

## Principal-component analysis of particle motion

H. Y. Chen,<sup>1,2</sup> Raphaël Liégeois,<sup>3</sup> John R. de Bruyn,<sup>2,\*</sup> and Andrea Soddu<sup>2,\*†</sup>

<sup>1</sup>*Department of Nuclear Science and Technology, Fudan University, Shanghai 200433, China*

<sup>2</sup>*Department of Physics and Astronomy, University of Western Ontario, London, Ontario, Canada N6A 3K7*

<sup>3</sup>*Montefiore Institute, Université de Liège, 4000 Liège, Belgium*

(Received 22 July 2014; published 15 April 2015)

We demonstrate the application of principal-component analysis (PCA) to the analysis of particle motion data in the form of a time series of images. PCA has the ability to resolve and isolate spatiotemporal patterns in the data. Using simulated data, we show that this translates into the ability to separate individual frequency components of the particle motion. We also show that PCA can be used to extract the fluid viscosity from images of particles undergoing Brownian motion. PCA thus provides an efficient alternative to more traditional particle-tracking methods for the analysis of microrheological data.

DOI: [10.1103/PhysRevE.91.042308](https://doi.org/10.1103/PhysRevE.91.042308)

PACS number(s): 83.85.Ns, 83.10.Pp, 05.45.Tp, 02.70.Rr

### I. INTRODUCTION

Principal-component analysis (PCA) [1,2], also known as Karhunen-Loève transformation, is a statistical method for extracting patterns from complex data sets and reducing the dimension of complex signals. PCA involves transforming a set of data into a linear superposition of orthogonal components, arranged such that the first principal component has the largest possible variance; i.e., it accounts for the largest contribution to the variation in the data. Successive principal components similarly have the maximum possible variance subject to the orthogonality condition. By construction, the first few principal components typically contain most of the information embedded in the data, making PCA the basis of very efficient data compression and spatiotemporal signal-processing methods [3]. PCA has important applications in fields such as biomedical science [4,5], neuroscience [6,7], environmental science [8–10], image compression [11,12], and image analysis [13].

In this paper, we demonstrate the application of PCA to the analysis of particle motion, primarily in the context of microrheological measurements. Microrheology involves the analysis of thermally driven or externally forced motion of micron-scale tracer particles suspended in a complex fluid to determine the local viscoelastic properties of the fluid or to study its microstructure [14–16]. In an active microrheological experiment, an external force is applied to tracer particles using, for example, optical tweezers, and the resulting response of the particles is analyzed to give the information about the viscoelastic properties of the fluid [14,16–21]. In a passive multiple-particle-tracking experiment, the positions of many tracer particles are recorded over time and the resulting time series of images is analyzed to determine the mean squared displacement of the particles as a function of lag time  $\tau$  [14–16,22,23]. The local frequency-dependent viscous and elastic moduli can be determined from the mean squared displacement using a generalized Stokes-Einstein relation [15,24]. Both types of experiment produce a set of

spatiotemporal data in the form of a time series of images which is well suited for analysis using PCA.

PCA differs from most commonly used particle-tracking analysis methods in that the entire spatiotemporal data set is analyzed as a whole, as opposed to identifying individual particles in each image and reconstructing their trajectories [25,26]. Using simulated data, we demonstrate in Sec. II A that PCA can resolve periodic particle motion with a range of frequencies, and in Secs. II B and II C we show that PCA provides an alternative to conventional particle-tracking algorithms for determining the viscosity of Newtonian fluids from microrheological data and for studying non-Newtonian behavior, respectively. Our results demonstrate that PCA is a promising tool for particle motion analysis.

Detailed introductions to PCA can be found in several excellent references [1,3,27,28]. The decomposition of a data set into its principal components is illustrated schematically in Fig. 1. We consider spatiotemporal data in the form of a time series of images. Each image consists of  $m$  pixels, and images are recorded at discrete times  $t = 1, 2, \dots, T$ . While in this paper we focus on the analysis of two-dimensional particle motion, there is no restriction on the spatial dimension of the data, and it is straightforward to apply our analysis to three- or higher-dimensional images. The entire data set can be represented as an  $m \times T$  matrix  $\mathbf{X}$ , each column of which is a vector  $\mathbf{x}_i$  containing the image data recorded at a particular time. Let  $\mathbf{W}$  be a matrix whose rows are the  $n$  principal-component vectors  $\mathbf{w}_i$ , normalized so that the modulus squared  $|\mathbf{w}_i|^2 = 1$ . A matrix  $\mathbf{Y}$  can be calculated via the transformation

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \end{bmatrix} = \mathbf{W}\mathbf{X} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \end{bmatrix} [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots], \quad (1)$$

and  $\mathbf{y}_i$ , the  $i$ th row of  $\mathbf{Y}$ , is the time-dependent amplitude of the principal component  $\mathbf{w}_i$ , as illustrated in Fig. 1.

