An iterative Bayesian filtering framework for fast and automated calibration of DEM models

Hongyang Cheng[a,*], Takayuki Shuku[b], Klaus Thoeni[c], Pamela Tempone[d], Stefan Luding[a], Vanessa Magnanimo[a]

[a] *Multi-Scale Mechanics (MSM), Faculty of Engineering Technology, MESA+, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[b] *Graduate School of Environmental and Life Science, Okayama University, 3-1-1 Tsushima naka, Kita-ku, Okayama 700-8530, Japan*
[c] *Centre for Geotechnical Science and Engineering, The University of Newcastle, Callaghan, NSW 2308, Australia*
[d] *Division of Exploration and Production, Eni SpA, Milano, Lombardy, Italy*

Highlights

- Fast generation of DEM packings through 3D feature-based image processing.
- An iterative Bayesian approach for a coarse-to-fine search in DEM parameter space.
- Posterior predictions are passed consistently over iterations via variational Gaussian mixtures.
- Interparticle stiffness and friction require less iterations to identify than rolling parameters.
- Increasing interparticle friction leads to larger macroscopic friction and smaller mean effective pressure during oedometric compression, whereas increasing stiffness has the opposite effect.

Abstract

The nonlinear, history-dependent macroscopic behavior of a granular material is rooted in the micromechanics between constituent particles and irreversible, plastic deformations reflected by changes in the microstructure. The discrete element method (DEM) can predict the evolution of the microstructure resulting from interparticle interactions. However, microme-chanical parameters at contact and particle levels are generally unknown because of the diversity of granular materials with respect to their surfaces, shapes, disorder and anisotropy.

The proposed iterative Bayesian filter consists in recursively updating the posterior distribution of model parameters and iterating the process with new samples drawn from a proposal density in highly probable parameter spaces. Over iterations the proposal density is progressively localized near the posterior modes, which allows automated zooming towards optimal solutions. The Dirichlet process Gaussian mixture is trained with sparse and high dimensional data from the previous iteration to update the proposal density.

As an example, the probability distribution of the micromechanical parameters is estimated, conditioning on the experimen-tally measured stress–strain behavior of a granular assembly. Four micromechanical parameters, i.e., contact-level Young's modulus, interparticle friction, rolling stiffness and rolling friction, are chosen as strongly relevant for the macroscopic behavior. The *a priori* particle configuration is obtained from 3D X-ray computed tomography images. The *a posteriori*

* Corresponding author.
  E-mail addresses: h.cheng@utwente.nl (H. Cheng), shuku@cc.okayama-u.ac.jp (T. Shuku), klaus.thoeni@newcastle.edu.au (K. Thoeni), pamela.tempone@eni.com (P. Tempone), s.luding@utwente.nl (S. Luding), v.magnanimo@utwente.nl (V. Magnanimo).

expectation of each micromechanical parameter converges within four iterations, leading to an excellent agreement between the experimental data and the numerical predictions. As new result, the proposed framework provides a deeper understanding of the correlations among micromechanical parameters and between the micro- and macro-parameters/quantities of interest, including their uncertainties. Therefore, the iterative Bayesian filtering framework has a great potential for quantifying parameter uncertainties and their propagation across various scales in granular materials.

1. Introduction

The Discrete Element Method (DEM) captures the collective behavior of a granular material by tracking the kinematics of the constituent grains [1]. DEM (coupled with other methods) can provide comprehensive cross-scale insights [2–7] that are difficult to obtain in either state-of-the-art experiments or sophisticated continuum models. DEM models allow to reproduce the macroscopic behavior of granular materials, starting from contact laws and particle characteristics. However, these microscopic quantities are not always directly measurable. Thus, the identification of micromechanical parameters has to be treated as an inverse problem.

Recently, it has become possible to measure micromechanical properties via ad-hoc experiments [8,9]. However, (i) often micro-properties greatly vary among grains of the same specimen and the average values are typically considered; (ii) for most industrial applications, fictitious non-physical parameters, e.g., rolling stiffnesses [10–12] are used in conjunction with physically based contact laws to match with experimental data. To tackle the issues mentioned above, it is often preferable to infer the micromechanical parameters from the history-dependent macroscopic measurements, following an "inverse-modeling calibration" approach [13].

Among a handful of optimization methods for the calibration of DEM models [14,15], the design of experiments (DOE) is efficient in searching for possible solutions with a manageable size of DEM model evaluations [16–20]. The efficiency of the DOE depends on orthogonal arrays constructed according to combinatorial theory [21,22]. The DOE helps in selecting a suitable number of *representative* trial sets of parameters and thus requires the knowledge of the interdependence between the parameters before conducting numerical "experiments". Moreover, the optimization methods generally make use of the characteristic bulk properties (e.g., Young's modulus, peak- and critical-state macroscopic friction), which could lead to local rather than global optima. Without the full stress–strain curves, the dependency of granular materials on stress and fabric history cannot be accounted for as well.

The Bayesian framework provides a promising and rigorous basis for solving the above-mentioned inverse problem [23,24]. Recently, Hadjidoukas et al. [25] employed the Transitional Markov Chain Monte Carlo (TMCMC) algorithm to sample the posterior distribution of microscopic parameters and quantify the associated uncertainties for DEM simulations of granular materials. A versatile computational framework Π4U was further developed for Bayesian uncertainty quantification and propagation of complex physical models [26]. The framework also allows for Bayesian model selection, i.e., the evaluation of the robustness of physical models (at various scales) conditioned on experimental data. Recent interest in atomistic-level uncertainties has led to several attempts for Bayesian calibration and model selection in molecular dynamics (MD) simulations [27–29]. For example, Kulakova et al. [30] inferred the repulsion exponent of the Lennard-Jones potential in MD simulations from experimental data. Nevertheless, the above approaches may not be directly applicable to DEM simulations of dense granular materials because the history-dependency was not considered therein.

Compared with optimization methods, sampling posterior distributions seems to be better suited for complex DEM models because (i) multi-modal and complex-shaped posterior distributions can be sampled, without knowing the interdependence between parameters; (ii) uncertainties are directly linked to the macroscopic quantities of interest (QOIs) through the ensemble. To take into account history-dependent material responses in Bayesian calibration, the inverse problem can be formulated by recursively applying the Bayes' theorem, e.g., via sequential Monte Carlo (SMC) methods [31,32]. Apart from Bayesian parameter estimation for physical models [33–36], SMC methods have been applied in various research fields such as target tracking, robotics and fault detection [37–42].

A drawback of the conventional SMC methods is that a significant number of model evaluations are needed, especially for high dimensional parameter space. To improve computational efficiency, a new iterative Bayesian

filtering framework is developed. Within each iteration, the posterior distribution of model parameters, conditioned on experimental data, is recursively updated by the SMC filter. The key to iterative Bayesian filtering is the ability to resample from a proposal density, that is, the posterior distribution or modes obtained from the previous iteration, for the following iteration. Over iterations, the proposal density is progressively localized near the posterior modes. The resampling algorithm makes use of the Dirichlet process Gaussian mixture to nonparametrically construct the proposal density function from the existing samples and their weights that approximate the posterior distribution. The computational power is thus allocated to model evaluations which could lead to high posterior probabilities. The approximation converges after a sufficient number of iterations and only the posterior modes where the most probable parameters are located remain.

The present approach allows to infer micromechanical parameters of DEM models from macroscopic measurements in a fast and automated manner. The efficiency of the Bayesian approach is demonstrated by calibrating DEM simulations of glass beads under cyclic oedometric compression. The initial particle configuration is obtained from 3D feature-based image analysis. A uniform proposal density is initially adopted to sample parameter space from an initial guess. The predictive capability of the DEM model is evaluated through the statistics on the micromechanical and macroscopic quantities [43–46]. After three iterations, the uncertainties in all QOIs reduce to less than 5%. The micro–micro and micro–macro correlations are extracted from the landscape of the posterior distribution. To the authors' knowledge, this work is the first attempt to bridge DEM simulations and the history-dependent macroscopic behavior of granular materials within an iterative Bayesian framework.

The remainder of this paper is organized as follows. Section 2 briefly explains DEM modeling of granular materials. Section 3 introduces the fundamental components of the iterative Bayesian filtering framework, including the state space model, sequential Bayesian filtering, and the multi-level sampling algorithm. Section 4 showcases the capability of the iterative Bayesian filter in estimating the micromechanical parameters for DEM modeling of glass beads under oedometric compression. Section 6 discusses the (re)sampling algorithm which plays an essential role in the "coarse-to-fine search", computational efficiency involved with the implementation, and uncertainty quantification and propagation for multiscale simulations. Conclusions are drawn in Section 7.

2. DEM modeling

The open-source DEM package YADE [47] is employed to perform 3D DEM simulations of dense granular materials. DEM represents granular materials as packings of solid particles with simplified geometries and vanishingly small interparticle overlaps. Governed by springs, dashpots and sliders upon collision, the kinematics of the particles are updated within the explicit time integration scheme, based upon the net forces and moments resulting from interparticle forces. To mimic the effect of surface roughness, rolling/twisting stiffness is typically adopted in addition to stiffnesses in normal and tangential directions. Both interparticle tangential forces and contact moments are bounded by Coulomb type yield criteria (see Section 2.2 for details).

2.1. Packing generation

A common practice to initialize DEM simulations of dense granular materials is through "dynamic random packing generation" [48] which typically requires a significant amount of computational time. However, even if the same macroscopic properties such as stress and porosity are satisfied, randomly generated packings may behave very differently due to a subtle difference in their microstructures [49]. 3D X-ray imaging of granular materials offers a detailed morphological description of grains and their configuration relations [50,51]. Using this prior information, the packing generation stage can be greatly accelerated with improved accuracy. For this task, 3D feature-based imaging processing is applied here to extract the particle configuration from 3D X-ray computed tomography (3DXRCT) images with high accuracy (see Fig. 1 and Section 4.2 for details).

2.2. Contact laws

The dynamical behavior of a DEM model can be described by contact-level force–displacement/moment–rotation laws, such as those defined in Eqs. (1)–(5). The contact laws that describe interparticle interactions are briefly explained as follows. For two contacting spheres with a normal overlap $\mathbf{u}_n$, a relative tangential displacement $d\mathbf{u}_s$
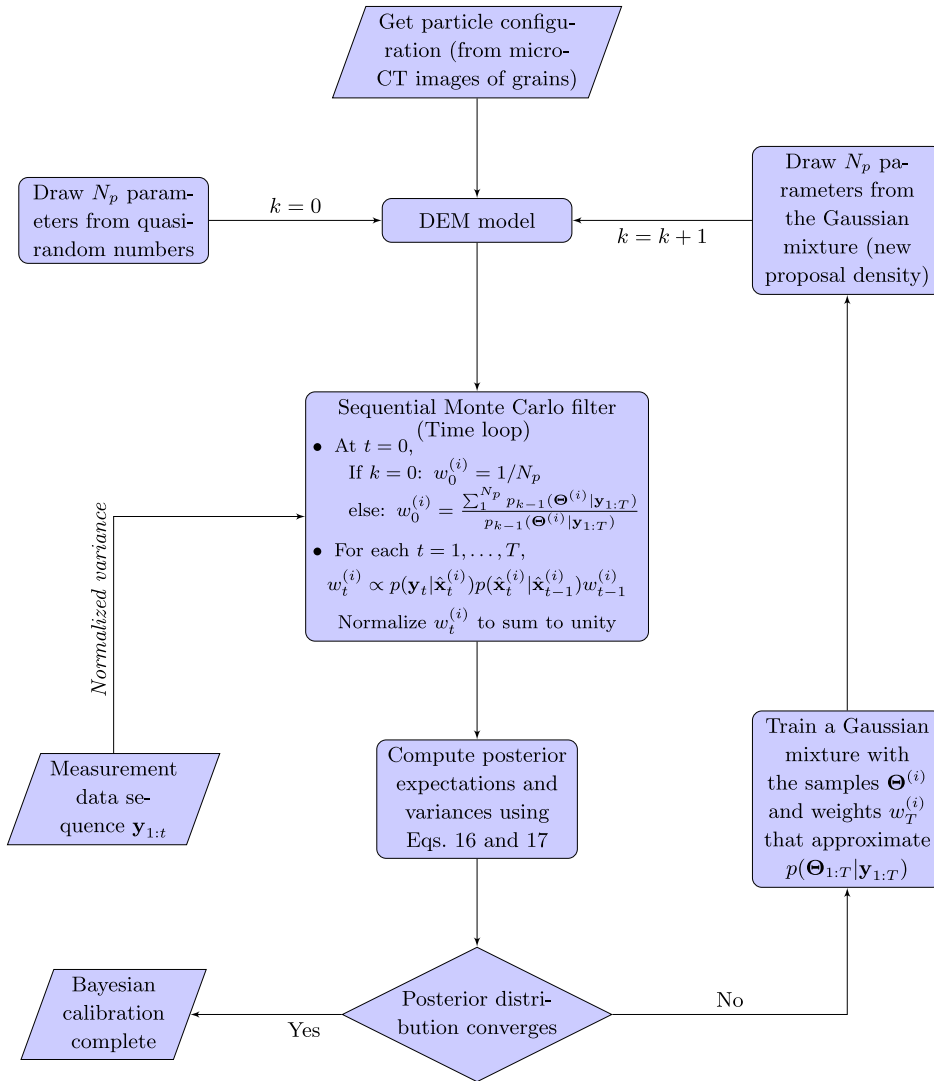
Fig. 1. Workflow of Bayesian calibration for DEM simulations of granular materials.

and a relative rotational angle $\boldsymbol{\theta}_c$ at the contact point, the interparticle normal and tangential forces $\mathbf{F}_n$ and $\mathrm{d}\mathbf{F}_s$, and contact moments $\mathbf{M}_c$ are calculated as

$$\mathbf{F}_n = \frac{2E_c \mathbf{u}_n}{3(1 - \nu_c^2)} \sqrt{r^* \|\mathbf{u}_n\|}, \tag{1}$$

$$\mathrm{d}\mathbf{F}_s = \frac{2E_c \, \mathrm{d}\mathbf{u}_s}{(1 + \nu_c)(2 - \nu_c)} \sqrt{r^* \|\mathbf{u}_n\|}, \tag{2}$$

$$\|\mathbf{F}_s\| \leq \tan\mu \|\mathbf{F}_n\|, \tag{3}$$

$$\mathbf{M}_c = k_m \boldsymbol{\theta}_c, \tag{4}$$

$$\|\mathbf{M}_c\| \leq \eta_m \|\mathbf{F}_n\|(r_1 + r_2)/2, \tag{5}$$

where E_c and ν_c are the contact-level Young's modulus and Poisson's ratio, μ is the interparticle friction, r^* is the equivalent radius defined as $1/(1/r_1 + 1/r_2)$, r_1 and r_2 are the radii of the two grains, k_m is the rolling stiffness, and η_m is the rolling friction which controls the plastic limit of $\mathbf{M}_c$.

The advantage of a DEM model is that it recovers the elastoplasticity of the granular material by capturing the micromechanics at the contact level, provided that the parameters (e.g., in Eqs. (1)–(5)) are properly calibrated. However, the micromechanical approach poses a great difficulty for Bayesian parameter estimation because it needs to deal with both material and geometrical nonlinearity and a nonlinear filtering problem is generally analytically intractable. In addition, the posterior distributions can have complex landscapes and evolutions, especially when fictitious parameters like k_m and η_m are used [36]. Hence, a new iterative Bayesian filtering framework is proposed to tackle the above challenges.

