Identification of continuous granular flow models

Dimitrios Papadopoulos

Delta Pi Systems Ltd., 57022 Thessaloniki, Greece

E-mail: `dimitris@delta-pi-systems.eu`

Abstract. Granular materials may exhibit different pattern forming behaviors, depending on the average energy per grain. Various granular flow PDE models exist, each capturing different behaviors of the physical phenomenon. In the present work we investigate the model and parameter identification problem of different continuous granular flow models as an encapsulated optimization problem. The identification problem is then split in a series of inverse problems. For the discrimination of the different models, the Fisher information matrix is used and different optimality criteria are discussed. Basic concepts of algorithmic differentiation (AD), which is used for the computation of the sensitivity matrix, are also given. The PDEs are discretized by the finite element method.

1. Introduction

Different approaches exist for the modeling of granular flow, from discrete to continuous ones, see [1, 6]. In this work we are going to present a process for identifying continuous models, see e.g. [2, 7, 9], from a certain class of nonlinear hyperbolic PDEs. In the following, the class of nonlinear hyperbolic PDEs is presented. In Section 2 the numerical solution approach is presented, i.e. the leapfrog method for the time integration and the finite element method for the space discretization. In Section 3 the process of model identification is presented, from the sensitivity and the covariance matrix to the Fisher information matrix, see e.g. [3], as well as the different optimality criteria and the data fitting problem. In Section 4 a short description of the algorithmic differentiation [5, 10] is given, which is necessary for the efficient computation of the sensitivity matrix. Finally, conclusions are given.

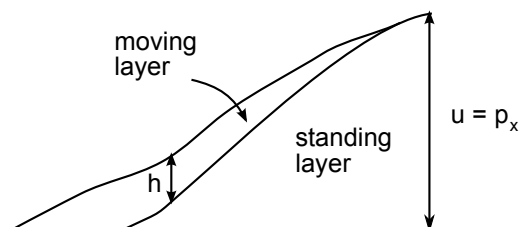In the following Figure 1 a model of granular matter piling is shown. The nonlinear hyperbolic

Figure 1. Granular matter

equations that model the continuous granular flow are given [6]:

$$h_t = \nabla(h\nabla u) + (|\nabla u| - 1)h + s, \tag{1}$$
$$u_t = (1 - |\nabla u|)h. \tag{2}$$

where u is the height of the standing layer, h is the height of the moving layer, and s a source term. In [1] it is assumed that $\nabla u > 0$ and the change of variables $\nabla u = p$ takes place. Here, we generalize the system of equations acquired in [1], by the addition of some to-be-identified parameters, into the following ones:

$$h_t = \mathbf{a}_{11} \cdot \nabla h + \mathbf{a}_{12} \cdot \nabla p + \mathbf{b}_{11} \cdot (p\nabla h) + \mathbf{b}_{12} \cdot (h\nabla p) + c_{11}h + c_{12}p + d_1 hp + s, \tag{3}$$
$$p_t = \mathbf{a}_{21} \cdot \nabla h + \mathbf{a}_{22} \cdot \nabla p + \mathbf{b}_{21} \cdot (p\nabla h) + \mathbf{b}_{22} \cdot (h\nabla p) + c_{21}h + c_{22}p + d_2 hp. \tag{4}$$

In the above system of PDEs, the parameters $\mathbf{a}_{ij}$, $\mathbf{b}_{ij}$, c_{ij}, d_i for $i, j = 1 \ldots, 2$, are to be determined during the model identification process. Again, s is a known source term.

2. Numerical solution of PDE
The previous system of nonlinear hyperbolic equations can be solved by the finite element method, see [4], for the space discretization and the leapfrog method for the time discretization.

