Length scales in Brownian yet non-Gaussian dynamics

José M. Miotto¹, Simone Pigolotti², Aleksei V. Chechkin³,⁴, and Sándalo Roldán-Vargas⁵,⁶ *

¹ Leiden Institute of Advanced Computer Science, 2333 CA Leiden, The Netherlands,
² Biological Complexity Unit, Okinawa Institute for Science and Technology and Graduate University, Onna, Okinawa 904-0495, Japan,
³ Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany,
⁴ Akhiezer Institute for Theoretical Physics, Kharkov 61108, Ukraine,
⁵ Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany,
⁶ Department of Applied Physics, Faculty of Sciences, University of Granada, 18071 Granada, Spain

According to the classical theory of Brownian motion, the mean squared displacement of diffusing particles evolves linearly with time whereas the distribution of their displacements is Gaussian. However, recent experiments on mesoscopic particle systems have discovered Brownian yet non-Gaussian regimes where diffusion coexists with an exponential tail in the distribution of displacements. Here we show that, contrary to the present theoretical understanding, the length scale \( \lambda \) associated to this exponential distribution does not necessarily scale in a diffusive way. Simulations of Lennard-Jones systems reveal a behavior \( \lambda \sim t^{1/3} \) in three dimensions and \( \lambda \sim t^{1/2} \) in two dimensions. We propose a scaling theory based on the idea of hopping motion to explain this result. In contrast, simulations of a tetrahedral gelling system, where particles interact by a non-isotropic potential, yield a temperature-dependent scaling of \( \lambda \). We interpret this behavior in terms of an intermittent hopping motion. Our findings link the Brownian yet non-Gaussian phenomenon with generic features of glassy dynamics and open new experimental perspectives on the class of molecular and supramolecular systems whose dynamics is ruled by rare events.

I. INTRODUCTION

In one of his celebrated 1905 papers, Einstein proposed a statistical interpretation of Robert Brown’s observations based on the corpuscular constitution of matter [1, 2]. Einstein’s theory predicted two concomitant properties for the probability density function (PDF) of displacements of the Brownian particles: its shape must be Gaussian and its variance, the mean squared displacement (MSD), must grow linearly (diffusively) with time. Since the seminal experiments conducted by Perrin more than one hundred years ago [3], these two predictions were routinely validated and the coexistence between Gaussianity and diffusivity became a paradigm. Exceptions to this long-standing paradigm were first observed in the realm of anomalous diffusion [4-8], where non-linear time dependencies of the MSD coexist with both Gaussian and non-Gaussian PDF’s of displacements [9-15].

Recent experiments have found a new class of counterexamples to this paradigm. Several mesoscopic particle systems present a time regime where linear diffusion coexists with a non-Gaussian PDF of displacements characterized by an exponential tail \( e^{-r/\lambda(t)} \) as a function of the displacement \( r \) [9,15]. The exponential tail is controlled by a time-dependent length scale, \( \lambda(t) \), which evolves as a power law: \( \lambda(t) \sim t^\beta \), with \( \beta > 0 \). This Brownian yet non-Gaussian regime appears in a variety of systems, including colloidal beads moving on the top of lipid tubes [9,10], nanospheres in entangled protein suspensions [9], binary mixtures of colloidal hard spheres [11], microspheres in biological hydrogels [12], and passive tracers in suspensions of eukaryotic swimmers [13]. Some of these works report values for \( \beta \) compatible with 1/2 [9-11,13] that have motivated theoretical models based on the idea of diffusing diffusivities [16-20]. These models assume a heterogeneous dynamics: particles move according to a time-dependent diffusion coefficient which leads to a Brownian yet non-Gaussian regime with exponential asymptotics [16-18]. In these models, the variance of the exponential, \( \lambda^2(t) \), is responsible for the total MSD: \( \lambda^2(t) \sim \text{MSD}(t) \sim t \). However, more recent experiments have measured a value of \( \beta \) significantly smaller than 1/2 [12], challenging an explanation based on diffusing diffusivities. Far from being an exotic phenomenon with an appealing historical background, this problem has deep implications in the understanding of a broad class of systems which are driven by rare events [10].

Here we study by computer simulations the equilibrium dynamics of representative models of glass-formers with isotropic and non-isotropic interactions [21,22]. Our main result is that the length scale \( \lambda \) associated with the Brownian yet non-Gaussian regime scales in a non-universal way. For systems where particles interact by an isotropic potential, the exponent controlling the evolution of the exponential tail is \( \beta = 1/d \), where \( d \) is the system dimension. We quantitatively explain this result by a scaling argument based on the idea of hopping motion [22,23]. In glass-formers where particles interact by a non-isotropic potential, the exponent \( \beta \) also depends on the temperature. This dependence becomes more obvious when the system, which behaves as a strong glass-
former [21], enters into the Arrhenius regime. We interpret this result as the consequence of an intermittent particle hopping motion at low temperatures. Finally, we show that the dynamics in the Brownian yet non-Gaussian regime for all the explored systems is the result of mixing the anomalous diffusion of individual particles.

II. EXPONENTIAL TAIL AND SYSTEM DIMENSION

We study by Molecular Dynamics simulations the equilibrium dynamics of isotropic Lennard-Jones particles in two [21] and three dimensions by the Kob-Andersen binary mixture [25] (see Methods). In both cases, the binary nature of the interaction potential is designed to resulting in a glassy system and, therefore, avoiding crystallization even at very low temperature. In particular, this system shows the super-Arrhenius behaviour characteristic of fragile glass-formers [21, 24]. This model also shares many fundamental dynamic and structural properties with other representative systems with isotropic interaction. For instance, its relaxation dynamic mechanism (including the emergence of dynamic heterogeneities) and structural patterns are similar to those observed in other molecular liquids [23, 20], colloidal hard spheres [27, 29], models of soft spheres [30], and even granular materials [23]. We cover for this system a wide range of temperatures, from the liquid state (slightly above the system onset temperature) to 1.03 \( T_c \), where \( T_c \) is the estimated system Mode Coupling Temperature (see Methods).

We first measure the particle MSD, see Figure 1a and 1b. For all temperatures, the MSD presents a characteristic short-time local ballistic motion (\( \sim t^2 \)) and a long-time diffusive regime (\( \sim t \)). Upon cooling, the MSD develops a plateau at intermediate times resulting from an increasingly long local residence time, a common feature of all glass-forming liquids [22]. We characterize the ensemble distribution of displacements by means of the self part of the van Hove function [31]

\[
G_s(\vec{r}; t) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \delta[\vec{r} - \Delta\vec{r}_i(t)] \right\rangle,
\]

where \( G_s(\vec{r}; t) \) is the fraction of particles (from a total number \( N \)) which have displaced by \( \Delta\vec{r}_i(t) = \vec{r}_i(t) - \vec{r}_i(0) \) in a time \( t \). Since both systems are isotropic, we explore the distribution of displacements as a function of the radial coordinate \( r = |\vec{r}| \) and define:

\[
P(r; t) \equiv W(r; t)/r^{d-1} = \phi_d G_s(r; t); \quad d \in \{2, 3\}
\]

where \( W(r; t) \) is the PDF to have a radial displacement \( r \), \( d \) the system dimension, and \( \phi_d \) the Jacobian angular prefactor (e.g., \( \phi_2 = 2\pi \) and \( \phi_3 = 4\pi \)) [31]. With this definition we count the fraction of particles which have displaced radially by \( r \) normalized (up to prefactors) by \( r^{d-1} \) to account for the volume of the shell within the range \( r \) to \( r + dr \). In this context, a purely Gaussian diffusion would result in \( P(r; t) \sim e^{-r^2/4D(T)t} \), being \( D(T) \) the temperature-dependent diffusion coefficient.

