Supplementary Materials for

"Evidence of Molecular Hydrogen in the N-doped LuH_3 System: a Possible Path to Superconductivity?"

Cesare Tresca,[1,*] Pietro Maria Forcella,[2] Andrea Angeletti,[3,4] Luigi Ranalli,[3,4] Cesare Franchini,[4,5] Michele Reticcioli,[4,†] and Gianni Profeta[1,2]

[1] *CNR-SPIN c/o Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio 10, I-67100 L'Aquila, Italy*

[2] *Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio 10, I-67100 L'Aquila, Italy*

[3] *University of Vienna, Vienna Doctoral School in Physics, Boltzmanngasse 5, 1090 Vienna, Austria*

[4] *University of Vienna, Faculty of Physics and Center for Computational Materials Science, Kolingasse 14-16, 1090 Vienna, Austria*

[5] *Dipartimento di Fisica e Astronomia, Università di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy.*

CONTENTS

LIST OF FIGURES

* cesare.tresca@spin.cnr.it

† michele.reticcioli@univie.ac.at

LIST OF TABLES

I. RAMPING-TEMPERATURE MLFF-MD CALCULATIONS

Fig. SF1 shows the results obtained from our machine-learning-accelerated molecular-dynamics calculations, with a temperature ramping from ∼0 K to 400 K ($50{\cdot}10^3$ steps). A portion of these data is shown in Fig. 1 in the main text (initial 200 fs), to highlight the formation of H_2 molecules at low temperature. As discussed in the main text, we started the simulation from a highly symmetric $Fm\bar{3}m$ $LuH_{2.875}N_{0.125}$ ($4 \times 4 \times 4$) unit cell (shown in Fig. SF2a); H_2 molecules start to form after a few tens of fs, at temperatures as low as 15 K (see also Fig. 1 in the main text). Fig. SF2b shows the structure with the H_2 molecules, as obtained from this MLFF-MD calculation. The steep temperature increase observed during the initial stage of the ramping process (at time $t << 1$ ps in Fig. SF1c) is attributed to the significant reduction in potential energy within the system (Fig. SF1b), due to the rapid formation of H_2 molecules.

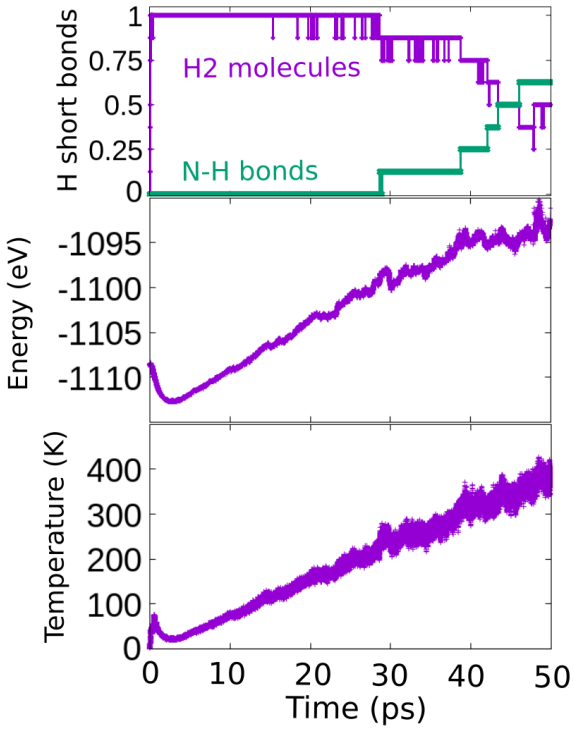

FIG. SF1. MLFF-MD simulations with temperature ramping from 0 K to 400 K. Number of H_2 molecules (purple) and N-H bonds (green) per N atom in the unit cell (top panel), free energy (central panel) and effective temperature (bottom panel). The formation of H_2 molecules in the initial 200 fs are shown in the Fig. 1 in the main text.

