Supplementary Information

A Data-Efficient Machine-Learning Approach for Modeling the Photodynamics of All-trans Hexatriene based on Multireference Configuration Interaction Calculations

Luan G. F. dos Santos,[1] Julio C. V. Chagas,[2] Mikołaj Martyka,[3] Pavlo O. Dral,[4,5,6] Mario Barbatti,[7,8] Francisco B.C. Machado,[9,10] Richard A. Messerly,[11,†] Hans Lischka[1,‡]

[†]messerlyra@ornl.gov

[‡]hans.lischka@ttu.edu

[1] Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX 79409, United States

[2] Department of Chemistry, Northwestern University, Evanston, IL 60208, United States

[3] University of Warsaw, Faculty of Chemistry, 02-093 Warsaw, Poland

[4] State Key Laboratory of Physical Chemistry of Solid Surfaces, Department of Chemistry and Chemical Engineering, College of Chemistry and Chemical Engineering, and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Xiamen University, Xiamen, Fujian, China

[5] Institute of Physics, Faculty of Physics, Astronomy, and Informatics, Nicolaus Copernicus University, Toruń, Poland

[6] Aitomistic, Shenzhen, China

[7] Aix Marseille University, CNRS, ICR, 13397 Marseille, France

[8] Institut Universitaire de France, 75231 Paris, France

[9] Department of Chemistry, Aeronautics Institute of Technology, São José dos Campos, SP 12228 900, Brazil

[10] Advanced Scientific Computing and Modeling Laboratory, Aeronautics Institute of Technology, São José dos Campos, SP 12228-900, Brazil

[11] National Center for Computational Sciences Division, Oak Ridge National Laboratory, Oak Ridge, 37830, TN, United States

Table of Contents

S.1 Ensemble Uncertainty Along MLIP-MRCI Trajectories

To assess the reliability of the MLIP-MRCI model in the regions of configurational space visited during the nonadiabatic dynamics, the ensemble standard deviation was evaluated along 100 MLIP-MRCI trajectories. For each geometry, all eight HIP-NN models of the ensemble were used to independently predict the energies of the three electronic states (S_0, S_1, and S_2), and the standard deviation across the eight predictions was computed as a measure of model uncertainty.