![Figure 1. MSD and exponential tails for the Kob-Andersen binary mixture. a) MSD for the 3D system at different temperatures from the liquid state to the deep supercooled regime: \( T = 0.7, 0.6, 0.54, 0.50, 0.475, \) and 0.45; b) MSD for the 2D system at different temperatures covering a similar range as in a): \( T = 0.6, 0.50, 0.45, 0.40, 0.36, \) and 0.34. The unit of length is equal to the particle diameter (see Methods for the definition of the time and temperature units). Normalized self part of the van Hove function, \( P(r; t) \), for the 3D system at \( T = 0.45 \) (c) and for the 2D system at \( T = 0.40 \) (d), where the selected times are marked dots in the corresponding MSD in a) and b). Dashed lines in c) and d) show the extent of the exponential range. Characteristic length \( \lambda \) as a function of time for the 3D (e) and 2D (f) systems: in both cases, temperatures are as in a) and b) respectively. Again, dots signal those times corresponding to \( P(r; t) \) in c) and d).

The distribution \( P(r; t) \) presents a Gaussian behaviour at short distances and an exponential decay at large distances: \( P(r; t) \sim e^{-r/\lambda(t)} \), where \( \lambda(t) \) is a characteristic length that increases with time (Fig. 1c and 1d). We observe this exponential decay both at low (3D, Fig. 1c) and intermediate temperatures (2D, Fig. 1d). Qualitatively similar behaviours to those presented in Fig. 1c and 1d appear at all the explored temperatures. This observation is compatible with previous works on
glassy systems [23], predicting exponential decays (with logarithmic corrections at large r) for the PDF’s of displacements.

To discriminate the r-ranges corresponding to the Gaussian and exponential regimes of $P(r; t)$ and, therefore, accurately estimate $\lambda(t)$, we implemented a non-parametric Kolmogorov-Smirnov test (see Methods). The distinction between the two regimes is more evident at low temperature (in particular at short and intermediate times), being the two regimes separated by an inflection point in $P(r; t)$. At longer times, the Gaussian range extends up to larger distance and therefore becomes dominant. In turn, the range of the exponential tail shrinks and thus becomes marginal at long times. Eventually, this takeover leads to a purely Gaussian distribution of displacements as prescribed by the Central Limit Theorem (CLT).

As for the MSD, the growth of $\lambda$ with time is characterized by three distinct regimes (Figs. 1e and 1f). At short times, when the MSD is ballistic, $\lambda(t)$ rapidly increases. Displacements within this short time are typically smaller than the particle diameter (Figs. 1c and 1d) and, therefore, reflect the particle local heterogeneous dynamics. At intermediate times, when the MSD manifests a plateau, $\lambda(t)$ increases in a slower way. In this regime, only a small fraction of particles has abandoned their initial local cage and are able to jump over distances on the order of few particle diameters (Fig. 1c and 1d). At long times, $\lambda(t)$ increases steeply again. In the same time regime, the MSD is compatible with diffusive behaviour (Fig. 1a and 1b). In the following, we focus on this Brownian yet non-Gaussian regime [10].

Our first goal is to explore the scaling of $\lambda(t)$ within the Brownian yet-non Gaussian regime. We define the time $t_0$ at which the 2D and 3D Lennard-Jones systems start to be diffusive ($\text{MSD}(t \geq t_0) \sim t$). In practice, we define $t_0$ as the time at which the exponent $\mu$ characterizing the scaling of the MSD with time is equal to 1 up to a tolerance of 5%, see Methods for details. We compare the scaling with time of $\text{MSD}(t)^{1/2}$ and $\lambda(t)$ in the time regime $t \geq t_0$ for different temperatures (Fig. 2). In both systems, the scaling of the MSD is diffusive, $\text{MSD}^{1/2}(t \geq t_0) \sim t^{1/2}$. In contrast, the scaling of $\lambda(t)$ is dimension dependent:

$$\lambda(t \geq t_0) \sim t^{\beta(d)}$$ (3)

The exponent $\beta$ does not appreciably depend on the temperature within the diffusive time window $t \geq t_0$. However, $\beta$ clearly depends on the system dimension, with a value compatible with $\beta = 1/2$ in two dimensions and $\beta = 1/3$ in three dimensions. While the value $\beta = 1/2$ in two dimensions is consistent with the diffusive scaling of the MSD, the scaling of $\lambda$ in three dimensions does not seem to be trivially related with that of the MSD. The fact that the scaling of $\lambda$ and the MSD are not in general trivially linked descends from the fact that the weight of the exponential tail compared to the rest of the distribution is time-dependent. While $\lambda$ is a property of the exponential tail only, the MSD is determined by the entire distribution. In this respect, the relative contribution of the exponential tail to the MSD diminishes, with the start of the exponential range moving to larger and larger displacements, as time increases. The scaling of $\lambda$ also contrasts with the interpretation of the Brownian yet non-Gaussian regime based on diffusing diffusivities [10, 16, 18], where $\lambda^2(t) \sim \text{MSD}(t) \sim t$. However, the time evolution we observe for $\lambda(t \geq t_0)$ in three dimensions agrees with recent experiments on microspheres diffusing in biological gels [12], where $\beta < 1/2$ and the total PDF of displacements presents a Gaussian core, which grows with time, and a marginal exponential tail, which tends to disappear with time.

![Figure 2](image-url)

**Figure 2. Brownian yet non-Gaussian regime for the Kob-Andersen binary mixture.** Double log plot of MSD$^{1/2}$ (squares) and $\lambda$ (circles) as a function of time for the 3D (a) and 2D (b) Kob-Andersen systems at different temperatures: $T = 0.45$, $0.5$, and $0.6$ (3D), and $T = 0.4$, $0.5$, and $0.6$ (2D). Here $t_0$ is the temperature- and system- dependent time at which each system reaches a $\mu$ value compatible with 1 within numerical uncertainty, i.e. $\text{MSD}(t \geq t_0) \sim t^{\mu \leq 1}$. To determine $t_0$ we numerically compute the local exponent $\mu(t)$ controlling the time evolution of the MSD$(t)$, i.e. MSD$(t) \sim t^{\mu(t)}$ (see Methods). Dashed lines in a) and b) serve as a reference.
These observations suggest a general dependence \( \beta(d) = 1/d \), with \( d \in \{2, 3\} \). We propose a theory for such scaling behaviour as an outcome of hopping motion, a signature of all glass-forming liquids below the onset temperature \cite{22}. The exponential tail of \( P(r; t) \) originates from a minority of particles able to escape from their initial cage and perform a large displacement for times of the order of \( t_0 \). We call these particles \textit{hoppers} and denote by \( N_h(t) \) their number at time \( t \). Since the system is at equilibrium, we assume that hoppers escape from their cage at a constant rate \( \omega \):

\[
N_h(t) \approx \omega N t \sim t. \tag{4}
\]