Once the molecules are formed and stabilized, the temperature rises following the expected linear ramping. At high temperature (approximately 250 K) the H_2 molecules start to dissociate (see Fig. SF1a at time steps $t > 25$ ps). As discussed in the main text, this process is associated to the formation of H-N bonds: Fig. SF1a shows the progressively increasing number of H-N bonds at high temperature (a representative structure is shown in Fig. SF2c).

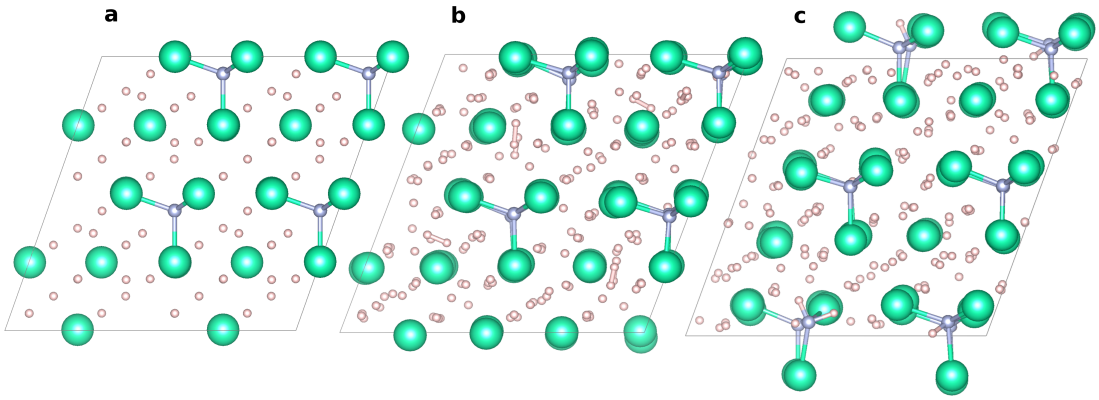

FIG. SF2. MLFF-MD representative structures from the simulation with ramping temperature. We show the initial structure with the crystal in the $Fm\bar{3}m$ phase (panel (a)), the structure after the formation of H_2 molecules at low temperature, T$\sim$ 50 K (panel (b)), and the structure showing the presence of N-H and N-H_2 bonds at high temperature, T$\sim$ 400 K (panel (c)).

II. ADDITIONAL MLFF-MD DATA

In this section we present additional data on N-doped LuH_3 as obtained from our MLFF-MD simulations.

Figure SF3 shows the distance between two H atoms forming a H_2 molecule as a function of time (in the MLFF-MD at 300 K). The average value of the H_2 bond-length is 0.83 Å (see solid line in the figure), approximately 10% larger than the bond-length value of 0.74 Å obtained for the gas phase by DFT calculations (dashed line).

