Topological analysis of polymeric melts: Chain length effects and fast-converging estimators for entanglement length

Robert S. Hoy, Katerina Foteinopoulou, and Martin Kröger

Materials Research Laboratory, University of California, Santa Barbara, CA 93106, U.S.A.

University Politécnica de Madrid (UPM), José Gutiérrez Abascal 2, E–28006 Madrid, Spain

Polymer Physics, ETH Zürich, Department of Materials, CH–8093 Zürich, Switzerland

(Dated: June 21, 2009)

The value of theories of dense polymeric systems [1].

I. INTRODUCTION

The features of polymer melt rheology are determined primarily by the random-walk-like structure of the constituent chains and the fact that chains cannot cross. The motion of sufficiently long chains is limited by “entanglements” which are topological constraints imposed by the other chains. These become important and dramatically change many melt properties (e. g. diffusivity and viscosity) as the degree of polymerization becomes larger than the “entanglement length” \( N_e \).

The value of \( N_e \) is both a key quantity measured in mechanical and rheological experiments and a key parameter in tube theories of dense polymeric systems [1].

\( N_e \) is usually considered to be a number set by chemistry and thermodynamic conditions (e. g. chain stiffness, concentration, and temperature). It has been empirically related to a “packing” length [2]: \( N_e \propto (\rho b^3)^{-2} \) [3], where \( \rho \) is monomer number density and \( b^3 = \langle R_{ee}^2/N - 1 \rangle \) is the statistical segment length of chains with end-to-end distance \( R_{ee} \) and mean degree of polymerization \( N \). In terms of individual entanglements, \( N_e \) is defined as the ratio between \( N \) and the mean number of entanglements per chain (\( \langle Z \rangle \)), in the limit of infinite chain length,

\[
N_e = \lim_{{N \to \infty}} \frac{N}{\langle Z \rangle}.
\]  

We call a function \( N_e(N) \) an \( N_e \)-estimate if it has the property

\[
\lim_{{N \to \infty}} N_e(N) = N_e,
\]  

where \( N_e \) is a system dependent but \( N \)-independent quantity. Comparing Eq. (1) with (2) does not imply choosing \( N_e(N) = N / \langle Z \rangle \). The typical experimental \( N_e \)-estimate uses the plateau modulus \( G'_N \) [1]:

\[
N_e(N) = \frac{4m\rho k_B T}{5G'_N},
\]

where \( m \) is monomer mass, \( k_B \) is the Boltzmann constant, and \( T \) is temperature.

A closely related theoretical construct is the primitive path (PP), defined by Edwards [4] as the shortest path a chain fixed at its ends can follow without crossing any other chains. Rubinstein and Helfand [5] realized that the entanglement network of a system could be obtained by reducing all chains to their PPs simultaneously. Such a reduction process is analytically intractable, but has recently been achieved by computer simulations [7, 10, 11, 12, 13] which generate networks of PPs from model polymer melts, glasses, random jammed packings and solutions. These simulations estimate \( N_e \) either from the chain statistics of the PPs [6, 8, 9] or from direct enumeration of entanglements (contacts between PPs) [7, 11, 12, 14, 15], which determines \( \langle Z \rangle \).

Chain-statistical and direct enumeration approaches produce different results for \( N_e \) for the same atomistic configurations, suggesting that “rheological” and “topological” entanglements are not equivalent [15]. This discrepancy has been attributed to the fact that chemical distances between entanglements are not uniform, but rather are drawn from broad distributions [7, 10, 14, 15, 16], even at equilibrium. Studies of how entanglement properties change with \( N \) are therefore

---

* Corresponding author. E–mail: robertscothoy@gmail.com

†Electronic address: http://www.complexfluids.ethz.ch
of obvious interest. Moreover, primitive path statistics enter recently developed slippink–based models [17].

In this paper we seek an “ideal” $N_c$–estimate which approaches $N_c$ at the smallest possible $N$. There have been several attempts in the literature, summarized below, to derive $N_c$–estimates, but these have all exhibited poor convergence (i.e. by approaching $N_c$ only at large $N \gg N_c$). Molecular dynamics simulation times increase with chain length $N$ approximately as $N^3$ at large $N$ (relaxation time $\tau \propto N^{3.5}$ times system size $\propto N^{3/2}$), so improved $N_c$-estimates have obvious benefits for computationally efficient determination of $N_c$. By analyzing a number of coarse-grained and atomistic systems, we find a rather general solution to this problem of setting up a $N_c$–estimator which allows to predict $N_c$ from weakly entangled linear polymer melts.

The organization of this paper is as follows. Section II presents the polymer models used here and the topological analysis methods which provide us with the entanglement network (primitive paths). Section III distinguishes between valid and quickly converging (ideal) $N_c$–estimators, and discusses some model– and method–independent issues with existing estimators. Examples are given which highlight systematic errors caused by improper treatment of chain ends and of the non-Gaussian statistics of chains and primitive paths. Section IV derives two (potentially) near–ideal estimators which extract $N_c$ from the variation of entanglement characteristics with $N$. Section V presents and discusses numerical results for these estimators for two very different model polymers. We verify that they are basically ideal, explain why this is so, and derive simplified forms which further illustrate the connection of $N_c$ to chain structure and entanglement statistics and are also near-ideal. Section VI contains conclusions, and two Appendices provide additional technical details.

II. POLYMER MODELS AND METHODS

A. Polymer model systems

We have created thoroughly equilibrated configurations for two very different (but commonly used) model polymer melts; monodisperse ‘Kremer–Grest’ bead–spring chains, and atomistic, polydisperse polyethylene. These two are chosen because they have similar values of $N_c$ but very different chain stiffness constants $C(\infty)$. Polyethylene is much more “tightly entangled” [18] in the sense of having a much lower value of $N_c/C(\infty)$; cf. Tabs. II and III.

The bead spring model [19] captures the features of polymers which are key to entanglement physics, most importantly chain connectivity/uncrossability. Each chain contains $N$ beads of mass $m$. All beads interact via the truncated and shifted Lennard-Jones potential $U_{LJ}(r) = 4\epsilon_{LJ}\left[(\sigma/r)^{12} - (\sigma/r)^{6}\right] + (\sigma/r_0)^6$, where $r_0 = \sigma^{1/6}$ is the cutoff radius and $U_{LJ}(r) = 0$ for $r > r_c$. Here $\sigma$ is the bead diameter and $\epsilon_{LJ}$ is the binding energy, which are both set to 1; all quantities will thus be dimensionless and given in the conventional Lennard–Jones units. Covalent bonds between adjacent monomers on a chain are modeled using the FENE potential

$$U(r) = -\frac{1}{2}kR_0^2 \ln[1 - (r/R_0)^2],$$

with the canonical parameter choices $R_0 = 1.5$ and $k = 30$ [19]. The equilibrium bond length is $l_0 \approx 0.96$. This model is hereafter referred to as the “LJ + FENE” model.

Values of the density and temperature ($\rho = 0.85$ and $T = 1.0$) are those typically used for melt simulations [6, 8, 19]. All systems contain 280,000 total beads. While all are monodisperse, we employ a wide range of chain lengths, $4 \leq N \leq 3500$. Those with $N \geq 100$ are equilibrated using the “double-bridging hybrid” (DBH) algorithm [20]. DBH uses molecular dynamics to update monomer positions and Monte Carlo chain-topology-altering moves [21] to overcome the slow diffusive dynamics of entangled chains. All equilibration simulations were performed using the LAMMPS molecular dynamics code. Ref. [19] predicted $N_c \approx 35$ at the above–mentioned state point using various “theoretical” measures applied to systems with $N \leq 400$, while a similar analysis in [23] predicted $N_c \approx 75$.