3. Iterative Bayesian parameter estimation

The posterior distribution of model states and parameters can be approximated by SMC methods [52]. To efficiently sample parameter space, a multi-level sampling algorithm is implemented. For the first iteration of Bayesian filtering ($k = 0$), the parameter values are uniformly sampled from quasi-random numbers. For the subsequent iterations ($k > 0$), new parameter values are drawn from the posterior distribution obtained from the previous iteration. The new Bayesian filtering framework allows to iteratively sample near potential posterior modes, with an increasing sample density over the iterations until the posterior expectations converge.

3.1. Nonlinear non-Gaussian state space model

A nonlinear, non-Gaussian state space model with unknown parameters $\mathbf{\Theta}$ can be written in the generic form [52]

$$\mathbf{x}_t = \mathbb{F}(\mathbf{x}_{t-1}, \mathbf{\Theta}_{t-1}) + \boldsymbol{v}_t, \tag{6}$$

$$\mathbf{y}_t = \mathbb{H}(\mathbf{x}_t, \mathbf{\Theta}_t) + \boldsymbol{\omega}_t, \tag{7}$$

$$\mathbf{\Theta}_t = \mathbf{\Theta}_{t-1}, \tag{8}$$

where $\mathbb{F}$ is the dynamical model for the states $\mathbf{x}_t$ which describe a hidden Markov process with the initial distribution $p(\mathbf{x}_0)$ and transition distribution $p(\mathbf{x}_t|\mathbf{x}_{t-1})$. If the measurements $\mathbf{y}_t$ are conditionally independent given $\mathbf{x}_t$, the measurement model $\mathbb{H}$ is reduced to the identity matrix $\mathbf{I}_d$ with d being the number of independent measurements. The unknown parameters $\mathbf{\Theta}$ are augmented as part of the model states, i.e., $\hat{\mathbf{x}}_t = (\mathbf{x}_t, \mathbf{\Theta}_t)$, which allows to use SMC methods for quantifying parameter uncertainty [52]. Note that the dynamical model in Eq. (8) indicates that the parameter space is sampled only at the initialization of the Bayesian filtering problem. Because $\mathbf{\Theta}$ is sampled only at the initialization, the transition distribution $p(\hat{\mathbf{x}}_t^{(i)}|\hat{\mathbf{x}}_{t-1}^{(i)})$ is fully deterministic. The modeling and measurement errors $\boldsymbol{v}_t$ and $\boldsymbol{\omega}_t$ are sampled from $N(0, \Sigma_t^M)$ and $N(0, \Sigma_t^D)$, respectively. The covariance matrices Σ_t^M and Σ_t^D which correlate the errors of the model states and the physical measurements are assumed to be diagonal.

In this work, the *DEM underlying* dynamical model $\mathbb{F}$ captures the evolution of the granular microstructure resulting from the interaction between grains (see Eqs. (1)–(5) and the parameters $\mathbf{\Theta} = \{E_c, \mu, k_m, \eta_m\}$). The irreversibility and history-dependence of granular materials are reflected by the mobilization of the microstructures due to interparticle sliding and/or rolling. The *macroscopic* model states $\mathbf{x}_t$ comparable to the physical measurements $\mathbf{y}_t$ are extracted from *microscopic* quantities such as interparticle forces and orientations. Both $\mathbf{x}_t$ and $\mathbf{y}_t$ include three independent variables, namely porosity n, mean stress p' and deviatoric stress q. The modeling error $\boldsymbol{v}_t$ is neglected for simplicity; the covariance matrix Σ_t^D of the measurement error is assumed to be $\sigma^2 \text{diag}(\mathbf{y}_t)$, with the normalized covariance σ being the only user-defined parameter. Note that other assumptions and correlation structures can be used for modeling the errors. However, a detailed study on these aspects is beyond the scope of the current work.

3.2. Bayesian Filtering

The aim of Bayesian filtering is to estimate the hidden augmented states $\hat{\mathbf{x}}_t = (\mathbf{x}_t, \mathbf{\Theta}_t)$ at time t given a sequence of measurements $\mathbf{y}_{1:t}$, namely the marginal **posterior** distribution $p(\hat{\mathbf{x}}_t|\mathbf{y}_{1:t})$ and the associated expectations of the form

$$\mathrm{E}[f_t(\hat{\mathbf{x}}_t)|\mathbf{y}_{1:t}] = \int f_t(\hat{\mathbf{x}}_t)p(\hat{\mathbf{x}}_t|\mathbf{y}_{1:t})\, d\hat{\mathbf{x}}_t, \tag{9}$$

where f_t is a $p(\hat{\mathbf{x}}_t|\mathbf{y}_{1:t})$-integrable function that describes an arbitrary quantity of interest as a function of the model states. Given the state space model (Eqs. (6)–(8)), we can write the full posterior distribution of the model parameters via Bayes' rule:

$$p(\mathbf{x}_{0:t}, \boldsymbol{\Theta}|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_{1:t}|\mathbf{x}_{0:t}, \boldsymbol{\Theta})p(\mathbf{x}_{0:t}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{p(\mathbf{y}_{1:t})}. \tag{10}$$

With the augmentation of the state space $\hat{\mathbf{x}}_t = (\mathbf{x}_t, \boldsymbol{\Theta}_t)$, Eq. (10) becomes

$$p(\hat{\mathbf{x}}_{0:t}|\mathbf{y}_{1:t}) \propto \prod_{t_i=1}^{t} p(\mathbf{y}_{t_i}|\hat{\mathbf{x}}_{t_i})p(\hat{\mathbf{x}}_{t_i}|\hat{\mathbf{x}}_{t_i-1})p(\hat{\mathbf{x}}_0). \tag{11}$$

For time $t > 0$, it is convenient to rewrite the posterior distribution in the recursive form

$$p(\hat{\mathbf{x}}_{0:t}|\mathbf{y}_{1:t}) \propto p(\mathbf{y}_t|\hat{\mathbf{x}}_t)p(\hat{\mathbf{x}}_t|\hat{\mathbf{x}}_{t-1})p(\hat{\mathbf{x}}_{1:t-1}|\mathbf{y}_{1:t-1}), \tag{12}$$

where the **likelihood** distribution $p(\mathbf{y}_t|\hat{\mathbf{x}}_t)$ and **transition** distribution $p(\hat{\mathbf{x}}_t|\hat{\mathbf{x}}_{t-1})$ can be easily evaluated by sampling from a given proposal density.

3.2.1. Bayesian sequential importance sampling

In sequential importance sampling (SIS), N_p samples, and thus N_p model evaluations, are drawn from a proposal density $\pi(\hat{\mathbf{x}}_{0:t}|\mathbf{y}_{1:t})$ which has a support no less than that of the target distribution, that is $p(\hat{\mathbf{x}}_{0:t}|\mathbf{y}_{1:t})$ in Eq. (11). To correct the approximation, each sample $\hat{\mathbf{x}}_{0:t}^{(i)}$, updated until time t, is associated with an importance weight $w_t^{(i)}$, that is

$$w_t^{(i)} = \frac{p(\hat{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{1:t})}{\pi(\hat{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{1:t})} = \prod_{t_i=1}^{t} p(\mathbf{y}_{t_i}|\hat{\mathbf{x}}_{t_i}^{(i)})p(\hat{\mathbf{x}}_{t_i}^{(i)}|\hat{\mathbf{x}}_{t_i-1}^{(i)})\frac{p(\hat{\mathbf{x}}_0^{(i)})}{\pi(\hat{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{1:t})}. \tag{13}$$

Because the dynamical model is history-dependent, the parameter values $\boldsymbol{\Theta}^{(i)}$ are sampled only at time $t = 0$. Given that $\mathbf{y}_0$ is typically known in a calibration, the augmented model states $(\mathbf{x}_0^{(i)}, \boldsymbol{\Theta}^{(i)})$ can be assembled with $\mathbf{x}_0^{(i)} = \mathbb{F}(\mathbf{x}_{ct}^p, \boldsymbol{\Theta}^{(i)})$ to exactly match $\mathbf{y}_0$ where $\mathbf{x}_{ct}^p$ represents particle positions and radii from 3DXRCT images. For time $t > 0$, the importance weights are recursively updated with the measurements $\mathbf{y}_t$:

$$w_t^{(i)} \propto p(\mathbf{y}_t|\hat{\mathbf{x}}_t^{(i)})p(\hat{\mathbf{x}}_t^{(i)}|\hat{\mathbf{x}}_{t-1}^{(i)})w_{t-1}^{(i)}. \tag{14}$$

With the covariance matrix of the measurement error $\boldsymbol{\Sigma}_t^D$, the likelihood $p(\mathbf{y}_t|\hat{\mathbf{x}}_t)$ can be evaluated by the multi-variate Gaussian distribution:

$$p(\mathbf{y}_t|\hat{\mathbf{x}}_t^{(i)}) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}_t^D|} \times$$
$$\exp\{-\frac{[\mathbf{y}_t - \mathbf{H}(\mathbf{x}_t^{(i)})]^T \boldsymbol{\Sigma}_t^{D^{-1}}[\mathbf{y}_t - \mathbf{H}(\mathbf{x}_t^{(i)})]}{2}\}, \tag{15}$$

where d is the dimension of the model state vector $\mathbf{x}_t^{(i)}$ and $\mathbf{H}$ is the measurement model in matrix form. The Monte Carlo approximations of the posterior expectations (see Eq. (9)) and variances are then computed as

$$\widehat{\mathrm{E}}[f_t(\hat{\mathbf{x}}_t)|\mathbf{y}_{1:t}] = \sum_{i=1}^{N_p} w_t^{(i)} f_t(\hat{\mathbf{x}}_t^{(i)}), \tag{16}$$

$$\widehat{\mathrm{Var}}[f_t(\hat{\mathbf{x}}_t)|\mathbf{y}_{1:t}] = \sum_{i=1}^{N_p} w_t^{(i)}(f_t(\hat{\mathbf{x}}_t^{(i)}) - \widehat{\mathrm{E}}[f_t(\hat{\mathbf{x}}_t)|\mathbf{y}_{1:t}])^2. \tag{17}$$

The posterior distribution represented by the empirical distribution formed with the augmented states and the associated weights is

$$p(\hat{\mathbf{x}}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N_p} w_t^{(i)}\delta(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^{(i)}), \tag{18}$$

with $\delta(\cdot)$ the Dirac delta function. Note that the importance weights $\{w_t^{(i)}; i = 1, \dots, N_p\}$ are normalized to sum to unity at each time step.

3.3. An efficient multi-level sampling algorithm

3.3.1. Initial sampling using quasi-random numbers

Because the posterior distribution of the model parameters is unknown, a non-informative proposal density is used to uniformly explore the parameter space in the *first* iteration ($k = 0$), namely,

$$\Theta^{(i)} \sim \text{U}(\Theta) \quad \text{and} \quad w_0^{(i)} = \frac{1}{N_p}. \tag{19}$$

The initial sampling is done with quasi-random numbers so as to identify "all" possible posterior modes. In fact, the use of deterministic quasi-random numbers leads to the so-called sequential quasi-Monte Carlo (SQMC) filter [53], which is recently proved to have better accuracy and smaller error rate than the SMC filter which uses pseudo-random numbers [54].

With the samples drawn from quasi-random numbers only at time $t = 0$, the evaluation of the importance weights with Eq. (14) becomes simple. However, the initial uniform sampling is inefficient because no experimental measurement is involved in choosing the proposal density. Resampling during the measurement updates is also not feasible because each sample trajectory has to be kept intact until the last time step T, for nonlinear, history-dependent dynamical models [55]. Since the randomness enters the augmented state space only through the parameters at time $t = 0$ (i.e., self-assembling state space in [56]), the SMC methods are ensured to degenerate namely, all but a few importance weights become negligible, as time approaches infinity. An efficient way to circumvent weight degeneracy is to reinitialize the samples and weights, and solve the same Bayesian filtering problem again with a more sensible proposal density, as described below.

3.3.2. Iterative resampling using the estimated posterior as proposal density

As can be seen in Eq. (13), the proposal density should already contain knowledge updated with the measurement data. Ideally, one would directly sample from the posterior distribution $p(\hat{\mathbf{x}}_{0:t}|\mathbf{y}_{1:t})$ or at least its approximation. Therefore, when repeating Bayesian filtering ($k > 0$), the posterior distribution of the parameters obtained at the end of the previous iteration $p_{k-1}(\Theta|\mathbf{y}_{1:T})$ is adopted as the proposal density. Thus, the model states propagating from $t = 0$ to T depend only on the randomness reintroduced in $\Theta^{(i)}$. Similar to Section 3.3.1, new parameter values are sampled from this new proposal density only at time $t = 0$ via,

$$\Theta^{(i)} \sim p_{k-1}(\Theta|\mathbf{y}_{1:T}) \quad \text{and} \quad w_0^{(i)} = \frac{\sum_1^{N_p} p_{k-1}(\Theta^{(i)}|\mathbf{y}_{1:T})}{p_{k-1}(\Theta^{(i)}|\mathbf{y}_{1:T})}. \tag{20}$$

As time loops from 1 to T, a new set of sample trajectories $\{\hat{\mathbf{x}}_t^{(i)}; t = 1, \ldots, T\}$ that evolve deterministically according to Eqs. (1)–(5) are generated. The measurement sequence $\mathbf{y}$ is reversed in time for the *odd-numbered* iterations, i.e., $\mathbf{Y} = \{\mathbf{y}_{1:T}, \mathbf{y}_{T:1}, \mathbf{y}_{1:T}, \ldots\}$, in order to ensure global continuity in the measurement updates, meaning that the resampling for the kth iteration takes place at the most adjacent time when the proposal density is estimated in the $(k-1)$th iteration.

For each iteration with $k > 0$, the proposal density is constructed by training a nonparametric mixture of Gaussian distributions with the old samples and importance weights from the previous iteration. The new samples and weights are initialized according to Eq. (20) and then updated by Eq. (14) as in the first iteration. The Bayesian approach to training Gaussian mixtures will be briefly explained in Section 3.5. Because the closer to the posterior modes the higher the sample density, resampling from the repeatedly updated proposal density allows to zoom into highly probable parameter subspaces in very few iterations. With the multi-level sampling and Bayesian SIS schemes (see Section 3.2.1), the iterative Bayesian filtering framework brings three major advantages:

 (i) The posterior distribution is iteratively estimated with increasing refinement in the posterior landscape.
 (ii) The multi-level sampling algorithm keeps allocating model evaluations in parameter subspaces where the posterior probabilities are expected to be high, thus significantly improving computational efficiency.
 (iii) Resampling from an updated proposal density happens only between two consecutive iterations, which can circumvent the weight degeneracy problem while keeping sample trajectories intact within each iteration.

Table 1
Iterative Bayesian filtering framework pseudocode.

