Accurate computation of shear viscosity from equilibrium molecular dynamics simulations

D. NEVINS*† and F. J. SPERA†

†Department of Earth Science, University of California, Santa Barbara, CA 93106, USA
*dean@nevins.org

The accuracy of the Green-Kubo formulation for computing shear viscosity from equilibrium Molecular Dynamics simulations depends on the quality of the potential used to model the material and on how the computation is carried out. We examine the role of the duration of the simulation, the number of particles used, and how the correlations are accumulated on the accuracy of the shear viscosity. We propose as a measure of the accuracy of the computed shear viscosity the standard deviation of five independently computed viscosity values based on components of the stress tensor. Using this measure, we examine the shear viscosity calculation for NaCl to determine the values of the run length, window width, and spacing between windows which are a good compromise between calculation time and viscosity quality. Significantly we note that even though viscosity can be calculated using relatively few particles, reducing state point uncertainty requires more, rather than less, particles.

Keywords: Viscosity; Molecular dynamics; NaCl; Green-Kubo

1. Introduction

The Green-Kubo (GK) expression for the shear viscosity is given by integration of the stress (pressure) autocorrelation function. In particular, the shear viscosity is computed according to

$$\eta = \frac{V}{3k_B T}\int_0^\infty \left\langle \sum_{x<y} P_{xy}(t)P_{xy}(0) \right\rangle dt \qquad (1)$$

where η is the shear viscosity, V is the volume of the system, T is the temperature, k_B is Boltzmann's constant, and P_{xy} refers to an independent component of the stress in the xy direction. The GK formulation utilizes a single summation that consolidates the contributions of all the atoms into a single autocorrelation function

$$C_{xy}(t) = \left\langle \sum_{x<y} P_{xy}(t)P_{xy}(0) \right\rangle \qquad (2)$$

This allows the formulation to be used with molecular dynamics simulations whereas alternative formulations based on particle displacement require translational invariance, an assumption violated in MD simulations utilizing periodic boundary conditions. The angle brackets around the summation in equation (2) refer to an average of a 'sufficiently large' number of samples [1]. The other independent estimates of the shear viscosity that are used are the remaining off-diagonal components of the stress .P_{xz} and P_{yz}, as well as two

[Insert figure1 about here]

forms of the first normal difference [2], P_{xx} - P_{yy} and P_{yy} - P_{zz}. To study the optimal way to implement equation (1), it is useful to refer to specific quantities illustrated schematically in Figure 1. These quantities include the duration of the MD simulation, t_D,

the time 'window' over which the autocorrelation C (t) is computed, t_W, and the time interval between the start of successive windows, t_S. These quantities are related to each other and the number of time origins, n_O, used in the summation according to

$$n_O = 1 + \left[\frac{t_D - t_W}{t_S} \right] \tag{3}$$

where the operation implied by the square brackets returns the integer part of the quotient. Since for GK computation of viscosity, t_D is of order a few ns, and t_W is a few ps, n_O reduces to

$$n_O = n_W \approx \left[\frac{t_D}{t_S} \right] \tag{4}$$

In this investigation we determine the dependence of the shear viscosity on t_D, t_S and t_W and N, the number of particles used in the simulation, in an effort to deduce the most efficient and accurate method for determining the shear viscosity by the GK method from equilibrium MD simulations. In order to gauge the effects of simulation duration, window width and number of origins on the precision of the viscosity, a measure of the error of the viscosity is introduced. We exploit the fact that *each* of the five independent components of the stress (i.e., P_{xy}, P_{xz},P_{yz}, P_{xx} - P_{yy}, and P_{yy} - P_{zz}) provides an independent estimate of the shear viscosity. We define the fractional error, ξ, as the root mean square of the 'component' deviations from the average shear viscosity, η_{AVE}, divided by the average viscosity according to

$$\xi = \frac{1}{\eta_{AVE}} \sqrt{\frac{(\eta_{XY} - \eta_{AVE})^2 + (\eta_{XZ} - \eta_{AVE})^2 + (\eta_{YZ} - \eta_{AVE})^2 + (\eta_{XXYY} - \eta_{AVE})^2 + (\eta_{YYZZ} - \eta_{AVE})^2}{5}} \tag{5}$$

where η_{AVE} is the arithmetic mean of the five independently determined viscosity estimates η_{xy}, η_{xz} , η_{yz}. η_{xxyy}, and η_{yyzz}.

The accuracy of the shear viscosity in the sense of comparison with laboratory values depends obviously on the quality of the potential. Here we study molten NaCl because it is a simple material for which a reasonably accurate effective pair potential exists. Although the point of this investigation is not to find a better description of the pair potential applicable to NaCl, it is informative to compare MD computed values with laboratory data[3]. The methodology developed in this study is directly applicable to determination of shear viscosity in other materials including molten geoliquids at conditions of elevated temperature (2000- 5000 K) and pressure (0-135 GPa) relevant to geophysical studies of the Earth's partially molten interior.