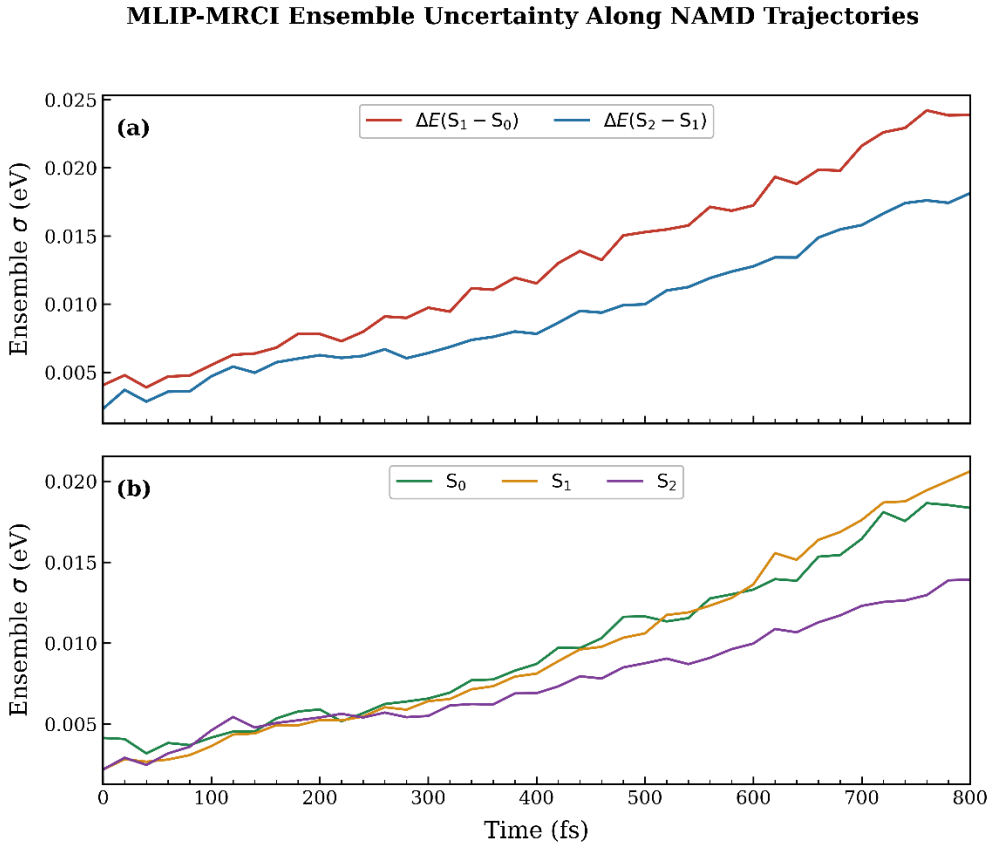

Figure S1: Mean ensemble standard deviation (σ) of the predicted energy gaps $\Delta E(S_1-S_0)$ and $\Delta E(S_2-S_1)$ (a) and state energies S_0, S_1, and S_2 (b) as a function of propagation time. For each geometry along 100 MLIP-MRCI nonadiabatic trajectories, the standard deviation was computed across the predictions of the eight independently trained HIP-NN models of the ensemble. The mean was then taken over all trajectories at each time step.

S.2 Selection of optimal HIP-NN architecture

Table S1: Hyperparameter combinations explored during HIP-NN optimization. Each model (hp_0–hp_{23}) corresponds to a unique combination of atomic features (N_{feat}), atom layers (N_{atom}), and interaction layers (N_{int}).

Combination ID	Atomic Features (N_{feat})	Atom Layers (N_{atom})	Interaction Layers (N_{int})
0	16	2	1
1	16	2	2
2	16	3	1
3	16	3	2
4	16	4	1
5	16	4	2
6	32	2	1
7	32	2	2
8	32	3	1
9	32	3	2
10	32	4	1
11	32	4	2
12	64	2	1
13	64	2	2
14	64	3	1
15	64	3	2
16	64	4	1
17	64	4	2
18	128	2	1
19	128	2	2
20	128	3	1
21	128	3	2
22	128	4	1
23	128	4	2