Equation (4) is valid in a time range where \( N_h(t) \ll N \), being, therefore, \( \omega^{-1} \gg t \). We now use that the functional form of the mass density of hoppers, \( \rho_h(r; t) \), as a function of the distance at time \( t \) coincides with that of \( P(r; t) \) for \( t \approx t_0 \) and is, therefore, exponential:

\[
P(r; t) \sim \rho_h(r; t) = \rho_0 e^{-r/\lambda(t)} ; \quad t \approx t_0 \tag{5}
\]

We consider \( \rho_0 \) to be independent of time for times of the order of \( t_0 \), i.e. the source of hoppers at \( r = 0 \) remains at constant density for \( t \approx t_0 \) since \( N_h(t) \ll N \). This is consistent with the behaviour of \( P(r; t) \) observed in Fig. 1c and 1d, and also clearly manifested by the \( P(r; t) \) shown in Appendix C, where all the exponential tails at different times cross at a common value at \( r \approx 0 \).

The number of hoppers at a time \( t \approx t_0 \) therefore scales as

\[
N_h(t) \sim \int_0^\infty r^{d-1} \rho_h(r; t) \, dr \sim \int_0^\infty r^{d-1} e^{-r/\lambda(t)} \, dr \sim \lambda^d(t) \tag{6}
\]

Combining Eq. (4) and Eq. (6) we finally obtain

\[
\lambda(t) \sim t^{1/d} \tag{7}
\]

The scaling argument embodied in Eq (4) to (6), and leading to Eq. (7) would, in principle, hold for all \( d \). In particular, we have also tested Eq. (7) in Appendix A for \( d = 4 \) for the Lennard-Jones system investigated, also confirming its validity. Nevertheless, the general extension of our scaling argument needs subsequent investigation. In particular, works based on a Mode Coupling Theory formalism predict changes in the dynamics of four-point correlators at higher dimensions whose hypothetical effect on the observables studied here should be investigated \cite{32}. Apart from that, and due to the observed emergence of finite size effects in 2D simulations \cite{33,35}, we have tested and confirmed Eq. (7) for a much more larger 2D-Lennard-Jones system in Appendix B, in accord with recent experiments in 2D systems \cite{36}.

Summarizing, our argument to explain \( \lambda(t) \sim t^{1/d} \) is based on three assumptions: (1) hoppers leave their local environment at a fixed hopping rate (Eq. (4)); (2) hoppers fill the space isotropically (Eq. (6)); (3) the restriction to a time regime \( t \approx t_0 \) where \( N_h(t) \ll N \). We test these assumptions individually for \( d = 2, 3 \) in Appendix C. Once the majority of the particles have abandoned their initial position \( (t \gg t_0) \), and present a statistically equivalent jumping record, the last assumption breaks. At such a long time, the PDF of displacements results from a large number of independent and equally distributed displacements and is, therefore, Gaussian by virtue of the CLT.

### III. EXPONENTIAL TAIL AND NON-ISOTROPIC INTERACTIONS

We now investigate the evolution of \( \lambda(t) \) when particles interact by a non-isotropic potential. To this purpose we consider a three-dimensional system of tetravalent \textit{patchy particles} \cite{37,39}. Such particles have four sticky spots tetrahedrally distributed on their surface which provide a strongly directional interaction with fixed valence. Upon cooling, the dynamics slows down and the system develops an amorphous, highly connected tetrahedral network \cite{40}. This model has already shown its capability for capturing some fundamental dynamic and structural features, such as the emergence of an amorphous tetrahedral order, present in different classical systems with non-isotropic interactions such as atomistic models of water, silicon, and silica \cite{37,40,43,46}. We investigate the equilibrium dynamics of this system, which shows the Arrhenius behaviour characteristic of a strong glass-former \cite{21,40}, at moderate density by Brownian Dynamics simulations (see Methods) and explore the emergence of the Brownian yet non-Gaussian regime. Our study covers a wide \( T \)-range from the liquid state (slightly above the percolation threshold) to the deep Arrhenius regime, where the great majority of the particles are tightly bound to their neighbors \cite{37}.

In contrast to the Lennard-Jones systems, the scaling of \( \lambda(t) \) clearly depends on \( T \) for the tetrahedral gelling system (Figure 3). While at high temperature \( \lambda(t) \sim t^{1/3} \), compatible with the 3D Lennard-Jones system, the exponent \( \beta \) decreases upon cooling the system below the Arrhenius temperature \cite{37,40}, reaching values \( \beta \approx 1/6 \) for temperatures deeply into the Arrhenius regime. Thus our argument leading to \( \beta(d) = 1/d \) independently of \( T \) (Eq. (4) to (7)) does not hold for the gelling system at low \( T \). To understand this discrepancy, we individually test the hypotheses underlying our scaling theory (see Appendix C). We
We summarize the behaviour of the exponent $\beta(T; d)$ for the three systems (see Figure 4). For the two Lennard-Jones systems, $\beta$ markedly depends on the system dimension (Eq. (3)) and is practically independent of $T$ (with, at most, a very modest decrease in 3D). Our theoretical argument (Eqs. (4) to (7)) predicts that other systems with isotropic interactions, such as hard and soft spheres, should present the same behaviour. Instead, $\beta$ clearly depends on temperature for the tetrahedral gelling system. This dependence is stronger once the system enters into the Arrhenius regime, where the diffusion coefficient decreases exponentially upon cooling [40] (inset in Fig.4), thereby revealing a connection between a defining feature of the dynamics observed in strong glass-formers and the emergence of rare events. At higher temperatures, where connectivity is low [37], $\beta$ attains a value compatible with that of the 3D Lennard-Jones system ($\approx 1/3$).

IV. NON-GAUSSIANITY AND DYNAMICS BY POPULATIONS

We now look further into the coexistence between linear diffusion and non-Gaussian distributions of displacements in 3D. To this aim, we discriminate the particles into populations according to their potential energy $E$ at $t = 0$ as in Ref. [40]. For the tetrahedral gelling system, this procedure leads to five different particle populations characterized by the number of bonds per particle ($i \in \{0, 1, 2, 3, 4\}$) at $t = 0$, being $E|_{i}$ the potential energy of population $i$ (where $E|i > E|j$ when $i < j$) [40]. For the 3D Lennard-Jones system we arbitrarily define two populations with a marked difference in their initial potential energies: one including the 1% of particles with the highest potential energy and the other one including the 1% of particles with the lowest potential energy, both at $t = 0$.