In all simulations of the atomistic polyethylene (PE) melt, the united atom (UA) representation is adopted. Accordingly, carbon atoms along with their bonded hydrogens are lumped into single spherical interacting sites. There is no distinction between methyl and methylene units in the interaction potentials. All bond lengths are kept constant ($l_0 = 1.54 \AA$), while bending and torsion angles are respectively governed by harmonic and sum-of-cosines potentials [24, 25]. Pair interactions between all intermolecular neighbors, and intramolecular neighbors separated by more than three bonds, are described by the 12–6 Lennard-Jones potential. The parameters of the mathematical formulas for the bonded and non-bonded interactions are given in Refs. [19, 21, 24, 25]. These interaction potentials yield accurate predictions of the volumetric, structural and conformational properties of PE melts over a wide range of chain lengths and temperatures [21, 24].

All atomistic PE systems were equilibrated through Monte Carlo (MC) simulations based on advanced chain-connectivity-altering algorithms: the end-bridging [26] and double bridging [21, 25] moves along with their intramolecular variants. The simulated systems are characterized by average chain lengths from $N = 24$ up to $N = 1000$, with a small degree of polydispersity. Chain lengths are uniformly distributed over the interval $[(1 - \Delta)N, (1 + \Delta)N]$. Here $\Delta$, the half width of the uniform chain length distribution reduced by $N$, is 0.5 and 0.4 for $24 \leq N \leq 224$ and 270 $\leq N \leq 1000$, respectively. More details about the MC scheme, including a full list of moves, attempt probabilities and acceptance rates, can be found elsewhere [24]. Equilibration at all length scales, which is essential to obtaining meaningful results from entanglement analyses [27], was verified using several metrics [24]. In this study, results are presented for $T = 400$ K and $T = 450$ K, both for $P = 1$ atm.

B. Entanglement network and primitive paths

For the melt configurations the reduction to primitive paths was performed using two methods, PPA and Z, using the procedures described in Refs. [6, 7, 8, 14]. PPA simulations used
LAMMPS and Z simulations used the Z1 code. Both PPA and Z1 analyses are performed for the LJ+FENE model, while only Z1 analysis is performed for PE. In both methods, all chain ends are fixed in space. Intrachain excluded volume interactions are disabled while chain uncrossability is retained. Both classical PPA and geometrical methods (Z1 or CReTA) provide the configuration of the entanglement network and the contour lengths of each primitive path. In PPA, disabling intrachain excluded volume produces a tensile force in chains which reduces the contour lengths. In Z1, contour lengths are monotonically reduced through geometrical moves in the limit of zero primitive chain thickness. In addition to \( L_{pp} \) and the configuration of the entanglement network, Z1 analysis also yields the number of interior “kinks” \( Z \), in the three-dimensional primitive path of each chain. \( \langle Z \rangle \) is considered to be proportional to the number of entanglements, regardless of the details of the definition used to define an entanglement.

Runs end when the mean length of the primitive paths, \( \langle L_{pp} \rangle \), and/or the mean number of interior kinks per chain, \( \langle Z \rangle \), converge. Self entanglements are neglected, but their number is inconsequential for the systems considered here. The CReTA method works similarly, and the conclusions reached here for Z1 analysis should apply similarly to CReTA results.

Table I summarizes chain and primitive path dimensions as well as \( \langle Z \rangle \) for LJ+FENE chains with \( 20 \leq N \leq 3500 \). Statistically independent initial states were used so that the random error on all quantities is\( \lesssim 2.5\% \). It is remarkable that PPA and Z1 data for \( \langle L_{pp} \rangle \) and also \( \langle L_{pp}^2 \rangle \) are so similar, considering the differences between the contour length reduction methods. Relative to Z1 results, PPA values of \( \langle L_{pp} \rangle \) are increased by finite chain thickness effects and decreased by chain end slipoff. Both these effects should decrease in strength as \( N \) increases, and indeed \( \langle L_{pp}^2 \rangle_{PPA}/\langle L_{pp}^2 \rangle_{Z1} \) decreases from \( \sim 1.17 \) to \( \sim 1.06 \) over the range \( 20 \leq N \leq 3500 \). A very comparable trend is offered by \( \langle L_{pp}^2 \rangle_{PPA}/\langle L_{pp}^2 \rangle_{Z1} \).

PPA results for the shortest chains \( (N < 20) \) are not presented. Standard PPA is unreliable for very short chains because the presence of a high concentration of fixed chain ends combined with the finite bead diameter effectively inhibits relaxation. These problems are even worse for topological analysis of lattice polymer systems, see e. g. Ref. [12]. In the following, where (as will be shown) accurate data from very short chains is important, we focus on Z1 results.

### III. TOWARDS VALID ESTIMATORS

A basic task of topological analysis is to calculate \( N_e \) from the full microscopic configuration of the entanglement network. The simplest approaches employ only the mean square end–to–end distance of chains, \( \langle R_e^2 \rangle \) and either the mean length of the primitive paths, \( \langle L_{pp} \rangle \) or the mean number of kinks, \( \langle Z \rangle \). Notice that \( \langle Z \rangle \) is not an integer, but semipositive, \( \langle Z \rangle \geq 0 \). In order to estimate \( N_e \) from weakly entangled systems one of course needs physical insight: when this is limited, a good \( N_e \)-estimator can only be guessed.

Some restrictions arise from a purely mathematical viewpoint. A valid estimator \( N_e(N) \) has the following properties.

(i) obeys Eq. (3) and uses information from polymer configurations whose mean chain length does not exceed \( N \);
(ii) either yields \( N_c(N) \geq N \) or leaves \( N_c(N) \) undefined for a system of completely unentangled \((Z) = 0\) chains.

An “ideal” estimator we define to

(iii) correctly predict \( N_c \) for all \( N \) exceeding \( N_c \), or for all \( \langle Z \rangle \) exceeding unity.

Accordingly, for an ideal estimator, the following weaker conditions hold. An ideal estimator

(iv) diverges for a system of rodlike chains possessing \( N_c = \infty \), and

(v) exhibits \( N_c(N) \leq N \) when each chain has in average more than a single entanglement, \( \langle Z \rangle > 1 \).

The following two subsections repeat earlier approaches to estimate \( N_c \). Basic considerations of finite chain length effects, errors from improper treatment of non-Gaussian structure, and the general behavior of quantities entering \( N_c \) are discussed. These subsections are meant to prepare the reader for the ideal estimators to be presented in Sec. [IV]. They reflect the chronology of our search for better estimators and help the reader to understand the magnitude of improvements presented in Section [V]. The arguments given here ultimately point the way to construct ideal estimators.

### A. Non-Ideal Estimators

Modelling primitive paths as random walks, Everaers et al. [6] developed an estimator (which we denote as “classical S–coil”) which operates on results for configurations (“ coils”) of a single (S) chain length,

\[
N_c(N) = (N - 1) \frac{\langle R_{cc}^2 \rangle}{\langle L_{pp} \rangle^2}.
\]

(4)

The classical S–coil estimate [4] is useful because (for long chains) it relates changes in chain structure to rheological trends [6, 18]. However, while it fulfills basic requirements (i) and (ii) (both unentangled and rodlike chains have \( R_{cc} = L_{pp} \)), it lacks properties (iii) and (iv). As the exact relation of \( \langle R_{cc}^2 \rangle /\langle L_{pp} \rangle^2 \) and \( \langle Z \rangle \) is unknown, it is a priori unclear whether it has property (v).

The corresponding estimator operating on the number of kinks, \( \langle Z \rangle \), and originally employed in [7], denoted here as “classical S–kink”, is

\[
N_c(N) = \frac{N(N - 1)}{\langle Z \rangle(N - 1) + N},
\]

(5)

which fulfills the basic requirements (i) and (ii), and also (v), but lacks (iii) according to Ref. [10] and (iv) by definition. The presence of both \( N = 1 \) and \( N \) in Eqs. (4), (5), and subsequent estimators reflects the fact that it is the existence of a bond rather than a bead which is responsible for the presence or absence of an entanglement between two chain contours.

