Time- and ensemble-average statistical mechanics of the Gaussian Network Model

Alessio Lapolla, Maximilian Vossel and Aljaž Godec
Mathematical bioPhysics group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Göttingen 37077, Germany
E-mail: agodec@mpibpc.mpg.de

Abstract. We present analytical results (up to a numerical diagonalization of a real symmetric matrix) for a set of time- and ensemble-average physical observables in the non-Hookean Gaussian Network Model (GNM) – a generalization of the Rouse model to elastic networks with links with a certain degree of extensional and rotational stiffness. We focus on a set of coarse-grained observables that may be of interest in the analysis of GNM in the context of internal motions in proteins and mechanical frames in contact with a heat bath. A C++ computer code is made available that implements all analytical results.

Submitted to: J. Phys. A: Math. Theor.

1. Introduction

Proteins utilize their unique dynamic character encoded in internal motions to execute a biological function [1]. These motions span fs to s time-scales and their study thus requires a multitude of experimental and/or computational methods [1]. The most detailed, atomically resolved information about these motions comes – with a grain of salt because of an underlying approximate, empirical potential energy function – from Molecular Dynamics (MD) simulations [2, 3]. However, even if the state-of-the-art hardware and highly parallel algorithms allow to reach ms time-scales [4] a substantial time-scale gap remains. In addition, the sheer amount of detail in such tour de force simulations [4] often poses a challenge if one aims at extracting minimal, “leading order” physical principles underlying protein internal motions. Moreover, physical or even topological properties alone may accurately predict selected features of protein dynamics [5, 6].

To describe internal motions in proteins on an effective, coarse-grained level disregarding chemical details Tirion introduced the so-called Elastic Network Model (ENM) [7] akin to the seminal works of Rouse [8] and Flory [9] in polymer physics. The basic idea underlying ENMs is an elastic network connecting those residues, more
The Gaussian Network Model

precisely the respective Ca atoms, that lie within a cutoff distance typically chosen in the range $7 - 16 \text{Å}$. Subsequent works considered various alternative models, e.g. so-called Gaussian Network Model (GNM) [10, 11] and the Anisotropic Network Model (ANM) [12, 13].

Up to date elastic network models in various forms have have been successfully applied (and extended) to refine NMR- [14] and X-ray crystallography-derived protein structures [15], derive NMR-structural order parameters [16], investigate structural correlations [17], function [18, 19, 20, 21], conformational transitions [22], and allosteric effects [23] in proteins, and to identify and decompose protein domains [24]. Further applications involve improving Molecular Dynamics simulations [25], the study of protein evolution [26], investigations of smart polymers [27, 28], viruses [29], membrane channels [30, 31], and nucleic acids [32], as well as the prediction of rupture points in single-molecule pulling experiments [33].

Most of these works rely on “standard” Normal Mode Analysis (NMA) [34, 35], i.e. on spectral characteristics of the underlying mechanical vibration spectrum. In the particular context of proteins NMA has been used predominantly to identify the large-scale collective motions encoded in the eigenvector corresponding to the principal eigenvalue of the Hessian. Notably, the low-frequency modes are quite insensitive to the precise value of the cutoff distance [36].

Here we go beyond and present analytical results for time- and ensemble-average characteristics of internal “reaction coordinates” in GNM in contact with a heat bath at a finite temperature. More precisely, we consider the non-Markovian dynamics of internal distances at equilibrium. Our results may be relevant for interpreting single-molecule spectroscopy data or Molecular Dynamics simulations.

2. The Gaussian Network Model

The Rouse model [8] is one of the earliest “elastic network” models of flexible linear polymers (later on extended to more general network structures [9]). It neglects excluded volume effects and hydrodynamic interactions. Within this theoretical framework beads are connected by ideal, Hookean springs with vanishing resting length (i.e. at $T = 0$ the beads’ positions would coincide). The strength of the springs is proportional to the temperature $T$ of the heat bath. The model does not accurately capture the features of molecules with a non-negligibly internal rigidity.

ENMs [7] extend these core ideas by including a non-zero resting length, i.e. at $T = 0$ the residues are assumed to have distinct positions that are fixed in space. This idea is consistent with the results of NMR and X-ray crystallography that yield a set of positions $\mathbf{R}^0 = \{\mathbf{r}_i^0\}$ of the $N + 1$ residues to which we refer as “the structure” of a protein (NMR experiments in fact yield an ensemble of such structures).

In GNM a pair of residues $i, j$ within a cutoff distance (i.e. $|\mathbf{r}_i^0 - \mathbf{r}_j^0| \leq r_c$) are assumed to be connected by identical (for sake of simplicity) but non-Hookean springs with a constant $K$. The interaction energy as a function of the particles’ positions
The Gaussian Network Model

\( \{r_i\} \) is written as

\[
U_{\text{GNM}}(\{r_{ij}\}) = \frac{K}{2} \sum_{(i,j)} (r_{ij} - r_{ij}^0)^T (r_{ij} - r_{ij}^0),
\]

(1)

where the sum spans all connected pairs. We now introduce for convenience the deviation from the (equilibrium) “structure”, \( \Delta \mathbf{R} = \{\Delta r_i \equiv r_i - r_i^0\} \). The main simplifying hypothesis of the GNM is that \( \Delta \mathbf{R} \) at temperature \( T \) corresponds to an isotropic Gaussian random super-vector, i.e.

\[
P(\Delta \mathbf{R}) = \left[(2\pi)^N \tilde{K} \det \Gamma^{-1}\right]^{-3/2} \exp \left( -\frac{\tilde{K}}{2} \Delta \mathbf{R}^T \Gamma \Delta \mathbf{R} \right),
\]

(2)

where \( \tilde{K} \equiv K/k_B T \) is the dimensionless strength (in units of thermal energy \( k_B T \)) and \( \Gamma \) is a \( 3(N+1) \times 3(N+1) \) block matrix in which each diagonal block is the connectivity (or Kirchoff) matrix \( \Gamma \) with elements

\[
\Gamma_{ij} = \begin{cases} 
-1, & \text{if } i \neq j \text{ and } |r_i^0 - r_j^0| \leq r_c \\
0, & \text{if } i \neq j \text{ and } |r_i^0 - r_j^0| > r_c \\
-\sum_{j,j \neq i} \Gamma_{ij}, & \text{if } i=j.
\end{cases}
\]

(3)

The dynamics of the beads’ positions (i.e. deviations from the equilibrium “structure”) is assumed to follow the Itô equation

\[
d\Delta \mathbf{R}(t) = -\xi K \Gamma \Delta \mathbf{R}(t) dt + \sqrt{2D} d\mathbf{W}(t),
\]

(4)

where \( D \) is the diffusion coefficient and \( \xi \equiv D/k_B T \) the mobility both assumed to be equal for all beads, and \( d\mathbf{W}(t) \) is the increment of the multi-dimensional Wiener process (i.e. Gaussian white noise) with zero mean and covariance \( \langle dW_i(t) dW_j(t') \rangle = \delta_{ij} \delta(t-t') \).

A discussion of ANM would require a different matrix \( \Gamma \) to take into account for anisotropic interactions between beads. Moreover, the potential energy \( U_{\text{ENM}} \) would depend only on distances between the residues. We do not treat this model here.

Henceforth we measure energy in units of thermal energy \( k_B T \) (i.e. \( U \to U/k_B T \)), distances in units of the cutoff distance \( r_c \) (i.e. \( \Delta R_i \to \Delta R_i / r_c \)) and time in units of the diffusion time, \( t_D \equiv r_c^2 / D \)—the time required for a bead with a diffusion coefficient \( D \) to diffuse a distance \( r_c \) (i.e. \( t \to t / t_D \)).

It is convenient to pass to normal super-coordinates \( \mathbf{Q} = \{q_k\} \) that diagonalize \( \Gamma \), i.e. \( Q^T \Gamma Q = \text{diag}(\mu) \) with \( (Q)_{ij} \equiv Q_{ij} \mathbb{1} \), \( \mathbb{1} \) being the \( 3 \times 3 \) identity matrix and where the matrix \( Q \) diagonalizes the Kirchoff matrix, i.e. \( Q^T \Gamma Q = \text{diag}(\mu_i) \), and therefore \( \text{diag}(\mu)_{ii} = \mu_i \mathbb{1} \). For convenience we let \( k \in \{0, \cdots, N\} \) with \( \mu_0 = 0 \) and \( Q_{i,0} \) referring to the center of mass motion, while \( i \in \{1, \cdots, N+1\} \) such that

\[
\Delta r_i = \sum_{k=0}^{N} Q_{ik} q_k, \forall i \{1, \cdots, N+1\}.
\]

(5)

In this notation the Itô equation corresponds to the Fokker-Planck equation describing \( N \) independent isotropic three-dimensional Ornstein–Uhlenbeck processes. Neglecting
The Gaussian Network Model

the center of mass motion we obtain the following equation for the Green’s function (i.e. transition probability density function)

$$\partial_t G(Q, t|Q_0) = \sum_{k=1}^{N} \left( \frac{\partial^2 q_k}{\partial q_k} + \mu_k \partial q_k q_k \right) G(Q, t|Q_0),$$

with localized initial condition $G(Q, t = 0|Q_0) = \delta(Q - Q_0)$ and natural boundary conditions $\lim_{|Q| \to \infty} G(Q, t|Q_0) = 0$. We solve Eq. (6) by means of an eigendecomposition yielding

$$G(Q, t|Q_0) = \sum_{N} \Psi^L_N(Q) \Psi^R_N(Q_0) e^{-\Lambda_N t},$$

where $\Lambda_N$ denote eigenvalues, $N$ being a multiset of integer-triples $\{n_1, \cdots, n_N\}$ with $n_i = \{n_{ix}, n_{iy}, n_{iz}\}$ such that

$$\Lambda_N = \sum_{i=1}^{N} (n_{ix} + n_{iy} + n_{iz}) \mu_i,$$

and $\Psi^L_N(Q)$ and $\Psi^R_N(Q)$ are the corresponding left and right eigenfunction given by

$$\Psi^L_N(Q) = \prod_{i=1}^{N} \psi_{n_i}(q_i), \quad \Psi^R_N(Q) = P_{eq}(Q) \prod_{i=1}^{N} \psi_{n_i}(q_i)$$

where $P_{eq}(Q) \prod_{i=1}^{N} (\mu_i/2\pi)^{3/2} e^{-\mu_i q_i^2/2}$ is the equilibrium probability density function of normal coordinates and

$$\psi_{n_i}(q_i) = H_{n_{ix}}(\mu_i q_{ix}^2/2) H_{n_{iy}}(\mu_i q_{iy}^2/2) H_{n_{iz}}(\mu_i q_{iz}^2/2) \sqrt{2^{n_{ix}+n_{iy}+n_{iz}} n_{ix}! n_{iy}! n_{iz}!},$$

where $q^x, q^y$ and $q^z$ are the components of the vector $q$, and $H_n(x)$ denotes the nth "physicist’s” Hermite polynomial. Using Mehler’s formula

$$\sum_{n=0}^{\infty} \frac{(y/2)^n}{n!} H_n(x) H_n(z) = \frac{1}{\sqrt{1 - y^2}} \exp \left( \frac{y^2 [x - z]^2}{1 - y^2} \right)$$

and recalling that $\mu_0 = 0$ we can also write Eq. (7) in a closed form

$$G(Q, t|Q_0) = \prod_{i=1}^{N} \left( \frac{\mu_i}{2 \pi (1 - e^{-2\mu_i t})} \right)^{3/2} \exp \left( -\frac{\mu_i (q_i - q_{i0} e^{-\mu_i t})^2}{2 (1 - e^{-2\mu_i t})} \right),$$

where the equilibrium probability density function corresponds

$$P_{eq}(Q) \equiv \lim_{t \to \infty} G(Q, t|Q_0)$$

In what follows we will use both forms of the Green’s function, i.e. Eq. (7) and Eq. (12).
3. Conformational dynamics

Throughout we are interested in conformational motions encoded in the dynamics of some internal distance \( d \), e.g. the distance between two beads \( i \) and \( j \), \( l = |r_i - r_j| \) or the distance between the center of masses of two sets of beads \( \Omega_1, \Omega_2 \) with \( \Omega_1 \cap \Omega_2 = \{0\} \),

\[
l_{\Omega_1,\Omega_2} = \left| \sum_{i \in \Omega_1} r_i / \text{card}(\Omega_1) - \sum_{j \in \Omega_2} r_j / \text{card}(\Omega_2) \right|
\]

where \( \text{card}(\Omega_i) \) is the cardinality the set \( \Omega_i \). Without loss of generality we may thus focus on the distance between two arbitrary beads. Note that in absence of any dynamics in an equilibrium at \( T = 0 \) such a distance is constant and equal to \( d_0 \). Expressed in normal coordinates we in turn have

\[
l \equiv r_i - r_j = \sum_{k=1}^{N} (Q_{ik} - Q_{jk})q_k + r_i^0 - r_j^0 \equiv \sum_{k=1}^{N} A_k q_k + d_0, \tag{14}
\]

where in the second equality we have defined \( A_k \) and \( d_0 \) and omitted the labels \( i, j \) to simplify the notation. Note, moreover, that \( l \equiv |l| \) and the generalization to \( l_{\Omega_1,\Omega_2} \) follows by linear superposition.

