

Cite this: *Soft Matter*, 2012, **8**, 6092

www.rsc.org/softmatter

PAPER

## Extracting vibrational modes from fluctuations: a pedagogical discussion

Silke Henkes,<sup>a</sup> Carolina Brito<sup>b</sup> and Olivier Dauchot<sup>\*c</sup>

Received 22nd December 2011, Accepted 27th March 2012

DOI: 10.1039/c2sm07445a

The study of the jamming transition of granular and colloidal systems, has lead to a proliferation of theoretical and numerical results formulated in the language of the eigenspectrum of the dynamical matrix for these disordered systems. Only recently however, have these modes been accessed experimentally in colloidal and granular media, by computing the eigenmodes of the covariance matrix of the particle positions. At the same time, new conceptual and methodological questions regarding the interpretation of these results have appeared. In the present paper, we first give an overview of the theoretical framework which is appropriate to interpret the eigenmodes and eigenvalues of the correlation matrix in terms of the vibrational properties of these systems. We then illustrate several aspects of the statistical and data analysis techniques necessary to extract reliable results from experimental data. Concentrating on the cases of hard sphere simulations, colloidal and granular experiments, we discuss how to test, in turn, for the existence of a metastable state and the statistical independence of the sampling, the effect of experimental resolution, and the harmonic hypothesis underlying the approach; highlighting both the promises and limitations of this approach.

### I. Introduction

Amorphous systems such as structural glasses, colloids, emulsions or granular matter still lack a satisfying description. This is particularly apparent when one considers the low temperature properties of glasses<sup>1,2</sup> or the sometime intriguing rheological properties of athermal amorphous systems, such as foams<sup>3,4</sup> or granular matter.<sup>5,6</sup> Of particular importance is understanding what guarantees the mechanical stability of such systems.<sup>7,8</sup> Information about the rigidity of a solid toward collective particle motion is contained in the density of vibrational modes: a system is stable—at least to linear order—if there are no unstable modes. In a continuous isotropic elastic medium, translation invariance implies that the vibrational modes are plane waves, and that the density of vibrational modes follows the Debye law  $D(\omega) \sim \omega^{d-1}$ ,<sup>9</sup> where  $\omega$  is the mode frequency and  $d$  is the spatial dimension. By contrast, disordered solids exhibit a set of common low-frequency vibrational properties that are completely unlike those of crystals, among which an excess of low vibrational modes, the so-called “boson peak”.<sup>10–13</sup> Several empirical facts suggest that the presence of these excess modes is related to many of the original properties of amorphous solids.<sup>2,14,15</sup>

The jamming transition of frictionless granular materials lies at the threshold of mechanical stability, also known as the

isostatic point. As a result of this equivalence, an excess of low-frequency vibrational modes arises at low frequencies.<sup>16</sup> The similarity with the case of thermal amorphous systems suggests that the zero-temperature jamming transition may provide a framework for understanding at least some of the properties of amorphous solids. This remark has triggered an extensive theoretical and experimental effort towards the study of the vibrational properties of various amorphous systems close to jamming.

The vibrational modes of a system of interacting particles are obtained by diagonalizing the Hessian  $K_{ij} = \frac{\partial^2 V}{\partial r_i \partial r_j}$ , defined as the matrix of the second derivatives of the pair interaction potential  $V(\{r_i\})$  with respect to the particles' displacements around a given metastable state  $\{r_i^0\}$ . Above the jamming transition, it has been shown that the onset frequency  $\omega^*$  of the anomalous modes scales linearly with the distance from isostaticity, and that it is related to the existence of a diverging length scale in the response to external perturbation.<sup>17,18</sup> Beyond the frictionless, non-dissipative case, the density of states has also been studied for non-spherical<sup>19–21</sup> and for frictional<sup>22,23</sup> soft particles. Below jamming, for purely repulsive particles whether they are hard or soft, one introduces an effective potential derived from the transfer of momentum during collisions, and then applies the same procedure.<sup>24</sup>

At thermal equilibrium, the density of states of molecular and colloidal solids is traditionally accessed through light scattering techniques<sup>2,25,26</sup> or by computing the Fourier transform of the velocity autocorrelation function.<sup>27</sup> This last method, which is straightforward to implement, has also been used in thermal soft

<sup>a</sup>Physics Department, Syracuse University, Syracuse, NY 13244, USA<sup>b</sup>Inst. de Física, Universidade Federal do Rio Grande do Sul, CP 15051, 91501-970 Porto Alegre, RS, Brazil<sup>c</sup>EC2M-Gulliver, ESPCI-ParisTech and CNRS UMR 7083, 75005 Paris, France. E-mail: oliver.dauchot@espci.fr

sphere simulations<sup>28</sup> and in some granular experiments.<sup>29</sup> However, such methods do not give access to the spatial structure of the vibrational modes and in principle only apply to thermally equilibrated systems.

For a-thermal systems, such as granular media, foams and suspensions of large colloidal particles, recent advances in experimental techniques now allow tracking the real space dynamics of the individual particles.<sup>30</sup> In principle, one should thus be able to follow the same procedure as in simulations, namely average the particle positions in order to obtain a reference state,  $\{r_i^0\}$ , and then diagonalize the Hessian around this state. However, the interaction potential  $V(\{r_i\})$  is usually unknown: for granular systems or foams, the elastic part of the interaction is well understood but the pairwise dissipation, either static friction or viscous damping, does not derive from a potential; in colloidal suspensions, the presence of hydrodynamic interactions prevents the derivation of a simple two-body interaction rule. Also, experimental resolution may as well be a limiting factor in performing such an analysis.

An alternative approach is to record the dynamics and perform a Principal Component Analysis (PCA) of the covariance matrix of the positions around the reference state of interest. This technique extracts the dominant components of the fluctuations, dominant in the sense that they concentrate most of the correlated motion. PCA has been applied successfully in a range of fields, from the pathways of protein folding<sup>31,32</sup> to financial mathematics.<sup>33</sup> Whatever the underlying dynamics is, PCA acts as a filter which separates the correlated significant part of the fluctuations from the uncorrelated noise.

For systems at thermal equilibrium and as long as the dynamics around the metastable states remains harmonic, the eigenmodes of the covariance matrix of the positions are identical to those of the Hessian, and the associated eigenvalues are closely related. Following this line, several teams have recently attempted to extract the density of states from various experimental systems such as hard-sphere colloids<sup>35–38</sup> and vibrated granular materials.<sup>39</sup> While contributing to this effort,<sup>24,35,39</sup> the authors of the present paper have realized that both the interpretation of the PCA analysis and its practical implementation require some attention. As already noted in the context of granular simulations,<sup>34</sup> the conditions of the underlying equivalence at the root of the approach are often not fulfilled by the system of interest.

The aim of this article is to provide the theoretical background and a practical guideline for extracting vibrational modes from the spatial fluctuations of the dynamics. We first discuss both the promise and the limitations of the approach, as well as the set of hypotheses it depends on and the interpretation one can give to the experimental results. We pay particular attention to the influence of the real dynamics on the interpretation of the modes, and also point out the fundamental differences between equilibrium and non-equilibrium systems. We discuss the differences between the idealized spring network to which the eigenfrequencies refer to and the actual physical system by developing the concept of *shadow system*, first introduced in ref. 35. We illustrate the discussion with simulations of jammed soft spheres obeying over-damped Langevin dynamics. On the way, we will see that:

- (i) the eigenspectrum of the covariance matrix of the position provides access to the density of state only in the long time limit and if equipartition of energy holds;
- (ii) the low frequency modes of the covariance matrix of the position matches those of the Hessian in almost all circumstances, except very close to jamming where non linearities dominate the dynamics;
- (iii) in all cases, the covariance matrix of the positions is in itself a rich source of information regarding the local rigidity properties of the system.

We then provide some practical guidelines, which we hope will spare some time to those who would like to follow the above approach. We again illustrate our purpose with several systems of interest. First, we use hard sphere simulations to address statistical issues, while experiments with granular systems close to jamming illustrate the possible anharmonicity of the dynamics. Finally, two experiments with colloidal suspensions<sup>13,35</sup> illustrate convergence and resolution issues, in particular the problems that arise from an insufficient number of samplings.<sup>35</sup>

At this point, let us state that the present paper is written in a pedagogical spirit. We aim to propose a reasonably complete and coherent treatment of the conceptual and practical issues surrounding the use of PCA in analyzing vibrational properties. In particular we do not dwell on the appropriateness of the vibrational properties to the real nonlinear dynamics of the systems presented here.

## II. Theoretical background

In this section, we present rigorous derivations of the link between the covariance matrix of the positions and the Hessian matrix. Let's consider a system of  $N$  particles following trajectories  $|r(t)\rangle$  and let's assume one can define a statistical reference state  $|r^0\rangle = \langle|r(t)\rangle$  during a time interval  $\tau$ , where  $\langle\cdot\rangle$  denotes the average over time. One then introduces the displacement vector

$|\delta r(t)\rangle = |r(t)\rangle - |r^0\rangle$ , the Hessian matrix  $K_{ij} = \frac{\partial^2 V}{\partial r_i \partial r_j}$ , defined as

the matrix of the second derivatives of the pair interaction potential  $V(|r\rangle)$  with respect to the particles' displacements around the reference state  $|r^0\rangle$ , and the covariance matrix of these displacements:

$$C_p = \langle|\delta r(t)\rangle \langle\delta r(t)|\rangle. \quad (1)$$

Although the displacement vector  $|\delta r(t)\rangle$  depends on  $t$ , we will hereafter omit this dependence to simplify the notation.  $C_p$  is the empirical correlation matrix (*i.e.* on a given realization of the trajectories), that one must not confuse with the true correlation matrix of the underlying statistical process. While in this first part we will assume that  $N_\tau$ , the number of independent samples taken during  $\tau$ , is much larger than  $N$  ensuring the equivalence between time and ensemble averages and thereby the equality between the empirical and the true correlations, we shall see later that this is usually not the case in practical situations. Note that ensemble average here refers to the statistical average over the thermal fluctuations around a given glassy state. We will then have to refer to the Marčenko–Pastur theorem<sup>40</sup> and its generalizations

to discuss the convergence properties of the spectrum of  $\mathbf{C}_p$  to the one of the true correlation matrix.

$\mathbf{C}_p$  is a real symmetric matrix that can always be diagonalized. We will see now how to relate its modes and eigenvalues to those of the Hessian, depending on the underlying microscopic dynamics. We first discuss two important classes of equilibrium dynamics, namely inertial Newtonian and fully over-damped Langevin dynamics. The generalization to inertial and partially damped dynamics in the presence of colored noise will allow us to discuss the case of athermal and mechanically excited systems.

### A. Equilibrium systems

For systems in thermal equilibrium, statistical physics tells us that for equal-time correlation properties, one can forget the dynamics and replace it by Gibbs statistics. One defines the partition function  $Z \propto \int \exp(-\beta V) d|\delta r\rangle$  and in the harmonic approximation  $V = \langle \delta r | \mathbf{K} | \delta r \rangle / 2$ , so that it is straightforward to compute the correlation functions from the partition function. In particular, the correlation of the displacements reads

$$\mathbf{C}_p = k_B T \mathbf{K}^{-1}. \quad (2)$$

Hence for equilibrium dynamics, the eigenvectors of  $\mathbf{C}_p$  are simultaneous eigenvectors of  $\mathbf{K}$ , while their eigenvalues are simply inversely proportional. We will now illustrate how one can recover the above relationship in the cases of Newtonian dynamics and over-damped Langevin dynamics. The goal of this calculation is twofold. First, it will highlight the underlying hypothesis satisfied by equilibrium dynamics. Second, it will provide a simple sketch of the more intricate calculation which we perform for the case of non equilibrium mechanically excited systems.

**1. Newtonian dynamics.** In the harmonic approximation, that is linearizing the dynamics around the reference state  $|r^0\rangle$ , one has for non-dissipative dynamics

$$|\delta \ddot{r}\rangle + \frac{\mathbf{K}}{m} |\delta r\rangle = 0. \quad (3)$$

This is the first and crucial approximation underlying the whole approach. Since the reference state is supposed to be mechanically stable, the eigenvalues  $\kappa_q$  of  $\mathbf{K}$  are positive. Now introduce  $\mathbf{D} = \mathbf{K}/m$ , also called the dynamical matrix, where for simplicity we have supposed that all particles have a same mass  $m$ . The eigenmodes of  $\mathbf{D}$ , also called the vibrational modes of the system, are then defined by  $\mathbf{D}|\lambda_q\rangle = \omega_q^2|\lambda_q\rangle$ , where  $\omega_q = \sqrt{\kappa_q/m}$  are the vibrational frequencies.

The solution of eqn (3) is given by:

$$|\delta r\rangle = e^{-i\sqrt{\mathbf{D}}t} |\delta r(0)\rangle, \quad (4)$$

where  $|\delta r(0)\rangle$  is the displacement field at time zero. Placing this solution in eqn (1) and writing  $|\delta r(0)\rangle$  in the eigenbasis of  $\mathbf{D}$ ,  $\{|\lambda_q\rangle\}$ , one easily shows (see Appendix VII A for derivation) that in the long time limit

$$\mathbf{C}_p |\lambda_q\rangle = \alpha_q^2 |\lambda_q\rangle. \quad (5)$$

Here  $\alpha_q = \langle \lambda_q | \delta r(0) \rangle$  is the amplitude of the projection of the initial condition on mode  $|\lambda_q\rangle$ .

The third and final key ingredient is: at equilibrium, the initial condition is thermalized and *the energy is uniformly distributed among the modes*. Each mode of frequency  $\omega_q$  and amplitude  $\alpha_q$  carries a kinetic energy  $m\alpha_q^2\omega_q^2/2 = k_B T/2$ . Hence the eigenvalues  $\lambda_q$  of  $\mathbf{C}_p$  and the vibrational frequencies  $\omega_q$  of the dynamical matrix are related through

$$\lambda_q = \alpha_q^2 = \frac{k_B T}{m\omega_q^2}; \quad \omega_q = \sqrt{\frac{k_B T}{m\lambda_q}}. \quad (6)$$

In summary, for Newtonian dynamics at equilibrium, the diagonalization of  $\mathbf{C}_p$  does provide the vibrational modes of the system, the frequencies of which are proportional to the square root of the inverse of the eigenvalues of  $\mathbf{C}_p$ . The largest eigenvalues of  $\mathbf{C}_p$ , that is the most coherent motion, correspond to the softest modes of the system.

**2. Overdamped Langevin dynamics.** Again, in the harmonic approximation, one obtains the following Langevin equation for over-damped dynamics:

$$|\delta \dot{r}(t)\rangle = -\frac{\mathbf{K}}{\mu} |\delta r\rangle + \frac{1}{\mu} |\eta(t)\rangle. \quad (7)$$

Here  $\mu$  is the viscous damping coefficient and  $|\eta(t)\rangle$  is a white noise of variance  $\langle \eta(t') | \eta(t'') \rangle = \Gamma \delta(t' - t'')$ , with  $\Gamma = 2\mu k_B T$  since we consider equilibrium dynamics, for which the fluctuation–dissipation relation holds.