The performance of the two classical estimators (4) and (5) for the two polymer models considered here is illustrated in Fig. 1. Values of \( N_c(N) \) converge very slowly with increasing \( N \). As expected from their form, but contrary to both rheological intuition and condition (ii), values of \( N_c(N) \) drop strongly with decreasing \( N \). For marginally entangled chains (where \( N \) is just large enough so that \( \langle Z \rangle \) is small but nonzero), both classical estimators yield \( N_c(N) \leq N - 1 \). For example, if \( N = 20 \), they both predict \( N_c(N) = 17 \), which is close to the (improper) upper bound \( N - 1 = 19 \). This prediction obviously has no connection to the actual topology of the system.

Thus Eqs. (4) and (5) always underestimate, but never overestimate \( N_c \). This feature of the two estimators in the limit of unentangled chains is particularly (if retrospectively) disappointing, as it is incompatible with goal (iii). Similar behavior was reported (but not analyzed as in this paper) in Refs. [31, 32, 33].

Other previously published \( N_c \)--estimators [9, 11, 16, 34] also have some, but not all, of properties (i)-(v). One of the most promising was proposed in Ref. [9]. It estimates \( N_c \) from the internal statistics of primitive paths, for a single \( N \). The squared Euclidean distances \( \langle R_{cc}^2 \rangle /\langle L_{pp} \rangle^2 \) between monomers separated by chemical distance \( n \geq N - 1 \) after topological analysis (i.e., the chain statistics of the primitive paths) were fit [9] to those of a freely rotating chain with fixed bond length fixed bending angle. \( N_c \) was then identified with the chain stiffness constant \( C(\infty) \) of the freely rotating chain [35]. This estimator does not obviously fail to meet any of conditions (i)-(v). In Ref. [9] it gave values of \( N_c(N) \) which decreased more slowly than Eq. (4) as \( N \) decreased. Unfortunately, its
predictions agree with Eq. (4) at moderate $N \geq 100$ and thus it fails condition (iii).

New S-estimators based on modifications to Eqs. (4) and (5) may be proposed. During the course of developing ideal estimators (to be introduced in Sec. IV), we developed two modified single chain length estimators which tend to approach $N_e$ from above rather than from below. These are the “modified S–kink” estimator

$$N_e(N) = \frac{N}{Z},$$

(6)

and the mathematically similar "modified S–coil" estimator

$$N_e(N) = (N - 1) \left( \frac{\langle L_{pp}^2 \rangle}{\langle R_{ee}^2 \rangle} - 1 \right)^{-1},$$

(7)

A motivation for the use of $\langle L_{pp}^2 \rangle$ rather than $\langle L_{pp} \rangle^2$ in Eq. (7) appears in Appendix A. Fig. 1 shows results for Eqs. (6) and (7) for the same systems analyzed in Fig. 1. Both modified single–chain estimators give $N_e(N) = \infty$ for unentangled chains, thus fulfilling criterion (iv) in addition to (i) and (ii), but they still fail to fulfill goal (iii) since they tend to overestimate $N_e$ for weakly entangled chains.

B. Errors from improper treatment of non-Gaussian structure and chain ends

Critically, none of the above-mentioned estimators seem to be able to predict $N_e(N) = N_e$ for weakly entangled systems with a slightly positive $\langle Z \rangle \lesssim 1$. All above–cited previous works as well as Eq. (7) have only produced convergence for $N \gg N_e$, and we are not aware of any studies where convergence has been achieved at $N \approx N_e$, i.e., we are not aware of the former existence of any ideal $N_e$–estimator. However, the failure of so many previous attempts both makes it worth examining the common reasons why they have failed, and in fact points the way to creating ideal $N_e$–estimators.

To leading order in $\epsilon \equiv (N - 1)^{-1}$ (i.e. the inverse number of bonds), data for a wide variety of model polymers (see e.g. Refs. [20, 24, 31]), as well as the data obtained in this study (see Fig. 3) are consistent with

$$\langle R_{ee}^2 \rangle(\epsilon) = D/\epsilon - Y;$$

(8)

where the relative magnitudes of the constant coefficients $Y$ and $D$ depend on factors such as chain stiffness, molecular details and thermodynamic conditions.

Also, orientations of successive PP segments are correlated [11], so $\langle L_{pp} \rangle^2$ should not be simply quadratic in chain length. The expected leading order behavior of $\langle L_{pp} \rangle^2$ is

$$\langle L_{pp} \rangle^2(\epsilon) = A/\epsilon^2 + B/\epsilon,$$

(9)

where $B$ contains contributions from non-Gaussian statistics and contour length fluctuations [1]. Relationships (8) and (9) are consistent with data reported elsewhere (e.g. Refs. [24, 32]) as well as our own data, as shown in Fig. 3.

At this point it is worthwhile to mention that we are going to make use of (9), which is able to capture our results for $\langle L_{pp} \rangle^2$ down to chain lengths $N$ small compared with $N_e$, to devise an ideal estimator in Sec. IV. Relationship (8) however, as we will see, will not be required to hold to devise an ideal estimator.

Inserting Eqs. (8) and (9) into the classical and modified S–coil Eqs. (4) and (7) respectively give, to leading order in $\epsilon$,

$$N_e(N) \quad \frac{D}{A} \frac{AY + BD}{A^2} \epsilon + O(\epsilon^2),$$

(10a)

$$N_e(N) \quad \frac{D}{A} \frac{D^2 - AY - BD}{A^2} \epsilon + O(\epsilon^2),$$

(10b)

Thus non-Gaussian structure of both chains and primitive paths naturally lead to systematic $O(\epsilon) \approx O(1/N)$ errors in

FIG. 2: (color online) (a) Performance of modified S–kink estimator (5) (lower two curves) and the modified S–coil (7) (upper two curves), which approach $N_e$ from above. Data are for the same systems analyzed in Fig. 1. The single-configuration estimator for kinks exhibits an improved convergence behavior compared with (5). Under circumstances discussed in Sec. V, application of both modified and original classical estimators allows one to obtain lower and upper bounds on $N_e$ which tighten with increasing $N$. (b) Shown are the relative differences (“gap [%]”) between $N_e(N)$ values shown in Fig. 1 and the ones plotted in part (a) of the current graph. Differences are smaller for $N_e$ estimated from kinks (lower two curves).
earlier estimators for $N_c$.

\[ \langle Z \rangle (e) = \frac{G}{e - H}, \quad (11) \]

where $G$ and $H$ are both positive. This assumption is actually consistent with the data in Tab. I and previous works [24]; see also Sec. IV. The classical and modified S–kink Eqs. (5) and (6) then become

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} - \frac{1 - G - H}{G^2} \epsilon + O(\epsilon^2), \quad (12a) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} + \frac{G + H}{G^2} \epsilon + O(\epsilon^2). \quad (12b) \]

Again, systematic $O(\epsilon)$ errors are predicted. In this case, however, the source is chains being too short to be in the asymptotic entangled limit defined by Eq. (11).

A key to understanding the failure of previous $N_c$–estimators is that differences in the prefactors of the $O(\epsilon)$ errors [Eqs. (10) and (12)] arise from different treatment of chain ends. The classical S–kink equation (5) strictly overestimates $N_c$ as long as $G + H < 0$, and the modified S–kink equation (6) overestimates $N_c$, since both $G$ and $H$ are positive. Similarly, the prefactor $(AY + BD - D^2)/A^2$ of the systematic $O(\epsilon)$ error in the modified S–coil equation (7) contains two contributions of different origins. $(AY + BD)/A^2$ arises from the Gaussian-chain approximation used, while $-D^2/A^2$ arises from the attempt to correct for chain end effects (i.e. the “-1”).