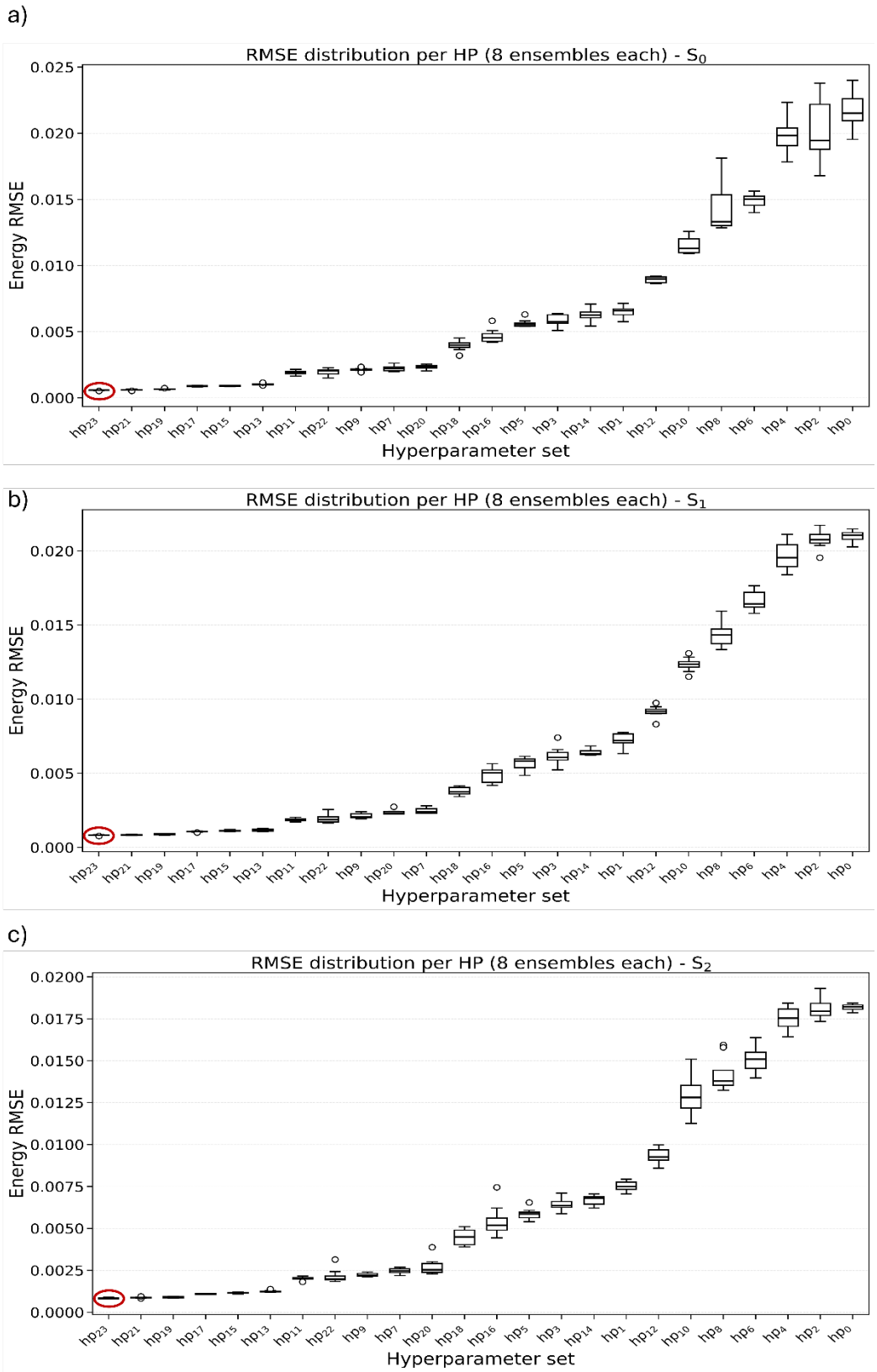

Figure S2: Distribution of validation energy RMSE (eV/atom) across the 24 tested hyperparameter combinations (hp_0–hp_{23}) for (a) S_0, (b) S_1, and (c) S_2. Each box plot represents the performance of an ensemble of eight independently trained models. The hyperparameter set yielding the lowest mean RMSE (optimal set) in each panel is highlighted with a red circle. The hyperparameter definitions corresponding to each hp index are provided in Table S1.

Table S2: Training hyperparameters used for all HIP-NN models. The same training protocol was applied across all hyperparameter combinations, ensuring that differences in model performance arise solely from architectural variations.

Parameter	Value
Optimizer	Adam
Learning rate	1×10^{-3}
Batch size (max)	256
Number of epochs	1000
Train/Validation/Test	90/10/10 (random)
Loss weights (E : F : Gaps)	5 : 1 :5
Regularization	L2 + hierarchicality penalty
Framework	PyTorch (HIP-NN implementation)

S.3 Comparison of Surface Hopping Protocols

The reliability of the NAC-free Landau–Zener–Belyaev–Lebedev (LZBL) scheme for describing the relaxation dynamics of the investigated systems was evaluated through a comparative analysis with other established surface hopping protocols. In this context, the LZBL-full implementation refers to the complete treatment of the surface hopping stochastic process, including the use of a kinetic energy reservoir to handle frustrated hops and ensure energy conservation, whereas reduced schemes typically neglect such reservoir effects or simplify the state-coupling transitions. As shown in Figure S3, the population evolution was obtained using different surface hopping methods, all employing the AIQM1 level of theory with identical initial conditions. Notably, although the Fewest Switches Surface Hopping (FSSH) method utilized analytical nonadiabatic couplings (NACs), the NAC-free LZBL algorithm yielded very similar results. This strong agreement validates the use of the LZBL protocol for the production dynamics.