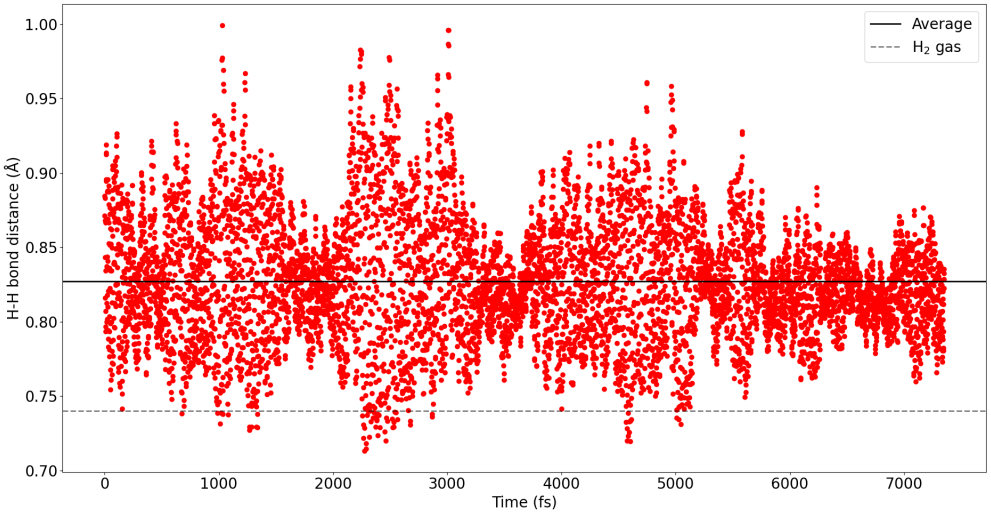

FIG. SF3. Distance between two hydrogen atoms in a H_2 molecule as a function of time. Data refer to the MLFF-MD simulation at 300 K. The solid line represents the average bond-length between two hydrogen atoms in a H_2 molecule, while the dashed line represents the bond-length calculated in the gas phase.

Figure SF4 shows the number of H_2 molecules per N as function of time in the MLFF-MD simulations at 100 and 200 K. At 100 K all molecules survive for the entire simulation. At 200 K some of the molecules dissociate, reaching a number similar as in the simulation at 300 K (see Figure 1 in the main text).

As discussed in the main text, the number of H_2 molecules and H-N bonds determines the metallic/insulating state of the system at every time step: When the total number of short hydrogen bonds (H_2 and N-H) is different from the total number of nitrogen atom in the unit cell, the density of states shows a metallic character. Figure SF5 shows the density of states of a representative metallic structure explored in the MLFF-MD simulation (the

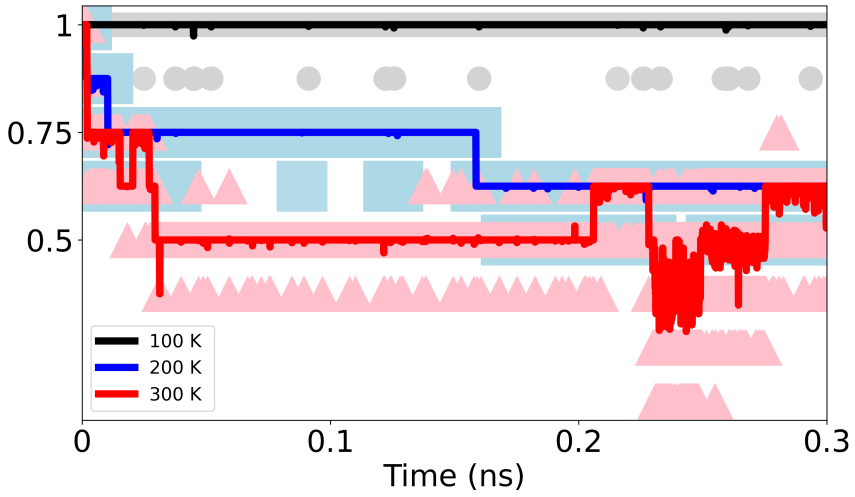

FIG. SF4. Number of H_2 molecules per N atom in the MLFF-MD simulations at 100, 200 and 300 K. We used a threshold distance of 1.0 Å to consider a H-H pair as a H_2 molecule. The circles, squares and triangles indicate the number of H_2 molecules found at every time step in the simulations at 100, 200 and 300 K, respectively. The solid lines represent the running average (calculated over 100 fs).

internal forces were not relaxed, *i.e.*, the atomic structure was kept at the MD positions). We have calculated the density of states for various structures explored in the MLFF-MD simulations, and marked the corresponding metallic regimes in Figure 1b in the main text.