We have determined the coefficients $A$, $B$, $D$, $G$, $H$, and $Y$ using all available data from our simulations; their values for both polymer models are shown in Tab. II. Coincidentally, for LJ+FENE chains, $(AY + BD)/A^2 \approx 5 \times 10^3$ and $D^2/A^2 \approx 7 \times 10^3$. The systematic $O(\epsilon)$ error for the modified S–coil (7) is actually small for LJ+FENE systems due to the near-cancellation of its contributing terms. There is no reason to believe this behavior is general, and tests on additional polymer models would be necessary to better characterize how rapidly the modified S–coil typically converges. However, it is reasonable to expect it typically converges more rapidly than the classical S–coil (4).

Before turning to ideal estimators, we mention that the modified S–kink (6) can be regarded as corrected version of classical S–kink (5), as it eliminates an $O(\epsilon)$ error from the latter, and thus converges faster.

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{1 - G - H}{G^2} \epsilon + O(\epsilon^2), \quad (12a) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{G + H}{G^2} \epsilon + O(\epsilon^2). \quad (12b) \]

\[ \langle Z \rangle (e) = \frac{G}{e - H}, \quad (11) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{1 - G - H}{G^2} \epsilon + O(\epsilon^2), \quad (12a) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{G + H}{G^2} \epsilon + O(\epsilon^2). \quad (12b) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{1 - G - H}{G^2} \epsilon + O(\epsilon^2), \quad (12a) \]

\[ N_c(N) \stackrel{[11]}{=} \frac{1}{G} \frac{G + H}{G^2} \epsilon + O(\epsilon^2). \quad (12b) \]
from (b) estimator implies, according to condition (iii), that are necessary to produce ear relationship between mathematical grounds at the beginning of this section.

Note that (a) is identical with the Nolds’ are located at [24].

For both models considered here, the ‘characteristic thresholds’ are located at (Z) < 1 and N < N_c, allowing us to make use of the ‘linearities’ to construct ideal N_c–estimators. We now derive two near-ideal N_c–estimators, for kinks and coils respectively. These estimators operate on multiple (M) systems with different chain lengths, rather than on a single configuration, and will be denoted as M-coil and M-kink in order to clearly distinguish between S– and M–estimators. Careful empirical tests of the new estimators’ validity is quite essential, and will be given in Sec. [V].

Below, the idea behind the different roles of Eqs. ([6], [9], and [11] is that the statistics of the entanglement network can be expected to be decoupled from the fractal dimension of the atomistic chain, because entanglements arise from inter–chain rather than intra–chain configurational properties. The estimator we develop in the following section will, in fact, potentially be applicable to non-Gaussian chains where ⟨R^2_e⟩ ∝ e^−μ (with 1 ≤ μ ≤ 2), as well as less-flexible polymers (like actin [37] or dendronized polymers [38]) for which N_c is [18] of the order of a “persistence length” of the atomistic chain.

### A. The M–kink estimator

Beyond some a priori unknown chain length N_1, we know that ⟨Z⟩ (as determined via Z1 or CReTA) varies linearly with N, i.e. ⟨Z⟩ = GN + Z_0 (with G > 0, and Z_0 = -(G + H) > −1 in the notation of Eq. [11]). We recall that an ideal N_c–estimator implies, according to condition (iii), that

(i) dN_c(N)/dN = 0 for N ≥ N_1, and

(ii) N_1 < N_c

are necessary to produce N_c = N_c(N_1). Uniquely, N_c = 1/G and N_c(N) = N_c for all N > N_1. Using the linear relationship between ⟨Z⟩ and N we thus propose (a) N_c(N) = N_c(⟨Z⟩ − Z_0), where Z_0 = Z_0(N) is the coefficient determined from data collected up to chain length N. Note that (a) is identical with the N_c–estimator suggested on mathematical grounds at the beginning of this section.

However, N_c(N) = 1/G can be equivalently obtained from (b) N_c(N) = dN/d⟨Z⟩. This is an estimator, denoted as “M-kink”, of extraordinary simplicity:

$$\frac{1}{N_c(N)} = \frac{d⟨Z⟩}{dN}. \quad (13)$$

M-kink is strictly an ideal estimator (i.e. it satisfies all five conditions proposed in Section [III] provided N_1 < N_c. It eliminates the unknown coefficient in the linear relationship, and identifies N_c to be responsible for the ultimate slope of ⟨Z⟩(N). This is analogous with measurements of diffusion coefficients, where one eliminates ballistic and other contributions by taking a derivative. Application of Eq. [13] requires studying more than a single chain length, which renders our M-kink estimator qualitatively different from the S-kink estimators. Data for ⟨Z⟩(N) for both polymer models, shown in Fig. 4 demonstrates that ⟨Z⟩ in fact becomes linear in N for ⟨Z⟩ below unity [24, 39], thus confirming N_1 < N_c. This suggests that N_c can be estimated using data for ⟨Z⟩ from chains of lengths even below N_c.

The occurrence of a nonvanishing N_1 is rooted in the fact that a minimum polymeric contour length (of the order of 2πℓ with polymer thickness ℓ, subsequently corrected by chemical details) is needed for geometrical reasons to form an entanglement (or tight knot) [40]. This length (ℓ) increases with the persistence length of the atomistic contour, and vanishes in the limit of infinitely thin polymers. This implies that determining N_c from the slope we correct for a thickness effect, and N_1 is proportional to the thickness of the atomistic polymer.

### B. The M–coil estimator

Next, we motivate and derive a near–ideal estimator for use with coil–properties ⟨R^2_e⟩ and ⟨L_{pp}⟩ (obtained via PPA, CReTA or Z1). Flory’s characteristic ratio C(N) is defined through the identity [41, 42]

$$⟨R^2_e⟩ ≡ (N − 1)⟨Z⟩^2/C(N). \quad (14)$$

Equation [14] is exact by construction; the N–dependence of C(N) characterizes the (non)-Gaussian structure of chains. In general, C(N) ≥ 1 if N > 1. For (mathematically) ideal chains, C(N) is related to the persistence length [3] of the chain stiffness constant C(∞) = lim_{N→∞} C(N) to be calculated from short chains for any sort of ideal chain, including random walks, freely rotating chains, wormlike chains, etc. Simulations on dense chain packings show [43] that the value of C(∞) = 1.48 is a universal lower limit for

| system          | l_0    | C(∞)   | D  | Y   | A    | B    | G    | H    |
|-----------------|--------|--------|----|-----|------|------|------|------|
| LJ + FENE       | 0.964  | 1.852  | 1.72| 3.55| 0.020| 1.04 | 0.020| 0.12 |
| polyethylene 450 K | 1.54 Å | 8.318  | 19.7 Å^2 | 131.4 Å^2 | 0.22 Å^2 | 8.58 Å^2 | 0.023 | 0.20 |
| polyethylene 400 K | 1.54 Å | 8.535  | 20.2 Å^2 | 85.3 Å^2 | 0.24 Å^2 | 9.37 Å^2 | 0.025 | 0.19 |

### TABLE II: Data obtained via Z1. The coefficients D, Y, A, and B have been obtained from a least square fit to the available data (covering N ≫ N_c) for ⟨R^2_e⟩ and ⟨L_{pp}⟩, according to Eqs. [6] and [9]. Similarly, coefficients G and H derive from the measured ⟨Z⟩ via [11].
excluded volume, flexible chain molecules. For real chains like polyethylene, chains much longer than \( N_p \) need to be studied to characterize \( C(N) \), cf. Ref. [24]. We assume knowledge of \( C(N) \) as function of \( N \) from the atomistic configurations.

