Normal Model Analysis & Elastic Network Models

A brief overview of methods based on “spring-network” models
Spring network models

Much of the methods that we’ve studied this semester provide insight into protein dynamics (i.e., MD, FIRST, and the DCM). However, all have issues that limit their utility.

For example, while MD does provide mobility information, the compute cost of the algorithm prevents us from accessing the longest timescale motions.

Conversely, while FIRST and the DCM are very fast (they both provide descriptions of quasi-stationary motions). Moreover, network rigidity describes flexibility (deformability); not mobility.

Spring-network models fill the gap between these two (and other) methods to probe macromolecular dynamics.

Generally speaking, there are two classes of spring-network models: (1) NMA and (2) ENM.
Common applications of NMA

- As you will see next, a complete description of the internal energy, $U(T)$, requires a complete description of the vibrational component of the internal energy, $U_{vib}(T)$, which can be achieved via NMA. This is one common application of NMA.

- However, it is much more common in large biomolecular systems to use NMA (or its derivatives) to study the dynamic equilibrium fluctuations (atomic mobilities).
Normal mode analysis

NMA is classical technique for studying the vibrational and thermal properties of various molecular structures at the atomic level.

Although this technique is widely used for molecular systems consisting of a small number of atoms, performing NMA on large systems is computational challenging.

Normal modes are useful because they correspond to collective motions of the atoms in a coupled system that can be individually excited.

The normal modes of CO$_2$ are:

![Normal modes of CO$_2$](image)

Figure 1. Normal Modes for a linear triatomic molecule. In the last bending vibration the motion of the atoms is in-and-out of the plane of the paper.
Normal mode analysis

The internal energy of a system at temperature $T$ is given by:

$$U(T) = U_{\text{trans}}(T) + U_{\text{rot}}(T) + U_{\text{vib}}(T)$$

If all translational and rotational modes are fully accessible in accordance with the equipartition theorem, then both are directly proportional to $T$.

However, the vibrational contribution requires knowledge of the actual vibrational frequencies.

The vibrational contribution, $U_{\text{vib}}(T) - U_{\text{vib}}(0)$, is given by:

$$U_{\text{vib}}(T) = \sum_{i=1}^{N_{nm}} \left( \frac{h\nu_i}{2} + \frac{h\nu_i}{\exp[h\nu_i / k_B T] + 1} \right)$$

where $h$ is Plank’s constant, and $\nu$ is the vibrational frequency, and $N_{nm}$ is the # of vibrational modes, which is $= 3N - 6$ ($3N - 5$ for a linear molecule).

Q: How do we identify the normal modes?
Harmonic oscillator review

Normal modes are calculated using Newton’s equations of motion.

Before diving into NMA, let’s consider a simple harmonic oscillator...

Consider a mass \( m \), supported on a string with force constant \( k \). Hooke’s law, \( F = -k\Delta x \), allows us to determine the force for compression/stretching of the spring from its equilibrium distance.

The potential energy is obtained by integrating \( F = -dV/dx = -kx \), to give \( V = 1/2*kx^2 \).

The force constant can be derived from the second derivative of the potential energy \( k = d^2V/dx^2 \).

The force constant is substituted into Newton’s eq, \( F = ma = m(d^2x/dt^2) = -(d^2V/dx^2)x \).

The second derivative of the extension, \( d^2x/dt^2 = -4\pi^2v^2x \), where \( v \) is the vibrational frequency. Substituting into the previous expression gives \( -4\pi^2v^2mx = -(d^2V/dx^2)x \), which is the basis of NMA.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
Normal mode analysis

For molecules the $x, y, z$ coordinates of each atom must be specified.

The extensions are the differences in the positions and the eq. positions (i.e., $x_1 = X_1 - X_{1eq}$).

MM calculations (sometimes MO) are used to find the PE of the molecule as a function of atomic positions, $V(x_1', y_1', z_1'... x_n', y_n', z_n')$.

The second derivative of the PE is used to calculate the force constants. There are $3Nx3N(!)$ possible second derivatives (which, as we saw before, are the force constants).

For example, $-(d^2V/dx_1^2) = k_{xx}^{11}$ is the change of the force on atom 1 in the x-direction when you move atom 1 in the x-direction.

Similarly, $-(d^2V/dx_1dy_2) = k_{xy}^{12}$ is the change of the force on atom 1 in the x-direction when you move atom 2 in the y-direction.

These force constants are not for individual bonds; rather, they describe the motion of a single atom subject to its neighbors (whether they are bonded or not).