To discuss this case, we introduce the operator  $\mathcal{L} = \mathbf{K}/\mu$  whose eigenvalue equation is  $\mathcal{L}|\lambda_q\rangle = \kappa_q/|\lambda_q\rangle$ . The eigenvalues have the dimension of inverse time, which we interpret as the relaxation time  $\tau_q = \mu/\kappa_q = \mu/(m\omega_q^2)$  of the system along the eigenmode  $|\lambda_q\rangle$ .

The solution of eqn (7) can be written as

$$|\delta r\rangle = e^{-\mathcal{L}t} |\delta r(0)\rangle + \frac{1}{\mu} \int_0^t e^{-\mathcal{L}(t-t')} |\eta(t')\rangle dt' \quad (8)$$

Following the same path as for the Newtonian dynamics, we show in Appendix VII B that  $\mathbf{C}_p$  is diagonal in the eigenbasis of  $\mathcal{L}$ :

$$\mathbf{C}_p |\lambda_q\rangle = \left[ \left( \alpha_q^2 - \frac{\Gamma \tau_q}{2\mu^2} \right) e^{-2t/\tau_q} + \frac{\Gamma \tau_q}{2\mu^2} \right] |\lambda_q\rangle, \quad (9)$$

where  $\alpha_q = \langle \lambda_q | \delta r(0) \rangle$  is the amplitude of the initial condition projected on the mode  $|\lambda_q\rangle$ . To derive the above relation, we assume that the initial conditions as well as the noise components on the modes are both uncorrelated. Note that under such assumptions,  $\mathbf{C}_p$  is diagonal in the basis of  $\mathcal{L}$ , even for finite time. However it is only in the long time limit  $t \gg \tau_q$  for the largest  $\tau_q$  that one recovers the simpler expression

$$\lambda_q = \frac{\Gamma \tau_q}{2\mu^2} = \frac{k_B T}{m\omega_q^2}, \quad (10)$$

where we have used the fluctuation–dissipation theorem  $\Gamma = 2\mu k_B T$ .

We conclude that for overdamped Langevin dynamics, the diagonalization of  $\mathbf{C}_p$  again allows us to compute the vibrational modes of the system. The eigenvalues  $\lambda_q$  are related to the relaxation time  $\tau_q$  of the dynamics in each mode, which are themselves related to the oscillating frequencies  $\omega_q$  of the undamped system.

## B. Athermal systems

For many systems of interest, such as grains, foams or suspensions of large particles, the thermal fluctuations are orders of magnitude too weak to drive the dynamics. Such systems are thus generically out of equilibrium; the interactions lead to dissipation and some external forcing must be provided to drive the dynamics. The central thermal equilibrium relation eqn (2) is based on assuming both ergodicity of the fluctuations around the reference state during the observation period, and an underlying thermal canonical ensemble. Both assumptions are *a priori* violated in a non-equilibrium system. Also, we have seen above that several hypotheses are necessary to derive a relation between the eigenvalues of  $\mathbf{C}_p$  and those of the Hessian. Most of these hypotheses are related to the correlation properties of the noise and these could, in principle, be extended to non equilibrium situations. However, one key assumption is equipartition of the energy on the modes and this property is very specific to equilibrium. It is well known that for out of equilibrium situations driven by a macroscopic forcing, the energy is in general not equidistributed. On the contrary, the generic situation is that non-linearities drive an energy cascade from the largest scale to the dissipative scale. For most out of equilibrium systems, the steady-state spectrum of the energy fluctuations is not even known theoretically.

Here we conduct the same kind of analysis as above for a homogeneously driven dissipative system. Vibrated grain experiments are typical realizations of this situation. To generalize from the two equilibrium cases of Newtonian dynamics and Langevin dynamics, we now consider a non-equilibrium system with both inertia and damping and a colored noise spectrum. We choose to stick to a Langevin type description of the dynamics, where the damping is linear and single particle, as appropriate for particles in a newtonian fluid bath, but not necessarily for a large scale mechanical excitation. The generic equations of motion for the linearized dynamics around a reference state  $|r^0\rangle$  are then:

$$m|\delta\ddot{r}(t)\rangle + \mu|\delta\dot{r}(t)\rangle = -\mathbf{K}|\delta r(t)\rangle + |\eta(t)\rangle \quad (11)$$

where  $\mu$  is the viscous damping coefficient and  $|\eta(t)\rangle$  is a *colored* noise defined in the basis of the modes as

$$|\eta(t)\rangle = \sum_q \eta_q(t)|\lambda_q\rangle \quad (12)$$

with  $\langle \eta_q(t)\eta_k(t') \rangle = \Gamma_q \delta_{qk} \delta(t-t')$ .

We assume that the noise correlations are fully described by their first and second moments, *i.e.* a normal distribution. Here  $\Gamma_q$  is the amount of energy which flows into the system through non-equilibrium processes along this mode. It should be emphasized that in general this function is unknown and difficult to determine.

Solving the dynamics projected on the modes of  $\mathbf{K}$ , using the same notation as for the Langevin case,  $\mathbf{K}|\lambda_q\rangle = \kappa_q|\lambda_q\rangle$  (see Appendix VII C), we find that  $\mathbf{C}_p$  is diagonal in the eigenbasis of  $\mathbf{K}$ . At finite time one obtains complicated eigenvalues where the damping and the oscillatory component of the dynamics interplay and one must distinguish the weak  $((\mu/m)^2 < 4\kappa_q/m)$  versus the strong  $((\mu/m)^2 > 4\kappa_q/m)$  damping limits (eqn (40) and (41) of the Appendix). We have also assumed that noise and initial conditions do not cross-correlate, and that in ensemble-average,

the initial conditions of different modes are independent of each other. These conditions are in principle not so easy to satisfy if the noise is produced by a regular external excitation with a well-defined coupling to the modes. Fortunately, *in the long time limit* the system loses memory of its initial conditions and one recovers an expression similar to that of the equilibrium result

$$\mathbf{C}_p|\lambda_q\rangle = \frac{\Gamma_q}{2\mu\kappa_q}|\lambda_q\rangle. \quad (13)$$

However, one does not get rid of the violation of equipartition. The eigenvalues of  $\mathbf{C}_p$  are the ratio of two amplitudes: one is the amount of energy the external forcing has put on the mode, the other is proportional to the mode stiffness. Hence the eigenvalues of  $\mathbf{C}_p$  do *not* give access to the density of states of the system.

## III. Interpretation

We have seen in the above section that for systems in thermal equilibrium and in linear response, one can in principle safely use the covariance matrix  $\mathbf{C}_p$  of the positions around an equilibrium state in order to obtain the stiffness matrix  $\mathbf{K}$ , or equivalently the dynamical matrix  $\mathbf{D}$ , and the eigenfrequencies of the system using the relations:

$$\mathbf{K} = k_B T \mathbf{C}_p^{-1} \quad (14)$$

$$\omega_q^2 = \frac{k_B T}{m\lambda_q}. \quad (15)$$

However, real systems can be vastly more complicated than the idealized Newtonian or Langevin equilibrium systems described above. Even a model system of jammed harmonic frictionless soft spheres obeying either Newtonian dynamics or treated with a conjugate gradient algorithm exhibits strong nonlinearities when approaching the jamming transition.<sup>41,42</sup> The description of purely repulsive soft or hard spheres below jamming requires the introduction of an effective potential as defined by the transfer of momentum among the particles.<sup>24</sup> In colloidal systems, electrostatic charges and the effect of solvent can significantly alter the pair interaction potential including the development of long range interactions. Finally, we have seen that for out of equilibrium systems the possible violation of the equipartition relation prevents us from using relations (14) and (15).

It is thus of primary importance to clarify the physical interpretation one can give to the eigenmodes and eigenvalues of  $\mathbf{C}_p$ . As we shall see below, there are two strategies. One may introduce a “shadow system”, which is *by definition* the thermally equilibrated system of which  $k_B T \mathbf{C}_p^{-1}$  is the rigidity matrix. As long as the hypothesis underlying relations (14) and (15) remain reasonable for the system of interest, this shadow system should in principle have the same eigenspectrum as the real one.

If these hypotheses are not valid, in particular if the harmonic approximation or the equipartition are strongly violated, one should alternatively stick to Principal Component Analysis without expecting to relate the eigenvalues of  $\mathbf{C}_p$  to the vibrational properties. Note that the comparison with thermal systems of reference can still be done, but by computing the spectrum of  $\mathbf{C}_p$  from the density of states of the thermal system instead of trying to compute the density of states of the athermal system

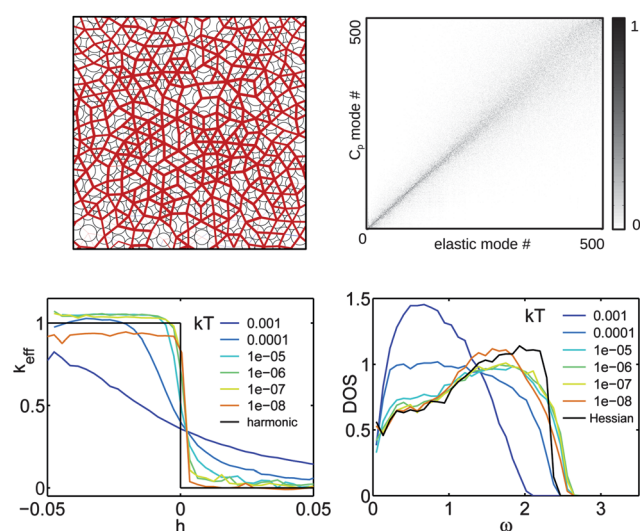
from  $\mathbf{C}_p$ . It can also be of interest, as we will illustrate below, to define, on the basis of the experimental data, a system of reference, the correlations of which have been suppressed.

### A. The “shadow system”: vibrational modes and density of states

Let us define the shadow system explicitly. Assume an experimental system of colloidal particles fluctuating around a metastable configuration  $|r^0\rangle$ . We first compute the effective spring constants  $k_{ij}^{\text{eff}}$  between particles from the empirical inverse stiffness matrix  $\mathbf{K}^{\text{eff}} = k_B T \mathbf{C}_p^{-1}$ . The shadow system is then defined as the Hamiltonian system of harmonic springs with stiffnesses  $k_{ij}^{\text{eff}}$ , strung between particles of mass  $m$  occupying the metastable configuration  $|r^0\rangle$ . Note that there is a distinction between the *shadow system* and the idealized *model system* the scientist has in mind when discussing observations of the *experimental system*. In the present case, despite possible complicated electrostatic and hydrodynamic interactions, the colloidal system is considered a good experimental realization of a model system: over-damped soft spheres in a thermal bath. Whereas here the model system obeys over-damped dynamics, the abstract shadow system obeys the Hamiltonian dynamics of perfectly harmonic springs.

For thermal systems at equilibrium, the spectrum of  $\mathbf{C}_p$  can be interpreted as the vibrational modes of the shadow system and both the shadow system and the model system share the same density of states. However, one must remember that the *vibrational properties* of the model system are still different from those of the shadow system. First, even in the linear regime, it is well known that resonance frequencies of a system of springs are shifted by damping and eventually completely eliminated in the over-damped limit. Second, the model system may well escape the linear regime. This is what happens for frictionless disks close to jamming, because of the vanishing range of linear response, as was shown in ref. 34,42. In the following we illustrate this last remark with two computational “experimental systems”. The first system consists of the poly-disperse packing of thermal frictionless soft spheres interacting through a harmonic potential, and simulated with over-damped Langevin dynamics<sup>43</sup> close to the jamming transition. The second one consists of thermal hard sphere packing simulated with event-driven Newtonian molecular dynamics just below the jamming transition. For this system, initial states are obtained by shrinking simulated packing at jamming by a factor  $\delta \in [10^{-6} - 10^{-2}]$  which controls the distance from jamming.<sup>24</sup>

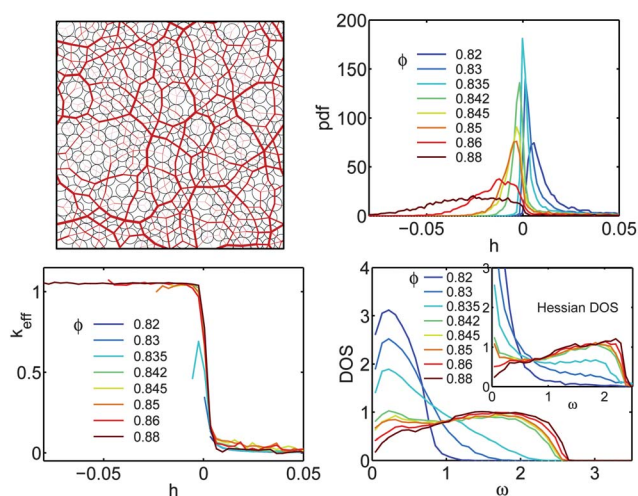
Fig. 1 (top left) displays the effective stiffnesses obtained for the soft sphere simulations slightly above jamming, at a very low temperature. Temperature here is measured as  $k_B T$  in units of the interparticle harmonic potential. We find effective stiffnesses consistent with the original connectivity and the original stiffness, which is a constant, here set to 1 for the harmonic potential. The only exceptions are occasional weak interactions with cage neighbors which are not contacts. We have cut off the visual representation in Fig. 1 (top left) at the nearest neighbor level for clarity, but any longer range interactions are at the noise level. In Fig. 1 (top right), we show the correlations between the modes of the dynamical matrix and the modes of  $\mathbf{C}_p$  for the same system. The nearly diagonal correlations show that, especially for the low energy modes, the correlation between the hessian and  $\mathbf{C}_p$  is



**Fig. 1** Top left: shadow system for a brownian system composed of soft harmonic spheres slightly above jamming at  $\phi = 0.86$  and  $k_B T = 10^{-8}$ . The contacts are plotted with widths proportional to their effective stiffness  $k_{\text{eff}}$ , red with positive  $k_{\text{eff}}$  and blue with negative  $k_{\text{eff}}$ . The representation has been cut off at the nearest neighbor level. Top right: cross-correlation coefficient between the modes of the dynamical matrix and the modes of  $\mathbf{C}_p$  for the same system. Bottom left: dependence of the effective stiffnesses  $k_{\text{eff}}$  on the interparticle gap for different temperatures  $k_B T$  and  $\phi = 0.86$ ; compared to the step function for the harmonic potential set in the simulation. Bottom right: density of states  $D(\omega)$  computed from the eigenspectrum of  $\mathbf{C}_p$  for the same range of temperatures, compared to the one derived from the Hessian of the soft sphere potential.

nearly perfect. Please also see Fig. 9 for a similar analysis of the hard sphere case. Fig. 1 bottom right displays the stiffness as a function of overlap and was obtained by binning the individual  $k_{ij}^{\text{eff}}$  by interparticle gap  $h_{ij}$  and then performing an average. At low temperatures, the stiffness as a function of overlap remains close to the  $T = 0$  step function: the shadow system is reasonably close to the “experimental” one and the density of states  $\mathbf{D}(\omega)$  computed from the Hessian matrix and the one computed from the spectrum of  $\mathbf{C}_p$  closely match (see Fig. 1, bottom right). When the temperature is increased, still remaining deeply within the glassy regime, we find that the stiffness as a function of the overlap broadens away from the  $T = 0$  step function, see Fig. 1 (bottom left). Equally, we find that the density of states obtained from  $\mathbf{C}_p$  begins to differ substantially from the Hessian density of states. The larger temperatures allow the system to explore the potential beyond the range of linear response.