The time evolution of the MSD for the different populations at fixed $T$ is presented in Figures 5a and 5b. Populations with high potential energy at $t = 0$ show a super-diffusive regime at short times which is compensated with the sub-diffusive motion associated to
the populations with low potential energy. In particular, for the tetrahedral system we see that this sub-diffusive motion is dominant and represents almost the total MSD. The reason is that, for the chosen temperature, the system is already highly percolated [31] and the populations having a low energy at \( t = 0\) are much more larger (number of particles) than the populations having a high energy at \( t = 0\). Thus, while particles with high potential energy are briefly connected to the gel network, low energy particles are retained for a longer time to the network, resulting in a plausible source for sub-diffusion. At intermediate times, the distinct populations reach different values of the MSD at their respective plateaus. This difference is more pronounced for the tetrahedral liquid due to its lower density. When abandoning the plateau, populations with high potential energy at \( t = 0\) show sub-diffusive motion while populations with low potential energy at \( t = 0\) present super-diffusive motion. This is particularly clear when looking at the \( \mu(t) \) exponent obtained from the slope of the MSD: when leaving the plateau, populations starting with a high potential energy present \( \mu(t) < 1\) whereas populations starting with a low potential energy show \( \mu(t) > 1\) (Figures 5c and 5d).

The mixing of anomalous diffusions is concomitant with a change in the energy of each population: particles starting with a high (low) potential energy show a decrease (increase) in their energy which slows down (boosts) their dynamics (Figures 5e and 5f). At long times, once the particles lose the memory of their initial state and pass through all the possible energy states, all the populations show the same average energy while their dynamics converge to a common diffusive trend. It is worth highlighting that this convergence happens at very long times pointing out to the existence of a long term memory of the potential energy. In addition, this convergence shows a common relaxation time which is independent of the particle population. All this phenomenology reduces to a simple intuition: when a particle is in a high (low) energy state it moves faster (slower) than the average. For large times, when all the particles have passed through all the possible fast and slow states, the sampled distributions of displacements for all populations are equivalent and the total PDF of displacements becomes Gaussian as dictated by the CLT.

Before collapsing into a common diffusive trend, there exists a time window where the different populations still show \( \mu(t)|i \neq 1 \forall i \), with \( \mu(t) \leq 1\) for the total number of particles (shaded region in Fig. 5). This is the Brownian yet non-Gaussian window where the exponential tail is still detectable despite the system as a whole shows \( \mu(t) \leq 1\). This observation helps understanding why within this diffusive time window, \( P(r; t)\) is not necessarily Gaussian: the distinct populations have not yet converged and, therefore, their sampled dynamic states are not yet equivalent, resulting in a non-Gaussian PDF of displacements. Our results show that the observed exponential tail originates from mixing anomalies: within the Brownian yet non-Gaussian regime each individual population shows its own anomalous diffusion \( (\mu(t)|i \neq 1)\). We have further shown that, within the Brownian yet non-Gaussian regime, the anomalous diffusion of each population (which is a dynamic feature) can be associated to its initial potential energy (which is a structural feature).

![Figure 5. Dynamics by particle populations.](image)

Left column corresponds to the 3D Kob-Andersen binary mixture and shows the time evolution of different observables discriminated into two particle populations: a) MSD(\( \mu(t)\) (MSD(\( t) \sim \mu(t)\)), and (c) difference between the average potential energy of each population, \( E[E|i]\), and the system average potential energy, \( E[E]\). Yellow (green) lines correspond to the population that includes the 1% of the particles with the highest (lowest) potential energy at \( t = 0\). Right column corresponds to the tetrahedral gelling system and shows the same observables as those appearing in the left column (b, d, and f). The five colored lines correspond to the five populations having a different number of bonds (\( i \in \{0, 1, 2, 3, 4\}\) at \( t = 0\) (lines from yellow to green). Black lines in the whole figure represent observables averaged over all particles. The shaded regions mark out the extent of the Brownian yet non-Gaussian time window, where the exponential tail is still detectable and \( \mu(t) \leq 1\) for the total number of particles. Inset in a) shows a detail of the MSD corresponding to the shaded time window. The chosen temperatures correspond to the deep supercooled regime (3D Kob-Andersen binary mixture) and the Arrhenius regime (tetrahedral liquid). For both systems, the different populations evolve in such a way that the corresponding observables pass from a short-time behaviour (where the particles still retain the information of their initial mechanical state) to a long-time regime, where the particles have lost memory of their initial state and the different populations converge to a common trend.
V. CONCLUSIONS AND PERSPECTIVES

Brownian yet non-Gaussian transport has recently attracted great interest for its unexpected but ubiquitous presence. It is at the heart of the more general problem of understanding, in a comprehensive way, the rare event dynamics present in many complex systems. Our study brings new insights into this appealing phenomenon. First, we have shown numerically that for canonical liquid models where particles interact by an isotropic potential, the exponent controlling the evolution of the observed exponential tail is not universal but depends on the system dimension: $\beta(d) = 1/d$. This important finding contrasts with the universality predicted by the current theoretical interpretations. We have rationalized this observation by a theoretical scaling argument which does not depend on the specific functional behaviour of the interaction potential. For that reason we expect other representative systems with significantly different potentials, such as hard and soft spheres, to show a similar scaling behaviour. Second, we have shown that dimension is not the only factor affecting the evolution of the exponential tail. In systems where particles interact by a non-isotropic potential, $\beta$ also depends on temperature. Contrary to the isotropic models investigated in this work (which behave as fragile glass-formers), the specific non-isotropic system investigated here behaves as a strong glass-former. Therefore, the temperature dependence shown by $\beta$ for the non-isotropic system reveals a new fundamental feature discriminating the dynamics of fragile and strong glass-formers, with far-reaching consequences for understanding and classifying glasses. It also establishes a connection between the emergence of the generic Arrhenius regime representative of strong glass-formers and their hopping dynamic mechanism. Third, we have shown that the time regime where the Brownian yet non-Gaussian transport occurs is characterized by a mixed anomalous diffusion of different particle populations. These populations are characterized by a configurational property: their potential energy. We expect our findings for the models investigated here to hold for a large variety of systems. In this respect, the canonical systems we have chosen to study share a common fundamental structural and dynamic phenomenology with other representative systems defined by different isotropic and non-isotropic potentials, e.g. hard or soft spheres [22, 29, 30], and water, silicon, or silica [13, 40].

Taken in a broad sense, our results bring new research perspectives, imposing severe constraints to future theories and calling for new experiments. Future theoretical models should account for the dimension and temperature dependencies that we observed in this work. For instance, the seminal work by Chaudhury and coworkers [23] and the large deviation model proposed by Barkai and Burov [48], predict exponential tails in the PDF of displacements (with logarithmic corrections at large $r$) using a Continuous Time Random Walk (CTRW) formalism. These two approaches do not suggest a dimension-dependent behavior of the characteristic length $\lambda \sim t^{1/d}$. Other works employing the idea of diffusing diffusivity [16–20, 49] lack a dependence on dimension as well. Similar conclusions hold for other approaches to Brownian yet non-Gaussian diffusion, e.g. models mixing CTRW and diffusing diffusivity [50], polymerization models where the fluctuating size of a diffuser generates a diffusing diffusivity process [51, 52], diffusion in a fluctuating corrugated channel [53], and colloidal Brownian yet non-Gaussian diffusion induced by nonequilibrium noise [54]. Future models should also explain the mixed anomalous diffusion observed in our numerical simulations during the Brownian yet non-Gaussian time regime. In our view, this is a crucial yet challenging task. We speculate on the idea of incorporating a CTRW formalism with correlated jumps, one feature that is usually missing in previous works.