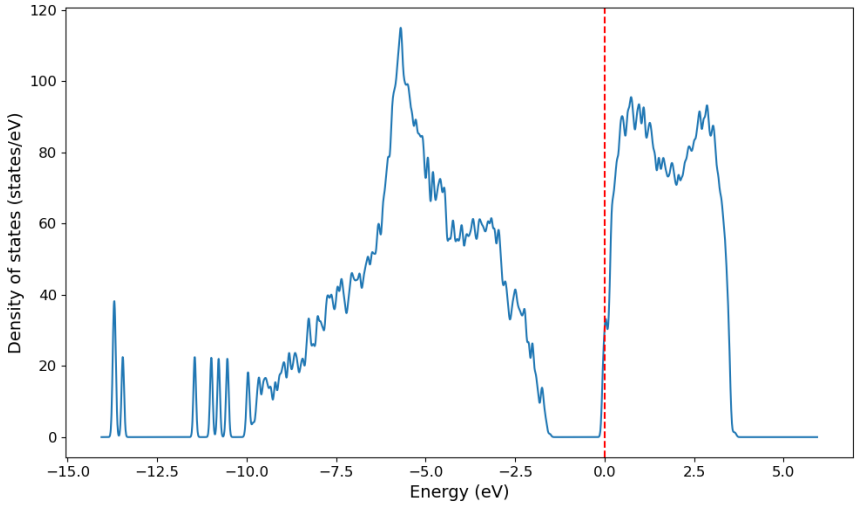

FIG. SF5. Density of states of a representative metallic configuration extracted from the MLFF-MD at time steps showing a number of H_2 molecules and H-N bonds different to the number of N atoms in the unit cell. The internal forces were not relaxed (the atomic structure was kept at the MD positions). The red dashed line indicates the Fermi energy.

III. BADER CHARGE ANALYSIS

Table ST1 shows the Bader charge analysis as obtained from `VASP` calculations for the N-doped LuH_3 system, comparing different cases.

In models including a number of short H-H ($\simeq$ 0.8 Å) and N-H ($\simeq$ 1.0 Å) bonds equal to the number of N dopants, the system show an insulating behavior. In such systems, we assign a valence state of $+3$ to Lu atoms, and -1 and -3 to H and N atoms not involved in the short bonds, respectively. The N-H bond does not perturb to a sizable extent the N valence state, while the H atom in this configuration tends to donate its electron to the crystal. The molecular H atoms do not share large amount of charge with the crystal, showing a small electron doping (Bader charge of 1.1 e per H atom in the H_2 molecule) responsible for the elongation of the molecular bond.

In case the number of short H-H and N-H bonds differ from the number of N impurities, then we observe metallic states. In these configurations, the excess electronic charge is accommodated on the Lu orbitals (the Bader charge goes from 7.3 e in the insulating systems to 7.4 e), in perfect agreement with the electronic band structure and density of states shown in the main text. Whether the N atom is substituted on tetragonal or octahedral sites of

size	model	el. charac.	Lu	H	H in H_2	H in H-N	N	N in H-N
2×	1 H_2, 1 N tetra	insulating	7.3 (+3)	1.6 (−1)	1.1 ($-\varepsilon$)	-	6.6 (−3)	-
2×	2 H_2, 1 N tetra	metallic	7.4 ($+3-\delta$)	1.6 (−1)	1.1 ($-\varepsilon$)	-	6.6 (−3)	-
2×	2 H_2, 1 N octa	metallic	7.4 ($+3-\delta$)	1.6 (−1)	1.1 ($-\varepsilon$)	-	6.6 (−3)	-
4×	3 H_2, 8 N, 5 N-H	insulating	7.3 (+3)	1.6 (−1)	1.1 ($-\varepsilon$)	0.6 ($+1-\theta$)	6.6 (−3)	6.6 (−3)

TABLE ST1. Bader Charge Analysis. Every line refer to a different model, built using either a $2\times2\times2$ (2×) or $4\times4\times4$ (4×) unit cell, including 1, 2 or 3 H_2 molecules, N substitutional doping in tetragonal (tetra) or octahedral (octa) sites, and 5 N-H bonds in the larger cell only (with all N in tetragonal sites). All structures were relaxed to minimize the internal forces (see Methods); the Bader charge and the (metallic or insulating) electronic characters were evaluated from the density of states of the relaxed structures. The average value of the Bader charge is reported for every element (in units of the electronic charge, e), separately; H and N atoms forming short (H_2 and N-H) bonds are distinguished from the rest of the atoms in the crystal. The corresponding valence state value in the ionic bond picture is reported in parenthesis (the reference number of valence electrons of every species in its neutral state reads 9 for Lu, 1 for H and 5 for N); the δ, ε and θ symbols refer to small charge deviation from the ideal values.

the Lu lattice seem not to affect the results for the Bader charge.

