A Fermi resonance and a parallel-proton-transfer overtone in the Raman spectrum of linear centrosymmetric N_4H^+: A polarizability-driven first principles molecular dynamics study

Oluwaseun Omodemi,[1] Ramsay Revennaugh,[1] Janiyah Riley,[1] Alexey L. Kaledin,[2,*] Martina Kaledin[1,*]

[1] *Department of Chemistry & Biochemistry, Kennesaw State University, 370 Paulding Ave NW, Box # 1203, Kennesaw, GA 30144, United States of America*

[2] *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, Georgia, 30322, USA*

Abstract

We present molecular dynamics (MD), polarizability driven MD (α-DMD) and pump-probe simulations of Raman spectra of the protonated nitrogen dimer N_4H^+, and some of its isotopologues, using the coupled-cluster single double triple CCSD(T)-F12b/aug-cc-pVTZ based potential energy surface in permutationally invariant polynomials (PIP) due to Yu *et al.* [J. Phys. Chem. A **119**, 11623, (2015)] and a corresponding PIP-derived CCSD(T)/aug-cc-pVTZ-tr (N:*spd,* H:*sp*) polarizability tensor surface (PTS), the latter reported here for the first time. To represent the PTS in terms of a PIP basis, we utilize a recently described formulation for computing the polarizability using a many-body expansion in the orders of dipole-dipole interactions while generating a training set using a novel approach based on linear regression for potential energy distributions. The MD/α-DMD simulations reveal: (i) a strong Raman activity at 260 and 2400 cm^{-1}, corresponding to the symmetric N−N···H bend and symmetric N−N stretch modes, respectively; (ii) a very broad spectral region in 500-2000 cm^{-1} range assignable to the parallel N···H^+···N proton transfer overtone, and (iii) the presence of a Fermi-like resonance in the Raman spectrum near 2400 cm^{-1} between the Σ_g^+ N−N stretch fundamental and the Π_u overtone corresponding to perpendicular N···H^+···N proton transfer.

Authors to whom correspondence should be addressed:
akaledi@emory.edu, mkaledin@kennesaw.edu

I. INTRODUCTION

Fermi resonances[1] appear both in IR and Raman spectra of a wide variety of systems, such as CO_2,[2-7] protonated water clusters,[8-13] asymmetric proton bound dimers,[14-16] or amines.[17,18] Since the IR and Raman spectroscopies provide information about the presence or absence of specific functional groups, they are widely used to identify structures of compounds.[19] For a vibrational mode to be IR active, there must be change in the dipole moment; Raman spectroscopy depends on change in the polarizability of a molecule. For linear centrosymmetric molecules, Raman and IR are mutually exclusive; therefore, Raman spectroscopy is considered to be complementary to IR spectroscopy. As such, Raman spectra can provide additional information on symmetric fundamental vibrations and anharmonic features such as combination bands and overtones that are IR inactive.[19]

Quantum chemical studies have routinely been used to interpret experimentally measured vibrational spectra. The standard approach based on the quantum mechanical double-harmonic approximation (quadratic potential and linear dipole/polarizability)[20] is generally considered inadequate for non-rigid systems, e.g., when large amplitude non-harmonic proton motion is involved,[21-35] which is the subject of the present work. Alternatively, molecular dynamics (MD) simulations and quantum variational calculations of vibrational structure traditionally serve as high-end solutions to the non-harmonicity problem.[36] Apart from the daunting task of running direct molecular dynamics simulations[15,30,37] or solving large-scale secular equations by means of direct vibrational configuration interaction (VCI) approaches,[38,39] surface fitting using permutationally invariant polynomials (PIPs), often in conjunction with machine learning methods, is a major alternative approach that has produced remarkable results for a variety of applications.[40-46] In many other applications, the PIP fitting tools have also been used for

generating accurate potential energy surfaces (PESs),[40] dipole moment surfaces (DMSs),[47] and polarizability tensor surfaces (PTSs).[48,49] Developing such surfaces by fitting an analytic (parameterized) polynomial function to a set of *ab initio* data is now possible for small, moderate, and as recently was reported fairly large polyatomic molecules.[31,40,43,47-51] Furthermore, user-friendly software that produces both the energy and its gradient in the form of F90 subroutines, called the Monomial-Symmetrization-Approach (MSA),[52] is now available for automatic and completely generic PIP fitting to a set of user-generated *ab initio* data points. Our calculations, described below, are partly based on the MSA software.

In our recent series of publications, we have described the spectra of linear centrosymmetric N-containing proton-bound complexes[15,30] $N_2H^+\cdots OC$ and N_4H^+ that are relevant to the interstellar clouds.[35,53-55] Since the nitrogen forms hydrogen bonds, N-H (a single bond) and N⋯H (a hydrogen bond) vibrations show up in the spectra as distinct features. Verdes *et al.* were first to report a high-resolution IR spectrum of the N_4H^+ ion and confirmed its linear centrosymmetric geometry.[34] The spectral feature at 2352 cm^{-1} was assigned to the asymmetric N−N stretch vibration.[34] It was concluded that the potential energy surface of N_4H^+ is strongly anharmonic with respect to the parallel proton stretching motion. Their reduced-dimensionality 2D and 3D quantum mechanical models predicted the parallel H^+ vibration (Σ_u^+ mode in the $D_{\infty h}$ point group) to be at 875 and 783 cm^{-1}, respectively. On the other hand, the simple harmonic approximation yields the frequency of 159 cm^{-1} at the highly respectable coupled-cluster single double triple CCSD(T)/aug-cc-pVTZ level of theory.[35] The true potential is obviously dominated by quartic and possibly higher terms in the parallel proton transfer motion. To address the issues stated above, Ricks *et al.* measured argon-tagged infrared spectra of protonated nitrogen dimer and its deuterium isotopologue in the 700-4000 cm^{-1} range.[56] Tagging the linear centrosymmetric N_4H^+ with an argon atom reduces the

point group symmetry to C_{2v} where the argon weakly binds to H^+ perpendicularly to the N_4H^+ molecular axis by 836 cm^{-1} [calculated at the CCSD(T)/aug-cc-pVTZ level of theory].[30,56] Ricks *et al.* assigned one of the more intense IR spectral features at 743 cm^{-1} to the parallel proton stretch vibration. Another region of IR activity in the N_4H^+·Ar experimental spectrum at 2349 cm^{-1} was assigned to the N−N asymmetric stretch. Other vibrational features, such as combination bands and overtones, were also IR active in the presence of the argon messenger atom. However, by symmetry rules,[20] these modes are IR inactive in the N_4H^+ bare ion.

Vibrational SCF and CI (VSCF/VCI) quantum studies using five-mode coupling[31] based on semiglobal and highly accurate analytical representations of the potential energy surface (PES) and dipole moment surface (DMS), presently referred to as **N4H+_PES** and **N4H+_DMS**, at the CCSD(T)-F12b/aug-cc-pVTZ and MP2/aug-cc-pVTZ levels of theory, respectively, were found to be in good agreement with the experimental IR measurements.[56] Another study of N_4H^+/N_4D^+ by Liao *et al.* using p-H_2 and n-D_2 matrix isolation IR absorption spectroscopy[57] in conjunction with discrete-variable representation (DVR) calculations on a reduced dimensionality derived PES, albeit at a lower level of electronic structure theory (CCSD/aug-cc-pVDZ), and its associated DMS using a four-mode coupling scheme (the $N_2\cdots H^+\cdots N_2$ symmetric stretch, the asymmetric "parallel proton transfer" $N\cdots H^+\cdots N$ stretch and two $N\cdots H^+\cdots N$ bending "perpendicular proton transfer" modes) provided very useful spectral assignments. However, some of the low-frequency $N-N\cdots H$ bending modes and the high-frequency N−N stretch modes were notably overestimated compared to the experimental measurements.[56,57] The above calculations suggest that dimensionality of the N_4H^+ vibrational Hamiltonian (the full dimension is 10) and the quality of the PES, i.e. the treatment of electron correlation, are very important for this system, thus presently we are highly motivated to use MD simulations, in full 3N Cartesian dimensions, using the best

available analytical representation of surface **N4H+_PES**. With this work, we build up on our recent N_4H^+ theoretical study,[30] where we examined the argon tag effects on the IR vibrational spectra using direct *ab initio* MD at the MP2/aug-cc-pVDZ level of theory. In that work, we also carried out driven molecular dynamics (DMD) simulations on the bare N_4H^+ ion using the PES and DMS surfaces and identified multiple anharmonic spectral features, which were in good agreement with quantum studies[31] and experimental measurements.[56,57]

We briefly remind that DMD has been shown to be capable of identifying fundamentals, combination bands, overtones, and their anharmonic shifts, even though it is based on classical trajectories.[15,30,37,58,59] Recently, we have expanded applicability of the DMD method to description of Raman spectra.[49,58] As with the dipole-driven MD method (μ-DMD), Raman active transitions beyond the usual fundamentals can be detected with the polarizability-driven MD method

(α-DMD) by scanning or gauging over the driving frequency.[58] DMD's facility of assignment[48,49,58] can greatly improve analysis of low-frequency vibrations that are difficult to calculate accurately using the quantum approaches without using a high-mode representation,[26,37,57] as well as symmetric vibrations that are IR inactive, while Raman active. Furthermore, developing PTS in analytical form, as shown by various groups,[42,43,48,60] circumvents the highly computationally demanding evaluation of polarizability, and its gradient, "on the fly" during direct MD simulations.

In the present work we describe an approach to construct analytical Cartesian representation of the PTS in terms of PIPs to be referred as **N4H+_PTS**. We further present a method for generating a training set of *ab initio* data points obtained from MD simulations at an electronic structure level comparable to the existing PES and DMS. We explore DMD's capabilities to assign

Raman spectra beyond the harmonic limit (overtones and combination bands). Most importantly, we describe a Fermi resonance between symmetric N−N stretch fundamental mode and an overtone of an asymmetric N···H^+···N bend fundamental mode of the symmetry and use isotopic substitutions and elements of 2D spectroscopy[61,62] to examine the origins of the resonance.

II. COMPUTATIONAL METHODS

A. The polarizability tensor surface (PTS)

We begin by developing a new PTS for N_4H^+. As is known, dipole polarizability of a collection of *N* atoms may be expressed using the point-dipole model of Applequist *et al*.[63]

$$\alpha_{ij} = \sum_{p,q}^{N} [(\mathbf{A}^{-1} + \mathbf{T})^{-1}]_{pi,qj} \qquad (1)$$

where *p*, *q* are atom labels with respective Cartesian components *i*, *j*.; **A** is a diagonal 3*N* square matrix of configuration-dependent isotropic atomic polarizabilities, and **T** is the dipole field tensor (see below). To avoid the well-documented numerical instabilities associated with inverting the matrix in Eq. 1 and to allow representation of α_{ij} using polynomials in internal coordinates, it was recently suggested that Eq. 1 be written as a weighted sum of many-body interactions as,

$$\alpha_{ij} = \sum_{n=1} \lambda_n \alpha_{ij}^{(n)} \qquad (2)$$

where λ_n are geometry dependent weights or correction factors, and $\alpha_{ij}^{(n)}$ are configuration dependent *n*-body polarizability tensors, starting at the diagonal one-body term.[48] The above form is a direct result of approximating the inverse in Eq. 1 using the power series $(\mathbf{A}^{-1} + \mathbf{T})^{-1} \cong \mathbf{A} - \lambda_2 \mathbf{ATA} + \lambda_3 \mathbf{ATATA} - \cdots$, where **T** is the perturbation. For $\lambda_n = 1$, Eq. 2 is an exact power series expansion of the polarizability tensor, Eq. 1, in the limit of weakly interacting atomic dipoles (**T** →

0).[64] The introduction of $\lambda_n \neq 1$ was intended[48] to regulate the behavior of α_{ij} in the regions of configurational space with short interatomic distances where polarizability given by Eq. 1 is otherwise discontinuous.[63,65,66] The individual terms are the one-body, which is the isotropic dipole induction at each atom directly due to the electric field,

$$\alpha_{ij}^{(1)} = \delta_{ij}\lambda_1 \sum_p \alpha_p \qquad (3a)$$

the two-body term, which is mutual dipole induction at *p(q)* due to the induced dipole at *q(p)*

$$\alpha_{ij}^{(2)} = -\lambda_2 \sum_{p\neq q} \alpha_p \alpha_q T_{pq}^{ij} \qquad (3b)$$

the three-body term, which is dipole induction at *p* due to the induced dipole at *s* induced by *q* (a *p-s-q* induction chain) or dipole induction at *p* due to the induced dipole at *s* induced by *p* itself (a *p-s-p* induction chain),

$$\alpha_{ij}^{(3)} = \lambda_3 \sum_{p,q} \alpha_p \alpha_q \sum_{s(\neq p,q)} \alpha_s \sum_{k=x,y,z} T_{ps}^{ik} T_{sq}^{kj} \qquad (3c)$$

the four-body term, correspondingly all possible *p-s-s'-q*, *p-s-s'-p*, *p-q-s'-q*, *p-q-s'-p*, *p-q-p-q* induction chains,

$$\alpha_{ij}^{(4)} = -\lambda_4 \sum_{p,q} \alpha_p \alpha_q \sum_{\substack{s(\neq p) \\ s'(\neq q,s)}} \alpha_s \alpha_{s'} \sum_{\substack{k=x,y,z \\ k'=x,y,z}} T_{ps}^{ik} T_{ss'}^{kk'} T_{s'q}^{k'j} \qquad (3d)$$

etc., with the *p*, *q*, *s*, *s'* summations running over the *N* atoms. We presently stop at the four-body limit and discuss its contribution to the full expression. The matrices $\mathbf{T}_{pq}$ are 3×3 dipole field tensors,

$$\left(\mathbf{T}_{pq}\right)_{ij} = \frac{\delta_{ij}}{r_{pq}^3} - 3\frac{r_i r_j}{r_{pq}^5} \qquad (4)$$

with r_{pq} being the magnitude of the p-q interatomic distance vector with the space-fixed Cartesian components $i, j = x, y, z$; and $\mathbf{T}_{pp}$=0. The isotropic atomic polarizabilities α_p and the correction factors are both parameterized functions of molecular geometry **r**,

$$\alpha_p(\mathbf{r}) = \alpha_p^{(0)} + \sum_{m=1}^{M_1} c_m^{(g_p)} u_{p,m}^{\mathrm{cv}}\{y_{ab}\} \qquad (5a)$$

$$\lambda_n(\mathbf{r}) = 1 + \sum_{m=1}^{M_2} c_{n,m}^{(\lambda)} u_m^{\mathrm{iv}}\{y_{ab}\} \qquad (5b)$$

where $u_{p,m}^{\mathrm{cv}}$ and u_m^{iv} are covariant and invariant polynomials,[40] respectively, in internuclear distances that assure that the total polarizability α_{ij} does not change upon permutation of any pair of like nuclei. The $\alpha_p^{(0)}$ are free atom "permanent" polarizabilities, and $\lambda_1(\mathbf{r}) \equiv 1$. The linear expansion coefficients for the atomic polarizabilities $c_m^{(g_p)}$ are the same for atoms within the same nuclear group g_p. The variables y_{ab} are functions of the internuclear distance r_{ab}. In the present work we have used the $y = r^{-1}$ form. In short, Eq. 2 represents the 3N-6 dimensional PTS, in analogy to the well-established descriptors for the potential energy and dipole moment surfaces, PES and DMS, respectively.