We will focus on four types of observables. The first one is the (non-Markovian) conditional probability density of the time series of the coordinate, \( l_t \), defined as

\[
G_{d_0}(l, t|l_0) \equiv \mathbb{P}(l_t \in dl|l_{t=0} = l_0) = \frac{\langle \delta(l(Q_t) - l) \delta(l(Q_0) - l_0) \rangle_{Q_t}}{\langle \delta(l(Q_0) - l_0) \rangle_{eq}}, \tag{15}
\]

with \( \lim_{t \to \infty} G_{d_0}(l, t|l_0) \equiv \mathcal{P}_{d_0}^{eq}(l) = \langle \delta(l(Q) - l) \rangle_{eq} \), and where in the second equality we have used the law of conditional probability and introduced the expectation over all Markovian paths of the full system evolving from equilibrium \( \langle \cdot \rangle_{Q_t} \), i.e.

\[
\langle B \rangle_{Q_t} \equiv \int dQ \int dQ_0 B(Q, Q_0) G(Q, t|Q_0) P_{eq}(Q_0). \tag{16}
\]

and the expectation of any observable \( B(Q) \) over the equilibrium measure \( \langle \cdot \rangle_{eq} \) is

\[
\langle B \rangle_{eq} \equiv \int dQ B(Q) P_{eq}(Q). \tag{17}
\]

The second observable is the normalized equilibrium autocorrelation function

\[
C_{d_0}(t) \equiv \frac{\langle l(t)l(0) \rangle - \langle l(t) \rangle \langle l(0) \rangle}{\langle l^2 \rangle_{eq} - \langle l \rangle_{eq}^2} \tag{18}
\]

where we have introduced \( t \) expectations

\[
\langle l(t)l(0) \rangle \equiv \langle l(Q_t)l(Q_0) \rangle_{Q_t} = \int_0^\infty dt \int_0^\infty dl_0 dl_0 G_{d_0}(l, t|l_0) \mathcal{P}_{d_0}^{eq}(l_0)
\]

\[
\langle l(t) \rangle \equiv \frac{\langle l(Q_t) \delta(l(Q_0) - l_0) \rangle_{Q_t}}{\langle \delta(l(Q_0) - l_0) \rangle_{eq}} = \int_0^\infty dl G_{d_0}(l, t|l_0)
\]

\[
\langle l^n \rangle_{eq} \equiv \langle l^n(Q) \rangle_{eq} = \int_0^\infty dl l^n \mathcal{P}_{d_0}^{eq}(l)
\]

The third observable is the \( 3(N + 1) \times 3(N + 1) \) position-covariance matrix [41] whose elements are defined as

\[
C_{\alpha,\beta}(t, t_0) = \langle (r_{i,\alpha}(t + t_0) - r_{i,\alpha}^0)(r_{j,\beta}(t_0) - r_{j,\beta}^0) \rangle_{Q_t}, \tag{19}
\]

where \( r_{i,\alpha} \) is the \( \alpha = \{x, y, z\} \) component of the position vector of bead \( i, r_i \).
The fourth, time-average observable is a functional of the projected path \( l_t \) evolving from \( l_{t=0} \) called the fraction of occupation time or “empirical density” \[ \theta_{d_0}(l; t) \equiv t^{-1} \int_0^t \delta(l_\tau - l) \, d\tau. \] (20)

Note that all observables defined above are assumed to evolve from equilibrium. However, except for \( C_{\alpha\beta}(t, t_0) \), the initial distribution in fact corresponds to equilibrium constrained to a given value of the tagged distance \( l_0 \), i.e., from all those equilibrium configurations drawn from \( P_{eq}(Q) \) that are compatible with \( l_0 \). This introduces memory in the dynamics of \( l_t \). \[ \text{[43]} \]

### 3.1. Projected propagator

The non-Markovian projected propagator \( \mathcal{G}_{d_0}(l, t|l_0) \) defined in Eq. (15) denotes the probability density that the distance between the two tagged beads is equal to \( l \) at time \( t \) given that it was initially equal to \( l_0 \). Introducing the auxiliary functions

\[ \eta_t \equiv \sum_{k=1}^N \frac{A_k^2}{2\mu_k} e^{-\mu_k t}, \quad \Xi_t(d_0, l, l') \equiv \text{erfi} \left( \frac{d_0(\eta_0 - \eta_t) + \eta_t(l + l')}{2\sqrt{\eta_t(\eta_0^2 - \eta_t^2)}} \right) \] (21)

we find (for details of the calculation see \[ \text{[Appendix A]} \])

\[ \mathcal{P}_{d_0}^{eq}(l_0)\mathcal{G}_{d_0}(l, t|l_0) = \frac{u_0 \exp \left( -\frac{(l_0^2 + d_0^2)\eta_0 + (\eta_0 - \eta_t)d_0^2}{4\eta_0(\eta_0 - \eta_t)} \right)}{8\sqrt{\pi \eta_0 d_0(\eta_0 - \eta_t)} \left[ \Xi_t(d_0, -l, -l_0) - \Xi_t(d_0, -l, l_0) + \Xi_t(l_0, l, l_0) + \Xi_t(-d_0, -l, l_0) \right]} \] (22)

where \( \text{erfi}(x) \) is the imaginary error function \[ \text{[38]} \] and

\[ \mathcal{P}_{d_0}^{eq}(l) = \frac{l}{d_0} \frac{e^{-(l^2 + d_0^2)/4\eta_0}}{\sqrt{\pi \eta_0}} \sinh \left( \frac{ld_0}{2\eta_0} \right), \] (23)

We also derive the spectral expansion of \( \mathcal{G}_{d_0}(l, t|l_0) \) that reads (see \[ \text{[Appendix B]} \])

\[ \mathcal{G}_{d_0}(l, t|l_0) = V_{00}(l_0; d_0)^{-1} \sum_N V_{0N}(l; d_0)V_{N0}(l_0; d_0)e^{-\Lambda_N t}, \] (24)

where the overlap elements \( V_{0N} \) and \( V_{N0} \) admit a closed-form expression that is, however, somewhat complicated and thus given in \[ \text{[Appendix B]} \]. Note that “the ground state” element is simple and corresponds to \( V_{00}(l; d_0) = \mathcal{P}_{d_0}^{eq}(l) \).

### 3.2. Equilibrium distance autocorrelation function

The (normalized) autocorrelation function defined in Eq. (17) is made explicit by means of the following results

\[ \langle l \rangle_{eq} = 2\sqrt{\frac{\eta_0}{\pi}} e^{-d_0^2/4\eta_0} + \left( d_0 + \frac{2\eta_0}{d_0} \right) \text{erf} \left( \frac{d_0}{2\sqrt{\eta_0}} \right), \] (25)

\[ \langle l^2 \rangle_{eq} = d_0^2 + 6\eta_0 \] (26)
where erf is the error function. Eqs. (26) follow from direct integration of the last line of Eq. (18) with the aid of Eq. (23). Conversely, an analytic computation of $\langle l(t)l(0) \rangle$ is possible only using the spectral expansion in Eq. (24) and yields,

$$\langle l(t)l(0) \rangle = \sum_{N} \nu_{0N}^{d_0} \nu_{N0}^{d_0} e^{-\Lambda t},$$

(27)

$$\nu_{0N}^{d_0} = \int_{0}^{\infty} dll V_{0N}(l; d_0), \quad \nu_{N0}^{d_0} = \int_{0}^{\infty} dll V_{N0}(l; d_0).$$

(28)

The analytic expression of the coefficients $\nu_{0N}^{d_0}$ is lengthy and can be found in Appendix D. Plugging Eqs. (28) and Eq. (26) into Eq. (17) delivers an exact analytical result for the equilibrium distance autocorrelation function $C_{d_0}(t)$. Alternatively one may also evaluate $C_{d_0}(t)$ by numerical integration of the first line of Eq. (16) using Eq. (22), which may in fact be numerically more convenient than implementing the analytical solution.

3.3. Position covariance matrix

In the analysis of atomistic Molecular Dynamics (MD) simulations one often focuses on the position covariance matrix $C_{ij}^{\alpha\beta}(t, t_0)$ and its eigendecomposition. The trajectory derived from an MD stimulation is then projected on the eigenvector (or principal component) corresponding the largest eigenvalue of the covariance matrix with the aim to identify the most important (potentially functionally relevant) motion in a protein. To facilitate a comparison between the aforementioned analysis of MD simulation with GNM we compute $C_{ij}^{\alpha\beta}(t, t_0)$ analytically. Passing as before to normal coordinates we find

$$C_{ij}^{\alpha\beta}(t, t_0) = \sum_{k=1}^{N} Q_{ik} q_{k\alpha}(t + t_0) \sum_{l=1}^{N} Q_{jl} q_{l\beta}(t_0),$$

(29)

where the matrix elements $Q_{ij}$ do not depend on the spatial coordinate because the GNM is isotropic. Each process $q_{k\alpha}$ corresponds to an independent Ornstein–Uhlenbeck process, i.e. the solution of the Itô integral (setting all constant to unity)

$$q_{k\alpha}(t) = \sqrt{2} \int_{0}^{t} e^{-\mu_k(t-s)} dW_{k\alpha}(s).$$

(30)

Since by construction (i.e. as a result of isotropy) only the elements of the same spatial coordinate for any given normal mode survive the averaging in Eq. (29), the elements of the covariance matrix read explicitly

$$C_{ij}^{\alpha\alpha}(t, t_0) = \sum_{k=1}^{N} \frac{Q_{ik} Q_{jk}}{\mu_k} e^{-\mu_k |t-t_0|}.$$  

(31)

Obviously $C_{ij}^{\alpha\alpha}(t, t_0)$ is stationary (i.e. depends only on the time difference, $C_{ij}^{\alpha\alpha}(t, t_0) = C_{ij}^{\alpha\alpha}(|t-t_0|)$).

The Gaussian Network Model
3.4. Fluctuations of occupation time

Single molecule experiments typically probe time-averaged observables. For example, Förster resonance energy transfer (FRET) [45] and plasmon ruler experiments [46] have been used to extract information about conformational motions of macro-molecules. A fundamental quantity to that underlies this kind of observables is the fraction of occupation time, \( \theta_{d_0}(l; t) \), defined in Eq. (20) [47, 48, 49, 50, 51, 42] – the random fraction of time a time-series (in our case an internal distance between two beads or between two center of masses) of length \( t \) attains a given value of \( l \).