The principal components  $\mathbf{w}_i$  are the eigenvectors of  $\mathbf{C} = \mathbf{X}\mathbf{X}^T$ , which is proportional to the covariance matrix of  $\mathbf{X}$ . The matrix  $\mathbf{D}$ , given by

$$\mathbf{D} = \mathbf{Y}\mathbf{Y}^T = \mathbf{W}\mathbf{X}\mathbf{X}^T\mathbf{W}^T = \mathbf{W}\mathbf{C}\mathbf{W}^T, \quad (2)$$

\*These two authors contributed equally to this work.

†asoddu@uwo.ca

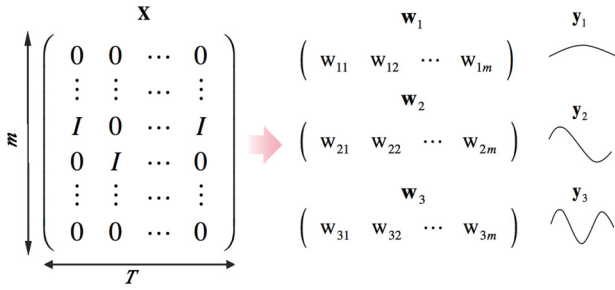


FIG. 1. (Color online) A schematic illustration of the decomposition of an  $m \times T$  data matrix  $\mathbf{X}$  into its principal components  $\mathbf{w}_i$ . The vectors  $\mathbf{y}_i$  are the time courses of the corresponding principal components.

is diagonal, and the diagonal elements of  $\mathbf{D}$  are its eigenvalues. The  $i$ th diagonal element of  $\mathbf{D}$ , which we label as  $\mathcal{D}^{(i)}$ , is the variance associated with  $\mathbf{w}_i$ . The principal components are sorted according to magnitude of  $\mathcal{D}^{(i)}$ , with  $\mathcal{D}^{(1)} = \max\{\mathcal{D}^{(i)}\}$  [1]. For low-noise signals, the amount of information embedded in each principal component is related to the corresponding variance: components with a larger variance carry a larger portion of the information from the original signal. Most of the information in the signal is therefore preserved by keeping only the first few principal components.  $\mathbf{X}$  is conventionally centered on zero by subtracting its mean before applying PCA, but this process inevitably results in the loss of the physical information embedded in the average signal. To avoid this, we did not center  $\mathbf{X}$  in the present work. In this case, the first principal component is just the average signal. The PCA calculations in this paper were performed using software [29] written in Matlab [30].

For concreteness in the following discussion, we assume that each image in the data set has a fixed number of particles  $n$  in the field of view, and that each particle occupies  $k$  pixels. Assuming that the particle concentration is small enough that particles do not overlap (as would normally be the case in a microrheological experiment [23]),  $p = kn$  pixels will be occupied. Occupied pixels have a value  $I$ , and vacant pixels are 0, as illustrated in Fig. 1. The pixel values change with time as the tracer particles move. As noted above, each two-dimensional image is reshaped to form one column of the data matrix  $\mathbf{X}$ , so that each row of  $\mathbf{X}$  corresponds to a given spatial location (pixel) and each column to a particular time.

## II. RESULTS

### A. Periodic motion

In a typical active microrheology experiment, tracer particles would be driven, and would all respond, at a single known frequency. In other applications, however, nonlinear coupling between individual oscillators can lead them to respond at different frequencies. One example comes from neuroscience, in which coupled oscillations in cortical circuits have been studied using the Kuramoto model [31]. We demonstrate here that such systems can be studied by taking advantage of the frequency-resolving ability of PCA, which we illustrate by analyzing images of particles undergoing periodic motion. We created a simulated data set comprising 300 images of

$600 \times 800$  pixels, with a simulated frame rate of  $10 \text{ s}^{-1}$ . Five groups of 10 particles were placed randomly in the first image. The particles moved in simple harmonic motion with random orientation and an amplitude of 50 pixels, with each group oscillating at a different frequency. The data were decomposed using PCA [29] into principal components  $\mathbf{w}_i$  and their respective time courses  $\mathbf{y}_i$ . The results are summarized in Fig. 2.