2.1. Time integration
By using the leapfrog method the time derivative is approximated by:

$$u(t^{n+1}) = u(t^{n-1}) + 2\Delta t u_t(t^n) + O(\Delta t^3). \tag{5}$$

The leapfrog method is a second-order accurate explicit time-stepping scheme. Moreover, the leapfrog method is symplectic in time. The system of nonlinear hyperbolic equations (3) and (4) is then discretized into:

$$\frac{h^{n+1} - h^{n-1}}{2\Delta t} = g_1(h^n, p^n, \nabla h^n, \nabla p^n) + s^n, \tag{6}$$
$$\frac{p^{n+1} - p^{n-1}}{2\Delta t} = g_2(h^n, p^n, \nabla h^n, \nabla p^n), \tag{7}$$

where g_1, g_2 are the right-hand-side functions of equations (3) and (4).
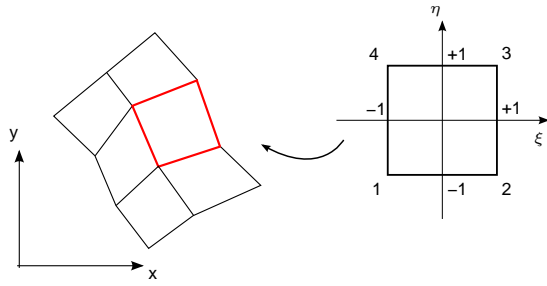
2.2. Finite Element Method
The computational domain Ω is discretized into elements Ω^e, such that:

$$\bar{\Omega} = \bigcup_{e=1}^{n_{\mathtt{el}}} \bar{\Omega}^e, \text{ and } \Omega^e \bigcap \Omega^f = \emptyset \text{ for } e \neq f,$$

where $\bar{\Omega}$ denotes the closure of Ω. The isoparametric representation of the solution is given with the help of the shape functions by:

$$\left\{ \begin{array}{c} x \\ y \end{array} \right\} = \sum_{a=1}^{n_{\mathtt{en}}} N_a(\xi, \eta) \left\{ \begin{array}{c} x_a \\ y_a \end{array} \right\}, \qquad u^h(x, y) \equiv u^h(\xi, \eta) = \sum_{a=1}^{n_{\mathtt{en}}} N_a(\xi, \eta) u_a.$$

The shape functions together with the transformation, for space dimensions $n_{\mathtt{sd}} = 2$, are given schematically in the following Figure 2.

Figure 2. Isoparametric mapping

Shape functions:
$N_1 = \frac{1}{4}(1-\xi)(1-\eta)$
$N_2 = \frac{1}{4}(1+\xi)(1-\eta)$
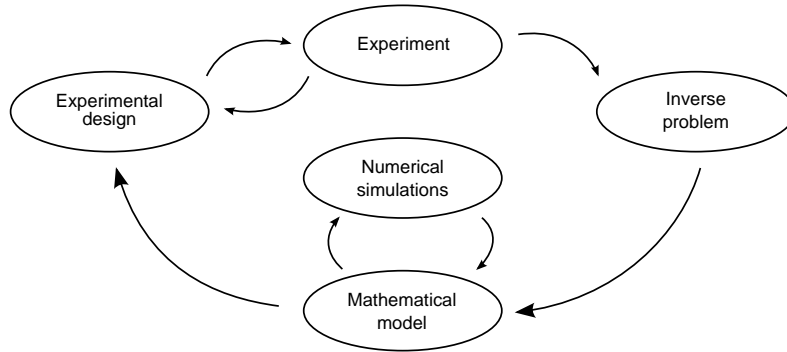$N_3 = \frac{1}{4}(1+\xi)(1+\eta)$
$N_4 = \frac{1}{4}(1-\xi)(1+\eta)$

3. Model identification

In this Section we are going to present the model identification process, see e.g. [2, 3, 7, 9], as shown in the Figure 3 below. Because experiments are time-consuming and in most cases expensive, we need methods where the number of experiments, needed for the identification of the best model describing a physical phenomenon, is minimized. This means that we know where to measure in order to increase the content of information of each experiment. Moreover, we need to quantify the influence of uncertainties on model parameters.

Assuming there are several possible modeling approaches, for $k = 1, \ldots, n_f$:

$$\mathbf{y}_k = \mathbf{f}_k(\mathbf{x}, \boldsymbol{\theta}), \tag{8}$$

where $\mathbf{x} = (x_1, \ldots, x_{n_x})$ are the independent variables, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{n_\theta})$ are the parameters of model k.