In previous publications we have shown how to obtain the mean and the variance of \( \theta_{d_0}(l; t) \) [51, 42]. Along these lines we here focus on the mean, \( \langle \theta_{d_0}(l; t) \rangle \), and the variance, \( \sigma_{\theta_{d_0}(l; t)}^2 \equiv \langle \theta_{d_0}^2(l; t) \rangle - \langle \theta_{d_0}(l; t) \rangle^2 \), of the occupation time fraction at equilibrium that read, respectively (for a derivation see Appendix F)

\[
\langle \theta_{d_0}(l) \rangle = P_{eq}^{d_0}(l),
\]

\[
\sigma_{\theta_{d_0}(l; t)}^2 = 2 \sum_{N \neq 0} \frac{V_{0N}(d; d_0)V_{N0}(d; d_0)}{\Lambda_N} \left( 1 - \frac{1 - e^{-\Lambda_N t}}{\Lambda_N t} \right).
\]

Note that \( \langle \theta_{d_0}(l; t) \rangle \) corresponds to the equilibrium probability density for all times \( t \) since we are considering an ergodic system evolving from equilibrium initial conditions. The variance of the occupation time fraction can equivalently be obtained from (see e.g. [52])

\[
\sigma_{\theta_{d_0}(l; t)}^2 = 2 t \mathcal{P}_{d_0}^{eq}(l) \left[ \int_0^t (1 - \tau/t) \mathcal{G}_{d_0}(l, \tau|l) - \mathcal{P}_{d_0}^{eq}(l) \right] d\tau.
\]

The integral in Eq. (34) does not admit an explicit solution. However, it can easily be computed via numerical quadrature. Moreover, it is possible to expand \( \mathcal{G}_{d_0}(l, \tau|l) \) for short times (details are given in Appendix E) yielding the small deviation limit

\[
\mathcal{G}_{d_0}(l, \tau|l) \overset{t \to 0}{\sim} 2 \sqrt{\frac{1}{\pi}} \left( \frac{2}{\sqrt{\kappa t}} + \frac{\sqrt{\kappa t}}{l^2} \right) + \mathcal{O}(t^{3/2})
\]

where we have introduced the shorthand notation \( \kappa = \sum_{k=1}^{N} A_k^2 \). Plugging Eq. (35) into Eq. (34) and performing the integral in turn yields

\[
\sigma_{\theta_{d_0}(l; t)}^2 \overset{t \to 0}{\sim} 2 \mathcal{P}_{d_0}^{eq}(l) \left( \frac{8}{3\sqrt{\kappa \pi t}} + \frac{4}{15l^2} \sqrt{\frac{\kappa t}{\pi}} - \mathcal{P}_{d_0}^{eq}(l) \right)
\]

Since the dynamics of every stable system at equilibrium can be “linearized” for sufficiently small times \( t \) the small deviation asymptotic in Eqs. (36) and (33) is in fact a general result for the (large) fluctuations of \( \theta_{d_0}(l; t) \) at sufficiently short times.

4. Examples

We now apply the result of the previous section to the analysis of a Gaussian Network Model of a protein called adenylate kinase and the analysis of toy-model mechanical frames.
Figure 1. Panels (a) and (b) depict a cartoon and the molecular surface (gray) of the two protein structures, called (a) “the closed” configuration 1AKE and (b) “the open” configuration 4AKE. Panels (c) and (d) show the corresponding connectivity matrices for 1AKE and 4AKE, respectively. The blue and cyan square enclose, respectively, the NMP and LID residues. The cutoff distance used to obtain these matrices was 8 Å.

4.1. Gaussian Network Model of adenylate kinase

Adenylate kinase (ADK) is an enzyme catalyzing the reversible phosphorylation reaction that transforms adenosine monophosphate (AMP) to adenosine triphosphate (ATP). The structure of ADK has been resolved using X-ray crystallography that uncovered two distinct conformations of the protein that are deposited in the Protein Data Bank (PDB ID: 1AKE [53] and PDB ID: 4AKE [54]) and shown in Fig. 1.

ADK consists of 214 residues divided in 3 macro-domains called CORE (residues 1 – 29, 68 – 116, and 160 – 214), LID (residues 118 – 160), and NMP (residues 30 – 67). Distinct studies suggest the function to be coupled to open-closed transitions of both,
Table 1. Distance between the center of masses of the three domains for both structures of ADK. All distances are expressed in units of the cutoff distance $r_c = 8$ Å.

|        | 1AKE | 4AKE |
|--------|------|------|
| CORE-LID | 2.6  | 3.8  |
| CORE-NMP | 2.3  | 2.7  |
| LID-NMP  | 2.6  | 4.5  |

LID and NMP domains with respect to the CORE domain [53, 54]. These transitions have been observed even in absence of nucleotides [53, 56]. However, there is a lively debate in the biophysical community about the precise mechanism and rate-limiting steps in the catalytic function of ADK [57].

Here we analyze the autocorrelation functions of distances between the center of mass of LID, NMP, and CORE using the results described in the previous sections. Note that each GNM describes only a single stable structure and therefore cannot capture transitions between the two structures. Nevertheless, the comparison between the two respective GNMs may highlight some differences of the dynamics around the two distinct stable minima.

We obtain the connectivity matrices (shown in Fig. 1) of the two GNMs using the Prody package [58] with a cutoff distance $r_c = 8$ Å. The static (zero-temperature) distances between the center of masses of the three domains in both structures are given in Table 1.

Fig. 2 shows the equilibrium probability density function $P_{d_0}^{\text{eq}}(l)$ (panels a and b) as well as the autocorrelation function $C_{d_0}(t)$ (panels c and d) for all considered distances of the two GNMs representing the two conformational states of ADK. The structure 1AKE is evidently more compact than 4AKE and its corresponding autocorrelation functions consistently decay faster. Moreover, the CORE-NMP distance autocorrelation function decays faster compared to the other two distances whose autocorrelation functions are almost identical (see Fig. 2d). This difference in relaxation is a result of differences in the respective projection, i.e. whereas the eigenvalues of the underlying generator are identical (see Eq. (28)) the numerical coefficients $V_{d_0 N}^{d_0 N}$ and $V_{d_0 N}^{d_0 N}$ depend strongly on the particular type of projection and thus modify the relaxation rate substantially [59].

The lines in Figs. 2c) and 2d) have been obtained by means of a numerical integration of the first line of Eq. (18) using the Gauss-Kronrod quadrature [60]. Unfortunately the evaluation of the integrand is challenging for very short-times because it is a function sharply peaked along the diagonal of the $l, l_0$-plane. This feature prohibits us to obtain reliably (that is, due to numerical imprecision) the autocorrelation function for very short times.

Next we inspect the covariance matrix in Eq. (31) to identify the dominant, potentially functional important, motions in ADK. In order to reduce the information content while retaining the most essential physics about the extent of local fluctuations and how much the motion of each bead correlated to the motion of other beads we
introduce the covariance-time defined as
\[ \tau_{ij\alpha} = \int_{0}^{\infty} C_{\alpha\alpha}^{ij}(t) dt, \] (37)
which may be interpreted in a manner analogous to the correlation time [61, 62, 63], i.e. as a measure of how much the motion between the beads \( i \) and \( j \) is correlated over time. To measure how much the motion of a single bead is correlated with the rest of the system we consider the total the total covariance-time \( \tau_{i\alpha}^{\text{tot}} \equiv \sum_{j \neq i} |\tau_{ij\alpha}| \). Conversely, the total variance-time is quantified directly by \( \tau_{i\alpha} \). Note that the model is isotropic and thus independent of \( \alpha \). The results are shown in Fig. 3.

Notably, one can immediately observe that those residues that are involved in the large-scale open-closed motion (i.e. residues with a large \( \tau_{i\alpha} \)) also participate in correlated motions denoted by large values \( \tau_{i\alpha}^{\text{tot}} \). For the open structure 4AKE (see Fig. 3a and b) the two ends of the LID and NMP domains move in a particularly correlated fashion. These residues are in fact those that move towards the core region in the functional open-closed motion of the protein [55, 56]. A remnant of this collective motion can also be observed in the CORE-NMP distance in panel (d) of Fig. 2.
Figure 3. (a) and (c) depict $\tau_{ii,j}$ and (b) and (d) $\tau_{i,j}^{\text{tot}}$ for each bead in the 4AKE and 1AKE structures, respectively. Notably, the beads in the LID and NMP domains in the 4AKE structure display a particularly large covariance-time. Motion can also be seen in the closed structure 1AKE (Fig. 3 c and d), where the same beads as in 4AKE have a larger $\tau_{i,j}^{\text{tot}}$. This is likely a result of a higher local connectivity.

4.2. Simple mechanical frames

Although GNMs were originally developed to describe proteins they can in fact be used to model any mechanical system in which some underlying network of links imposes constraints on the position of nodes while allowing small, Gaussian fluctuations driven by thermal noise. Examples may include nano-machines such as piezoelectric actuators that move probe-tips in atomic force microscopes [64, 65].

In the generic context of “mechanical frames” the theory of structural rigidity deals with the question of whether frames are rigid or not [66]. A frame is said to be rigid if one cannot change the distance between pairs of nodes without simultaneously altering the
length of at least one connection. A structure that is not rigid is in turn said to allow for inextensional mechanisms. These arise due to a too low number or a particular arrangement of links. In addition, in frames with redundant links there exist states of self-stress. Under given circumstances these states of self-stress impart stiffness to inextensional mechanisms [67].

As anticipated by Maxwell such a classification of mechanical frames is often non-trivial and may require more information than encoded in the topology of the network [68]. A complete analysis of the mechanisms of a given frame can be obtained by a “singular value decomposition“ of the respective Equilibrium Matrix $A$ [69] that relates forces $f$ on the nodes with tensions $t$ in the links

$$\mathbf{At} = \mathbf{f}. \quad (38)$$

Singular value decomposition of $A$ allows (amongst other things) to determine the rank $r$ of $A$ and thereby the number of inextensional mechanisms $m$ and states of self-stress $s$ via $s = b - r$ and $m = 3j - 6 - r$, where $j$ is the number of joints and $b$ the number of links in the structure, and note that there are in general 6 rigid-body motions in 3 spatial dimensions. Maxwell’s well-known formula $b = 3j - 6$ is then extended to:

$$b - 3j + 6 = s - m. \quad (39)$$

To illustrate the concept we consider two toy-model frames depicted in Fig. 4. Both have $j = 4$ nodes and $s = 0$ states of self-stress. The rigid structure with $b = 6$ links has no inextensional mechanism (i.e. $6 - 12 - 6 = 0 - 0$) while the structure with $b = 5$ links has exactly $m = 1$ mechanism (i.e. $5 - 12 - 6 = 0 - 1$).

To highlight the rôle of rigidity and to investigate the effect of a heat-bath we first analyze the autocorrelation function between the blue beads (see Fig. 4) as a function of the rest-length $d_0$. Notably, in a GNM such distance fluctuations do not depend on the equilibrium structure $\mathbf{R}_0$. Only the equilibrium distance between the tagged beads, $d_0 = |\mathbf{r}_i^0 - \mathbf{r}_j^0|$, is relevant. In turn there is a redundancy – many distinct equilibrium
Figure 5. Distance autocorrelation function $C_{d_0}(t)$ for various values of the rest length $d_0$ for the rigid (top panel) and non-rigid (bottom panel) frames depicted in Fig. 4. The black dots depict $C_{d_0}(t)$ in the Rouse limit $d_0 = 0$ (see Appendix D for details). The vertical dashed lines correspond to the time $t_c$ at which $C_{d_0}(t_c) = e^{-1}$. Note that the unstable structure relaxes slower.

structures $R^0$ may yield the same result that depends only on the connectivity matrix $\Gamma$ and $d_0$.