From a practical perspective, our results can a priori be tested in several real systems regulated by different control parameters, not necessarily temperature. For instance, we expect colloidal glass- and gel-forming liquids [22, 23] to present a similar phenomenology as the one described in this paper when studied as a function of other parameters, e.g. packing fraction. In addition, non-Gaussian tails have also been reported recently in granular systems [55], where shear stress plays the role that temperature plays in the systems we studied here. Complex biological media, controlled by pH, have also shown Brownian yet non-Gaussian transport, where the measured value of $\beta$ is compatible with our results [12]. Finally, we expect a similar phenomenology as the one reported in this work in non-equilibrium active systems, where transport is controlled by an energy of metabolic origin [56, 57]. In this respect, non-Gaussian regimes have indeed been found in active biological systems constituted by animate, energy-consuming, particles, for example in cell migration processes [58] and in the transport of different organelles in the cytoplasm of animal cells [59].

VI. METHODS

A. Lennard-Jones system

We performed two and three dimensional Molecular Dynamics simulations of a Kob-Andersen binary mixture [24]. In both cases we first run simulations in the canonical ensemble to equilibrate the system at a fixed temperature. From them, we run simulations in the microcanonical ensemble to evaluate the dynamic observables presented in this article. The interaction between a particle of species $\alpha$ and a particle of species $\beta$ is given by the Lennard-Jones potential:
where $A$ and $B$ are the labels for the two species and $r$ is the distance between the centers of mass of the two particles. For both systems, $\sigma_{AA} = 1$, $\sigma_{AB} = 0.8$, $\sigma_{BB} = 0.88$, $\epsilon_{AA} = 1$, $\epsilon_{AB} = 1.5$, and $\epsilon_{BB} = 0.5$. The potential is truncated and shifted at $r = 2.5\sigma_{\alpha\beta}$ [25]. All the results are given in reduced units, where $\sigma_{AA}$ is the unit of length, $\epsilon_{AA}$ the unit of energy, and $\sigma_{AA}\sqrt{m/48\epsilon_{AA}}$ the unit of time (being $m = 1$ the mass of the particles). Temperature, $T$, is controlled during the equilibration process by an Andersen thermostat with an effective mass of 48 reduced units [24] with Boltzmann’s constant set to 1.

The number of particles for each species of the 2D system are $N_{AA}^D = 6500$ and $N_{BB}^D = 3500$ with a total number density $\rho^{2D} = (N_{AA}^D + N_{BB}^D)/L^2 = 1.16$, being $L = 92.78$ the length of the square simulation box [24]. For the 3D system we used a different composition with $N_{AA}^D = 6400$ and $N_{BB}^D = 1600$ for a total number density $\rho^{3D} = (N_{AA}^D + N_{BB}^D)/L^3 = 1.20$, being $L = 18.80$ the length of the cubic box [25]. These compositions avoid the emergence of a crystal structure even at very low temperature. We covered temperature ranges $T \in [0.34,0.75]$ (2D) and $T \in [0.45,0.9]$ (3D). For both systems (2D and 3D), the lowest temperature investigated is 1.03 $T_c$, where $T_c$ is the Mode Coupling Temperature of the glass transition as estimated from data coming from numerical simulations [24, 60]. The onset temperatures for the 2D and 3D systems are: $T^{2D}_s = 0.75$ [61] and $T^{3D}_s = 0.9$ [62] (Fig. 4 in the main text). Both systems show at low $T$ the super-Arrhenius dynamic behaviour characteristic of fragile glass-formers [24]. We used for both systems a Velocity Verlet algorithm with a time step depending on the temperature: $\delta t^{2D} = 0.02$ ($\delta t^{3D} = 0.01$) for $T \leq 0.6$ ($T > 0.6$) and $\delta t^{3D} = 0.02$ ($\delta t^{3D} = 0.01$) for $T \leq 0.7$ ($T > 0.7$). The runs extended over $10^7$ and $5 \cdot 10^7$ time steps for the 2D and 3D system respectively.

The results shown in this article correspond in both cases to particles of the species $A$. For the 2D system, we averaged over 50 independent simulations (around 325.000 individual particle trajectories) for each value of the temperature running on a high-end CPU processor cluster with a total amount of CPU time of 11.5 years. For the 3D system, we averaged over 70 independent simulations (around 450.000 individual particle trajectories) for each value of the temperature with a total amount of CPU time of 18.5 years.

### B. Tetrahedral system

We performed three dimensional Brownian Dynamics simulations of tetravalent patchy particles in the canonical ensemble. The number of particles is $N = 10000$ with a length of the cubic simulation box $L = 25.98\sigma$, being $\sigma$ the particle hard sphere-like diameter, here taken as the unit of length. The number density is $\rho = N/L^3 = 0.57$. For this density, the system develops a homogeneous amorphous tetrahedral network even at very low temperature [40, 63]. The interaction potential comprises a spherical steep repulsion and a short-range attraction. The interaction between a generic pair of particles 1 and 2 is given by:

$$V(1,2) = V_{CM}(1,2) + V_P(1,2)$$

where $V_{CM}(1,2)$ is the repulsive part of the potential between particles 1 and 2 whereas $V_P(1,2)$ is the attractive part of the potential between the patches of particles 1 and 2. These potentials are modeled as follows:

$$V_{CM}(12) = \left(\frac{\sigma}{r_{12}}\right)^p$$

$$V_P(12) = -\sum_{i=1}^{M} \sum_{j=1}^{M} \epsilon \exp \left[-\frac{1}{2} \left(\frac{r_{ij}}{\alpha}\right)^\eta\right]$$

Here $r_{12}$ is the distance between the center of mass of particles 1 and $2$, $r_{ij}$ is the distance between patch $i$ on particle 1 and patch $j$ on particle 2, and $M = 4$ is the number of patches per particle. The four patches are tetrahedrally distributed on the surface of the particles. Exponents in $V_{CM}(1,2)$ and $V_P(1,2)$ are taken as $p = 200$ and $q = 10$ to resemble the functional behaviours of a hard sphere and a square well interaction respectively. We selected $\alpha = 0.12$ as the patch diameter to avoid having more than one bond per patch and $\epsilon = 1.001$, for which the minimum of the attractive part of the potential energy in a bounded configuration is $u_0 = \min V_P(12) = -1$. Temperature $T$ is measured in units of the potential well, taking Boltzmann’s constant as 1. The unit of time is $\sigma\sqrt{m/|u_0|}$, being $m = 1$ the mass of the particles. To integrate the equations of motion we used a Velocity Verlet algorithm with a fixed time step $\delta t = 0.001$ using a modified Brownian thermostat which explicitly avoids unphysical decorrelations in the particle velocity [68, 64]. The longest runs extended over $7 \cdot 10^9$ time steps for the lowest temperatures ($T = 0.105$ and $T = 0.1025$). All simulations were performed using the oxDNA simulation package running on GPUs [65].