2. Model and simulation parameters

For this investigation, NaCl was simulated using the potential form

$$\phi_{ij}(r_{ij}) = \frac{q_i q_j}{r_{ij}} + A_{ij}\exp\left(\frac{-r_{ij}}{B_{ij}}\right) - \frac{C_{ij}}{r^6} - \frac{D_{ij}}{r^8} \qquad (6)$$

This form has been shown to describe the alkali halides quite well [4, 5, 6]. The numerical values of the parameters used are given in Table 1. The MD code utilized was a modified version of the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [7]. The pair potential has contributions from Coulombic forces, Born-

[Insert table1 about here]

Huggins exponential electron repulsion, dipolar and quadrupolar terms. Long-range Coulomb terms were computed using the Particle-Particle Particle-Mesh K-space solver with a precision of one part in 10 000. Short-range Coulombic interactions, Born-

Huggins repulsions, and dipolar and quadrupolar forces are calculated directly within a 0.6 nm (6 Å) radial cutoff in direct space.

The simulations were carried out in the NEV microcanonical ensemble with the numbers of particles ranging from 258 to 25800. Ion positions are updated using the leapfrog Verlet scheme with a timestep of 1 femtosecond (fs). Initial conditions were developed using the skew-start methodology [8] with an initial temperature distribution of 5000 K. Upon removal of net momentum, temperature was held to 5000 K by velocity scaling for 10 ps. The system was cooled from 5000 K to the target temperature of 1400 K using the "slow" cooling schedule (0.5×10^{13} K s^{-1}) adopted by Matsui and Kawamura [9]. The liquid was then held at 1400 K for 150 ps to achieve thermal equilibrium by velocity scaling. At thermal equilibrium, velocity scaling was turned off and production runs of duration t_D 2 to 10 ns were carried out. Potential and kinetic energy, temperature, isochoric heat capacity and the stress components P_{xx}, P_{yy}, P_{zz}, P_{xy}, P_{xz} and P_{yz}. were computed and saved every timestep.

For N = 25 800 particles (12 900 NaCl 'molecules') at 1379.3 kg m^{-3} and using t_D = 2 ns, t_W = 2 ps, and t_S of 10 fs, the simulated temperature was 1410 ± 19 K. Pressure was 197.6 ± 21.0 MPa and the viscosity was computed to be 0.692 mPa s. This compares favourably with the laboratory results for NaCl summarized extensively by Janz [3] who reports a density of 1377.6 kg m^{-3} and a viscosity of 0.614 mPa s at 1400 K and 0.1 MPa. Our higher computed density is consistent with the higher pressure of the MD simulation compared to the laboratory measurement.

3. Results and discussion

1. Selection of window width (t_W)

When computing the autocorrelation function C(t), t_W (Figure 1) must be chosen. t_W should be long enough to capture the decay of C(t) in its entirety but not so long that noise is added to the correlation signal by the correlation value doing a random walk away from the true value when the noise of the correlation function approaches the intrinsic value. Figure 2 shows a typical autocorrelation C(t) for [Insert figure2 about here]

the off-diagonal pressure P_{xy}. C(t) decays rapidly towards zero and after about 1 ps, C(t) exhibits small amplitude oscillations around zero. The rate of the descent toward zero differs for each of the off-trace pressure C (t) and for each run. We have chosen to cut off the contribution of C(t) at the time step where the slowest decaying C(t) functions falls below 0.5×10^{-2}. In practice, this value varied from about 0.56 ps to 1.35 ps. In order to have a uniform window width, a conservative value of 2 ps seconds is adopted and used throughout the rest of this study.

2. Role of simulation duration (t_D)

To examine the effects of changing run duration t_D on computed viscosity, the window width (t_W) and spacing (t_S) were set equal to 2 ps and 0.01 ps, respectively. The number of windows and the simulation duration are linked through equation (2) and are not independent at fixed t_S. Figure 3 shows the variation of ξ, the fractional error defined by equation (5) plotted against t_D. The average deviation in the t_D interval 20-100 ps is ≈ 25% whereas for the longest simulation duration (10 ns), the deviation is < 5%. Simulations of durations of less than ~1.0 ns can have significant changes in the value of the viscosity calculated from each autocorrelation and are consequently unsuitable for

[Insert figure3 about here]

the determination of viscosity. Alternatively, not much additional reduction in ξ occurs for durations greater than about 2 ns, although the computation time increases considerably. The effect of the run duration, t_D, is also apparent in the value of the average viscosity. As shown in the inset to Figure 3, the average viscosity computed for the short duration t_D=0.02 ns run does not approach an asymptotic limit like the viscosity value computed for t_D=10 ns. In summary, the combination of a stable η_{AVE} combined with small ξ suggests a run length of ~2 ns is a good compromise between accuracy and computational cost. In terms of n_O, the number of origins, t_D = 2 ns corresponds to 200 000 time origins (windows) when the interval between successive windows (t_S) is 0.01 ps.

3. Role of temporal spacing between time origins (t_S)

In order to test the effect of changing t_S, the interval between successive time origins on the fractional error ξ, a series of simulations was carried out at constant t_D=10 ns. The window spacing t_S was varied from 1fs to 100 ps. From equation (4) it is clear that at const t_D, the number of time origins (n_O) is not independent of the window spacing, t_S.

[Insert figure4 about here]

Figure 4 shows that to obtain ξ < 2% the window spacing must be t_S smaller than ~100 fs. Smaller values do not lead to any reduction in ξ. A value of t_S = 10 fs is a good choice to maintain a small error, ξ, while minimizing computational time.

4. Role of system size (particle number)

It has been claimed [10] that increasing the numbers of particles (N) has no effect on improving the quality of the Green-Kubo shear viscosity calculation. We have tested this

by studying the effect of varying N at fixed values of t_D (10 ns), t_W (2 ps) and t_S (10 fs). The relationship between fractional error, ξ and particle number is depicted in Figure 5.