The (normalized) distance autocorrelation function $C_{d_0}(t)$ (see Eq. (17)) for the two frames is shown Fig. 5. For $d_0 \lesssim 0.5$ (in dimensionless units) $C_{d_0}(t)$ depends only very weakly on $d_0$. For larger values of $d_0$ the relaxation time (see dashed vertical lines in Fig. 5) increases. This observation may be explained by noticing that entropy dominates the motion for small $d_0$. That is, in the limit of small $d_0$ the rest length may be neglected and the “Rouse limit” suffices to explain the dynamics essentially quantitatively. Conversely, as $d_0$ increases a certain “stiffness” emerges in the frame and the (random) oscillations become localized around the equilibrium value $d_0$. Note that the entropic contribution to $C_{d_0}(t)$ is more important for the non-rigid frame (see right panel in Fig. 5) as we increase the value of $d_0$ (see Fig. 5b)). Conversely, the departure from the Rouse limit towards the “large stiffness” case is faster in the stable frame (see Fig. 5a)). A larger $d_0$ leads to a slower decay of the autocorrelation function $C_{d_0}(t)$.

Next we consider the fraction of occupation time $\theta_{d_0}(l; t)$ [42]. We assume that the
Figure 6. Panels a-c show the equilibrium probability density $P_{eq}^{d_0}(l)$ for the stable (full lines) and unstable (dashed lines) structure for several values of $d_0$. Panels e-f depict the variance of the occupation time $\sigma_{\theta; d_0}^2(l, t)$ for the stable (full lines) and unstable (dashed lines) structure, respectively, for different values of $d_0$. The length of the trajectory $t$ increases from d to f.

initial condition evolves from equilibrium and therefore $\langle \theta_{d_0}(l; t) \rangle = P_{eq}^{d_0}(l)$ whereas $\sigma_{\theta; d_0}^2(l, t)$ depends on time (see Eq. (33) as well as \cite{51, 42}). The aforementioned dominance of the entropic (heat bath) contribution at small values of $d_0$ is also noticeable the the fluctuations of $\theta_{d_0}(l; t)$ as depicted in Fig. 6. Notably, as $d_0$ increases the support of $\sigma_{\theta; d_0}^2(l, t)$ progressively shifts towards larger $l$ and concentrates near $d_0$.

Notably, the variance of the occupation time fraction $\sigma_{\theta; d_0}^2(l, t)$ changes shape from unimodal shape at short times $t$ to bimodal at long $t$. Such a behavior is characteristic for stochastic process in spatial confinement \cite{42}, i.e. fluctuations of $\theta_{d_0}(l; t)$ are larger in the vicinity of confining boundaries (even if these boundaries are “soft”).

Moreover as $d_0$ increases the shape of both, $P_{eq}^{d_0}(l)$ as well as $\sigma_{\theta; d_0}^2(l, t)$ becomes more symmetric. The reason seems to be that the effect of the confining boundary at $l = 0$ becomes irrelevant as the support of $\sigma_{\theta; d_0}^2(l, t)$ begins to concentrate near a substantial $d_0$. In other words although the projection of the dynamics of a link in 3-dimensional space onto a (1-dimensional) distance destroys the Gaussian behaviour, the latter becomes (partially) restored at large values of $d_0$. 
5. Conclusions

We presented analytical results (up to a numerical diagonalization of a symmetric matrix) for a selection of relevant time- and ensemble-average physical observables in the Gaussian Network Model (GNM). One may think of GNM as a certain generalization of the Rouse model to networks with links with a certain degree of extensional and rotational stiffness. We determined a set of coarse-grained observables – internal distances – that may be of interest in the analysis of GNM in the context of internal motions in proteins or mechanical frames in contact with a heat bath. We hope that our results will enable and motivate a more systematic analysis of GNM derived from proteins [58]. To this end a C++ computer code is provided in the Supplementary material that implements all results (for more details about the implementation see Appendix G).

Acknowledgments

The authors thank David Hartich and Lars Bock for the useful discussions. The financial support from the German Research Foundation (DFG) through the Emmy Noether Program GO 2762/1-1 to AG is gratefully acknowledged.

Appendix A. Derivation of the equilibrium probability density

The equilibrium probability density function of any link-vector \( \mathbf{l} \) is defined by

\[
P^{eq}_{d_0}(\mathbf{l}) = V_{00}(\mathbf{l}; d_0) \equiv \int dQ \Psi^R_0(Q) \delta(\sum_{k=1}^N A_k q_k + d_0 - 1) \Psi^L_0(Q)
\]

Applying the Fourier transform \( \tilde{f}(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx f(x) e^{-i sx} \) component-wise to Eq. (A.1) (i.e. \( \mathbf{l} \to s \)) yields

\[
\frac{1}{(2\pi)^3} \int dQ \prod_{k=1}^N \left( \frac{\mu_k}{2\pi} \right)^{3/2} \exp \left( -\sum_{k=1}^N \frac{\mu_k}{2} q_k^2 + i(A_k q_k + d_0)s \right) =
\]

\[
\frac{1}{(2\pi)^3} e^{-s^2 \sum_{k=1}^N A_k^2 / 2\mu_k + id_0 s}.
\]

Inverting the Fourier transform we obtain, defining \( \eta_0 = \sum_{k=1}^N A_k^2 / (2\mu_k) \),

\[
P^{eq}_{d_0}(\mathbf{l}) = V_{00}(\mathbf{l}; d_0) = \frac{1}{(2\pi)^3} \left( \frac{\pi}{\eta_0} \right)^{3/2} e^{-(1-d_0)^2 / 4\eta_0}.
\]

Since we are only interested in the distance and not the direction we need to marginalize over angles, i.e.

\[
\int_{0}^{\infty} dx x^2 \int_{0}^{2\pi} d\phi \int_{-1}^{1} d(cos \theta) V_{00}(x) \delta(|x|-x),
\]
where \( \phi \) is the polar angle, \( \theta \) is the azimuthal angle and without loss of generality we choose a frame of reference such that the vector \( \mathbf{d}_0 \) is parallel to the z-axis. The solution of this integral finally gives Eq (23):

\[
\mathcal{P}_{d_0}^{eq}(l) = V_{00}(l; d_0) = \frac{1}{\sqrt{\pi \eta_0}} \frac{l}{d_0} e^{-(l^2 + d_0^2)/4\eta_0} \sinh \left( \frac{ld_0}{2\eta_0} \right).
\] (A.6)

**Appendix B. Spectral solution for \( G_{d_0} \)**

In the spectral solution for the Green’s function in Eq. (24) we have defined the elements \( V_{0N}(l; d_0), V_{N0}(l; d_0) \) which are derived as follows. Let

\[
V_{0N}(l; d_0) = \int dQ \Psi_N^R(Q) \delta \left( \sum_{k=1}^{N} A_k q_k + d_0 - 1 \right) \Psi_0^L(Q),
\] (B.1)

\[
V_{N0}(l; d_0) = \int dQ \Psi_N^R(Q) \delta \left( \sum_{k=1}^{N} A_k q_k + d_0 - 1 \right) \Psi_N^L(Q).
\] (B.2)

Fortunately, the above elements \( V_{0N} \) and \( V_{N0} \) are identical (cf. Eq (9)). Therefore what we need to solve for is

\[
V_{N0}(l; d_0) = \prod_{k=1}^{N} \int dq_k \ (\frac{\mu_k}{2\pi})^\frac{3}{2} \sqrt{\frac{1}{2^{n_kx+n_ky+n_kz} n_{kx}! n_{ky}! n_{kz}!}} \times
\]

\[
H_{n_{kx}} \left( \sqrt{\frac{\mu_k}{2} q_k^x} \right) H_{n_{ky}} \left( \sqrt{\frac{\mu_k}{2} q_k^y} \right) H_{n_{kz}} \left( \sqrt{\frac{\mu_k}{2} q_k^z} \right) \times
\]

\[
e^{-\mu_k q_k^0 / 2} \delta \left( \sum_{k=1}^{N} A_k q_k + d_0 - 1 \right). \] (B.3)

It is convenient to define the auxiliary variables \( \left\{ q_k' \right\} \equiv \left\{ q_k^x - d_0^x, q_k^y - d_0^y, q_k^z - d_0^z \right\} \), and then perform the Fourier transform \( l \rightarrow s \) to obtain

\[
\frac{1}{(2\pi)^3} \prod_{k=1}^{N} \int dq_k' \ (\frac{\mu_k}{2\pi})^\frac{3}{2} \sqrt{\frac{1}{2^{n_kx+n_ky+n_kz} n_{kx}! n_{ky}! n_{kz}!}} \times
\]

\[
H_{n_{kx}} \left( \sqrt{\frac{\mu_k}{2} q_k'^x} \right) H_{n_{ky}} \left( \sqrt{\frac{\mu_k}{2} q_k'^y} \right) H_{n_{kz}} \left( \sqrt{\frac{\mu_k}{2} q_k'^z} \right) e^{-\mu_k q_k'^0 / 2 - i A_k s \cdot q_k'}. \] (B.4)

Factorizing in the three spatial dimensions, completing the square in the exponential, and changing the variable to \( t_k^h = \sqrt{\mu_k q_k'^h} / \sqrt{2} \) (where the subscript \( h \) denotes the respective spatial coordinate) we find

\[
\frac{1}{(2\pi)^3} \prod_{k=1}^{N} \left( \frac{1}{\pi} \right)^\frac{3}{2} \prod_{h=1}^{3} \sqrt{\frac{1}{2^{n_{kh} h \cdot h} n_{kh}!}} e^{-\frac{1}{2} (A_k)^2 / 2\mu_k} \times
\]

\[
\int_{-\infty}^{\infty} dt_h^k H_{n_{kh}} (t_h^k) \exp \left( - \left[ t_h^k - \left( - \frac{i A_k}{\sqrt{2\mu_k}} s_h \right) \right]^2 \right). \] (B.5)
whose solution is [70]

$$\frac{1}{(2\pi)^3} \prod_{k=1}^{N} \prod_{h=1}^{3} \sqrt{\frac{2^{n_{kh}}}{n_{kh}!}} (-\frac{iA_k}{\sqrt{2\mu_k}})^{n_{kh}} e^{-\frac{s_h^2}{2\mu_k}}.$$  \hfill (B.6)

It turns out to be convenient to write Eq. (B.6) as

$$\frac{1}{(2\pi)^3} \left[ \prod_{k=1}^{N} \prod_{h=1}^{3} \sqrt{\frac{2^{n_{kh}}}{n_{kh}!}} (-iA_k)^{n_{kh}} \right] \sum_{k=1}^{N} \sum_{h=1}^{3} \frac{n_{kh}^{2}}{2\mu_k} e^{-s_h^2 \sum_{k=1}^{N} (A_k)^2 / 2\mu_k},$$  \hfill (B.7)

and to define

$$M = \prod_{k=1}^{N} \prod_{h=1}^{3} \sqrt{\frac{2^{n_{kh}}}{n_{kh}!}} (-iA_k)^{n_{kh}} = \frac{\sqrt{\prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh}} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh}}.$$  \hfill (B.8)

We now invert the Fourier transform

$$M \frac{(2\pi)^3}{1 \prod_{h=1}^{3} \int_{-\infty}^{\infty} ds_h^2 s_h \sum_{k=1}^{N} n_{kh} e^{-\eta_0 s_h^2 / 2 \eta_0}}.$$  \hfill (B.9)