The differences between the shadow system and the model system become strongly apparent if we cross the jamming transition at finite temperature. Fig. 2 shows different aspects of the shadow system as a function of packing fraction  $\phi$  at finite (low) temperature  $k_B T = 10^{-6}$ . From the plot of the shadow system at  $\phi = 0.835$ , it is clear that the stiffness now deviates significantly from a step function. The gap distribution (top right) moves from distributions with only negative gaps (overlaps) above jamming, to a situation with only positive gaps below jamming. Around jamming, we find both over- and underlaps within the same packing, and both contribute to the effective stiffness. Even though the stiffness-gap function has only slightly broadened from a step function, the different packing fractions sample

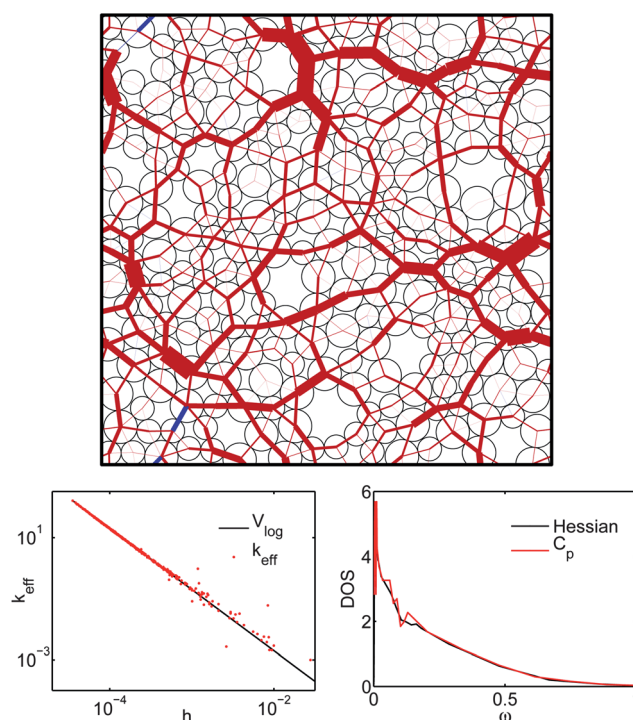


**Fig. 2** Top Left: shadow system for a packing of thermal soft sphere just below jamming at  $\phi = 0.835$  and  $k_B T = 10^{-6}$ . The contacts are plotted with widths proportional to their effective stiffness  $k_{\text{eff}}$ , red with positive  $k_{\text{eff}}$  and blue with negative  $k_{\text{eff}}$ . Top right: interparticle gap distribution for a range of packing fractions; with a negative mean gap above jamming and a positive gap below jamming. Bottom left: dependence of the effective stiffnesses  $k_{\text{eff}}$  on the interparticle gap for different  $\phi$ ; the different densities sample different regions of the stiffness curve. Bottom right: density of states  $D(\omega)$  computed from the eigenspectrum of  $\mathbf{C}_p$  compared to the one derived from the Hessian of the soft sphere potential with an added log-potential (inset; see text). All curves for  $k_B T = 10^{-6}$ .

different parts of the curve (bottom left), and the resulting shadow system is significantly different from the model system. At high densities, we do recover a shadow system density of states which resembles the Hessian of the soft sphere potential (bottom right, compare to inset). At lower densities, the similarity breaks down, most clearly below jamming, where the soft sphere density of states has a large number of zero modes, and it is mechanically unstable, unlike the shadow system one.

In the inset to the bottom right figure, we compare the shadow system density of states to the simplest effective model for thermal soft spheres:  $V_{\text{eff}} = V_{\text{harm}} + k_B T V_{\text{log}}$ , that is we have added an effective logarithmic potential for hard spheres to the soft sphere potential (see below for a discussion of the log-potential). While  $V_{\text{eff}}$  captures part of the shift of the density of states to lower frequencies below jamming, it is not a good fit: all the shadow system density of states have a peak at intermediate frequencies, unlike the density of states obtained through  $V_{\text{eff}}$ .

The effective log-potential we just used was first introduced for hard spheres slightly below jamming where the situation is again not so clear: the hard spheres potential is singular and, as for the soft spheres case, the system is *a priori* not mechanically stable any more. On one hand one can derive an effective pair potential and effective stiffnesses by computing the forces amongst the particles *via* the exchange of momentum during the collisions.<sup>24</sup> This potential scales like the logarithm of the interparticle gap. On the other hand, one can compute the effective stiffnesses derived from the inversion of  $\mathbf{C}_p$ . Fig. 3 compares both effective stiffnesses for a system of  $N = 256$  hard disks just below jamming ( $\phi = 0.83878$ ,  $\phi_J = 0.83894$ ). We indeed find an effective connectivity of the packing even though in their mean positions, the particles are not touching. We also find  $k_{\text{eff}} \sim 1/h^2$ , where  $h$  is



**Fig. 3** Top: shadow system for a packing of  $N = 256$  thermal hard spheres just below jamming ( $\phi = 0.83878$ ,  $\phi_J = 0.83894$ ). The contacts are plotted with widths proportional to their effective stiffness  $k_{\text{eff}}$ , red with positive  $k_{\text{eff}}$  and blue with negative  $k_{\text{eff}}$ . Bottom Left: effective stiffnesses as a function of the gaps separating the particles: in black the effective stiffnesses derived from the effective logarithmic potential; in red those extracted from the inversion of  $\mathbf{C}_p$ . Bottom right: density of states  $D(\omega)$  computed from the eigenspectrum of  $\mathbf{C}_p$  compared to the one derived from the Hessian of the effective potential.

the interparticle gap, consistent with a logarithmic potential. Hence the shadow system remains a good approximation in this case too. As a consequence one can again compare the density of states  $D(\omega)$  computed from  $\mathbf{C}_p$  and the one derived from the Hessian of the effective potential. One observes on Fig. 3 that they match pretty well. Given the singular character of the hard sphere potential and the fact that the reference state is only a metastable state, this is a non trivial result which validates in a self consistent way both the approach in terms of an effective potential introduced in ref. 24 and the use of  $\mathbf{C}_p$  as a tool for extracting the density of states for slightly unjammed systems.

Before ending this section, let us mention that we have followed the same analysis in the case of a real experimental system composed of NIPA colloidal particles at densities above jamming.<sup>13</sup> Conceptually, the NIPA system is very similar to the soft sphere simulation we have just discussed. However, as was already mentioned, the presence of a solvent and electrostatic interactions could alter the description in terms of a simple repulsive pair potential. We found that the effective stiffness far exceeds the noise only for adjacent particles. This tells us that elastic interactions between neighboring particles dominate the data, as they should. Furthermore, the effective stiffnesses as a function of the interparticle separation could be measured and was found to be compatible with harmonic or Hertzian interactions. We find that the shadow system captures the essence of the

vibrational properties of the real system under consideration here as well. We refer the reader interested by this experimental study to the original paper.<sup>13</sup>

## B. Principal component analysis

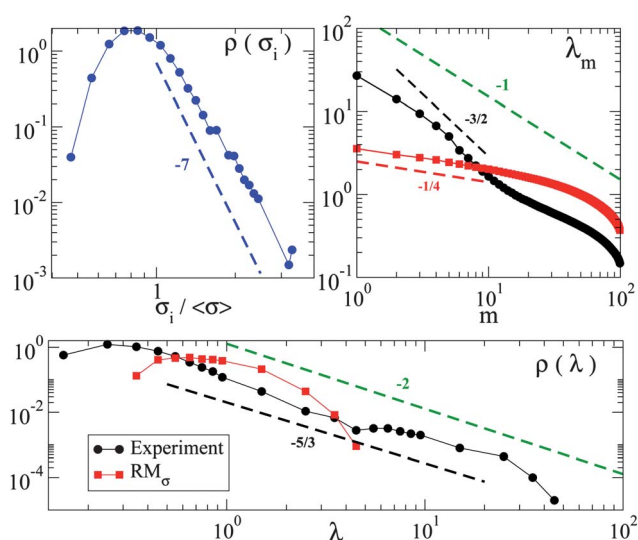
So far we have discussed how to interpret the eigenmodes and eigenvalues of  $C_p$  for thermally equilibrated systems. We have seen in Section II B that for non-equilibrium systems, the distribution of the energy on each mode must be known before extracting the density of states, and in general this is not the case. However, the spectrum of  $C_p$  and the associated modes still convey a lot of information about the dominant modes of relaxations.

Such a path has been followed in analyzing different systems for which the underlying microscopic dynamics is unknown. In ref. 33, the price fluctuations in financial markets are analyzed by comparing the spectrum of eigenvalues of the covariance matrix with the results from random matrix theory. The deviations from the purely random case contain the true information about the prices. In ref. 31 the protein near-native motion is characterized. Although molecular dynamics simulations based on an all-atom potential allows the identification of a protein functional motion on a wide range of timescales, the very large time scales are not easily accessible in simulations. A way to bypass this problem is to study the correlations of the fluctuations of the deviation of the protein backbone from its native configuration. The principal components of this matrix correspond to the collective modes of the protein which happen on larger time scales.

Two of the present authors have used a similar approach in the case of vibrated granular media.<sup>39</sup> The experimental system consists of a 1 : 1 bidisperse monolayer of brass cylinders. Mechanical energy is injected in the bulk of the system by horizontally vibrating the glass plate on which the grains stand and dissipated by solid friction. The experimental protocol produces very dense steady states with a packing fraction up to  $\phi = 0.8457$ . The stroboscopic motion of a set of 1500 grains is tracked in the center of the sample. The set-up, the quench protocols and the main properties of the system are described in detail in ref. 44,45.

We now briefly illustrate the methodology adapted in ref. 39 to discuss the spectral properties of  $C_p$ . We have considered a subset of  $N = 350$  grains acquired during  $N_\tau = 100$  time steps at a packing fraction  $\Phi = 0.844$  so that the metastable state of reference is well defined (no rattlers, no structural rearrangements). Because  $r = dN/N_\tau$ , where  $d$  is the space dimension, is larger than one, strictly zero eigenvalues appear in the spectrum of  $C_p$ . We shall come back to this practical issue in the next section and restrict the discussion to the positive eigenvalues. The fluctuations of each grain  $i$  around its metastable position are Gaussian but the width,  $\sigma_i$ , of the distribution is widely distributed amongst the particles according to the fat tailed distribution  $\rho(\sigma) \sim \sigma^{-(1+\mu)}$  plotted in Fig. 4 (top-left), where  $\mu \approx 6$ .

The Marčenko–Pastur theorem relates the spectrum computed at finite  $r$  to the one of the true correlation matrix, please see Appendix VII D for details. It could in principle be used to obtain predictions, but it provides no explicit analytical form for the shape of the eigenvalue spectrum of  $C_p$ . One alternative strategy is to define a Random Model of *uncorrelated* Gaussian variables with variance  $\sigma_i^2$  equal to the experimental one. We will



**Fig. 4** Eigenvalues of  $C_p$  for jammed grains under horizontal vibration above jamming. Top-left: probability density of the rms displacement of each grain around its average position in the metastable state. Top-right: positive eigenvalues sorted in decreasing order. Bottom: probability density of the positive eigenvalues. On the center and right plots the black dots correspond to the real dynamics and the red square to the  $RM_\sigma$  model; the dotted lines correspond to the expected distribution for a crystal according to the Debye law.

refer to this model as the  $RM_\sigma$  case. The eigenmodes extracted from the real dynamics with eigenvalues larger than the largest one of the  $RM_\sigma$  model contain the relevant information about the correlations.

Fig. 4 (top-right) and Fig. 4 (bottom) display, respectively, the eigenvalues  $\lambda_m$  sorted in decreasing order and the probability density  $\rho(\lambda)$  extracted from the real dynamics, compared to those obtained for the  $RM_\sigma$  model. Note that there is a straightforward correspondence between these two curves since  $\frac{1}{N}m(\lambda)$  is exactly the inverse cumulated distribution of  $\lambda$  and a power-law behavior  $\lambda_m \sim m^{-1/\alpha}$  translates into a density  $\rho(\lambda) \sim \lambda^{-(1+\alpha)}$ . The five largest eigenvalues for the experimental system are clearly larger than the largest one obtained for the  $RM_\sigma$  model. This comparison shows unambiguously that the top eigenvalues of  $C_p$  contain useful information about the dynamics of the system, and are not drowned in noise. It also demonstrates the existence of strong spatial correlations: by moving together, particles achieve large collective fluctuations that would not develop otherwise.

If one should not attempt to convert these eigenspectra into densities of states, it is conversely always possible to convert the expected density of states of a thermal crystal into the spectrum of its dynamical matrix. A 2d crystal has a density of states  $\rho(\omega) \sim \omega^{d-1} \sim \omega$ . Using  $\omega \propto \lambda^{1/2}$ , one obtains  $\rho(\lambda) \sim \lambda^{-2}$ , that is  $\alpha = 1$  and  $\lambda_m \sim m^{-1}$  as indicated on the figures by the green dotted line. The experimental data are consistent with an estimated  $\alpha \approx 2/3$ , that is a slower decay of the spectrum than for the crystal case. One can interpret this result by stating that there is a larger fraction of modes participating in the dynamics than in the crystalline case.

One can also analyze the structure of the most significant modes. This was done in ref. 39 for this system of bi-disperse grains when approaching jamming from above, and it could be

demonstrated quantitatively that there is a redistribution of spectral weight towards larger eigenvalues corresponding to softer modes and that the softest mode becomes more coherent and spatially organized. However, as it was discussed above, for such a system where the energy is not distributed evenly over the modes, one cannot convert this information into the vibrational properties and the density of states of the system. The results presented here as well as in ref. 39 illustrate well how PCA of the dynamics remains a fruitful tool of analysis without referring to the vibrational properties.

#### IV. Practical guidelines

In the previous section, we have assumed that the data at hand was “perfect” in the sense that neither statistical limitations nor resolution issues came into play. Also, we did not consider the possibility of a strong anharmonicity of the dynamics and we have not discussed removing rattlers, those particles which do not participate in the collective motion but instead “rattle” in their cages.

Measuring  $C_p$  to a sufficient level of precision to extract reliable information from it is not a simple affair. The goal of this section is to disentangle all the possible sources of practical problems that one faces when computing  $C_p$ , and to propose methods and alternatives to obtain faithful information from its spectrum.

