

# 1 Normal Mode mass re-weighting issues.

## 1.1 Mapping between spaces.

For the Normal Mode methods we obtain the Eigenvalue/Eigenvector sets by minimizing the structure of interest and then diagonalizing the mass re-weighted Hessian. Since we wish to order the system by frequency the mass re-weighting is necessary so that the resulting system is ‘massless’ and hence the frequencies are only dependent on the potential energy function. The eigenvectors  $\mathbf{Q}$  then provide the following mapping from the Hessian  $\mathcal{H}$  to the matrix  $\mathbf{D}$  with the corresponding eigenvalues on its diagonal (zero elsewhere)

$$\mathbf{M}^{-\frac{1}{2}}\mathcal{H}\mathbf{M}^{-\frac{1}{2}}\mathbf{Q} = \mathbf{Q}\mathbf{D}, \quad (1)$$

where  $\mathbf{M}$  is the diagonal matrix of the system masses. The mapping from the real positions  $\mathbf{x}$  to the ‘massless’ positions  $\hat{\mathbf{x}}$  is not clear from this equation since the Hessian was calculated at the minimized structure  $\mathbf{x}_0$ . If we consider the Hamiltonian

$$\mathbf{H} = \frac{1}{2}\dot{\mathbf{x}}^T\mathbf{M}\dot{\mathbf{x}} + U(\mathbf{x}), \quad (2)$$

we obtain the ‘massless’ Hamiltonian by substituting  $\hat{\mathbf{x}} = \mathbf{M}^{\frac{1}{2}}\dot{\mathbf{x}}$ , which comes from the mapping  $\hat{\mathbf{x}} = \mathbf{M}^{\frac{1}{2}}(\mathbf{x} - \mathbf{x}_0)$ , to give

$$\hat{\mathbf{H}} = \frac{1}{2}\dot{\hat{\mathbf{x}}}^T\dot{\hat{\mathbf{x}}} + U\left(\mathbf{M}^{-\frac{1}{2}}\hat{\mathbf{x}} + \mathbf{x}_0\right). \quad (3)$$

Differentiating  $U$  twice w.r.t.  $\hat{\mathbf{x}}$  then leads to the mass re-weighted Hessian since  $\mathbf{x} = \mathbf{x}_0$ , the correct mappings between the real and normal mode spaces are then

$$\mathbf{c} = \mathbf{Q}^T\mathbf{M}^{\frac{1}{2}}(\mathbf{x} - \mathbf{x}_0), \quad (4)$$

$$\mathbf{x} = \mathbf{M}^{-\frac{1}{2}}\mathbf{Q}\mathbf{c} + \mathbf{x}_0, \quad (5)$$

where  $\mathbf{c}$  represents the modes.

## 1.2 Normal mode subspace forces.

From Newtons second law we have

$$\mathbf{M}\ddot{\mathbf{x}} = \mathbf{F}, \quad (6)$$

for force  $\mathbf{F}$ . We can map this into the normal mode subspace using (4)

$$\ddot{\mathbf{c}} = \mathbf{Q}^T\mathbf{M}^{-\frac{1}{2}}\mathbf{F}. \quad (7)$$

To evolve the system in the real space we can map (7) back to the real space using (5) giving, after multiplying through by  $\mathbf{M}$ ,

$$\mathbf{M}\ddot{\mathbf{x}} = \mathbf{M}^{\frac{1}{2}}\mathbf{Q}\mathbf{Q}^T\mathbf{M}^{-\frac{1}{2}}\mathbf{F}. \quad (8)$$

We can then define the normal mode subspace force in the real space,  $\mathbf{f}_s$ , as

$$\mathbf{f}_s = \mathbf{M}^{\frac{1}{2}} \mathbf{Q} \mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \mathbf{F}. \quad (9)$$

The forces in the complement space,  $\mathbf{f}_c$ , will then be

$$\begin{aligned} \mathbf{f}_c &= \mathbf{F} - \mathbf{M}^{\frac{1}{2}} \mathbf{Q} \mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \mathbf{F} \\ &= \mathbf{M}^{\frac{1}{2}} (\mathbf{I} - \mathbf{Q} \mathbf{Q}^T) \mathbf{M}^{-\frac{1}{2}} \mathbf{F}. \end{aligned} \quad (10)$$