Input:

N_p: sample size

Θ_{min} and Θ_{max}: initial guess of parameter ranges

α: precision parameter of the Dirichlet process Gaussian mixture model

Output:

Posterior distribution/modes of micromechanical parameters

Steps:

1. Draw N_p parameter values $\{\Theta^{(i)}; i = 1, \ldots, N_p\}$ from a proposal density, run model evaluations, and assemble augmented model states $\hat{\mathbf{x}}_{0:T} = (\mathbf{x}_{0:T}, \Theta)$

   - If $k = 0$, $\Theta^{(i)} \sim \mathrm{U}(\Theta)$ and set the corresponding important weights $w_0^{(i)}$ to $1/N_p$
   - If $k > 0$, $\Theta^{(i)} \sim p_{k-1}(\Theta|\mathbf{y}_{1:T})$ and set $w_0^{(i)}$ to $\sum_1^{N_p} p_{k-1}(\Theta^{(i)}|\mathbf{y}_{1:T})/p_{k-1}(\Theta^{(i)}|\mathbf{y}_{1:T})$

2. Assume the normalized covariance parameter σ

3. For each $t = 1, \ldots, T$, update importance weights as $w_t^{(i)} \propto p(\mathbf{y}_t|\hat{\mathbf{x}}_t^{(i)})w_{t-1}^{(i)}$ where the likelihood $p(\mathbf{y}_t|\hat{\mathbf{x}}_t^{(i)})$ follows a multi-variate Gaussian distribution

4. Compute the effective sample size $ESS = 1/(\sum_1^{N_p} (w_t^{(i)})^2 N_p)$

   - For $k = 0$, repeat from step 2 if $ESS > 20\%$
   - For $k > 0$, repeat from step 2 if ESS is not at its maximum

5. Approximate the posterior distribution $p_k(\Theta|\mathbf{y}_{1:T})$ with the importance weights at time T, i.e., $p(\Theta|\mathbf{y}_{1:T}) \approx \sum_1^{N_p} w_T^{(i)} \delta(\Theta - \Theta^{(i)})$

6. Compute the posterior expectations $\widehat{\mathrm{E}}_k[\Theta|\mathbf{y}_{1:T}]$ and variances $\widehat{\mathrm{Var}}_k[\Theta|\mathbf{y}_{1:T}]$ with Eqs. (16) and (17)

7. Repeat from step 1 if $\mathrm{abs}(\widehat{\mathrm{E}}_k[\Theta|\mathbf{y}_{1:T}] - \widehat{\mathrm{E}}_{k-1}[\Theta|\mathbf{y}_{1:T}])/\widehat{\mathrm{E}}_k[\Theta|\mathbf{y}_{1:T}] > 10^{-2}$

3.4. Iterative sequential Monte Carlo filter

A multi-level sampling algorithm is implemented for iterative Bayesian filtering. For the first iteration ($k = 0$), a non-informative uniform proposal density is considered. For the subsequent iterations ($k > 0$), the posterior distribution $p_{k-1}(\Theta|\mathbf{y}_{1:T})$ from the previous iteration is chosen as the proposal density. Irrespective of the (re)sampling method, no sample (and model evaluation) is rejected/resampled during any iteration of Bayesian filtering. The multi-level sampling algorithm and the SMC filter for (re)initializing the samples and approximating the posterior distribution, respectively, are implemented in the Bayesian calibration toolbox ''GrainLearning'', as illustrated schematically in Fig. 1. Table 1 shows the pseudo-code of the iterative Bayesian filter.

Fig. 2 illustrates the resampling of parameters using a Dirichlet process Gaussian mixture model (DPGMM) that is trained with the initial quasi-random samples and their importance weights. After obtaining the first approximation of the posterior distribution at $k = 0$, a DPGMM is trained with the ensemble (samples and weights) shown in Fig. 2(b). An efficient variational inference method [57] is employed to determine the number of Gaussian components needed to approximate the distribution, their parameters (e.g., means and variances), and mixing proportions. Fig. 2(c) demonstrates how new parameters for the subsequent iteration are drawn from the Gaussian mixture. Note that although the illustration is in 2D, DPGMMs can perform comparably well in high dimensional space, as will be discussed later in Section 4.

The normalized covariance parameter σ for the first iteration is chosen such that the effective $ESS = 1/(\sum_1^{N_p}(w_t^{(i)})^2 N_p)$ [52] is larger than $ESS = 20\%$. The goal is to have a sufficient number of effective samples for estimating the proposal density. For the following iterations, the appropriate σ which maximizes ESS is selected. Fig. 3 shows the variation of ESS at time $t = T$ with an increasing σ. The same Bayesian filtering problem is solved repeatedly, with the proposal density and samples updated for each iteration, until the posterior expectations do not vary between two iterations. The other essential parameters except for σ are the *hyperparameters* that optimize the number of Gaussian components and mixing proportions in training the nonparametric mixture model.

3.5. Bayesian nonparametric density estimation via the Dirichlet process Gaussian mixture

Choosing an appropriate proposal density is pivotal to the effectiveness of any SMC filters. Because the posterior distribution on micromechanical parameters can be multimodal [36], a Gaussian mixture model (GMM) can help characterize the complex-shaped probability density function. However, the determination of the number of mixture components and their parameters can be challenging, especially when sparse and high dimensional data are involved.
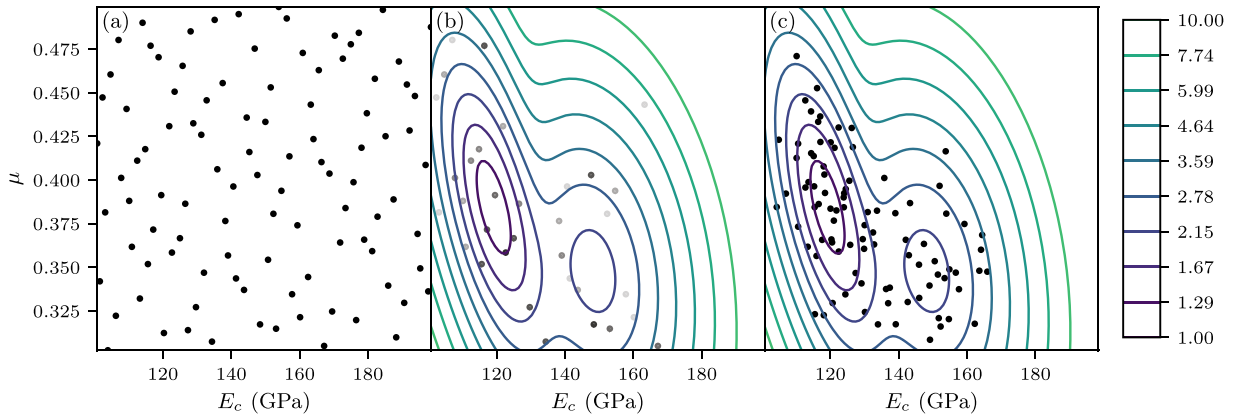
Fig. 2. (a) Initial parameter sets drawn from quasi-random numbers, (b) A mixture model trained with updated posterior probabilities and (c) new parameters resampled from the trained mixture. Note that all weighted samples in (b) are utilized in the training process; a few samples are not clearly visible because their weights are small but not negligible.
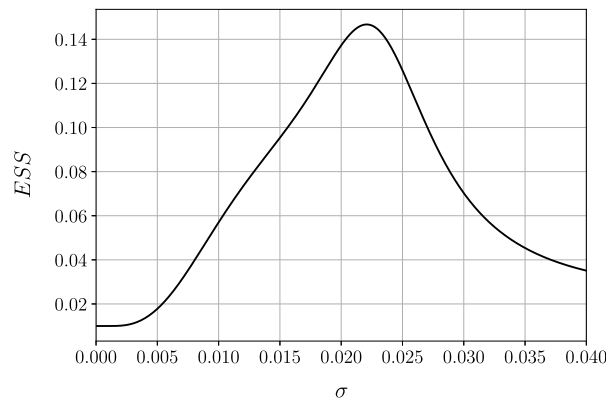

Fig. 3. Variation of the effective sample size (ESS) at time $t = T$ with an increasing normalized covariance parameter σ.

In this work, we take advantage of Bayesian nonparametrics, the Dirichlet process (DP) mixture in particular, to allow inferring an approximate posterior distribution over the mixture parameters using the *variational inference method* of Blei and Jordan [57]. The basic idea of variational inference is to formulate the computation of posterior distributions into an optimization problem which depends on a number of unknown hyperparameters. In contrast to *finite* GMMs that take as input the number of mixture components and thus are prone to "overfitting" [58,59], DPGMMs select the active components from an *infinite* Gaussian mixture by assigning the mixing proportions in a "stick-breaking" process.

In this section, Bayesian nonparametric density estimation is briefly reviewed. The Gaussian mixture model formulation is outlined. The Dirichlet process which generates the conjugate priors on the mixture parameters is introduced below, with its convenient properties and *stick-breaking* representation highlighted in Section 3.5.2.

3.5.1. Dirichlet process Gaussian mixture models

Let $\{\hat{\mathbf{x}}^{(1)}, \ldots, \hat{\mathbf{x}}^{(N)}\}$[1] be the statistically exchangeable samples of the model states in Section 3.2, drawn from the posterior distribution $F(\cdot) := p(\hat{\mathbf{x}}_{1:T}|\mathbf{y}_{1:T})$ in Section 3.3.2. A nonparametric mixture model can be constructed to estimate $F(\cdot)$ as

$$F(\cdot) = \int_{\boldsymbol{\phi} \in \boldsymbol{\Phi}} f(\cdot|\boldsymbol{\phi}) \, \mathrm{d}\mathbf{G}(\boldsymbol{\phi}), \tag{21}$$

---

[1] The sample size is enlarged by copying each sample for $N \cdot w_T^{(i)}$ times ($N > N_p$). The subscript for time t is dropped for simplicity.

where $\boldsymbol{\phi}$ contains the mixture parameters (e.g., mean and covariance), $f(\cdot|\boldsymbol{\phi})$ is the conditional density on mixture components (with parameters $\boldsymbol{\phi}$), and $\mathbf{G}(\boldsymbol{\phi})$ is the mixing distribution [60]. In Bayesian nonparametric methods, $\mathbf{G}$ is treated as an unknown "infinite-dimensional" parameter which requires a prior distribution in the usual way [61]. The draws from this prior are a collection of *random* probability distributions that belong to the same family G_0 on the $\boldsymbol{\phi}$ parameter space.

A mixture model with K Gaussian components may be written as

$$p(\hat{\mathbf{x}}|\phi_1, \ldots, \phi_K) = \sum_{j=1}^{K} \pi_j \mathcal{N}(\hat{\mathbf{x}}|\boldsymbol{\mu}_j, \mathbf{S}_j^{-1}), \tag{22}$$

where $\phi_j = \{\boldsymbol{\mu}_j, \mathbf{S}_j^{-1}\}$ is the mixture parameter for component j, $\boldsymbol{\pi}$ contains the nonzero mixing proportions that sum to unity, and $\boldsymbol{\mu}_j$ and $\mathbf{S}_j$ are the means and precisions (inverse covariances). We can now introduce a stochastic indicator variable, $c_i, i = 1, \ldots, N$, associated each sample $\hat{\mathbf{x}}^{(i)}$, to encode which component has generated the sample. The indicators can conveniently take values from $1, \ldots, K$, where K is the total number of mixture components. Suppose that we know the mixing proportions $\boldsymbol{\pi}$. Given that the likelihood of finding n_j samples belonging to component j is multinomial [62], the distribution of the indicators can be written as

$$p(c_1, \ldots, c_n|\boldsymbol{\pi}) = \prod_{j=1}^{K} \pi_j^{n_j}. \tag{23}$$

The key to generalize the parametric GMM to the nonparametric DPGMM is the use of (conditional) conjugate priors on the mixture parameters [62]. A typical choice is a Dirichlet distribution, conjugate to the multinomial distribution, which means that the posterior resulting from a Dirichlet prior and a multinomial likelihood is also a Dirichlet distribution. The Dirichlet process (DP), introduced by Ferguson [63], is a stochastic process that generalizes the Dirichlet distribution to infinite dimensions and thus can be conveniently used to set priors on unknown parameters.

3.5.2. The Dirichlet process

Given a Dirichlet process (Dir), parameterized by a scaling parameter $\alpha > 0$ and a base measure G_0, a random probability distribution G distributed according to $\mathrm{DP}(G_0, \alpha)$, satisfies for any K and K-partitions $\{A_1, \ldots, A_k\}$ of the sample space,

$$(G(A_1), \ldots, G(A_K)) \sim \mathrm{Dir}(\alpha G_0(A_1), \ldots, \alpha G_0(A_K)), \tag{24}$$

where the base measure G_0 defines the expectation $\mathrm{E}(G) = G_0$ and α is a precision parameter that controls the variance. Eqs. (23) and (24) show that samples drawn from the nonparametric mixture model that is constructed with a DP prior can be partitioned according to the distinct values of the parameters. Through the DP construction with K-partitions, the DGPMM is mathematically convenient to update the posterior distribution on unknown parameters. Let $\phi_1, \ldots, \phi_n$ be n samples independently drawn from G and $G \sim \mathrm{DP}(G_0, \alpha)$. Because of conjugacy, the posterior distribution of $G|\phi_1, \ldots, \phi_n$ is also a DP, namely

$$G|\phi_1, \ldots, \phi_n \sim \mathrm{DP}(\frac{\alpha}{\alpha + n} G_0 + \frac{1}{\alpha + n} \sum_{k=1}^{n} \delta_{\phi_k}, \alpha + n). \tag{25}$$

Another important property is that realizations of the DP are discrete, with probability one. One can show that G admits the so-called *stick-breaking* representation, established by Sethuraman [64].

$$G = \sum_{j}^{\infty} \pi_j \delta_{\phi_j}. \tag{26}$$

This representation makes clear that G is discrete, with the support encompassing an infinite set of ϕ_j, drawn independently from G_0. The mixing proportions $\boldsymbol{\pi}$ are assigned by successively breaking a unit length "stick" into an infinite number of pieces, following $\pi_j = \beta_j \prod_{l=1}^{j-1}(1 - \beta_l)$ and $\beta_j \sim \mathrm{Beta}(1, \alpha)$. For the sake of brevity, the priors on the means and precisions are not detailed here. Interested readers are encouraged to consult [57–59].

3.5.3. Variational inference for Dirichlet process mixtures

Markov Chain Monte Carlo (MCMC) sampling allows for the computation of likelihoods and posteriors involved in estimating DPGMMs. However, MCMC methods can be slow and difficult to diagnose. Here, the variational inference method introduced by Blei and Jordan [57] is utilized for the sake of efficiency. Variational inference consists in reformulating the computation of conditional distributions as an optimization problem, relaxing the problem by including information from a prior distribution, and computing optimum estimates of the hyperparameters of the prior distribution. It can be understood as an extension of expectation–maximization that maximizes a lower bound on the marginal likelihood instead of data likelihood. We make use of the publicly available open-source package *scikit-learn* [65] and in particular the `BayesianGaussianMixture` class, in which the coordinate ascent algorithm for mean-field variational inference of Blei and Jordan [57] is implemented.

The hyperparameter α plays an essential role in estimating the posterior distribution of $\{\phi_i; i = 1, \ldots, N\}$, along with the base measure G_0 which is Gaussian in `BayesianGaussianMixture`. It can be proved from Eq. (26) that the probability of ϕ_n being different from any parameters in $\{\phi_1, \ldots, \phi_{N-1}\}$ is proportional to α. Therefore, a lower value of α activates only a few components apart from many negligible ones, whereas a higher value of α leads to many unique ϕ_j and a larger K. The other parameter is the upper bound on the number of Gaussian components K_{max} for the truncation of the stick-breaking construction. For the application in Section 4, K_{max} is set to $N_p/10$, assuming clustering of 10 samples on average; α is set to 0.01 because a small number of Gaussian components is preferred. The elements in each sample, i.e., the micromechanical parameters $\Theta = \{E_c, \mu, k_m, \eta_m\}$ and the macroscopic quantities $\mathbf{x} = \{n, p', q\}$, are assumed to be correlated via a full covariance matrix.

4. Application: DEM modeling of a glass bead packing under oedometric compression

The Bayesian calibration tool is applied to DEM modeling of glass beads subjected to cyclic oedometric compression. The micromechanical parameters considered here are E_c, μ, k_m and η_m in Eqs. (1)–(5). Parameters values, *randomized* according to the multi-level sampling algorithm in Section 3.3.2, are assigned to model evaluations for identifying the optimal parameters and quantifying the micro- and macro-uncertainties conditioned on the experimental data.