$\mathbf{w}_1$ , the principal component with the largest variance, is shown in the left-hand panel of Fig. 2(a). Its time course  $\mathbf{y}_1$  is shown in the inset in the right panel of Fig. 2(a), and the

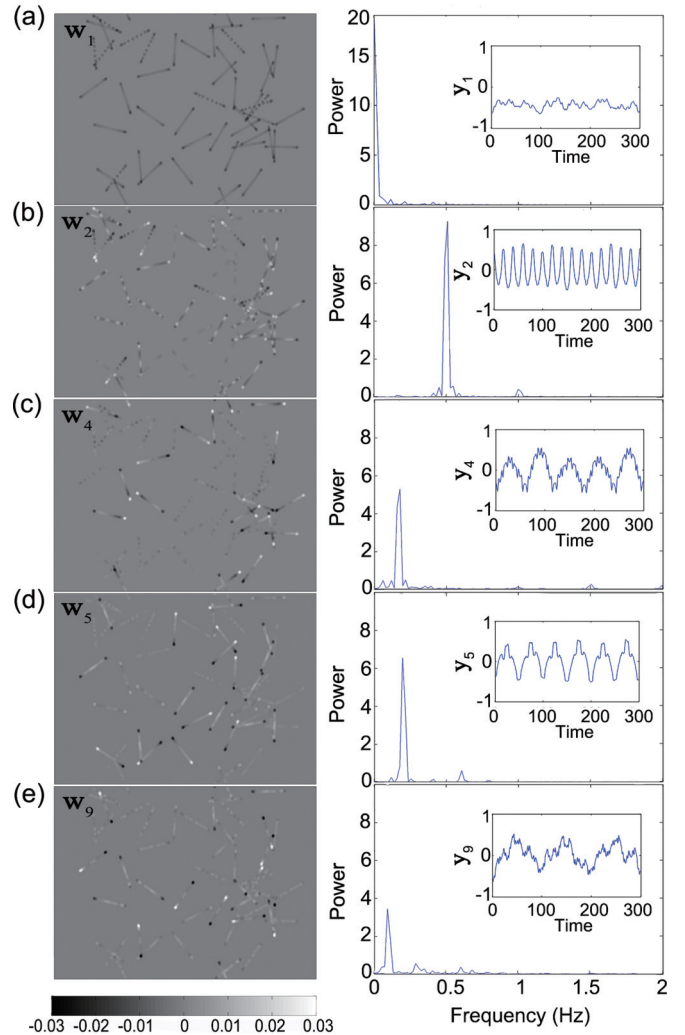


FIG. 2. (Color online) The results of PCA of simulated data for particles undergoing simple harmonic motion. The left-hand panels of (a)–(e) show the first, second, fourth, fifth, and ninth principal components, respectively. The inset on the corresponding right-hand panel shows the time course of the principal component, and the power spectrum is plotted in the main graph of the right-hand panel, with all  $y$ -axis units arbitrary. (a) The first principal component is the time average of the data. Its time course is roughly constant, and there is no strong peak (other than that at zero frequency) in the power spectrum. The other principal components shown in (b)–(e) selectively pick out oscillations at frequencies 0.5 Hz, 0.167 Hz, 0.2 Hz, and 0.1 Hz respectively. Oscillations at 0.056 Hz were not isolated by any principal component.

Fourier power spectrum of  $\mathbf{y}_1$  in the main plot on the right of Fig. 2(a). This principal component represents the average state of the system and, as discussed above, could be eliminated by subtracting the time average of  $\mathbf{X}$  from  $\mathbf{X}$ . The probability distribution function of the position  $z$  of a particle undergoing harmonic oscillation about the origin with an amplitude  $A$  is [32]  $f(z) = \pi^{-1}(A^2 - z^2)^{-1/2}$ , and the image of  $\mathbf{w}_1$  reflects this distribution for each particle in the data set. Since it includes information from all particles, its time course  $\mathbf{y}_1$  is roughly constant; its power spectrum is peaked at 0 frequency and shows no other significant features.

The other leading principal components tend to pick out oscillations of a certain frequency. As illustrated in Fig. 2(b)–2(e),  $\mathbf{w}_2$ ,  $\mathbf{w}_4$ ,  $\mathbf{w}_5$ , and  $\mathbf{w}_9$  are dominated by motion at frequencies 0.5, 0.167, 0.2, and 0.1 Hz, respectively, these being the oscillation frequencies of four of the five groups of particles in the data. The most prominent features in the images of these principal components are the end points of the trajectories of particles oscillating at the selected frequency, and the power spectra of the corresponding time courses  $\mathbf{y}_2$ ,  $\mathbf{y}_4$ ,  $\mathbf{y}_5$ , and  $\mathbf{y}_9$  have single strong peaks at the selected frequency. For our simulated data, PCA is able to cleanly separate motion at frequencies that differ by only 20%, and its frequency selectivity would presumably be higher for a longer data set. The fifth set of particles in the data oscillated at 0.056 Hz. This low frequency was not detected by our analysis, presumably because the length of the data set, which corresponded to 1.68 oscillation periods at this frequency, was too short. Other low-index principal components not shown in Fig. 2 contain either harmonics or combinations of the fundamental particle frequencies. As seen in the results shown in Fig. 2, PCA tends to rank the principal components from high to low frequency, indicating that higher-frequency oscillations contribute more variance to the data. Similar results were obtained from data in which small stochastic perturbations were added to the oscillatory motion, indicating that the analysis is robust with respect to small amounts of noise.

