Coefficients of multivariate pdfs in a tensor Hermite polynomial expansion

A. Kabakçıoğlu

Department of Physics, Koç University, Sarıyer 34450 İstanbul, Turkey

(Dated: April 22, 2009)

Abstract

I review the expansion of a multivariate probability distribution in terms of tensor Hermite polynomials and then discuss how the anharmonic terms of the form $x_1^\alpha x_2^\beta$ can be calculated in terms of the expectation values $\langle x_1^i x_2^j \rangle$. The result reduces the problem to calculation of Hermite polynomials $H_n(q)$ in one dimension, the coefficients of which one can write down in closed form. The method can be generalized to include three-point correlations $\langle x_1^i x_2^j x_3^k \rangle$.

Given a multivariate probability distribution function $f(\mathbf{\Delta R})$ of the variables ΔR_i $(i = 1 \ldots N)$, one can express this distribution in terms of its deviation from a multi-dimensional Gaussian distribution as

$$f(\mathbf{\Delta R}) \;=\; \frac{1}{\sqrt{(2\pi)^N \det(\Gamma)}} e^{-\frac{1}{2}\mathbf{\Delta R^T \Gamma^{-1} \Delta R}} \left[1 + \text{corrections}\right] , \tag{1}$$

where $\Gamma \equiv \left\langle \mathbf{\Delta R \Delta R}^T \right\rangle$ is the covariance matrix, whose inverse defines a rotation+scaling that aligns the distribution with the principle axes of the best fitting ellipsoid which is then transformed to a sphere. In this orthonormal basis with

$$\mathbf{\Delta r} \;=\; \left\langle \mathbf{\Delta R \Delta R}^T \right\rangle^{-1/2} \mathbf{\Delta R} \tag{2}$$

$f(\mathbf{\Delta r})$ has the simpler form

$$f(\mathbf{\Delta r}) \;=\; \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2} \left[1 + \sum_{\nu=3}^{\infty} \mathbf{C}_\nu \cdot H_\nu(\mathbf{\Delta r})\right] , \tag{3}$$

where $\mathbf{C}_\nu$ (constant) and $\mathbf{H}_\nu$ (derived below) are tensors of rank ν, and $(\cdot)$ refers to $(C_\nu)^{ij..k}(H_\nu)^{ij..k}$. The fluctuations in the "normal" basis are meanless ($\langle \Delta r_i \rangle = 0$) and decoupled in second order, i.e., $\langle \Delta r_i^T \Delta r_j \rangle = \delta_{ij}$ (even when the anharmonic terms are included). The tensor Hermite polynomials used as a basis for the anharmonic corrections is particulary suitable due to the fact that the kernel of the inner product is the Gaussian function itself. For example, in one dimension H_ν reduces to the usual Hermite polynomial of rank ν with the inner product defined as the integral

$$\int_{-\infty}^{\infty} dx\, H_\nu(x) H_\mu(x) e^{-x^2/2} = \begin{cases} \nu! \;, & \nu = \mu \\ 0 \;, & \nu \neq \mu \end{cases} \tag{4}$$

Generalized to arbitrary dimension, $\mathbf{\Delta r} = (\Delta r_1, \ldots, \Delta r_d)$, the orthogonality condition takes the form

$$\int_{-\infty}^{\infty} \mathbf{dx} \left[\mathbf{C}_\nu \cdot \mathbf{H}_\nu(x)\right] \mathbf{H}_\mu(x) e^{-\mathbf{x}^2/2} = \begin{cases} \mathbf{C}_\nu \nu! \;, & \nu = \mu \\ 0 \;, & \nu \neq \mu \end{cases} \tag{5}$$

true for any constant tensor $\mathbf{C}_\nu$. The advantage of using the Hermite basis is that, the coefficients $\mathbf{C}_\nu$ of the anharmonic expansion in (3) can be obtained directly by sampling $f(\mathbf{\Delta r})$, since we have from (3)&(5)

$$\left\langle \mathbf{H}_\nu(\mathbf{\Delta r}) \right\rangle \;=\; \int_{-\infty}^{\infty} \mathbf{H}_\nu(\mathbf{x}) f(\mathbf{\Delta r})\, \mathbf{d\Delta r} = \nu!\, \mathbf{C}_\nu \;. \tag{6}$$

In [2], one dimensional Hermite polynomials are used to estimate the anharmonic contribution of the *pure moments* $\langle \Delta {r_i}^n \rangle$ to $f(\mathbf{\Delta r})$, where $\mathbf{\Delta r}$ are obtained by sampling the fluctuations of C_α atoms in a protein (giving $\mathbf{\Delta R}$) through Monte-Carlo or molecular dynamics and then applying (2).

