relies on two complementary diffusion models : one to generate individual voxel blocks, and one to stitch them into seamless assemblies.

The first model is a 3D denoising diffusion probabilistic model [2] with a UNet backbone adapted to volumetric data. During training, a forward process gradually corrupts equilibrated voxel blocks by adding Gaussian noise ; the network learns to reverse this corruption and thereby captures the distribution of packed configurations. At inference, new blocks are sampled in parallel by denoising from pure noise.

However, independently generated blocks do not match at their interfaces, preventing direct assembly. To address this, a second diffusion model performs 3D inpainting on the gaps between blocks. Generated blocks are placed with small overlaps along the desired direction ; the inpainting model receives the known voxel context from neighbouring blocks and reconstructs only the missing region, ensuring geometric continuity at the seams.

By repeating this stitching operation, hundreds of blocks can be joined into metre-scale voxel grids. A final watershed segmentation converts the volume into individual grain meshes compatible with DEM solvers. Figure 1 shows an example of a ballast layer obtained by this process.

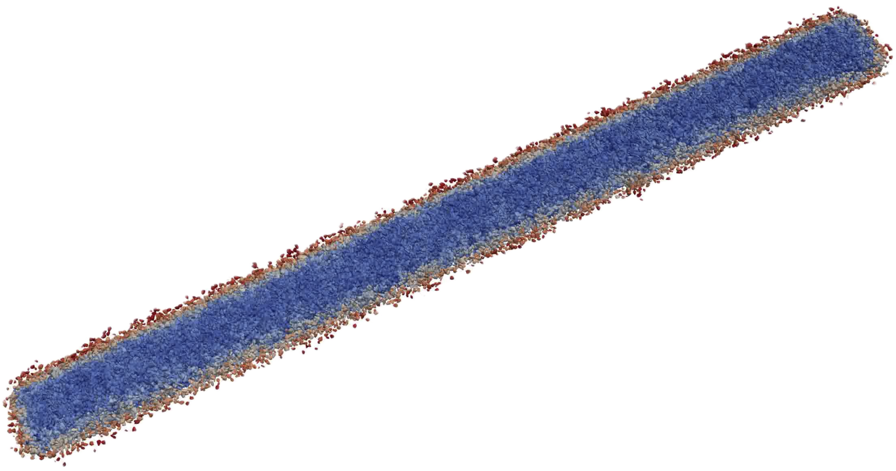

FIGURE 1 – Example of a ballasted track segment with 128k grains obtained by stitching hundreds of voxel blocks with the 3D inpainting model, then segmenting into individual grains.

3 Validation and computational cost

3.1 Statistical comparison with DEM

To assess physical realism, we compare diffusion-generated assemblies to reference DEM packings using standard macroscopic and microstructural indicators such as global void ratio, first- and second-order coordination numbers, fabric anisotropy and grain size distribution.

On ballast samples with about 10^4 grains, the void ratio of generated samples matches the DEM reference to within 10^{-4}. Mean coordination numbers differ only moderately (e.g. an increase of less than 0.2 contacts for first neighbours), leading to a small change in a simple stiffness proxy proportional to the square of the coordination number. The grain size distributions of DEM and generated assemblies are statistically indistinguishable as well. Similar conclusions hold for the lunar simulants, where the target packing fraction and size distribution are recovered by the generative model.

3.2 Speed-up over DEM initialization

The main practical benefit of the diffusion pipeline is the drastic reduction in sample preparation time. Figure 2 compares wall-clock times for traditional DEM initialization and for the proposed generative pipeline, as a function of the number of grains. DEM costs grow roughly linearly with the number of contacts and the number of time steps, whereas diffusion sampling remains almost constant once the models are trained. On the largest examples considered (around 2×10^5 grains), the proposed pipeline is more than two orders of magnitude faster than a full DEM initialization.