We investigated temperatures within the range $T \in [0.1025,0.16]$, covering a slowing down of the
dynamics of four orders of magnitude. The dynamics of this system exhibits Arrhenius behaviour at low temperature, a signature of strong glass-formers \cite{40}. The temperature at which the system enters into the Arrhenius regime is $T_A \approx 0.115$ \cite{37, 40} (Fig. 4 in the main text). We take as the onset temperature for this system $T_0 = 0.16$ (Fig. 4 in the main text). Above this temperature the system is not completely percolated into an infinite cluster and the MSD does not present any detectable plateau \cite{37}.

We evaluated dynamic observables by an average of 50 independent simulations (500,000 individual particle trajectories) per each value of the temperature. The total amount of GPU time was approximately 2 months, which for this system corresponds to more than 3 years of CPU time on a high-end processor \cite{65}.

C. Estimation of $\lambda(t)$

We determined $\lambda(t)$ via maximum likelihood assuming an exponential probability density of displacements from a given threshold $r_t$. The likelihood at time $t$ is given by:

$$L_t = \prod_{i=1}^{M_t} \exp \left( -\frac{r_i - r_t}{\lambda_t} \right) / \lambda_t$$

where $M_t$ is the number of individual particle trajectories with a displacement $r_i$ larger than $r_t$ at time $t$. We considered for convenience the minus logarithm of the likelihood,

$$l_t = -\log L_t = \sum_{i=1}^{M_t} \frac{r_i - r_t}{\lambda_t} + M_t \log \lambda_t.$$

We then obtained the $\hat{\lambda}_t$ that minimizes $l_t$ with respect to $\lambda_t$:

$$\frac{\partial l_t}{\partial \lambda_t} \bigg|_{\lambda_t = \hat{\lambda}_t} = 0 \Rightarrow \hat{\lambda}_t = \frac{1}{M_t} \sum_{i=1}^{M_t} (r_i - r_t)$$

It is desirable to have the lowest possible threshold $r_t$ so that more data are included in the estimation while still having a statistically significant exponential tail. In order to achieve this, we used a standard iterative procedure for power-law tail evaluation \cite{66}. We started with a very large threshold $r_t$ and computed $\hat{\lambda}_t$ by means of Eq. 14. At this point, we performed a test of statistical significance: if the test is passed, \textit{i.e.} the data are well represented by the estimated exponential distribution, we decreased the threshold $r_t$ by 0.025$\sigma$ (being $\sigma$ the particle diameter) and repeated the procedure; if the test fails, we stopped the procedure and took $\lambda(t) = \hat{\lambda}_t$.

The statistical test we used is the non-parametric Kolmogorov-Smirnov test \cite{67}, which measures how extreme is the distance between the theoretical cumulative distribution under the null hypothesis (where the data are sampled from the theoretical distribution) and the empirical cumulative distribution. The distance between distributions is measured by the Kolmogorov-Smirnov statistics,

$$KS_{\lambda_t} = \sup_{r_t(t)} \left| F_{\text{null}}(r_t(t)) - F_{\text{emp}}(r_t(t)) \right|$$

where $F_{\text{null}}(x)$ is the cumulative distribution under the null hypothesis (exponential distribution with exponential rate $\lambda_t$) and $F_{\text{emp}}(x)$ the empirical cumulative distribution. The null hypothesis was accepted (rejected) when the $p$-value was greater (smaller) than 0.05 \cite{67}.

D. Estimation of $\mu(t)$, $t_0$, and $\beta$

We define $y_t \equiv \text{MSD}(t)$. Since each $y_t$ is obtained as an average over many independent, and equivalent, simulations (see previous sections), we assumed that it is normally distributed. We denote by $\bar{y}_{1,T}$ and $\sigma_{1,T}$ the mean and standard deviation of this distribution for a discrete time interval $\{t_1, ..., t_T\}$. The corresponding likelihood is

$$L_{1,T} = \prod_{t=t_1}^{t_T} \exp \left( -\frac{(y_t - \bar{y}_{1,T})^2}{2\sigma_{1,T}^2} \right) / \sqrt{2\pi\sigma_{1,T}^2}$$

As in the previous section, we considered the minus logarithm of the likelihood

$$l_{1,T} = -\log L_{1,T} =$$

$$= \sum_{t=t_1}^{t_T} \frac{(y_t - \bar{y}_{1,T})^2}{2\sigma_{1,T}^2} + T/2 \log 2\pi + T \log \sigma_{1,T}$$

We then assumed that $\bar{y}_{1,T}$ and $\sigma_{1,T}$ scale with the same power law exponent, $\mu_{1,T}$, within the discrete time interval $\{t_1, ..., t_T\}$: $\bar{y}_{1,T} = at^{\mu_{1,T}}$ and $\sigma_{1,T} = ct^{\mu_{1,T}}$, with $a, c \in \mathbb{R}^+$. This assumption is valid when deviations from normality are small, where the variance $\sigma_{1,T}^2$ scales as $2y_{1,T}^2/(T-1)$ and, therefore, $\sigma_{1,T} \sim \bar{y}_{1,T}$. This results in

$$l_{1,T} = \sum_{t=t_1}^{t_T} \frac{(y_t - at^{\mu_{1,T}})^2}{2c^2t^{2\mu_{1,T}}} +$$

$$+ \frac{T}{2} \log 2\pi + T \log c + \mu_{1,T}T \log t$$
This function was minimized with respect $\mu_{1,T}$, $a$, and $c$ via the L-BFGS-B algorithm implemented in scipy.optimize.minimize [68], and following the procedure detailed in [69].

In Eq. 18 $\mu_{1,T}$ corresponds to the discrete window given by $\{t_1,...,t_T\}$, which was chosen to cover one decade in time. To associate $\mu_{1,T}$ to a specific time, we chose the middle time in the sequence, i.e. $\mu_{1,T}(t_{(1+T)/2})$, being $T$ an odd natural number. In this way, the value $t_0$ associated to the entrance to the diffusive regime is the middle time in that sequence $\{t_1,...,t_T\}$ for which $\mu_{1,T} = \mu(t_{(1+T)/2}) \equiv \mu(t_0) = 0.95$. The reason for this choice is that $\mu = 1$ is an asymptotic limit ($t \to \infty$) which is approached from below but never reached. The criterion $\mu = 0.95$ ensures that the exponent $\mu$ is compatible with $\mu = 1$ up to our measured standard error, which is equal to 0.05 on average.

Finally, to estimate $\beta$ we considered time windows of one decade starting at $t_0$ (see Figs. 2 and 3), with the exception of those simulations where the Brownian yet non-Gaussian regime does not cover one decade in time (see for instance $T = 0.1025$ in Fig. 3). Then we fitted $\lambda(t \geq t_0)$ by a least-squares regression assuming a power law $\lambda(t \geq t_0) = a_1 t^2 + a_2$, with $a_1, a_2 \in \mathbb{R}^+$ as free parameters.

ACKNOWLEDGEMENTS

We thank Walter Kob for fruitful exchanges about this work and Lorenzo Rovigatti for his assistance with the tetrahedral system simulations. We also thank the anonymous reviewers for their careful reading of our manuscript and their insightful comments and suggestions. A.V.C. acknowledges partial financial support from the Deutsche Forschungsgemeinschaft (DFG Grant ME 1535/7-1). S.R.-V. acknowledges support from the European Commission through the Marie Skłodowska-Curie Individual Fellowship 840195-ARIADNE.
[62] D. Coslovich and G. Pastore, Understanding fragility in supercooled Lennard-Jones mixtures. I. Locally preferred structures, J. Chem. Phys. 127, 124504 (2007).