To proceed, we make use of our finding that \( \langle L_{pp} \rangle^2 / (N - 1) \) is linear in \( N \) above a certain characteristic \( N_0 \), before \( \langle Z \rangle(N) \) has reached unity, i.e., we assume \( N_0 \leq N_e \) to derive an ideal estimator (15). The linear relationship clearly holds for both polymer models considered here (Ref. [24], Table I Figs. 3, and 9), and has already been formulated in Eq. (9). Next we relate \( \langle L_{pp} \rangle \) and \( \langle Z \rangle \) for large \( N \gg N_e \) by a simple argument: the length of the primitive path, \( L_{pp} \), is the number of “entanglement nodes”, \( N/N_e \), times the mean Euclidean distance \( \ell_e \) between such nodes. This distance \( \ell_e \) equals the mean end–to–end distance of the atomistic chain with \( N_e \) monomers. We thus expect that to a factor of order unity (related to fluctuations in \( \ell_e \)),

\[
\lim_{N \to \infty} \langle L_{pp} \rangle^2 = \left( \frac{N}{N_e} \right)^2 (N_e - 1) \frac{d}{dN} C(N_e).
\]

FIG. 4: \( Z_1 \) results for the two model polymer melts. Testing the applicability of Eq. (11) which predicts linear behavior in this representation (slope \( G_e \), offset \( H_e \)). Clearly \( \langle Z \rangle(N) \) becomes linear at an \( N \) for which \( \langle Z \rangle < 1 \). This implies \( N_1 < N_e \) and that Eq. (13) can be an ideal estimator. An interpretation for \( N_1 \) is given in Sec. [IVA].

Data for larger \( N \) is not shown here, but the slope \( d(Z)/dN \) does not significantly change with increasing \( N \).

By following the procedure of Section [IV A] we arrive at an \( N_e \)-estimator, denoted as “M–coil”, using coil properties alone:

\[
\left( \frac{C(x)}{x} \right)_{x=N_e} = \frac{d}{dN} \left( \frac{\langle L_{pp} \rangle^2}{R_{coil}^2} \right),
\]

where \( R_{coil}^2 = (N - 1) \ell_0^2 \), and \( C(x) \) is the characteristic ratio for a chain with \( N_e \) monomers. This estimator fulfills all conditions from our above definition of an ideal estimator. As for M–kink, the derivative in the M–coil Eq. (15) signals that we have to measure \( \langle L_{pp} \rangle \) as function of \( N \) rather than a single value to estimate \( N_e \). The convergence properties are not as clear \( a \) priori as they are for the M-kink estimator Eq. (13), as this derivation required an approximation. In practice, one must simulate systems with increasing \( N \) until the M–coil converges. There is no apparent way to come up with an \( N_e \)-estimator from coil quantities which converges before \( N \) reaches \( N_e \). This is a noticeable difference between the estimators from coils and kinks (M–kink). Technical considerations in the application of Eq. (15) are discussed in Appendix [B].

V. NUMERICAL RESULTS AND DISCUSSION

The data in Tab. I and a similar set for atomistic polyethylene (configurations from Ref. [24]), will now be used to test the M–estimators. Figure 5 shows results for the M-kink estimator (Eq. (13) and M-coil estimator (Eq. (15) for the same systems analyzed in Figs. I and 2. Comparison of these figures shows that the M–estimators indeed converge faster than the S–estimators (Eqs. (4)–(7)). Moreover, comparison to Fig. 4 shows that the M–estimators converge for marginally entangled systems; values of \( N_e(N) \) appears to be diverging as \( N \to 0 \). The convergence properties are important. For example, \( N_e = 20 \) chains have \( \langle Z \rangle = 0.127 \), and the vast majority have zero entanglements, so the prediction \( N_e(20) = 192 > 20 \) of modified S–kink (6) just signals that we are deep in the unentangled regime, where \( N_e \) cannot yet be estimated.

The fast convergence of the M-kink estimator can be better understood by plugging Eq. (11) into M–kink (13). This produces a special case of the M–kink estimator, which is only asymptotically correct, and can be used when (11) holds. We refer to it as the “approximate M–kink” estimator:

\[
N_e(N) \approx \frac{1}{G},
\]

Here, \( G \) is the coefficient in the linear relationship between \( \langle Z \rangle \) and \( N \) obtained from data collected up to chain length \( N \), and thus \( N_e(N) \) depends on \( N \). Note that the derivative with respect to \( N \) in Eq. (13) removes the \( O(\epsilon) \) errors! This is a major difference with respect to all S–estimators (the estimator used in [9] can be considered as intermediate between S– and M–estimators).

In a similar attempt to rationalize the fast convergence of the M–kink estimator, we insert Eqs. (8) and (9) into Eq. (15). This yields, accordingly, the “approximate M–coil” estimator,

\[
N_e(N) \approx 1 + \frac{D + \sqrt{D^2 - 4A}}{2A}.
\]

Like Eq. (16), Eq. (17) has no \( O(\epsilon) \) corrections. Again, this arises from the “M” approach of taking derivatives with respect to \( N \). In both cases, the use of the derivatives removes undesirable effects related to proper treatment of chain
where $l$ is given by

The simplified M-coil has a simple connection to polymer large enough such that quickly, it cannot be ideal. For the systems under study, $D$ estimators should not be used.

When the variation of the coefficients is large, these three estimators, Eqs. (16)–(18), stems from the variation of $A$, $D$, $G$, and $Y$ with $N$; these coefficients, which are obtained by linear interpolation, must generally be assumed be considered to depend on the available range of studied chain lengths.

The only dependence on $N$ of the approximate and simplified estimators, Eqs. (16)–(18), stems from the variation of $A$, $D$, $G$, and $Y$ with $N$: these coefficients, which are obtained by linear interpolation, must generally be assumed be considered to depend on the available range of studied chain lengths. When the variation of the coefficients is large, these three estimators should not be used.

Note that the simplified M-coil does not agree with M-coil if $C(N_e)$ has not reached $C(\infty)$; though it may converge quickly, it cannot be ideal. For the systems under study, $N_e$ is large enough such that $C(N_e)$ is quite close to $C(\infty)$.

The simplified M-coil has a simple connection to polymer structure and the tube model [1], $D = C(\infty)l_0^2 = l_0l_K$, where $l_K$ is the Kuhn length [35]. The tube diameter $d_T$ is given by $d_T^2 = l_0l_KN_e$, and hence $A = (d_T/N_e)^2$.

Table III quantifies the performance of the new M-estimators. The two presented values for each estimator $N_e(N)$ are the final $N_e$, obtained by analyzing all available chain lengths, together with the value predicted by the estimator at $N = N_e$ (i. e. at the border between unentangled and entangled regimes, using only chains of length up to $\sim N_e$). For an ideal M-estimator these two numbers should be the same within statistical errors, here $\sim 2.5\%$, and $N_e$ needs to be estimated from coil (kink) information (see also Appendix B). The fact that for all these estimators $N_\text{e}(N_e) \approx N_e$ gives sufficient evidence that these are in fact ideal estimators, in sharp contrast to most S-estimators, quantitatively discussed in Tab. IV. Note that the very similar values of $N_e$ reported for LJ+FENE and PE systems are a pure coincidence arising from their similar values of $D/A$ (Tab. [III] cf. Eq. [18]).

| system          | $N_e$ | $N_e(N_e)$ | $N_e$ | $N_e(N_e)$ | $N_e$ | $N_e(N_e)$ | $N_e$ | $N_e(N_e)$ |
|-----------------|------|-----------|------|-----------|------|-----------|------|-----------|
|                 | M-coil | approximate M-coil | simplified M-coil | M-kink | approximate M-kink |
| LJ + FENE       | 86.1 | 87.8      | 85.1 | 89.6      | 86.2 | 90.1      | 48.9 | 46.3      |
| polyethylene 450 K | 84.0 | 83.4      | 84.4 | 84.5      | 90.6 | 90.1      | 44.2 | 42.2      |
| polyethylene 400 K | 82.3 | 80.1      | 80.5 | 77.8      | 83.9 | 84.1      | 41.5 | 38.5      |