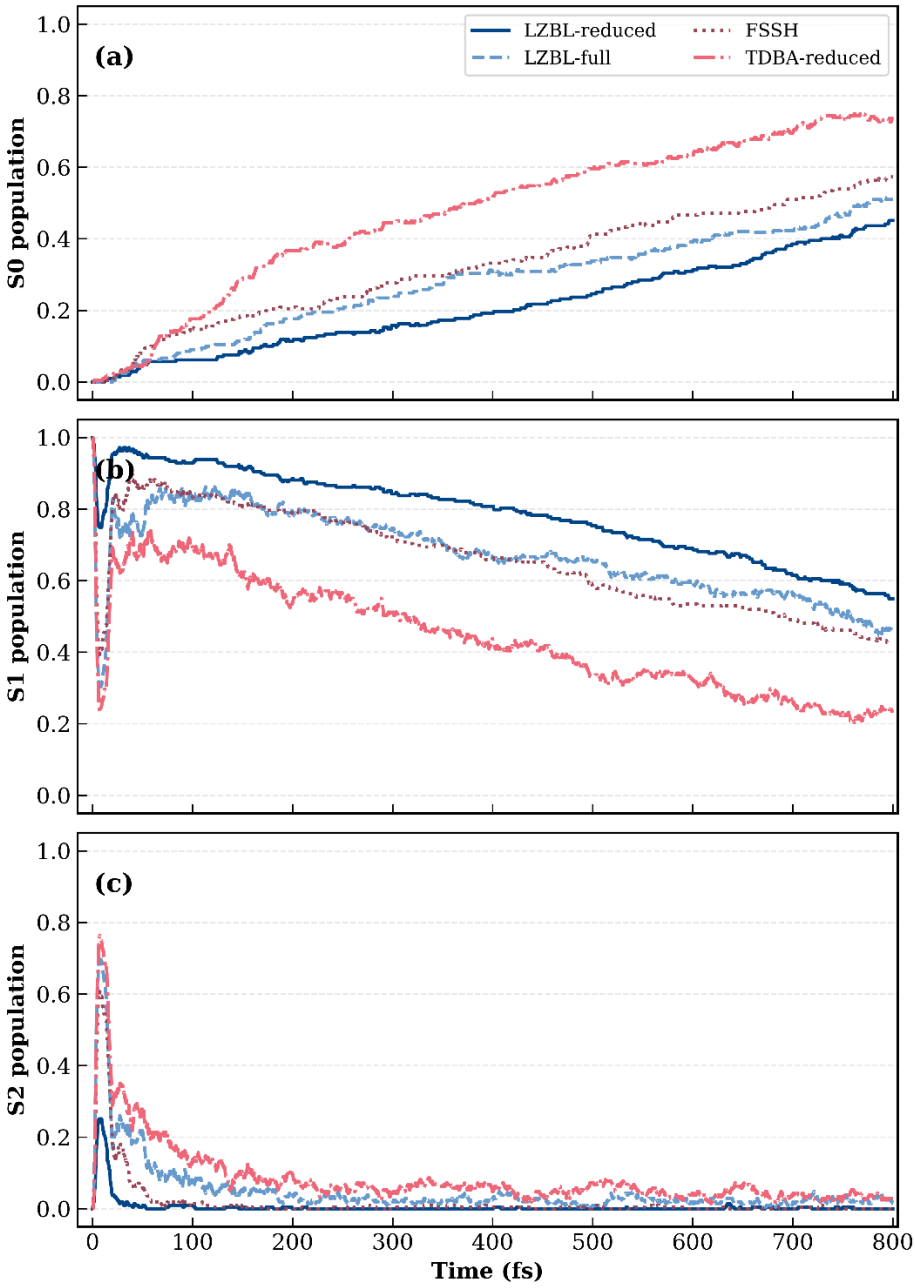

Figure S3: Comparison of the electronic population dynamics for the S_0 (a), S_1 (b), and S_2 states (c) over 800 fs. All simulations were performed at the AIQM1 level of theory using identical initial conditions and a time step of 0.2 fs.

S.4 Vertical Excitation Energies and State Characterization

Table S3: Vertical excitation energies (eV) and oscillator strengths at the S_0 equilibrium geometry for different electronic structure methods. States are labeled according to their wavefunction character (covalent $2^1A_g^-$ and ionic $1^1B_u^+$). Oscillator strengths highlight the bright character of the ionic state and the dark character of the covalent state.

Method	State	Energy (eV)	Oscillator Strength
AIQM1	$2^1A_g^-$	5.830[a]	0.0000[a]
AIQM1	$1^1B_u^+$	5.318[a]	1.2456[a]
MR-CISD	$2^1A_g^-$	5.470[a]	0.0000[a]
MR-CISD	$1^1B_u^+$	6.294[a]	1.2729[a]
CASSCF	$2^1A_g^-$	5.799[b]	0.0000[b]
CASSCF	$1^1B_u^+$	7.609[b]	1.4498[b]
MR-CISD+P	$2^1A_g^-$	5.819[b]	-
MR-CISD+P	$1^1B_u^+$	5.827[b]	-
MR-AQCC	$2^1A_g^-$	5.818[b]	-
MR-AQCC	$1^1B_u^+$	5.833[b]	-
CASPT2	$2^1A_g^-$	5.761[b]	-
CASPT2	$1^1B_u^+$	5.564[b]	-
NEVPT2 (QD-SC2)	$2^1A_g^-$	4.341[c]	-
NEVPT2 (QD-SC2)	$1^1B_u^+$	5.265[d]	-
NEVPT2 (QD-PC2)	$2^1A_g^-$	4.347[c]	-
NEVPT2 (QD-PC2)	$1^1B_u^+$	5.165[d]	-

[a] Results obtained in this work.