[63] C. De Michele, S. Gabrielli, P. Tartaglia, and F. Sciortino, Dynamics in the Presence of Attractive Patchy Interactions, J. Phys. Chem. B 110, 8064 (2006).

[64] J. Russo, P. Tartaglia, and F. Sciortino, Reversible gels of patchy particles: Role of the valence, J. Chem. Phys. 131, 014504 (2009).

[65] L. Rovigatti, P. Sulc, Z. R. Istv, and F. Romano, A Comparison between Parallelization Approaches in Molecular Dynamics Simulations on GPUs, J. Comput. Chem. 36, 1 (2015).

[66] A. Clauset, C. R. Shalizi, and M. E. Newman, Power-law distributions in empirical data, SIAM review 51, 661 (2009).

[67] F. A. Graybill, D. C. Boes, and A. M. Mood, Introduction to the Theory of Statistics (McGraw-Hill, 1974).

[68] SciPy-Community, Scipy optimize minimize (2019). URL https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html#scipy.optimize.minimize

[69] J. Leitão, J. Miotto, M. Gerlach, and E. G. Altmann, Is this scaling nonlinear?, R. Soc. Open Sci. 3, 150649 (2016).
APPENDIX A: SCALING IN A 4D LENNARD-JONES SYSTEM

In this Appendix we provide evidence of the scaling behavior $\lambda(t) \sim t^{1/4}$ for the Lennard-Jones system in $d = 4$ from the analysis of the exponential tails of $P(r; t) = \phi_4 G(r; t)$, where $\phi_4 = 2\pi^2$ (see Eq. (2)). We performed simulations for a 4D-Kob-Andersen binary mixture with the same interaction parameters and units defined in the Methods for $d = 2, 3$. We simulated a system with $N = 10^4$ particles (five times the number of particles used in Ref. [24]), with $N_A = 6500$ and $N_B = 3500$. The linear size of the simulation box was $L = 8.493581$. The total simulation time was $1.6 \cdot 10^5$ and the total number of independent trajectories of $A$ particles analyzed $1.3 \cdot 10^4$. The composition and number density is the same as in Ref. [24] for $d = 4$. We investigated two temperatures, using an integration time step $\delta t^{4D} = 0.02$. The higher $T$ is an intermediate temperature which presents an incipient plateau at intermediate times for the MSD, whereas the lower temperature presents a marked plateau at intermediate times (Fig. 6a)). With the chosen temperatures we enclose a wide range from the beginning of the glassy dynamics to the inner supercooled regime, leaving a more detailed study of the functional behavior of $\beta(T)$ in 4D to future studies.

As in Fig. 2, we fit in Fig. 6b the scaling of $\lambda(t)$ for the 4D-Lennard-Jones system within a time window where $t \geq t_0$ for the two temperatures investigated. This scaling is compatible with $\lambda(t) \sim t^\beta$. The fitted values of $\beta$ were $\beta(T = 0.750) = 0.29 \pm 0.05$ and $\beta(T = 0.675) = 0.28 \pm 0.04$. Both of these values are compatible with the theoretical value $\beta = 1/4$ as predicted by the argument embodied in Eqs. (4) to (6).

Due computer limitations, our statistics for the 4D-Lennard-Jones system resulted in a time window smaller than one decade. For $t \geq 4t_0$ the data become noisy and the exponential tails were in general not accepted by the Kolmogorov-Smirnov test within a sufficiently large $r$-range. To verify the reliability of the results, we performed an additional chi-squared goodness-of-fit test using three families of fundamental increasingly monotonic functions, all of them with two free parameters. First we fitted $\lambda(t)$ within the explored time window using the power law predicted by our scaling argument $\alpha t^\beta$, with $\alpha > 0$ and $\beta > 0$ as free parameters. Second we fitted $\lambda(t)$ using the family $ae^{bt}$, exactly for the same time window and the same number of points (50 in the plot). Finally, we fitted $\lambda(t)$ using the family $\alpha + \beta \log(t)$, again for the same time window and number of points. Thus, for the same number of degrees of freedom, the goodness of the test for the power law was greater than 99% for the two temperatures investigated, whereas the goodness of the exponential family was less than 30% and the logarithmic family less than 1%. This statistical test supports the power law behavior and the value $\beta \cong 1/4$ predicted by our scaling argument.

Figure 6. MSD and Brownian yet non-Gaussian regime for the 4D-Kob-Andersen binary mixture. a) Double log plot of the MSD for the two temperatures investigated, covering the whole time regime from ballistic to diffusive motion. We mark by vertical dashed lines the time windows corresponding to the analyzed Brownian yet non-Gaussian regime for each temperature; b) Double log plot of MSD $1/2$ (squares) and $\lambda$ (circles) as a function of time for the 4D-Kob-Andersen binary mixture at different temperatures: $T = 0.750$ and 0.675. Here $t_0$ is the temperature-dependent time at which the system reaches a $\mu$ value compatible with 1 within numerical uncertainty, i.e. $\text{MSD}(t \geq t_0) \sim t^{\beta \cong 1}$ (see Methods). The analyzed time windows correspond to the temperature regimes signaled by vertical dashed lines in a). Dashed lines representing power law behaviors in a) and b) serve as a reference.

APPENDIX B: FINITE SIZE EFFECTS IN A 2D LENNARD-JONES SYSTEM

In this Appendix we explore whether finite size effects influence our results for the 2D-Lennard-Jones system. We compare the system presented in Methods with an equivalent larger system (same composition, number...
density, and interaction parameters). We choose an intermediate temperature $T = 0.6$ already presented in Fig. 2b. The larger system has a total number of particles $N = 4 \cdot 10^4$, with $N_A = 2.6 \cdot 10^4$ and $N_B = 1.4 \cdot 10^4$, i.e. four times the number of particles in the system described in Methods. The linear size of the simulation box was $L = 185.56246$ (twice the size in the system described in Methods). The total simulation time was $2 \cdot 10^4$ and the total number of independent trajectories of A particles $6 \cdot 10^5$ (the same number of trajectories in the system described in Methods at $T = 0.6$).

In Figure 7a we show a comparison of the MSD for the two systems. We only appreciate a tiny difference in the diffusive regime, where the larger system seems to present a slightly greater diffusion coefficient than the small system. Thus, as shown in Fig. 7b, while the two systems arrive at the diffusive regime almost at the same time, there is a small shift in both the MSD and $\lambda$ within the Brownian yet non-Gaussian regime. In any case, when comparing the dynamics of the two systems within this regime we see that both systems present a common power law behavior compatible with $\beta = 1/2$, as already documented for the small system in Section II. The fitted values were $\beta(large) = 0.45 \pm 0.05$ and $\beta(small) = 0.47 \pm 0.04$. This comparison confirms that finite size effects negligibly affect our main result for 2D-Lennard-Jones systems: the power law dependence $\lambda(t) \sim t^{1/2}$ in the Brownian yet non-Gaussian regime.