While the conventional PIPs have been well described in the literature in the applications to PES fitting,[52] their covariant counterparts appearing in DMS fitting (and presently in PTS fitting) are less common and are generally not uniquely defined. Here we use a standard scheme to generate the $u_{p,m}^{\mathrm{cv}}$ polynomials in Eq. 5a using symmetrized linear combinations of products of the internuclear distance functions (monomials), as follows,

$$u_{p,m}^{\mathrm{cv}} = \sum_j b_{p,m,j} y_{12}^{a_{m,j}} y_{13}^{b_{m,j}} \cdots y_{(N-1)N}^{z_{m,j}} \qquad (6)$$

Equation 6 renders the m-th covariant PIP of the power $0 < a_{m,j} + b_{m,j} + \cdots + z_{m,j} \leq \mathrm{MAX}$ describing atom p, where MAX is the maximal power of the PIP representation. The binary coefficient, or phase factor, $b_{p,m,j} = \pm 1$ takes on one of these two values and is a function of the monomial index j. It is the only element responsible for covariance between a pair of identical nuclei p and q. That is, without symmetrization, or equivalently with $b_{p,m,j} = +1$ for all j in the summation, Eq. 6 is reduced to a conventional PIP. A properly chosen phase assignment ensures covariance of $u_{p,m}^{\mathrm{cv}}$, meaning that a permutation of two identical nuclei, e.g., p and q, leaves the electron density unchanged but results in a corresponding exchange of α_p and α_q so that the overall polarizability tensor is invariant under the said permutation. A complete procedure for determination of $b_{p,m,j}$ is described in the supplementary material.

B. Optimization of the training set using NVE trial sets

Presently, for fitting of the PTS, and in the future applications in general, our motivation is to construct a training set of configurations such that the distribution of their potential energies is as close to uniform as possible. In other words, the general requirement is that all configurations with potential energies up to some maximal value V_{max} be equally represented in the set. Our approach is to first generate a few trial sets using NVE ensembles, i.e., classical MD trajectories of same size, and then to "contract" them into a single set, while maximizing the criterion of the potential energy uniformity, as stated above,

$$g(V) = \sum_i b_i f_i(V) \tag{7}$$

That is, the contracted set $g(V)$ is a linear combination of the trial NVE sets $f_i(V)$ with the coefficients chosen to make $g(V)$ uniform. The total energies of each of the trial sets are selected

according to an exponential cooling schedule. The coefficients $\{b_i\}$ in Eq. 7 are solved using a linear regression. This procedure is described in detail in the supplementary material, and here, we show the final result for N_4H^+ on [0,6200] cm^{-1}, up to the dissociation limit of $N_2H^+ + N_2$ fragments for **N_4H^+_PES**, using five trial NVE sets with the total energies 388, 755, 1550, 3100 and 6200 cm^{-1}. The NVE trajectories were propagated using the **N_4H^+_PES** potential for 100 ps each with a time step of 1fs. From them, we constructed an optimized training set of 10 000 points, illustrated in Figure 1. Given the five trial NVE sets, the curve is the best uniform distribution in the least-squares sense. Up to ~4200 cm^{-1} the potential energies are nearly equally represented, while a smooth tapering of the tail occurs between 4500 and 6200 cm^{-1}. We then used the optimized training set configurations to calculate the polarizability tensors needed for Raman scattering analysis.

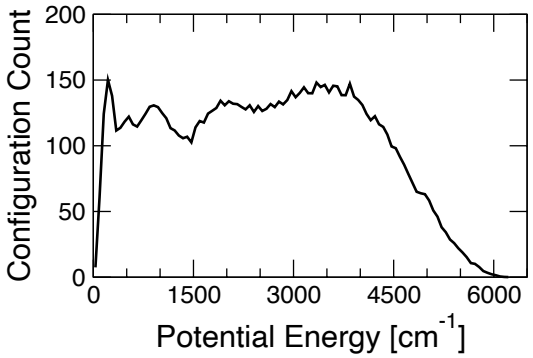


Figure 1. An optimized training set of length *M*=10 000 sourced from five trial NVE sets, each of length *N*=100001 configurations, for N_4H^+ calculated at the **N4H+_PES** level of theory. See supplementary material for all definitions and contraction procedure.

C. Calculation of Raman spectra from MD simulations

Calculation of Raman spectra involves decomposition of the full polarizability tensor into a spherical part $\bar{\alpha} \equiv (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3$ and a traceless anisotropic part $\beta_{ij} \equiv \alpha_{ij} - \delta_{ij}\bar{\alpha}$, for *i,j=x,y,z*. The respective polarized and depolarized components of the Raman spectrum are then given by[67]

$$I_{\mathrm{pol}}(\omega) = \frac{1}{\pi}\int_0^{\infty} dt\, e^{-i\omega t}\langle \overline{\alpha}(0)\overline{\alpha}(t)\rangle \qquad (8a)$$

$$I_{\mathrm{depol}}(\omega) = \frac{1}{\pi}\int_0^{\infty} dt\, e^{-i\omega t}\,\langle \sum_{i,j=x,y,z} \beta_{ij}(0)\beta_{ij}(t)\rangle \qquad (8b)$$

The brackets <...> involve the averaging over both the initial positions and momenta and over time by employing a signal window.[30,68,69] Additionally, we use the depolarization ratio[20,67] $\rho = 3I_{\mathrm{depol}}/\left(4I_{\mathrm{depol}} + 30I_{\mathrm{pol}}\right)$ as a tool to help identify weak resonances.

In total, 20 trajectories were generated randomly for isotopologues $^{14}N_4H^+$, $^{14}N_4D^+$ and $^{15}N_4H^+$. The starting point was the global minimum structure (Table S5) optimized on the analytical **N4H+_PES** trained on CCSD(T)-F12b/aug-cc-pVTZ data, and randomly selected velocities. Each trajectory was propagated as an NVE ensemble corresponding to the temperature of 200 K, up to 50 ps using the velocity-Verlet integrator.[70] Calculations of convergence tests showed that larger Δt time steps, e.g., 0.5 fs, resulted in appreciable peak position errors in the high-frequency end of the spectra, and much better results were obtained with a time step of Δt=0.2 fs. Furthermore, to better describe peak intensity in those high-frequency regions, as has been done previously,[69] the classically derived spectral functions I_{pol} (Eq. 8a) and I_{depol} (Eq. 8b) were additionally scaled by the quantum mechanical frequency-dependent factor $\kappa(\omega)=\omega/[1-\exp(-\omega/k_{\mathrm{B}}T)]$, where k_{B} is the Boltzmann constant.[71]

2.4 The α-DMD equations of motion

The equations of motion adapted presently for the DMD are

$$\dot{\mathbf{r}} = \mathbf{p}/m \qquad (9a)$$

$$\dot{\mathbf{p}} = -\nabla V(\mathbf{r}) + \mathbf{f}(\mathbf{r}, t) \qquad (9b)$$

where *m* are atomic masses, **p** and **r** are the 3*N* Cartesian momenta and coordinates, respectively, and *V*(**r**) is the molecular interaction potential. The time-dependent driving force is **f**(**r**,*t*), and its component with respect to Cartesian "*k*" on atom "*a*", omitting the non-resonant sum-frequency terms, is given by[58]

$$f_{a,k}(\mathbf{r},t) = -\frac{\varepsilon_0^2}{2}\sum_{ij}\left[P_i S_j \cos(\omega t) + \frac{P_i P_j + S_i S_j}{2}\right]\frac{\partial \alpha_{ij}(\mathbf{r})}{\partial r_{a,k}} \qquad (10)$$

where ε_0 is the field strength, P_i and S_j are the normalized incident and Stokes scattered light vector components, respectively, in the laboratory frame where *i, j = x, y, z*. The first term in the square brackets corresponds to vibrational energy absorption at the Stokes frequency shift ω, and the second term is responsible for rotational scattering, i.e., the molecule may be rotationally excited by the field. Note that the rotational scattering term does not depend explicitly on time but resonates with the molecular vibrational motion as a perturbation to the potential energy via the field-polarizability interaction. In the calculations below, we examine both the *parallel scattering* regime, i.e., **P** = **S**, and the *perpendicular scattering* regime where $\mathbf{P} \cdot \mathbf{S} = 0$, since the dynamical differences between the two can help identify difficult cases, such as overtones.

Atomic coordinates, forces, dipoles, polarizabilities, and total energies, dipole, and polarizability derivatives were collected along each trajectory to identify and assign Raman active frequencies. The average absorbed energy monitored along the trajectory indicates Raman activity and is defined as a time integral of the unperturbed molecular Hamiltonian H_0

$$I_{DMD}(\omega) = \frac{1}{t}\int_0^t H_0(\mathbf{q}(t'),\mathbf{p}(t');\omega)\,dt' \qquad (11)$$

For inactive modes, the average absorbed energy is small and oscillatory, while increases rapidly at Raman resonances.[58]

Raman mode assignment is done similar to what we have done previously for the IR of N_4H^+.[30] The symmetry coordinates, s_i were analyzed (see supplementary material information for the definition of s_i and additional information) and plotted as functions of time for the resonant trajectories. In addition, we also monitor averaged symmetrized displacements defined as a time integral of the difference of symmetry coordinates from their equilibrium values to better identify dominant vibrational motion,

$$\Delta s_i(t) = \sqrt{\frac{1}{t}\int_0^t [s_i(t') - s_i(0)]^2 \, dt'} \qquad (12)$$

III. RESULTS AND DISCUSSION

A. Molecular polarizability benchmark study

Here we describe the comparative analysis of the quantum mechanical methodologies in calculating the polarizability tensor α done with MP2, CCSD, and CCSD(T) methods with several basis sets: aug-cc-pVDZ (AVDZ), d-aug-cc-pVDZ (d-AVDZ), and cc-pVTZ (VTZ), and aug-cc-pVTZ excluding *d* functions on hydrogen and *f* functions on nitrogen atoms, respectively (AVTZ-tr), aug-cc-pVTZ (AVTZ), and d-aug-cc-pVTZ (d-AVTZ), respectively. We report atomic and molecular polarizabilities in $Å^3$ units and corresponding percent errors with the respect to the available experiment,[67,72] and high-level theoretical calculations[73] (Tables I and Tables S1-S4). All molecular polarizabilities were evaluated at the equilibrium geometries optimized at the CCSD(T)/AVTZ level of theory and C_{2v} symmetry transition state[30] of N_4H^+. The best results relative to the corresponding benchmark values of N_4H^+, N_2H^+, N_2 and $N(^4S)$ and considering the central processing unit (CPU) times required for a single point calculation were obtained with CCSD(T) and a truncated AVTZ-tr basis set. Such basis set reduction yields relatively small percent errors ~ 3% for N and less than 1% for N_2, N_2H^+, and N_4H^+. Also, for N_4H^+ the computer

times are manageable for generating the necessary data points (Table S4). Thus, we use this level of theory to calculate the polarizability values for the 10,000 training set points with the CCSD(T)[74,75] as implemented in MOLPRO.[76]

TABLE I. Benchmark calculations of the fragments, and a comparison of N_4H^+ non-zero polarizability tensor components[a] calculated at the *ab initio* CCSD(T)/AVTZ-tr level of theory and the analytical **N4H+_PTS**.

Atom/Molecule	α_{xx} (Å^3)	α_{yy} (Å^3)	α_{zz} (Å^3)	α_{iso} (Å^3)	α_{aniso} (Å^3)
Ab initio					
N	1.062	1.062	1.062	1.062	0.000
N_2	1.514	1.514	2.225	1.751	0.712
N_2H^+	1.151	1.151	2.085	1.462	0.934
N_4H^+ min, $D_{\infty h}$	2.434	2.434	5.536	3.468	3.102
N_4H^+ TS, C_{2v}	2.516	3.143	4.317	3.325	1.583
N4H+_PTS					
N_4H^+ min, $D_{\infty h}$	2.431	2.431	5.353	3.466	3.104
N_4H^+ TS, C_{2v}	2.427	3.333	4.290	3.350	1.613

[a] See Tables S1-S4 for additional information.

B. Fitting of the PTS

Presently, we use a PIP of order 5 to fit the PTS. The unknown coefficients $\{c_m\}$ of Eq. 5 are searched for by a large-scale L-BFGS minimization procedure[77] applied to a least-squares functional with the starting point of a non-interacting system of atoms $\{c_m = 0\}$. For large training sets we find it necessary to improve the performance of this approach by evaluating the least-squares function along with its gradient using shared-memory parallelization, as described in the SI.

We take the permanent atomic polarizabilities $\alpha_{\mathrm{N}}^{(0)} = 1.0617$ Å^3 calculated for N(^{4}S) at the CCSD(T)/aug-cc-pVTZ(N:*spd,* H:*sp*) level of theory and $\alpha_{\mathrm{H}}^{(0)} = 0$ for H^+. The total number of variables for the $\alpha_{ij}^{(3)}$ and $\alpha_{ij}^{(4)}$representations are 832 and 1040. Using the non-linear optimization procedure, the final converged RMSE are 0.0387 Å^3 (0.85%) and 0.0239 Å^3 (0.53%) for the three-body and four-body polarizability representations, respectively.

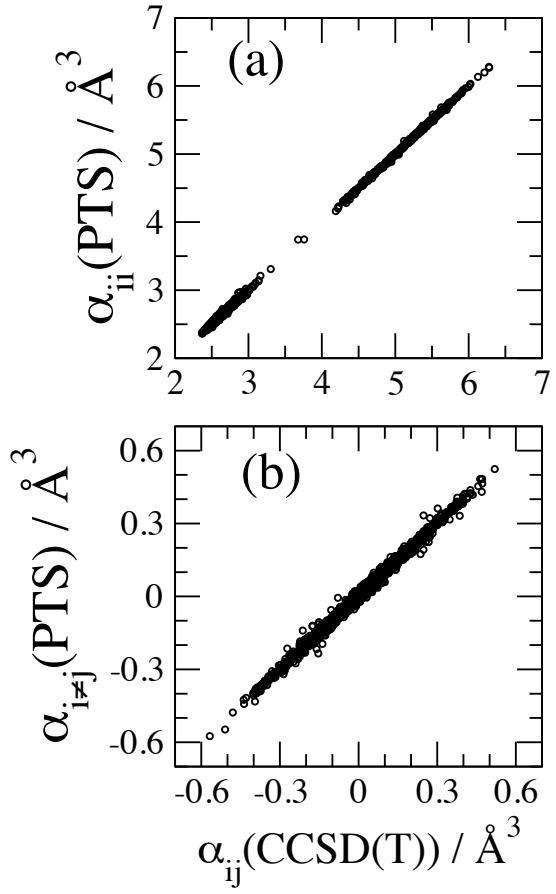


Figure 2. Correlation plot of the fourth order PTS fit using the training set of 10 000 points for (a) the diagonal elements and (b) for the off-diagonal elements of the (3×3) polarizability tensor. The RMSE of the fit is 0.024 Å^3 or ~0.53%.