Figure 3. Model identification process

In our case $\mathbf{y}_k = \mathbf{f}_k(\mathbf{x}, \boldsymbol{\theta})$ is not an explicitly given model, but comes from the solution of the system of PDEs. The objective of this model identification process is to find which modeling approach with which model parameters describes the experiment most accurately. The parameters of each model are identified by solving a least-squares problem, which is the inverse problem in the above Figure 3, with the following objective function:

$$\min_{\theta_k} \Phi_k(\theta_k) = \sum_{i=1}^{m} (\mathbf{y}_{exp} - \mathbf{y}_k(\mathbf{x}_i, \theta_k))^T \mathbf{V}^{-1} (\mathbf{y}_{exp} - \mathbf{y}_k(\mathbf{x}_i, \theta_k)). \tag{9}$$

In the above minimization problem, Φ_k is the weighted residual for model k after m experiments, $\mathbf{y}_{exp}$ are the experimental data, $\mathbf{y}_k$ are the data from model k, $\mathbf{x}_i$ are the independent variables,

θ_k are the parameters of model k, $\mathbf{V}$ is the covariance matrix of measurement error. The minimization method used can be either a general SQP or a more specialized Levenberg-Marquardt method.

The model dependence with respect to the parameters $\boldsymbol{\theta}$ at different measurement points $\mathbf{x}_i$ for $i = 1, \ldots, m$ is given by the sensitivity matrix:

$$\mathbf{Q}_k = \begin{pmatrix} \left.\frac{\partial y_k}{\partial \theta_1}\right|_{(\mathbf{x}_1)} & \left.\frac{\partial y_k}{\partial \theta_2}\right|_{(\mathbf{x}_1)} & \cdots & \left.\frac{\partial y_k}{\partial \theta_{n_\theta}}\right|_{(\mathbf{x}_1)} \\ \left.\frac{\partial y_k}{\partial \theta_1}\right|_{(\mathbf{x}_2)} & \left.\frac{\partial y_k}{\partial \theta_2}\right|_{(\mathbf{x}_2)} & \cdots & \left.\frac{\partial y_k}{\partial \theta_{n_\theta}}\right|_{(\mathbf{x}_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial y_k}{\partial \theta_1}\right|_{(\mathbf{x}_m)} & \left.\frac{\partial y_k}{\partial \theta_2}\right|_{(\mathbf{x}_m)} & \cdots & \left.\frac{\partial y_k}{\partial \theta_{n_\theta}}\right|_{(\mathbf{x}_m)} \end{pmatrix}, \tag{10}$$

where n_θ is the number of model parameters, and m the number of measurement points.

The covariance matrix, containing the standard deviations of experimental responses is given, for $i, j = 1, \ldots, m$ measurement points, by:

$$\mathbf{V} = \begin{pmatrix} \sigma_{1,1}^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,m}^2 \\ \sigma_{2,1}^2 & \sigma_{2,2}^2 & \cdots & \sigma_{2,m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m,1}^2 & \sigma_{m,2}^2 & \cdots & \sigma_{m,m}^2 \end{pmatrix}. \tag{11}$$

It can be safely assumed that the error in measurement i is independent from measurement j. In this case the inverse covariance matrix is the diagonal matrix:

$$\mathbf{V}^{-1} = \begin{pmatrix} \frac{1}{\sigma_{1,1}^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_{2,2}^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_{m,m}^2} \end{pmatrix}. \tag{12}$$

The information measure for model k is then given by the Fisher information matrix:

$$\mathbf{F}_k = \mathbf{Q}_k^T \mathbf{V}^{-1} \mathbf{Q}_k, \tag{13}$$

where $\mathbf{V}$ is the covariance matrix of the measurement error and $\mathbf{Q}_k$ is the sensitivity matrix of model k. The Fisher information matrix $\mathbf{F}_k$ must be *"optimal"* and for this reason different optimality criteria exist [9], according to the invariant of $\mathbf{F}$:

$$\mathbf{I}_F^{(1)} = \lambda_1 + \lambda_2 + \cdots + \lambda_p = \mathrm{tr}(\mathbf{F}), \longrightarrow \text{A-Optimality},$$

$$\mathbf{I}_F^{(2)} = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \cdots + \lambda_{p-1} \lambda_p,$$

$$\vdots$$

$$\mathbf{I}_F^{(k)} = \sum_{o_1 < o_2 < \cdots < o_k}^{p} \lambda_{o_1} \lambda_{o_2} \ldots \lambda_{o_k},$$