**Figure 7.** MSD and Brownian yet non-Gaussian regime for two 2D-Kob-Andersen binary mixtures of different size. a) Double log plot of the MSD at $T = 0.6$ for the large and small system, covering the whole time regime from ballistic to diffusive motion. We mark by vertical dashed lines the time window corresponding to the analyzed Brownian yet non-Gaussian regime; b) Double log plot of MSD$^{1/2}$ (squares) and $\lambda$ (circles) as a function of time for the large and small system at $T = 0.6$. Here $t_0$ is the temperature-dependent time at which the systems reached a $\mu$ value compatible with 1 within numerical uncertainty, i.e. MSD($t \geq t_0$) $\sim t^{\mu-1}$ (see Methods). The analyzed time window corresponds to the time regime signaled by vertical dashed lines in a). Dashed lines representing power law behaviors in a) and b) serve as a reference.

**Figure 8.** Potential energy and pressure for two 2D-Kob-Andersen binary mixtures of different size. a) Probability density of the Potential energy per particle for the large and small system at $T = 0.6$ monitored every 20 time units for a total time $2 \cdot 10^4$; b) Probability density of Pressure for the large and small system at $T = 0.6$ monitored as in a).
We perform an additional comparison for two static observables: potential energy and pressure. The probability densities of the potential energy for the small and large systems are peaked at about the same mean value: $-3.139$ for the large system and $-3.140$ for the small system, see Fig. 8a. The difference between the two distributions is in their standard deviations: $\delta E_l = 0.007$ and $\delta E_s = 0.015$. As expected by the law of large numbers, the standard deviation decreases at increasing the number of particles as $\delta E_s / \delta E_l \approx \sqrt{N_l/N_s} = 2$. In Fig. 8b we show similar results for the pressure: a common mean value of 6.60 for the two systems and a factor two in the standard deviations, $\delta P_l = 0.06$ and $\delta P_s = 0.12$.

APPENDIX C: TESTING THE SCALING ARGUMENT

We test the scaling argument embodied in Eqs. (4) to (6) of the Main Text. We first consider the 3D Lennard-Jones and 3D tetrahedral gel systems at their respective lowest temperatures. We wish in particular to show why the scaling argument fails for the tetrahedral system at low $T$, given as a result a $\beta$ exponent significantly smaller than 1/3 (see Figure 4 in the Main Text).

We first focus on $P(r;t)$ at different times within the diffusive yet non-Gaussian regime ($t \geq t_0$), see Figure 9. The different exponential tails extend up to $r = 0$ (straight lines in Fig. 9) and all intersect approximately at a common value at $r = 0$. Therefore, Eq. (5), which assumes a non-evolving $\rho_0$, is well satisfied for both systems. At short distances, $P(r;t)$ decreases with time for the 3D Lennard-Jones. In contrast, $P(r;t)$ for the tetrahedral gelling system presents two favorable distances (inset in Fig. 9, right) at which the probability increases with time, at least within the time regime investigated. One distance corresponds to the average first neighbour distance ($r \approx 1$) and the other one corresponds to the tetrahedral neighbour distance ($r \approx 1.7$). This result is consistent with the underlying amorphous tetrahedral structure at low $T$ which still manifests itself, despite many particles have already jumped. In other words, some particles have moved to a position at time $t$ which was occupied by another particle at $t = 0$.

Two different, but non-mutually exclusive, mechanisms could explain why Eq. (4) fails for the tetrahedral system. First, the tetrahedral system maintains a underlying structure at times greater than $t_0$. This structure is revealed by the two peaks in the inset of Fig. 9 (right). This means that some particles that have already jumped to a distance corresponding to one of these peaks, could in principle jump back to their original positions, thus abandoning the exponential tail. Such a mechanism would slow down the growth of $N_h(t)$ with time. Second, we also speculate that the tetrahedral system could produce hoppers intermittently in time. In this way, the tetrahedral system would present avalanches in the production of hoppers alternated with periods of scarce production. As a result, the empirical total production of hoppers given by $N_h(t)$ would not be linear before reaching the final Gaussian regime, at which all the particles have already jumped. We leave a detailed test of these hypotheses for future studies.

Finally, we directly test Eq. (6) of the Main Text by plotting $N_h$ versus $\lambda$ in Figures 10c and d. We obtain $N_h(\lambda)$ from $N_h(t)$ at a fixed time $t$ and take the $\lambda$ value corresponding to this time $t$. In this case, the behavior of both systems is compatible with Eq. (6): $N_h(\lambda) \sim \lambda^{d-3}$. In summary, the results of Fig. 10 show: $N_h(t) \sim t$ and $N_h(\lambda) \sim \lambda^3$ for the 3D Lennard-Jones system; and $N_h(t) \sim t^{1/2}$ and $N_h(\lambda) \sim \lambda^3$ for the tetrahedral system. This results in $\lambda \sim t^{2=1/3}$ for the 3D Lennard-Jones system; and $\lambda \sim t^2=1/6$ for the tetrahedral system (see the $\beta$ exponent at lowest temperature for the 3D Lennard-Jones and the tetrahedral systems in Fig. 4 in the Main Text).

For completeness, we test our scaling argument in the 2D Lennard-Jones system at $T = 0.36$. The common intersection of the exponential tails is shown in Fig. 11a, the predicted behavior $N_h(t) \sim t$ is presented in Fig. 11b, and the behavior $N_h(\lambda) \sim \lambda^3$ in Fig. 11c. All these partial results lead to $\lambda \sim t^{1/2}$, as predicted by the scaling argument (Eqs. (4) to (6)).
Figure 9. Left: $P(r;t)$ at different times $t > t_0$ for the 3D-Kob-Andersen binary mixture at the lowest temperature investigated ($T = 0.45$). Right: $P(r;t)$ at different times $t \geq t_0$ for the tetrahedral gelling system at the lowest temperature investigated ($T = 0.1025$). The inset in the right figure shows a detail of $P(r;t)$ at distances corresponding to the first and tetrahedral neighbours. Straight lines for both systems correspond to the exponential tails of the different $P(r;t)$, extended to $r = 0$.

Figure 10. $N_h(t)$ normalized by $N_h(t_0)$ for the 3D-Kob-Andersen binary mixture at $T = 0.45$ (a) and the tetrahedral gelling system at $T = 0.1025$ (b), both obtained by integrating the exponential tails shown in Fig. 9. $N_h(\lambda)$ normalized by $N_h(\lambda t_0)$ for the 3D-Kob-Andersen binary mixture at $T = 0.45$ (c) and the tetrahedral gelling system at $T = 0.1025$ (d). Dashed lines in all the figures are different power laws, shown as references.
Figure 11. Testing the scaling behavior for the 2D-Kob-Andersen binary mixture at $T = 0.36$: a) $P(r; t)$ at different times $t$ ($> t_0$); b) $N_h(t)$ normalized by $N_h(t_0)$, showing the expected linear behavior in time; c) $N_h(\lambda)$ normalized by $N_h(\lambda t_0)$, with the predicted $\lambda^2$ behavior.