[Insert figure5 about here]

Indeed, there is no correlation between N and ξ. N varies by a factor of 100 from N=258 to 25 800 and there is no systematic improvement (decrease) in ξ. The same effect is seen for the average viscosity; increasing N does give rise to asymptotic behaviour in η_{AVE}. Although tempting to conclude that increasing N has little effect on the computed viscosity value, this conclusion is not warranted. The reason is the as follows. It is well known [11] that the fluctuations in pressure (σ_P) and temperature (σ_T) during a MD

[Insert figure6 about here]

simulation (microcanonical ensemble) scale as $N^{-1/2}$. This is shown in Figure 6 by plotting $\sigma_T/\hat{T}$, the temperature fluctuation (σ_T) divided by the mean temperature $\hat{T}$ of the run, and the analogous quantity for pressure, $\sigma_P/\hat{P}$, against N. It is clear from Figure 6 that the uncertainty in temperature and pressure of an MD simulation depends on N. For example, for N = 500 the temperature and pressure of the simulation are known to within 3 % and 50 %, respectively. For a simulation carried out at a typical geophysically relevant state point of, say, 3500 K and 10 GPa, the uncertainties in temperature and pressure are ±100 K and ±4.8 GPa, respectively. In contrast, a simulation with N = 10 000 at the same state point carries with it an error in temperature and pressure of ±45 K and ±1 GPa, respectively. Because there is a one-to-one mapping between a state point and the shear viscosity, simulations run with large N are clearly superior than those with small N since the uncertainties of the state point are smaller in the former than in the latter case.

5. Conclusions

Systematic study of implementation of the GK method for determination of the shear viscosity reveals the parameters that are most cost-effective. Because each independent component of the stress tensor enable one to compute an estimate of the shear viscosity, and because these values must, in a real fluid, converge to a single value, one may develop a criterion of accuracy based on the convergence of the estimates from the three components. Using this metric, we then studied the effect of simulation duration, t_D, width of the correlation window, t_W and the temporal spacing between time origins, t_S. We find that a set of optimal values that effectively trade off accuracy with computational cost are $t_D = 2$ns, $t_W = 2$ps and $t_S = 10$ fs. Although ξ does not depend on N, the number of particles in the simulation, the fluctuations in temperature and pressure of the run, and hence the uncertainty of the state scale as $N^{-1/2}$. Hence, a viscosity computed by the GK method cannot be reliably placed in the context of material behaviour unless the state point is known well. This requires a relatively large number of particles.

6. Acknowledgments

This research was supported by an allocation of advanced computing resources supported by the National Science Foundation and the Office of Science of the U.S. Department of Energy. The computations were performed in part on DataStar at the San Diego Supercomputer Center and Seaborg at the National Energy Research Scientific Computing Center. We also thank Nuno Galamba, Toshiro Tanimoto Mark Ghiroso, and an anonymous reviewer for useful discussions on various parts of this work. We acknowledge support from NSF grants EAR-0440057 and ATM-0425059 and DOE grant DE-FG-03-91ER-14211 to FJS.

7. References

[1] D.C. Rapaport. *The art of molecular dynamics simulation*, pp. 116-117, Cambridge University Press, Cambridge (1995).

[2] Alfe, Dario, and Gillan, Michael, (1998) First-Principles Calculation of Transport Coefficients, Physical Review Letters, pp. 5161- 5164.

[3] G. J. Janz, Molten salts data as reference standards for density, surface tension, viscosity and electrical conductance: KNO_3 and NaCl, *J. Phys. Chem. Ref. Data*, **9** No. 4 (1980)

[4] N. Galamba, C. A. Nieto de Castro, and J. F. Ely. Molecular dynamics simulation of the shear viscosity of molten alkali halides. *Journal of Physical Chemistry*, **108**, 3658 (2004).

[5] F. G. Fumi, M. P. Tosi. Ionic sizes and born repulsive parameters in the NaCl-type alkali halides-I. *J. Phys. Chem. Solids*, **25**, 31 (1964).

[6] J. W. E. Lewis, K. Singer, and L. V. Woodcock. Thermodynamic and structural properties of liquid ionic salts obtained by Monte Carlo computation. Part 2. Eight alkali metal halides. *J. Chem. Soc., Faraday Trans. 2* **71**, 301 (1975).

[7] S. J. Plimpton. Fast parallel algorithms for short-range molecular dynamics, *J Comp Phys*, **117**, 1 (1995).

[8] K. Refson. *Moldy user's manual*, Oxford Department of Earth Sciences, Oxford (2001).

[9] Y. Matsui, K. Kawamura, and Y. Syono. Molecular dynamics calculations applied to silicate systems: molten and vitreous $MgSiO_3$ and Mg_2SiO_4 under low and high pressures. In *High pressure research in geophysics: Advances in Earth and planetary sciences*. S. Akimoto, and M.H. Manghnani (Eds.), pp. 511-24, Reidel , Boston, Massachusetts (1982).

[10] A. Dario and M. J. Gillan, First-principles calculation of transport coefficients, *Physical Review Letters*, 81, no. 23, 1998.

[11] D. A. McQuarrie, *Statistical Mechanics*, Harper Collins, New York (1976).