TABLE III: Data obtained via Z1. Selected results for $N_e(N)$ for all near-ideal M–coil and M–kink estimators defined in this manuscript. For each estimator, two characteristic values are shown: $N_e$ uses all available $N$ (up to $N = 3500$ and $N = 1000$ for the LJ+FENE and PE models, respectively), and $N_\text{e}(N_e)$ uses only data from short chains with $N \leq N_e$ (cf. Tab. I). Values of $N_\text{e}(N_e)$ are thus obtained at moderate computational cost, and are all in overall agreement with $N_e$. Approximate M–coil (M–kink) results should coincide with M–coil (M–kink) results, if the relationships (8), (9) and (11), respectively, accurately hold. The simplified M–coil does not take into account the effect (see also Appendix B). The fact that for all these estimators $N_\text{e}(N_e) \approx N_e$ gives sufficient evidence that these are in fact ideal estimators, in sharp contrast to most S–estimators, quantitatively discussed in Tab. IV. Note that the very similar values of $N_e$ reported for LJ+FENE and PE systems are a pure coincidence arising from their similar values of $D/A$ (Tab. [III] cf. Eq. [18]).

| system          | $N_e$ | $N_e(N_e)$ | $N_e$ | $N_e(N_e)$ | $N_e$ | $N_e(N_e)$ |
|-----------------|------|-----------|------|-----------|------|-----------|
|                 | classical S-coil | modified S-coil | classical S-kink | modified S-kink |
| LJ + FENE       | 86.1 | 40.0      | 86.1 | 129.7     | 48.9 | 31.7      |
| polyethylene 450 K | 84.0 | 39.7      | 84.0 | 192.8     | 44.2 | 30.4      |
| polyethylene 400 K | 82.3 | 37.3      | 82.3 | 191.8     | 41.5 | 28.8      |

TABLE IV: Data obtained via Z1. For comparison with Tab. III Performance of previous S–coil and S–kink estimators. Accurate $N_e$–values have been overtaken from M–coil and M–kink in Tab. III. Obviously, $N_\text{e}(N_e)$ is far from being close to $N_e$ in all cases, while the deviations are strongest for the $N_\text{e}$–estimates based on coils; the two kink measures seem to at least bracket the true $N_e$ (for the deeper reason that $Z_0$, introduced in Sec. IV.A, must obey $Z_0 \in [-1, 0]$).

ends. The approximate M–coil estimator is related only to the (in general, non-Gaussian) structure of chains and primitive paths. Finally, if the assumptions which lead to Eq. (10) hold, and in order to quantify the contributions to Eq. (17), the above analysis combined with tube-theoretic considerations suggests another estimator, which we refer to as “simplified M–coil”:

$$N_\text{e}(N) \approx \frac{D}{A}. \quad (18)$$
improvement over the method of Ref. [9]. Data is for the same systems analyzed in Figs. 1 and 2. Clearly, \( \langle Z \rangle \) has converged for \( N \ll 100 \), and as shown by comparison to Fig. 4 \( N_\epsilon(N) \) approaches \( N_\epsilon \) before \( (Z) \) exceeds unity. This allows us to estimate \( N_\epsilon \) from mostly unentangled systems. (b) Same data as in (a) vs. \( \log_{10} N \), which allows the full range of \( N \) to be presented. For comparison, blue broken and red dashed lines for PE and LJ+FENE, respectively, show reference data for S–estimators, already presented in Figs. 1 and 2.

FIG. 5: (color online) (a) Performance of proposed estimators M–kink [23] (lower two curves with large symbols) and M–coil [18] (upper two curves with large symbols); see also Appendix B. (b) Same data as in (a) vs. \( \log_{10} N \), which allows the full range of \( N \) to be presented. For comparison, blue broken and red dashed lines for PE and LJ+FENE, respectively, show reference data for S–estimators, already presented in Figs. 1 and 2.

suggest that the “best” estimate of the entanglement length for flexible chains is well above the previously reported value. This is significant e. g. for calculating the ratio \( N_e/N_c \), where \( N_c \) is the rheological crossover chain length where zero shear viscosity changes its scaling behavior from Kouse to reptation, and has been estimated as \( N_c \approx 100 [13, 14] \).

One could imagine fitting the squared contour length \( \langle L_{pp}^2(n) \rangle \) of primitive path subsections [46] to \( \langle L_{pp}^2(n) \rangle = A n^2 + C n \), and attempting to calculate \( N_e(N) = D/A \) by also fitting to \( \langle R^2(n) \rangle = D n - Y \), or developing other improved estimators for \( N_e \) based on \( \langle R_{pp}^2(n) \rangle \) and \( \langle L_{pp}^2(n) \rangle \). However, analysis along these lines failed to produce any estimators better than those described above. In particular, no improvement over the method of Ref. [9] was found.

It is important to notice that our Eq. (13) is not compatible with some earlier definitions of \( \langle Z \rangle \) from coil quantities, because of the prefactor \( C(\infty)/C(N_c) \). This prefactor had usually been omitted or not mentioned, since random walk statistics were clearly a convincing starting point. Assuming Gaussian statistics (constant \( C(N) \) for all \( N \)) hence underestimates values of \( N_e \) calculated from coil properties. This issue is also one of the reasons why the \( N_e \) estimates between PPA and geometrical approaches differ. Another reason is given in [13].

VI. CONCLUSIONS

Very significantly improved, near–ideal, and apparently polymer–model–independent estimators for \( N_e \) were derived in this paper, M–coil (Eq. (15) to be used with PPA, Z1 or CreTA) and M–kink (Eq. (13) Z1 and CreTA only). They reduce, under further assumptions which seem valid for the model systems studied here, to approximate M–coil (Eq. (17), simplified M–coil (Eq. (18), and approximate M–kink (Eq. (16)). These estimators require simulation of multiple chain lengths, but have eliminated systematic \( O(\epsilon) \) errors present in previous methods. This is important for the design of efficient simulation methods in the field of multiscale modeling of polymer melts.

Furthermore, we have proposed variants of the original estimators. The two main problems with existing estimators were identified as: i) improper treatment of chain ends, and ii) non–treatment of the non-Gaussian statistics of chains and primitive paths [36]. Improper handling of thermal fluctuations was an additional problem relevant to very short chains. Issues i) and ii) lead to separate, independent \( O(\epsilon) \) errors. Estimators based on direct enumeration of entanglements lack issue ii), and so are fundamentally advantageous for estimation of \( N_e \). The new “M” estimators proposed here formally correct for the errors arising from effects i) and ii). The values of the M–coil and M–kink–estimators can be taken as “best estimates” for \( N_e \) when results are available for multiple chain lengths. The best estimator when only a single chain length is available is the modified S–kink, Eq. (6).

We have shown that \( \epsilon \langle L_{pp}^2 \rangle \), \( \langle Z \rangle \), and also \( \langle R_{pp}^2 \rangle \) are all linear in \( 1/\epsilon \) (thus linear in \( N \)) down to the mostly unentangled regime, and have used this information to derive the M–estimators and to improve the earlier ones. All coefficients in these linear relationships have been evaluated and listed in Tab. III. The prefactors for the above mentioned \( O(\epsilon) \) errors can be large, and depend both on the polymer model and method of topological analysis. These errors can produce
large changes in estimates of $N_c$ for values of $N$ typically considered in previous studies (e.g. Refs. [4] [12] [33]). This is significant in light of attempts to compare PPA results for $N_c$ to values obtained by other methods [6] [23] [33] [47] such as direct rheological measurement of the plateau modulus $C \langle q^2, t \rangle$, evolution of the time-dependent structure factor $S(q, t)$, and estimation of the disentanglement time $\tau_d \propto (N/N_c)^3$ [1]. Some conclusions of those studies may need to be reevaluated in light of the new data.

The proposed M-estimators are to our knowledge the first estimators which exhibit all features required for an ideal estimator (a term which we made precise in Sec. II), and they have been physically motivated. They converge to $N_c$ for weakly entangled systems $(N \leq N_c)$. They leave $N_c$ either undefined or infinite for rodlike chains (because $C(N) = N$ for a rod). They predict $N_c(N) \geq N$ for a completely unentangled system, which is characterized by $\langle Z \rangle = 0$ and $L_{pp} = R_{ee}$ in accord with the definition of the primitive path which we have adopted in this work (see [29]). The appearance of the coefficient $N_1$ suggests that there might be a minimum amount of material, $N_1$, needed to form a single entanglement (as observed for phantom chains [9]). If so, it can be expected to depend on the thickness of the atomic chain and its stiffness as well as particle density. We expect our findings to be universal in the sense that they should apply to all sorts of real linear polymer chains in the melt state, and we have verified the assumptions underlying the M-estimators by direct comparison with both atomistic semiflexible and coarse-grained flexible polymer melts.