### B. Viscosity determination

PCA can be used to determine the viscosity of a Newtonian fluid from a series of images of tracer particles diffusing in the fluid. These tracers undergo Brownian motion, and their mean squared displacement  $\langle \Delta r^2 \rangle$  grows linearly with time  $t$ :  $\langle \Delta r^2(t) \rangle = 2dDt$ , where  $d$  is the dimensionality and  $D$  is the diffusion coefficient. For Brownian motion, the distribution of particle displacements  $\Delta r$  over a time interval period  $\Delta t$  is Gaussian with a standard deviation

$$\sigma = \sqrt{2dD\Delta t}. \quad (3)$$

For a single particle, the viscosity  $\eta$  is related to  $D$  by the Stokes-Einstein relation,

$$\eta = k_B T / 2\pi a d D, \quad (4)$$

with  $k_B$  the Boltzmann constant,  $T$  the absolute temperature, and  $a$  the radius of the particle [33]. Although hydrodynamic interactions between particles cause  $\eta$  to increase with particle volume fraction [34], we neglect any such effects in this work.

From Eq. (3), a particular particle  $j$  whose original position is  $\mu_j$  will appear at position  $\xi_i$  in a subsequent frame with a

probability given by the distribution

$$f(\xi_i, \mu_j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\xi_i - \mu_j)^2}{2\sigma^2}\right], \quad (5)$$

where  $\sigma$  is a function of  $\Delta t$  as in Eq. (3). To maximize the variance  $\mathcal{D}^{(1)}$ , the elements of the first principal component  $\mathbf{w}_1$  should have values determined by Eq. (5), summed over the  $p$  occupied pixels. If the position  $\xi_i$  corresponds to the  $i$ th row of the data matrix  $\mathbf{X}$ , then the  $i$ th component of  $\mathbf{w}_1$  will be

$$\mathbf{w}_{1i} = \sum_{j=1}^p \alpha \exp\left[-\frac{(\xi_i - \mu_j)^2}{2\sigma^2}\right], \quad (6)$$

where  $\alpha$  is a normalization factor to be determined. From Eq. (2), the variance associated with the first principal component is then

$$\mathcal{D}^{(1)} = \mathbf{w}_1 \mathbf{X} \mathbf{X}^T \mathbf{w}_1^T = I^2 \sum_{t=1}^T \left( \sum_{i=occ} \mathbf{w}_{1i} \right)^2, \quad (7)$$

where the second sum is over the  $p$  occupied pixels in the image.

For simplicity we now consider the motion of a single particle occupying a single pixel, in which case  $p = 1$  and we can drop the  $j$  subscript. The variance in the first principal component is then

$$\mathcal{D}^{(1)} = I^2 \sum_{t=1}^T \left\{ \alpha \exp\left[-\frac{(\xi - \mu)^2}{2\sigma^2}\right] \right\}^2 \quad (8)$$

$$\approx I^2 T \int \left\{ \alpha \exp\left[-\frac{(\xi - \mu)^2}{2\sigma^2}\right] \right\}^2 f(\xi, \mu) d\xi \quad (9)$$

$$= I^2 T \frac{\alpha^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{ \exp\left[-\frac{(\xi - \mu)^2}{2\sigma^2}\right] \right\}^3 d\xi \quad (10)$$

$$= \frac{I^2 T \alpha^2}{\sqrt{3}}. \quad (11)$$

The normalization condition is

$$|\mathbf{w}_1|^2 = \sum_{i=1}^m \left\{ \alpha \exp\left[-\frac{(\xi_i - \mu)^2}{2\sigma^2}\right] \right\}^2 \quad (12)$$

$$\approx \alpha^2 \int_{-\infty}^{\infty} \left\{ \exp\left[-\frac{(\xi - \mu)^2}{2\sigma^2}\right] \right\}^2 d\xi \quad (13)$$

$$\approx \alpha^2 \sigma \sqrt{\pi} = 1, \quad (14)$$

from which  $\alpha \approx [1/(\sigma\pi^{1/2})]^{1/2}$ . This finally gives

$$\mathcal{D}^{(1)} \approx \frac{I^2 T}{\sigma\sqrt{3\pi}}. \quad (15)$$

More generally, when  $p > 1$

$$\mathcal{D}^{(1)} \approx \left( \frac{1}{\sqrt{3}} + \frac{p-1}{2} \right) \frac{I^2 T}{\sigma\sqrt{\pi}}. \quad (16)$$