In Subsection IV A, we start with a method to test the metastability of the reference state. This is crucial since the entire modes approach is based on a displacement field that fluctuates around a steady reference state. We illustrate this issue in the case of hard spheres, which by definition sit below jamming, where metastable states only have a finite lifetime. It is thus of crucial importance to control that the system does not escape the metastable state during the time window of the analysis. In Subsection IV B, using the case of vibrated granular media, we then illustrate anharmonicity setting in when approaching the jamming transition from above. The reader may also refer to ref. 42, where the same issue is discussed for the case of numerical simulations of soft spheres. As an immediate consequence of the above effects, the time window of the analysis may be seriously shortened compared to the full duration of the numerical or real experiment, and statistical issues come into play. More specifically, we shall see in Sections IV C and IV D how to deal with (the lack of) statistical independence of the data as well as convergence issues. In the case of experimental data sets, resolution issues may also alter the analysis. In Subsection IV E, we analyze an NIPA colloidal experiment to illustrate how convergence and experimental resolution interplay to affect the computation of the density of states. Finally, identifying rattlers is usually a challenging matter. It emerges that calculating the eigenspectrum itself is a good “filter” for rattlers: their motion appears only in a few low energy eigenvalues which do not mix with the remainder of the motion. This is discussed in Section IV F.

##### A. Metastability

Before starting any of the analysis discussed in the previous section, one must confirm that the dynamics are purely composed of fluctuations around a well defined reference state. This can be done both *a priori*, in order to select a good time

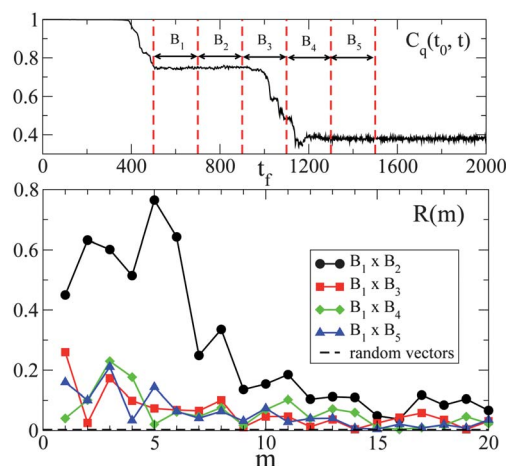
window to perform the analysis, and *a posteriori* based on the properties of  $C_p$ . A natural tool to characterize the dynamics is the self-intermediate correlation function, defined as

$$C_q(t_0, t) = \langle \cos(\vec{q} \cdot \Delta \vec{r}_i(t, t_0)) \rangle_j, \quad (16)$$

where the average is over the particles but not over the initial time. Here  $\Delta \vec{r}_i(t, t_0) = \vec{r}_i(t) - \vec{r}_i(t_0)$  is the displacement of the particle  $j$  between  $t_0$  and  $t + t_0$ ,  $\vec{q}$  is a vector whose amplitude is given by  $q = \pi/a$  and  $a$  is the displacement length scale above which the particle motion induces a change of metastable state. There is obviously an arbitrary part in such a definition and traditionally  $a$  is set to a fraction of the particle diameter, so that a metastable state more or less corresponds to a given configuration of neighbors.

Fig. 5 (top) displays  $C_q(t_0, t)$  computed for a system of  $N = 1024$  bidisperse hard disks the dynamics of which has been simulated using molecular dynamics slightly below jamming ( $\phi = 0.841947$ ;  $\phi_J = 0.841959$ ).<sup>24,46</sup> Here  $a$  has been set to the radius of the smaller particles. One observes very well defined plateaus, during which the system is trapped in a metastable state, separated by a quick relaxation event that we call a *crack*. Usually, the size of the plateau defines  $\tau$ , the interval of time available to compute  $C_p$ . There are however some cases where one cannot identify the plateaus so easily. This will happen each time the system size is large compared to the typical size of a crack or spatial relaxation event. When this is the case, one can always cut the system into smaller subsystems and do the analysis for each subsystem as illustrated in ref. 39.

As long as the system remains in the same reference state, the basis of eigenvectors should essentially remain unchanged. This can be checked by dividing the time interval of duration  $\tau$  into smaller subintervals and computing an indicator of the modes robustness, defined as:<sup>39</sup>



**Fig. 5** Probing metastability. Top:  $C_q(t_0, t)$  for a system of hard spheres below jamming ( $N = 1024$  and  $\phi = 0.841947$ ;  $\phi_J = 0.841959$ ). The time unit is  $t_f$ , the time interval at which particle positions are stored, and corresponds to 800 collisions per particle. Each interval delimited by dashed lines has 200 snapshots, during which we compute  $C_p$  and its basis  $B_i$  of eigenvectors  $\{\lambda_m\}$ . Bottom: robustness of the first 20 modes as defined in eqn (17) between pairs of intervals  $B_i \times B_j$  as indicated in the legend. The dashed line indicates the value of  $R(m)$  for purely random eigenvectors.



$$R(m) = \sum_{j=-M}^{j=+M} \langle \lambda_m | \lambda'_{m+j} \rangle^2, \quad (17)$$

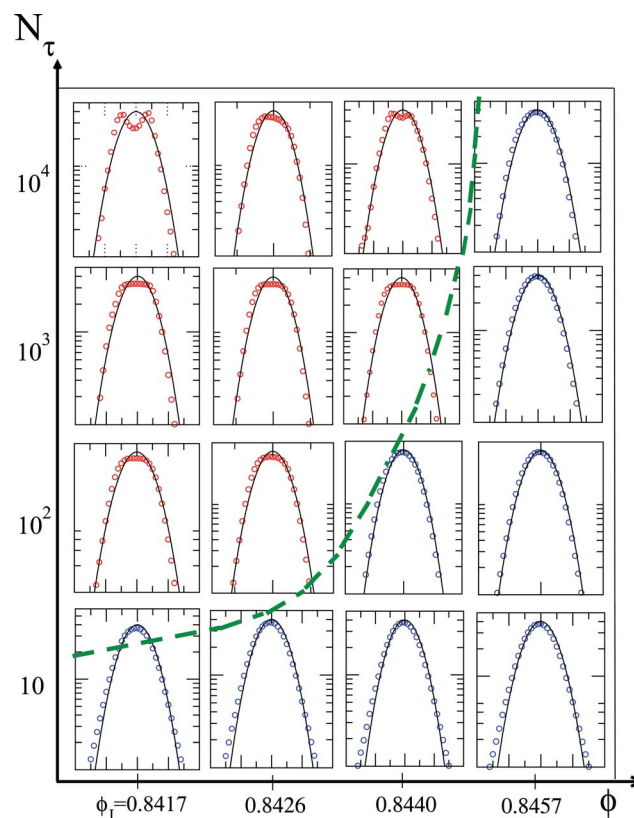
where  $\langle \lambda_m | \lambda'_{m+j} \rangle$  is the scalar product between the modes computed during two successive observation windows of duration  $N_\tau < N_\tau$ . Setting  $M$  to a small but strictly positive value allows neighboring modes to possibly exchange their rank. For the largest eigenvalues, we find that the results are basically independent of  $M$  for  $M \geq 2$  and we fix  $M = 2$ . Fig. 5 (bottom) displays  $R(m)$  for the twenty largest eigenvalues computed between intervals defined as indicated on Fig. 5 (top) by vertical dashed lines. The robustness of the first modes is higher when computed amongst intervals belonging to the same plateau. As soon as one computes the modes on an interval including the crack, the robustness decreases sharply and rapidly reaches the baseline level  $R(m) = (2M + 1)/2N$  expected for purely random vectors. Note also that the robustness is not restored when the two intervals belong to different plateaus indicating that the system has indeed relaxed from one metastable state to another.

## B. Harmonicity of the dynamics

In the previous section, we have identified a metastable state and checked *a posteriori* that the basis of eigenvectors remains unchanged during the lifetime of this state. Additionally we must also check that the dynamics around this state are purely harmonic. For an equilibrium system close to jamming, the structure of the packing is frozen and particle fluctuations are thermal, so that assuming Gaussian fluctuations around the metastable state sounds reasonable. However, it was recently claimed that repulsive contact interactions make jammed particles systems inherently anharmonic.<sup>42</sup> The authors argue that at very low temperature the breaking and forming of inter-particles contact is responsible for anharmonicity in the sense that the response does not remain confined to the original mode of excitation. Such a definition of harmonicity is very strict, and how the result depends on system size, temperature and distance to jamming is still a matter of debate. On the other hand, it is clear that for mechanically excited athermal and dissipative systems, such as shaken macroscopic grains, there is no reason why the dynamics should be harmonic.

While studying horizontally shaken grains, two of the present authors proposed to check the harmonicity of the dynamics in a naïve but experimentally accessible way, namely by computing the distribution of the *individual* fluctuations:  $\rho(\delta x_i / \sigma_i)$ , where  $\delta x_i$  is the position fluctuation along a given direction and  $\sigma_i$  is the root mean square displacement of particle  $i$ . Fig. 6 presents such distributions obtained for a system of  $N = 1550$  brass disks shaken horizontally on a oscillating plate just above jamming ( $\phi_J = 0.8417$ , see ref. 44 for details on the experimental set up and protocol). The distributions are computed for four different packing fractions  $\phi$  and four durations  $\tau$  of the window of observation. They are then ensemble averaged over the  $10^4/N_\tau$  intervals provided by the full dataset.

The distributions shown in Fig. 6 highlight some important characteristics of the dynamics. The parameter space  $(\phi, N_\tau)$  can be divided into two regions, as illustrated by the green hatched line: for small enough observation duration  $\tau$  or large enough packing fractions, the distributions are unimodal with a Gaussian



**Fig. 6** Distributions of the position fluctuations  $\rho(\delta x_i / \sigma_i)$  for four values of the packing fraction and four durations of observation  $\tau$ . We focus on the top of the distribution, which is compared to a Gaussian (continuous black line). The green hatched line separates the distributions well described by a Gaussian from those deviating from this harmonic description.

core: particles jiggle around a well defined average position; for longer  $\tau$  or smaller densities, the distribution starts developing a flat top, with a poorly defined maximum. This suggests that on these longer observation times, the average position of a significant part of the particles is not well defined any more. Particles either drift slowly or even find (collectively) another metastable position, as suggested by the double peak observed in the case  $\phi = 0.8417$  and  $N_\tau = 10^4$ , *i.e.* for the loosest packing fraction and the longest observation time. This means that over long time scales, the evolution of the average position becomes comparable or even larger than the fluctuations, and it becomes meaningless to describe the system in terms of small vibrations around a fixed metastable state. For an infinite size system, some rearrangement always happens somewhere, and the covariance matrix  $\mathbf{C}_p$  is always ill-defined. The “allowed” time scale  $(N_\tau)_{\max}(\phi, L)$  is expected to scale inversely with the system size  $L$ .

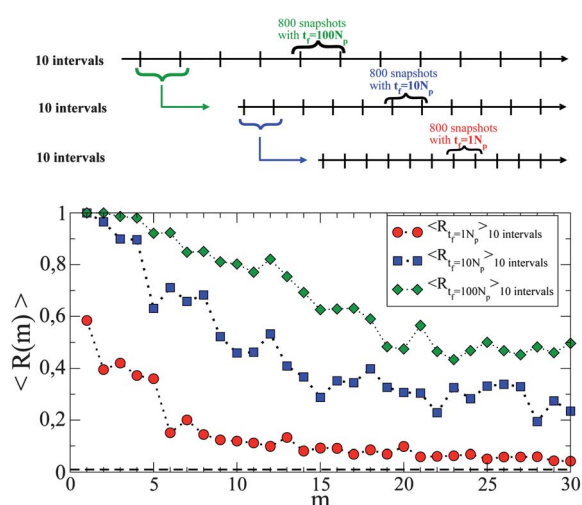
## C. Statistical independence

Once a reference state has been identified for a time window of duration  $\tau$ , one needs to know whether the particle positions stored in successive snapshots are independent, or at least sufficiently uncorrelated, to compute  $\mathbf{C}_p$ . Again, computing the robustness of the basis of eigenmodes is a good way to check the validity of the analysis: if the particle positions are not

independent the correlator  $C_p$  is badly estimated and the eigenmodes are mostly composed of noise. To illustrate this point we again use the hard discs system just below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ), now with  $N = 256$ , for which we could identify a long lasting metastable state the total duration of which,  $\tau$ , corresponds to  $1.6 \times 10^6$  collisions per particle.

Fig. 7 (top) sketches the procedure. The total acquisition is divided in 10 intervals. Each of them contains 800 snapshots of the positions, separated by  $t_f = 200$  collisions per particle.  $C_p$  is computed and diagonalized on each interval, and the average robustness among successive intervals  $\langle R(m) \rangle$  is evaluated for the first 20 largest eigenvalues. To preserve the same statistics, together with a reduced interval of time between the snapshots, we then subdivide each interval into 10 subintervals, each of which contain again 800 snapshots of the positions, but now separated by  $t_f = 20$  collisions per particle. Finally, the procedure is iterated once more leading to 800 snapshots of the positions separated by  $t_f = 2$  collisions per particle in each interval. Fig. 7 (bottom) compares  $\langle R(m) \rangle$  for the three successive levels of time discretization. Although the system remains in the same metastable state and the number of snapshots and the averaging used to compute  $C_p$  are identical in all three cases, one clearly observes the convergence of  $\langle R(m) \rangle$  towards larger values when the snapshots are well separated. In the present case the finest discretization is obviously irrelevant and a separation of successive snapshots larger than about 50 collisions per particle is necessary.

In summary, whether it arises from experimental limitations of the acquisition rate or from a microscopic timescale inherent to the dynamics, there is always a minimal timescale separating the accessible and useful snapshots of the particle positions. Since we have also seen that the analysis is confined to the lifetime of the reference state, we are now in the position to face the convergence issue of the eigenspectrum as discussed by the Marčenko–Pastur theorem.



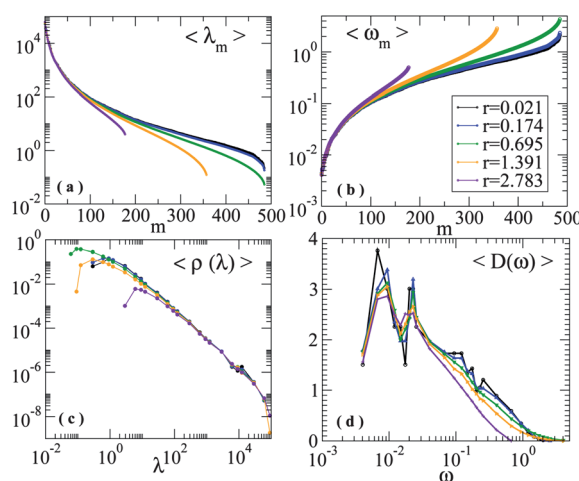
**Fig. 7** Testing statistical independence during a given metastable state for a system of  $N = 256$  hard disks below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ). Top: sketch of the time discretization procedure (see text for details). Bottom:  $\langle R(m) \rangle$  for the three discretization schemes as indicated by the legend.