[b] Chagas *et al.* (Ref. 18, geometry optimized at the ωB97X-D/cc-pVTZ level; single-point energy at the CAS(6,6), MR-CISD+P/CAS(6,6), MR-AQCC/CAS(6,6) and CASPT2/CAS(6,6) levels).

[c] Guareschi and Angeli (Ref. 13), geometry optimized at the NEVPT2/CAS(6,6) level; single-point energy at the SC- and PC-NEVPT2/CAS(6,14) levels, respectively).

[d] Guareschi and Angeli (Ref. 13), geometry optimized at the RASSCF/RAS(12)CAS(6,6)AUX(12) level; single-point energy at the SC- and PC-NEVPT2/CAS(6,14) levels, respectively.

S.5 Convergence of Population Dynamics with Respect to the Number of Trajectories

Figure S2 shows the electronic-state populations of S_0, S_1, and S_2 computed from subsets of N = 50, 100, 150, and 200 trajectories, ran using MLIP-MRCI. The population curves become progressively more stable as N increases, with the N = 150 and N = 200 results showing excellent mutual agreement throughout the entire 800 fs simulation window. The largest deviations are observed for S_1 at intermediate times (~250–400 fs). Nevertheless, even in this region the N = 200 ensemble yields smooth and well-converged populations. These results confirm that 200 trajectories provide a statistically converged description of the photodynamics of all-trans-hexatriene at the MLIP-MRCI level.

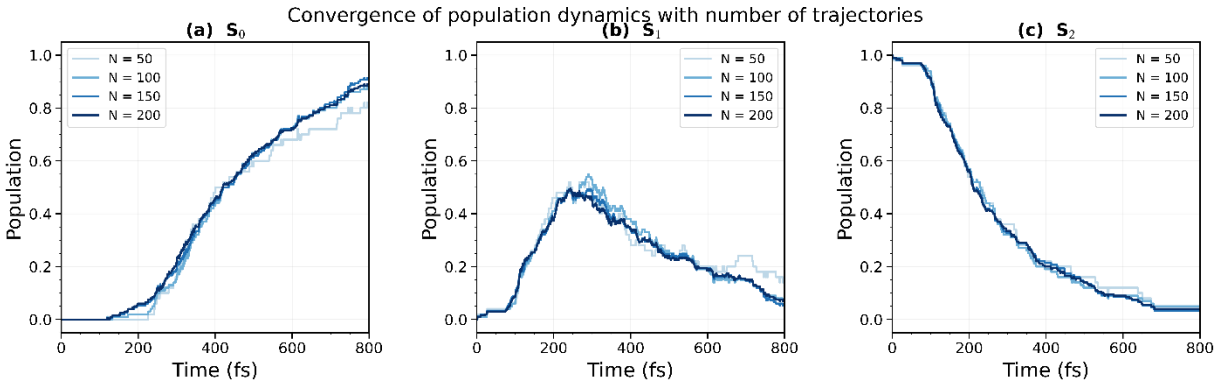

Figure S4: Convergence of the electronic-state population dynamics with respect to the number of trajectories. Populations of (a) S_0, (b) S_1, and (c) S_2 are shown for ensembles of N = 50, 100, 150, and 200 trajectories (light to dark blue). The close agreement between the N = 150 and N = 200 curves across all three states confirms that 200 trajectories are sufficient to obtain well-converged population dynamics.

S.6 Principal Component Analysis of $S_1 \rightarrow S_0$ Hopping Geometries

To further investigate the structural diversity of the $S_1 \rightarrow S_0$ hopping geometries, a principal component analysis (PCA) was performed on a set of six internal coordinates: the three absolute torsional angles ($|\varphi_1|$, $|\varphi_2|$, $|\varphi_3|$) around each C=C double bond, and the three individual double bond lengths (d_{12}, d_{34}, d_{56}), all labeled according to the hexatriene backbone shown in Figure 1. Prior to PCA, all coordinates were standardized to have zero mean and unit variance, ensuring that differences in physical units between torsional angles (degrees) and bond lengths (Å) do not bias the analysis.