The Bader charge analysis contributes to clarify the stabilization of the metallic-vs-insulating character of the system upon formation of short H-H and H-N bonds. The 3 electrons donated by the Lu^{+3} atoms are transferred to the H^{-1} atoms in the undoped system. By substituting one H^{-1} atom, the N^{-3} dopant can accommodate two additional electrons as compared to the H atom in the pristine cell: thus, two other H atoms are now not involved in any sizable charge transfer, and can form a molecular-like H_2 complex (with a Bader charge of 1.1 e per H atom, *i.e.*, one electron per H atom as in the gas phase, plus a small $-\varepsilon$ perturbation induced by the crystal). In this situation, the system is insulating. Similarly, H-N short bonds might form instead of the H_2 molecule, keeping the system insulating as the result of the H atom donating its electron to the crystal (the H atom tends to a H^{+1} valence state, showing a Bader charge of 0.6 e). In case the number of short H-H and H-N bonds is not commensurate with the number of N impurities, our systems show a

metallic character, with excess electronic charge occupying Lu orbitals ($Lu^{+3-\delta}$).

IV. MLFF-MD ON THE UNDOPED COMPOUND

In addition to the MLFF-MD simulations described in detail in the main text, we performed further calculations to inspect the role of nitrogen atoms in the stabilization of the H_2 molecules. To this purpose, we performed MLFF-MD calculations of the undoped LuH_3 compound, starting from the system either in the $Fm\bar{3}m$ phase or in a structure including an high number of manually created H_2 molecules (16 molecules in the $4 \times 4 \times 4$ unit cell, corresponding to 2 molecules per $2 \times 2 \times 2$).

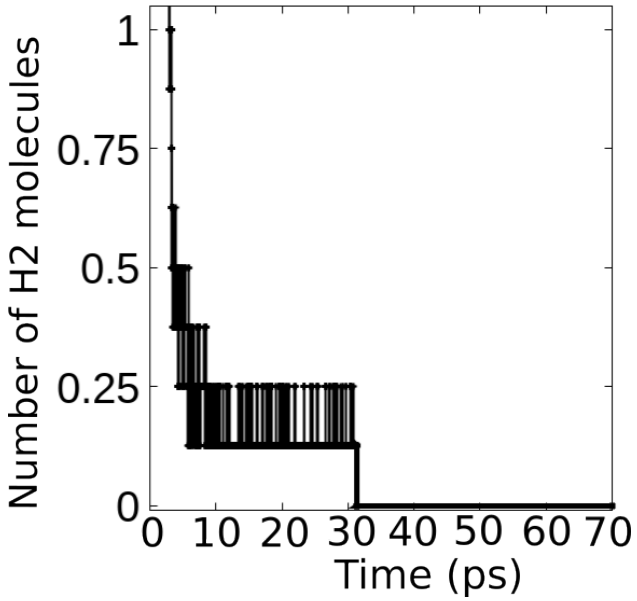

FIG. SF6. Number of H_2 molecules in the undoped LuH_3 $4 \times 4 \times 4$ unit cell in the MLFF-MD simulation at 300 K. For better comparison with the N-doped case, the number of H_2 molecules is reported per $2 \times 2 \times 2$ cell.

Starting from the $Fm\bar{3}m$ phase we do not observe formation of any molecules (results not shown), while in Figure SF6 we report the results obtained from an initial guess in which molecules were artificially formed. Relaxing the internal forces to the local energy minimum, we observe that the H_2 molecules remain still unbroken. However, during the molecular dynamics simulations all molecules have rapidly dissociated (as shown in the Figure for the simulation at 300 K) and after $\sim$ 30 ps the molecules have completely disappeared.