The complete list of second derivatives is called a Hessian.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
Normal mode analysis

With the force constants in hand...

\[-4\pi^2 \nu^2 m_1 x_1 = -k_{xx}^{11} x_1 - k_{xy}^{11} y_1 - k_{xz}^{11} z_1 - k_{xx}^{12} x_2 - k_{xy}^{12} y_2 - k_{xz}^{12} z_2 - \ldots - k_{xz}^{1N} Z_N\]
\[-4\pi^2 \nu^2 m_1 y_1 = -k_{yx}^{11} x_1 - k_{yy}^{11} y_1 - k_{yz}^{11} z_1 - k_{yx}^{12} x_2 - k_{yy}^{12} y_2 - k_{yz}^{12} z_2 - \ldots - k_{yz}^{1N} Z_N\]
\[-4\pi^2 \nu^2 m_1 z_1 = -k_{zx}^{11} x_1 - k_{zy}^{11} y_1 - k_{zz}^{11} z_1 - k_{zx}^{12} x_2 - k_{zy}^{12} y_2 - k_{zz}^{12} z_2 - \ldots - k_{zz}^{1N} Z_N\]
\[-4\pi^2 \nu^2 m_2 x_2 = -k_{xx}^{21} x_1 - k_{xy}^{21} y_1 - k_{xz}^{21} z_1 - k_{xx}^{22} x_2 - k_{xy}^{22} y_2 - k_{xz}^{22} z_2 - \ldots - k_{xz}^{2N} Z_N\]
\[-4\pi^2 \nu^2 m_2 y_2 = -k_{yx}^{21} x_1 - k_{yy}^{21} y_1 - k_{yz}^{21} z_1 - k_{yx}^{22} x_2 - k_{yy}^{22} y_2 - k_{yz}^{22} z_2 - \ldots - k_{yz}^{2N} Z_N\]
\[-4\pi^2 \nu^2 m_2 z_2 = -k_{zx}^{21} x_1 - k_{zy}^{21} y_1 - k_{zz}^{21} z_1 - k_{zx}^{22} x_2 - k_{zy}^{22} y_2 - k_{zz}^{22} z_2 - \ldots - k_{zz}^{2N} Z_N\]
\[\ldots\]
\[-4\pi^2 \nu^2 m_N x_N = -k_{xx}^{N1} x_1 - k_{xy}^{N1} y_1 - k_{xz}^{N1} z_1 - k_{xx}^{N2} x_2 - k_{xy}^{N2} y_2 - k_{xz}^{N2} z_2 - \ldots - k_{xz}^{NN} Z_N\]
\[-4\pi^2 \nu^2 m_N y_N = -k_{yx}^{N1} x_1 - k_{yy}^{N1} y_1 - k_{yz}^{N1} z_1 - k_{yx}^{N2} x_2 - k_{yy}^{N2} y_2 - k_{yz}^{N2} z_2 - \ldots - k_{yz}^{NN} Z_N\]
\[-4\pi^2 \nu^2 m_N z_N = -k_{zx}^{N1} x_1 - k_{zy}^{N1} y_1 - k_{zz}^{N1} z_1 - k_{zx}^{N2} x_2 - k_{zy}^{N2} y_2 - k_{zz}^{N2} z_2 - \ldots - k_{zz}^{NN} Z_N\]

The above equations state that the total force on atom \(i\) is the sum of all the pairwise interactions in each of the three constituent directions.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
Normal mode analysis

It is common to rewrite the previous equations in terms of mass weighted positions where $X_{i}^{\sim} = m^{1/2}x_{i}$, thus they now have the form of...

$$\begin{pmatrix}
\frac{k_{11}}{\sqrt{m_{1}/m_{1}}} & \frac{k_{12}}{\sqrt{m_{1}/m_{2}}} & \frac{k_{13}}{\sqrt{m_{1}/m_{3}}} \\
\frac{k_{21}}{\sqrt{m_{2}/m_{1}}} & \frac{k_{22}}{\sqrt{m_{2}/m_{2}}} & \frac{k_{23}}{\sqrt{m_{2}/m_{3}}} \\
\frac{k_{31}}{\sqrt{m_{3}/m_{1}}} & \frac{k_{32}}{\sqrt{m_{3}/m_{2}}} & \frac{k_{33}}{\sqrt{m_{3}/m_{3}}}
\end{pmatrix}
\begin{pmatrix}
x_{1} \\
x_{2} \\
x_{3}
\end{pmatrix}
= 4\pi^{2}v^{2}
\begin{pmatrix}
x_{1} \\
x_{2} \\
x_{3}
\end{pmatrix}$$

where mass weighted force constants result in a symmetrical matrix.

Notice that the above equation is an eigenvalue-eigenvector equation. Each eigenvector is referred to as a normal mode, whose vibrational frequency is given by the eigenvalue.

The overall dynamics of the system can be described by a superposition of a number of linearly independent normal modes.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
NMA vs. MD

• MD and NMA are really quite similar; both include the kinetic and potential energy of the molecule.

• The ff is generally the same.

• They both calculate the Hessian and then integrate Newton’s laws of motion.

• The motions that you see in MD are, in fact, the normal modes and the atomic position fluctuations can be used to extract the normal mode frequencies.

• The difference between MD and NMA is that the equations of motion are integrated numerically in MD, but sinusoidal solutions are assumed in NMA.

• In addition, in MD all normal mode motions are studied simultaneously, whereas they are studied one mode at a time in NMA.

• In practice, it can be easier to extract thermodynamic properties from NMA because of the long MD simulation times that are needed to ensure adequate sampling.

• On the other hand, MD more easily handles anharmonicity and explicit solvation.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
NMA of small molecules

NMA is particularly important in molecular spectroscopy.