As can be seen, the correction to the three-body representation made at the $\alpha_{ij}^{(4)}$ representation is ~37%, a valuable improvement considering only a marginal increase in the computational complexity. To note, the number of operations is $O(27N^3)$ in $\alpha_{ij}^{(3)}$ and $O(54N^4)$ in $\alpha_{ij}^{(4)}$ for the number of atoms $N = 5$. This result suggests that the four-body representation is quite suitable for small gas phase systems, similar to the one considered in the present case. In Figure 2 we show correlation plots of the training set points fitted with the $\alpha_{ij}^{(4)}$ representation. The uniformity of data spread around the target line (the diagonal) points to the more-or-less equal goodness of the fit for all the configurations on the broad energy range of [0,6200] cm^{-1}.

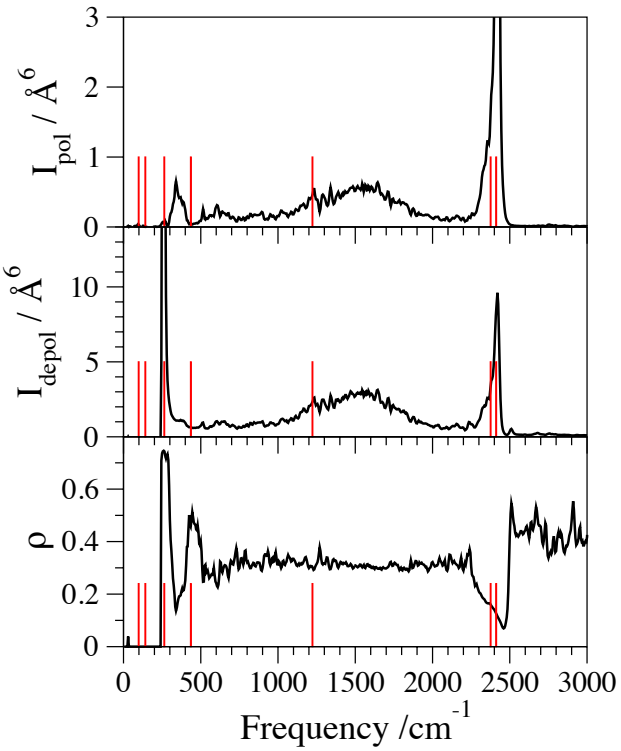


Figure 3. MD Raman spectra of $^{14}N_4H^+$ at 200 K. The polarized I_{pol} (top panel, Eq. 8a), depolarized I_{depol} (middle panel, Eq. 8b) components, and the depolarization ratio ρ (bottom panel) were calculated using the analytical **N4H+_PES** and the presently fitted **N4H+_PTS** surfaces. The harmonic frequencies are shown as red sticks.

C. **The MD simulations**

In the calculated spectra we identified the fundamentals (Figs. 3 and S5) and the broad region of the non-harmonic features (Figure 3 and Figure 4), roughly between 500 and 2000 cm^{-1} for the three isotopologues, that are complementary to the IR spectra studied previously using theoretical[30,31] and experimental methods[56,57] Tables S9-S11 in the supplementary material summarize the vibrational frequencies and point group symmetries obtained from a normal mode analysis. In addition, we present the Raman intensities calculated at the MP2/AVTZ level of theory as implemented in Gaussian.[78]

The N_4H^+ complex has three Raman active symmetric fundamental modes, namely, the N−N···H bend (ν_5, Π_g), $N_2\cdots H^+\cdots N_2$ stretch (ν_2, Σ_g^+), and N−N stretch (ν_1, Σ_g^+), with the frequency labels (ν_i) adopted from the experiment[56] and our previous work.[30] The corresponding harmonic frequencies on **N4H+_PES** are 264, 436, and 2412 cm^{-1} (Table S9), respectively. The MD spectrum calculated at 200 K shows two dominant peaks appearing on both the polarized and depolarized signals (Figure 3) and the peaks are assigned to ν_5 at 260 cm^{-1} and ν_1 at 2420 cm^{-1}., respectively. There is a weak peak in the polarized spectrum between 300 and 400 cm^{-1} that may be loosely assigned to the totally symmetric ν_2. Concurrently, the depolarization ratio curve clearly indicates a sharp drop near 345 cm^{-1}, suggesting the presence of a highly symmetric mode which we think is the $N_2\cdots H^+\cdots N_2$ stretch vibration, ν_2. High level VCI/VSCF calculations using MULTIMODE, 5MR level of treatment[31] predicted several modes, in particular, ν_5 = 260 cm^{-1} and a pair of two strongly mixed states with the energies 2335/2376 cm^{-1}, where one state is ν_1 although with some minor ambiguity of assignment. We partially reproduced these calculations using a lower level of theory: 3MR on a CCSD(T)/AVDZ surface in a direct MOLPRO implementation.[38,39,76] We found two highly mixed states of Σ_g^+ vibrational symmetry at 2268 cm^{-}

^{1} and 2340 cm^{-1}, with the corresponding eigenvectors for state-1: 19% |100000>(ν_1) + 44% |000002>($2\nu_6$), and for state-2: 50% |100000>(ν_1). This result appears to agree with the previously reported ones and suggests a presence of a strongly mixed and nearly resonant pair of a fundamental and an overtone. Aiding this issue to some degree, the IR experimental study by Ricks *et al*[56] made an indirect assignment of these IR-inactive modes by matching fundamentals and combination bands while assuming additivity of fundamentals, resulting in 240 cm^{-1} for ν_5 and 2228 cm^{-1} for ν_1.

The broad and weak intensity super band spanning 500-2000 cm^{-1} in the N_4H^+ Raman spectrum contains only one fundamental mode, ν_6, which is Raman inactive. The band is thus interpreted as a dense superposition of non-harmonic features such as symmetry allowed combination bands and overtones of any of the seven fundamentals listed in Table S9. For instance, any combination of the *gerade* modes and any even-valued combination of the *ungerade* modes will be Raman active.

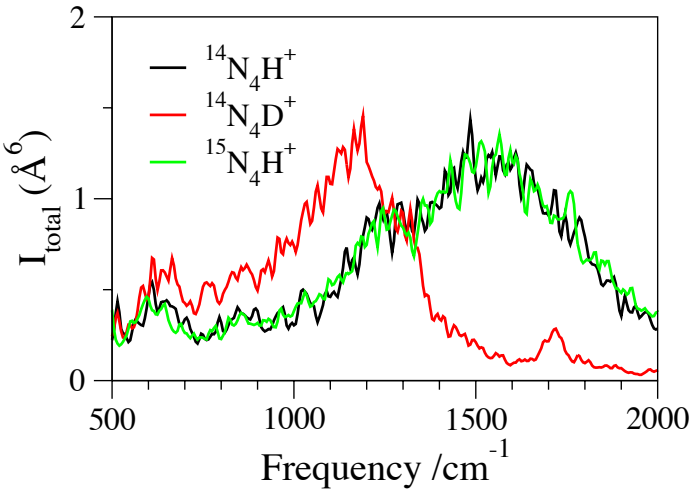


Figure 4. The Raman MD spectra (showing the total intensity) for $^{14}N_4H^+$,$^{14}N_4D^+$, and $^{15}N_4H^+$ at 200 K calculated with analytical **N4H+_PES** and the presently fitted **N4H+_PTS**.

Targeted application of isotopic substitution provided additional insights. We constructed $^{14}N/^{15}N$ and H/D isotopic analogs of N_4H^+ and repeated the MD simulations. The calculations reveal that the broad spectral feature is massively red shifted upon the H/D substitution, as can be seen in Figure 4**.** The $^{14}N/^{15}N$ isotopic shift revealed the most substantial changes in the symmetric N−N stretch being redshifted by ~90 cm^{-1}, the symmetric $N_2\cdots H^+\cdots N_2$ stretch redshifted by ~15 cm^{-1}, and the symmetric N−N···H bend redshifted by ~9 cm^{-1}. The broad band remained virtually unaffected by the heavy nitrogen; therefore, it is presently attributed to vibrations dominated by the proton motion. (See the supplementary material for more details on the isotope spectra.)

D. α-DMD simulations of Raman activity

Here we discuss mode assignment and the nature of non-harmonic spectral features identified in the MD Raman spectrum. Each DMD trajectory was propagated at a scanned frequency ω_n up to 10 ps with a 0.2 fs time step. We used the frequency resolution of $\Delta\omega_n$= 25 cm^{-1}. The starting point was always the global minimum structure optimized on the analytical **N4H+_PES** (Table S5) and zero velocities. The electric field strength is tuned to elicit appreciable absorption of energy at a given frequency range, i.e., some of the low frequency modes required a stronger field. An α-DMD spectrum is generated at the ω_n points from the averaged absorbed energies (Eq. 11) collected at the end of each of the driven trajectories.

First, we calibrate the field strength by examining the rate of energy absorption using a strongly Raman active mode as a reference. Several values of the field strength were tested, from 100 to 500 mV/bohr, and Figure 5 shows a well-behaved and clearly identifiable set of resonances in the high frequency symmetric N−N stretch range between 2200 and 2600 cm^{-1} with isotopic shifts upon $^{14}N/^{15}N$ and H/D substitutions and with field strength of 350 mV/bohr. The maximum

peak position is 2400 cm^{-1} for $^{14}N_4H^+$ and $^{14}N_4D^+$, and 2325 cm^{-1} for $^{15}N_4H^+$ isotope, which are near their harmonic values (Tables S9-S11). One can see that the total absorbed energy (I_{DMD}) at the end of the driving is roughly equal to the frequency of the vibrational resonance and well below the $N_2H^+ + N_2$ dissociation limit of ~6000 cm^{-1}.[30] This serves as a 'rule-of-thumb' for identifying true resonances while helping distinguish them from spurious ones.

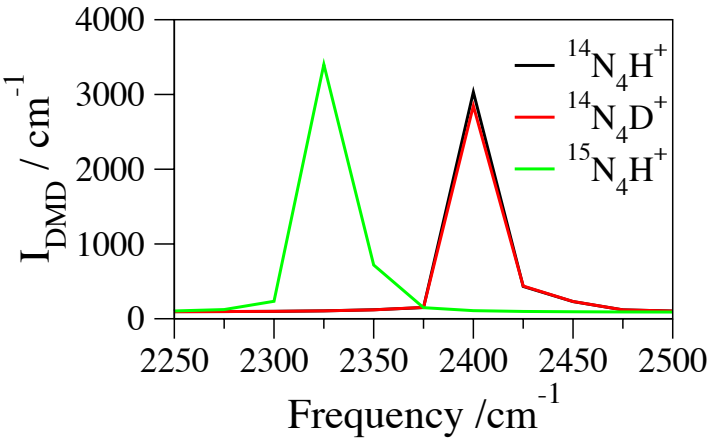


Figure 5. The parallel scattering α-DMD spectrum of $^{14}N_4H^+$,$^{14}N_4D^+$ and $^{15}N_4H^+$ in the high frequency range calculated with the analytical **N4H+_PES** and the presently fitted **N4H+_PTS**. The electric field strength is 350 mV/bohrs with the initial orientation along the molecular axis.

We proceed by investigating a possible Fermi resonance in the Raman regime between the *gerade* symmetric N−N stretch (ν_1) and the *ungerade* N···H^+···N bend overtone ($2\nu_6$). This overtone was identified, but not discussed, as a Fermi-like resonance in the previous experimental study[56] and the VSCF/VCI quantum vibrational study.[31] Yet, these two transitions, ν_1 and $2\nu_6$, are nearly degenerate in $^{14}N_4H^+$. The corresponding harmonic values on the analytical **N4H+_PES** are 2412 cm^{-1} and 2 × 1223 cm^{-1}, (Table S9), respectively, with a gap of only 34 cm^{-1} and are easily resolvable by the 10 ps DMD trajectories employed presently. Driving the polarizability at

the peak of the ν_1 DMD absorption, 2400 cm^{-1}, in a parallel scattering regime yields a highly suggestive interplay of intermodal energy transfer. Strong evidence of a Fermi resonance, in our interpretation of the DMD dynamics, is seen in Figure 6 (the black curve) and Figures S6-S8 as large N−N stretch displacements in the early stages of the driving 0-7 ps. The energy and the Δs_1 coordinate (the *gerade* N−N stretch, ν_1) are excited immediately at the onset of the electric field indicating a resonant absorption, while the coordinates characteristic of the ν_6 mode, namely, Δs_7 and Δs_9, are weakly excited or are unexcited at all. Their amplitudes remain small and only begin to sharply increase at the later stage of the trajectory, near the 7 ps mark, indicative of an onset of a second resonance. At exactly the same time the Δs_1 coordinate has stopped being excited, seen in Figure 6 as the flattening of the curve. In other words, after 7 ps of electric field exposure, the vibrational energy absorption in the ν_1 fundamental mode has been saturated, while simultaneously the ν_6 mode overtone, described primarily by Δs_7, has begun to absorb energy resulting in what we call a bi-modal excitation. That the energy is being continuously absorbed, as opposed to being oscillatory and partially released back to the field due to detuning, is suggested by the smoothly increasing absorbed energy curve. There is a question of whether the overtone absorption at 7 ps arises due to mode ν_6 directly interacting with the field (the direct absorption pathway) or via an efficient two-step mechanism: field => ν_1 => $2\nu_6$ (the resonant vibrational energy transfer "VET" pathway). In either case, beginning at the 7 ps mark both the ν_1 fundamental and the $2\nu_6$ overtone are excited and are interacting or exchanging energy, a signature of a Fermi resonance in the classical regime.

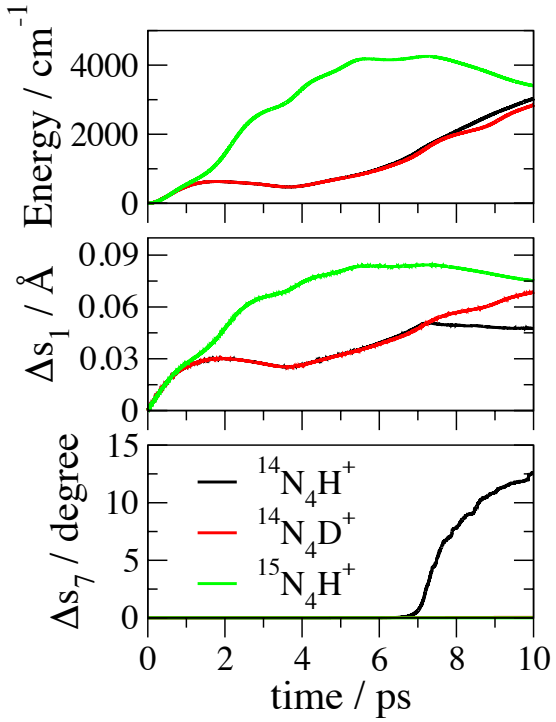


Figure 6. The average absorbed energy (top panel, Eq. 11) and averaged symmetrized displacements (Eq. 12) of the ν_1 fundamental (Δs_1, middle panel) and the $2\nu_6$ overtone (Δs_7, bottom panel) along each of the α-DMD trajectories driven at the corresponding symmetric N≡N stretch frequencies of $^{14}N_4H^+$ (2400 cm^{-1}), $^{14}N_4D^+$ (2400 cm^{-1}) and $^{15}N_4H^+$ (2324 cm^{-1}). The strength of the electric field was 350 mV/bohr in a parallel scattering regime.