The total variance of all  $m$  principal components is

$$\mathcal{D}^{\text{tot}} = \sum_{i=1}^m \mathcal{D}^{(i)} = \text{tr}\{\mathbf{C}\} = \sum_{i=1}^m \sum_{t=1}^T \mathbf{x}_{it}^2 = I^2 T p, \quad (17)$$

where  $x_{it}$  is the  $(i, t)$  element of  $\mathbf{X}$ .  $\mathcal{D}^{\text{tot}}$  is constant for a given data set, regardless of the details of the motion of the particles. Using this, the fraction of the total variance accounted for by the first principal component is

$$\delta^{(1)} = \frac{\mathcal{D}^{(1)}}{\mathcal{D}^{\text{tot}}} \approx \left( \frac{1}{\sqrt{3}} + \frac{p-1}{2} \right) \frac{1}{p\sigma\sqrt{\pi}}. \quad (18)$$

In a typical multiple-particle-tracking experiment [23],  $p \gg 1$ , implying that

$$\delta^{(1)} \approx 1/(2\sigma\sqrt{\pi}). \quad (19)$$

This behavior is illustrated in Fig. 3(a), which shows  $\delta^{(1)}$  calculated from simulated data for 25 particles undergoing random walks. Here  $\sigma$  is an input to the simulations and determines the magnitude of the particle displacements at each time step, and the results were calculated for a time interval between images of  $\Delta t = 1$  time step (i.e., all images were used in the analysis). Equivalently, for a fixed particle step size (or in a real experiment),  $\sigma$  would be proportional to  $\Delta t^{1/2}$  according to Eq. (3). The data in Fig. 3 show the predicted  $1/\sigma$  dependence for  $\sigma \gtrsim 0.3$ .  $\delta^{(1)}$  approaches 1 for small  $\sigma$ ,

indicating that the first principal component accounts for most of the variance in the data set in this regime. As  $\sigma$  increases, the other components become progressively more significant.

The low- $\sigma$  plateau is not captured by the approximate expression Eq. (18) (or Eq. (19)). This is due to the approximations made in Eqs. (8) and (12), and due to the assumption that the particles are pointlike. In fact, each particle in our simulations (and in real experiments) occupies a number of pixels, and when  $\sigma$  is much smaller than the radius of the particles, most of the occupied pixels do not change with time. This implies that  $\delta^{(1)}$  does not depend strongly on  $\sigma$ , leading to the low- $\sigma$  plateau observed in Fig. 3(a). As  $\sigma$  is increased and the particles move more,  $\delta^{(1)}$  decreases as  $1/\sigma$ , in accord with the prediction of Eq. (18).

Figure 3(b) shows the fraction of the total variance accounted for by the first  $s$  principal components,

$$\delta^s = \frac{\sum_{i=1}^s \mathcal{D}^{(i)}}{\mathcal{D}^{\text{tot}}}, \quad (20)$$

for several values of  $\sigma$  for the same simulations. For small  $\sigma$ , essentially all of the variance is in the first principal component. As  $\sigma$  increases, the variances of the higher principal components become comparable to  $\mathcal{D}^{(1)}$ . In the extreme case that  $\sigma \rightarrow \infty$ , the data set is completely random and all principal components have the same variance. In this limit,  $\delta^s$  grows linearly with  $s$ .

Equation (18) and the results plotted in Fig. 3 provide a means of determining the diffusion constant  $D$ , and hence the viscosity  $\eta$ , from a sequence of images of particles diffusing in a liquid using PCA rather than the usual particle-tracking methods. As an example, we constructed a simulated video comprising 200 images of 25 particles undergoing Brownian motion and analyzed it using PCA. The first two principal components calculated from these data are shown in Fig. 4(a).