Although the full PCA yields six components (equal to the number of input coordinates), only the first four are reported here, as they collectively explain ~76% of the total variance

(Figure S3d); the remaining two components each contribute less than 13% and do not alter the conclusions. Scatter plots are shown for PC1 vs PC2 and PC1 vs PC3, as these pairs involve the components of greatest variance and are therefore the most informative projections.

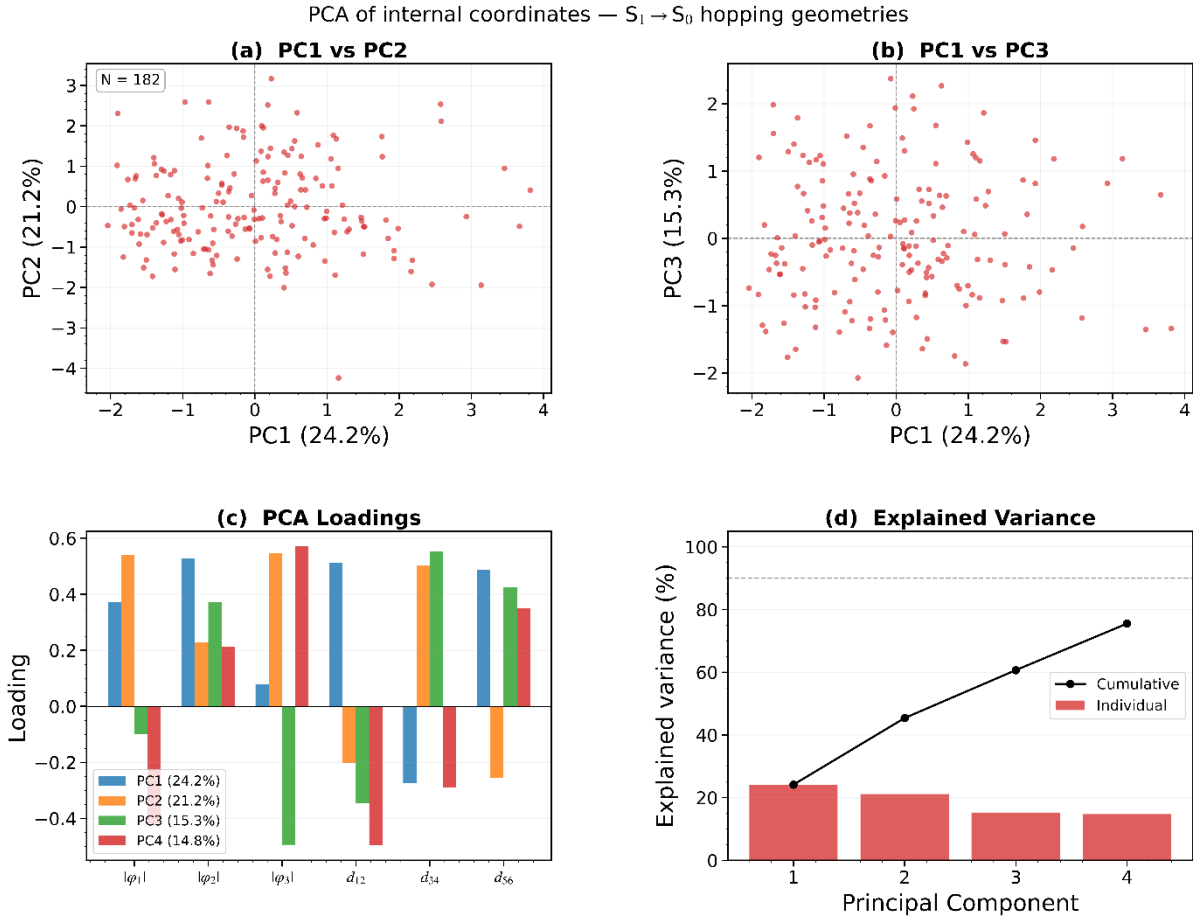

Figure S5: Principal component analysis (PCA) of internal coordinates at the $S_1 \rightarrow S_0$ hopping geometries (N = 182). The internal coordinate set includes the three absolute torsional angles ($|\varphi_1|$, $|\varphi_2|$, $|\varphi_3|$), and the C–C bond lengths (d_{23}, d_{45}, d_{34}). Panels (a) and (b) show the distribution of hopping geometries projected onto PC1 vs PC2 and PC1 vs PC3, respectively. Panel (c) shows the PCA loadings for the first four principal components. Panel (d) shows the individual and cumulative variances as a function of the number of principal components.