#### D. Convergence of the spectrum (small sample limit)

The Marčenko–Pastur (M–P) theorem<sup>33,40,47,48</sup> applies to the ensemble of Wishart random matrices; that is the ensemble of  $p \times p$  matrices constructed by the addition of  $n$  independent samples  $x_i x_j$ , where the  $x_i$  ( $i \in 1 \dots p$ ) are distributed according to either a multivariate Gaussian distribution or a distribution with finite higher moments, similar to the conditions of the central limit theorem. The M–P theorem then predicts the eigenvalue distribution of such matrices in the limit  $n \rightarrow \infty$ ,  $p \rightarrow \infty$ ,  $r = p/n$  finite. However, as shown in Appendix VII D, there is no explicit form for the distribution. Also, the theorem does not touch the convergence of the eigenmodes. Consequently, we concentrate here on the numerical characterization of the convergence in a specific and well controlled situation. This will allow us to discuss both the shape of the spectrum and the relevance of the modes themselves. Also, it will allow us to examine the case where  $r > 1$ , when a finite part of the spectrum is strictly zero. The reader interested in analytical predictions may refer to the original papers,<sup>33,40,47,48</sup> as well as to Appendix VII D.

We consider the same system as above, with  $N = 256$  hard disks just below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ) within a long lasting metastable state of total duration  $4.6 \times 10^6$  collisions per particles and we choose to retain independent snapshots separated by 200 collisions per particle, so that  $n = N\tau = 23\,000$ . In Section III A, we showed that the spectrum of  $C_p$  and that of the effective dynamical matrix match when the whole set of data is used to compute  $C_p$ . Here we discuss how much the spectrum and the modes of  $C_p$  deviate from their asymptotic behavior when the time interval used to compute  $C_p$  is artificially reduced such that  $r = 2N/N\tau$  varies in the range  $[0.021-2.28]$ .

Fig. 8 displays the spectral properties of  $C_p$  for increasing values of  $r$ . As soon as  $r > 0.5$ , the smallest eigenvalues are seriously underestimated, respectively the largest frequencies are overestimated. For  $r > 1$ , strictly zero eigenvalues replace the lowest eigenvalues. On the contrary, the largest eigenvalues,



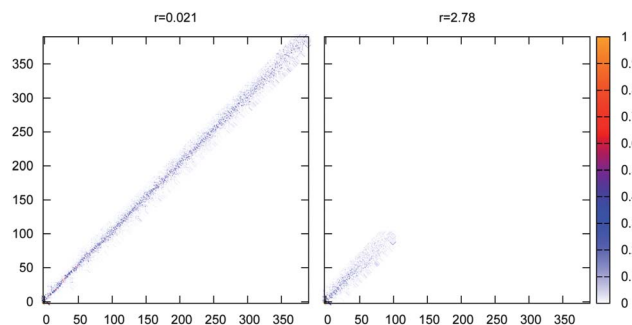
**Fig. 8** Convergence of the spectrum of eigenvalues of  $C_p$  for a system of  $N = 256$  hard disks below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ) for decreasing values of  $r = 2N/n$  as indicated in the legend. Top: eigenvalues (left) and corresponding eigenfrequencies (right) sorted in decreasing, respectively increasing order. Bottom: eigenvalues spectra (left) and corresponding density of states (right).

respectively the lowest eigenfrequencies, are surprisingly robust, even in the cases where  $r > 1$ . Because of this robustness, when one is interested in the low frequency part of the density of states, we recommend normalizing the density such that the area under the curve equals the fraction of non-zero eigenvalues. As a result, the density of states  $D(\omega)$  conserves its shape as long as  $r < 1.5$ , and only for larger  $r$  the large frequencies part of the density of states starts to be significantly depleted.

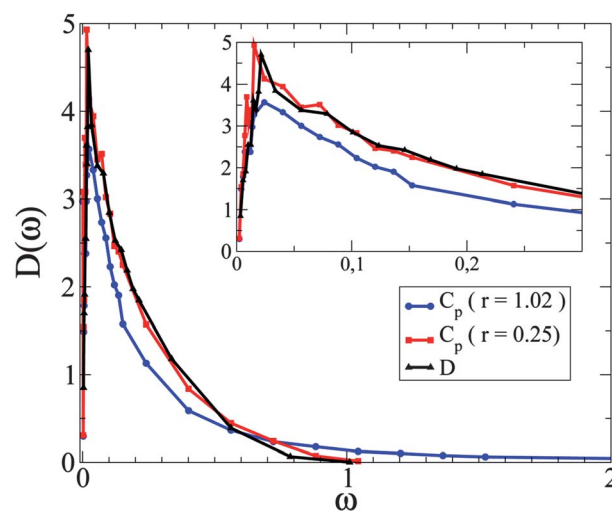
Finally, Fig. 9 illustrates the robustness of the modes themselves. We define the matrix of scalar products between the two eigenbases: the modes of  $C_p$  and the modes of the effective dynamical matrix. If those modes were strictly the same, this matrix should be the identity matrix. Fig. 9 (left) shows the matrix of the scalar products for  $r = 0.021$ , from where we observe that the modes of  $C_p$  project on a very small number of modes of the effective dynamical matrix, as already pointed at in Section III A. We see that even for  $r = 2.78$ , the modes of  $C_p$  associated with the largest eigenvalues still project on a small number of modes of the effective dynamical matrix, while the higher frequency modes of  $C_p$  are spread across the whole spectrum of the effective dynamical matrix.

Altogether, as long as  $r < 1.5$ , the most significant eigenmodes and eigenvalues of the dynamic matrix as well as the qualitative features of the density of states are well captured by the analysis of  $C_p$ . Practically speaking, when very long lasting metastable states are available it is thus often advantageous to average the spectrum or the density of states on successive time windows, provided that the duration of each of them satisfies  $r \approx 1$ .

When there are not enough snapshots to recover the whole spectrum of modes, an alternative strategy is to reduce the system size in order to increase  $r$ . In principle one can divide the system into smaller subsystems, and compute the density of states in each of them. This strategy allows better access to the low frequency part of the density of states, although the lowest accessible frequency increases as the inverse of the system size. Fig. 10 and Fig. 11 illustrate the application of such a strategy to a system of  $N = 1024$  particles below jamming ( $\phi = 0.841947$ ;  $\phi_J = 0.841959$ ). In this example, the metastable state has a total duration of  $1.6 \times 10^5$  collisions per particles and we store independent particle positions separated by 80 collisions per particle. This means that  $n = 2000$  snapshots and  $r = 1.02$ . From Fig. 8 we



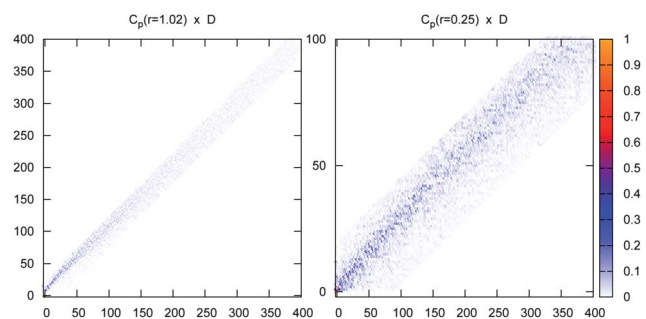
**Fig. 9** Matrices of the scalar product between the modes of  $C_p$  and the modes of the dynamical matrix for a system of  $N = 256$  hard disks below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ). The rows are the indices of the modes in increasing order of their frequency and the color bar indicates the value of the scalar product. Left:  $r = 2N/N_\tau = 0.021$  Right:  $r = 2N/N_\tau = 2.78$ .



**Fig. 10** Density of states for a system of hard disks below jamming ( $\phi = 0.841947$ ;  $\phi_J = 0.841959$ ) extracted from  $C_p$  either computed using the whole system of  $N = 1024$  hard disks (in blue) or cutting the system into 4 sub-systems with  $N_p^{\text{sub}} = 256$  each (in red) as compared to the density of states obtained from the effective dynamical matrix computed for the whole system (in black). Inset: zoom on the lowest frequency part of the density of states.

know that for this value of  $r$  the large frequencies are underestimated. By cutting the system into 4 boxes, each one has  $r = 0.25$  and we are now able to recover the whole spectrum. In Fig. 10 we compare the density of states  $D(\omega)$  extracted from  $C_p$  either using the whole system of  $N = 1024$  hard disks ( $r = 1.02$ ) or cutting the system into 4 sub-systems with  $N_p^{\text{sub}} = 256$  each ( $r = 0.25$ ) to the density of states obtained from the effective dynamical matrix computed for the whole system. Whereas the spectrum obtained from  $C_p$  computed with the whole system only captures the very low frequencies, the one computed for the sub-systems matches the spectrum from the dynamical matrix on the whole range of frequencies.

Concerning the robustness of the modes, once the modes have been computed in each subsystem, we first reconstruct artificial modes for the whole system by pasting together the modes of



**Fig. 11** Matrices of the scalar products between the modes of  $C_p$  and the modes of the dynamical matrix for a system of  $N = 1024$  hard disks below jamming. Left: modes of  $C_p$  computed on the whole system. Right: modes of  $C_p$  reconstructed from the modes computed separately in four sub-systems of size  $N_p^{\text{sub}} = 256$ . The horizontal (resp. the vertical) indices label the modes of the dynamical (reps. the covariance) matrix. The color bar indicates the value of the scalar product.

each subsystem. Note that the number of modes in these 2 bases is not the same: for the whole system there are  $2N_p = 2048$  modes while in each subsystem has  $2N_p^{\text{sub}} = 512$  modes. Fig. 11 shows the matrices of the scalar products between the modes of the dynamical matrix and those of  $C_p$ , either computed on the whole system (left) or reconstructed as described above from the modes computed on the subsystems (right). Despite the artificial nature of the modes obtained from the computation of  $C_p$  in each subsystem, their projection on the modes of the dynamical matrix remains rather concentrated, especially when one remembers that each mode of  $C_p$  in this case projects approximately on four modes of the dynamical matrix.

Let us end this section by citing recent work<sup>49</sup> which shows that in general the modes are modified by the truncation of the full real-space correlations, and proposes windowing strategies to best avoid these boundary condition artifacts. Implementing such strategies on top of the segmentation proposed here would certainly deserve proper attention.

### E. Convergence and experimental resolution (large sample limit)

Even in the most favorable case of all experimental situations discussed above, we have not yet addressed the issue of the experimental resolution. As detailed in Appendix VII D, the Marčenko–Pastur theorem allows us to compute the impact of finite resolution on the spectral properties of  $C_p$  at a finite  $r = 2N/n$ , where  $n = N_r$  is again the number of independent snapshots of the particle positions.

To do so we assume that the applicability conditions of the theorem detailed in Appendix VII D are met. The finite moments condition corresponds broadly speaking to a long-lived metastable state with short-time fluctuations of the particles in well-defined cages. The metastability condition can be tested through the methods of Section IV A. For an equilibrium system, the short-time fluctuations of the particles in their cages are thermal and can hence well be approximated as Gaussian within the range of linear response. For a non-equilibrium system such as shaken granular systems, this condition needs to be verified experimentally. The independence of the different samplings of the system can be tested through the methods of Section IV C.

If we have now constructed  $C_p$  in these conditions, how does the experimental resolution affect the properties of the matrix and how does it interplay with finite  $r$ ? How does it modify the spectral properties of the true correlation matrix, the one would obtain in the limits of infinite resolution and  $r \rightarrow 0$ ? We model the resolution as an additive Gaussian noise of variance  $\varepsilon^2$ , where  $\varepsilon$  is the resolution-induced uncertainty of the particle positions. If  $C_0$  is the true correlation matrix, *i.e.* the one of the shadow system, one easily shows that for Gaussian fluctuations of the particle positions, the finite resolution correlation matrix is  $\tilde{C}_0 = C_0 + \varepsilon^2 I$ , where  $I$  is the identity matrix. Then the measured  $C_p$  is  $\tilde{C}_r$ , such that  $\lim_{n \rightarrow \infty} \tilde{C}_r = \tilde{C}_0$ . The Marčenko–Pastur theorem is still valid and one finally relates the eigenvalues distribution of  $\tilde{C}_r$  and the true correlation matrix  $C_0$ .

As for the case without noise, there is no explicit form for the distributions, but we show in Appendix VII D that the mean and variance of the eigenvalue distribution of  $C_p$  converge with  $r \rightarrow 0$  and  $\varepsilon/\mu_0 \rightarrow 0$  in the following way:

$$\begin{aligned} \tilde{\mu}_r &= \mu_0 + \varepsilon^2 \\ \tilde{\sigma}_r^2 &= \sigma_0^2 + r\mu_0^2 + r\varepsilon^4, \end{aligned} \quad (18)$$

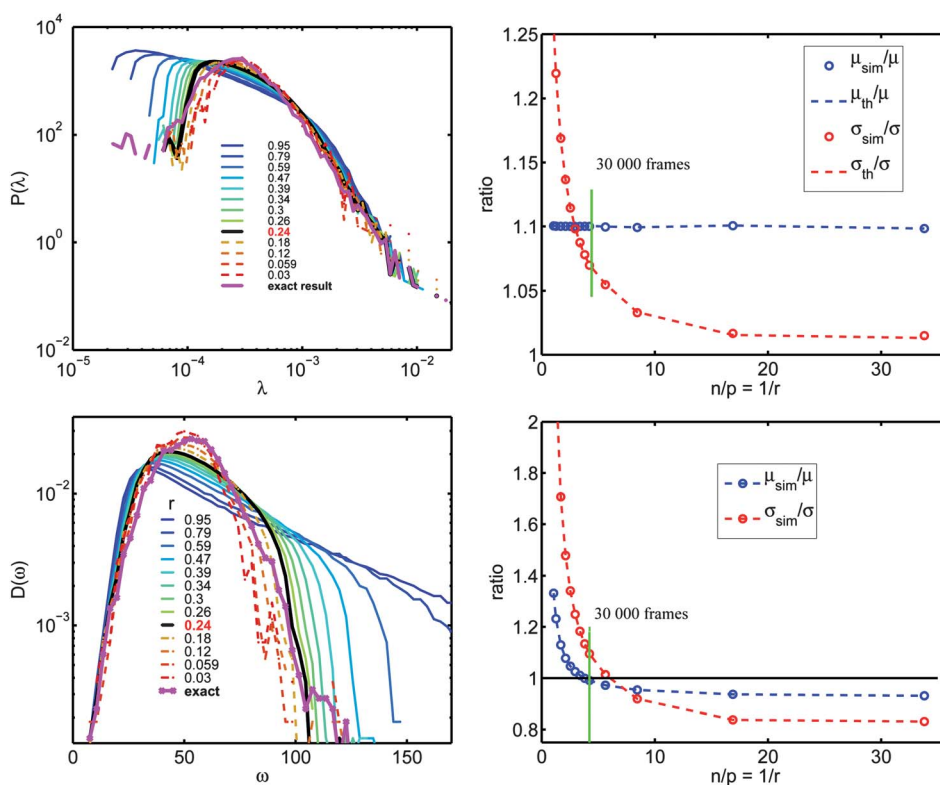
where the quantities labeled 0 refer to the exact eigenvalue distribution and the ones labeled  $r$  to the measured quantities. Note that for uncorrelated particles  $\mu_0$  would be the typical cage size squared, *i.e.* the value of the plateau in the mean square displacement curves, preceding the onset of diffusive behavior. Interestingly, while the resolution directly impacts the mean, it only changes the width at  $\mathcal{O}(\varepsilon^4)$ , and the correction is moreover multiplied by  $r$ . As a result, the noise dependence of the eigenvalue distribution is in fact much weaker than one might have expected.