Verifying the direct overtone absorption pathway, we scanned the frequencies slightly to the blue of the fundamental ν_1 peak, that is, in the 2400 and 2450 cm^{-1} range using the same field parameters: 350 mV/bohr in a parallel scattering regime. No resonances were identified in this case or with the stronger field, 700 mV/bohr. Switching to a perpendicular scattering regime, i.e., with **P** along the molecule axis and with **S** perpendicular to it, we identified a single peak at ~2420 cm^{-1} using the stronger field, 700 mV/bohr (Figure S10). Coordinate analysis revealed weakly excited stretching coordinates, ~0.05 Å, in the early stages 0-5 ps, followed by a sharp increase in the bending coordinates at 7 ps, suggesting the presence of a Raman active bending transition at

this frequency, which can only be the $2\nu_6$ overtone (Figure S11). Since it appears to scatter at 2420 cm^{-1} only in the perpendicular direction, the 2400 cm^{-1} parallel scattering trajectory described above must likely involve the VET $\nu_1/2\nu_6$ coupling pathway. From this we can also deduce that the level splitting should be of the order of ~20 cm^{-1}.

Nevertheless, it is instructive to test the VET hypothesis explicitly. To this end we employed elements of two-dimensional "pump-probe" spectroscopy. Using the 2400 cm^{-1} α-DMD trajectory as a pump acting to excite ν_1 up to ~7 ps, the time necessary to active the transition to $2\nu_6$, we 'turn on' the probe by switching off the electric field and recording (probing) the resulting dynamics. A probe duration of 2 ps was sufficient to observe a decay of the Δs_1 coordinate (the ν_1 fundamental) accompanied by rapid excitation occurring in the Δs_7 coordinate (the $2\nu_6$ overtone). These curves are shown in Figure S12. There is a clear indication of a rather fast vibrational energy flow from the excited fundamental to the overtone. Fitting an exponential decay formula to the fundamental mode coordinate: $\Delta s_1(t) = A\exp(-t/\tau) + B$ yielded a value of $\tau = 466$ fs for the lifetime (see Figure S12). If we assume a resonant energy level scheme and apply the golden rule treatment to this process, the lifetime can be used to find the Fermi resonance's coupling matrix element as $V = \sqrt{\hbar/2\pi\rho(E)\tau}$. The density of states can be approximated as the number of final states $N = 4$, ($\Pi_u \times \Pi_u = 2\Sigma_g + \Delta_g$), divided by the total energy absorbed by the trajectory E = 1500 cm^{-1}, as seen at the 7 ps mark of Figure 6, treating all other states below E as unavailable. Thus, writing $\rho(E) = N/E$ gives an upper bound for the coupling matrix element V = 26 cm^{-1}. The high-quality quantum mechanical calculations of Yu *et al.*[31] give the level splitting δE = 41 cm^{-1} or equivalently $V = \delta E / 2$ = 20 cm^{-1}, assuming a two-level interaction picture. This pump-probe derived estimate appears to agree well with the quantum result and the DMD calculated peak difference described above.

As a control simulation of the presence of the Fermi resonance, we employ isotopic substitution, namely, for the isotopic substituted species $^{14}N_4D^+$ and $^{15}N_4H^+$ such a mode mixing along with an energy transfer interplay is not seen as taking place (Figs. 6, and S9 red and green curves, respectively). In the $^{14}N_4D^+$ case, the symmetric N−N stretch fundamental is virtually unchanged while the N-N-D bend overtone is expected to be at $2 \times 893 = 1786$ cm^{-1} (the harmonic values, Table S9), a ~600 cm^{-1} gap. In the $^{15}N_4H^+$ case, the symmetric N−N stretch is isotopically shifted to 2330 cm^{-1} and $2\nu_6 = 2 \times 1220 = 2440$ cm^{-1} (the harmonic values, Table S10) a 110 cm^{-1} gap. If we drive these systems at their respective ν_1 fundamental frequency, corrected for anharmonicity by performing several exploratory drives, we see evidence of regular uni-modal excitation profiles. That is, the respective Δs_1 coordinates corresponding to ν_1 in both isotopologues, $^{15}N_4H^+$ and $^{14}N_4D^+$ are increasing relatively smoothly, without indication of an abrupt change. Meanwhile, the bend coordinates Δs_5-s_{10} are unexcited and the complementary stretch coordinates (Δs_2-s_4) are slightly excited very early but remain stationary for the duration of the driven trajectories. These calculations strongly suggest that the $\nu_1/2\nu_6$ resonance is absent in the two isotopologues.

As a final remark, we note that our α-DMD simulations do in fact reveal the Raman active overtone of the asymmetric N···H$^+$···N stretch (ν_4) identified in the MD spectrum (Figure 4**)** as part of the broad spectral feature between 500-2000 cm^{-1} that shifts upon H/D isotopic substitution. We remind that in our previous N_4H^+ work[30] on assigning the IR spectrum using μ-DMD we identified the N···H$^+$···N asymmetric stretch fundamental at 775 cm^{-1}. The high level VSCF/VCI calculations[31] predicted it at 758 cm^{-1}. Attempting to find a characteristic resonance in the 500-2000 cm^{-1} feature of the Raman spectrum, we drive near the center of the broad band, which happens to roughly correspond to $2\nu_4$ at $2 \times 775 = 1550$ cm^{-1}. It is known that due to the very

small intensity of the band one must use a much stronger electric field to incite an inherently weak overtone transition, as we observed previously in related applications.[15,30] A 700 mV/bohr value proved to be sufficient. Figure 7 summarizes this calculation. We note an early bump in the absorbed energy accompanied by a similar bump in the $N\cdots H^{+}\cdots N$ displacement Δs_4. It is followed by the fast and monotonically increasing Δs_4 with the other Σ symmetry displacements, Δs_1-s_3, remaining smaller by an order of magnitude. The angular displacements are virtually zero suggesting that the Π symmetry modes are not being excited. Only after ~13 ps of field exposure when the $2\nu_4$ overtone has clearly been excited, do the other modes begin to follow as all other possible combination bands begin to respond to the field. This result, in conjunction with our isotope-derived conclusions stated above, provides solid evidence that the 500 and 2000 cm^{-1} region is dominated by the parallel-proton-transfer bands. Based on VSCF/VCI calculations[31] and symmetry rules,[19] we expect additional N_4H^{+} Raman active combination bands and overtones between 500-2000 cm^{-1}, e.g. $\nu_4+\nu_7$, $\nu_2+\nu_5$, $\nu_6+\nu_7$, and $2\nu_2$.

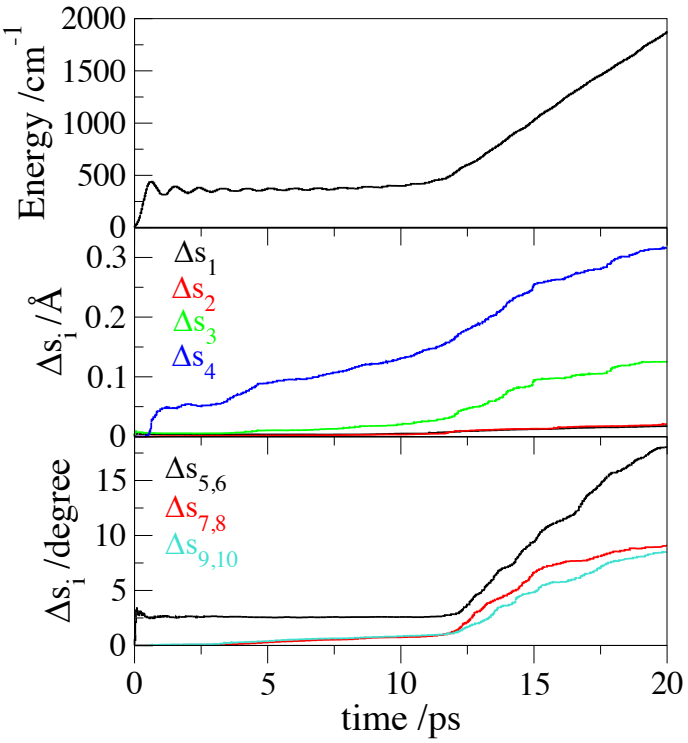


Figure 7. The average absorbed energy (top panel) and averaged symmetrized stretch displacements, $\Delta s_{1\text{-}4}$ (middle panel), and averaged symmetrized bend displacements, $\Delta s_{5\text{-}10}$ (bottom panel), (Eq. 12) along the parallel scattering α-DMD trajectory for the asymmetric N···H$^+$···N stretch overtone $2\nu_4$ of N_4H^+ driven at 1550 cm^{-1}. The intensity of the electric field is 700 mV/bohr.

IV. CONCLUSIONS

We have presented new calculations of the Raman spectra of the linear centrosymmetric N_4H^+ ion using first principles molecular dynamics simulations, namely, employing the previously published CCSD(T)-F12b/aug-cc-pVTZ potential energy surface (**N_4H^+_PES**) and a newly fitted CCSD(T)/aug-cc-pVTZ-tr polarizability tensor surface (**N_4H^+_PTS**). To represent the PTS in terms of a permutationally invariant polynomial (PIP) basis, we utilize a recently described formulation for computing the polarizability using a many-body expansion in the orders of dipole-dipole interaction pairs. We examined three- and four-body expansions, and showed that the four-

body representation, with the relative RMSE of 0.53%, is quite suitable for N_4H^+. Future developments related to PTS fitting are currently being investigated in our ongoing work on linear parameterization of PTS and associated developments on tensor parameterization approaches using linear regression techniques.

We have also described a new approach for generating training sets, for fitting PTS, DMS and PTS, by superposing and pruning (or contracting) trial sets obtained from NVE trajectories so that the distribution of potential energy points is as close to uniform as possible. This new approach allows to take the strategy of first using a low-level method for the generation of the extended NVE trial sets, e.g., at a DFT level, and then using the contracted set configurations to calculate the energy, gradient, polarizability, and its gradient at a higher level of theory, e.g., MP2 or CCSD(T).

The extensive MD and polarizability-driven (α-DMD) simulations reveal several major features in the Raman spectrum:

(i) A strong Raman activity at 260 cm^{-1} and at 2400 cm^{-1} corresponding to the symmetric fundamental N−N···H bend (Π_u) and symmetric fundamental N−N stretch (Σ_g^+) modes, respectively. Weaker activity was identified at 345 cm^{-1} and assigned to the fundamental symmetric $N_2\cdots H^+\cdots N_2$ (Σ_g^+) stretch vibration. These assignments are in close agreement with the experiment.

(ii) A very broad Raman active region on 500-2000 cm^{-1} assignable as a dense superposition of non-harmonic features involving proton transfer motion, such as symmetry allowed combinations ($\nu_4+\nu_7$, $\nu_2+\nu_5$, and $\nu_6+\nu_7$) and overtones ($2\nu_2$ and $2\nu_4$). The calculations suggest that this region is dominated by parallel proton transfer bands, such as one identified as the overtone of the N···H^+···N asymmetric stretch Σ_u^+ fundamental of 775 cm^{-1}.

(iii) Existence of a Fermi-like resonance in the Raman spectrum near 2400 cm^{-1} between one quantum of the Σ_g^+ N−N stretch and two quanta of the Π_u fundamental, corresponding to the perpendicular $N\cdots H^+\cdots N$ proton transfer. This resonance was interrogated in detail by two unrelated calculations, (1) differentiating the parallel and perpendicular Raman scattering regimes, which located, respectively, the fundamental excitation at 2400 cm^{-1} and the overtone excitation at 2420 cm^{-1}, suggesting the interaction of the order of 20 cm^{-1}; and (2) by setting up a two-dimensional (pump-probe) "experiment" which showed fairly efficient energy transfer from the excited fundamental to the overtone with the transition lifetime of 466 fs, corresponding to the (upper bound of) Fermi interaction matrix element of ~26 cm^{-1}. Both above estimates are in good agreement with the previously reported high-level VSCF/VCI calculations.

SUPPLEMENTARY MATERIAL

See supplementary material for additional information on formation of a training set by pruning trial data, least-square optimization procedure, N_4H^+ optimized geometry, harmonic frequencies, IR and Raman intensities, molecular polarizability data, MD spectra, driven molecular dynamics coordinate analysis, a covariant symmetrization scheme, and details of the polarizability tensor.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CHE–1855583. This work was also supported in part by research computing resources and technical expertise via a partnership between Kennesaw State University's Office of the Vice President for Research and the Office of the CIO and Vice President for Information Technology.[79]

ALK acknowledges the use of computational resources of the Cherry L. Emerson Center for Scientific Computation.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

References

(1) E. Fermi, Z. Phys. **71**, 250, (1931).

(2) S. Roy, P. J. Wrzesinski, D. Pestov, M. Dantus, and J. R. Gord, J. Raman Spectrosc. **41**, 1194, (2010).

(3) R. B. Wright and C. H. Wang, J. Chem. Phys. **58**, 2893, (1973).

(4) R. Lemus, M. Sánchez-Castellanos, F. Pérez-Bernal, J. M. Fernández, and M. Carvajal, J. Chem. Phys. **141**, 054306, (2014).

(5) G. Tejeda, B. Maté, and S. Montero, J. Chem. Phys. **103**, 568, (1995)

(6) M. Basire, F. Mouhat, G. Fraux, A. Bordage, J.-L. Hazemann, M. Louvel, R. Spezia, S. Bonella, and R. Vuilleumier, J. Chem. Phys. **146**, 134102, (2017).

(7) I. A. Verzhbitskiy, A. P. Kouzov, F. Rachet, and M. Chrysos, J. Chem. Phys. **134**, 194305, (2011).

(8) L.R. McCunn, J. R. Roscioli, M. A. Johnson, and A. B. McCoy, Phys. Chem. B **112**, 321, (2008).

(9) M. Kaledin, A. L. Kaledin, J. M. Bowman, J. Ding, and K. D. Jordan, J. Phys. Chem. A **113**, 7671, (2009).

(10) Q.-R. Huang, Y.-C. Li, T. Nishigori, M. Katada, A. Fujii, and J. Kuo, J. Phys. Chem. Lett. **11**, 10067, (2020).

(11) N. R. Samala and N. Agmon, Chem. Phys. **514**, 164, (2018).

(12) M. Kaledin, A. L. Kaledin, and J. M. Bowman, J. Phys. Chem A **110**, 2933, (2006).

(13) D. C. McDonald II; J.P. Wagner, A. B. McCoy, and M. A. Duncan, J. Phys. Chem. Lett. **9**, 5664, (2018).

(14) Q.-R. Huang, R. Shishido, C.-K. Lin, C.-W. Tsai, J. A. Tan, A. Fujii, and J. -L. Kuo, Angew. Chem., Int. Ed. **60**, 1936, (2021).

(15) D. Boutwell, O. Okere, O. Omodemi, A. Toledo, A. Barrios, M. Olocha, M. Kaledin, J. Phys. Chem. A **124**, 7549, (2020).

(16) Q.-R. Huang, T. Endo, S. Mishra, B. Zhang, L.-W. Chen, A. Fujii, L. Jiang, G. N. Patwari, Y. Matsuda, and J.-L. Kuo, Phys. Chem. Chem. Phys. **23**, 3739, (2021).

(17) G. Zundel, W.D. Lubos, and K. Kolkenbeck, Can. J. Chem, **49**, 3795, (1971).

(18) S. Mishra, H.-Q. Nguyen, Q.-R. Huang, C.-K. Lin, J.-L. Kuo, and G. N. Patwari, J. Chem. Phys. **153**, 194301, (2020).