4.1. Packing preparation, 3DXRCT imaging and oedometric experiment

In the experiment, two glass bead specimens were prepared with extreme care, such that their sizes, porosities and particle size distribution (PSD) were almost the same. Both specimens were isotropically compressed until reaching an effective confining pressure of 5 MPa. Subsequently, one specimen was solidified with resin to perform 3D X-ray computed tomography (3DXRCT) imaging (see Fig. 5), whereas the other was kept in the triaxial cell for the mechanical tests. The representative volume shown in Fig. 4(b) is selected, such that the material enclosed is homogeneous from the viewpoint of local voxel-based porosity (see the yellow cube highlighted in Fig. 4(a)). The cyclic oedometric experiment was conducted at a strain rate of 2.0×10^{-4} %/s, with the axial strain ε_a increasing from 0% to 1.75% and then varying between approximately 1.75% and 1.0% for two cycles. The experimental data contains approximately 100 stress and strain measurements. The three measurement vectors that consist of mean effective stress p', deviatoric stress q and porosity n, respectively, are used for the calibration.

4.2. Initial particle configuration from image analysis

Fig. 5 shows the workflow of 3D feature-based image processing. Empowered by the Trainable Weka Segmentation [66] and `MorphoLibJ` library [67], the workflow is ad hoc developed to (i) segment multiphase 3DXRCT images (Figs. 5(d) and 5(e)), (ii) separate touching objects (Figs. 5(f)–5(h)), and (iii) characterize the morphological features of the objects, e.g., particle size distribution as shown in Fig. 4(c). The novelty here is the combination of morphological operations and machine learning algorithms to extract image features for segmenting multiphase images. Note that image segmentation becomes extremely difficult when defects are present and adjacent to particle-void boundaries (Fig. 5(a)). Because of the similar attenuation coefficients of voids and defects (and the partial-volume effect [68]), the particles encompassing defects tend to be thresholded as "porous" objects (Fig. 5(c)) and thus overly segmented into small fragments, by the watershed algorithm. While the non-local means (NLM) denoising (Fig. 5(b)) preserves image sharpness at the boundaries [50], the machine learning based classifier takes
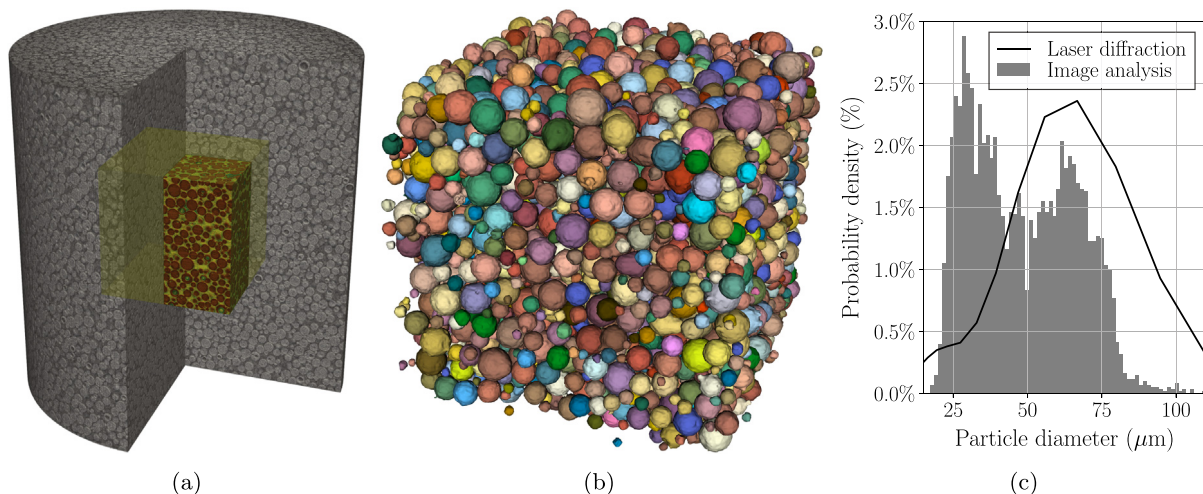
Fig. 4. Input and output of 3D feature-based image processing: (a) representative volume selected from the specimen (volume-rendered 3DXRCT images), (b) reconstruction of individual glass beads in the selected representative volume, and (c) particle size distributions obtained from morphological analysis and laser diffraction spectroscopy.
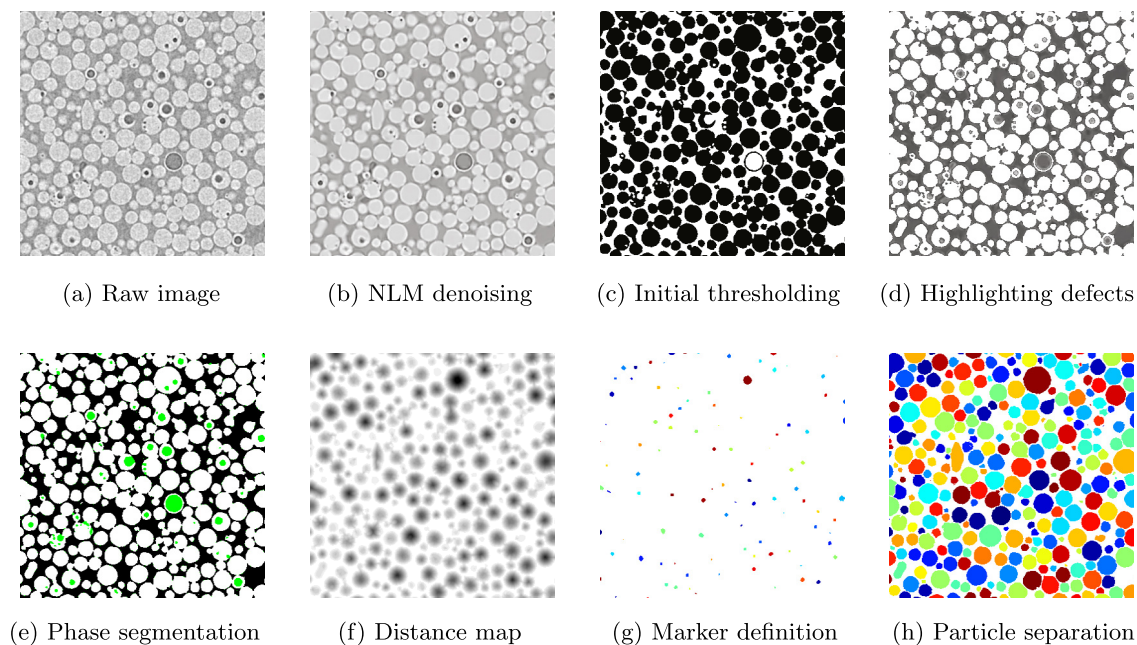

Fig. 5. Workflow of 3D feature-based image processing.

the pixels of highlighted morphological features, such as plate- and blob-like structures, as training data. Once trained, the classifier is able to handle the rest of the similar images.

The resulting representative volume contains 4526 fully reconstructed glass beads, as displayed in Fig. 4(b). By fitting the reconstructed glass beads into perfect spheres, one can estimate the equivalent diameters and other morphological characteristics. Fig. 4(c) shows the PSD obtained from the 3D image analysis. The second peak at around 62.5 µm seems to coincide with the unimodal distribution from the laser diffraction analysis. The difference between the two PSDs can be attributed to agglomerates that form when dispersing small particles into the carrier

Table 2
Upper and lower limits of the parameters to generate quasi-random numbers for the first iteration.

	E_c (GPa)	μ	k_m (10^{-6} Nm)	η_m
Θ_{max}	200	0.5	10	0.5
Θ_{min}	100	0.3	0	0.1

fluid during the laser diffraction analysis. However, a detailed morphological characterization of the glass beads is outside the scope of this paper.

4.3. DEM Simulation of oedometric compression

The glass bead specimen subjected to cyclic oedometric compression is modeled using the DEM, with the initial particle configuration obtained from the 3DXRCT images. "Random" model evaluations are managed by the multi-level sampling algorithm (Section 3.3.2), and their results are sequentially processed according to Bayesian updating (Section 3.3.1).

Although the geometrical representation of the DEM model is facilitated by the 3DXRCT images, the DEM packing still needs to undergo an initial "dynamic relaxation" stage in order to recover the periodicity of the representative volume. At this stage, the standard material properties of glass beads, e.g., $E_c = 70$ GPa, are used, and the rolling stiffness k_m and rolling friction η_m are set to zero. The interparticle friction μ is gradually reduced to match both initial porosity ($n_0 = 0.391$) and effective confining pressure ($p_0 = 5$ MPa), as in the experiment. It is worth noting that the 3DXRCT image acquisition is performed at the same confining pressure. During the initial relaxation stage, the particles readjust their positions towards equilibrium at $p_0 = 5$ MPa, while a high background damping is used to maximize kinetic energy dissipation [69,70]. When the equilibrium state is reached, the porosity is checked against the experimental value ($n_0 = 0.391$). If the porosity is too big, the interparticle friction μ is reduced by 1%, and hence another relaxation substage continues.

After the DEM packing stabilizes at the desired initial porosity and confining pressure, the model evaluations are initialized, with the rolling resistance activated and the parameters drawn from the proposal density. From the initial guess of E_c, μ, k_m and η_m in Table 2, a four-dimensional Halton sequence [71] is generated and employed as the parameter samples for the *first* iteration of Bayesian filtering. Adopting the randomized parameter values causes n_0 and p_0 in the simulations to slightly deviate from their experimental values. Therefore, each model evaluation starts with a second "dynamic relaxation" stage to correct n_0 and p_0. Until the correct initial porosity and confining pressure are restored, the interparticle friction is periodically reduced while keeping p_0 constant. After the correction finishes, the interparticle friction is raised back to the "randomized" value before the cyclic oedometric loading commences.

In each model evaluation, the DEM packing is uniaxially compressed and then subjected to two unloading–reloading cycles, as for the experiment in Section 4.1. The DEM simulation is strain-controlled, so that n, p' and q can be obtained at the exact levels of ε_a where the measurements were made. When applying a load increment, a global damping ratio of 0.2 is adopted to ensure numerical stability [72]. To obtain quasistatic simulation results, the global damping ratio is increased to 0.9 after each load increment so as to maximize kinetic energy dissipation. The time step dt is fixed to 1/10 of the critical time step as defined in [47]. Note that it is also possible to leave dt as a free parameter to optimize, such that the computational cost is reduced to the largest extent for a given level of accuracy [19,73]. It may be tempting to already use the randomized parameter values from the time when the particle configuration is imported into each model evaluation. Although this is feasible, the computational cost will increase dramatically because each model evaluation would include a complete "dynamic relaxation" stage which is more expensive than simulating the oedometric experiment.

5. Bayesian calibration of micromechanical parameters

5.1. Evolution of identified parameters

As all DEM simulations of a single iteration proceed along the oedometric loading path, the measurement data $\mathbf{y}_t$ and the model state samples $\{\mathbf{x}_t^{(i)}; i = 1, \ldots, N_p\}$ at time t are passed to the SMC filter to update
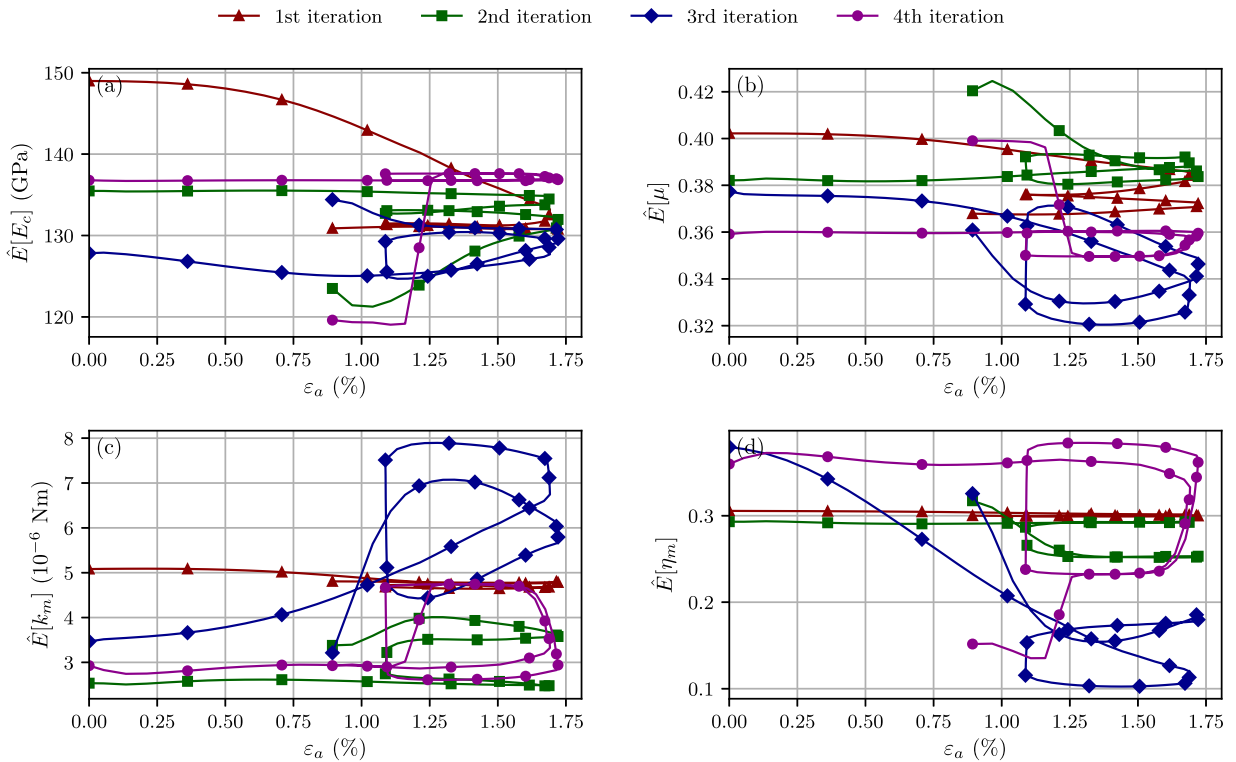
Fig. 6. Evolution of identified parameters over axial strain ε_a during each sequential Bayesian filtering. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the posterior distribution $p(\hat{\mathbf{x}}_t|\mathbf{y}_{1:t})$. To track the evolution of the ensemble (samples and weights), the posterior expectations and variances are calculated along the loading path [25,27]. Given the parameter samples and associated importance weights, the posterior expectations $\widehat{\mathrm{E}}[\boldsymbol{\Theta}_t]$ and variances $\widehat{\mathrm{Var}}[\boldsymbol{\Theta}_t]$ can be evaluated according to Eqs. (16) and (17).[2] $\widehat{\mathrm{E}}[\boldsymbol{\Theta}_t]$ contains the posterior expectations of the respective parameters $\widehat{\mathrm{E}}[E_c(t)]$, $\widehat{\mathrm{E}}[\mu(t)]$, $\widehat{\mathrm{E}}[k_m(t)]$ and $\widehat{\mathrm{E}}[\eta_m(t)]$ updated until time step t, whereas $\widehat{\mathrm{Var}}[\boldsymbol{\Theta}_t]$ consists of the corresponding variances. Note that the posterior expectations $\widehat{\mathrm{E}}[\boldsymbol{\Theta}_t]$ are defined as the "identified micromechanical parameters" rather than individual parameter sets that possess high posterior probabilities. $\sqrt{\widehat{\mathrm{Var}}[\boldsymbol{\Theta}_t]}$ are divided by $\widehat{\mathrm{E}}[\boldsymbol{\Theta}_t]$ to obtain the coefficients of variation namely $COV_{E_c}(t)$, $COV_{\mu}(t)$, $COV_{k_m}(t)$ and $COV_{\eta_m}(t)$.