In practice we require the forces in the space complement to that spanned by  $\bar{\mathbf{Q}}$ , the eigenvectors in  $\mathbf{Q}$  and the six conserved degrees of freedom, giving

$$\bar{\mathbf{f}}_c = \mathbf{M}^{\frac{1}{2}} (\mathbf{I} - \bar{\mathbf{Q}} \bar{\mathbf{Q}}^T) \mathbf{M}^{-\frac{1}{2}} \mathbf{F}. \quad (11)$$

### 1.3 Langevin thermostat.

As previously observed, to use the Langevin thermostat correctly in the subspace requires that the random variables meet the requirements of the fluctuation/dissipation theorem in that subspace. This leads to the generation of  $m$  random variables if the dimension of the subspace is  $m$ . We consider the Langevin equation in the subspace using the projected real forces and with fixed  $\gamma$

$$\ddot{\mathbf{c}} = \mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \mathbf{F} - \gamma \dot{\mathbf{c}} + \mathcal{R}, \quad (12)$$

where  $\mathcal{R}$  represent the random forces. Mapping back to the real space gives

$$\ddot{\mathbf{x}} = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{Q}^T \mathbf{M}^{-\frac{1}{2}} \mathbf{F} - \gamma \dot{\mathbf{x}} + \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathcal{R}. \quad (13)$$

In the massless subspace we require that

$$\langle \mathcal{R}(t)^T \mathcal{R}(t') \rangle = 2\gamma k_B T \delta(t - t'). \quad (14)$$

In the latest implementation of the software  $m$  random variables are drawn from  $N(0, 1)$ , projected to massless space by  $\mathbf{Q}$  and the multiplied by  $\sqrt{2\gamma k_B T} \mathbf{M}^{-\frac{1}{2}}$  to meet (14) and (13).

## 2 Eigenvalue and Eigenvector ‘lifetime’.

Our current scheme evolves the system in the  $3N$  cartesian space with the selected Charm force field, but achieves course-graining in time by constraining the *forces* (to zero) in a linear subspace. This is chosen so that it coincides with the space where the numerical method would not converge for the chosen step-size. The choice of sub-space and step-size is determined from a Normal Mode analysis of the system at the initial, minimized, conformation.

Since the eigenvalues and vectors from a normal mode analysis of the system will change as it evolves then the ‘lifetime’ of those from the initial conformation is of interest. To extend the lifetime of the constrained subspace we have adopted the concept of a ‘compliment subspace’, where the eigenvectors of this subspace are not considered but it is determined as the subspace which is the compliment of a set of ‘low frequency’ eigenvectors, which are known to have considerably longer lifetime. The lifetime of these low frequency eigenvectors will determine the ability of the method to allow large conformation changes and also the length of time where the numerical method is operating within its stability region. It is the latter problem that I will try and address here.

## 2.1 Rayleigh quotient.

The eigenvalues of a Normal Mode analysis represents the square of the frequency associated with it’s mode, and is generally accurate for medium to high frequencies. It is these values that we use to chose the optimal step-size for a given size set of low frequency eigenvectors. If the eigenvector/value set is no longer valid then the stability of the numerical method is not guaranteed, usually leading to complimentary sub-space structure which cannot be effectively minimized. It is important to determine the end of life of the eigenvector set so that a new normal mode analysis can be carried out.

The Rayleigh Quotient allows us to estimate an approximate eigenvalue  $\alpha_i$  given an approximate eigenvector  $v_i$ , since our Hessian  $\mathcal{H}$  is symmetric. Then

$$\alpha_i = \frac{v_i^T \mathcal{H} v_i}{v_i^T v_i}. \tag{15}$$

Furthermore we can find bounds for the nearby eigenvalue

$$\begin{aligned} \lambda_i &\in [\alpha_i - \epsilon, \alpha_i + \epsilon], \\ \epsilon &= \frac{\|(\mathcal{H} - \alpha_i I) v_i\|_2}{\|v_i\|_2}. \end{aligned} \tag{16}$$

It seems reasonable to assume that the stability of the numerical method will be compromised when one of the low frequency eigenvectors has an associated eigenvalue which is greater than the limit imposed by the choice of time-step. To this end I implemented the Rayleigh Quotient method in Protomol and studied the blocked Alanine Dipeptide model.