Tensor Hermite polynomials are defined by the generating function

$$\exp\big[\frac{1}{2}\mathbf{x}^2 - \frac{1}{2}(\mathbf{x}-\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})\big] \;=\; \sum_{\{\nu_k\}} \prod_{k=1}^{d} \frac{a_k^{\nu_k}}{\nu_k!} \mathbf{H}_{\nu_k}(\mathbf{x}) \tag{7}$$

and can be obtained by successive differentiation using Rodrigues' formula:

$$\mathbf{H}_\nu(\mathbf{x}) \;=\; \frac{(-1)^\nu}{\omega(\mathbf{x})} \nabla^\nu \omega(\mathbf{x}) \;, \tag{8}$$

where $\omega(\mathbf{x}) = (2\pi)^{d/2} \exp[-\mathbf{x}^2/2]$ and $\nabla^m \equiv \nabla_{i_1}\nabla_{i_2}\ldots\nabla_{i_m}$ with $\nabla_i = \partial/\partial x_i$ is the gradient tensor of rank m. Written explicitly,

$$\begin{aligned}
\mathbf{H}_2(\mathbf{x}) \;&=\; x_i x_j - \delta_{ij} \\
\mathbf{H}_3(\mathbf{x}) \;&=\; x_i x_j x_k - \delta_{ij} x_k - \delta_{jk} x_i - \delta_{ki} x_j
\end{aligned}$$

where i, j, k run over the indices of the vector $\mathbf{x}$. Higher orders can be obtained by summing over all $0, 1, .., \nu/2$ pair-wise contractions.

Although the number of terms in $\mathbf{H}_\nu(\mathbf{x})$ becomes large rather quickly, we will focus on the case when $\mathbf{x}$ is a 2 dimensional vector (or alternatively, we will ignore the terms involving more than two different indices). With $\mathbf{x}^T = [x_1, x_2]$, we have

$$\begin{aligned}
\mathbf{H}_2(\mathbf{x})^{11} \;&=\; x_1^2 - 1 \\
\mathbf{H}_2(\mathbf{x})^{12} \;&=\; x_1 x_2 = \mathbf{H}_2(\mathbf{x})^{21} \\
\mathbf{H}_2(\mathbf{x})^{22} \;&=\; x_2^2 - 1 \\
\mathbf{H}_3(\mathbf{x})^{111} \;&=\; x_1^3 - 3x_1 \\
\mathbf{H}_3(\mathbf{x})^{112} \;&=\; x_1^2 x_2 - x_2 = \mathbf{H}_3(\mathbf{x})^{121} = \mathbf{H}_3(\mathbf{x})^{211} \\
\mathbf{H}_4(\mathbf{x})^{1111} \;&=\; x_1^4 - 6x_1^2 + 6 \\
\mathbf{H}_4(\mathbf{x})^{1112} \;&=\; x_1^3 x_2 - 3x_1 x_2 = \mathbf{H}_3(\mathbf{x})^{1121} = \mathbf{H}_3(\mathbf{x})^{1211} = \mathbf{H}_3(\mathbf{x})^{2111} \\
\mathbf{H}_4(\mathbf{x})^{1122} \;&=\; x_1^2 x_2^2 - x_1^2 - x_2^2 + 1 = \mathbf{H}_4(\mathbf{x})^{1212} = \cdots = \mathbf{H}_4(\mathbf{x})^{2211}
\end{aligned} \tag{9}$$

Each term in (9) can be graphically represented and computed as shown in Fig.1, which agrees with Rodrigues's formula in (8) (verified). As these examples suggest, tensor elements