In this study, four iterations of sequential Bayesian filtering are sufficient for the posterior expectations to converge. Each iteration contains 100 DEM model evaluations ($N_p = 100$) that constitute the ensemble. The parameters are (re)sampled from either the initial or the updated proposal density (see Section 3.3). For each iteration, the normalized covariance parameter σ is selected according to Section 3.4, such that the effective sample size is sufficiently large.

Figs. 6 and 7 show the evolutions of the posterior expectations and COVs during the cyclic compression. The identified parameters are plotted as functions of axial strain ε_a which controls the quasistatic loading/unloading as in the experiment. The evolution of the identified parameters with ε_a reflects the underlying Bayesian updating mechanism during the loading history; it does not mean that the parameters keep changing in the simulations.

During the first iteration, it is clear that all identified parameters $\widehat{\mathrm{E}}[E_c]$, $\widehat{\mathrm{E}}[\mu]$, $\widehat{\mathrm{E}}[k_m]$ and $\widehat{\mathrm{E}}[\eta_m]$ keep changing until the unloading–reloading cycles. Interestingly, the change of the identified parameters are negligible during the unloading–reloading cycles, except for $\widehat{\mathrm{E}}[\mu]$. This is in line with the fact that the interparticle friction dominates the plastic deformation of granular materials.

$\widehat{\mathrm{E}}[k_m]$ is another key parameter for the shear strength of granular materials [35,36]. Similar to μ, $\widehat{\mathrm{E}}[k_m]$ during the unloading–reloading cycle at the second iteration (Fig. 6(c)) also suggests that the identified parameters readjust

---

[2] The measurement sequence $\mathbf{y}_{1:T}$ which the parameter vector $\boldsymbol{\Theta}_t$ is conditioned on are dropped here for brevity.
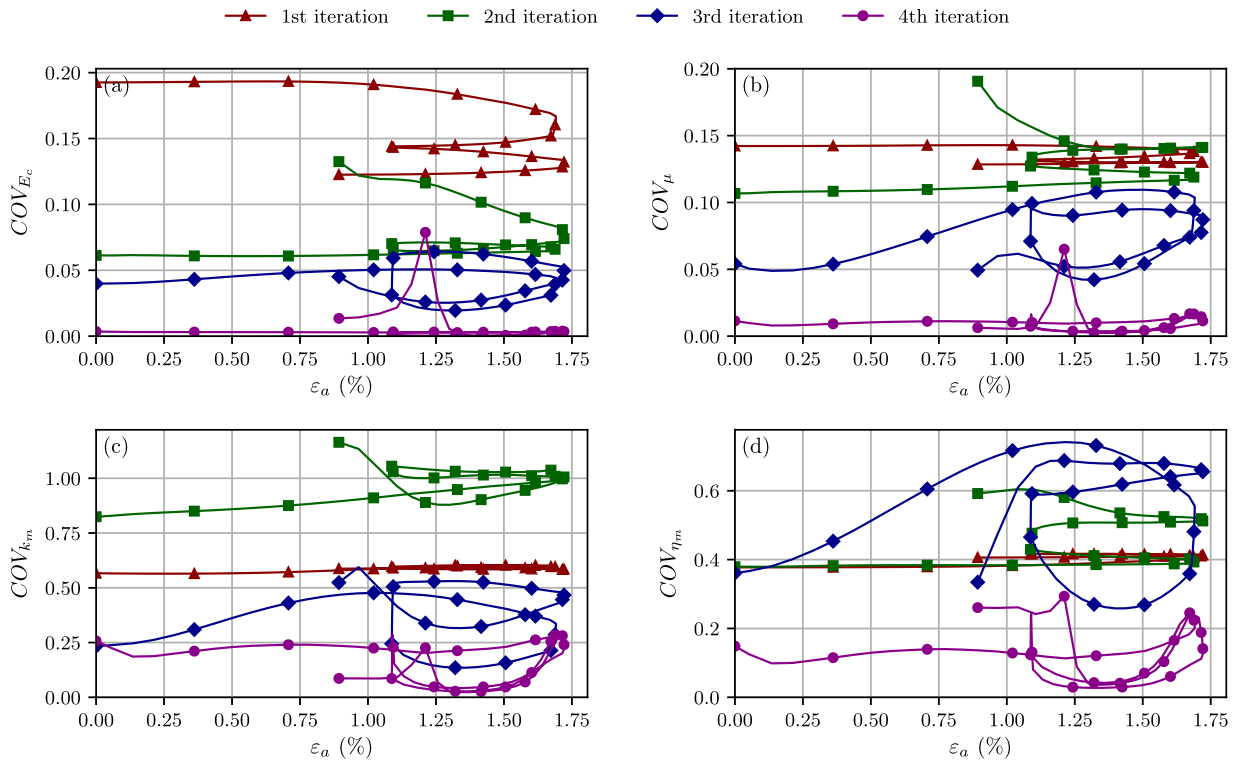
Fig. 7. Evolution of coefficients of variation over axial strain ε_a during each sequential Bayesian filtering.

further as more relevant experimental data become available. Interestingly, the change of $\widehat{E}[k_m]$ and $\widehat{E}[\eta_m]$ during the unloading–reloading cycles at the first iteration appears to be less pronounced (red line in Fig. 6(c), (d)), compared with those in the second and third iterations (green and blue lines). The difference in the identified rolling parameters between the iterations reveals that the quality of the ensemble is improved iteratively, as can be observed from the decreasing uncertainty in Fig. 7(c), (d). Fig. 7 shows that both COV_{E_c} and COV_μ decrease almost monotonically during all iterations, whereas COV_{k_m} and COV_{η_m} increase only in the first two iterations. The increasing coefficients of variation suggest that multiple modes emerge in the posterior distributions on k_m and η_m, and the initial quasi-random sampling is insufficient to find the optimal solutions. Because the DPGMM constructs a smooth and continuous distribution over the identified posterior models, potential optima can be efficiently explored at later iterations. Interestingly, in the last two iterations, all parameter uncertainties, particularly COV_{k_m} and COV_{η_m}, appear to fluctuate between certain values during the unloading–reloading cycles. The hystereses in Figs. 7 suggest that the importance weights on the samples within different posterior modes are updated differently, depending on the unloading or reloading stage. The hystereses again proves the significance of including the unloading–reloading branches of stress–strain curves in Bayesian calibration and the multimodal posterior landscape on the rolling parameters.

Note that the "identified" parameters at the end of the first iteration are almost the same as those at the beginning of the second iteration[3] (red and green lines in Fig. 6). The continuity can also be observed between the other pairs of consecutive iterations as well. The continuity in the coefficients of variation (except for k_m) is also shown in Fig. 7. This means that the knowledge updated from the previous iteration is successfully passed to the iterations afterwards, through the DP Gaussian mixture trained with the previously approximated posterior distribution. Because of the negligible change of the posterior expectations during the fourth iteration, the Bayesian calibration is finished after four iterations of sequential Bayesian filtering. Further details on the convergence will be discussed in Sections 5.3 and 6.1.

---

[3] To ensure continuity in the measurement updates, the measurement sequence is reversed in time at the second and fourth iterations.
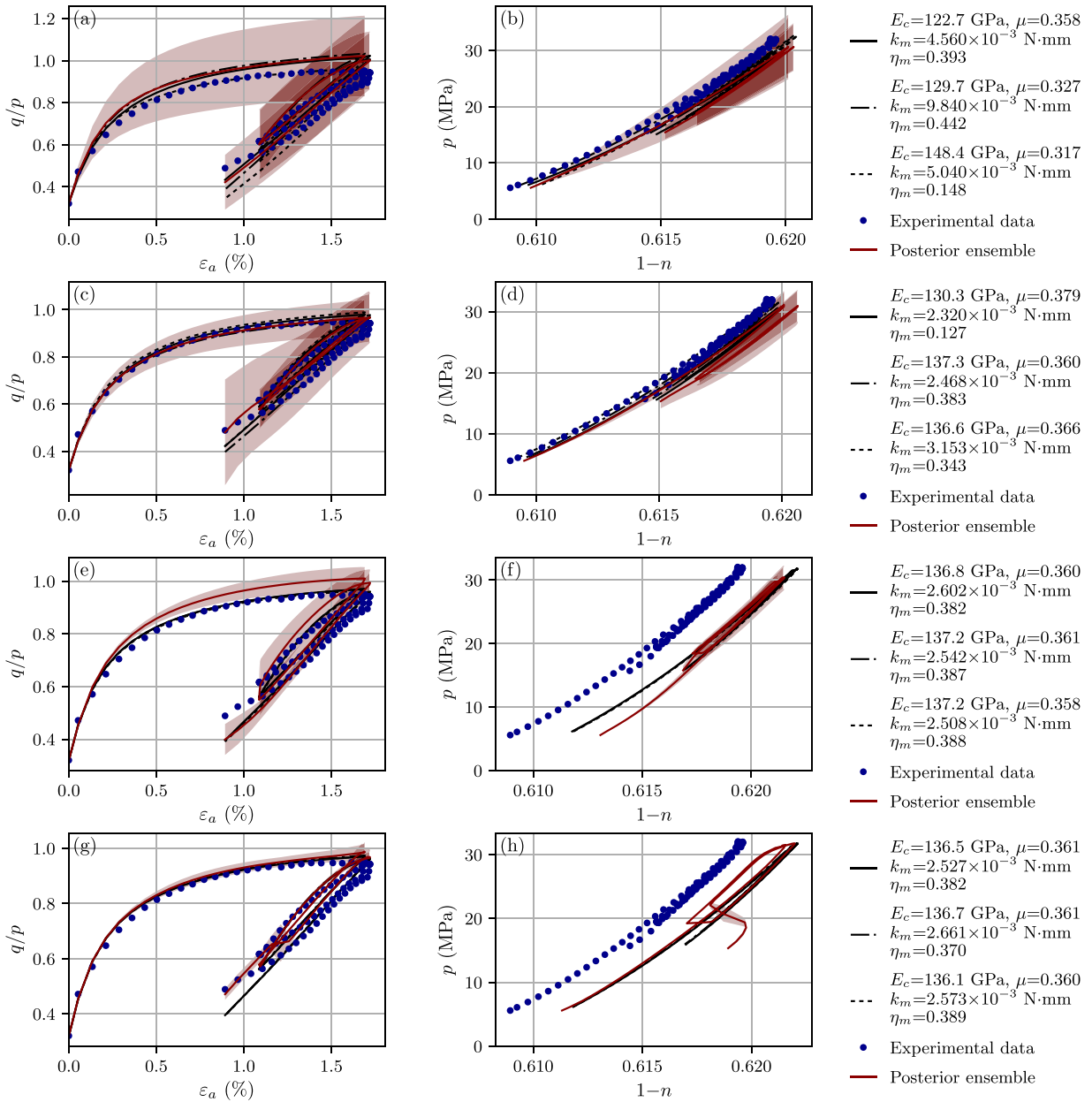
Fig. 8. Comparison of experimental data and numerical predictions given by the posterior ensemble and the parameter sets associated to the top three highest posterior probabilities at the first (a, b), second (c, d), third (e, f) and fourth (g, h) iterations. The red shaded areas indicate an uncertainty range of $\pm 2\sigma$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.2. Numerical predictions versus experimental data

To understand how the quality of the identified parameters is improved over the iterations of sequential Bayesian filtering, the numerical predictions of the macroscopic variables, namely, porosity n, mean stress p' and deviatoric stress ratio q/p' are compared with the experimental data in Fig. 8. The numerical predictions therein are obtained using the parameter sets associated with the three highest posterior probabilities. The prediction of the posterior ensemble for p' and q/p', also plotted in Fig. 8, is represented by the respective expectations and standard deviations (dark red lines and shaded areas), calculated in a similar fashion as $\widehat{E}[\Theta_t]$ and $\widehat{Var}[\Theta_t]$ in Eqs. (16) and (17).

As shown in Fig. 8(a), (b), the numerical predictions obtained at the first iteration do not match well with the experimental stress–strain ε_a–q/p and porosity–pressure n–p responses. The uncertainty is minimal at the beginning because the packing preparation ensures all model evaluations start from the same state. As stress anisotropy increases, the uncertainties in p' and q/p' keep increasing. Because the initial porosity n_0 in the simulations is servo-controlled to be precisely the same as in the experiment (see Section 4.3), the uncertainty in n never exceeds 0.5%. Fig. 8(c), (d) and (e), (f) show that the accuracy of both the ensemble and the individual model evaluations increases over the iterations. After two iterations, $COV_{p'}$ becomes less than 5%, whereas the maximum $COV_{q/p'}$ remains bigger than 10%. After the third iteration, the best three numerical predictions lie perfectly on the top of the experimental data in Fig. 8(e), with $COV_{p'}$ vanishingly small and $COV_{q/p'}$ at a minimum of 3%. Note that in Fig. 8(f) the ensemble expectation of n initially deviates from the three model evaluations. This is because the samples located further away from the posterior modes are initialized with higher importance weights, according to Eq. (20). It takes a few Bayesian updating steps before the correct modes can reemerge in the posterior distribution and the ensemble expectation rematches with the three model evaluations. The mismatch becomes more pronounced in Fig. 8(h) at the end of the loading history, that is, at the beginning of the fourth iteration because the measurement sequence is reversed (see Section 3.3.2).

At the fourth iteration, we observe a negligible change in the posterior distribution, with the majority of the samples located on the posterior modes (see Fig. 9(d)) and the uncertainty in each macroscopic QOI as low as 1%, throughout the loading history. The Bayesian approach works well because it extracts and makes use of the hidden relations between the micro- and macro-quantities given the measurement data, without reproducing every experimental detail. For example, simulating surface asperities with non-spherical particles in DEM requires a significant computational cost. Alternatively, the rolling resistance can mimic the effect of surface roughness and asperities on the macroscopic stress–strain behavior to a large extent. By means of Bayesian updating, the intrinsic relationships between the physical and non-physical parameters are identified. The micro–macro correlations are implicitly utilized in the iterative (re)sampling of the parameter/state space.

5.3. Evolution of the posterior distribution over iterations

The dependence between micromechanical parameters can be understood from the landscape of the posterior distribution. Fig. 9 compares the snapshots of the posterior distribution approximated at the beginning and the end of each iteration. While the diagonal panels indicate the marginals of the joint posterior distribution, the above- and below-diagonal panels present the 2D projections of the posterior distribution at the beginning (blue) and the end (red) of each iteration. The projections in the above- and below-diagonal 2D parameter spaces are colored by the associated posterior probability densities. The posterior distribution at the end of the first iteration (Fig. 9(a)) appears to be multimodal. Each mode of the distribution suggests a highly probable parameter subspace which may contain a potential posterior estimate. The distribution updated at the end (below-diagonal panels of Fig. 9(a)) is adopted as the new proposal density for initializing the second iteration.

Similarity between the posterior distributions that link two consecutive iterations (e.g., the below-diagonal and above-diagonal panels of Figs. 9(a) and 9(b)) confirms that the DPGMM allows to resample between the highly probable parameter subspaces, so that multiple peaks resulting from the previous iteration are merged into a smooth complex-shaped proposal density for reinitializing the succeeding iteration. Although the posterior distributions of two consecutive iterations are approximated with different ensembles, their ensemble expectations and variances are mostly consistent as shown in Figs. 6 and 7. Over the iterations, the proposal density and thus the posterior distribution become progressively localized and degenerate into multiple peaks with fluctuating probabilities, which give rise to the hystereses in Fig. 6.