(19) P. Atkins, J. de Paula, J. Keeler. *Atkins' Physical Chemistry*, 11th ed. (Oxford University Press, UK, 2018).

(20) E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular vibrations* (Dover, New York, 1980).

(21) H. J. Zeng and M. A. Johnson, Ann. Rev. Phys. Chem. **72**, 667, (2021).

(22) R. Roscioli, L. R. McCunn, and M. A. Johnson, Science **316**, 249, (2007).

(23) M. V. Vener and J. Sauer, Chem. Phys. Lett. **312**, 591, (1999).

(24) M. V. Vener, O. Kuhn, and J. Sauer, J. Chem. Phys. **114**, 240, (2001).

(25) Q. Yu and J. M. Bowman, J. Chem. Phys. **146**, 121102, (2017).

(26) J. Dai, Z. Bacic, X. Huang, S. Carter, and J. M. Bowman, J. Chem. Phys. **119**, 6571, (2003).

(27) T. K. Esser, H. Knorke, K. R. Asmis, W. Schollkopf, Q. Yu, C. Qu, J. M. Bowman, and J. Kaledin, J. Phys. Chem. Lett. **9**, 798, (2018).

(28) A.B. McCoy, X. C Huang, S. Carter, M. Y. Landerweer, and J. M. Bowman, J. Chem. Phys. **122**, 061101, (2005).

(29) C. H. Duong, O. Gorlova, N. Yang, P. J. Kelleher, M. A. Johnson, A. B. McCoy, Q. Yu, and J. M. Bowman, J. Phys. Chem. Lett. **8**, 3782, (2017).

(30) R. Hooper, D. Boutwell, and M. Kaledin, J. Phys. Chem. A **123**, 5613, (2019).

(31) Q. Yu, J. M. Bowman, R. C. Fortenberry, J. S. Mancini, T. J. Lee, T. D. Crawford, W. Klemperer, and J. S. Francisco, J. Phys. Chem. A **119**, 11623, (2015).

(32) K. Terrill and D. J. Nesbitt, Phys. Chem. Chem. Phys. **12**, 8311, (2010).

(33) M. Kaledin and C. A. Wood, J. Chem. Theory Comput. **6**, 2525, (2010).

(34) D. Verdes, H. Linnartz, J. P. Maier, P. Botschwina, R. Oswald, P. Rosmus, and P. J. Knowles, J. Chem. Phys. **111**, 8400, (1999).

(35) M. Mladenovic and E. Roueff, Astron. Astrophys. A **144**, 566, (2014).

(36) M. E. Tuckerman and G. J. Martyna, J. Phys. Chem. B **104**, 159, (2000).

(37) D. Boutwell, D. Pierre-Jacques, C. Tyler, J. Dyke, and M. Kaledin, J. Phys. Chem. A **126**, 583, (2022).

(38) M. Neff and G. Rauhut, J. Chem. Phys. **131**, 124129, (2009).

(39) M. Neff, T. Hrenar, D. Oschetzki, and G. Rauhut, J. Chem. Phys. **134**, 064105, (2011).

(40) B. J. Braams and J. M. Bowman, Int. Rev. Phys. Chem. **28**, 577, (2009).

(41) S. Manzhos and T. Carrington, Chem. Rev. **121**, 10187, (2021).

(42) G. M. Sommers, M. F. Calegari Andrade, L. Zhang, H. Wang, and R. Car, Phys. Chem. Chem. Phys. **22**, 10592, (2020).

(43) Y. Zhang, S. Ye, J. Zhang, C. Hu, J. Jiang, and B. Jiang, J. Phys. Chem. B **124**, 7284, (2020).

(44) A. Nandi, C. Qu, P. L. Houston, R. Conte, Q. Yu, and J. M. Bowman, J. Phys. Chem. Lett. **12**, 10318, (2021).

(45) P. L. Houston, C. Qu, A. Nandi, R. Conte, Q. Yu, and J. M. Bowman, J. Chem. Phys. **156**, 044120, (2022).

(46) Q. Yu, C. Qu, P. L. Houston, R. Conte, A. Nandi, J. M. Bowman, J. Phys. Chem. Lett. **13**, 5068, (2022).

(47) X. Huang, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **122**, 044308, (2005).

(48) O. Omodemi, S. Sprouse, D. Herbert, M. Kaledin, and A. L. Kaledin, J. Chem. Theory Comput. **18**, 37, (2022).

(49) O. Omodemi, M. Kaledin, and A. L. Kaledin, J. Comput. Chem. **43**, 1495, (2022).

(50) A. B. McCoy, B. J. Braams, A. Brown, X. Huang, Z. Jin, and J. M. Bowman, J. Phys. Chem. A **108**, 4991, (2004).

(51) A. Brown, A. B. McCoy, B. J. Braams, Z. Jin, and J. M. Bowman, J. Chem. Phys. **121**, 4105, (2004).

(52) Z. Xie and J. M. Bowman, J. Chem. Theory Comput. **6**, 26, (2010).

(53) D. Smith, Chem. Rev. **92**, 1473, (1992).

(54) E. Herbst, J. Payzant, H. Schiff, and D. Bohme, Astrophys. J. **201**, 603, (1975).

(55) E. Vigren, V. Zhaunerchyk, M. Hamberg, M. Kaminska, J. Semaniak, M. af Ugglas, M. Larsson, R. D. Thomas, and W. D. Geppert, Astrophys. J. **757**, 34, (2012).

(56) A. M. Ricks, G. E. Douberly, and M. A. Duncan, J. Chem. Phys. **131**, 1045312, (2009).

(57) H. -Y. Liao, M. Tsuge, J. A. Tan, J. -L. Kuo, and Y. -P. Lee, Phys. Chem. Phys. Chem. **18**, 20484, (2017).

(58) D. Pierre-Jacques, C. Tyler, J. Dyke, A. L. Kaledin, and M. Kaledin, Mol. Phys. **119**, e1939453, (2021).

(59) K. M. Christoffel and J. M. Bowman, J. Phys. Chem. **85**, 2159, (1981).

(60) G. R. Medders and F. Paesani, J. Chem. Theory Comput. **11**, 1145, (2015).

(61) A. Tokmakoff, M. J. Lang, D. S. Larsen, G. R. Fleming, V. Chernyak, and S. Mukamel, Phys. Rev. Lett. **79**, 2702, (1997).

(62) M. Kaledin, A. L. Kaledin, A. Brown, and J. M. Bowman, In Normal-Mode Analysis: Theory and Applications to Biological and Chemical Systems, Q. Cui, I. Bahar, Eds. (CRC Press: Boca Raton, FL, 2005).

(63) J. Applequist, J. R. Carl, and K. -K. Fung, J. Am. Chem. Soc. **94**, 2952, (1972).

(64) J. Applequist, P. Rivers, and D. E. Applequist, J. Amer. Chem. Soc. **91**, 5705, (1969).

(65) B. T. Thole, Chem. Phys. **59**, 341, (1981).

(66) P. T. van Duijnen and M. Swart, J. Phys. Chem. A **102**, 2399, (1998).

(67) D. A. McQuarrie, *Statistical Mechanics* (Harper & Row: New York, 1976).

(68) M. Kaledin, J. M. Moffit, C. R. Clark, and F. Rizvi, J. Chem. Theory Comput. **5**, 1328, (2009).

(69) M. Kaledin and D. T. Adedeji, J. Phys. Chem. A **119**, 1875, (2015).

(70) L. Verlet, Phys. Rev. **159**, 98, (1967).

(71) P. H. Berens and K. R. Wilson, J. Chem. Phys. **74**, 4872, (1981).

(72) T. M. Miller and B. Bederson, Adv. At. Mol. Phys. **13**, 1, (1977).

(73) Y. N. Kalugina and N. Cherepanov, N. Atmos. Ocean. Opt. **28**, 406, (2015).

(74) C. Hampel, K. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1, (1992).

(75) M. J. O. Deegan and P. J. Knowles. J. Chem. Phys. Lett. **227**, 321, (1994).

(76) H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, and R. Lindh, *et al.* MOLPRO, version 2019.2, a package of ab initio programs, https://www.molpro.net.

(77) D. Liu and J. Nocedal, Math. Program. B **45**, 503, (1989).

(78) Gaussian 16, Revision A.03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, and H. Nakatsuji. *et al.* Gaussian, Inc., Wallingford CT, 2016.

(79) Research Computing, Kennesaw State University, 2021, Digital Commons Training Materials. 10. https://digitalcommons.kennesaw.edu/training/10 (accessed August 3, 2022).

Supplementary Information

A Fermi resonance and a parallel-proton-transfer overtone in the Raman spectrum of linear centrosymmetric N_4H^+: A polarizability-driven first principles molecular dynamics study

Oluwaseun Omodemi,[1] Ramsay Revennaugh,[1] Janiyah Riley,[1] Alexey L. Kaledin,[2,*] Martina Kaledin[1,*]

[1] *Department of Chemistry & Biochemistry, Kennesaw State University, 370 Paulding Ave NW, Box # 1203, Kennesaw, GA 30144, United States of America*

[2] *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, GA, 30322, United States of America*

S.1 Formation of a training set by pruning extensive trial data

Our present task is to generate as diverse and global a set of geometries as possible using only classical trajectories. To this end, we may invoke the ergodicity hypothesis and propagate a "very long" NVE trajectory of energy E_{tot} and then select every n-th geometry to generate an arguably high-quality training set for PES/DMS/PTS fitting. Even if assuming the entire coordinate space has been visited by the trajectory, such a set would necessarily sample significantly more geometries with the higher potential energies, centered about ~$E_{tot}/2$, and "neglect" many of those with the lower potential energies. (One is referred to the classical harmonic oscillator.) The regions near the low energy stationary points, which contain chemically important structures, may be severely underrepresented in this set. To amend this, one may add to the training set a trajectory propagated for the same total time, but with a lower total energy, e.g., $E_{tot}/2$. Yet, simply combining these two sets from the two trajectories will put substantially more weight on the structures with the potential energies near the $E_{tot}/4$ region than on all other regions, making a combined set better than the original but still poorly balanced, and so on.

Below, we propose a pruning method to approximately eliminate these types of imbalances and make equal representation of configurations by the potential energy, i.e., assign appropriate weights to configurations in a consistent and simple manner. This prescription achieves equal representation of the points regardless of the energy of the configuration.

For a set of trial potential energy distributions $\{f_i(V)\}$ derived from microcanonical trajectories of different total energies, here assumed ergodic for the sake of argument, numbered i = 1, 2, ... each of length N time steps, we seek such coefficients in the "pruned" distribution

$$g(V) = \sum_i c_i f_i(V) \qquad S.1$$

that

$$\frac{\partial}{\partial \mathbf{c}} \int_0^{V_{max}} dV \left(g(V) - \alpha \frac{M}{V_{max}} \right)^2 = 0 \qquad S.2$$

where α is a constant to be determined by the $g(V)$ normalization requirement (see below), and where $M << N$ is the size of the pruned set. This condition ensures that $g(V)$ is as close to a uniform, or a top-hat, distribution as possible by assigning approximately equal weights to the

configurations with the potential energies on [0, V_{max}]. One may interpret $g(V)$ as a quasi-canonical ensemble depending on the nature of the mixing coefficients as functions of the total energy; this will be examined below.

Variation with respect to c_i yields the unique solution

$$\mathbf{c} = \alpha \frac{MN}{V_{max}} \mathbf{S}^{-1} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \qquad S.3a$$

or

$$c_i = \alpha \frac{MN}{V_{max}} \sum_j [\mathbf{S}^{-1}]_{ij} \qquad S.3b$$

where the elements of the overlap matrix **S** are

$$S_{ij} = \int_0^{V_{max}} dV f_i(V) f_j(V) \qquad S.4$$

We define the pruned set to be of size

$$M = \int_0^{V_{max}} dV g(V) = \sum_i c_i \int_0^{V_{max}} dV f_i(V) = N \sum_i c_i \qquad S.5a$$

with all trial distributions by construction containing the same number of points N,

$$\int_0^{V_{max}} dV f_i(V) = N \qquad S.5b$$

which gives by substitution of S.3b

$$M = N\alpha \frac{MN}{V_{max}} \sum_{ij} [\mathbf{S}^{-1}]_{ij} \qquad S.6a$$

or

$$\alpha = \frac{V_{max}}{N^2 \sum_{ij} [\mathbf{S}^{-1}]_{ij}} \qquad S.6b$$

Therefore, the coefficient of set "*i*" in the pruned set is

$$c_i = \frac{M}{N} \frac{\sum_j [\mathbf{S}^{-1}]_{ij}}{\sum_{ij} [\mathbf{S}^{-1}]_{ij}} \qquad S.7$$

independent of the energy cutoff parameter V_{max} and the grid on which $f_i(V)$ is defined. The number of points taken from set "*i*" is simply

$$M_i = N c_i \qquad S.8$$

while the normalization condition for the "pruned" set "*i*" holds true

$$\sum_i M_i = N \sum_i c_i = M \frac{\sum_{ij} [\mathbf{S}^{-1}]_{ij}}{\sum_{ij} [\mathbf{S}^{-1}]_{ij}} = M \qquad S.9$$

Finally, the frequency of pruned points (sampling intervals) in each trial set is

$$\nu_i = \frac{N}{M_i} = c_i^{-1} \qquad S.10$$

which is to say, we pick every ν_i-th point from each of the corresponding trial sets "*i*" to build the pruned set of size *M* that will best resemble a uniform distribution of potential energy points on [0, V_{max}]. Note that the relative weights, defined by the ratios c_i/c_j, formally depend only on the overlap matrix, i.e., the shapes of the trial distributions. However, a certain amount of noise in the pruned distribution will be present due to a somewhat arbitrary way of selecting the M_i points using formula S.10. Still, for large values of *M* and *N* the amount of noise is expected to be insignificant, and a greater number of trial sets will produce a better fit to a uniform distribution.

And since most computing facilities and workstations are equipped with multiple processors, propagation of several NVE trajectories can be easily carried out in parallel without consuming excessive wall-clock time.

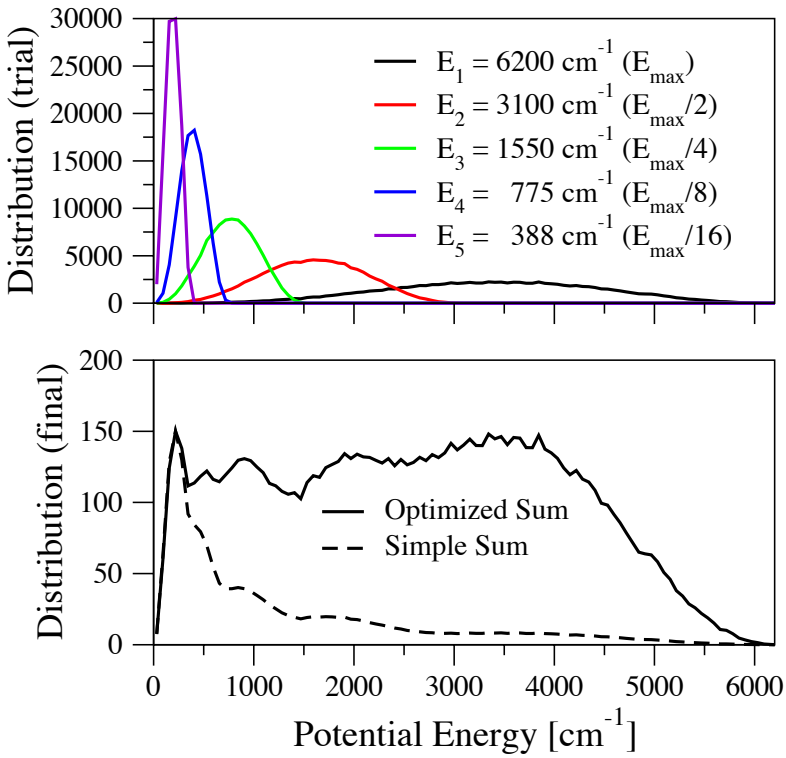


Figure S1. A five-trial-set contraction scheme of length N=100001 (each set), of N_4H^+ calculated at the N4H+_PES level of theory, to a pruned set of size M=10000 using the (a=1, b=2, E_{max}=6200 cm^{-1}) schedule. The pruning intervals are 15, 60, 96, 273 and 283 for the E_1, E_2, E_3, E_4 and E_5 trial sets, respectively. Note the difference in the shapes of the Simple Sum (unpruned) set and the Optimized Sum (pruned) set.