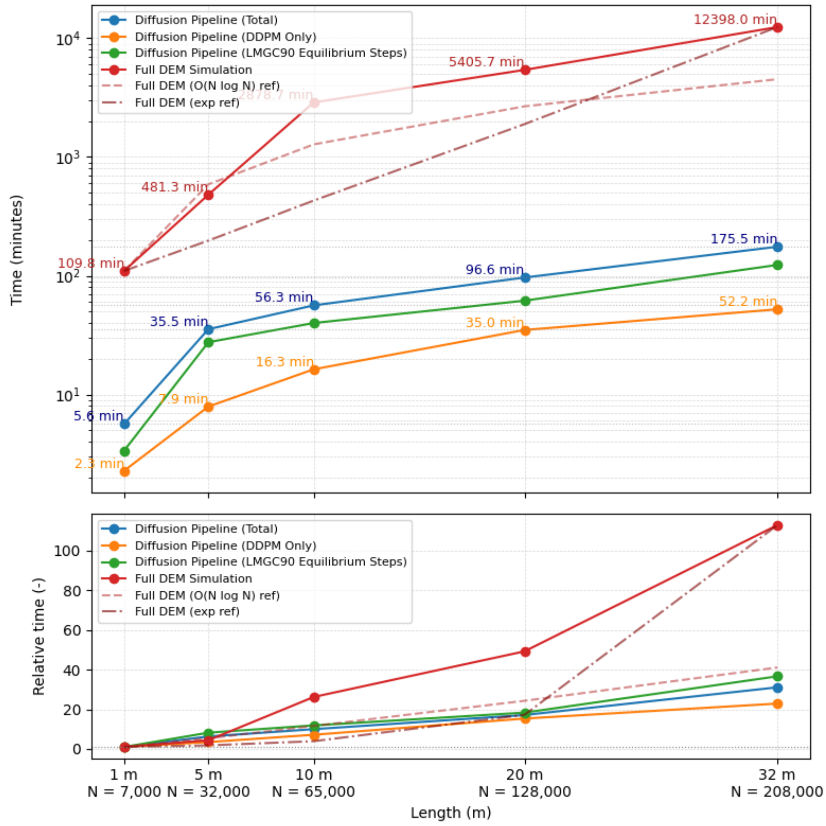

FIGURE 2 – Speed-up of the diffusion-based synthesis compared with DEM initialization for railway ballast and lunar regolith simulants.

4 Towards a continuum surrogate

Once generative initialization becomes cheap, it is possible to run DEM campaigns on long ballasted tracks that would otherwise be out of reach. Diffusion-generated sleeper setups are used to simulate several tens of track configurations under quasi-static sleeper loading. The domain is partitioned into a regular grid of bins that play the role of representative volume elements, and standard micro-macro relations are applied to compute homogenised stress and strain tensors [7, 8, 9].

The strain tensor is obtained using the method proposed by Bagi [8], based on the displacement gradient averaged over the space cells of the granular assembly :

$$\mathrm{d}\bar{e}_{ij} = \frac{1}{V} \sum_{c} \mathrm{d}\Delta u_j^c d_i^c, \tag{1}$$

where $\mathrm{d}\Delta u_j^c$ is the relative translation of particles at contact c and d_i^c is the complementary area vector. The microstructural strain tensor is then defined as the symmetric part :

$$\bar{\varepsilon}_{ij} = \tfrac{1}{2}\left(\mathrm{d}\bar{e}_{ij} + \mathrm{d}\bar{e}_{ji}\right). \tag{2}$$

The stress tensor is computed using the Love-Weber formula [9] :

$$\sigma_{ij} = \frac{1}{V} \sum_{c \in C} f_i^{(c)} \, l_j^{(c)}, \tag{3}$$

where $\underline{f}^{(c)}$ is the contact force vector and $\underline{l}^{(c)}$ is the branch vector connecting the centers of the two particles in contact. From these tensors, scalar invariants such as pressure, von Mises stress and volumetric strain are derived.

The resulting continuum fields form a database on which a neural operator is trained to emulate the DEM response at the track scale. The operator takes as input the discretised configuration and loading conditions, and outputs the evolution of continuum fields on the bin grid. Coupled with the diffusion-based initialization, this yields a two-stage surrogate : diffusion provides realistic packed assemblies at negligible cost, and the neural operator delivers fast continuum predictions suitable for parametric studies or uncertainty analyses.

In summary, the present paper presents a 3D diffusion pipeline for fast synthesis of realistic and *infinitely* long granular assemblies. On railway ballast and lunar simulants, the method reproduces key DEM statistics while reducing sample preparation time by more than two orders of magnitude. With the ability to generate arbitrarily long assemblies, and many in parallel, the pipeline unlocks the possibility of running large-scale DEM campaigns on long ballasted tracks and generating large databases of continuum fields for machine learning.

Références

[1] P. A. Cundall, O. D. L. Strack. A discrete numerical model for granular assemblies. Géotechnique, 29(1), 1979.

[2] J. Ho, A. Jain, P. Abbeel. Denoising diffusion probabilistic models. Neural Information Processing Systems, 2020.

[3] F. Meyer, S. Beucher. Morphological segmentation. Journal of Visual Communication and Image Representation, 1(1), 1990.

[4] O. Kafka, N. Moser, A. Chiaramonti. Three dimensional shape and size X ray CT database for lunar regolith simulants. NIST data set, 2021.

[5] S. Baidya, M. Melius, A. M. Hassan, A. Sharits, A. N. Chiaramonti, T. Lafarge, J. D. Goguen, E. J. Garboczi. Optical scattering characteristics of 3-D lunar regolith particles measured using X-ray nano computed tomography. IEEE Geoscience and Remote Sensing Letters, 2022.

[6] L. de Abreu Corrêa, J. C. Quezada, R. Cottereau, S. S. Costa D'Aguiar, C. Voivret. Randomly-fluctuating heterogeneous continuum model of a ballasted railway track. Computational Mechanics, 60(5), 2017.

[7] K. Bagi. Stress and strain in granular assemblies. Mechanics of Materials, 22(3) :165–177, 1996.

[8] K. Bagi. Analysis of microstructural strain tensors for granular assemblies. Int. J. Solids and Structures, 43(10) :3166–3184, 2006.

[9] J. Christoffersen, M. M. Mehrabadi, S. Nemat-Nasser. A micromechanical description of granular material behavior. J. Appl. Mech., 48(2) :339–344, 1981.