Note that the ranges of E_c and μ explored in the third iteration (Fig. 9(c)) are much smaller than the initially sampled ones (see Table 2 and Fig. 9(a)), whereas the ranges of k_m and η_m are almost the same between these two iterations. Although the sample density for E_c and μ is high, the posterior distribution remains multimodal throughout the third iteration. The complexity of the posterior distribution is mostly attributed to the importance weights associated with the nonphysical parameters k_m and η_m. This suggests that most computational time is used in identifying the nonphysical parameters k_m and η_m whose posterior distributions are mostly multimodal.

The Bayesian calibration could be stopped after three iterations because the numerical predictions using the optimal parameters match the experimental data very well, as shown in Fig. 8(c), (d). Nevertheless, an additional
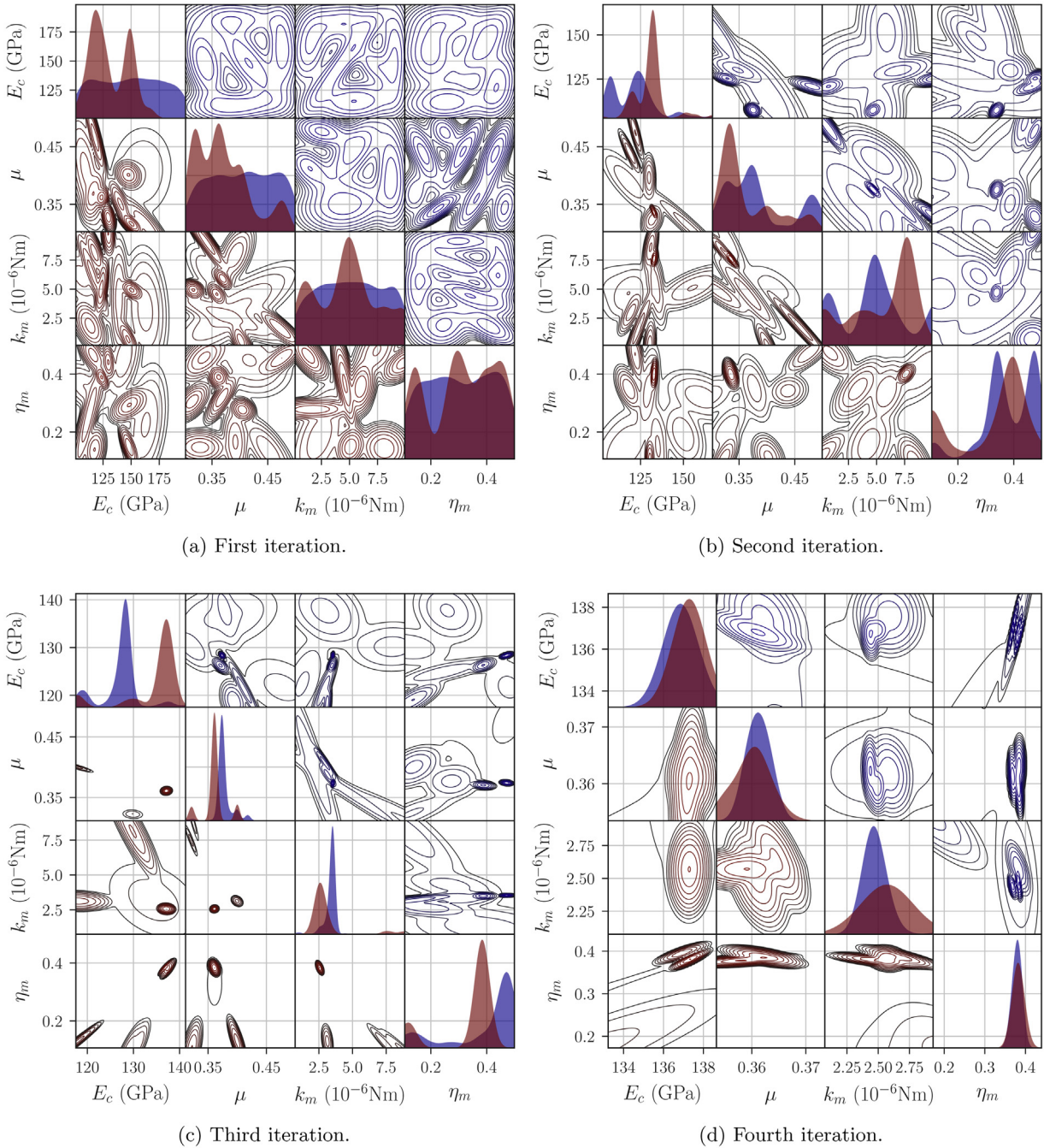
(a) First iteration.

(b) Second iteration.

(c) Third iteration.

(d) Fourth iteration.

Fig. 9. Posterior distribution approximated at the beginning (blue) and the end (red) of sequential Bayesian filtering. 2D projections of the posterior distribution in the above- and below-diagonal panels are colored by the posterior probability densities. Note that in order to clearly observe the convergence, only the primary mode of the posterior is shown in the subplot (d). See the supplementary videos for the complete evolutions of the posterior distribution over the loading history. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

iteration is performed in order to estimate the posterior distribution near the modes for checking convergence. Fig. 9(d) also shows that the "stopping criterion" is satisfied at the fourth iteration. The excellent agreement between the initial and final posterior distributions in Fig. 9(d) also suggests that the accuracy is already sufficient in the third

iteration for the dual purpose of parameter identification and uncertainty quantification. Therefore, in Section 5.4 the posterior distribution obtained at the third iteration will be used to discuss the uncertainties associated with the micromechanical parameters.

5.4. Uncertainty propagation from micro to macro

Sequential Bayesian filtering can quantify the evolution of uncertainties propagating from model parameters to macroscopic QOIs (e.g., Figs. 7 and 8), conditioned on a given history of physical states (e.g., stress anisotropy in this work). Through the samples and associated weights, the uncertainties of the micromechanical parameters can be directly linked to the uncertainties of the macroscopic quantities, such as stress, porosity and coordination number C^* (the average number of contacts per particle). This is of great interest for multiscale characterization of granular materials because uncertainty propagation from microscale to macroscale is pivotal to the design optimization of granular-based products. Uncertainty quantification and propagation allow us to understand how important a microscopic property is to a macroscopic QOI, subjected to various physical processes.

Fig. 10 shows the evolution of individual pairs of micro–macro coefficients of variation for the first three iterations. The macroscopic uncertainties increase until the peak stress ratio is reached because the covariance matrix is scaled with the measurement data. During the first iteration, $COV_{p'}$ is reduced from 10% to 5%, whereas $COV_{q/p'}$ oscillates around 10%. The uncertainty reduction at the first iteration is mostly owing to the 8% decrease of COV_{E_c}, with $COV_{p'}$ decreasing from 10% to 4% at the macroscale. During the second iteration, COV_{E_c} again decreases more than COV_μ, followed by COV_{k_m} and COV_{η_m} which hardly change. After two iterations, while COV_{k_m} and COV_{η_m} remain larger than 35%, COV_{E_c} and COV_μ drop down to approximately 7%, which jointly lead to the 5% decrease in all macroscopic uncertainties. Further decrease of the macroscopic uncertainties at the third iteration is mostly attributed to the optimization of k_m and η_m, which significantly affects n and C^*. Because the normalized covariance parameter on n is scaled by a factor of 0.01, the uncertainties in n and C^* are very low from the beginning. However, enforcing small errors on n and C^* does not guarantee low uncertainties in the predictions of p' and q. Note that COV_n is consistently smaller than COV_{C^*}, which probably relates to the fact that the coordination number can be different between DEM packings that have the same global porosity. It appears that the uncertainties in n and C^* are reduced by first locating the posterior modes on E_c and μ at the first two iterations and then optimizing the plausible combinations of k_m and η_m at the third iteration.

Similar to the proposal density (e.g., Fig. 2), the posterior distribution on the macroscopic QOIs can be obtained and plotted against that on the micromechanical parameters, using the DPGMM density estimator. Figs. 11(a)–11(c) show the micro–macro correlations at the three characteristic states, namely the initial and peak anisotropy states and the end of first unload, during the first iteration of Bayesian filtering. As expected, there exists a strong correlation between E_c and p' already at the initial state. It is interesting to see that the correlation between E_c and q/p' changes from proportional to inversely proportional, as the stress anisotropy increases. The dependence of q/p' on μ and k_m becomes much clearer at the peak stress anisotropy state. The porosity n appears to be independent of any micromechanical parameters because $n_0 \approx 0.391$ is strictly satisfied during the packing preparation. These correlations still remain valid, even after the end of the first unload (Fig. 11(c)).

The second iteration gives a more accurate overview of the landscape of the posterior distribution because of the updated proposal density. In the second iteration, already at the initial state (Fig. 11(d)), the correlations between q/p', μ and k_m are clearer than in the first iteration. Interestingly, C^* appears to be proportional to k_m, which means the increase of surface asperities, taken into account by k_m, leads to a larger coordination number. The new correlation that appears in Figs. 11(e) and 11(f) is the inverse proportionality of C^* to μ. The other correlations remain the same as in the first iteration, but with much clearer structures.

As shown in Figs. 11(g)–11(i), the micro–macro correlations are localized in the vicinity of the posterior modes, with all macroscopic uncertainties lower than 3%, Nevertheless, similar correlation structures such as the dependence of q/p' on μ and k_m and the inverse proportionality of q/p' to E_c are still observable. However, since the resampling only takes place near the posterior modes, the correlation structures estimated by the DPGMM at the third iteration are incomplete.
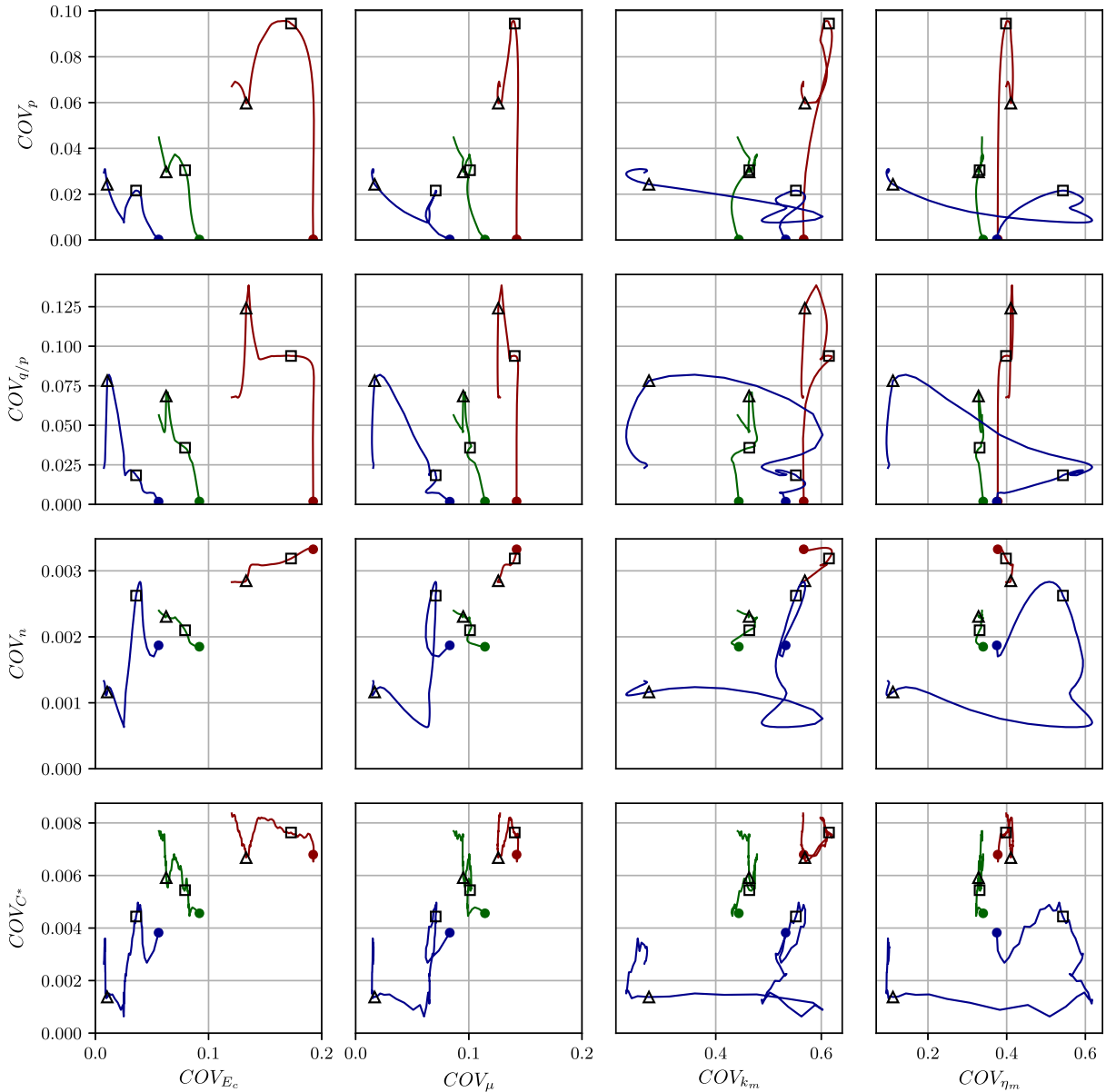
Fig. 10. Relationship between the pairs of micro–macro coefficients of variation (COV) during the first (red), second (green) and third (blue) iterations of Bayesian filtering. The beginning of each iteration is marked with a solid circle. Open squares and triangles represent the peak stress ratio state and the end of the second loading branch. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Discussion