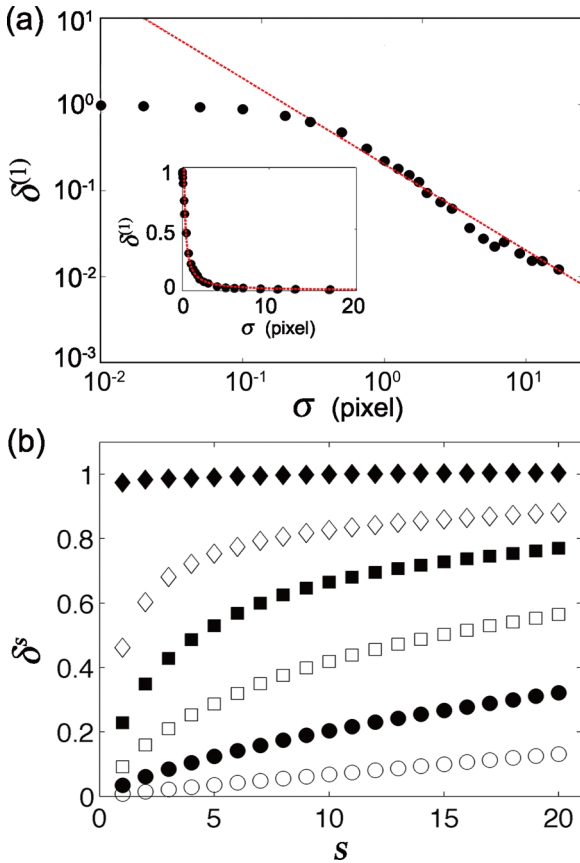


FIG. 3. (Color online) (a)  $\delta^{(1)}$ , the fraction of the total variance in the data accounted for by the first principal component as a function of  $\sigma$  for a simulation of 25 particles undergoing Brownian motion. The inset shows the same data with linear axes. In both cases the line is the prediction of Eq. (18). Panel (b) shows the fraction of the total variance in the first  $s$  principal components for  $\sigma = 0.01$  (solid diamonds), 0.5 (open diamonds), 1 (solid squares), 2 (open squares), 4 (solid circles), and 20 (open circles) pixels.

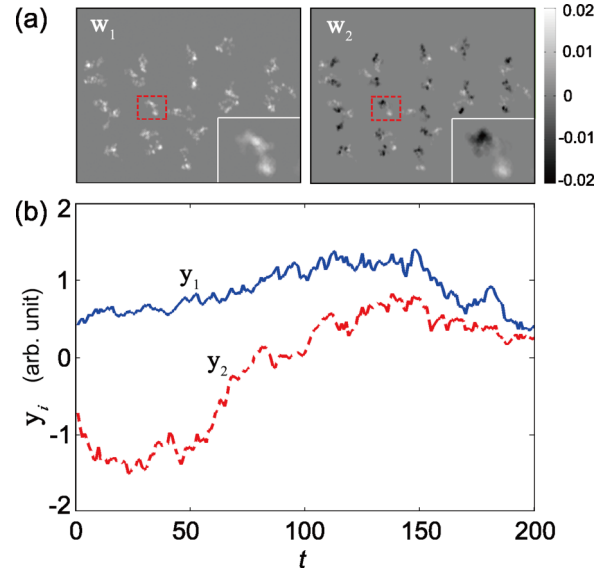


FIG. 4. (Color online) (a) The first two principal components and (b) their time courses, calculated from simulated data on particles undergoing Brownian motion. The insets at the bottom right of the images in (a) are magnified views of the small regions outlined by rectangles.

As above,  $\mathbf{w}_1$  is the time average of the complete data set. The positive regions of  $\mathbf{w}_1$  correspond to regions occupied more frequently by particles, and their size is an indication of how far the particles diffuse over the duration of the run. The time course  $y_1$  of the first principal component, shown in Fig. 4(b), is roughly constant. The image of  $\mathbf{w}_2$  (and higher components not shown) contains localized positive and negative regions representing the end points of the particles' trajectories, and  $y_2$  varies quasisinusoidally as the particle diffuses from its initial to its final location. By calculating  $\delta^{(1)}$ , the fraction of the total variance contained in  $\mathbf{w}_1$ ,  $\sigma$  could be calculated using Eq. (18) [or Eq. (19)], or from Fig. 3(a) as appropriate, and  $\eta$  could then be determined using Eqs. (3) and (4). Note that in order to obtain an accurate value for  $\sigma$ , one should work in the regime where  $\delta^{(1)}$  shows a significant dependence on  $\sigma$ , that is, away from the plateau around  $\delta^{(1)} = 1$ . In this analysis, we have considered only the first principal component; a more detailed analysis involving higher-order components would be expected to give a more accurate result.

**C. More complex fluids**

In complex fluids, the motion of suspended tracer particles is affected by viscoelasticity and/or the microstructure of the material [15,23,35–37]. In a material undergoing gelation, for example, changes in material microstructure can lead to temporally and spatially heterogeneous behavior which has been studied using particle tracking techniques [36–39]. PCA is also sensitive to these effects. Trapping or other factors that restrict particle motion locally will result in variance-rich regions in the data images, which appear as regions of high intensity in the low-order principal components. This allows trap locations and trapping times to be identified from analysis of the principal components and their time courses. As an example, Fig. 5 illustrates the results of PCA of a simulation in which the step size of the diffusing particles decreases linearly with time. This can be viewed as a crude model for the progressive localization of the tracer particles due to the evolution of material microstructure, for example, due to

gelation. Figure 5(a) shows the first principal component  $\mathbf{w}_1$  obtained from the simulated data; its time course is plotted as a solid line in Fig. 5(b). The dashed line in Fig. 5(b) shows the rms displacement of the particles in the simulation as a function of time. In contrast to the case of simple Brownian motion shown in Fig. 4,  $y_1$  is clearly not constant in the present case, but rather increases with time as the particles become more confined. The image of the principal component provides information about local structure in the fluid by indicating areas where the particle motion is more restricted, while its time course gives information about the change in variance over time. Together, they can be used to determine the changes with time in the behavior of the tracer particles and the properties of the fluid.