Completing the square in the exponential and defining $t_h = \sqrt{\eta_0} s_h$, we can write:

$$M \frac{(2\pi)^3}{1 \prod_{h=1}^{3} \int_{-\infty}^{\infty} dt_h \sum_{k=1}^{N} n_{kh} e^{-(t_h - i2\sqrt{\eta_0})^2}},$$  \hfill (B.10)

the integral in the previous equation can be solved analytically [70]

$$M \frac{(2\pi)^3}{1 \prod_{h=1}^{3} \int_{-\infty}^{\infty} dt_h \sum_{k=1}^{N} n_{kh} e^{-(t_h - i2\sqrt{\eta_0})^2}},$$  \hfill (B.11)

Using the definition of $M$ in Eq. (B.8), defining $N_h = \sum_{k=1}^{N} n_{kh}$ and $N = N_x + N_y + N_z$, and going back to the original, non-shifted coordinates we arrive at the following form of Eq. (B.3)

$$V_{ON}(l; d_0) = \frac{1}{(2\sqrt{\pi})^3} \sqrt{\prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh} \prod_{k=1}^{N} \sum_{h=1}^{3} n_{kh}} \times$$

$$\frac{1}{\sqrt{\eta_0}^{N+3}} H_{N_x}(l^x - d_0^x / 2\sqrt{\eta_0}) H_{N_y}(l^y - d_0^y / 2\sqrt{\eta_0}) H_{N_z}(l^z - d_0^z / 2\sqrt{\eta_0}) e^{-(1-d_0)^2 / 4\eta_0}.$$  \hfill (B.12)

To integrate over the angular part it is convenient to use the following expansion of the Hermite polynomials [38]:

$$H_n(x) = \sum_{m=0}^{n} \frac{(-1)^m}{m!(n-2m)!} (2x)^{n-2m}.$$  \hfill (B.13)
The Gaussian Network Model

If we rotate our frame of reference such that \( \hat{z} \parallel d_0 \), (i.e.: \( d_0^x = 0 \), \( d_0^y = 0 \) and \( d_0^z = d_0 \)) we find

\[
V_{ON}(l; d_0) = \frac{1}{(2\sqrt{\pi})^3} \sqrt{\frac{1}{2^N \prod_{k=1}^N n_{kx}!n_{ky}!n_{kz}!}} \prod_{k=1}^N \left( \frac{A_k}{\sqrt{2\mu_k}} \right)^{n_{kx}+n_{ky}+n_{kz}} \times \frac{1}{\sqrt{\eta_0}^{N+3}} \frac{N_x!N_y!N_z!e^{-\frac{d^2}{4\eta_0}}}{\sqrt{\eta_0}} \times \frac{1}{\sqrt{\eta_0}^{N_x/2}} \frac{1}{\sqrt{\eta_0}^{N_y/2}} \frac{1}{\sqrt{\eta_0}^{N_z/2}} \sum_{a=0}^{N_x} \sum_{b=0}^{N_y} \sum_{c=0}^{N_z} a!b!c!(N_x - 2a)!(N_y - 2b)!(N_z - 2c)! \times \left( \frac{l_x}{\sqrt{\eta_0}} \right)^{N_x-2a} \left( \frac{l_y}{\sqrt{\eta_0}} \right)^{N_y-2b} \left( \frac{l_z - d_0}{\sqrt{\eta_0}} \right)^{N_z-2c} e^{l_z - d_0/2\eta_0}. \tag{B.14}\]

Using the binomial expansion of \( (l_z - d_0)/\sqrt{\eta_0} \) we obtain

\[
V_{ON}(l; d_0) = \frac{1}{(2\sqrt{\pi})^3} \sqrt{\frac{1}{2^N \prod_{k=1}^N n_{kx}!n_{ky}!n_{kz}!}} \prod_{k=1}^N \left( \frac{A_k}{\sqrt{2\mu_k}} \right)^{n_{kx}+n_{ky}+n_{kz}} \times \frac{1}{\sqrt{\eta_0}^{N+3}} \frac{N_x!N_y!N_z!}{\sqrt{\eta_0}} e^{-(l^2 + d_0^2)/4\eta_0} \times \frac{1}{\sqrt{\eta_0}^{N_x/2}} \frac{1}{\sqrt{\eta_0}^{N_y/2}} \frac{1}{\sqrt{\eta_0}^{N_z/2}} \sum_{a=0}^{N_x-2c} \sum_{b=0}^{N_y-2b} \sum_{c=0}^{N_z-2c} a!b!c!(N_x - 2a)!(N_y - 2b)!(N_z - 2c)! \times \left( \frac{l_x}{\sqrt{\eta_0}} \right)^{N_x-2a} \left( \frac{l_y}{\sqrt{\eta_0}} \right)^{N_y-2b} \left( \frac{l_z - d_0}{\sqrt{\eta_0}} \right)^{N_z-2c} m! \cdot \eta_0^{N_x-2a} \times \tag{B.15}\]

At this point we can integrate over the angles of \( l \) fixing the length, hence

\[
V_{ON}(l, d_0) = \frac{1}{(2\sqrt{\pi})^3} \sqrt{\frac{1}{2^N \prod_{k=1}^N n_{kx}!n_{ky}!n_{kz}!}} \prod_{k=1}^N \left( \frac{A_k}{2\mu_k} \right)^{n_{kx}+n_{ky}+n_{kz}} \times \frac{1}{\sqrt{\eta_0}^{N+3}} \frac{N_x!N_y!N_z!}{\sqrt{\eta_0}} e^{-(l^2 + d_0^2)/4\eta_0^2} \times \frac{1}{\sqrt{\eta_0}^{N_x/2}} \frac{1}{\sqrt{\eta_0}^{N_y/2}} \frac{1}{\sqrt{\eta_0}^{N_z/2}} \sum_{a=0}^{N_x-2c} \sum_{b=0}^{N_y-2b} \sum_{c=0}^{N_z-2c} a!b!c!(N_x - 2a)!(N_y - 2b)!(N_z - 2c)! \times \left( \frac{l_x}{\eta_0} \right)^{N_x-2a} \left( \frac{l_y}{\eta_0} \right)^{N_y-2b} \left( \frac{l_z - d_0}{\eta_0} \right)^{N_z-2c} \times \int_0^{2\pi} d\phi \sin^\phi \left( \cos^\phi \right)^{N_x-2a} \times \int_0^{2\pi} \frac{1}{m!} \left( \frac{l}{\eta_0} \right)^{N_x-2c-m} \left( \frac{1}{\eta_0} \right)^{N_z-2c-2m} \times \int_0^\pi d\theta \left( \sin \theta \right)^{N_x-2a+1} \left( \cos \theta \right)^{N_z-2c-m} e^{2\cos \theta d_0/2\eta_0^2}. \tag{B.16}\]

The first integral is

\[
\int_0^{2\pi} d\phi \cos^n \phi \sin^m \phi = \frac{\pi n! m!}{2^n m! (\frac{n}{2}) (\frac{m}{2}) (\frac{n+m}{2})!}. \tag{B.17}\]
and is non-zero only if $n$ and $m$ are even \(\left[\right]\). Therefore $N_x$ and $N_z$ must be even. While the second integral reads \(\left[\right]\)
\[
\int_0^\pi d\theta (\sin \theta)^n (\cos \theta)^m e^{k\cos \theta} =
\]
\[
\frac{\sqrt{\pi}}{4} \gamma \left( \frac{1 + n}{2} \right) \left[ 2(1 + (-1)^m) \gamma \left( \frac{1 + m}{2} \right) _1\tilde{F}_2 \left( \frac{1 + m}{2}; \frac{1}{2}, \frac{2 + m + n}{2}; \frac{k^2}{4} \right) \right. 
- \left. (1 + (-1)^m) k \gamma \left( \frac{1 + m}{2} \right) _1\tilde{F}_2 \left( \frac{2 + m}{2}; \frac{3 + m + n}{2}; \frac{k^2}{4} \right) \right],
\] (B.18)
where we have introduced the Euler’s gamma function $\gamma(x)$ as well as the regularized hypergeometric function $_p\tilde{F}_q(a_1, \cdots, a_p; b_1, \cdots, b_q; x)$ \(\left[\right]\). Putting all together we finally arrive at
\[
V_{0N}(l; d_0) = \frac{1}{16} \sqrt{\frac{2}{\pi}} \prod_{k=1}^{N} \frac{1}{n_{kx}!n_{ky}!n_{kz}!} \prod_{k=1}^{N} \left( \frac{A_k}{\sqrt{2\mu_k}} \right)^{n_{kx}+n_{ky}+n_{kz}}
\]
\[
\frac{1}{\sqrt{\hbar_0}}^{N+1} N_x! N_y! N_z! e^{-(d^2+d_0^2)/4\hbar_0} \times
\]
\[
\sum_{a=0}^{N_x/2} \sum_{b=0}^{N_y/2} \sum_{c=0}^{N_z/2} \frac{(-1)^{a+b+c}}{a!b!c!(N_x-2a)!/(N_x-2a)!} \frac{(-1)^{N_x-N_y-2a}}{2^{N_x-N_y-2a}} \times
\]
\[
\sum_{m=0}^{N_x-2c} \frac{1}{m!(N_x-2c-m)!} \left( \frac{l}{\sqrt{\hbar_0}} \right)^{N_x-2(a+b+c)-m+2} \left( -\frac{d_0}{\sqrt{\hbar_0}} \right)^m \times
\]
\[
_1\tilde{F}_2 \left( \frac{1 + N_x - 2c - m}{2}; \frac{1}{2}, \frac{3 + N_x - 2(a + b + c) - m}{2}; \frac{l^2 d_0^2}{16\hbar_0^2} \right) \times
\]
\[
_1\tilde{F}_2 \left( \frac{1 + N_x - 2c - m}{2}; \frac{1}{2}, \frac{3 + N_x - 2(a + b + c) - m}{2}; \frac{l^2 d_0^2}{16\hbar_0^2} \right) \times
\]
\[
- \left[ 2(1 + (-1)^m) k \gamma \left( \frac{1 + m}{2} \right) _1\tilde{F}_2 \left( \frac{2 + m}{2}; \frac{3 + m + n}{2}; \frac{k^2}{4} \right) \right]
\] (B.19)
which finally allows us to write down the non-Markovian Green’s function expressed as an infinite series in Eq. \(\left[\right]\). In addition, the series expansion allows us the compute the cross conditioned Green’s function
\[
G_{d_0,d'_0}(l, t | l', t) = V_{00}(l'; d'_0)^{-1} \sum_N V_{0N}(l; d_0) V_{N0}(l'; d'_0) e^{-\Lambda N t}
\] (B.20)
that is the probability that the distance between the beads $i$ and $j$ is equal to $l$ at time $t$ conditioned to the fact that the distance between the beads $k$ and $l$ at time $0$ was equal to $l'$, assuming that these two distances have rest lengths $d_0$ and $d'_0$, respectively. In particular in order to evaluate $V_{N0}(l'; d'_0)$ we need to consider that the distance $l'$ is expressed via the normal coordinates as
\[
d' = r_k - r_l = \sum_{i=1}^{N} B_i q_i,
\] (B.21)
and we in turn use these coefficients to define \( \zeta_t = \sum_{k=1}^{N} B_k^2/2\mu_k e^{-\mu_k t} \) and \( \zeta_0 = \sum_{k=1}^{N} B_k^2/2\mu_k \) instead of \( \eta_t \) and \( \eta_0 \).