Refs. [11] [15] pointed out that primitive paths are not random walks, and that there appears to be more than one “topological” entanglement per “rheological” entanglement; thus it is unsurprising that $N_c$ from coils is significantly larger than $N_c$ from kinks (for details see Ref. [18]). The utility of any topological analysis of chains shorter than $N_c$ remains highly questionable, because the chains’ dynamics are well described by the Rouse model [1] [19] and so they cannot be considered “fully entangled” in any meaningful way. However, it seems that the M-estimators developed in this work have the ability to extract information from a partial or even marginal degree of entanglement.

The $M$-estimators could be applied in a post-processing step on existing configurations. For example, it should be of interest to study the effect of flow and deformation on entanglement network characteristics in order to establish equations of motion for relevant coarse-grained variables characterizing the polymer melt. Shear and elongational flows have been studied for both polymer models considered here, but either Z1 was not yet available at the time of these studies [48], or the chains were [49] [50] “too short”, i.e. had $\langle Z \rangle \ll 1$.

The apparent ability to accurately estimate $N_c$, even for weakly entangled systems may be useful for atomistic models whose computational cost prohibits equilibrating large-$N$ systems, such as polymers containing bulky side groups. The procedure for removal of the $O(\epsilon)$ systematic errors, while clearly described here, requires performing analyses on a limited number of configurations on a range of chain lengths, which is most easily undertaken for systems composed of “short, but not too short” chains.

VII. ACKNOWLEDGEMENTS

R. H. thanks Alexei E. Likhtman for pointing out that the non-Gaussian statistics of chains and primitive paths produce $O(1/N)$ systematic errors in the old estimators for $N_c$. Steven J. Plimpton integrated the DBH algorithm into LAMMPS [22]. Gary S. Grest and Nikos Karayiannis provided helpful discussions. Gary also provided an equilibrated $N = 3500$ state, and Nikos was deeply involved in all PE developments. This work was supported by the MRSEC Program of the National Science Foundation under Award No. DMR05-20415, as well as through EU-NSF contract NMP3-CT-2005-016375 and FP6-2004-NMP-TI-4 STRP 033339 of the European Community. All atomistic simulations were conducted in the “magerit” supercomputer of CeSViMa (UPM, Spain).

APPENDIX A: TREATMENT OF THERMAL FLUCTUATIONS

Ref. [4] and other studies have typically used $\langle L_{pp} \rangle^2$ rather than $\langle L_{pp}^2 \rangle$ in estimators for $N_c$, such as the analogue for the modified S-coil [7] which reads

$$N_c(N) = (N - 1) \left( \frac{\langle L_{pp} \rangle^2}{\langle L_{pp}^2 \rangle} - 1 \right)^{-1}.$$  \hspace{1cm} (A1)

However, Eq. (A1) gives pathological results for short chains due to existing thermal fluctuations of $L_{pp}$. Consider the unentangled limit, where the entanglement density (denoted as $\rho_e$) vanishes. For an “ideal” topological analysis, $L_{pp} \to R_{ee}$ (from above) for each and every chain as $\rho_e \to 0$. However, chain dimensions fluctuate in thermodynamic equilibrium [1]. To leading order in the fluctuations, $\langle L_{pp} \rangle^2 = \langle L_{pp} \rangle^2 - \langle \Delta L_{pp} \rangle^2 = \langle R_{ee}^2 \rangle - \langle \Delta R_{ee} \rangle^2$, where $\Delta$ is “variance of”. So, even for an ideal topological analysis procedure, Eq. (A1) would predict a negative $N_c(N) \to -N^{-1} \langle R_{ee}^2 \rangle \langle \Delta R_{ee} \rangle^2$ as $\rho_e \to 0$. Negative $N_c(N)$ are of course useless, but indeed, are predicted using our data in Tab. 1. For $N = 20$ (LJ+FENE melt), application of Eq. (A1) yields negative $N_c(20)$. A term identical to the term in parenthesis in Eq. (A1) was found to be negative for short chains in Ref. [49], but was not used to directly calculate $N_c(N)$ in their work, as its negative value was considered to signal (and to only occur in) the mostly unentangled regime.

The reason to fix chain ends during PPA or Z1 analysis is the assumption, implicit in Edwards’ definition of the primitive path [4], that chains are entangled. In this context it is worthwhile mentioning that there are other definitions of PP’s, for example one [51] where the length of the PP goes down to zero for the unentangled chain, and where chain ends are not fixed.
APPENDIX B: TECHNICAL CONSIDERATIONS IN USE OF THE M–COIL ESTIMATOR

FIG. 6: The graph demonstrates how to graphically evaluate $N_e(N)$ according to the M–coil estimator [15]. Shown are both the left hand side (lhs) $C(N)/N$, and right hand side (rhs) of Eq. (15) for both types of polymer melts. The dotted red path makes an example on how to obtain $N_e(N) \approx 87$ for given $N = 48$. Obviously, this value is quite identical with both, $N_e$ and $N_e(N_e)$, cf. Tab. III. The ratio $C(N)/N$ (small points) monotonically decreases with increasing $N$, while the rhs (large symbols) reaches a plateau at the time $N$ has approached $N_e$ (at the crosspoint), which is a distinguished feature of an ideal estimator.

While the M–kink estimator (Eq. [13]) is explicitly evaluated from the local derivative $d\langle Z \rangle/dN$ around $N$, our M–coil expression, Eq. [14], is only an implicit expression for the estimator $N_e(N)$. Formally, we need the inverse of $C(N)/N$ to calculate $N_e(N)$. In the following, we describe the procedure in order to prevent any ambiguities upon applying M–coil in practice. Fig. 6 shows both the left (lhs) and right hand (rhs) sides of Eq. [13] versus $N$ for our data. For any given $N$ (say, $N = 48$ for the PE data, where the dotted red line starts in Fig. 6), the $N_e(N)$ estimate is the value at the ordinate for which the abscissa values for lhs and rhs coincide (end of the red curve is at $N_e(48) \approx 87$). The same procedure is repeated for all $N$ to arrive at Fig. 5 and particular values collected in the M–coil row of Tab. III. The difference between lhs and rhs can be used to estimate the difference between the largest $N$ available and $N_e$. If only short chains had been studied, only a part of this plot could have been drawn.

Note that this procedure requires $C(N)/N$ to be monotonically decreasing with $N$, and access to $C(N)$ at sufficiently large $N$. While the former is essentially valid for all polymer models, the latter may pose a problem. Without reliable values for $C(N)$ for $N = N_e$, there is no apparent way to come up with an M–coil which converges before $N$ reaches $N_e$. However, since $C(N)/N$ decreases with increasing $N$ and ultimately reaches $C(\infty)/N$ behavior, in practice (and formally for ideal chains) $C(N)$ can be estimated by extrapolation, and the necessary $C(N)/N$ values could be added for chain lengths exceeding those studied.

This issue disappears by construction when the largest simulated $N$ exceeds $N_e(N)$, so that the conditions for an ideal estimator are met in any case. Still, this is a noticeable and principal difference between the estimators from coils (M–coil) and kinks (M–kink).