Figure Captions

Figure 1: Schematic diagram illustrating the relationship between the time intervals used to compute the shear viscosity from MD simulation data using the Green-Kubo formulation. t_D refers to the total duration of the simulation. The autocorrelation functions for the stress components P_{xy}, P_{xz},P_{yz}, P_{xx} - P_{yy}, and P_{yy} - P_{zz} are each partitioned into multiple windows of duration t_W and have their origins separated from one another by t_S. The number of time origins (equal to the number of windows) is given exactly by equation (3) and approximately by equation (4).

Figure 2: Off-diagonal stress autocorrelation function $C_{xy}(t)$ *versus* time for N = 2580 particles at 1418 K (±18 K), 202.9 MPa (±46.6 MPa) and density of 1379.3 kg m^{-3}.

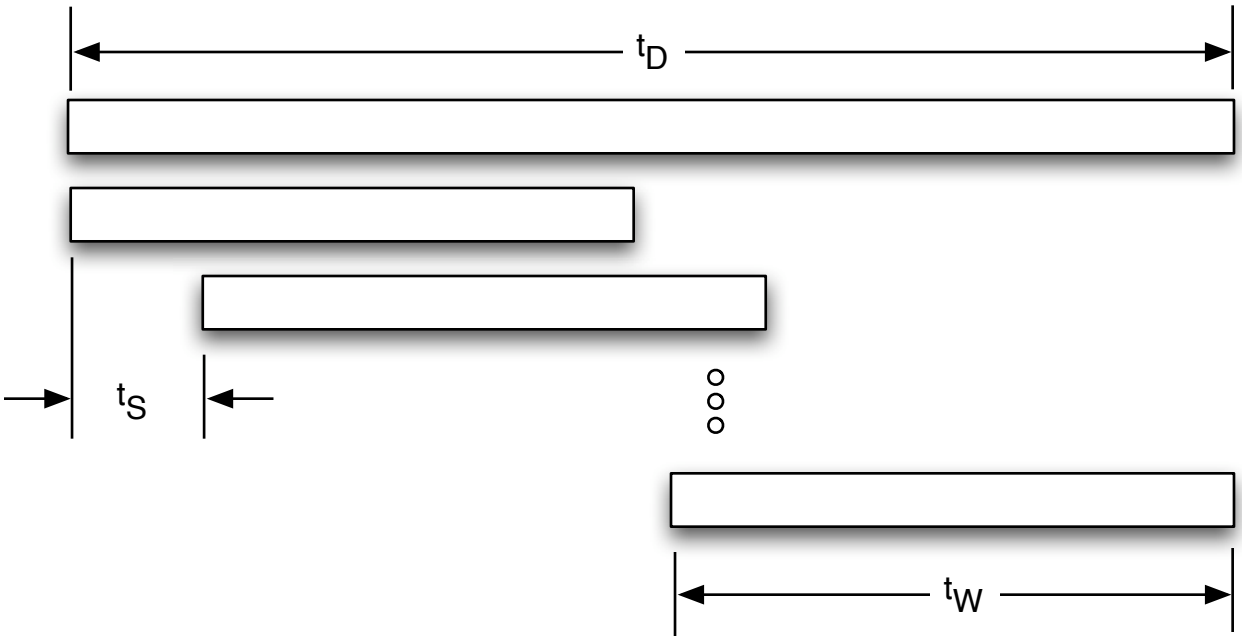
Figure 3: Fractional RMS error, ξ and component–averaged shear viscosity, η_{AVE} for 12 960 particles (N=12 960) at 1424 K (±17 K), 175.8 MPa (±24.6 MPa) and ρ = 1379.3 kg m^{-3} for simulation durations t_D ranging from 0.02 to 10 ns. A simulation duration t_D = 2 ns is a good compromise between accuracy and computational cost. The inset shows that asymptotic shear viscosity values are not obtained in short duration experiments.