**Appendix C. Closed form solution for \( G_{d_0} \)**

In order to obtain the equivalent result in a closed form solution we should consider the following integral:

\[
\mathcal{J}_{d_0}(l, t; l_1) = \int dQ \int dQ_1 G(Q, t|Q_1) P_{eq}(Q_1) \times \delta \left( \sum_{k=1}^{N} A_k q_{1k} + d_0 - l_1 \right) \delta \left( \sum_{k=1}^{N} A_k q_k + d_0 - l \right).
\]

(C.1)

Performing the first Fourier transform, between \( l_1 \rightarrow u \) the above integral becomes

\[
\int dQ \delta \left( \sum_{k=1}^{N} A_k q_k + d_0 - l \right) e^{-i d_0 \cdot u} \times \\
\left( \frac{1}{2\pi} \right)^3 \int dQ_{1k} \prod_{k=1}^{N} \left( \frac{\mu_k^2}{2\pi} \right)^{3/2} \left( \frac{\mu_k}{2\pi(1 - e^{-2\mu_k t})} \right)^{3/2} \times \\
\exp \left[ -\frac{\mu_k}{2} \left( q_{1k}^2 + q_{1k} e^{2\mu_k t} - 2q_k \cdot q_{1k} e^{-\mu_k t} \right) \right] \times \\
e^{-i A_k q_{1k} \cdot u - \mu_k q_{1k}^2/2},
\]

(C.2)

and the integration yields

\[
\left( \frac{1}{2\pi} \right)^3 \int dQ \delta \left( \sum_{k=1}^{N} A_k q + d_0 - l \right) e^{-i d_0 \cdot u} \times \\
\prod_{k=1}^{N} \left( \frac{\mu_k^2}{2\pi} \right)^{3/2} \exp \left[ -\frac{\mu_k}{2} \left( q_k^2 + 2i \frac{A_k}{\mu_k} e^{-\mu_k t} u \cdot q_k + \frac{A_k^2}{\mu_k^2}(1 - e^{-2\mu_k t}) u^2 \right) \right]
\]

(C.3)

Performing the second Fourier transform \( l \rightarrow v \) we find

\[
e^{-i d_0 \cdot (u + v)} \left( \frac{1}{2\pi} \right)^6 \prod_{k=1}^{N} \int dq_k \left( \frac{\mu_k}{2\pi} \right)^{3/2} \times \\
\exp \left[ -\frac{\mu_k}{2} \left( q_k^2 + 2i \frac{A_k}{\mu_k} (e^{-\mu_k t} u + v) \cdot q_k + \frac{A_k^2}{\mu_k^2}(1 - e^{-2\mu_k t}) u^2 \right) \right],
\]

(C.4)

that reads

\[
e^{-i d_0 \cdot (u + v)} \left( \frac{1}{2\pi} \right)^6 \prod_{k=1}^{N} \exp \left[ -\frac{A_k^2}{2\mu_k} (u^2 + v^2) + 2 \frac{A_k^2}{2\mu_k} e^{-\mu_k t} u \cdot v \right].
\]

(C.5)

It is convenient to define

\[
\sum_{k=1}^{N} \frac{A_k^2}{2\mu_k} e^{-\mu_k t} = \eta_t \rightarrow \sum_{k=1}^{N} \frac{A_k^2}{2\mu_k} = \eta_0
\]

(C.6)
so the Fourier transform of the joint-density is:
\[ \mathcal{J}_d(v, t; u) = \frac{1}{(2\pi)^3} \exp \left( -\eta_0 u^2 - \eta_0 v^2 + 2\eta_0 u \cdot v + i\delta_0 \cdot (u + v) \right). \] (C.7)

The inversion of the two Fourier transforms gives straightforwardly
\[ \mathcal{J}_d(l, l'; l) = \frac{1}{2\pi^3} \left( \frac{1}{\eta_0^2 - \eta_t^2} \right)^{3/2} \times \exp \left[ -\frac{\eta_0 (l - d_0)^2 + \eta_0 (l_1 - d_0)^2 - 2\eta_0 (l - d_0) \cdot (l_1 - d_0)}{4(\eta_0^2 - \eta_t^2)} \right], \] (C.8)

We now marginalize over the angles
\[ \mathcal{J}_d(l, l'; l) \equiv \int dd \int dd_0 \delta(|d_0| - d_0) \delta(|l_1| - l_1) \delta(|l| - l) \mathcal{J}_d(l, l'; l), \] (C.9)

by moving to a frame of reference where \( d_0 \) is parallel to the the \( z \) axis, and express all the vectors in spherical coordinates. This removes all delta-functions and \( d_0 \) in the new frame of reference is just a scalar. By doing so we obtain
\[ \mathcal{J}_d(l, l'; l) = \frac{1}{2\pi^3} \left( \frac{1}{\eta_0^2 - \eta_t^2} \right)^{3/2} \times \int_0^{2\pi} d\phi \int_0^{2\pi} d\phi' \int_{-1}^{1} d(cos \theta) \int_{-1}^{1} d(cos \theta') \times \exp \left[ \frac{(\eta_0 - \eta_1)l_0d_0}{2(\eta_0^2 - \eta_t^2)} \cos \theta + \frac{2(\eta_0 - \eta_1)l_1d_0}{2(\eta_0^2 - \eta_t^2)} \cos \theta' + \frac{\eta_ll_1}{2(\eta_0^2 - \eta_t^2)} \times (\cos \phi \cos \phi' \sin \theta \sin \theta' + \sin \phi \sin \phi' \sin \theta \sin \theta' + \cos \theta \cos \theta') \right]. \] (C.10)

The two integrals over \( \phi \) and \( \phi' \) (keeping in mind that \( \cos(\phi - \phi') = \cos \phi \cos \phi' + \sin \phi \sin \phi' \)) give us
\[ \frac{1}{16\pi} \left( \frac{1}{\eta_0^2 - \eta_t^2} \right)^{3/2} \times \int_{-1}^{1} d(cos \theta) \int_{-1}^{1} d(cos \theta') \exp \left[ \frac{(\eta_0 - \eta_1)l_0d_0}{2(\eta_0^2 - \eta_t^2)} \cos \theta + \frac{(\eta_0 - \eta_1)l_1d_0}{2(\eta_0^2 - \eta_t^2)} \cos \theta' \right. \\
\left. + \frac{\eta_ll_1}{2(\eta_0^2 - \eta_t^2)} \cos \theta \cos \theta' \right] I_0 \left( \frac{\eta_ll_1}{2(\eta_0^2 - \eta_t^2)} \sqrt{1 - \cos^2 \theta} \sqrt{1 - \cos^2 \theta'} \right), \] (C.11)

where \( I_0(x) \) is the modified Bessel function of the first kind. The first integral in \( \cos \theta' \) is solvable [70], and by changing the variable \( \cos \theta \to x \) we are left with
\[ \frac{1}{8\pi} \left( \frac{1}{\eta_0^2 - \eta_t^2} \right)^{3/2} \times \int_{-1}^{1} dxe^{2(\eta_0^2 - \eta_t^2) x} \sinh \left( \sqrt{\frac{(\eta_0 - \eta_1)^2 l_0^2 d_0^2 + \eta_1^2 l_1^2 d_0^2}{4(\eta_0^2 - \eta_t^2)^2}} \right), \] (C.12)
And the final integral yields \[71\]

$$J_{d_0}(l, t; l_1) = \frac{1}{16\sqrt{\pi}} \left( \frac{1}{\eta_0^2 - \eta_l^2} \right)^{3/2} \exp \left( -\frac{\eta_0^2 l^2 + \eta_l^2 l_1^2 + 2(\eta_0 - \eta_l)d_0^2}{4(\eta_0^2 - \eta_l^2)} \right) \times \right.$$ 

$$l_1^2 l_{t1}^2 e^{-ab/c-c/a} \sqrt{ac} \left[ \text{erfi} \left( \frac{2a\sqrt{b-c} - c}{2\sqrt{ac}} \right) - \text{erfi} \left( \frac{2a\sqrt{b-c} + c}{2\sqrt{ac}} \right) \right]$$

\[\text{erfi} \left( \frac{c - 2a\sqrt{b+c}}{2\sqrt{ac}} \right) + \text{erfi} \left( \frac{c + 2a\sqrt{b+c}}{2\sqrt{ac}} \right) \]  \hfill (C.13)

having defined

\[a = \frac{(\eta_0 - \eta_l)d_0}{2(\eta_0^2 - \eta_l^2)}; \] \hfill (C.14)

\[b = \frac{(\eta_0 - \eta_l)^2 l_1^2 d_0^2 + \eta_l^2 l_2 l_1^2}{4(\eta_0^2 - \eta_l^2)^2}; \] \hfill (C.15)

\[c = \frac{\eta_l(\eta_0 - \eta_l)l_2 l_1^2 d_0}{2(\eta_0^2 - \eta_l^2)^2}; \] \hfill (C.16)

the direct substitution of these auxiliary variables gives, upon division by \(\mathcal{P}_{d_0}^\text{eq}\) and some simplification, Eq. \[(22).\]

**Appendix D. Derivation of equilibrium autocorrelation function**

In order to compute the autocorrelation function in Eq. \[(28)\] the following integrals must be evaluated

$$\mathcal{V}_{oN}^{d_0} = \int_0^\infty dx V_{oN}(x, d_0) x, \quad \mathcal{V}_{NO}^{d_0} = \int_0^\infty dx V_{NO}(x, d_0) x. \quad (D.1)$$

These two integrals are identical and the integration yields \[70\]

$$\mathcal{V}_{oN}^{d_0} = \frac{1}{16} \sqrt{\frac{1}{2^N N! \prod_{k=1}^N n_{kx}! n_{ky}! n_{kz}!}} \prod_{k=1}^N \left( \frac{A_k}{\sqrt{2\mu_k}} \right)^{n_{kz} + n_{ky} + n_{kz}} \frac{1}{\sqrt{\eta_0}^{N+1}} N_x! N_y! N_z! e^{-d_0^2/4\eta_0} \times$$

$$\sum_{a=0}^{N_x/2} \sum_{b=0}^{N_y/2} \sum_{c=0}^{N_2/2} \frac{(-1)^{a+b+c}}{a!b!c!((N_z-2a)!((N_y-2b)!2N_x+N_y-2(a+b))} \sum_{l=0}^{N_z-2c} \frac{1}{l!(N_z-2c-l)!} \left( -\frac{d_0}{\sqrt{\eta_0}} \right)^l \times$$

$$\left[ (1 + (-1)^{N_x-2c-l}) \gamma \left( \frac{1 + N_z - 2c - l}{2} \right) 2^{N-2(a+b+c)-l+4} \eta_0 \gamma \left( \frac{N - 2(a + b + c) - l + 4}{2} \right) \frac{1}{N_2} \right]$$

$$2 \tilde{F}_2 \left( \frac{1 + N_z - 2c - l}{2}, \frac{N - 2(a + b + c) - l + 4}{2}; \frac{3}{2}, \frac{N - 2(a + b + c) - l}{2}; \frac{d_0^2}{4\eta_0} \right) -$$

$$(-1 + (-1)^{N_x-2c-l}) d_0 \sqrt{\eta_0} \gamma \left( \frac{1 + N_z - 2c - l}{2} \right) \times$$

$$2^{N-2(a+b+c)-l+3} \gamma \left( \frac{N - 2(a + b + c) - l + 5}{2} \right) \times$$

$$2 \tilde{F}_2 \left( \frac{2 + N_z - 2c - l}{2}, \frac{N - 2(a + b + c) - l + 5}{2}; \frac{3}{2}, \frac{4 + N - 2(a + b + c) - l}{2}; \frac{d_0^2}{4\eta_0} \right) \right] \quad (D.2)$$
If we are instead interested in the cross-correlation the more general Eq (B.20) must be used and the two integrals differ in terms of some constants, i.e. they are obtained by changing the following variables \( d_0 \rightarrow d'_0 \), \( \{A_k\} \rightarrow \{B_k\} \) and \( \eta_t \rightarrow \zeta_t \).