Our approach to trial set generation is based on partitioning the total energy of the system in reference to the harmonic ZPE, a dissociation limit or some key transition state, with the NVE trajectory total energies selected according to the schedule,

$$E_i = \frac{aE_{max}}{b^{i-1}}, \qquad i = 1,2,\ldots \qquad S.11$$

where the choice of (a, b, E_{max}) is guided by particular goals of the fit and system's properties. For example, (a=1, b=2, E_{max}=1/2 ZPE) is expected to produce a trial distribution suitable for most

MD applications at low/room temperature, while (a=3, b=1.5, E_{max}=3/2 ZPE) yields a set suitable for both high temperature MD and quantum vibrational calculations, and so on. Figure S1 illustrates this approach with four trial sets of 5001 points each to make a pruned set of 1000 optimally selected points. Unlike the highly imbalanced Simple Sum distribution, the pruned distribution has a well-defined top-hat shape.

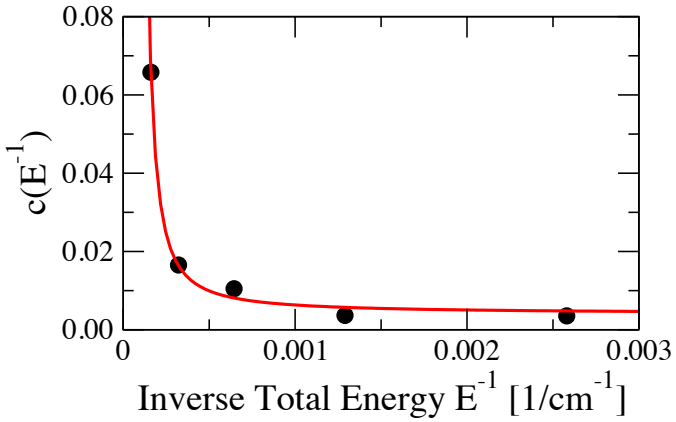


Figure S2. Trial set mixing coefficients (open circles) appearing in Eq. S.1 and given by Eq. S.7. The total energies (inverted) are those seen in Figure S1. The red line is an exponential fit to these points of the form $c(E^{-1}) = c(\infty)\exp(1/(TE^{-1}))$, with the characteristic temperature $T = 2200$ cm^{-1}.

Analysis of the mixing coefficients c_i, which are related to the pruned points per set M_i, suggests an increasing exponential function with the total energy of the microcanonical trial ensembles. Or, in other words, the nature of the pruned ensemble $g(V)$ may be interpreted as an "inverse" Boltzmann distribution where the lower energy trial NVE ensembles make exponentially smaller contribution to the pruned ensemble. This is exposed in Figure S2 on the distributions presented in Figure S1.

S.2 Atomic and molecular polarizability benchmark study

The benchmark study for molecular polarizabilities (Eq. S.12) of N (^{4}S) atom and N_2, N_2H^+, N_4H^+ molecules was carried out using the MOLPRO program. The polarizabilities were calculated numerically by the differentiation with the respect to the electric field. MP2, CCSD, and CCSD(T) methods were tested with various basis sets. All polarizabilities for N_2, N_2H^+, N_4H^+ were evaluated

at the CCSD(T)/aug-cc-pVTZ optimized geometries and compared to the to the available experiment[1,2] and high-level theoretical calculations.[3] In the N_4H^+ case, we used the reference point at the CCSD(T)/d-aug-cc-pVTZ level of theory. Computer times were reported for the single point molecular polarizability calculation on a single processor (Table S4).

$$\overline{\alpha} = \alpha_{\mathrm{iso}} = (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3 \qquad (S.12a)$$

$$\alpha_{\mathrm{aniso}} = \frac{1}{\sqrt{2}}\left[\left(\alpha_{xx} - \alpha_{yy}\right)^2 + \left(\alpha_{yy} - \alpha_{zz}\right)^2 + (\alpha_{zz} - \alpha_{xx})^2 + 6\left(\alpha_{xy}^2 + \alpha_{xz}^2 + \alpha_{yz}^2\right)\right]^{1/2} \qquad (S.12b)$$

Table S1: Error analysis for the N (^{4}S) atomic polarizabilities.

Method	α/ Å^3	% error
MP2/AVDZ	1.009	8.243
MP2/dAVDZ	1.073	2.436
MP2/VTZ	0.720	34.519
MP2/AVTZ-tr	1.052	4.374
MP2/AVTZ	1.055	4.061
MP2/dAVTZ	1.068	2.914
CCSD/AVDZ	1.008	8.336
CCSD-dAVDZ	1.075	2.259
CCSD/VTZ	0.724	34.214
CCSD/AVTZ-tr	1.054	4.159
CCSD/AVTZ	1.063	3.400
CCSD/dAVTZ	1.076	2.198
CCSD(T)/AVDZ	1.013	7.881
CCSD(T)/dAVDZ	1.082	1.611
CCSD(T)/VTZ	0.723	34.285
CCSD(T)/AVTZ-tr	1.062	3.476
CCSD(T)/AVTZ	1.070	2.771
CCSD(T)/dAVTZ	1.084	1.485
CCSD(T)/AVQZ	1.081	1.723
CCSD(T)/AV5Z	1.080	1.795
Experiment[1]	1.100	0.000

Table S2: Error analysis for the N_2 molecular polarizabilities.

Method	$\alpha_{xx,yy}$ / Å^3	α_{zz}/ Å^3	$\bar{\alpha}$/ Å^3	% error
MP2/AVDZ	1.475	2.135	1.695	3.701
MP2/dAVDZ	1.506	2.153	1.722	2.184
MP2/VTZ	1.105	2.005	1.405	20.159
MP2/AVTZ-tr	1.494	2.139	1.709	2.878
MP2/AVTZ	1.505	2.136	1.715	2.534
MP2/dAVTZ	1.510	2.141	1.720	2.248
CCSD/AVDZ	1.481	2.199	1.720	2.266
CCSD-dAVDZ	1.511	2.218	1.747	0.743
CCSD/VTZ	1.102	2.081	1.428	18.858
CCSD/AVTZ-tr	1.496	2.202	1.731	1.627
CCSD/AVTZ	1.502	2.197	1.734	1.492
CCSD/dAVTZ	1.507	2.201	1.738	1.240
CCSD(T)/AVDZ	1.497	2.224	1.739	1.200
CCSD(T)/dAVDZ	1.529	2.243	1.767	0.381
CCSD(T)/VTZ	1.107	2.092	1.436	18.437
CCSD(T)/AVTZ-tr	1.514	2.225	1.751	0.524
CCSD(T)/AVTZ	1.520	2.218	1.753	0.408
CCSD(T)/dAVTZ	1.525	2.222	1.758	0.130
Experiment[2]	1.450	2.380	1.760	0.000

Table S3: Error analysis for the N_2H^+ molecular polarizabilities.

Method	$\alpha_{xx,yy}$ / Å^3	α_{zz}/ Å^3	$\bar{\alpha}$/ Å^3	% error
MP2/AVDZ	1.147	2.020	1.438	1.017
MP2/dAVDZ	1.153	2.025	1.444	0.610
MP2/VTZ	0.985	2.007	1.326	8.746
MP2/AVTZ-tr	1.144	2.016	1.435	1.249
MP2/AVTZ	1.153	2.017	1.441	0.788
MP2/dAVTZ	1.149	2.018	1.438	0.997
CCSD/AVDZ	1.150	2.076	1.458	0.392
CCSD-dAVDZ	1.155	2.081	1.464	0.780
CCSD/VTZ	0.984	2.067	1.345	7.408
CCSD/AVTZ-tr	1.144	2.073	1.454	0.068
CCSD/AVTZ	1.153	2.073	1.459	0.457
CCSD/dAVTZ	1.148	2.073	1.456	0.228
CCSD(T)/AVDZ	1.156	2.090	1.468	1.027
CCSD(T)/dAVDZ	1.162	2.096	1.474	1.436
CCSD(T)/VTZ	0.988	2.077	1.351	7.024
CCSD(T)/AVTZ-tr	1.151	2.085	1.462	0.643
CCSD(T)/AVTZ	1.160	2.084	1.466	1.041
CCSD(T)/dAVTZ	1.155	2.084	1.465	0.815
CCSD(T)/AV5Z[3]	1.148	2.061	1.453	0.000

Table S4: Error analysis for the N_4H^+ molecular polarizabilities and computer times per single processor.

Method	$\alpha_{xx,yy}$ / Å^3	α_{zz}/ Å^3	α_{iso}/Å^3	% error	α_{aniso}/Å^3	% error	Time/ sec
MP2/AVDZ	2.403	5.297	3.368	2.985	2.894	6.031	34
MP2/dAVDZ	2.433	5.306	3.391	2.330	2.873	6.710	130
MP2/VTZ	2.026	5.216	3.089	11.016	3.191	3.619	91
MP2/AVTZ-tr	2.414	5.287	3.372	2.881	2.873	6.707	120
MP2/AVTZ	2.430	5.286	3.382	2.585	2.856	7.263	540
MP2/dAVTZ	2.428	5.285	3.380	2.634	2.858	7.199	1897
CCSD/AVDZ	2.409	5.485	3.434	1.068	3.076	0.122	578
CCSD-dAVDZ	2.438	5.493	3.457	0.429	3.055	0.801	2017
CCSD/VTZ	2.022	5.407	3.150	9.253	3.385	9.929	1240
CCSD/AVTZ-tr	2.415	5.473	3.434	1.074	3.058	0.701	1935
CCSD/AVTZ	2.427	5.463	3.439	0.932	3.035	1.424	7462
CCSD/dAVTZ	2.424	5.462	3.437	1.001	3.038	1.356	21312
CCSD(T)/AVDZ	2.427	5.553	3.469	0.064	3.126	1.517	1527
CCSD(T)/dAVDZ	2.458	5.562	3.493	0.615	3.104	0.790	5730
CCSD(T)/VTZ	2.032	5.460	3.174	8.557	3.429	11.349	4338
CCSD(T)/AVTZ-tr	2.434	5.536	3.468	0.096	3.102	0.733	4084
CCSD(T)/AVTZ	2.448	5.525	3.474	0.057	3.078	0.055	22115
CCSD(T)/dAVTZ	2.445	5.524	3.472	0.000	3.079	0.000	182229

S.3 N_4H^+ geometries and corresponding polarizability tensors

Figures and Tables below show N_4H^+ molecular geometries and corresponding polarizability tensors.

Figure S3. N_4H^+ linear geometry and atom numbering. The molecular axis is along Z.

Table S5: XYZ coordinates (Å) for the N_4H^+ global minimum structure optimized on the analytical **N_4H^+-PES** (CCSD(T)-F12b/aug-cc-pVTZ), ($D_{\infty h}$ symmetry).
The minimum energy is -219.03515571855 Hartree.

Atom	X	Y	Z
N	0.0000000000	0.0000000000	2.3736107464
N	0.0000000000	0.0000000000	-2.3736107464
N	0.0000000000	0.0000000000	1.2766043851
N	0.0000000000	0.0000000000	-1.2766043851
H	0.0000000000	0.0000000000	0.0000000000

Table S6: Polarizability, α_{ii} data in atomic units for N_4H^+ evaluated using analytical **N4H+_PTS** at the optimized geometry listed in the Table S5.

	X	Y	Z
X	16.3665705316	0.0000000000	0.0000000000
Y	0.0000000000	16.3665705316	0.0000000000
Z	0.0000000000	0.0000000000	37.1473788311

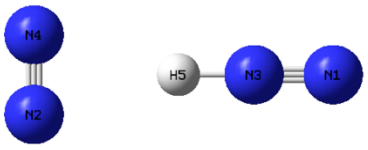
Table S7: XYZ coordinates (Å) for the N_4H^+ global minimum structure ($D_{\infty h}$ symmetry) optimized using MOLPRO at CCSD(T) /aug-cc-pVTZ level of theory.
The minimum energy is -218.98619507 Hartree.

Atom	X	Y	Z
N	0.0000000000	0.0000000000	2.3767595533
N	0.0000000000	0.0000000000	-2.3767595533
N	0.0000000000	0.0000000000	1.2758541851
N	0.0000000000	0.0000000000	-1.2758541851
H	0.0000000000	0.0000000000	0.0000000000

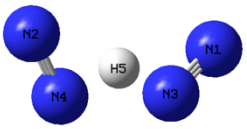
Table S8: Polarizability, α_{ii} data in atomic units for N_4H^+ evaluated at the optimized geometry listed in the Table S7.

X	Y	Z
16.517523384	0.000000000	0.000000000
0.000000000	16.517523385	0.000000000
0.000000000	0.000000000	37.286578143

C_{2v} (*194i*)
E=4741 cm^{-1}



D_{2d} (*554i, 554i*)
E=19887 cm^{-1}



D_{2h} (*73i, 325i, 825i*)
E=20188 cm^{-1}

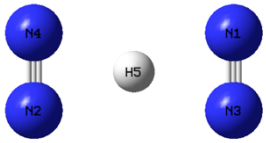


Figure S4. CCSD(T)/aug-cc-pVTZ N_4H^+ transition states. Imaginary frequencies (cm^{-1}) and energies relative the global minimum (cm^{-1}) are also given.

S.4 Harmonic frequencies, IR, and Raman intensities for N_4H^+, N_4D^+, and MD spectra

We run normal mode analysis at the MP2, CCSD(T) levels of theory, and analytic PES surface for various nitrogen and hydrogen isotopes, namely $^{14}N_4H^+$, $^{15}N_4H^+$, $^{14}N_4D^+$ to predict harmonic IR and Raman intensities and to identify $^{14}N/^{15}N$ and H/D isotopic substitution related frequency shifts. The corresponding Raman active harmonic frequencies on the analytical **N4H+_PES** for the N−N symmetric stretch in $^{14}N_4H^+$ vs. $^{15}N_4H^+$ (Tables S9 and S11) are 2412 and 2330 cm^{-1}, for $N_2\cdots H^+\cdots N_2$ stretch 436 and 421 cm^{-1}, and the symmetric N-N-H bend 264 and 255 cm^{-1}, respectively. Raman MD spectrum obtained using the analytical **N4H+_PTS** is shown on Figure S5. Figure 4 in the main text shows the Raman MD spectra in the range between 500-2000 cm^{-1}.