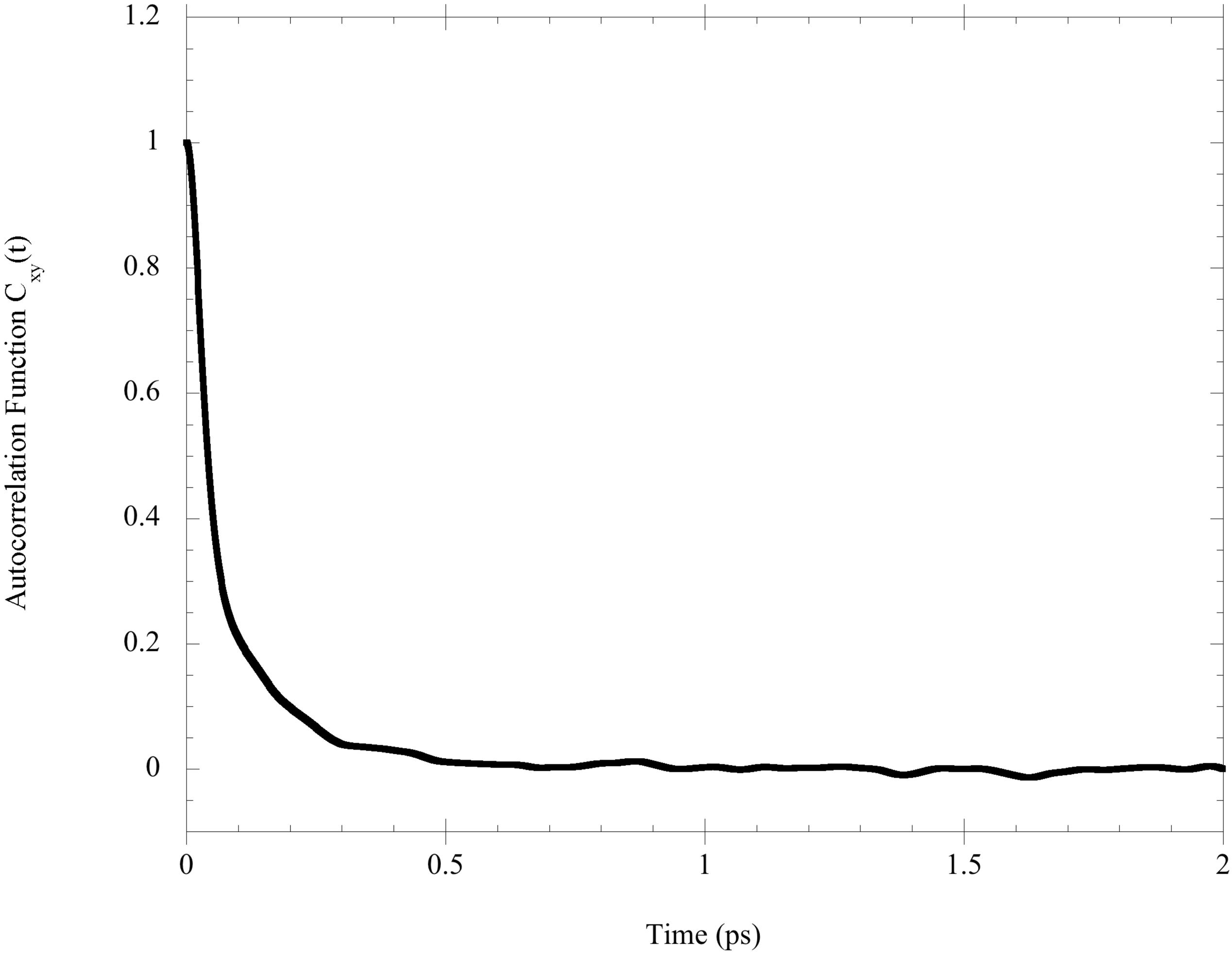
Figure 4: Fractional RMS error, ξ for N= 12 960 at 1424 K (±17 K), 175.8 MPa (±24.6 MPa) and ρ = 1379.3 kg m^{-3} *versus* the number of time origins (n_O) or the time spacing, t_S, between successive time origins used in calculation of the autocorrelation function. At least 100000 origins or (t_S = 10 fs) are needed to develop adequate statistics for determination of shear viscosity. Run duration is t_D= 10 ns and window width t_W is at 2 ps.

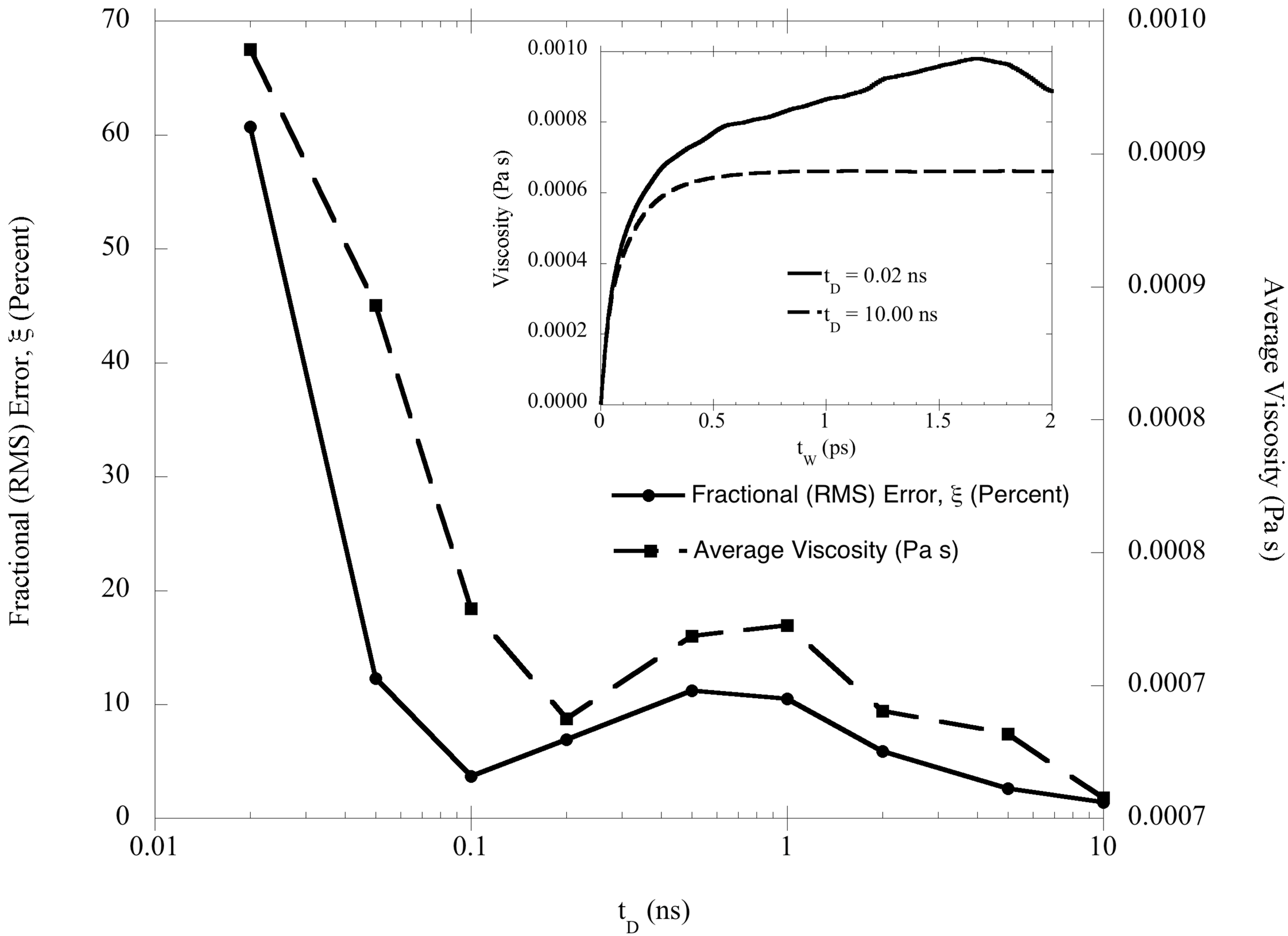
Figure 5: Fractional RMS error, ξ and average shear viscosity as a function of N (particle number) at 1399 K (±79 K), 182.4 MPa (±165.7 MPa) and ρ = 1379.3 kg m^{-3}. Simulation duration t_D = 2 ns, t_W = 2 ps and t_S= 10 fs for all calculations. Although there is no obvious dependence of ξ on N, it should be emphasized that the fluctuation in temperature (σ_T) and pressure (σ_P) scale as $N^{-1/2}$. In order to assign a shear viscosity to a specific well-known temperature and pressure, large N simulations are required (see Figure 6).