**Appendix D.1. Rouse-limit autocorrelation function**

In Fig. 5 we showed how the autocorrelation for a GNM compares to the autocorrelation in the Rouse limit (i.e. \( d_0 \rightarrow 0 \)). The latter can be obtained in a closed form [59]

\[
C(t) = \frac{\langle l(t)l(0) \rangle - \langle l \rangle^2}{\langle l^2 \rangle - \langle l \rangle^2};
\]

\[
\langle l(t)l(0) \rangle = \frac{4\left[3\eta_t\sqrt{\eta_0^2 - \eta_t^2} + 2(\eta_0^2 + \eta_t^2)\arctan(\eta_t/(\eta_0^2 - \eta_t^2))\right]}{\pi\eta_t},
\]

\[
\langle l \rangle = 4\sqrt{\eta_0/\pi}, \quad \langle l^2 \rangle = 6\eta_0.
\]

**Appendix E. Short-time expansion of \( G_{d_0} \)**

Introducing the auxiliary variable \( \phi(t) = \eta_t/\eta_0 \) in Eq. (22) we can write the return joint-density as and expanding to linear order in \( t \) using

\[
\phi(t) \overset{t \rightarrow 0}{\approx} 1 - \sum_{k=1} A_k^2 \frac{t}{\eta_0} \phi^2(t) \overset{t \rightarrow 0}{\approx} 1 - \sum_{k=1} A_k^2 \frac{t}{\eta_0};
\]

we find the partial limits

\[
\exp\left(-\frac{2d^2\phi(t) + (1 - \phi(t))d_0^2}{4\eta_0\phi(t)(1 - \phi(t))}\right) \overset{t \rightarrow 0}{\approx} e^{-1/t} \rightarrow 0,
\]

\[
\text{erfi}\left(\pm 2d\phi(t) + d_0(1 - \phi(t))\right) \overset{t \rightarrow 0}{\approx} \text{erfi}(\pm t^{-1/2}) \rightarrow \pm \infty,
\]

\[
\text{erfi}\left(\frac{d_0(1 - \phi(t))}{2\sqrt{\eta_0\phi(t)(1 - \phi(t)^2)}}\right) \overset{t \rightarrow 0}{\approx} \text{erfi}(\sqrt{t}) \rightarrow 0;
\]

where all the convergences are of exponential order. Therefore, while we can neglect the second erfi, we need to retain the product between the exponential and the two diverging erfis and only then plug them into in Eq. (E.1). Thus considering the expansion for large and real arguments of erfi [38]

\[
\text{erfi}(x) \overset{x \rightarrow \pm \infty, x \in \mathbb{R}}{\approx} \mp i + \left(\frac{1}{x} + \frac{1}{2x^3} + O(x^{-5})\right) \frac{e^{x^2}}{\sqrt{\pi}},
\]

and explicitly, multiplying by the remaining exponentials Eq. (22) becomes (note that \( \mathcal{P}_{d_0}^\infty(l) \mathcal{G}_{d_0}(l, t|l) \equiv \mathcal{J}_{d_0}(l, t; l) \))

\[
\mathcal{J}_{d_0}(d, t; d) \overset{t \rightarrow 0}{\approx} \frac{d^2}{8\pi d_0} e^{-l^2 + d_0^2/2\eta_0(1 + \phi(t))} \times
\]
The Gaussian Network Model

\[
\left\{ \begin{align*}
\frac{2\sqrt{1 + \phi(t)}}{\sqrt{1 - \phi(t)}\eta_0(-2d\phi(t) + d_0(1 - \phi(t)))} + 4\frac{\phi(t)\sqrt{1 - \phi(t)(1 + \phi(t))^{3/2}}}{(-2d\phi(t) + d_0(1 - \phi(t)))^3} e^{-ld_0/\gamma_0(1+\phi(t))} \\
+ \frac{2\sqrt{1 + \phi(t)}}{\sqrt{1 - \phi(t)}\eta_0(2d\phi(t) + d_0(1 - \phi(t)))} + 4\frac{\phi(t)\sqrt{1 - \phi(t)(1 + \phi(t))^{3/2}}}{(2d\phi(t) + d_0(1 - \phi(t)))^3} e^{ld_0/\gamma_0(1+\phi(t))}
\end{align*} \right. 
\]

(E.6)

Using Eq. (E.1) and expanding \( t = 0 \) and introducing \( \kappa = \sum_{k=1}^{N} A_k^2 \) we finally arrive at Eq. (35).

Appendix F. Evaluation of the variance of the occupation time fraction

The direct implementation of Eq. (33) suffers from slow convergence issues. We suspect that this problem has its roots in the (well-known) slow convergence of series involving Hermite polynomials \[72\]. We therefore combine the analytical short-time asymptotics in Eq. (36) with the spectral solution. Defining a small cutoff time \( t_s \ll 1 \) and rewriting Eq. (34) (using the linearity of integration) as

\[
\sigma_{d_0}^2(l,t) = \frac{2P_{d_0}^{eq}(l)}{t} \int_0^{t_s} d\tau (1-\tau/t)G_{d_0}(l,\tau|l) + \frac{2P_{d_0}^{eq}(l)}{t} \int_{t_s}^{t} d\tau (1-\tau/t)[G_{d_0}(l,\tau|l)-P_{d_0}^{eq}(l)]. 
\]

(F.1)

We can explicitly evaluate the first addend using Eq. (36) and evaluate the second term using the spectral expansion \[24\]. Note that the first term in the series (with \( \Lambda_0 = 0 \)) must be treated in a manner different than the rest. Therefore \( \sigma_{d_0}^2(l,t) \) can be conveniently written (and implemented) in the form

\[
\sigma_{d_0}^2(d,t) = 2P_{d_0}^{eq}(l) \left( \frac{8}{3\sqrt{\kappa \pi t}} + \frac{4}{15t^2} \sqrt{\frac{\kappa t}{\pi}} - \frac{P_{d_0}^{eq}(l)}{t} \right)
+ \frac{2}{t^2} \sum_{N\neq 0} V_{N0}(l,d_0)V_{0N}(l,d_0) \left[ (t-t_s)\frac{e^{-\Lambda N t_s}}{\Lambda_N} - \frac{e^{-\Lambda N t_s} - e^{-\Lambda N t}}{\Lambda_N^2} \right]
+ P_{d_0}^{eq}(l)^2 \left( \frac{t_s}{t} - 2 \right) \frac{t_s}{t}. 
\]

(F.2)

Appendix G. Notes on the numerical implementation of the results

Accompanying this article there is a C++ implementation of all analytical results. The code allows the computation the Green's function \( G_{d_0} \), the mean \( \langle \theta_l(l,d_0) \rangle \) and variance \( \sigma_{d_0}^2(d,t) \) of the occupation time fraction, as well as the autocorrelation function \( C_{d_0}(t) \) for a generic Gaussian Network. The connectivity matrix of the network \( \Gamma \) must be provided as a plain text file and is diagonalized using the Armadillo library \[73, 74\].

A closed-form expression of the joint density in Eq. (22) is implemented in the available C++ code. However, for numerical stability and speed of computation it is convenient to implement Eq. (C.12) and perform the final integral numerically using a Gauss-Kronrod quadrature routine \[60\].
Our main results are based on the evaluation of both, Eq. (B.19) and Eq. (D.2). Both require the evaluation of the less common regularized hypergeometric functions $pF_q$. A notable exception is the Arblib library [75], that implements several "special" functions using arbitrary precision arithmetic. The reliable evaluation of such functions is challenging and often requires several different methods to cover the entire domain [76]. Unfortunately this higher reliability comes with a higher computational cost compared to machine precision arithmetic. However hypergeometric functions converge on the entire complex plane if $p \leq q$ [76]. In addition, we only need to evaluate them when all the parameters are positive real numbers. Therefore we implemented the series definitions of these function directly since in our case these converge reasonably fast to a desired accuracy as long as the parameters are not too large.

Many of our results, in particular the autocorrelation function and the variance of the fraction of occupation time, can only be expressed analytically using the eigendecomposition of the Fokker-Plank operator. Unfortunately the computational effort required in the generation of all necessary terms to achieve convergence is huge. In addition, this number scales non-polynomially with the number of beads in the network. Therefore the attached program should be used with care as it does not generate reliable results when the size of the network becomes too large.
The Gaussian Network Model

[1] Katherine Henzler-Wildman and Dorothee Kern. Dynamic personalities of proteins. *Nature*, 450(7172):964–972, Dec 2007.

[2] Wilfred F. van Gunsteren and Herman J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed.*, 29(9):992–1023, Sep 1990.

[3] John L Klepeis, Kresten Lindorff-Larsen, Ron O Dror, and David E Shaw. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struc. Biol.*, 19(2):120–127, Apr 2009.

[4] David E. Shaw, Paul Maragakis, Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, Michael P. Eastwood, Joseph A. Bank, John M. Jumper, John K. Salmon, Yibing Shan, and Willy Wriggers. Atomic-level characterization of the structural dynamics of proteins. *Science*, 330(6002):341–346, 2010.

[5] Marco Baiesi, Enzo Orlandini, Flavio Seno, and Antonio Trovato. Exploring the correlation between the folding rates of proteins and the entanglement of their native states. *J. Phys. A: Math. Theor.*, 50(50):504001, December 2017.

[6] Cristian Micheletti, Jayanth R. Banavar, Amos Maritan, and Flavio Seno. Protein structures and optimal folding from a geometrical variational principle. *Phys. Rev. Lett.*, 82:3372–3375, Apr 1999.

[7] Monique M. Tirion. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Phys. Rev. Lett.*, 77(9):1905–1908, August 1996.

[8] Prince E. Rouse. A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.*, 21(7):1272–1280, July 1953.

[9] Flory, P. J. Statistical thermodynamics of random networks. *Proc. R. Soc. Lond. A*, 351(1666):351–380, November 1976.

[10] Ivet Bahar, Ali Rana Atilgan, and Burak Erman. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding and Design*, 2(3):173–181, June 1997.

[11] Turkan Haliloglu, Ivet Bahar, and Burak Erman. Gaussian Dynamics of Folded Proteins. *Phys. Rev. Lett.*, 79(16):3090–3093, October 1997.

[12] Pemra Doruker, Ali Rana Atilgan, and Ivet Bahar. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins*, 40(3):512–524, 2000.

[13] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, and I. Bahar. Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophys. J.*, 80(1):505–515, January 2001.

[14] M. Delarue and P. Dumas. On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. *Proc. Natl. Acad. Sci. USA*, 101(18):6957–6962, May 2004.

[15] Gunnar F. Schröder, Axel T. Brunger, and Michael Levitt. Combining Efficient Conformational Sampling with a Deformable Elastic Network Model Facilitates Structure Refinement at Low Resolution. *Structure*, 15(12):1630–1641, December 2007.

[16] Dengming Ming and Rafael Brüschweiler. Reorientational Contact-Weighted Elastic Network Model for the Prediction of Protein Dynamics: Comparison with NMR Relaxation. *Biophys. J.*, 90(10):3382–3388, May 2006.

[17] Qian-Yuan Tang, Yang-Yang Zhang, Jun Wang, Wei Wang, and Dante R. Chialvo. Critical Fluctuations in the Native State of Proteins. *Phys. Rev. Lett.*, 118(8):088102, February 2017.

[18] Kay Hamacher, Joanna Trylska, and J. Andrew McCammon. Dependency Map of Proteins in the Small Ribosomal Subunit. *PLoS Comput Biol*, 2(2):e10, February 2006.