**III. DISCUSSION AND CONCLUSION**

PCA is well suited to the extraction of information from spatiotemporal data sets. In this paper we have demonstrated the potential of PCA as a tool for the analysis of sequences of images of tracer particles suspended in fluids, such as would be generated in microrheological studies.

Analysis of simulated data for harmonically oscillating particles showed that the first principal component provides information about the probability distribution of the particles' position. The orthogonality of the principal components gives PCA the ability to decompose the particle motion into its individual component frequencies. In this case PCA behaves similarly to Fourier analysis, although, apart from the orthogonality condition, the two methods are mathematically quite distinct. In the case of Fourier analysis, the basis vectors are imposed *a priori*, while in PCA, the optimal eigenvectors are determined as part of the calculation. It can be shown that the Fourier transform is in fact the optimal decomposition for a periodic time series [40,41]: the orthogonal principal components  $\mathbf{w}_i$  obtained by singular value decomposition of the covariance matrix ( $\mathbf{X}\mathbf{X}^T$ ) turn out to be the eigenstates of the Fourier matrix. Projection of  $\mathbf{w}_i$  onto  $\mathbf{X}$  then gives a sinusoidal time course  $y_i$ , as in Fourier transformation.

We showed that PCA can be used as an alternative to particle tracking methods to determine fluid viscosity from images of particles undergoing Brownian motion through Eq. (18) [or Eq. (19)]. Our analysis of diffusion in Sec. II B was approximate, being based only on the first principal component. Accuracy would likely be improved if the contributions of higher components were explicitly accounted for. We also demonstrated the ability of PCA to capture time-varying fluid properties through simulation of progressive localization behavior of tracer particle.

PCA has certain advantages over standard particle-tracking algorithms. It treats the entire data set globally, rather than identifying particles in each image and reconstructing trajectories from frame to frame. Because of this, PCA is not subject to the ambiguities that can arise in particle-tracking methods when particles move a distance comparable to their separation between images, or when trajectories cross. On the other hand, fluctuations in particle number and changes in intensity that are unavoidable in experimental data may influence the PCA results. The application of the methods described in this paper

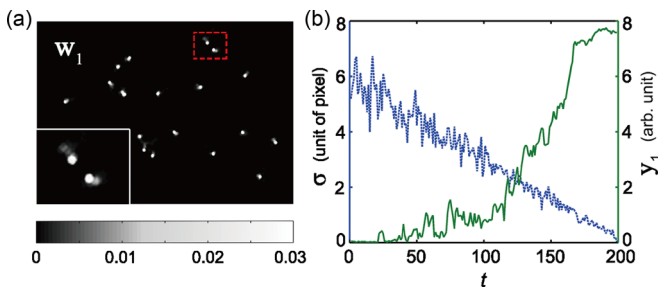


FIG. 5. (Color online) PCA of simulations in which diffusing particles become progressively more confined over time. (a) The first principal component  $\mathbf{w}_1$ . The inset at the bottom left is a magnified view of the small region outlined by the rectangle. (b) Its time course  $y_1$  (solid line) and the rms displacement  $\sigma$  of the particles in a single time step over the course of the simulations (dashed line).  $y_1$  increases over time as the particles become less mobile, while the image of  $\mathbf{w}_1$  shows the areas to which the particles become confined.

to experimental particle-tracking data will be the subject of future work. Our preliminary attempts in this direction have been encouraging but are not yet quantitatively accurate, due largely to experimental noise and the above issues. In general, the interpretation of PCA results is less straightforward than for particle tracking, and more detailed theoretical analysis is required to guide the application of PCA to data from real microrheological experiments. In particular, it should

be straightforward to extend our methods to viscoelastic fluids.

#### ACKNOWLEDGMENTS

This research was supported by Natural Science and Engineering Research Council of Canada and the Mitacs Globalink program.