We now present the results of a statistical model specifically developed to test the validity of the experimental results for colloidal NIPA particles.<sup>13</sup> In the experiment,  $N = 3600$  particles forming quasi two dimensional packings deep in the glassy phase were imaged with a confocal microscope, yielding  $n = 30\,000$  independent samples of their positions per run. In the language of the Marčenko–Pastur theorem, this corresponds to  $r = 0.24$ . The combination of optical resolution and sub-pixel accuracy particle tracking led to an estimated uncertainty in the particle positions of  $\varepsilon = 0.007 \mu\text{m}$ .

We build a ‘random vector’ model by constructing a model  $C_p$  where we repeatedly sample from an uncorrelated multivariate normal distribution corresponding to the experimentally measured cages sizes, and with added noise at the experimental amplitude. The covariance matrix of the model is then  $\tilde{C}_{ij} = (\lambda_i + \varepsilon^2)\delta_{ij}$ , where the  $\lambda_i$  are the diagonal elements of one of the measured  $C_p$  from ref. 13. While this model resembles the ‘Random Model’ introduced in Section III B detailing PCA, our aim here is different: we only intend to illustrate eqn (18) with a model where we know the exact eigenvalue distribution in the absence of noise and for  $r \rightarrow 0$ . By construction, this is just  $P(\lambda_i)$ , the distribution of diagonal elements or cage sizes which enter the model.

Fig. 12 (top-left) shows the convergence of the numerical eigenvalue distribution  $\tilde{P}_r(\lambda)$  as  $r \rightarrow 0$  at fixed noise. The eigenvalue distribution corresponding to the experimental value  $r = 0.24$  is shown in black. It appears to be relatively close to the exact distribution shown in pink. The numerically obtained first two moments of the distributions exactly match the theoretical result from eqn (18) (Fig. 12 (top-right)). The curves indicate that for the experimental value of  $r$  (green vertical line), the error made compared to the true converged noiseless distribution is of the order of 10%.

For equilibrium systems, one is mostly interested in the density of states. Unfortunately, when performing the change of variables  $\omega = \sqrt{k_B T/m\lambda}$ , there is no straightforward and reliable way to calculate the moments of the distribution and we have to rely mostly on numerical results. Only the asymptotic values in the limit  $r \rightarrow 0$  could be obtained (see Appendix VII D). In the bottom row of Fig. 12 we have numerically applied the transformation  $\omega = 1/\sqrt{\lambda}$  to obtain the density of states and its first two moments for the ‘random vector’ model. The density of states obtained for the experimental value of  $r = 0.24$  again matches pretty well the asymptotic one. Note that we find a fortuitous compensation in the moments of the distributions: in the limit  $r \rightarrow 0$ , one can show that a finite



**Fig. 12** Statistical convergence analysis for the NIPA system,<sup>13</sup> using the ‘random vector’ model (see text). Top left: numerical eigenvalue distribution for different values of  $r$  as indicated in the legend. The distribution associated with the experimental value  $r = 0.24$  is plotted in black and the asymptotic  $r \rightarrow 0$  one is plotted in pink. Bottom left: same for the density of states. Top right: numerical (dots, subscript ‘sim’) and theoretical (dashed lines, subscript ‘th’) results for the mean and variance of  $P(\lambda)$  normalized by the exact result. The experimental value of  $r = 0.24$  is indicated by a green line. Bottom right: same for the density of states, numerical result only (dots and dashed lines, subscript ‘sim’). The green line marked ‘30 000 frames’ corresponds to the number of samples in the experimental NIPA system.<sup>13</sup>

resolution systematically leads to an underestimation of both the mean and the variance. Since working at finite  $r$  always leads to an overestimation of these moments, it compensates the above underestimation. In the present case, the experimental value of  $r$  (green line) is very close to the point where the compensation is perfect. Note that this is pure coincidence related to the specific values of the moments of the true distribution, and that even so it does not validate the detailed shape of the distribution.

In the case of the experimental data obtained with NIPA particles,<sup>13</sup> and for which correlations are present, the authors tested convergence of the density of states based on the methods outlined here and found scaling of mean and width consistent with our example. The error from experimental and statistical resolution was estimated in the 10% range at  $r = 0.24$ .

We can make a similar comparison for the hard sphere colloid setup of Ghosh *et al.*<sup>35</sup> For this experiment, the estimated number of particles is  $N = 2000$ , and approximately  $n = 4000$  samples at a resolution of  $\varepsilon = 0.03 \mu\text{m}$  were taken to obtain the DOS. While the resolution effect is comparable to the NIPA setup, we obtain  $r \approx 1$ . This introduces a large error into the measurement of the DOS, similar to the  $r = 0.95$  curves in the left panels of Fig. 12. We estimate an error of 40% and 200% on the mean and the width of the DOS, respectively. However, as detailed in the previous section, even at this  $r$  it is still possible to obtain the relevant low energy modes.<sup>50</sup>

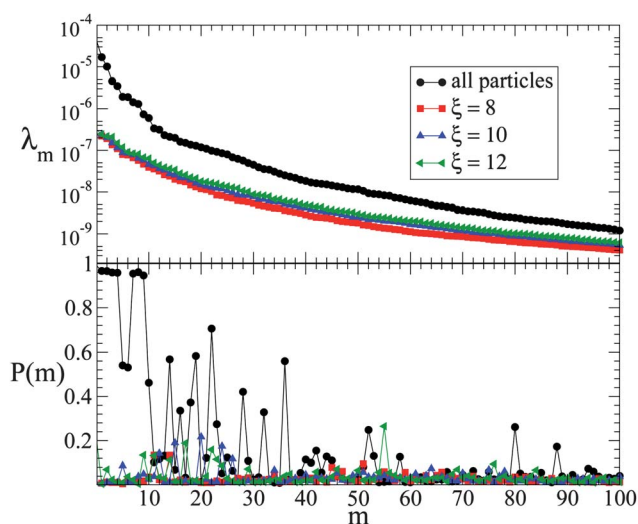
## F. Rattlers

Above jamming, there are particles, the so-called rattlers, which do not participate to the rigidity of the structure: removing them does not alter mechanical stability. Rattlers compose a finite and increasing fraction of particles when approaching jamming. The vibrational properties of rattlers are well known: they coincide with modes of zero frequency entirely localized on the rattler. When computing, for instance, the average coordination of the packing, these rattlers need to be identified and excluded. For simulations of soft spheres, the procedure is straightforward. The number of contacts is known for each particle and those which have less than two contacts are considered as rattlers. For simulations of hard spheres just below jamming, the contacts are defined by computing the transfer of momentum during successive collisions. Again, particles with less than two contacts are defined as rattlers, but additionally also those which have a significant smaller momentum transfer per contact as compared to the average. Such a criterion was proposed in ref. 51, where it was shown that particles with momentum transfer per contact smaller than 2% of the average value could safely be considered as rattlers. Varying this threshold in the range 0.5–5% did not alter the results.

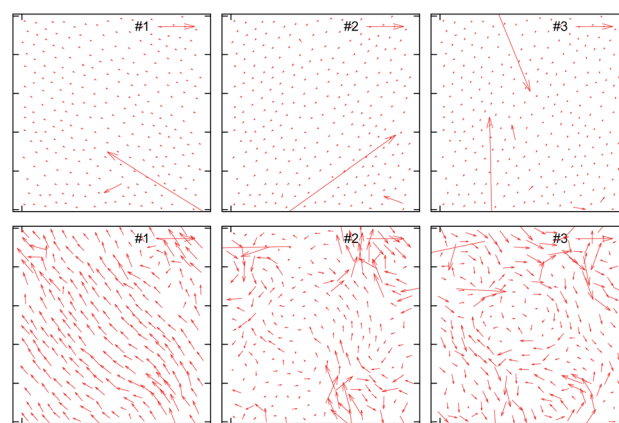
We now offer an alternative way to identify rattlers in experiments by using the Principal Component Analysis (PCA) itself

and compare it to the above method used for hard spheres. As a matter of fact, the PCA diagonalizes the dynamics on its dominant features. Since the rattlers are not tightly confined by their neighbors, they move significantly more than other particles. Also such motion is by definition localized on a single particle. Hence rattlers correspond to a very singular kind of excitation and thereby tend to jeopardize the modes structure. Diagonalizing  $C_p$  and analyzing the modes with high eigenvalues and a very high participation ratio is thus a good strategy to identify rattlers. Here the participation ratio of a given mode  $m$  is simply defined as  $P(m) = \sum_{i=1}^N (\delta r_i^m)^4$ , where  $\delta r_i^m$  is the displacement of the particle  $i$  on the mode  $m$ . Given that the modes are normalized,  $\sum_{i=1}^N (\delta r_i^m)^2 = 1$  so that a perfectly delocalized mode has a participation ratio of  $1/N$ , whereas a mode fully localized on a single particle would have a participation ratio of order one.

The method is illustrated in the case of a system of  $N = 256$  hard disks just below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ) within a long lasting metastable state of total duration of  $46 \times 10^6$  collisions per particles with  $n = 23\,000$  independent snapshots and  $r = 0.021$ . Fig. 13 shows the spectrum and the participation ratio of the modes of  $C_p$  computed with and without rattlers. One easily identifies the modes with a simultaneously large eigenvalue and high participation ratio. Three of these modes are plotted on Fig. 14 (top). Here also, one immediately identifies the rattlers with huge displacements as compared to the other particles. The rattlers in each modes are then defined as the particles having a displacement amplitude larger than  $\xi$  times the average displacement inside the mode:  $\delta r_i^m \geq \langle \delta r^m \rangle + \xi \sigma_{\delta r^m}^m$ , with  $\sigma_{\delta r^m}^m$  being the standard deviations of the displacement on the mode  $m$ . One must fix a threshold; however because the modes selected by the procedure concentrate most of the motion on the rattlers, it is very easy to fix the threshold and the results are robust to the choice of  $\xi$ , as we verified in Fig. 13. Once the rattlers are identified—sometime in



**Fig. 13** Eigenvalues of  $C_p$  and participation ratio as defined in the text for a system of  $N = 256$  hard disks below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ) with its rattlers (black curves) and having removed them according to different values of the threshold criteria  $\xi$  as indicated in the legend. See text for the definition of  $\xi$ .



**Fig. 14** Eigenmodes of  $C_p$  for a system of  $N = 256$  hard disks below jamming ( $\phi = 0.83878$ ;  $\phi_J = 0.838865$ ) with its rattlers (top row) and having removed them (bottom row).

several mode—one eliminates them and recomputes  $C_p$ . One then easily checks that the extremely localized modes have disappeared, together with the rattlers, Fig. 14 (bottom). Note that most of the particles identified as rattlers using this criterion are identical to those identified with the previous criterion used in ref. 51. The mismatch between both criteria is of the order of 10% of the rattlers identified.

## V. Conclusions

We have reviewed various situations for which the study of  $C_p$ , the covariance matrix of the positions of a system of particles, can be used to infer dynamical properties of the system. It has been emphasized that for systems without the equipartition of energy, or without an independent knowledge of the energy distribution amongst the modes, the spectrum of  $C_p$  cannot be transposed into the density of states of the system. Still, it conveys relevant information about the dynamics. We also insist on the physical interpretation of the modes obtained *via* the analysis in terms of the vibrational properties of the *shadow system*, which must not be confused with the experimental system, even an idealized one.

We have then reviewed the numerous steps one has to go through to implement the analysis. Metastability of the reference state, harmonicity of the dynamics, statistical independence, convergence and experimental resolutions are all key elements. How to deal with rattlers was finally discussed. Following the whole procedure requires, as always, a bit of physical insight, a good knowledge of the system of interest and patience.

We have *not* discussed the new physics one can eventually extract from this analysis. It depends on the specifics of each system, especially the range of linear response and the particular dynamics; but it also depends on the aim of the investigation. If one insists on predicting the dynamics precisely, the approach is limited. However, as detailed above, the largest eigenvalues, along with their associated eigenvectors, keep their significance and even at the level of PCA, a wealth of useful information can be extracted from  $C_p$  or the dynamical matrix. We believe dynamical matrix calculations in the harmonic approximation will remain an important tool for the analysis of

systems near jamming. We hope that in this light the present review of the methods associated to the correlation matrix approach will be useful to the community.

## VI. Appendix

### A. Newtonian dynamics

Starting from the Newton equations linearized around the reference state  $|r^0\rangle$ ,

$$|\delta\ddot{r}\rangle + \mathbf{D}|\delta r\rangle = 0, \quad (19)$$

where  $\mathbf{D} = \mathbf{K}/m$  is called the *dynamical matrix*, the eigenmodes of  $\mathbf{D}$ , also called the *vibrational modes* of the system are defined by  $\mathbf{D}|\lambda_q\rangle = \omega_q^2|\lambda_q\rangle$ , where  $\omega_q$  are the vibrational frequencies. The solution of eqn (19) is given by:

$$|\delta r\rangle = e^{-i\sqrt{\mathbf{D}}t}|\delta r(0)\rangle, \quad (20)$$

where  $|\delta r(0)\rangle$  is the displacement field at time zero. Replacing this solution in eqn (1), one has

$$\mathbf{C}_p = \left\langle e^{i\sqrt{\mathbf{D}}t}|\delta r(0)\rangle \left\langle \delta r(0)|e^{i\sqrt{\mathbf{D}}t}\right\rangle; \quad (21)$$

then writing  $|\delta r(0)\rangle$  in the eigenbasis of  $\mathbf{D}$ ,  $\{|\lambda_q\rangle\}$ :

$$|\delta r(0)\rangle = \sum_q \alpha_q |\lambda_q\rangle, \quad (22)$$

where  $\alpha_q = \langle \lambda_q | \delta r(0) \rangle$ , is the amplitude of the initial condition on the mode  $|\lambda_q\rangle$ , one obtains

$$\mathbf{C}_p = \left\langle \sum_{q,k} \alpha_q \alpha_k^* e^{-i(\omega_q - \omega_k)t} |\lambda_q\rangle \langle \lambda_k| \right\rangle \quad (23)$$

Provided that the time average  $\langle \cdot \rangle$  is performed on a large enough interval, large as compared to the inverse of the minimal gap between adjacent frequencies ( $\omega_q - \omega_m$ ), one has

$$\left\langle e^{-i(\omega_q - \omega_m)t} \right\rangle = \delta_{q,m}, \quad (24)$$

and

$$\mathbf{C}_p = \sum_k \alpha_k^2 |\lambda_k\rangle \langle \lambda_k|. \quad (25)$$

Using the orthogonality of the eigenvectors  $\{|\lambda_q\rangle\}$ ,  $\langle \lambda_k | \lambda_m \rangle = \delta_{k,m}$ , one finally obtains the eigenvalue equation of  $\mathbf{C}_p$ :

$$\mathbf{C}_p |\lambda_q\rangle = \alpha_q^2 |\lambda_q\rangle. \quad (26)$$

At equilibrium, the initial condition is thermalized and the energy is equally distributed among the modes. Each mode of frequency  $\omega_q$  and amplitude  $\alpha_q$  carries an energy  $m\alpha_q^2\omega_q^2/2 = k_B T/2$ . Hence the eigenvalues  $\lambda_q$  of  $\mathbf{C}_p$  and the vibrational frequencies  $\omega_q$  of the dynamical matrix are related through

$$\lambda_q = \alpha_q^2 = \frac{k_B T}{m\omega_q^2}. \quad (27)$$

### B. Overdamped Langevin dynamics

Starting from the Langevin equation for an overdamped linearized dynamics around the reference state  $|r^0\rangle$ :

$$|\delta\dot{r}(t)\rangle = -\frac{\mathbf{K}}{\mu}|\delta r\rangle + \frac{1}{\mu}|\eta(t)\rangle, \quad (28)$$

where  $\mu$  is the viscous damping and  $|\eta(t)\rangle$  is a white noise of amplitude  $\Gamma$ :  $\langle \eta(t')\eta(t'') \rangle = \Gamma\delta(t' - t'')$ , one introduces the operator  $\mathcal{S} = \mathbf{K}/\mu$  whose eigenvalue equation is:  $\mathcal{S}|\lambda_q\rangle = m\omega_q^2/\mu|\lambda_q\rangle$ .  $\tau_q = \mu/\kappa_q = \mu/(m\omega_q^2)$  is the relaxation time of the system along the eigenmode  $|\lambda_q\rangle$ .