6.1. Sampling parameter space with an iteratively updated proposal density

As discussed in Sections 5.1 and 5.3, the main advantage of iterative Bayesian filtering is the ability to update the proposal density through the Bayes theorem and then pass it to the following iteration consistently. The transfer of the updated posterior over iterations is accomplished by the resampling steps which enable the coarse-to-fine search, as shown in Fig. 12. The importance weights on the samples are taken into account by the DPGMM density
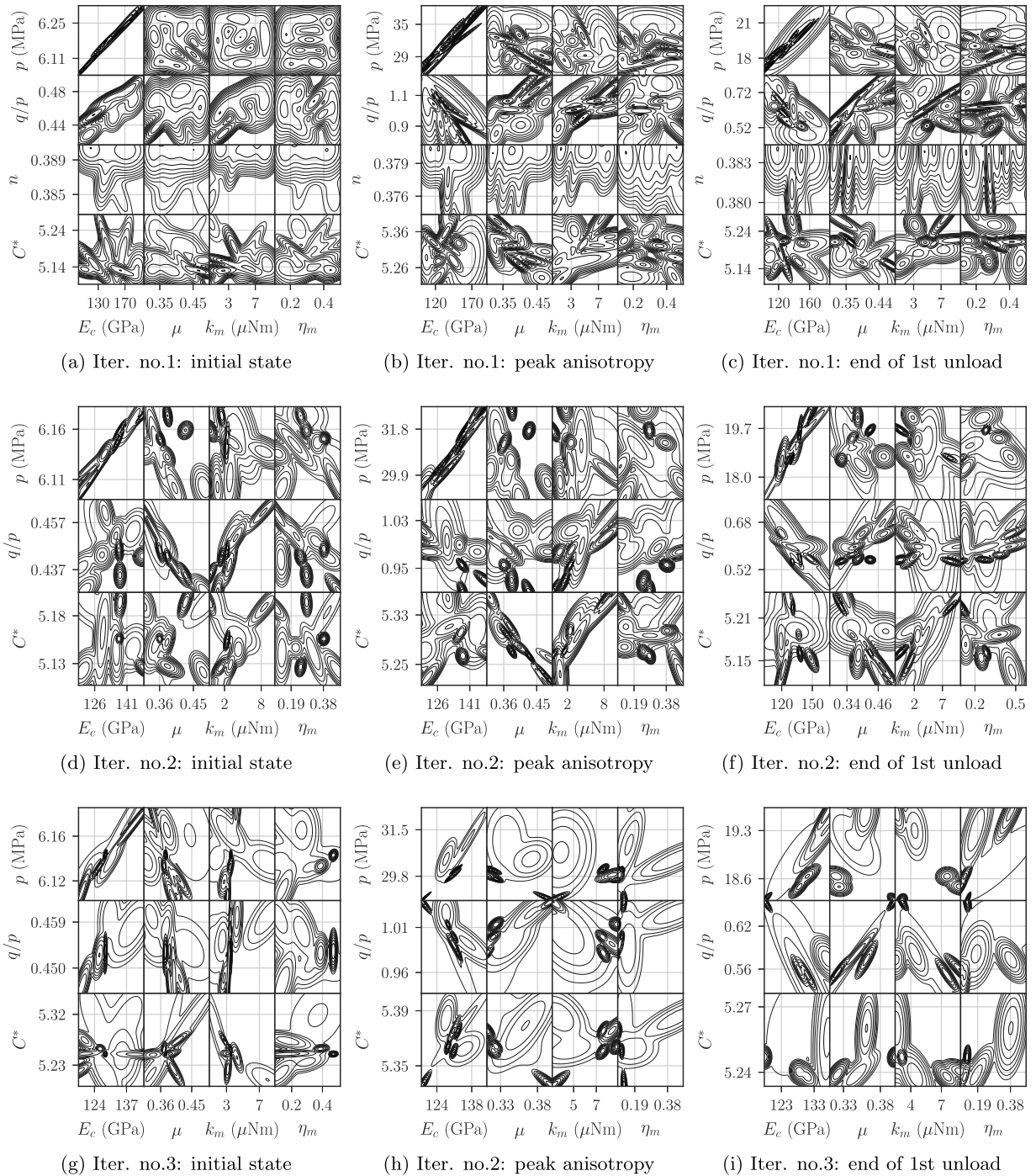
Fig. 11. Approximated posterior distributions projected on all pairs of micromechanical parameters and macroscopic quantities of interest at the three characteristic states. Note that the scales vary between the subfigures for different iterations. The posterior distributions associated with porosity n are not shown in Figs. 11(d)–11(i) because porosity becomes uncorrelated to all parameters after the first iteration.

estimator (Fig. 2(b)), which allows the approximation to localize in the highly probable regions, as shown in Figs. 12(a)–12(h). Note that only the 2D projections on the E_c–μ and k_m–η_m planes are plotted in Fig. 12 for the sake of brevity.
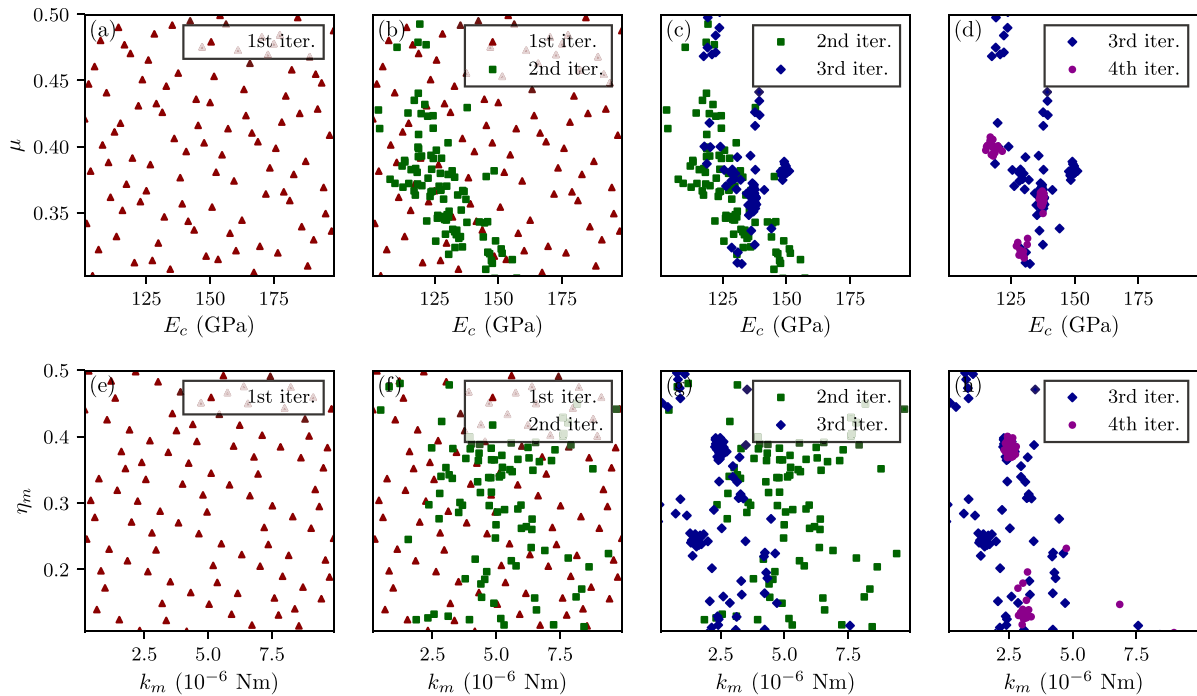
Fig. 12. Parameter samples that are progressively localized near the posterior modes over the iterations.

As demonstrated in Fig. 12(a), the initial quasi-random sampling in the to-be-explored parameter space gives uniformly distributed samples ($N_p = 100$). A larger sample size could be employed to obtain a more accurate global picture of the posterior distribution [35,36]. However, to have the same sample density as in Fig. 12(d), it would require at least 5000 DEM model evaluations initialized by the quasi-random sampler. This is neither practical nor necessary for the purpose of parameter estimation. It is more sensible to allocate the computational power to the highly probable parameter subspaces. Note that the current DEM model has four dominant micromechanical parameters. For contact models that have more parameters, one needs a larger sample size because the convergence rate of the quasi-MC error decreases with increasing dimensionality [54]. Fortunately, the number of parameters in DEM models is typically small, compared with those in most macroscopic constitutive models.

Another important issue relevant to the performance of the coarse-to-fine search is the construction of the Dirichlet process Gaussian mixture model (DPGMM). As explained in Section 3.5, the DP mixtures for the proposal density estimation are constructed using the variational inference method [57]. Compared with the conventional Gaussian mixture model (GMM), the DPGMM is advantageous in constructing a smooth and continuous probability distribution on a relatively small number of samples. First, the estimated proposal density is not subjected to the number of components initially assumed by the user because this parameter is also iterated by variational inference. When fitting a conventional GMM with samples, the expectation–maximization algorithm tends to use as many Gaussian components as possible. The resulting probability distribution will be very different, depending on the assumed number of components. On the other hand, the variational inference for the Gaussian mixture using a Dirichlet process prior does not depend on the maximum number of components initially assumed, and therefore reduces the uncertainty involved in the construction of the proposal density. In addition, unlike the conventional GMM, no singularity/over-fitting problems arise in the DPGMM, even though a large number of Gaussian components is assumed. It should be noted that no constraints are put on the parameter space, and the DPGMM can construct distributions even beyond the existing parameter space because the DPGMM tends to use the least number of active components. If the initial guess was not able to capture at least one posterior mode, the resampling scheme could explore outside the current parameter space, at the cost of more iterations and model evaluations.

The marginals and 2D projections of the posterior distribution in Fig. 9(d) and the samples in Fig. 12(d) show that the complex posterior landscape which suggests the interdependence of the parameters is filtered out after four iterations. This is because the current resampling algorithm generates fewer samples in the regions where optimal solutions are less likely to be located. These samples and the potential ones nearby, which may later be picked up by the multi-level sampler, tend to be, unfortunately, overlooked by the DPGMM density estimator. In order to retain the quality of the multimodal distribution, one can perform quasi-random sampling within the complex-shaped parameter subspace, bounded by a certain threshold of posterior probability. In doing so, the sample density near and away from the posterior modes is almost the same.

6.2. Computational efficiency

The iterative Bayesian filtering framework is implemented as a calibration tool for the open-source DEM package YADE [47]. At each iteration, "random" parameter sets are passed to YADE to initiate model evaluations (DEM simulations) that can be parallelized on a multi-core processor. In this work, the computational time needed for one DEM simulation of glass beads to undergo the cyclic oedometric loading path is approximately 3–4 h, on a dual-core system (Intel(R) Core(TM) i7-4790 CPU @ 3.60 GHz). Because the model evaluations are all independent, the computational speed can be significantly increased by simply distributing them on a computer cluster, e.g., parallelizing 3×100 model evaluations on a 100-core computer cluster allows the calibration to finish within one day. The (re)sampling of parameters for each iteration takes place after the model evaluations results are processed by the SMC filter. Through iterative resampling, the parameter space is explored at multiple levels of sample density, depending on the proposal distribution that targets the modes of the posterior distribution. Only 3×100 model evaluations are needed to accurately approximate the complex-shaped posterior distribution. The fourth iteration is needed only for checking convergence. The non-iterative Bayesian filter [36], however, would require more than 5000 DEM model evaluations to maintain the same level of sample density at the fourth iteration. In addition, the target level of accuracy also determines the number of model evaluations. In fact, the calibration could have finished after two iterations, i.e., 2×100 model evaluations because the numerical predictions obtained using the most probable combination of parameters are already in an excellent agreement with the experimental data, as shown in Fig. 8(c), (d).

Note that in the current work, each DEM simulation undergoes the entire loading history, before the macroscopic variables are passed to the SMC filter. Alternatively, all model evaluations could be stopped at a certain time step, provided that the approximation of the posterior distribution has already converged. This would definitely improve the computation efficiency, but requires the "on-line" application of SMC filtering to the model states, as soon as they become available in each model evaluation. Besides reducing the computational cost involved in DEM simulations, improving the resampling algorithm such that the dimensionality can be lowered after a few time steps/iterations will significantly improve the convergence rate. As discussed in Section 5.1, the dynamic updating of the identified E_c and μ only occurs during the first and second iterations of Bayesian filtering. Therefore, it may be more sensible to consider only k_m and η_m as the unknown parameters in the third/fourth iteration because the convergence speed increases with decreasing dimensionality.

It is very important to discuss the efficiency of the proposed Bayesian framework for high dimensional problems. According to Gerber and Chopin [54], the overall complexity of the sequential quasi-MC sampling is $\mathcal{O}\{N \log(N)\}$, which is relevant for the first iteration. If the same resampling algorithm was adopted by the remaining iterations, it would require approximately 581 model evaluations for calibrating six unknown parameters. However, in this work, after the first iteration, the resampling is performed using the DPGMM as the proposal density. Because the sample density becomes increasingly high near the posterior modes over the iterations, the number of model evaluations necessary for a six-dimensional problem should be lower than 581. Nevertheless, the relationship between the number of unknown parameters and the required size of the ensemble will be further investigated in a future dedicated work.

6.3. Bayesian uncertainty quantification for multiscale simulations of granular material

In the present work, the identification of the micromechanical parameters and the quantification of the associated uncertainties are accomplished by approximating the posterior distribution and its evolution over the loading

history. Using the identified parameters, the DEM packing can either be employed as a representative volume element to pass the local constitutive behavior to a macroscale finite element solver or stacked up to assemble a large-scale DEM simulation domain. Both approaches can lead to the Bayesian uncertainty quantification of granular mechanics across scales. The Bayesian framework is able to bridge the gap between multiscale simulations and experiments where numerical predictions and experimental measurements are obtained at various scales. For example, with the particle kinematics measured by 3D X-ray tomography, in addition to the macroscopic stress–strain measurements, it would be possible to quantify model uncertainties and their propagation from micro to macro within the Bayesian filtering framework [74]. The propagation of microscopic uncertainties to a continuum description of granular materials would also provide new insights into the micro–macro transition problem [75]. Furthermore, stochastic multiscale simulations [76] can be conducted conveniently, using material parameters drawn from the posterior distribution, e.g., estimated by the current Bayesian framework. For example, implementing a heterogeneous distribution of material properties in multiscale simulations is a typical approach to identify the sources of aleatory uncertainty [77]. The epistemic uncertainty associated with the formulation of contact laws can be evaluated by the so-called Bayes factor, which can then guide the development of new contact laws [25].

7. Conclusions

A fast, efficient and automated Bayesian calibration procedure for DEM modeling of dense granular materials is developed, based on an iteratively improved/localized approximation of the posterior distribution of micromechanical parameters, conditioned on experimental data. The previously developed calibration tool [36] is extended to enable an automated coarse-to-fine search in high-dimensional parameter space, by introducing a resampling step between iterations of nonlinear, non-Gaussian Bayesian filtering. Within each iteration, sequential Monte Carlo filtering is applied to the augmented model states and measurement data, along a given loading path. The iterative resampling scheme takes advantage of the Dirichlet process Gaussian mixture model to update a proposal density that targets the posterior modes where optimal parameters are located. The multi-level sampling algorithm allows to zoom into these highly probable parameter subspaces over iterations and characterize the local posterior landscape with great detail. The agreement of the posterior expectations before and after one iteration of Bayesian filtering is adopted as the convergence criterion.

To demonstrate its performance, the iterative Bayesian filter is applied to estimate the micromechanical parameters for DEM modeling of glass beads under cyclic oedometric compression. The DEM packing is initiated with a particle configuration acquired by 3D feature-based image processing, specially developed for segmenting 3DXRCT images of multiphase materials (particles, defects and voids in this work). The parameter sets, either sampled from the initial quasi-random sequence or resampled from the updated proposal density, are passed to the model evaluations that are parallelized on a computer cluster. The macroscopic state variables, i.e., porosity, stress and strain, along the cyclic oedometric path are returned to the sequential Monte Carlo filter. The quasi-random sequence generates samples uniformly to avoid overlooking global optima within the to-be-explored parameter space. The updated proposal density is constructed by estimating a nonparametric Gaussian mixture from the ensemble obtained at the end of the previous iteration. For identifying the posterior distribution of the parameters conditioned on the experimental stress–strain curves, 3×100 DEM model evaluations, that is three iterations of sequential Bayesian filtering, are sufficient. A fourth iteration is employed for convergence check.

In addition to the major goal of parameter identification, the iterative Bayesian filter is able to quantify correlations and uncertainties across scales. Key findings relevant to the uncertainties in micro–micro and micro–macro correlations are summarized as follows:

  (i) The posterior expectations of the contact-level Young's modulus E_c and interparticle friction μ which dominate the macroscopic stress–strain behavior converge within very few iterations. The decrease of uncertainties in the rolling parameters k_m and η_m requires more iterations, compared with E_c and μ that have clear physical origins.
 (ii) The uncertainties in micromechanical parameters drop more significantly in the unloading–reloading cycles than in the initial loading branch of the oedometric compression, which confirms the importance of incorporating elastoplastic behavior in parameter identification.
(iii) The macroscopic uncertainties could be efficiently reduced by first identifying E_c and μ, and then search for the posterior modes on k_m and η_m in later iterations.

(iv) Our micro–macro correlations from the first three iterations support what intuition suggests: an increase of E_c leads to a larger effective pressure p' and a smaller macroscopic friction q/p'; an increase of μ gives rise to a larger q/p' and a smaller coordination number C^*; and both q/p' and C^* increase with an increasing k_m. However, our method allows for a quantification of such intuitive trends.