[19] A. J. Rader, G. Anderson, B. Isin, H. G. Khorana, I. Bahar, and J. Klein-Seetharaman. Identification of core amino acids stabilizing rhodopsin. *Proc Natl Acad Sci USA*, 101(19):7246–7251, May 2004.
[20] Ozlem Keskin, Stewart R. Durell, Ivet Bahar, Robert L. Jernigan, and David G. Covell. Relating Molecular Flexibility to Function: A Case Study of Tubulin. *Biophys. J.*, 83(2):663–680, August 2002.

[21] Jingwei Weng, Jianpeng Ma, Kangnian Fan, and Wenning Wang. The Conformational Coupling and Translocation Mechanism of Vitamin B12 ATP-Binding Cassette Transporter BtuCD. *Biophys. J.*, 94(2):612–621, January 2008.

[22] Ines Putz and Oliver Brock. Elastic network model of learned maintained contacts to predict protein motion. *PLoS ONE*, 12(8):e0183889, August 2017.

[23] Wenjun Zheng and Bernard Brooks. Identification of Dynamical Correlations within the Myosin Motor Domain by the Normal Mode Analysis of an Elastic Network Model. *J. Mol. Bio.*, 346(3):745–759, February 2005.

[24] Sibsankar Kundu, Dan C. Sorensen, and George N. Phillips. Automatic domain decomposition of proteins by a Gaussian Network Model. *Proteins*, 57(4):725–733, December 2004.

[25] Zhiyong Zhang, Yunyu Shi, and Haiyan Liu. Molecular Dynamics Simulations of Peptides and Proteins with Amplified Collective Motions. *Biophys. J.*, 84(6):3583–3593, June 2003.

[26] K. Hamacher. Relating sequence evolution of HIV1-protease to its underlying molecular mechanics. *Gene*, 422(1-2):30–36, October 2008.

[27] Jinjun Zhang, Bonsung Koo, Yingtao Liu, Jin Zou, Aditi Chattopadhyay, and Lenore Dai. A novel statistical spring-bead based network model for self-sensing smart polymer materials. *Smart Mater. Struct.*, 24(8):085022, August 2015.

[28] Jinjun Zhang, Bonsung Koo, Nithya Subramanian, Yingtao Liu, and Aditi Chattopadhyay. An optimized cross-linked network model to simulate the linear elastic material response of a smart polymer. *J. INTEL. MAT. SYST. STR.*, 27(11):1461–1475, July 2016.

[29] Florence Tama and Charles L. Brooks. The Mechanism and Pathway of pH Induced Swelling in Cowpea Chlorotic Mottle Virus. *J. Mol. Bio.*, 318(3):733–747, May 2002.

[30] Indira H. Shrivastava and Ivet Bahar. Common Mechanism of Pore Opening Shared by Five Different Potassium Channels. *Biophysical Journal*, 90(11):3929–3940, June 2006.

[31] Nicolas Bocquet, Hugues Nury, Marc Baaden, Chantal Le Poupon, Jean-Pierre Changeux, Marc Delarue, and Pierre-Jean Corringer. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature*, 457(7225):111–114, January 2009.

[32] Giovanni Pinamonti, Sandro Bottaro, Cristian Micheletti, and Giovanni Bussi. Elastic network models for RNA: a comparative assessment with molecular dynamics and SHAPE experiments. *Nucleic Acids Res.*, 43(15):7260–7269, September 2015.

[33] Joanna I. Sulikowska, Andrzej Kloczkowski, Taner Z. Sen, Marek Cieplak, and Robert L. Jernigan. Predicting the order in which contacts are broken during single molecule protein stretching experiments. *Proteins*, 71(1):45–60, April 2008.

[34] Herbert Goldstein, Charles P. Poole, and John L. Safko. *Classical mechanics*. Addison Wesley, San Francisco, NJ, 3. ed edition, 2002. OCLC: 248389949.

[35] Ivet Bahar, Timothy R. Lezon, Ahmet Bakan, and Indira H. Shrivastava. Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins. *Chem. Rev.*, 110(3):1463–1497, March 2010.

[36] S. Nicolay and Y.-H. Sanejouand. Functional Modes of Proteins Are among the Most Robust. *Phys. Rev. Lett.*, 96(7):078104, February 2006. Publisher: American Physical Society.

[37] Gerald Wilemski and Marshall Fixman. Diffusion-controlled intrachain reactions of polymers. I Theory. *J. Chem. Phys.*, 60(3):866–877, February 1974.

[38] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Dover books on mathematics. Dover Publ, New York, NY, 9. dover print. edition, 2013. OCLC: 935935300.

[39] Dominique Foata. Some hermite polynomial identities and their combinatorics. *Adv. Appl. Math.*, 2(3):250–259, 1981.

[40] Hannes Risken and Till Frank. *The Fokker-Planck Equation: Methods of Solution and
The Gaussian Network Model

Applications. Springer Series in Synergetics. Springer-Verlag, Berlin Heidelberg, 2 edition, 1996.

[41] Andrea Amadei, Antonius B. M. Linssen, and Herman J. C. Berendsen. Essential dynamics of proteins. *Proteins*, 17(4):412–425, December 1993.

[42] Alessio Lapolla, David Hartich, and Aljaž Godec. Spectral theory of fluctuations in time-average statistical mechanics of reversible and driven systems. *Phys. Rev. Research*, 2(4):043084, October 2020.

[43] Alessio Lapolla and Aljaž Godec. Manifestations of Projection-Induced Memory: General Theory and the Tilted Single File. *Front. Phys.*, 7, 2019.

[44] Gardiner, C.W. *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences*. Springer-Verlag, second edition, 1985.

[45] Kevin Truong and Mitsuhiro Ikura. The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo. *CURR OPIN STRUC BIOL*, 11(5):573–578, September 2001.

[46] Weixiang Ye, Markus Götz, Sirin Celiksoy, Laura Tüting, Christoph Ratzke, Janak Prasad, Julia Ricken, Seraphine V. Wegner, Rubén Ahijado-Guzmán, Thorsten Hugel, and Carsten Sönnichsen. Conformational Dynamics of a Single Protein Monitored for 24 h at Video Rate. *Nano Lett.*, 18(10):6633–6637, October 2018.

[47] Kac, M. On distributions of Certain Weiner functionals. *Trans. Amer. Math. Soc*, 65:1–13, 1949.

[48] Ju-Yi Yen and Marc Yor. *Local times and excursion theory for Brownian motion: a tale of Wiener and Itô measures*. Number 2088 in Lecture notes in mathematics. Springer, Cham, 2013. OCLC: 862993594.

[49] Satya N. Majumdar and Alain Comtet. Local and Occupation Time of a Particle Diffusing in a Random Medium. *Phys. Rev. Lett.*, 89(6):060601, July 2002.

[50] Satya N. Majumdar. Brownian Functionals in Physics and Computer Science. In *The Legacy of Albert Einstein*, pages 93–129. WORLD SCIENTIFIC, December 2006.

[51] Alessio Lapolla and Aljaž Godec. Unfolding tagged particle histories in single-file diffusion: exact single- and two-tag local times beyond large deviation theory. *New J. Phys.*, 20(11):113021, November 2018.

[52] Irina V. Gopich and Attila Szabo. Single-Macromolecule Fluorescence Resonance Energy Transfer and Free-Energy Profiles. *J. Phys. Chem. B*, 107(21):5058–5063, May 2003.

[53] Christoph W. Müller and Georg E. Schulz. Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor Ap5A refined at 1.9 Å resolution. *J. Mol. Bio.*, 224(1):159–177, March 1992.

[54] Cw Müller, Gj Schlauderer, J Reinstein, and Ge Schulz. Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure*, 4(2):147–156, February 1996.

[55] Katherine A. Henzler-Wildman, Vu Thai, Ming Lei, Maria Ott, Magnus Wolf-Watz, Tim Fenn, Ed Pozharski, Mark A. Wilson, Gregory A. Petsko, Martin Karplus, Christian G. Hübner, and Dorothee Kern. Intrinsic motions along an enzymatic reaction trajectory. *Nature*, 450(7171):838–844, December 2007.

[56] Yuqing Zheng and Qiang Cui. Multiple Pathways and Time Scales for Conformational Transitions in apo-Adenylate Kinase. *J. Chem. Theory Comput.*, 14(3):1716–1726, March 2018.

[57] A. Bakan, L. M. Meireles, and I. Bahar. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics*, 27(11):1575–1577, June 2011.

[58] Alessio Lapolla and Aljaž Godec. A Toolbox for Quantifying Memory in Dynamics Along Reaction Coordinates. *arXiv:2101.05237 [cond-mat, physics:physics]*, January 2021. arXiv: 2101.05237.

[59] https://www.boost.org/doc/libs/1.73.0/libs/math/doc/html/math_toolkit/gauss_kronrod.html, August 2020.

[60] Giovanni Lipari and Attila Szabo. Model-free approach to the interpretation of nuclear magnetic
The Gaussian Network Model

resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, 104(17):4546–4559, 1982.

[62] Angelo Perico, Roberto Pratolongo, Karl F. Freed, Richard W. Pastor, and Attila Szabo. Positional time correlation function for one-dimensional systems with barrier crossing: Memory function corrections to the optimized Rouse–Zimm approximation. *J. Chem. Phys.*, 98(1):564–573, 1993.

[63] Alessio Lapolla and Aljaž Godec. Single-file diffusion in a bi-stable potential: Signatures of memory in the barrier-crossing of a tagged-particle. *J. Chem. Phys.*, 153(19):194104, 2020.

[64] Y. Tian, B. Shirinzadeh, and D. Zhang. Design and dynamics of a 3-DOF flexure-based parallel mechanism for micro/nano manipulation. *Microelectron Eng.*, 87(2):230–241, February 2010.

[65] Danelle P. Sierra, Nathan A. Weir, and James Frank Jones. *A review of research in the field of nanorobotics*. Sandia National Laboratories, October 2005.

[66] Henry Crapo. Structural rigidity. *Structural topology, 1979, núm. 1*, 1979. Publisher: Université du Québec à Montréal.

[67] C.R. Calladine. Buckminster Fuller’s “Tensegrity” structures and Clerk Maxwell’s rules for the construction of stiff frames. *Int J Solids Struct*, 14(2):161–172, 1978.

[68] J. Clerk Maxwell. L. On the calculation of the equilibrium and stiffness of frames. *Lond.Edinb.Dubl.Phil.Mag.*, 27(182):294–299, April 1864.

[69] S. Pellegrino and C.R. Calladine. Matrix analysis of statically and kinematically indeterminate frameworks. *Int J Solids Struct*, 22(4):409–428, 1986.

[70] Gradshteyn, I. S. and Ryzhik, I. M. *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, seventh edition, 2007.

[71] Wolfram Research, Inc. Mathematica, Version 12.0, 2019.

[72] John P. Boyd. The rate of convergence of Hermite function series. *Math. Comp.*, 35(152):1309–1309, January 1980.

[73] Conrad Sanderson and Ryan Curtin. Armadillo: a template-based C++ library for linear algebra. *J. Open Source Softw.*, 1(2):26, 2016. Publisher: The Open Journal.

[74] Conrad Sanderson and Ryan Curtin. A User-Friendly Hybrid Sparse Matrix Class in C++. In James H. Davenport, Manuel Kauers, George Labahn, and Josef Urban, editors, *Mathematical Software – ICMS 2018*, pages 422–430, Cham, 2018. Springer International Publishing.

[75] Fredrik Johansson. Arb: Efficient Arbitrary-Precision Midpoint-Radius Interval Arithmetic. *IEEE T COMPUT*, 66(8):1281–1292, August 2017.

[76] Fredrik Johansson. Computing Hypergeometric Functions Rigorously. *ACM Trans. Math. Softw.*, 45(3), August 2019. Place: New York, NY, USA Publisher: Association for Computing Machinery.