The solution of the eqn (28) can be written as

$$|\delta r\rangle = e^{-\mathcal{S}t}|\delta r(0)\rangle + \frac{1}{\mu} \int_0^t e^{-\mathcal{S}(t-t')}|\eta(t')\rangle dt' \quad (29)$$

Replacing this solution 29 in eqn (1) one obtains an expression with four terms, only two of which are not zero because the noise  $|\eta(t)\rangle$  is assumed not to be correlated with the position field  $|\delta r(0)\rangle$ . One ends up with:

$$\begin{aligned} \mathbf{C}_p &= \left\langle e^{-\mathcal{S}t}|\delta r(0)\rangle \left\langle \delta r(0)|e^{-\mathcal{S}t}\right\rangle \right. \\ &+ \frac{1}{\mu^2} \left\langle \int_0^t dt'' \int_0^t dt' e^{-\mathcal{S}(t-t')}|\eta(t')\rangle \langle \eta(t'')|e^{-\mathcal{S}(t-t'')} \right\rangle \end{aligned} \quad (30)$$

Expressing  $|\delta r(0)\rangle$  and  $|\eta(t)\rangle$  in the eigenbasis  $\{|\lambda_q\rangle\}$  of  $\mathcal{S}$ :  $|\delta r(0)\rangle = \sum_q \alpha_q |\lambda_q\rangle$  and  $|\eta(t)\rangle = \sum_q \beta_q(t) |\lambda_q\rangle$ , with  $\alpha_q = \langle \lambda_q | \delta r(0) \rangle$  and  $\beta_q = \langle \lambda_q | \eta(t) \rangle$ , the first term of the previous equation, which we note  $T_1$  turns into:

$$T_1 = \left\langle \sum_{q,k} e^{-(1/\tau_q + 1/\tau_k)t} \alpha_q \alpha_k^* |\lambda_q\rangle \langle \lambda_k| \right\rangle = \sum_k e^{-2t/\tau_k} \alpha_k^2 |\lambda_k\rangle \langle \lambda_k|,$$

where the last equality is justified by the fact that the components on each mode of the initial conditions are uncorrelated  $\langle \alpha_q \alpha_k^* \rangle = \alpha_q \alpha_k^* \delta_{k,q}$ .

We now turn our attention to the second term in the eqn (30), referred to as  $T_2$ :

$$\begin{aligned} T_2 &= \frac{1}{\mu^2} \left\langle \int_0^t dt'' \int_0^t dt' \sum_{q,k} e^{-1/\tau_q(t-t')} \beta_q(t') |\lambda_q\rangle \langle \lambda_k | \beta_k(t'') e^{-1/\tau_k(t-t'')} \right\rangle \\ &= \int_0^t dt'' \int_0^t dt' \sum_{q,k} e^{-(t-t')/\tau_q - (t-t'')/\tau_k} \beta_q(t') \beta_k(t'') |\lambda_q\rangle \langle \lambda_k| \end{aligned}$$

Now, assuming that the components of the noise on the modes are also uncorrelated,  $\langle \beta_q(t') \beta_k(t'') \rangle = \Gamma \delta_{q,k} \delta(t' - t'')$  one obtains for  $T_2$ :

$$\begin{aligned} T_2 &= \frac{\Gamma}{\mu^2} \sum_k \int_0^t dt' e^{-2(t-t')/\tau_k} |\lambda_k\rangle \langle \lambda_k| \\ &= \frac{\Gamma}{\mu^2} \sum_k \frac{\tau_k}{2} (1 - e^{-2t/\tau_k}) |\lambda_k\rangle \langle \lambda_k| \end{aligned}$$

Finally applying  $\mathbf{C}_p$  on the eigenvector  $|\lambda_q\rangle$  one obtains the eigenvalue equation for  $\mathbf{C}_p$ :

$$\mathbf{C}_p |\lambda_q\rangle = \left[ \left( \alpha_q^2 - \frac{\Gamma \tau_q}{2\mu^2} \right) e^{-2t/\tau_q} + \frac{\Gamma \tau_q}{2\mu^2} \right] |\lambda_q\rangle \quad (31)$$

### C. Out of equilibrium stochastic forcing

To generalize on the two equilibrium cases of Newtonian dynamics and Langevin dynamics, we now consider

a non-equilibrium system with both inertia and damping and a colored noise spectrum. We choose to stick to a Langevin type description of the dynamics, where the damping is linear and single particle, as appropriate for particles in a Newtonian fluid bath, but not necessarily for a large scale mechanical excitation. The generic equations of motion will then be:

$$m|\delta\dot{r}(t)\rangle + \mu|\delta r(t)\rangle = -\mathbf{K}|\delta r(t)\rangle + |\eta(t)\rangle \quad (32)$$

We now expand in the modes of  $K$ , using the same notation as for the Langevin case,  $\mathbf{K}|\lambda_q\rangle = \kappa_q|\lambda_q\rangle$ , and  $|\delta r(t)\rangle = \sum_q \alpha_q(t)|\lambda_q\rangle$ . We define the colored noise in the basis of the modes as

$$|\eta(t)\rangle = \sum_q \eta_q(t)|\lambda_q\rangle \quad (33)$$

with  $\langle \eta_q(t)\eta_k(t') \rangle = \Gamma_q \delta_{qk} \delta(t-t')$ .

In the eigenbasis of the modes, the equations of motion are:

$$\ddot{\alpha}_q(t) + \frac{\mu}{m}\dot{\alpha}_q(t) + \frac{\kappa_q}{m}\alpha_q(t) = \eta_q(t). \quad (34)$$

We first solve the homogeneous equation, in the absence of noise. The solutions are either oscillatory or overdamped depending on the relative importance of inertia and damping.

For  $(\mu/m)^2 < 4\kappa_q/m$ , i.e. the low damping limit, we find two oscillating solutions

$$\alpha_q^s(t) = \alpha_q^s(0)e^{-\frac{\mu}{2m}t} \sin\left[\frac{1}{2}\tilde{\Omega}_q t\right]$$

$$\alpha_q^c(t) = \alpha_q^c(0)e^{-\frac{\mu}{2m}t} \cos\left[\frac{1}{2}\tilde{\Omega}_q t\right]$$

$$\tilde{\Omega}_q = \sqrt{4\frac{\kappa_q}{m} - \left(\frac{\mu}{m}\right)^2} \quad (35)$$

Clearly, the limit  $\mu/m \rightarrow 0$  corresponds to Newtonian dynamics.

For  $(\mu/m)^2 > 4\kappa_q/m$ , i.e. the strong damping limit, we obtain two decaying solutions

$$\alpha_q^\pm(t) = \alpha_q^\pm(0) \exp\left\{\frac{1}{2}\left[\frac{-\mu}{m} \pm \Omega_q\right]t\right\} \quad (36)$$

$$\Omega_q = \sqrt{\left(\frac{\mu}{m}\right)^2 - 4\frac{\kappa_q}{m}}$$

It can be shown that in the limit  $m/\mu \rightarrow 0$ , the '+' solution corresponds to the Langevin homogeneous solution  $\alpha_q(t) = \alpha_q^+(0)e^{-\kappa_q t/m}$ , while the '-' solution decays infinitely fast.

To solve the inhomogeneous equation, we find a particular solution  $\alpha_q^p(t)$  through the method of variation of constants based on the homogeneous solutions. In the oscillatory regime, we find

$$\alpha_q^p(t) = \int_0^t \frac{\eta_q(t')}{m\tilde{\Omega}_q} e^{\frac{\mu}{2m}(t-t')} 2\sin\left(\frac{\tilde{\Omega}_q}{2}(t-t')\right) dt' \quad (37)$$

while for the damped regime, the particular solution is given by:

$$\alpha_q^p(t) = \int_0^t dt' \frac{\eta_q(t')}{m\Omega_q} e^{\frac{\mu}{2m}(t-t')} 2\sinh\left(\frac{\Omega_q}{2}(t-t')\right) \quad (38)$$

Both solutions are similar in spirit to the particular solution of the Langevin equation: the source of continued motion in the system is the colored noise, and in the long time limit, *only* the

motion due to the noise remains. From this point of view, the memory a non-dissipative system such as the Newtonian case retains of the initial conditions is singular.

We can now calculate  $C_p$  by performing an ensemble average over initial conditions and the noise. In particular, we assume that we can replace the noise correlations by their expectation value  $\langle \eta_q(t)\eta_k(t') \rangle = \Gamma_q \delta_{qk} \delta(t-t')$ . We also assume that noise and initial conditions do not cross-correlate, and that in ensemble-average, the initial conditions of different modes are independent of each other. Then only the diagonal terms in the modes remain and, schematically,  $C_p$  is given by:

$$C_p(t) = \sum_q \langle \alpha_q^1(t)^2 \rangle + \sum_q \langle \alpha_q^2(t)^2 \rangle + \sum_q \langle \alpha_q^p(t)^2 \rangle \quad (39)$$

For the oscillatory case, we obtain

$$C_p = \sum_q e^{-\mu t/m} \left[ \langle (\alpha_q^s(0))^2 \rangle \sin^2(\tilde{\Omega}_q t/2) + \langle (\alpha_q^c(0))^2 \rangle \cos^2(\tilde{\Omega}_q t/2) \right] + \frac{\Gamma_q}{2\mu\kappa_q} - \frac{2\Gamma_q e^{-\mu t/m}}{m\mu\tilde{\Omega}_q^2}$$

$$- \frac{\mu}{m} \frac{\Gamma_q e^{-\mu t/m}}{2\mu\kappa_q \tilde{\Omega}_q^2} \left[ \tilde{\Omega}_q \sin(\tilde{\Omega}_q t) - \frac{\mu}{m} \cos(\tilde{\Omega}_q t) \right] |\lambda_q\rangle \langle \lambda_q| \quad (40)$$

For the damped case, we find

$$C_p = \sum_q e^{-\mu t/m} \left[ \langle (\alpha_q^+(0))^2 \rangle e^{\Omega_q t} + \langle (\alpha_q^-(0))^2 \rangle e^{-\Omega_q t} \right]$$

$$+ \frac{\Gamma_q}{2\mu\kappa_q} (1 + e^{-\mu t/m}) - \frac{\mu}{m} \frac{\Gamma_q e^{-\mu t/m}}{4\mu\kappa_q \Omega_q^2} \left[ \Omega_q (e^{\Omega_q t} - e^{-\Omega_q t}) \right]$$

$$+ \frac{\mu}{m} (e^{\Omega_q t} + e^{-\Omega_q t} - 2) \left] |\lambda_q\rangle \langle \lambda_q| \quad (41)$$

which in the over-damped limit  $m/\mu \rightarrow 0$ , defaults to the result for Langevin dynamics if we assume equilibrium noise  $\Gamma_q = \Gamma$ ,  $\forall q$ .

In both cases, in the long time limit we recover the an expression similar to that of the equilibrium result

$$C_p |\lambda_q\rangle = \frac{\Gamma_q}{2\mu\kappa_q} |\lambda_q\rangle \quad (42)$$

where the initial conditions cease to matter.

#### D. Sampling, resolution and the Marčenko–Pastur theorem

We derive relations which characterize the deviation induced by finite sampling *and finite resolution* on the spectral properties of  $C_p$ . For a system of  $N$  particles the positions of which are sampled independently  $n$  times with a resolution  $\varepsilon$ , we will establish an adapted form of the Marčenko–Pastur theorem, which then allows the derivation of relations between the moments of the experimentally measured eigenvalue distribution and the moments of the eigenvalue distribution of the true correlation matrix. The sampling number  $r = Nd/n$ , is defined as the total number of degrees of freedom  $(Nd)^2$  divided by the number of measurements  $n \times Nd$ .

In its most general form, the Marčenko–Pastur theorem<sup>40,47,48,52</sup> is valid on the ensemble of Wishart random matrices, that is matrices constructed as  $C_r = X^* X$ , where



$X$  is a  $n \times p$  matrix which can be written as  $X = YC_0^{1/2}$  where the elements of the  $n \times p$  matrix  $Y$  are identically independently distributed (i.i.d), with mean 0, variance 1 and a finite fourth moment and  $C_0$  is a  $p \times p$  positive definite matrix. In the following, we shall use the slightly more specialized derivation for Gaussian distributions found in Burda *et al.*<sup>52</sup> We thus assume that  $C_r$  is the average of  $n$  samplings of  $x_i x_j$ , where the variables  $\{x_i\}$ ,  $i = 1 \dots p$  are identically and independently distributed according to a multivariate normal distribution:

$$P(x_1, \dots, x_p) = \left[ (2\pi)^p \det(C_0) \right]^{-1/2} \exp \left[ -\frac{1}{2} x_i C_{0ij}^{-1} x_j \right]. \quad (43)$$

In the limit  $n \rightarrow \infty$ , at fixed  $p$ , one then clearly has from eqn (43)  $\lim_{r \rightarrow 0} C_r = C_0$ .