Table S9: $^{14}N_4H^+$ harmonic vibrational frequencies (in cm^{-1}), IR intensities (km/mol), and Raman intensities ($Å^4$/amu) calculated using optimized structure on the analytical PES, *ab initio* MP2, and CCSD(T) with aug-cc-pVTZ basis set. The frequency labels (ν_i) correspond to the experimental work[4] and our previous work.[5]

Symmetry Label	MP2 freq.	MP2 IR int.	MP2 Raman int.	Analytical N_4H^+-PES	CCSD(T) freq.	CCSD(T) IR int.
$\nu_7(\Pi_u)$	139	7.48	0.00	141	143	6.55
$\nu_5(\Pi_g)$	258	0.00	1.09	264	265	0.00
$\nu_2(\Sigma_g^+)$	438	0.00	3.84	436	438	0.00
$\nu_4(\Sigma_u^+)$	493	5170.42	0.00	99	159	5421.06
$\nu_6(\Pi_u)$	1227	86.36	0.00	1223	1235	86.98
$\nu_3(\Sigma_u^+)$	2212	241.51	0.00	2376	2366	270.21
$\nu_1(\Sigma_g^+)$	2246	0.00	32.33	2412	2402	0.00

Table S10: $^{14}N_4D^+$ harmonic vibrational frequencies (in cm^{-1}), IR intensities (km/mol), and Raman intensities ($Å^4$/amu) calculated using optimized structure on the analytical PES, *ab initio* MP2 and CCSD(T) with aug-cc-pVTZ basis set.

Symmetry Label	MP2 freq.	MP2 IR int.	MP2 Raman int.	Analytical N_4H^+-PES	CCSD(T) freq.	CCSD(T) IR int.
ν_7 ($^1\Pi_u$)	136	6.65	0.00	137	140	5.81
ν_5 ($^1\Pi_g$)	258	0.00	1.09	264	265	0.00
ν_2 ($^1\Sigma_g^+$)	438	0.00	3.84	436	439	0.00
ν_4 ($^1\Sigma_u^+$)	352	2596.84	0.00	70	114	2740.81
ν_6 ($^1\Pi_u$)	896	36.21	0.00	893	905	37.16
ν_3 ($^1\Sigma_u^+$)	2211	199.18	0.00	2376	2365	240.32
ν_1 ($^1\Sigma_g^+$)	2247	0.00	32.33	2412	2402	0.00

Table S11: $^{15}N_4H^+$ harmonic vibrational frequencies (in cm^{-1}), IR intensities (km/mol), and Raman intensities ($Å^4$/amu) calculated using optimized structure on the analytical PES, *ab initio* MP2 and CCSD(T) with aug-cc-pVTZ basis set.

Symmetry Label	MP2 freq.	MP2 IR int.	MP2 Raman int.	Analytical N_4H^+-PES	CCSD(T) freq.	CCSD(T) IR int.
ν_7 ($^1\Pi_u$)	134	7.03	0.00	136	138	6.17
ν_5 ($^1\Pi_g$)	250	0.00	1.02	255	256	0.00
ν_2 ($^1\Sigma_g^+$)	423	0.00	3.59	421	424	0.00
ν_4 ($^1\Sigma_u^+$)	493	5168.68	0.00	98	160	5419.59
ν_6 ($^1\Pi_u$)	1224	87.36	0.00	1220	1233	87.94
ν_3 ($^1\Sigma_u^+$)	2137	231.57	0.00	2296	2286	256.44
ν_1 ($^1\Sigma_g^+$)	2170.9	0.00	30.18	2330	2321	0.00

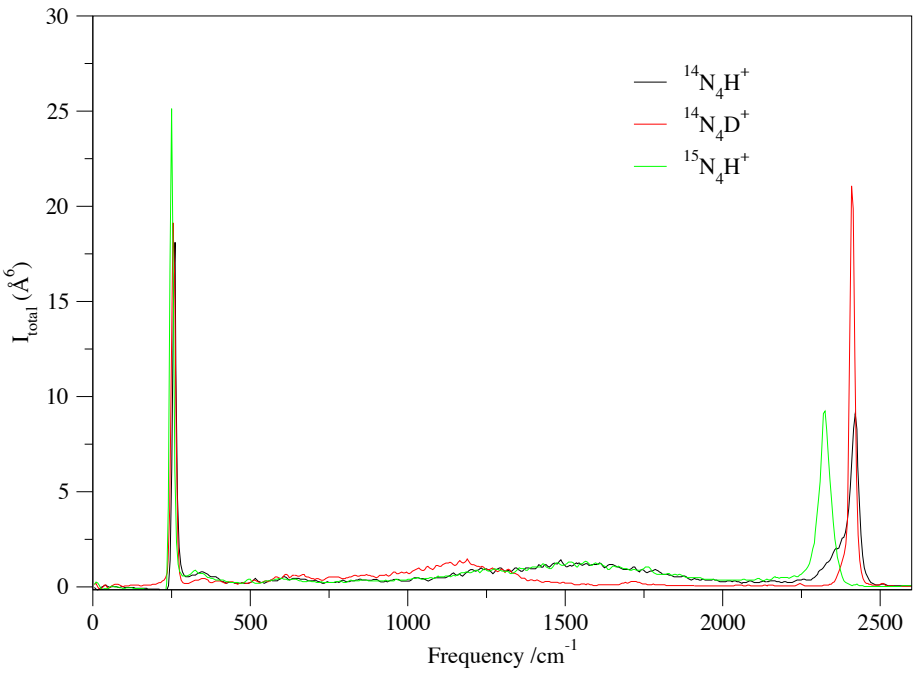


Figure S5. The MD spectra (the total intensity) for $^{14}N_4H^+$,$^{14}N_4D^+$ and$^{15}N_4H^+$ at 200 K calculated with analytical **N4H+_PES** and the presently fitted **N4H+_PTS**.

S.5 Driven molecular dynamics coordinate analysis

The following set of bond-angle coordinates (Eq. S. 13) were used in our previous work [5]. These symmetric internal coordinates, *s* were monitored along the DMD trajectories to visualize and assign the vibrational modes. The atom numbering is described in the Figure S3.

$$
\begin{aligned}
&s_1(\sigma_g) = r(N_1 - N_3) + r(N_2 - N_4)\\
&s_2(\sigma_u) = r(N_1 - N_3) - r(N_2 - N_4)\\
&s_3(\sigma_g) = r(H_5 - N_3) + r(H_5 - N_4)\\
&s_4(\sigma_u) = r(H_5 - N_3) - r(H_5 - N_4)\\
&s_{5,6}(\pi_g, xz, yz) = a(N_1 - N_3 - H_5)_{y,x} + a(N_2 - N_4 - H_5)_{y,x}\\
&s_{7,8}(\pi_u, xz, yz) = a(N_3 - H_5 - N_4)_{y,x}\\
&s_{9,10}(\pi_u, xz, yz) = a(N_1 - N_3 - H_5)_{y,x} - a(N_2 - N_4 - H_5)_{y,x}
\end{aligned}
\qquad \text{(S. 13)}
$$

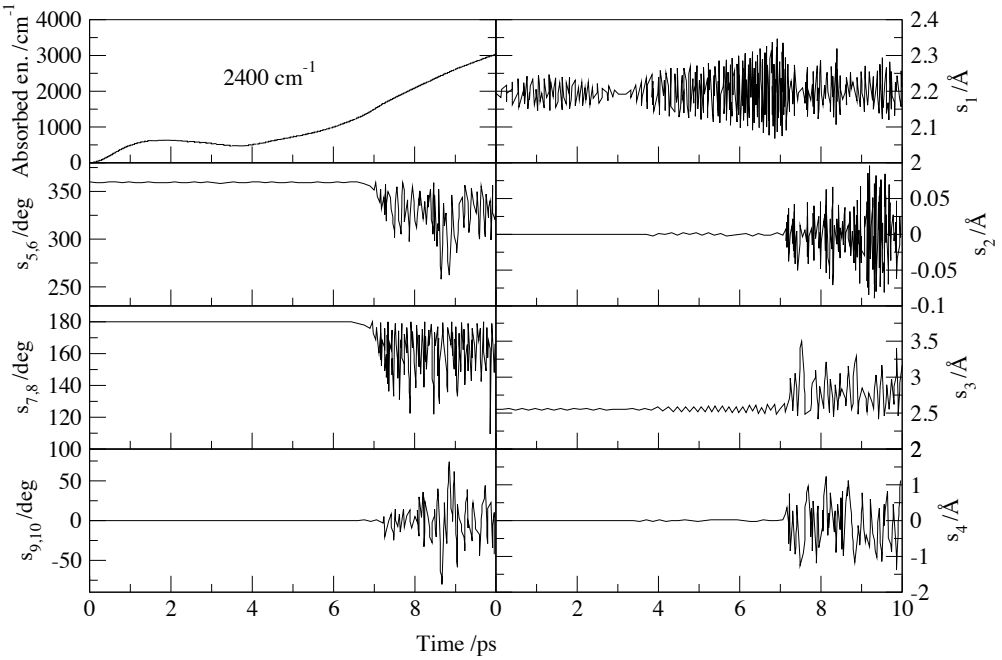


Figure S6. Average absorbed energy (in cm^{-1}) and symmetric internal coordinates $s_{1\text{-}10}$ along the parallel scattering α-DMD trajectories of $^{14}N_4H^+$ for the symmetric N−N stretch frequency 2400 cm^{-1}. The electric field strength is 350 mV/bohr with the initial orientation parallel to the molecular axis. The Fermi resonance is confirmed by large displacements of the $s_{5\text{-}10}$ coordinates from their equilibrium values at later stage of the trajectory, 7 ps.

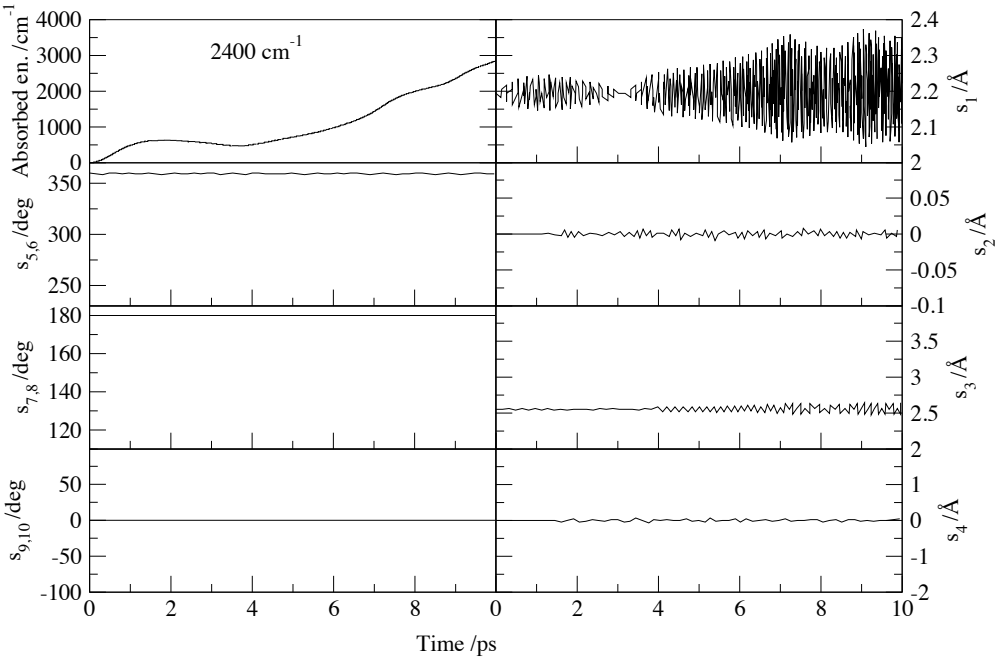


Figure S7. Average absorbed energy (in cm^{-1}) and symmetric internal coordinates $s_{1\text{-}10}$ along the parallel scattering α-DMD trajectories of $^{14}N_4D^+$ for the symmetric N−N stretch frequency 2400 cm^{-1}. The electric field strength is 350 mV/bohr with the initial orientation parallel to the molecular axis. The Fermi resonance is absent here, due to small displacements of the $s_{5\text{-}10}$ coordinates from their equilibrium values.

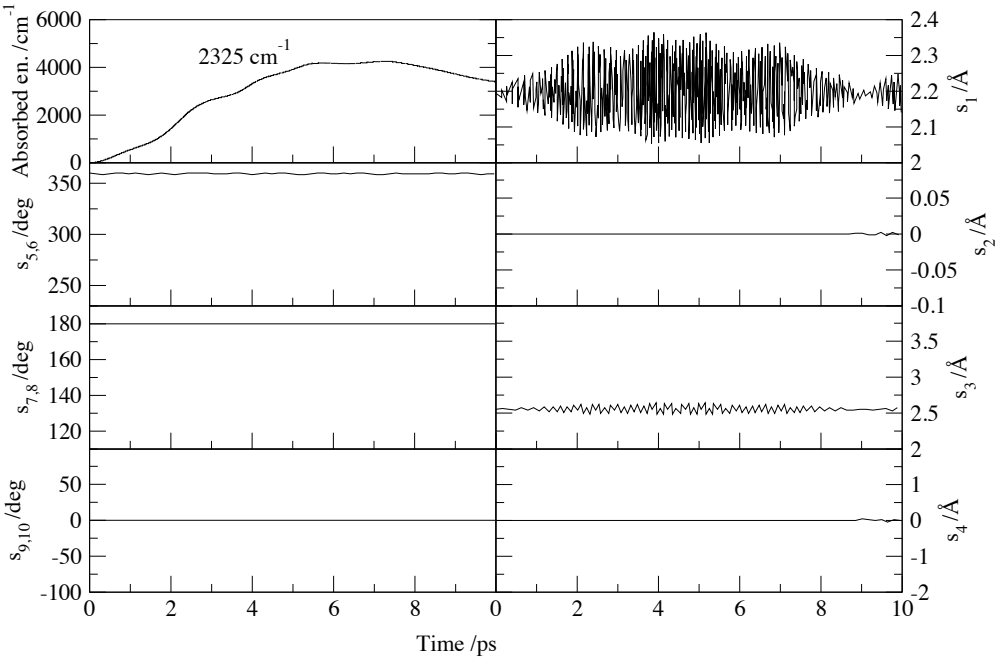


Figure S8. Average absorbed energy (in cm^{-1}) and symmetric internal coordinates $s_{1\text{-}10}$ along the parallel scattering α-DMD trajectories of $^{15}N_4H^+$ for the symmetric N−N stretch frequency 2325 cm^{-1}. The electric field strength is 350 mV/bohr with the initial orientation parallel to the molecular axis. The Fermi resonance is absent here, due to small displacements of the $s_{5\text{-}10}$ coordinates from their equilibrium values.