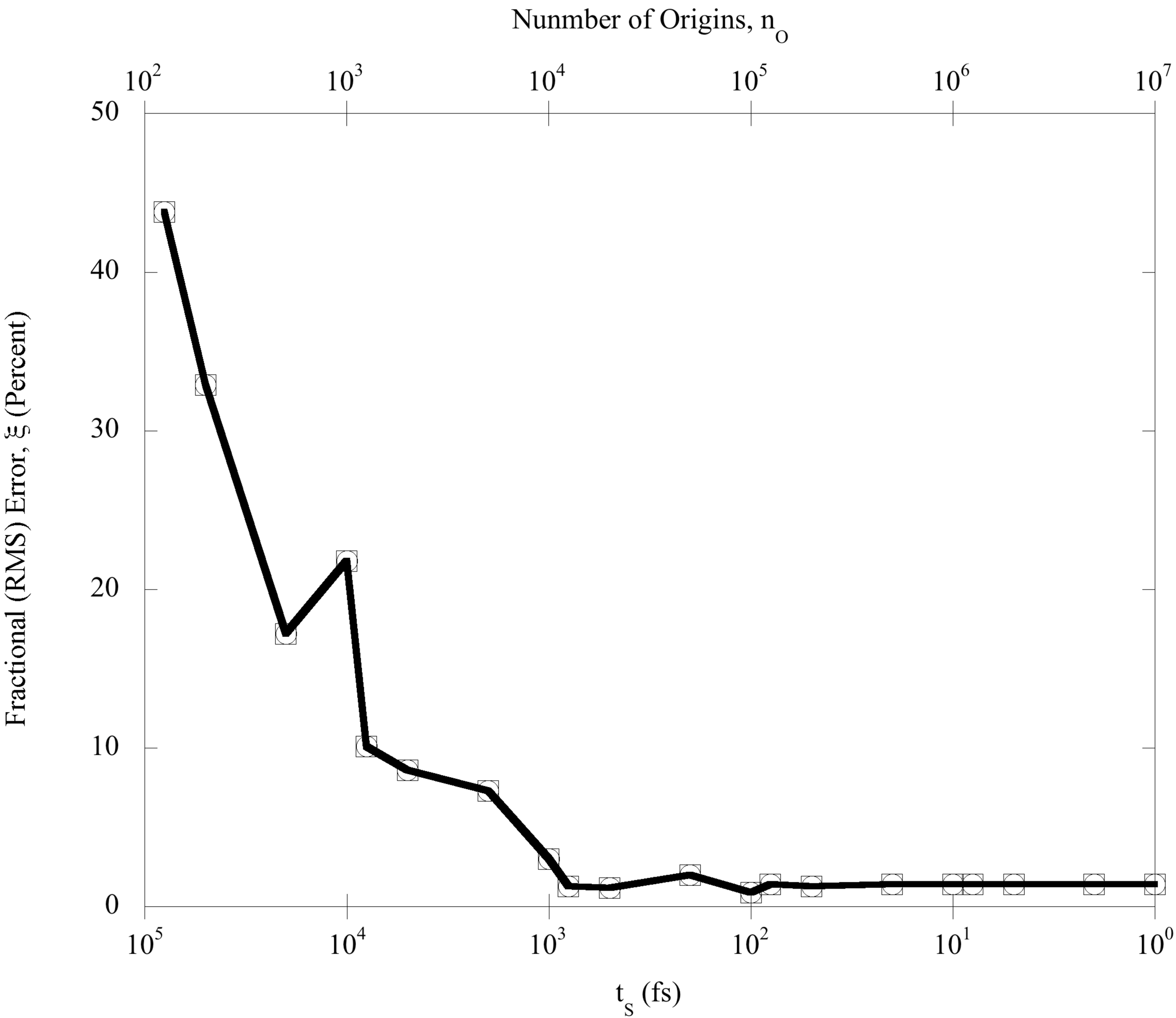
Figure 6: Normalized temperature ($\sigma_T / \hat{T}$) and pressure ($\sigma_P / \hat{P}$) fluctuations versus particle number, N. In the limit of large N, σ_T and σ_P scale as $N^{-1/2}$. Since one must associate a shear viscosity to a specific state point, large N simulations are needed to decrease σ_T and σ_P to acceptable levels. Although the GK retrieval of shear viscosity does not depend explicitly upon N, the uncertainty of the state point T and P clearly do.