- 
- [1] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. (Springer, New York, 2002).
- [2] K. Pearson, *Phil. Mag.* **2**, 559 (1901).
- [3] H. Hotelling, *J. Educ. Psychol.* **24**, 417 (1933).
- [4] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak, *J. Cerebral Blood Flow Metab.* **13**, 5 (1993).
- [5] J. C. Lindon, E. Holmes, and J. K. Nicholson, *Prog. Nucl. Mag. Res. Spect.* **39**, 1 (2001).
- [6] R. Viviani, G. Grön, and M. Spitzer, *Human Brain Mapping* **24**, 109 (2005).
- [7] A. H. Andersen, D. M. Gash, and M. J. Avison, *Magn. Res. Imaging* **17**, 795 (1999).
- [8] D. A. Saucy, J. R. Anderson, and P. R. Buseck, *J. Geophys. Res.: Atmos.* **96**, 7407 (1991).
- [9] S. M. Almeida, C. A. Pio, M. C. Freitas, M. A. Reis, and M. A. Trancoso, *Sci. Total Environ.* **368**, 663 (2006).
- [10] T. W. Chan and M. Mozurkewich, *Atmos. Chem. Phys.* **7**, 887 (2007).
- [11] Q. Du and J. E. Fowler, *IEEE Geosci. Remote Sensing Lett.* **4**, 201 (2007).
- [12] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 3rd ed. (Prentice-Hall, Upper Saddle River, NJ, 2008).
- [13] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, *IEEE Trans. Pattern Anal. Machine Intell.* **26**, 131 (2004).
- [14] M. L. Gardel, M. T. Valentine, and D. A. Weitz, "Microrheology," in *Microscale Diagnostic Techniques* (Springer, Berlin, 2005), p. 1.
- [15] T. M. Squires and T. G. Mason, *Annu. Rev. Fluid Mech.* **42**, 413 (2010).
- [16] T. A. Waigh, *Rep. Prog. Phys.* **68**, 685 (2005).
- [17] E. M. Furst, *Curr. Opin. Colloid Interface Sci.* **10**, 79 (2005).
- [18] T. M. Squires and J. F. Brady, *Phys. Fluids* **17**, 073101 (2005).
- [19] R. N. Zia and J. F. Brady, *J. Fluid Mech.* **658**, 188 (2010).
- [20] R. N. Zia and J. F. Brady, in *Theoretical Microrheology*, edited by S. Spagnolie (Springer, New York, 2014), Chap. 3.
- [21] A. Meyer, A. Marshall, B. G. Bush, and E. M. Furst, *J. Rheol.* **50**, 77 (2006).
- [22] T. G. Mason and D. A. Weitz, *Phys. Rev. Lett.* **74**, 1250 (1995).
- [23] M. T. Valentine, P. D. Kaplan, D. Thota, J. C. Crocker, T. Gisler, R. K. Prud'homme, M. Beck, and D. A. Weitz, *Phys. Rev. E* **64**, 061506 (2001).
- [24] T. G. Mason, *Rheol. Acta* **39**, 371 (2000).
- [25] J. C. Crocker and D. G. Grier, *J. Colloid Interf. Sci.* **179**, 298 (1996).
- [26] See K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. J. Schmid, and G. Danuser, *Nature Methods* **5**, 695 (2008) for an example of a particle-tracking method that determines particle trajectories from analysis of the entire data set.
- [27] J. E. Jackson, *A User's Guide to Principal Components* (Wiley, New York, 1991).
- [28] G. Saporta and N. Niang, Principal component analysis: application to statistical process control, in *Data Analysis*, edited by G. Govaert (Wiley, London, 2009), p. 123.
- [29] R. Liégeois, Structured sparse principal component analysis for fMRI imaging, M. Eng. thesis, University of Liège, 2011 (unpublished). [http://www.montefiore.ulg.ac.be/~rligeois/documents/TFE\\_Raphael\\_Liegeois.pdf](http://www.montefiore.ulg.ac.be/~rligeois/documents/TFE_Raphael_Liegeois.pdf).
- [30] Matlab, MathWorks Inc., Natick, MA.
- [31] M. Breakspear, S. Heitmann, and A. Daffertshofer, *Front. Hum. Neurosci.* **4**, 190 (2010).
- [32] R. W. Robinett, *Am. J. Phys.* **63**, 823 (1995).
- [33] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
- [34] A. Einstein, *Ann. Physik* **324**, 289 (1906).
- [35] E. R. Weeks and D. A. Weitz, *Chem. Phys.* **284**, 361 (2002).
- [36] F. K. Oppong, P. Coussot, and J. R. de Bruyn, *Phys. Rev. E* **78**, 021405 (2008).
- [37] F. K. Oppong and J. R. de Bruyn, *Rheol. Acta* **50**, 317 (2011).
- [38] J. P. Rich, G. H. McKinley, and P. S. Doyle, *J. Rheol.* **55**, 273 (2011).
- [39] T. H. Larsen and E. M. Furst, *Phys. Rev. Lett.* **100**, 146001 (2008).
- [40] R. J. Clarke, *IEE Proc.-F* **128**, 359 (1981).
- [41] R. Kouassi, P. Gouton, and M. Paindavoine, *Signal Proc.-Image Comm.* **16**, 541 (2001).