[1] M. Doi and S.F. Edwards, *The theory of polymer dynamics* (Clarendon Press, Oxford, NY, 1986).
[2] L.J. Fetters, D.J. Lohse, D. Richter, T.A. Witten, and A. Zirkel, Macromolecules 27, 4639 (1994); L.J. Fetters, D.J. Lohse, S.T. Milner, and W.W. Graessley, Macromolecules 32, 6847 (1999).
[3] M. Kröger and S. Hess, Phys. Rev. Lett. 85, 1128 (2000).
[4] S.F. Edwards, Br. Polym. J. 9, 140 (1977).
[5] M. Rubinstein and E. Helfand, J. Chem. Phys. 82, 2477 (1985).
[6] R. Everaers, S.K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer, Science 303, 823 (2004).
[7] M. Kröger, Comput. Phys. Commun. 168, 209 (2005).
[8] Q. Zhou and R.G. Larson, Macromolecules 38, 5761 (2005).
[9] S.K. Sukumaran, G.S. Grest, K. Kremer, and R. Everaers, J. Polym. Sci. Part B-Polym. Phys. 43, 917 (2005).
[10] K. Foteinopoulou et al., Macromolecules 39, 4207 (2006).
[11] C. Tzoumanekas and D.N. Theodorou, Macromolecules 39, 4592 (2006).
[12] S. Shanbhag and R.G. Larson, Macromolecules 39, 2413 (2006).
[13] K. Foteinopoulou et al. Phys. Rev. Lett. 101, 265702 (2008); M. Laso et al. Soft Matter 5, 1762 (2009).
[14] S. Shanbhag and M. Kröger, Macromolecules 40, 2897 (2007).
[15] C. Tzoumanekas and D.N. Theodorou, Curr. Opin. Solid State Mater. Sci. 10, 61 (2006).
[16] M. Kröger, *Models for polymeric and anisotropic liquids* (Springer, Berlin, 2005).
[17] D. Curcó and C. Alemán, Chem. Phys. Lett. 436, 189 (2007); J. Comput. Chem. 28, 1929 (2007); J.D. Schieber, D.M. Nair, and T.J. Kitkraitard, J. Rheol. 51, 1111 (2007); J.D. Schieber, J. Chem. Phys. 118, 5162 (2003).
[18] N. Uchida, G.S. Grest, and R. Everaers, J. Chem. Phys. 128, 044902 (2008).
[19] K. Kremer and G.S. Grest, J. Phys. Chem. 92, 5057 (1990).
[20] R. Auhl, R. Everaers, G.S. Grest, K. Kremer, and S. J. Plimpton, J. Chem. Phys. 119, 12718 (2003).
[21] N.C. Karayiannis, V.G. Mavrantzas, and D.N. Theodorou, Phys. Rev. Lett. 88, 105503 (2002).
[22] http://lammps.sandia.gov/
[23] M. Pütz, K. Kremer, and G.S. Grest, Europhys. Lett. 49, 735 (2000).
[24] K. Foteinopoulou, N.C. Karayiannis, M. Laso, and M. Kröger, J. Phys. Chem. B 113, 442 (2009).
[25] N.C. Karayiannis, A.E. Giannoussaki, V.G. Mavrantzas, and D.N. Theodorou, J. Chem. Phys. 117, 5465 (2002).
[26] P.V.K. Pant and D.N. Theodorou, Macromolecules 28, 7224.
For the case of wormlike (non-gaussian) chains, S. Leon, N. van der Vegt, L. Delle Site, and K. Kremer, Macromolecules 40, 8389 (2007).

In Ref. [6], the primitive path had been defined as the path with minimum elastic energy, rather than the one with shortest length. Larson et al. [52] discussed this topic and concluded that methods producing the shortest length paths were preferable. The geometric approaches Z1 [14] and CReTA [11] minimize total contour length rather than an elastic energy, and provide us with a shortest, parameter-free path.

In Ref. [24], reported \( N_c > 40 \) and that \( \langle Z \rangle \) to become linear at \( N > 24 \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

A counterexample would be LJ+FENE bead spring chains made semiflexible by the addition of a bending potential \( U = -k_0 \cos(\theta) \); when \( k_0 = 2 \epsilon_1 \), \( C(N) \) is still well below \( C(\infty) \).

Unfortunately, this approach can not be used with standard PPA, because the tension in the primitive chains produces \( \langle L_{pp}(n) \rangle \approx b_n^2 \pi^2 \). In contrast, Z1 and similar methods (e.g. [11]) produce a wide range of local fractional contour length reductions.

A.E. Likhtman, private communication. See also Refs. [7, 11].

A.E. Likhtman, S.K. Sukumaran, and J. Ramirez, Macromolecules 40, 710 (2007).

For the case of wormlike (non-gaussian) chains, \( C(N) \) is analytically related to the persistence length \( l_p \) of short chains; \( C(N+1) = C(\infty) - 2\alpha(1-\alpha^N)/N \), where \( \alpha = \exp(-l_0/l_p) \), and \( C(\infty) = \lim_{N \to \infty} C(N) = (\alpha + 1)/(\alpha - 1) \). Note that \( \lim_{N \to \infty} C(N) = N \) and \( \lim_{N \to 0} C(N) = 1 \).

In the limit \( N \gg l_0 \), the mean square end-to-end distance of a wormlike chain consequently reads \( \langle R_2^2 \rangle = 2l_0L(1-\langle -L/l_0 \rangle) \) with contour length \( L \equiv (N-1)l_0 \). In the further limit \( L \gg l_p \), \( \langle R_2^2 \rangle/(N-1) \approx 2l_0l_k = C(\infty)l_k^2 \equiv l_Kl_k \), where \( l_K \) denotes Kuhn length. These limits are not employed in this work, as we are seeking for corrections in the regime where these limits have not necessarily been reached. For freely rotating chains with a fixed bending angle \( \theta \) the same result is obtained upon identifying \( \alpha = \langle \cos \theta \rangle \). The above expressions for \( C(N) \) are directly obtained from a bond-vector correlation function which exponentially decreases with distance between bonds. However, simulations have shown [20] that flexible bead spring chains in melts do not follow this form at small and moderate \( N \). The ratio \( C(N) \) must sublinearly and monotonically increase with \( N \) in order to uniquely determine \( N_c \) from [15], which limits its use to systems which do not change their statistics on a length scale larger than entanglement length. For the wormlike chain as well as equilibrated bead change spring chains [20], \( C(N)/N \) decreases monotonically with \( N \).

A.E. Likhtman, private communication. See also Refs. [7, 11].

J. Kas et al., Biophys. J. 70, 609 (1996); R. Granek, J. Phys. II (France) 7, 1761 (1997); S. Ramanathan and D.C. Morse, Phys. Rev. E 76, 010501(R) (2007).

A.D. Schlüter, Top. Curr. Chem. 245, 151 (2005); D.K. Christopoulos, A.F. Terzis, A.G. Vanakaras, and D.J. Photinos, J. Chem. Phys. 125, 204907 (2006); Y. Ding et al., J. Chem. Phys. 127, 094904 (2007). M. Kröger, O. Peleg, Y. Ding, and Y. Rabin, Soft Matter 5, 1762 (2009).

Ref. [24] reported \( N_c > 40 \) and that \( \langle Z \rangle \) to become linear at \( N > 24 \).

G. Buck and E.J. Rawdon, Phys. Rev. E 70, 011803 (2004).

P.J. Flory, Statistical mechanics of chain molecules (Wiley, New York, 1969).

M. Rubinstein and R.H. Colby, Polymer Physics (Oxford University Press, New York, 2003) p. 53.

N.C. Karayiannis and M. Laso, Phys. Rev. Lett. 100, 050602 (2008); N.C. Karayiannis, K. Foteinopoulou, and M. Laso, J. Chem. Phys. 130, 164908 (2009).

A.E. Likhtman, private communication. See also Refs. [7, 11].

J. Kas et al., Biophys. J. 70, 609 (1996); R. Granek, J. Phys. II (France) 7, 1761 (1997); S. Ramanathan and D.C. Morse, Phys. Rev. E 76, 010501(R) (2007).

R.G. Larson, Q. Zhou, S. Shanbhag, and S. Park, J. AIChE J. 53, 542 (2007).