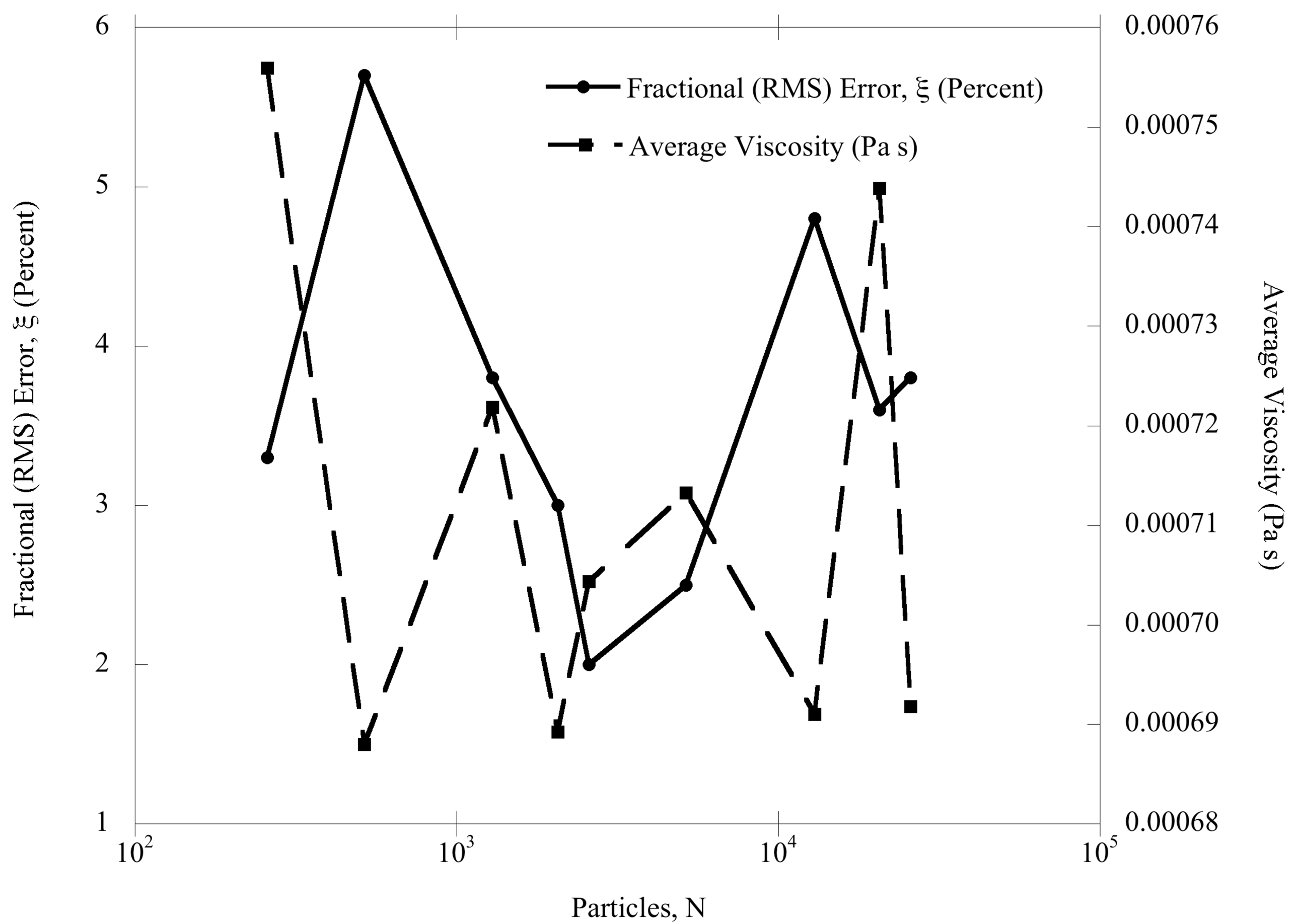
Table 1: Potential parameters for NaCl used in this study