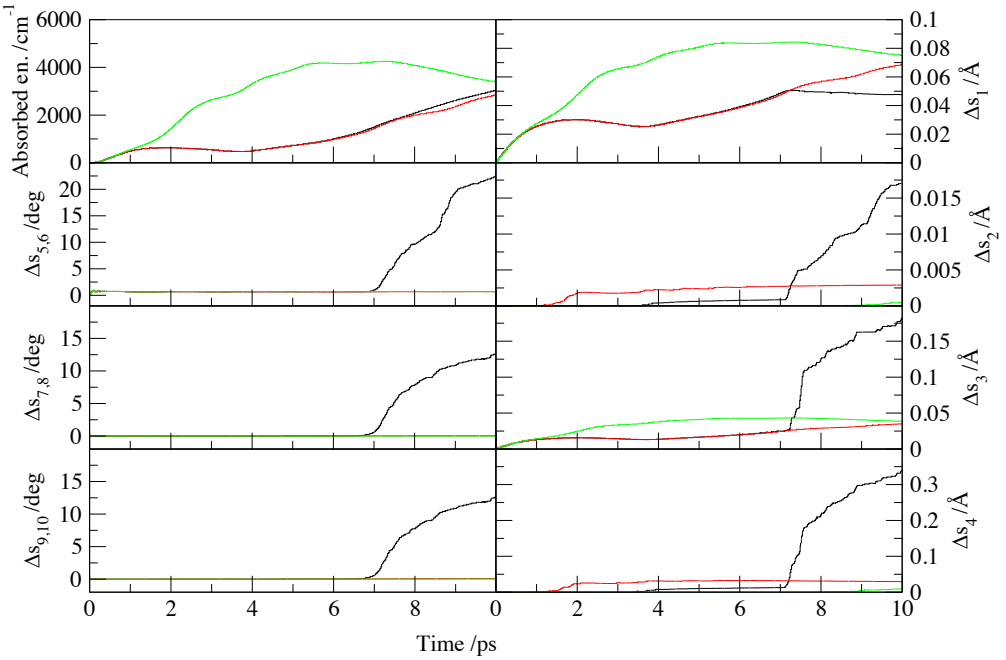


Figure S9. The average absorbed energy (in cm^{-1}) and averaged symmetrized displacements, Δs_i (Eq. 12, main text) along the parallel scattering α-DMD trajectories for the symmetric N−N stretch frequencies $^{14}N_4H^+$ 2400 cm^{-1} (in black), $^{14}N_4D^+$ 2400 cm^{-1} (red), and $^{15}N_4H^+$ 2325 cm^{-1} (green). The electric field strength is 350 mV/bohr with the initial orientation parallel to the molecular axis.

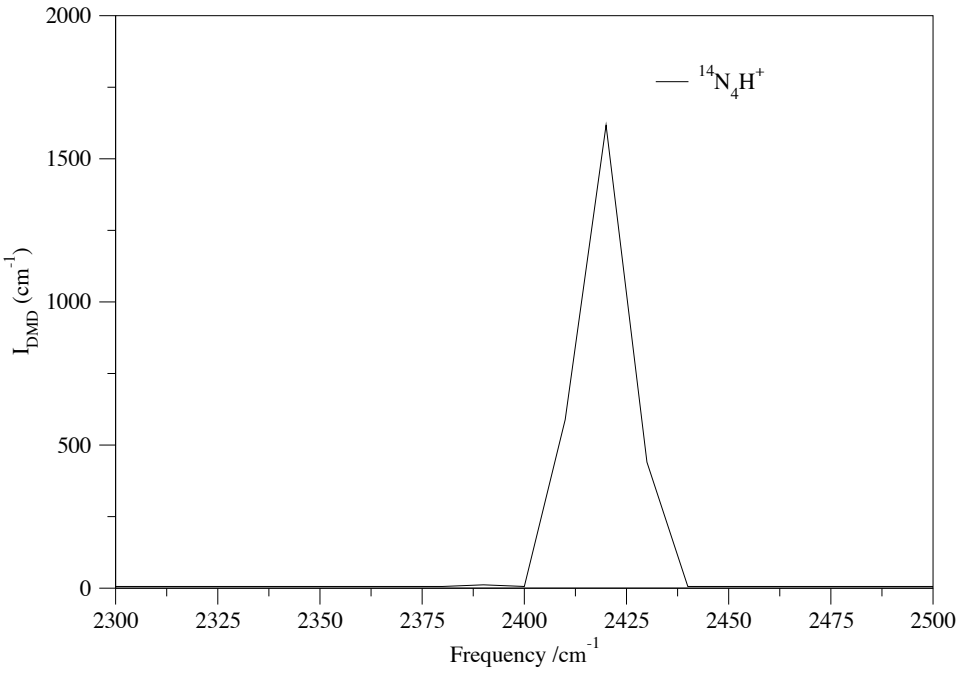


Figure S10. The perpendicular scattering α-DMD spectrum of $^{14}N_4H^+$ in the high frequency range calculated with the analytical **N4H+_PES** and the presently fitted **N4H+_PTS**. The electric field strength is 700 mV/bohr with the initial orientation perpendicular to the molecular axis.

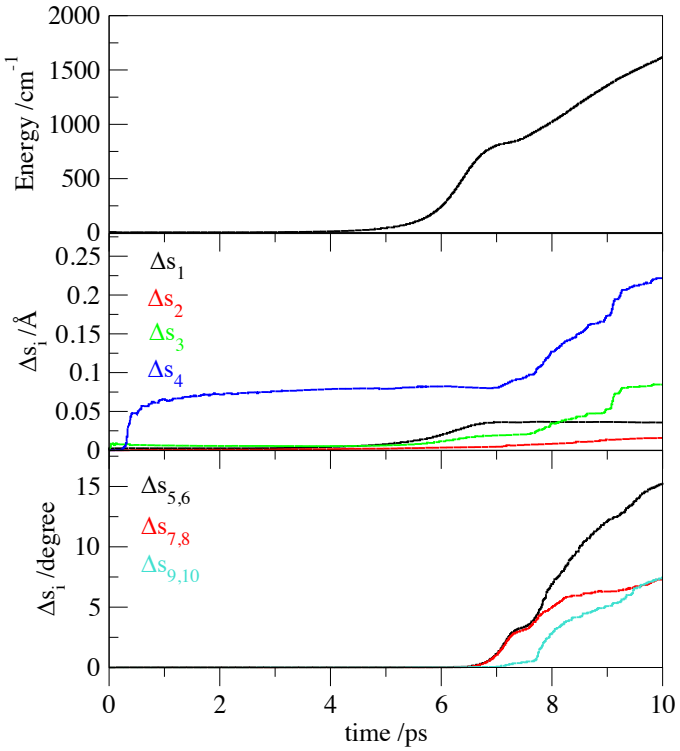


Figure S11. The average absorbed energy (in cm^{-1}) and averaged symmetrized displacements, Δs_i (Eq. 12, main text) along the perpendicular scattering α-DMD trajectory for the asymmetric $N\cdots H^+\cdots N$ bend overtone frequency 2420 cm^{-1} of $^{14}N_4H^+$. The electric field strength is 700 mV/bohr with the initial orientation perpendicular to the molecular axis.

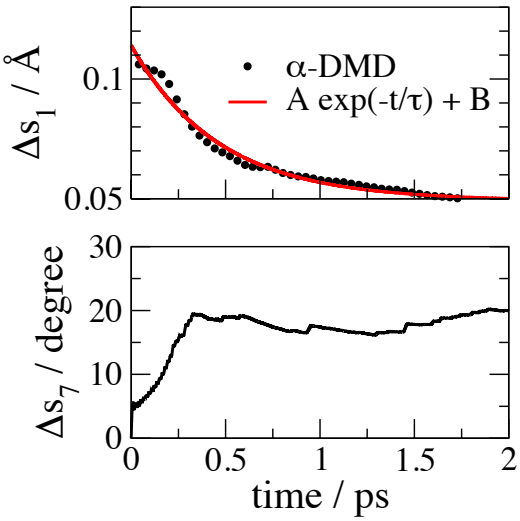


Figure S12. Probe dynamics. τ = 466 fs.

S.6 Least-squares optimization

We define the least-squares function for polarizability fitting in the following way,

$$W(\mathbf{c}) = \frac{1}{2K}\sum_{k=1}^{K}\sum_{i,j=1}^{3}\left(\alpha_{ij}(\mathbf{r}_k;\mathbf{c}) - \alpha_{ij}^{\mathrm{REF}}(\mathbf{r}_k)\right)^2 \qquad (S.14)$$

where k runs over the total of K configurations; i, j are the polarizability tensor's Cartesian components X, Y, Z. "REF" refers to reference data obtained in an electronic structure calculation, which may be of any level of theory, via trajectory propagation or any other comparable sampling methods. In this form the coefficients of all six independent tensor components are simultaneously optimized. This formulation is advantageous. The absolute RMS error of the fit is $\sqrt{W}$, and the relative RMS error is defined as

$$\mathrm{RMS}(\%) = 100\left[\frac{\sum_{k=1}^{K}\sum_{i,j=1}^{3}\left(\alpha_{ij}(\mathbf{r}_k;\mathbf{c}) - \alpha_{ij}^{\mathrm{REF}}(\mathbf{r}_k)\right)^2}{\sum_{k=1}^{K}\sum_{i,j=1}^{3}\left(\alpha_{ij}^{\mathrm{REF}}(\mathbf{r}_k)\right)^2}\right]^{\frac{1}{2}} \qquad (S.15)$$

To carry out an actual optimization, we differentiate W with respect to $\mathbf{c}$ analytically and pass both the function and its gradient

$$\nabla W(\mathbf{c}) = \frac{1}{K}\sum_{k=1}^{K}\sum_{i,j=1}^{3}\left(\alpha_{ij}(\mathbf{r}_k;\mathbf{c}) - \alpha_{ij}^{\mathrm{REF}}(\mathbf{r}_k)\right)\frac{\partial\alpha_{ij}(\mathbf{r}_k;\mathbf{c})}{\partial\mathbf{c}} \qquad (S.16)$$

to the BFGS optimizer. With ∇W being highly non-linear, in a difficult case a typical number of iterations to reach a tight convergence, e.g. RMS(%) < 1, can be $O(10^5)$ with CPU times measured in days on a single processor workstation. Obviously, this is a major bottleneck when the total polynomial order and size of the training set need to be increased. However, noticing that both W and ∇W can be calculated by partitioning the training set K into any number of independent blocks, we apply the usual MPI 'send-receive' parallelization routine to Eqs. S.14 and S.16 and nest it inside the main BFGS driver. With the overhead of a BFGS step being small, the parallelization is highly scalable, permitting us to run efficient optimizations on a 56-core CPU Intel Xeon "Gold 6132" 2.6GHz node.

S.7 A covariant symmetrization scheme

1. determine the number of non-zero powers in a monomial: M_NZ{monomial}

   Ex: M_NZ$\{y_{12}^{2}y_{13}^{1}y_{14}^{0}y_{15}^{0}y_{23}^{1}y_{24}^{0}y_{25}^{3}\}$ = 4

2. define quantity MAX_SYM = min[N_ATOM - 1, M_NZ]

   where N_ATOM is the number of atoms

3. define quantity MIN_SYM = M_NZ - MAX_SYM

4. compute the number of times (F_k) the atom index (k) shows up in the monomial

   Ex: for $y_{12}^{2}y_{13}^{1}y_{14}^{0}y_{15}^{0}y_{23}^{1}y_{24}^{0}y_{25}^{3}$

   F_1=3, F_2=3, F_3=2, F_4=0, F_5=1

5. Compare F_k against the quantity F^* = max[MAX_SYM - 1, 1]:

   a) set by default b_k = 1

   b) if F_k > MIN_SYM; then

   if $F_k \geq F^*$; then b_k = -1

S.8 Details of the polarizability tensor

XYZ coordinates in Å of configuration I.

```
N1     0.00000000    0.00000000    1.27836800
N2     0.00000000    0.00000000    2.37975600
N3     0.00000000    0.00000000   -1.27836800
N4     0.00000000    0.00000000   -2.37975600
H5     0.00000000    0.00000000    0.00000000

 ISOTROPIC POLARIZABILITIES (Ang^3)
  1 N     1.255024
  2 N     0.242878
  3 N     1.255024
  4 N     0.242878
  5 H     0.157721

#LAMBDA-2 LAMBDA-3 LAMBDA-4
    0.295457   -0.056623    0.427348

 # 1-body tensor
      3.1535239284      0.0000000000      0.0000000000
      0.0000000000      3.1535239284      0.0000000000
      0.0000000000      0.0000000000      3.1535239284
 # 2-body tensor
     -1.5174468517      0.0000000000      0.0000000000
      0.0000000000     -1.5174468517      0.0000000000
      0.0000000000      0.0000000000      3.0348937034
 # 3-body tensor
      0.9926978062      0.0000000000      0.0000000000
      0.0000000000      0.9926978062      0.0000000000
      0.0000000000      0.0000000000      3.9707912249
 # 4-body tensor
     -0.5040854970      0.0000000000      0.0000000000
      0.0000000000     -0.5040854970      0.0000000000
      0.0000000000      0.0000000000      4.0326839757

 POLARIZABILITY TENSOR [Ang^3]

      2.4335549933      0.0000000000      0.0000000000
      0.0000000000      2.4335549933      0.0000000000
      0.0000000000      0.0000000000      5.5487280203
```

XYZ coordinates in Å of configuration II upon a cyclic permutation of X.

```
N2     0.00000000    0.00000000    2.37975600
N3     0.00000000    0.00000000   -1.27836800
N4     0.00000000    0.00000000   -2.37975600
N1     0.00000000    0.00000000    1.27836800
H5     0.00000000    0.00000000    0.00000000

 ISOTROPIC POLARIZABILITIES ([Ang^3])
  1 N     0.242878
  2 N     1.255024
  3 N     0.242878
  4 N     1.255024
  5 H     0.157721

#LAMBDA-2 LAMBDA-3 LAMBDA-4
    0.295457   -0.056623    0.427348

 # 1-body tensor
      3.1535239284      0.0000000000      0.0000000000
      0.0000000000      3.1535239284      0.0000000000
      0.0000000000      0.0000000000      3.1535239284
 # 2-body tensor
     -1.5174468517      0.0000000000      0.0000000000
      0.0000000000     -1.5174468517      0.0000000000
      0.0000000000      0.0000000000      3.0348937034
 # 3-body tensor
      0.9926978062      0.0000000000      0.0000000000
      0.0000000000      0.9926978062      0.0000000000
      0.0000000000      0.0000000000      3.9707912249
 # 4-body tensor
     -0.5040854970      0.0000000000      0.0000000000
      0.0000000000     -0.5040854970      0.0000000000
      0.0000000000      0.0000000000      4.0326839757

 POLARIZABILITY TENSOR [Ang^3]

      2.4335549933      0.0000000000      0.0000000000
      0.0000000000      2.4335549933      0.0000000000
      0.0000000000      0.0000000000      5.5487280203
```

S.9 References

[1] T. M. Miller, B. Bederson, *Adv. At. Mol. Phys*. **1977**, *13*, 1.

[2] D. A. McQuarrie, *Statistical Mechanics*; Harper & Row: New York, p. 481, 1976.

[3] Y. N. Kalugina, N. Cherepanov, *Atmos. Ocean. Opt.* **2015**, *28*, 406.

[4] Ricks, A. M.; Douberly, G. E.; Duncan, M. A.; *J. Chem. Phys.* **2009**, *131*, 1045312.

[5] R. Hooper, D. Boutwell, M. Kaledin, *J. Phys Chem. A,* **2019,** *123*, 5613.