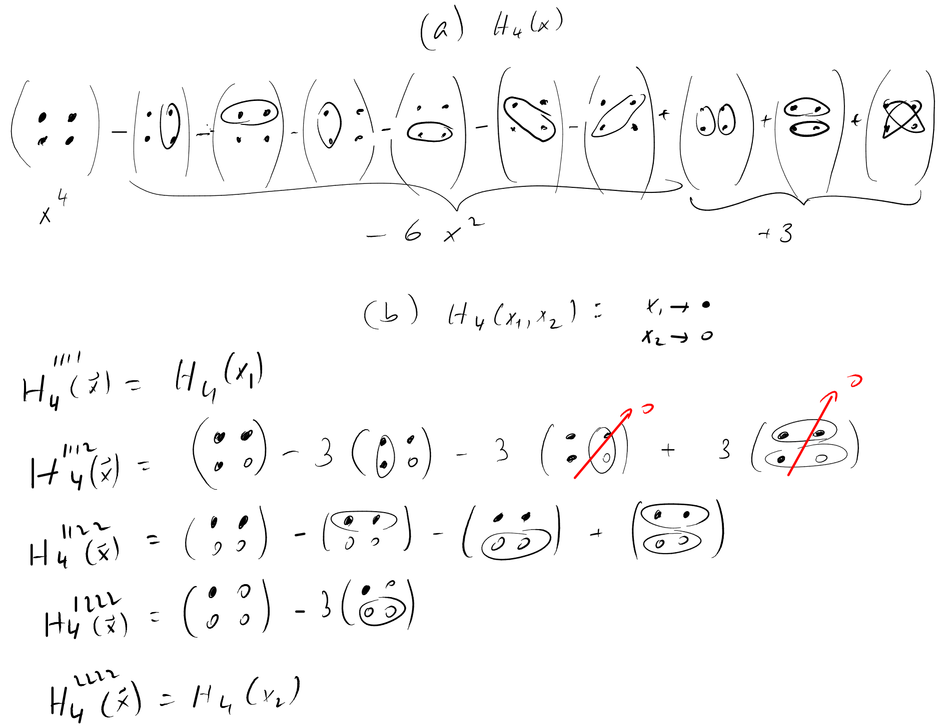

FIG. 1: Graphical representation of $H_4(x)$.

depend only on the *composition* of the index and not on the permutation. The question we would like to answer first is whether we can calculate the coefficient of each term in the expansion given in (3). Past work [2] calculated as subset of the terms which include $\mathbf{H}_\nu^{ii\ldots i}(\mathbf{\Delta r}) = H_\nu(\Delta r_i)$ only. It is also feasible to calculate higher order terms that couple *different* modes. Below, we show how the tensor elements of the form $\mathbf{H}_\nu^{i_1 i_2 \ldots i_\nu}(\mathbf{\Delta r})$ with $i_m \in \{k, l\}$ can be included.

First let's calculate for a scalar x, $H_\nu(x)$, i.e., the usual Hermite polynomials: the number of m pairwise contractions of N variables is given by the number of ways $2m$ out of N points can be clustered into sets of size 2 (Fig.1), which is given by

$$N! \,/\, (N-2m)! \, m! \, 2^m \ .$$

Then,

$$
\begin{aligned}
H_\nu(x) &= \sum_{m=0}^{\nu/2} (-1)^m \frac{\nu!}{(\nu-2m)!m!2^m}\, x^{\nu-2m} \\
&= \sum_{m=0}^{\nu/2} (-1)^m \binom{\nu}{2m} m!!\, x^{\nu-2m} \;, \qquad (10)
\end{aligned}
$$

where $m!! \equiv \frac{(2m)!}{2^m m!} = (2m-1)(2m-3)\cdots 3\cdot 1$. Equation (10) correctly reproduces the coefficients listed in the Appendix of [2]. In order to next include two-mode coupling, let's consider the case when $\mathbf{x}^T = (x_1, x_2)$ (we set $\mathbf{\Delta r} \to \mathbf{x}$ for brevity). The fact that our $\mathbf{x}$ is a higher dimensional vector ($\simeq 3\times$ number of C^α atoms in the protein) doesn't make a difference, since each mode-coupling term $x_i^\alpha x_j^\beta$ will still appear with the coefficient calculated below.