For a given correlation matrix  $C$ , define the (negative) moment generating function of the eigenvalue spectrum  $P(\lambda)$

$$m(z) = \sum_{k=1}^{\infty} \frac{m_k}{z^k}, \quad (44)$$

where  $m_k = \langle \lambda^k \rangle$ . Also define the Stieltjes transform of the eigenvalue distribution  $P(\lambda)$  written as a sum over the poles of  $(C - zI)^{-1}$  at the eigenvalues

$$s(z) = \frac{1}{p} \text{Tr} \left( (C - zI)^{-1} \right), \quad z \in C^+. \quad (45)$$

Both are related by

$$m(z) = -zs(z) - 1. \quad (46)$$

To see this, note that in the eigenbasis of  $C$ , we can write the Laurent series of the right hand side as

$$\sum_{k=1}^{\infty} \frac{1}{p} \sum_{l=1}^p \frac{\lambda_l^k}{z^k}. \quad (47)$$

Let us denote  $P_0(\lambda)$ ,  $m_0(z)$ ,  $s_0(z)$  the eigenvalue spectrum, the generating function and the Stieltjes transform of the true correlation matrix  $C_0$ , and  $P_r(\lambda)$ ,  $m_r(z)$ ,  $s_r(z)$  the corresponding quantities for the experimental  $C_r$ . Then the Marčenko–Pastur theorem stated in Burda *et al.*<sup>52</sup> provides an explicit conformal transformation relating  $m_r(z)$  to  $m_0(z)$ :

$$m_r(z_r) = m_0(z_0) \quad \text{where} \quad z_0 = \frac{z_r}{1 + r m_r(z_r)}. \quad (48)$$

Burda *et al.* then derive relations between the moments  $m_k^r$  and  $m_k^0$  through writing the Laurent series in  $z_r$  on both sides, and then equating the coefficients of the powers of  $z_r^{-k}$ .

We can adapt this approach to include an experimental Gaussian noise term as follows. If  $C_0$  is the true correlation matrix, let  $y_i = \delta x_i + \xi_i$  be the particle fluctuations with the resolution-induced noise included, so that we have  $P(\xi_i) = [2\pi\epsilon^2]^{-1/2} \exp(-\xi_i^2/\epsilon^2)$ . The probability distribution of a sum of two random variables is obtained through a convolution of their respective probability distributions, so that we have

$$P(y_1, \dots, y_p) = \prod_{i=1}^p \int d\xi_i P(\xi_i) P(y_1 - \xi_1, \dots, y_p - \xi_p). \quad (49)$$

This integral can be calculated by completing the square in the eigenbasis of  $C_0$ , and the resulting distribution is simply Gaussian with variance matrix  $\tilde{C}_0 = C_0 + \epsilon^2 I$ .

Then the correlation matrix which determines the probability distribution in our system is  $\tilde{C}_0 = C_0 + \epsilon^2 I$  and what is measured is  $\tilde{C}_r$ , such that  $\lim_{n \rightarrow \infty} \tilde{C}_r = \tilde{C}_0$ . The Marčenko–Pastur theorem is still valid and we have  $\tilde{m}_r(z_r) = \tilde{m}_0(z_0)$ . We can then relate the modified moment generating function  $\tilde{m}_0(z_0)$  to the real moment generating function  $m_0(\tilde{z}_0)$  through a change of variables  $\tilde{z}_0 = z_0 - \epsilon^2$ . The Stieltjes transform of the new eigenvalue distribution becomes:

$$\tilde{s}(z) = \frac{1}{p} \text{Tr}(\tilde{C} - zI)^{-1} = \frac{1}{p} \text{Tr}(C - \tilde{z}I)^{-1} = s(\tilde{z}). \quad (50)$$

We can then derive the moment generating function:

$$\tilde{m}(z) = -(z\tilde{s}(z)) - 1 = -(\tilde{z}s(\tilde{z})) - 1 - \epsilon^2 s(\tilde{z}). \quad (51)$$

The first two terms are nothing but  $m(\tilde{z})$ , while one can show through a Laurent expansion that the last term is given by  $\frac{\epsilon^2}{\tilde{z}} [1 + m(\tilde{z})]$ . The relation between the moment generating functions becomes then

$$\tilde{m}(z) = m(\tilde{z}) \left[ 1 + \frac{\epsilon^2}{\tilde{z}} \right] + \frac{\epsilon^2}{\tilde{z}}, \quad (52)$$

from which we derive an adapted form of the Marčenko–Pastur theorem:

$$\tilde{m}_r(z_r) = m_0(\tilde{z}_0) \left[ 1 + \frac{\epsilon^2}{\tilde{z}_0} \right] + \frac{\epsilon^2}{\tilde{z}_0}, \quad (53)$$

$$\text{where} \quad \tilde{z}_0 = \frac{z_r}{1 + r m_r(z_r)} - \epsilon^2$$

Finally, this allows us to derive relations between the moments  $\tilde{m}_{rk}^r$  of the experimentally measured eigenvalue distribution and the moments of the eigenvalue distribution of the shadow system  $m_k^0$ . This can be done through a Laurent expansion in  $z_r$  on both sides, together with a Taylor expansion in  $\epsilon^2$ . After a considerable amount of algebra we find to second order in  $\epsilon^2$ :

$$\begin{aligned} \tilde{m}_r^1 &= m_0^1 + \epsilon^2 \\ \tilde{m}_r^2 &= m_0^2 + 2m_0^1 \epsilon^2 + r(m_0^1)^2 + \epsilon^4 \end{aligned} \quad (54)$$

The central moments, *i.e.* the mean  $\mu$  and the variance  $\sigma^2$ , are more convenient and we find

$$\begin{aligned} \tilde{\mu}_r &= \mu_0 + \epsilon^2 \\ \tilde{\sigma}_r^2 &= \sigma_0^2 + r\mu_0^2 \end{aligned} \quad (55)$$

To this order, the second moment does not depend on the noise, however it comes in at higher orders. For a system with pure noise eqn (55) gives  $\mu_n = \epsilon^2$  and  $\sigma_n^2 = r\epsilon^4$  (note that for  $r \rightarrow 0$ , the second moment of the noise eigenvalue distribution vanishes as it should). Since for large enough  $\epsilon$ , this has to be consistent with eqn (55), we obtain:

$$\begin{aligned} \tilde{\mu}_r &= \mu_0 + \epsilon^2 \\ \tilde{\sigma}_r^2 &= \sigma_0^2 + r\mu_0^2 + r\epsilon^4 \end{aligned} \quad (56)$$

To convert these relations into useful relations for the density of states one needs to perform the change of variable  $\omega \propto \lambda^{-1/2}$ . Unfortunately, we could not find a straightforward and reliable way to calculate the moments of this distribution and we had to rely mostly on numerical results. However in the limit  $r \rightarrow 0$ , the

relation  $\tilde{C}_0 = C_0 + \varepsilon^2 I$  directly translates to the eigenvalues, *i.e.*  $\tilde{\lambda}_j^0 = \lambda_j^0 + \varepsilon^2$ , and we recover eqn (56) for  $r = 0$ . Here we can perform the change of variables explicitly, and we find to order  $\varepsilon^2$  (written in the most convenient mixture of direct and central moments):

$$\begin{aligned}\tilde{\mu}_0^\omega &= \mu_0^\omega - \frac{\varepsilon^2}{2} \langle \omega^3 \rangle_0 \\ (\tilde{\sigma}_0^2)^\omega &= (\sigma_0^2)^\omega - \varepsilon^2 (\langle \omega^4 \rangle_0 - \langle \omega^3 \rangle_0 \langle \omega \rangle_0)\end{aligned}\quad (57)$$

It can be shown that the factor multiplying  $\varepsilon^2$  in the equation for  $\sigma^2$  is strictly positive, so that in the limit  $r \rightarrow 0$ , both mean and variance are reduced proportional to  $\varepsilon^2$ .

## Acknowledgements

We would like to thank J.-P. Bouchaud and G. Biroli for inspiring the correlation matrix approach and pointing to the Marčenko–Pastur theorem and acknowledge the soft matter groups at U. Penn. for collaborating and sharing their data on the NIPA system. None of the present work would have been done without the numerous discussions, we have had at various level with those involved in the analysis in terms of vibrational modes of experimental and numerical data. In particular we thank Wouter Ellenbroek, Andrea Liu and Corey O’Hern for helpful discussions on the interpretation of modes. A special thought to Wim van Sarloos and Martin van Hecke who hosted us in Leiden and always encouraged us to persevere in this delicate game, the rules and traps of which we have tried to expose in this paper.

## References

- 1 W. Phillips, *J. Non-Cryst. Solids*, 1978, **31**, 267; proceedings of the Topical Conference on Atomic Scale Structure of Amorphous Solids.
- 2 T. Nakayama, *Rep. Prog. Phys.*, 2002, **65**, 1195.
- 3 G. Katgerl and M. van Hecke, *Europhys. Lett.*, 2010, **92**, 34002.
- 4 B. P. Tighe, *Phys. Rev. Lett.*, 2011, **107**, 158303.
- 5 M. van Hecke and J. Phys., *J. Phys.: Condens. Matter*, 2010, **22**, 033101.
- 6 J. A. Dijksman, G. H. Wortel, L. T. H. van Dellen, O. Dauchot and M. van Hecke, *Phys. Rev. Lett.*, 2011, **107**, 108303.
- 7 S. Alexander, *Phys. Rep.*, 1998, **296**, 65.
- 8 M. Wyart, *Ann. Phys. (Paris, Fr.)*, 2006, **30**, 1.
- 9 N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Brooks Cole, 1976.
- 10 A. I. Chumakov, I. Sergueev, U. van Bürck, W. Schirmacher, T. Asthalter, R. Rüffer, O. Leupold and W. Petry, *Phys. Rev. Lett.*, 2004, **92**, 245508.
- 11 M. Wyart, S. Nagel and T. Witten, *Europhys. Lett.*, 2005, **72**, 486.
- 12 N. Xu, M. Wyart, A. J. Liu and S. R. Nagel, *Phys. Rev. Lett.*, 2007, **98**, 175502.
- 13 K. Chen, W. G. Ellenbroek, Z. Zhang, D. T. N. Chen, P. J. Yunker, S. Henkes, C. Brito, O. Dauchot, W. van Saarloos, A. J. Liu and A. G. Yodh, *Phys. Rev. Lett.*, 2010, **105**, 025501.
- 14 G. Baldi, A. Fontana, G. Monaco, L. Orsingher, S. Rols, F. Rossi and B. Ruta, *Phys. Rev. Lett.*, 2009, **102**, 195502.
- 15 H. Shintani and H. Tanaka, *Nat. Mater.*, 2008, **07**, 870.

- 16 L. E. Silbert, A. J. Liu and S. R. Nagel, *Phys. Rev. Lett.*, 2005, **95**, 098301.
- 17 M. Wyart, L. E. Silbert, S. R. Nagel and T. A. Witten, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **72**, 051306.
- 18 W. G. Ellenbroek, E. Somfai, M. van Hecke and W. van Saarloos, *Phys. Rev. Lett.*, 2006, **97**, 258001.
- 19 M. Mailman, C. F. Schreck, C. S. O’Hern and B. Chakraborty, *Phys. Rev. Lett.*, 2009, **102**, 255501.
- 20 Z. Zeravcic, N. Xu, A. J. Liu, S. R. Nagel and W. van Saarloos, *Europhys. Lett.*, 2009, **87**, 26001.
- 21 C. F. Schreck, N. Xu and C. S. O’Hern, *Soft Matter*, 2010, **6**, 2960.
- 22 E. Somfai, M. van Hecke, W. G. Ellenbroek, K. Shundyak and W. van Saarloos, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2007, **75**, 020301.
- 23 S. Henkes, M. van Hecke and W. van Saarloos, *Europhys. Lett.*, 2010, **90**, 14003.
- 24 C. Brito and M. Wyart, *Europhys. Lett.*, 2006, **76**, 149.
- 25 J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed., Academic Press, Amsterdam, 2006.
- 26 G. Greaves, *Adv. Phys.*, 2007, **56**, 1.
- 27 T. Keyes, *J. Phys. Chem. A*, 1997, **101**, 2921.
- 28 N. Xu, *Structural signature of jamming transition in thermal amorphous systems*, 2009, arXiv: 0911.1576v1 [cond-mat.soft].
- 29 E. Owens and K. Daniels, in 2011 March Meeting, Dallas, TX (2011).
- 30 E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield and D. A. Weitz, *Science*, 2000, **287**, 627.
- 31 C. Micheletti, P. Carloni and A. Maritan, *Proteins: Struct., Funct., Bioinf.*, 2004, **55**, 635.
- 32 L. Petitjean, M. Reffay, E. Grasland-Mongrain, M. Poujade, B. Ladoux, A. Buguin and P. Silberzan, *Biophys. J.*, 2010, **98**, 1790.
- 33 L. Laloux, P. Cizeau, J.-P. Bouchaud and M. Potters, *Phys. Rev. Lett.*, 1999, **83**, 1467.
- 34 G.-J. Gao, J. Bławdziewicz and C. S. O’Hern, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2006, **74**, 061304.
- 35 A. Ghosh, V. K. Chikkadi, P. Schall, J. Kurchan and D. Bonn, *Phys. Rev. Lett.*, 2010, **104**, 248305.
- 36 D. Kaya, N. L. Green, C. E. Maloney and M. F. Islam, *Science*, 2010, **329**, 656.
- 37 P. J. Yunker, K. Chen, Z. X. Zhang and A. G. Yodh, *Phys. Rev. Lett.*, 2011, **106**, 225503.
- 38 P. J. Yunker, K. Chen, Z. X. Zhang, W. G. Ellenbroek, A. J. Liu and A. G. Yodh, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2011, **83**, 011403.
- 39 C. Brito, O. Dauchot, G. Biroli and J.-P. Bouchaud, *Soft Matter*, 2010, **6**, 3013.
- 40 V. A. Marčenko and L. A. Pastur, *Math. USSR-Sb.*, 1967, **1**, 457.
- 41 M van Hecke, in 2011 March Meeting, Dallas, TX (2011).
- 42 C. F. Schreck, T. Bertrand, C. S. O’Hern and M. D. Shattuck, *Phys. Rev. Lett.*, 2011, **107**, 078301.
- 43 M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Clarendon Press, 1999.
- 44 F. Lechenault, O. Dauchot, G. Biroli and J.-P. Bouchaud, *Europhys. Lett.*, 2008, **83**, 46003.
- 45 F. Lechenault, O. Dauchot, G. Biroli and J.-P. Bouchaud, *Europhys. Lett.*, 2008, **83**, 46002.
- 46 C. Brito and M. Wyart, *J. Stat. Mech.: Theory Exp.*, 2007, **2007**, L08003.
- 47 J. W. Silberstein and Z. D. Bai, *J. Multivariate Anal.*, 1995, **54**, 175.
- 48 N. E. Karoui, *Ann. Stat.*, 2008, **36**, 2757.
- 49 M. Schindler and A. C. Maggs, *Truncated correlations in video microscopy of colloidal solids*, 2011, arXiv: 1111.1419v1 [cond-mat.soft].
- 50 A. Ghosh, V. Chikkadi, P. Schall and D. Bonn, *Phys. Rev. Lett.*, 2011, **107**, 188303.
- 51 C. Brito and M. Wyart, *J. Chem. Phys.*, 2009, **131**, 024504.
- 52 Z. Burda, A. Görlich, A. Jarosz and J. Jurkiewicz, *Phys. A*, 2004, **343**, 295.