The proposed iterative Bayesian filter is expedient because it enables an automated coarse-to-fine search for the most probable regions in high dimensional parameter space. The major advantage of implementing the iterative version of the sequential Monte Carlo filter is that the posterior distribution, which is difficult to acquire or assume, can now be learned through few iterations. With the iteratively updated proposal density, the computational power is allocated to samples which may potentially possess high posterior probabilities, thus greatly improving the efficiency of Bayesian calibration. The iterations are handled by a Python script which interfaces to DEM packages and manages the model evaluations. The current implementation is easily parallelizable because of the "off-line" estimation of the posterior distribution. Although the "off-line" implementation is simple, an "on-line" iterative Bayesian filter, i.e., incorporating the multi-level sampling and convergence check during each iteration can further improve the computational efficiency. In addition to the "on-line" implementation, future work will involve the quantification and propagation of uncertainties from micro to macro to facilitate predictive multiscale modeling of granular materials.

Acknowledgments

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cma.2019.01.027.

References

[1] P.A. Cundall, O.D.L. Strack, A discrete numerical model for granular assemblies, Géotechnique 29 (1) (1979) 47–65.
[2] X. Li, K. Wan, A bridging scale method for granular materials with discrete particle assembly – Cosserat continuum modeling, Comput. Geotech. 38 (8) (2011) 1052–1068.
[3] C. Wellmann, P. Wriggers, A two-scale model of granular materials, Comput. Methods Appl. Mech. Engrg. 205–208 (2012) 46–58.
[4] N. Guo, J. Zhao, 3D multiscale modeling of strain localization in granular media, Comput. Geotech. 80 (2016) 360–372.
[5] A. Effeindzourou, K. Thoeni, A. Giacomini, C. Wendeler, Efficient discrete modelling of composite structures for rockfall protection, Comput. Geotech. 87 (2017) 99–114.
[6] C. Yuan, B. Chareyre, A pore-scale method for hydromechanical coupling in deformable granular media, Comput. Methods Appl. Mech. Engrg. 318 (2017) 1066–1079.
[7] H. Cheng, H. Yamamoto, N. Guo, H. Huang, A simple multiscale model for granular soils with geosynthetic inclusion, in: Proc. 7th Int. Conf. Discret. Elem. Methods, Springer Singapore, Singapore, 2017, pp. 445–453.
[8] R. Fuchs, T. Weinhart, J. Meyer, H. Zhuang, T. Staedler, X. Jiang, S. Luding, Rolling, sliding and torsion of micron-sized silica particles: Experimental, numerical and theoretical analysis, Granul. Matter 16 (3) (2014) 281–297.
[9] L. Gilson, S. Kozhar, S. Antonyuk, U. Bröckel, S. Heinrich, Contact models based on experimental characterization of irregular shaped, micrometer-sized particles, Granul. Matter 16 (3) (2014) 313–326.
[10] K. Iwashita, M. Oda, Rolling resistance at contacts in simulation of shear band development by DEM, J. Eng. Mech. 124 (3) (1998) 285–292.
[11] M. Jiang, H.-S. Yu, D. Harris, A novel discrete model for granular material incorporating rolling resistance, Comput. Geotech. 32 (5) (2005) 340–357.
[12] C. Wensrich, A. Katterfeld, Rolling friction as a technique for modelling particle shape in DEM, Powder Technol. 217 (2012) 409–417.
[13] P. Kulatilake, B. Malama, J. Wang, Physical and particle flow modeling of jointed rock block behavior under uniaxial loading, Int. J. Rock Mech. Min. Sci. 38 (5) (2001) 641–657.
[14] C.J. Coetzee, Review: Calibration of the discrete element method, Powder Technol. 310 (2017) 104–142.
[15] L. Benvenuti, C. Kloss, S. Pirker, Identification of DEM simulation parameters by artificial neural networks and bulk experiments, Powder Technol. 291 (2016) 456–465.
[16] J. Yoon, Application of experimental design and optimization to PFC model calibration in uniaxial compression simulation, Int. J. Rock Mech. Min. Sci. 44 (6) (2007) 871–889.
[17] M.W. Johnstone, Calibration of DEM Models for Granular Materials using Bulk Physical Tests (Ph.D. thesis), The University of Edinburgh, 2010.

[18] K.J. Hanley, C. O'Sullivan, J.C. Oliveira, K. Cronin, E.P. Byrne, Application of Taguchi methods to DEM calibration of bonded agglomerates, Powder Technol. 210 (3) (2011) 230–240.

[19] M. Rackl, K.J. Hanley, A methodical calibration procedure for discrete element models, Powder Technol. 307 (2017) 73–83.

[20] S.K. Wilkinson, S.A. Turnbull, Z. Yan, E.H. Stitt, M. Marigo, A parametric evaluation of powder flowability using a freeman rheometer through statistical and sensitivity analysis: A discrete element method (DEM) study, Comput. Chem. Eng. 97 (2017) 161–174.

[21] R. Lewis, Y. Zheng, Coarse optimization for complex systems: An application of orthogonal experiments, Comput. Methods Appl. Mech. Engrg. 94 (1) (1992) 63–92.

[22] H.-G. Beyer, B. Sendhoff, Robust optimization – a comprehensive survey, Comput. Methods Appl. Mech. Engrg. 196 (33) (2007) 3190–3218.

[23] J.T. Oden, R. Moser, O. Ghattas, Computer predictions with quantified uncertainty, Part I, SIAM News 43 (9) (2010).

[24] J.T. Oden, R. Moser, O. Ghattas, Computer predictions with quantified uncertainty, Part II, SIAM News 43 (10) (2010).

[25] P. Hadjidoukas, P. Angelikopoulos, D. Rossinelli, D. Alexeev, C. Papadimitriou, P. Koumoutsakos, Bayesian Uncertainty quantification and propagation for discrete element simulations of granular materials, Comput. Methods Appl. Mech. Engrg. 282 (2014) 218–238.

[26] P. Hadjidoukas, P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Π4U: A high performance computing framework for Bayesian uncertainty quantification of complex models, J. Comput. Phys. 284 (2015) 1–21.

[27] P. Angelikopoulos, C. Papadimitriou, P. Koumoutsakos, Bayesian Uncertainty quantification and propagation in molecular dynamics simulations: a high performance computing framework, J. Chem. Phys. 137 (14) (2012) 144103.

[28] K. Farrell, J.T. Oden, D. Faghihi, A Bayesian framework for adaptive selection, calibration, and validation of coarse-grained models of atomistic systems, J. Comput. Phys. 295 (2015) 189–208.

[29] K. Farrell, J.T. Oden, CaLibration and validation of coarse-grained models of atomic systems: Application to semiconductor manufacturing, Comput. Mech. 54 (1) (2014) 3–19.

[30] L. Kulakova, G. Arampatzis, P. Angelikopoulos, P. Chatzidoukas, C. Papadimitriou, P. Koumoutsakos, Data driven inference for the repulsive exponent of the Lennard-Jones potential in molecular dynamics simulations, Sci. Rep. 7 (1) (2017) 16576.

[31] G. Evensen, Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res.: Oceans 99 (C5) (1994) 10143–10162.

[32] S. Nakano, G. Ueno, T. Higuchi, Merging particle filter for sequential data assimilation, Nonlinear Process. Geophys. 14 (2007) 395–408.

[33] G. Manoli, M. Rossi, D. Pasetto, R. Deiana, S. Ferraris, G. Cassiani, M. Putti, An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment, J. Comput. Phys. 283 (2015) 37–51.

[34] M. Rossi, G. Manoli, D. Pasetto, R. Deiana, S. Ferraris, C. Strobbia, M. Putti, G. Cassiani, Coupled inverse modeling of a controlled irrigation experiment using multiple hydro-geophysical data, Adv. Water Resour. 82 (2015) 150–165.

[35] H. Cheng, T. Shuku, K. Thoeni, H. Yamamoto, CaLibration of micromechanical parameters for DEM simulations by using the particle filter, EPJ Web Conf. 140 (12011) (2017).

[36] H. Cheng, T. Shuku, K. Thoeni, H. Yamamoto, Probabilistic calibration of discrete element simulations using the sequential quasi-Monte Carlo filter, Granul. Matter 20 (1) (2018) 11.

[37] H. Ruiz, H.J. Kappen, Particle smoothing for hidden diffusion processes: adaptive path integral smoother, IEEE Trans. Signal Process. 65 (12) (2017) 3191–3203.

[38] H.J. Kappen, H.C. Ruiz, Adaptive importance sampling for control and inference, J. Stat. Phys. 162 (5) (2016) 1244–1266.

[39] P. Chavali, A. Nehorai, Hierarchical particle filtering for multi-modal data fusion with application to multiple-target tracking, Signal Process. 97 (2014) 207–220.

[40] Z. Fan, H. Ji, Y. Zhang, Iterative particle filter for visual tracking, Signal Process. Image Commun. 36 (2015) 140–153.

[41] S. Yin, X. Zhu, Intelligent particle filter and its application to fault detection of nonlinear system, IEEE Trans. Ind. Electron. 62 (6) (2015) 3852–3861.

[42] I. Yoshida, T. Shuku, Particle filter with Gaussian mixture model for inverse problem, in: Proc. 6th Asian-Pacific Symp. Struct. Reliab. its Appl., 2016, pp. 643–648.

[43] Z. Medina-Cedina, H.D.V. Khoa, Probabilistic calibration of discrete particle models for geomaterials, in: Proc. of the 17th Int. Conference on Soil Mechanics and Geotechnical Engineering, 2009, pp. 704–707.

[44] J.T. Oden, E.E. Prudencio, P.T. Bauman, Virtual model validation of complex multiscale systems: Applications to nonlinear elastostatics, Comput. Methods Appl. Mech. Engrg. 266 (2013) 162–184.

[45] K. Farrell-Maupin, J.T. Oden, Adaptive selection and validation of models of complex systems in the presence of uncertainty, Res. Math. Sci. 4 (1) (2017) 14.

[46] J.T. Oden, I. Babuška, D. Faghihi, Predictive computational science: computer predictions in the presence of uncertainty, in: E. Stein, R. Borst, T.J. Hughes (Eds.), Encyclopedia of Computational Mechanics, second ed., 2017.

[47] V. Šmilauer, et al., Yade Documentation, second ed., 2015, http://dx.doi.org/10.5281/Zenodo.34073, http://yade-dem.org.

[48] K. Bagi, An algorithm to generate random dense arrangements for discrete element simulations of granular assemblies, Granul. Matter 7 (2005) 31–43.

[49] V. Magnanimo, L.L. Ragione, J.T. Jenkins, P. Wang, H.A. Makse, Characterizing the shear and bulk moduli of an idealized granular material, Europhys. Lett. 81 (3) (2008) 34006.

[50] I. Vlahinić, E. Andò, G. Viggiani, J.E. Andrade, Towards a more accurate characterization of granular media: extracting quantitative descriptors from tomographic images, Granul. Matter 16 (1) (2014) 9–21.

[51] B. Zhao, J. Wang, 3D quantitative shape analysis on form, roundness, and compactness with uCT, Powder Technol. 291 (January) (2016) 262–275.

[52] S. Särkkä, Bayesian Filtering and Smoothing. Institute of Mathematical Statistics Textbooks, Cambridge University Press, 2013.

[53] J.H. Halton, Sequential monte carlo techniques for the solution of linear systems, J. Sci. Comput. 9 (2) (1994) 213–257.

[54] M. Gerber, N. Chopin, Sequential quasi Monte Carlo, J. R. Stat. Soc. Ser. B Stat. Methodol. 77 (3) (2015) 509–579.

[55] P.J. van Leeuwen, Nonlinear data assimilation in geosciences: an extremely efficient particle filter, Q. J. R. Meteorol. Soc. 136 (653) (2010) 1991–1999.

[56] G. Kitagawa, S. Sato, Monte Carlo Smoothing and Self-Organising State-Space Model, Springer New York, New York, NY, 2001, pp. 177–195.

[57] D.M. Blei, M.I. Jordan, Variational inference for dirichlet process mixtures, Bayesian Anal. 1 (1) (2006) 121–143.

[58] C.E. Rasmussen, The infinite gaussian mixture model, in: Proceedings of the 12th International Conference on Neural Information Processing Systems, NIPS'99, MIT Press, Cambridge, MA, USA, 1999, pp. 554–560.

[59] C.M. Bishop, Pattern Recognition and Machine Learning, in: Information Science and Statistics, Springer-Verlag, Berlin, Heidelberg, 2006.

[60] F. Caron, M. Davy, A. Doucet, E. Duflos, P. Vanheeghe, Bayesian Inference for linear dynamic models with dirichlet process mixtures, IEEE Trans. Signal Process. 56 (1) (2008) 71–84.

[61] J.D. McAuliffe, D.M. Blei, M.I. Jordan, Nonparametric empirical bayes for the dirichlet process mixture model, Stat. Comput. 16 (1) (2006) 5–14.

[62] D. Görür, C. Edward Rasmussen, Dirichlet process gaussian mixture models: Choice of the base distribution, J. Comput. Sci. Tech. 25 (4) (2010) 653–664.

[63] T.S. Ferguson, A bayesian analysis of some nonparametric problems, Ann. Statist. 1 (2) (1973) 209–230.

[64] J. Sethuraman, A constructive definition of dirichlet priors, Statist. Sinica 4 (2) (1994) 639–650.

[65] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[66] I. Arganda-Carreras, V. Kaynig, C. Rueden, K.W. Eliceiri, J. Schindelin, A. Cardona, H. Sebastian Seung, Trainable weka segmentation: a machine learning tool for microscopy pixel classification, Bioinformatics 9 (15) (2017) 676–682.

[67] D. Legland, I. Arganda-Carreras, P. Andrey, Morpholibj: integrated library and plugins for mathematical morphology with ImageJ, Bioinformatics 13 (22) (2016) btw413.

[68] A. Tengattini, E. Andò, Kalisphera: an analytical tool to reproduce the partial volume effect of spheres imaged in 3D, Meas. Sci. Technol. 26 (9) (2015) 095606.

[69] H. Cheng, H. Yamamoto, K. Thoeni, Numerical study on stress states and fabric anisotropies in soilbags using the DEM, Comput. Geotech. 76 (2016) 170–183.

[70] H. Cheng, H. Yamamoto, K. Thoeni, Y. Wu, An analytical solution for geotextile-wrapped soil based on insights from DEM analysis, Geotext. Geomembranes (2017).

[71] F.-M. De Rainville, C. Gagné, O. Teytaud, D. Laurendeau, Evolutionary optimization of low-discrepancy sequences, ACM Trans. Model. Comput. Simul. 22 (2) (2012) 9:1–9:25.

[72] C. O'Sullivan, Particulate Discrete Element Modelling: A Geomechanics Perspective, Taylor & Francis, Hoboken, NJ, 2011.

[73] H.Q. Do, A.M. Aragón, D.L. Schott, A calibration framework for discrete element model parameters using genetic algorithms, Adv. Powder Technol. 29 (6) (2018) 1393–1403.

[74] A. Chernatynskiy, S.R. Phillpot, R. LeSar, Uncertainty quantification in multiscale simulation of materials: a prospective, Annu. Rev. Mater. Res. 43 (1) (2013) 157–182.

[75] C. Zhao, Z. Yin, A. Misra, P. Hicher, Thermomechanical formulation for micromechanical elasto-plasticity in granular materials, Int. J. Solids Struct. 138 (2018) 64–75.

[76] Y. Feng, T. Zhao, J. Kato, W. Zhou, Towards stochastic discrete element modelling of spherical particles with surface roughness: A normal interaction law, Comput. Methods Appl. Mech. Engrg. 315 (2017) 247–272.

[77] K. Matouš, M.G. Geers, V.G. Kouznetsova, A. Gillman, A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials, J. Comput. Phys. 330 (2017) 192–220.