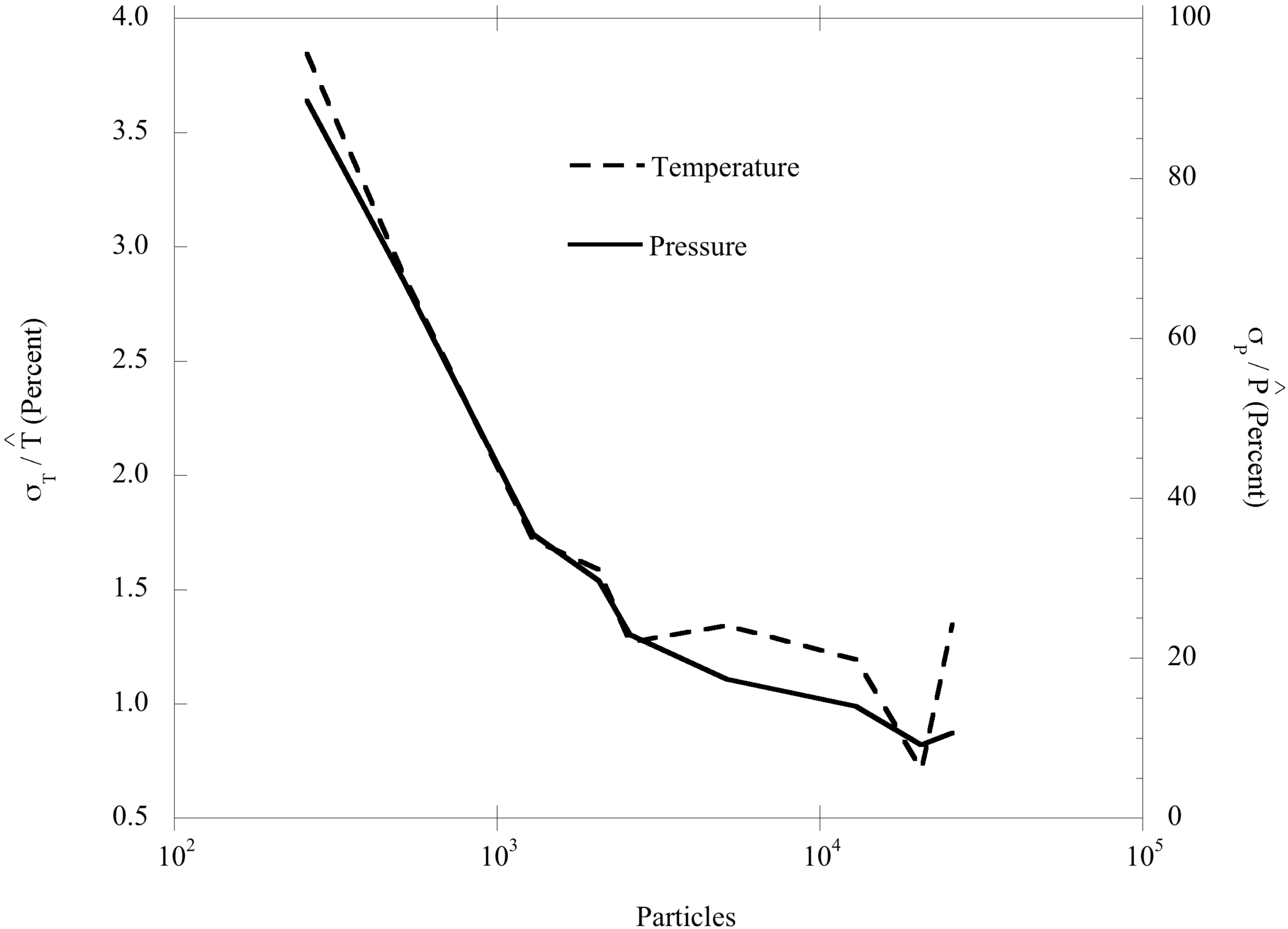

t_D
t_S
t_W


Autocorrelation Function $C_{xy}(t)$
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
Time (ps)


Fractional (RMS) Error, ξ (Percent)
Average Viscosity (Pa s)
t_D (ns)
0.01
0.1
1
10
0
10
20
30
40
50
60
70
0.0007
0.0008
0.0009
0.0009
0.0008
0.0010
Fractional (RMS) Error, ξ (Percent)
Average Viscosity (Pa s)
Viscosity (Pa s)
t_W (ps)
0.0000
0.0002
0.0004
0.0006
0.0008
0.0010
0
0.5
1
1.5
2
t_D = 0.02 ns
t_D = 10.00 ns


Nunmber of Origins, n_O
10^2
10^3
10^4
10^5
10^6
10^7
50
40
30
20
10
0
Fractional (RMS) Error, ξ (Percent)
10^5
10^4
10^3
10^2
10^1
10^0
t_S (fs)


Fractional (RMS) Error, ξ (Percent)
Average Viscosity (Pa s)
Fractional (RMS) Error, ξ (Percent)
Average Viscosity (Pa s)
Particles, N
6
5
4
3
2
1
0.00076
0.00075
0.00074
0.00073
0.00072
0.00071
0.00070
0.00069
0.00068
10^2
10^3
10^4
10^5


4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
100
80
60
40
20
0
Temperature
Pressure
$\sigma_T / \hat{T}$ (Percent)
$\sigma_P / \hat{P}$ (Percent)
10^2
10^3
10^4
10^5
Particles

Species	A_{ij} ($x10^{-7}$ J mol^{-1})	B_{ij} ($x10^{11}$ m)	C_{ij} ($x10^{52}$ J m^6 mol^{-1})	D_{ij} ($x10^{73}$ J m^8 mol^{-1})
Na-Na	4.040947	3.174603	1.011070	4.814617
Na-Cl	11.948257	3.174603	6.740466	83.653976
Cl-Cl	33.120536	3.174603	69.811965	1402.257558