Using NMA, it is actually possible to reproduce the vibrational spectra of simple molecules.

This is because that spectroscopic methods (i.e., IR and Raman) are based on the excitation of specific vibrational frequencies.

Figure 2. The Infrared spectrum of air. This spectrum is the background scan from an FT-IR spectrometer.

Based on: http://www.colby.edu/chemistry/PChem/notes/NormalModesText.pdf
Normal modes

Ref: http://situs.biomachina.org/sd03/flo.pdf
Normal modes

Collective motions
## Problems of NMA in large systems

<table>
<thead>
<tr>
<th>NMA requires:</th>
<th>Problems for large systems:</th>
</tr>
</thead>
</table>
| → Energy minimization | → Time consuming due to size of system  
<→ Can distort the structure |
| → Diagonalization of the Hessian  
(2nd derivative of PE) | → Time consuming due to size of matrix |

Ref: http://situs.biomachina.org/sd03/flo.pdf
## Problems of NMA in large systems

<table>
<thead>
<tr>
<th>NMA requires:</th>
<th>Problems for large systems:</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ Energy minimization</td>
<td>→ Time consuming due to size of system</td>
</tr>
<tr>
<td></td>
<td>→ Can distort the structure</td>
</tr>
<tr>
<td></td>
<td><strong>No energy minimization b/c protein is simplified to beads on a string</strong></td>
</tr>
<tr>
<td>→ Diagonalization of the Hessian (2nd derivative of PE)</td>
<td>→ Time consuming due to size of matrix</td>
</tr>
<tr>
<td></td>
<td><strong>ENM matrix is much smaller</strong></td>
</tr>
</tbody>
</table>

Meaning, ENMs resolve many of the practical problems associated with NMA.

Ref: http://situs.biomachina.org/sd03/flo.pdf
Elastic network models resolve many of the practical problems associated with NMA

Connectivity (Kirchhoff) matrix is defined by:

\[
\Gamma_{ij} = \begin{cases} 
-1, & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\
0, & \text{if } i \neq j \text{ and } R_{ij} > r_c \\
-\sum_{j,j \neq i}^{N} \Gamma_{ij}, & \text{if } i = j
\end{cases}
\]
## ENM spring constants

**ENM spring constants:**

→ A single off-diagonal value is used throughout:

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

→ *Q: What value should be used?*
### ENM spring constants:

- A single off-diagonal value is used throughout:

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

→ Q: What value should be used?

The spring constants are fit to experimental B-factors.

---

**Q:** Explain what problems are associated with this approach.
The slowest mode of Cdc2B catalytic domain from an ENM

Ref: http://en.wikipedia.org/wiki/Gaussian_network_model
ENMs have been used to probe catalytic site dynamics

Color-coded ribbon diagrams for two enzymes, β-lactamase (1BLC) (a) and penicillopepsin (1BXO) (b), illustrating the mobility of residues in the first (lowest frequency) GNM mode. The color code is blue-red-yellow-green in the order of increasing mobility. Both enzymes contain an inhibitor (shown in space filling, gray) bound near the most constrained (lowest mobility) region. (c) and (d) Corresponding square fluctuation profiles and positions of catalytic and inhibitor-binding residues. See table 1 for the list of chemical (from experiments) and mechanical (from computations) key residues. Residues directly involved in catalytic function at active sites are shown by the green open circles, inhibitor-binding residues are shown by the red squares and residues serving both catalytic and inhibitor-binding functions are marked by the orange diamond.

Based on the compute efficiency of ENMs, the dynamics of very large systems can be considered.

Dynamics of the HK97 bacteriophage viral capsid using ENMs.

Predicting allosteric sites via DPA (dynamics perturbation analysis)

Ref: Ming et al. (2008), BMC Struct Biol, 8:5.
Ming & Wall (2005), PROTEINS, 59:697-707.
Ming & Wall (2005), Phys Rev Let, 95:337-345.
There are myriad ENM servers on the web

**GNM servers**
- cGNM: Online calculation of structural dynamics using GNM [http://gnm.ccb.pitt.edu/GNM_Online_Calculation.htm](http://gnm.ccb.pitt.edu/GNM_Online_Calculation.htm)
- GNM server [http://gor.bb.iastate.edu/gnm/gnm.htm](http://gor.bb.iastate.edu/gnm/gnm.htm)

**ANM servers**
- Anisotropic Network Model web server [http://www.ccb.pitt.edu/anm](http://www.ccb.pitt.edu/anm) [28]
- ANM server [http://gor.bb.iastate.edu/anm/anm.htm](http://gor.bb.iastate.edu/anm/anm.htm)

**ENM servers**
- eLنمو: Web-interface to The Elastic Network Model [http://www.igs.cnrs-mrs.fr/elnemo](http://www.igs.cnrs-mrs.fr/elnemo)

**Other relevant servers**
- HingeProt: An algorithm for protein hinge prediction using elastic network models [http://www.prc.boun.edu.tr/appserv/prc/hingeprot](http://www.prc.boun.edu.tr/appserv/prc/hingeprot)
- Protein Data Bank (PDB) [http://www.pdb.org](http://www.pdb.org)