For and ‘averaged’ Hessian the Alanine Dipeptide at a step size of 4 fs gave the results in Figure 1. Here the highest frequency mode in the set of low frequency eigenvectors is used. We see that the estimated eigenvalue rapidly jumps from 27 to 38, a 1.2 times increase in frequency. The interesting change occurs around step 4070 where the estimated eigenvalue jumps to 54. If we study the conformations in Figure 2 we see that all that has happened is that the minimizer has rotated the ALA:HB1-3 triplet around through  $\frac{2}{3}\pi$  radians (RHS of figure). Although this indicates that we can detect conformation change in this way, clearly this particular change is within the complimentary space and so should not affect out measure. The second plotted line in Figure 1 shows

$$\hat{\alpha}_i = \frac{v_i^T \mathbf{Q} \mathbf{Q}^T \mathcal{H} \mathbf{Q} \mathbf{Q}^T v_i}{v_i^T \mathbf{Q} \mathbf{Q}^T v_i}, \tag{17}$$

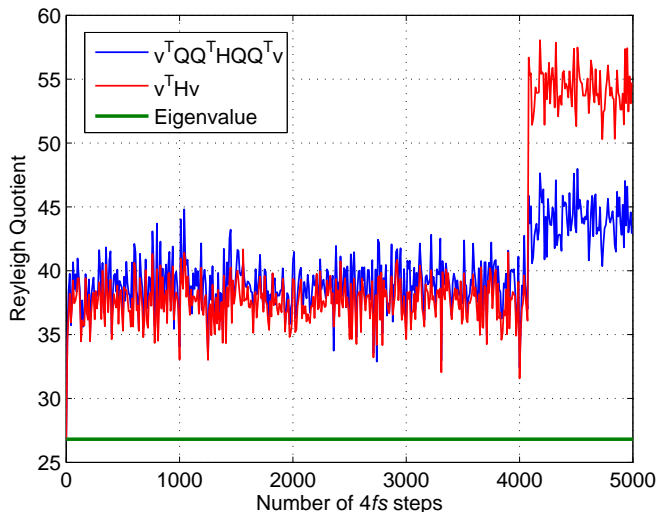


Figure 1: 20ps trajectory for Alanine Dipeptide showing the Rayleigh Quotient for the highest frequency eigenvector.

as expected this removes most of the change and is our candidate for the measure.

The bound of the Rayleigh Quotient seems less useful for the method as we see in Figure 3, although the bound is much less when the compliment space is removed.

I will now test the measure on a WW-domain model which is known to fail after a given length of trajectory.

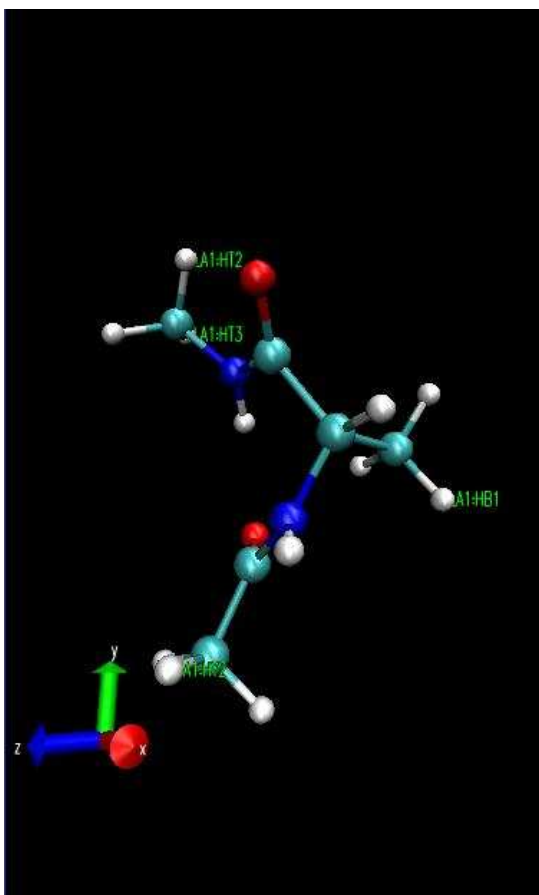
## 2.2 Averaged Hessian.

We have always observed better stability for the method when using an averaged Hessian in the diagonalization. We had initially assumed that this was due to the removal of negative eigenvalues, having an imaginary frequency. Use of the Rayleigh Quotient is illuminating here if we consider Figure 4. For the non-averaged Hessian the Rayleigh estimate is much higher, indicating that the high frequency components of the Hessian have a disproportionately large effect on the analysis. I will investigate this further as it may provide a theoretical basis for our use of the averaged Hessian.