Panel (Figure S3d) shows that four principal components are required to explain ~76% of the total structural variance, with individual contributions of 24.2%, 21.2%, 15.3%, and 14.8% for PC1–PC4, respectively. These contributions are remarkably similar across all four PCs, indicating that the structural variance is distributed nearly uniformly across multiple independent modes. Inspection of the loadings in panel (c) confirms this picture: no single coordinate dominates any of the principal components. In PC1, φ_1 (+0.37), φ_2 (+0.53), d_{12} (+0.51), and d_{56} (+0.49) all contribute substantially. In PC2, φ_1 (+0.54), φ_3 (+0.55), and d_{34} (+0.50) are the main contributors. PC3 and PC4 show similarly distributed loadings across the

coordinate set. This absence of a dominant coordinate confirms that the $S_1 \rightarrow S_0$ transition does not proceed through a single well-defined structural mode.

Panels (a) and (b) show the projection of the 182 hopping geometries onto PC1 vs PC2 and PC1 vs PC3, respectively. In both panels, the geometries form a diffuse, unstructured cloud with no identifiable clusters or preferential regions, consistent with the broad and uniform distribution observed in the torsional angle histograms (Figure 8 of the main text) and with the highly scattered distribution of hopping points in Figure 7b. Together, these results support the finding that internal conversion to the ground state in all-trans-hexatriene is not governed by a single nuclear coordinate or a narrow region of configurational space, but rather occurs across a structurally diverse ensemble of geometries spanning multiple torsional and bond-length deformation modes.

S.7 Geometries (Å)

Hexatriene equilibrium geometry obtained at ωB97X-D/cc-pVTZ level:

C	1.199785	2.813176	0.000000
C	1.199785	1.483143	0.000000
C	-0.001156	0.668414	0.000000
C	0.001156	-0.668414	0.000000
C	-1.199785	-1.483143	0.000000
C	-1.199785	-2.813176	0.000000
H	2.120872	3.378561	0.000000
H	0.273669	3.375164	0.000000
H	2.145268	0.949588	0.000000
H	-0.950722	1.196588	0.000000
H	0.950722	-1.196588	0.000000
H	-2.145268	-0.949588	0.000000
H	-2.120872	-3.378561	0.000000
H	-0.273669	-3.375164	0.000000

Representative $S_2 \rightarrow S_1$ hopping geometry ($t \approx 234$ fs) extracted from MLIP-MRCI dynamics:

C	1.30335581	2.77925565	0.27704951
C	1.03906277	1.37110984	0.39003123
C	0.04224895	0.80630050	-0.73069395
C	-0.10958283	-0.72707285	-0.56053369
C	-1.09406776	-1.38704637	0.26128606
C	-1.17345982	-2.87649558	0.41028838
H	1.99806634	3.41082415	0.88052706
H	1.00985395	3.26936007	-0.73166739
H	1.47112879	0.59799652	1.00924140
H	-0.59617766	1.33696663	-1.32568801
H	0.22719596	-1.29480389	-1.50813720
H	-1.88063590	-0.67129120	0.67936833
H	-1.99023301	-3.19541286	0.97664815
H	-0.30467380	-3.44938289	0.36001521

Representative $S_1 \rightarrow S_0$ hopping geometry ($t \approx 410$ fs) extracted from MLIP-MRCI dynamics:

C	1.05791510	2.78395585	-0.53761244
C	1.09799008	1.28676139	-0.35892153
C	0.25333819	0.57123755	0.68337798
C	-0.16203292	-0.91241798	0.64084896
C	-1.32747215	-1.36574945	0.32263060
C	-1.04830465	-2.41099554	-0.52861893
H	2.06927751	3.25622274	-0.46335431
H	0.48400210	3.15234520	-1.34182590
H	1.69636044	0.51765284	-1.01271806
H	0.59167116	1.08881454	1.61545962
H	0.72789511	-1.51280299	1.24319824
H	-2.19937262	-0.66102110	0.37211083
H	-0.98088567	-3.34135138	0.11788285
H	-0.74421083	-2.71906626	-1.47397421