$\underline{\mathbf{x}^T = (x_1, x_2):}$

As observed above, different tensor elements with identical index composition are all equal, so let $\mathbf{H}_\nu^{p,q}(\mathbf{x})$ represent the tensor element with p of the indices equal to 1 and $q = \nu - p$ of the indices equal to 2. Furthermore, the contractions of x_i and x_j for $i \neq j$ do not contribute to $\mathbf{H}_\nu(\mathbf{x})$. Then, we simply have

$$
\mathbf{H}^{p,q}(\mathbf{x}) = H_p(x_1) \times H_q(x_2)
$$

and (3) becomes

$$
\begin{aligned}
f(\Delta r_1, \Delta r_2) &= \frac{e^{-(\Delta r_1^2 + \Delta r_2^2)/2}}{2\pi} \times \\
& \left[1 + \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \sum_{p=0}^{\nu} \binom{\nu}{p} \Big\langle H_p(x_1) H_{\nu-p}(x_2) \Big\rangle H_p(x_1) H_{\nu-p}(x_2) \right] \qquad (11)
\end{aligned}
$$

Let for simplicity, $C_p^\nu \equiv \Big\langle H_p(x_1) H_{\nu-p}(x_2) \Big\rangle$. Using (10) an setting $[\cdots] \equiv 1 + U$ above, we obtain

$$
\begin{aligned}
U &= \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \sum_{p=0}^{\nu} \binom{\nu}{p} C_p^\nu \left[\sum_{m_1=0}^{p/2} (-1)^{m_1} \binom{p}{2m_1} m_1!!\, x_1^{p-2m_1} \right] \\
& \times \left[\sum_{m_2=0}^{(\nu-p)/2} (-1)^{m_2} \binom{\nu-p}{2m_2} m_2!!\, x_2^{\nu-p-2m_2} \right] \qquad (12)
\end{aligned}
$$

After gathering terms with the the same power, (12) can be expressed in the more useful form

$$
U = \sum_{\alpha,\beta} W_{\alpha\beta} x_1^\alpha x_2^\beta \qquad (13)
$$

which unveils the total contribution of $x_1^\alpha x_2^\beta$ to the anharmonic sum U. Of course, we should get $W_{00} = W_{10} = W_{20} = W_{10} = 0$, thanks to the transformation in (2).

$W_{\alpha\beta}$ can be calculated after a bit of massaging. First note that, from (12) we require $p - 2m_1 = \alpha$ and $\nu - p - 2m_2 = \beta$. Since $p \geq \alpha$ and $\nu \geq \alpha + \beta$ (smaller p and ν terms in (12) do not contribute $W_{\alpha\beta}$), let's set $p = 2q + \alpha$ and $\nu = 2n + \alpha + \beta$ (i.e., $m_1 = q$ and $m_2 = n - q$). Then, picking up the terms with the right power in (12), we get

$$\begin{aligned} W_{\alpha\beta} &= \sum_{n=0}^{\infty} \frac{1}{(2n+\alpha+\beta)!} \sum_{q=0}^{n} \binom{2n+\alpha+\beta}{2q+\alpha} C_{2q+\alpha}^{2n+\alpha+\beta} (-1)^q \binom{2q+\alpha}{2q} q!! \times \\ &\quad \times (-1)^{n-q} \binom{2n-2q+\beta}{2n-2q} (n-q)!! \\ &= \frac{1}{\alpha!\beta!} \sum_{n=0}^{\infty} \frac{(-2)^{-n}}{n!} \sum_{q=0}^{n} \binom{n}{q} C_{2q+\alpha}^{2n+\alpha+\beta} \,. \end{aligned} \tag{14}$$

There appears to be no advantage in expanding $C_{2q+\alpha}^{2n+\alpha+\beta} = \Big\langle H_{2q+\alpha}(x_1) H_{2n-2q+\beta}(x_2) \Big\rangle$ further by means of (10). These coefficients should be calculated numerically by a direct multiplication of the two Hermite sums.

[1] Flory, 1976.

[2] Yogurtcu *et al.* (2009).