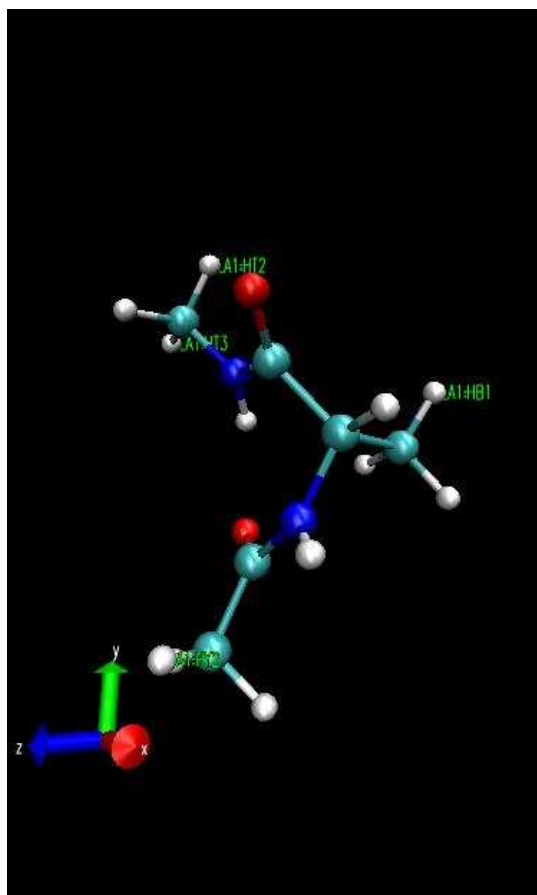
1

---

<sup>1</sup>Chris Sweet. 12/2/2006.



Conformation at step 4075.



Conformation at step 4076.

Figure 2: Conformations before and after the step change in Rayleigh Quotient.

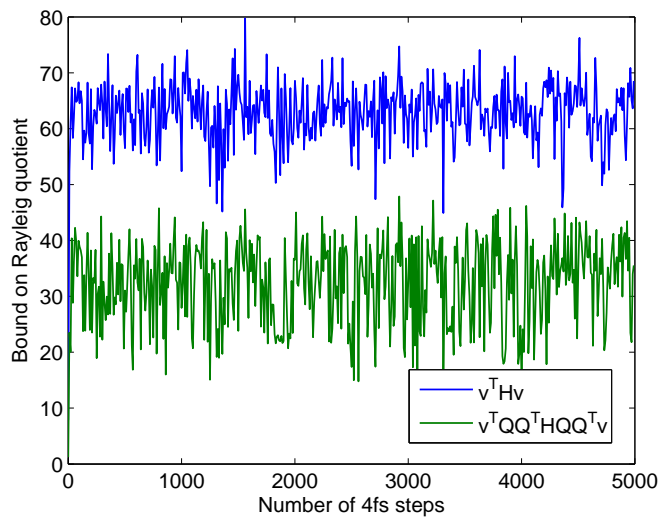


Figure 3: 20ps trajectory for Alanine Dipeptide showing the bounds on the Rayleigh Quotient.

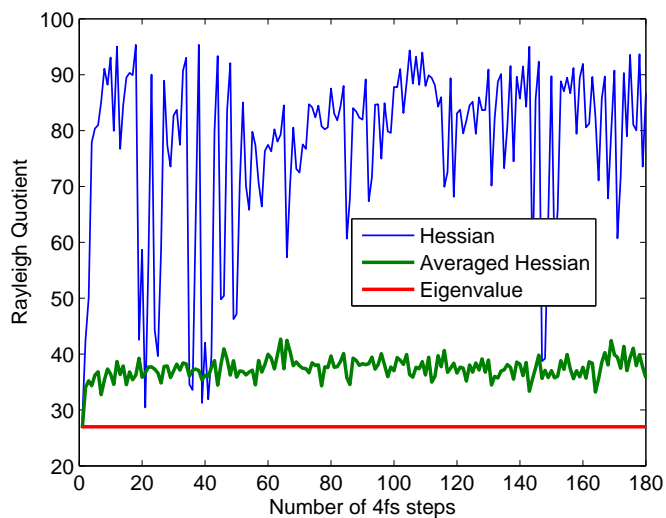


Figure 4: Comparison of Rayleigh Quotient for normal and averaged Hessians.