$$\vdots$$

$$\mathbf{I}_F^{(p)} = \lambda_1 \lambda_2 \ldots \lambda_p = \det(\mathbf{F}), \longrightarrow \text{D-Optimality}.$$
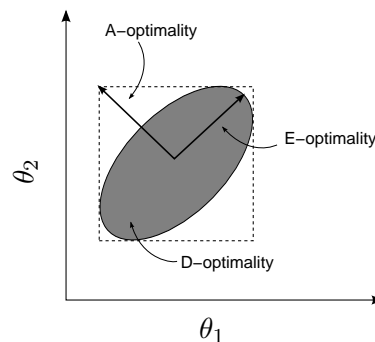
Figure 4. Geometric interpretation of different optimality criteria.

The geometric interpretation of some of the above optimality criteria is shown in Figure 4. The D-Optimality represents the volume of ellipsoid, and the E-Optimality represents the longest diagonal. Generally, one could also use a linear combination of the the above criteria, leading to a multiobjective optimization problem, but such a case is not assumed in this work.

The overall efficiency of the models is given by [2]:

$$\zeta_k(\theta) = \left[\frac{\eta_k(\theta)}{\eta^*(\theta)} \right]^{1/n_\theta}, \tag{14}$$

where $\eta^* = \max_k(\eta_k(\theta))$, and $1/n_\theta$ is a scaling factor for models with different number of parameters. Following this approach, one can also "design" the point at which the next measurement takes place, in order to maximize the information content of each experiment.

4. Algorithmic Differentiation (AD)

The main question that has to be answered regarding the computation of the jacobian matrix needed for the inverse problem and the sensitivity matrix is which method to use for their computation. One can choose from implementing the analytical derivatives manually, using a symbolic differentiation (via Computer Algebra Systems), using divided differences, and finally using an automatic (algorithmic) differentiation method. In [8] a symbolic differentiation approach with automated code generation was followed. In this work, an algorithmic differentiation [5, 10] is applied. The dilemma when selecting the step size of the divided differences is that the truncation error is small, when h small, and the loss of significance is small, when h big. However, in exact arithmetic AD is also exact! This approach [10] is based on a combination of univariate Taylor polynomial arithmetic and matrix calculus in the combined forward/reverse mode.

5. Conclusions

In this paper we have presented a framework for the model identification of continuous granular flow models. The overall process is driven by the Fisher information matrix. The model parameters are identified by the solution of a least-squares inverse problem. For the computation of the jacobian matrix of the inverse sensitivity matrix an algorithmic differentiation methodology is applied. Finally, the system of PDEs is discretized by the finite element method and solved with the leapfrog method.

6. Bibliography

[1] D. Amadori, W. Shen. A Hyperbolic Model of Granular Flow. Nonlinear partial differential equations and hyperbolic wave phenomena: 2009–2009, Research program on Nonlinear Partial Differential Equations, Centre for Advanced Study of the Norwegian Academy of Sciences and Letters, Oslo, Norway.

[2] A. Bardow. Model-based experimental analysis of multicomponent diffusion in liquids. Doctoral dissertation, VDI Verlag.

[3] E. Bertakis, M. Kalem, A. Pfenning. Model-based geometry optimization of a Nitsch cell using the Fisher information matrix. *Chemical Engineering Science* 63 (2008) 4881–4887.

[4] J. Donea, A. Huerta. Finite Element Methods for Flow Problems. John Wiley & Sons Ltd, 2003.

[5] A. Griewank. Evaluating derivatives: principles and techniques of algorithmic differentiation. SIAM, 2000.

[6] K.P. Hadeler, C. Kuttler. Dynamical models for granular matter. *Granular Matter* 2, 9-18, 1999.

[7] W. Marquardt. Model-based experimental analysis of kinetic phenomena in multi-phase reactive systems. *Chemical Engineering Research and Design* 83, 561–573.

[8] D. Papadopoulos. Optimal control of coupled PDE networks with automated code generation. *International Conference of Numerical Analysis and Applied Mathematics*, 2012.

[9] E. Walter, L. Pronzato. Identification of Parametric Models from Experimental Data. Springer, Berlin.

[10] S.F. Walter, L.Lehmann. Algorithmic differentiation in Python with AlgoPy. *Journal of Computational Science* 4(5) 2013, 334–344.