V. MAPPING TO THE REPRESENTATIVE $2 \times 2 \times 2$ SUB-SYSTEM

To study in more details the effect of H_2 molecules, we have modelled two structures representative of the stable MLFF-MD phase, by selecting sub-units in a $2 \times 2 \times 2$ unit cell containing one or two H_2 molecules per N atom and then relaxing the structures in their local energy minimum (see Fig.SF7 and SF2 for comparison).

The phase with one H_2 molecule results lower in energy with respect to the $Fm\bar{3}m$ structure (however, it is dynamically unstable as discussed in the next Section). On the other hand, the one containing two H_2 molecules results energetically higher (0.3 eV/Lu) than the $Fm\bar{3}m$ structure (because of the smaller unit cell with respect to the one used in the MLFF-MD calculations, which does not allow for a complete description of structural disorder). Consistently with the Bader charge analysis, we have found that the system with one H_2

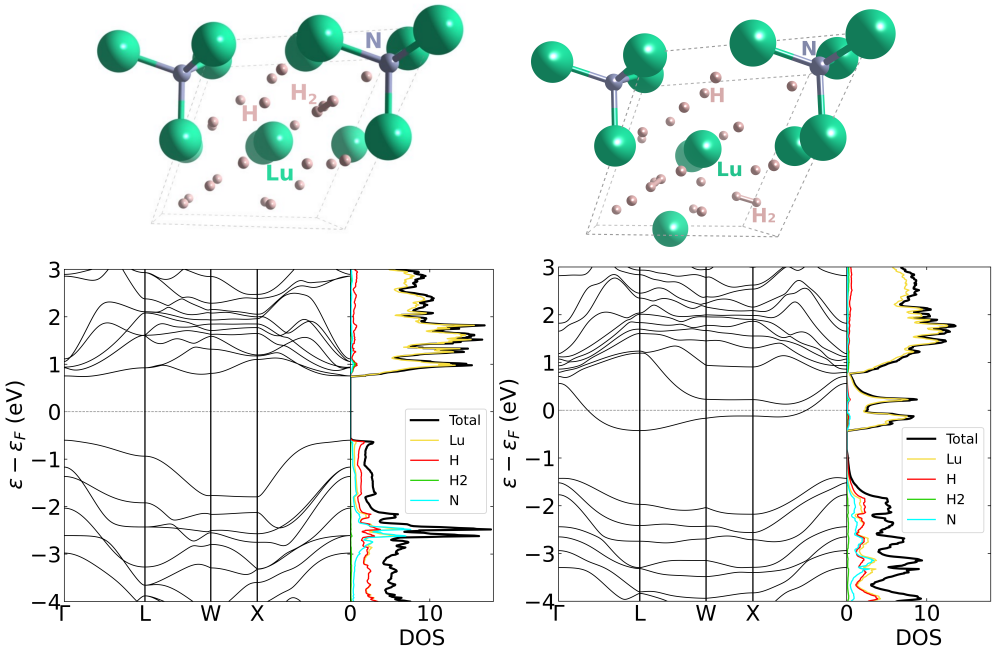

FIG. SF7. In the top panel we report a perspective sketch of both the representative sub-system containing one (left) and two (right) H_2 molecules. In the bottom left panel we report the electronic band structure over the BZ and the projected DOS for $LuH_{2.875}N_{0.125}$ with one H_2 molecule. In the bottom right panel, the same system in a higher-energy configuration involving two H_2 molecules. The atomic contribution to the DOS is highlighted with different colours, see legend.

molecule formed is an insulator, presenting nearly flat bands and van Hove singularities (see

Fig.SF7(bottom-left)) at -2.5 eV. On the other hand, as discussed in the main text, with two H_2 per unit cell the system is metallic and thus explored for possible superconducting properties.

VI. MORE ON THE DYNAMICAL PROPERTIES

In the following we will briefly discuss the dynamical properties of both the representative sub-systems described in the previous section.

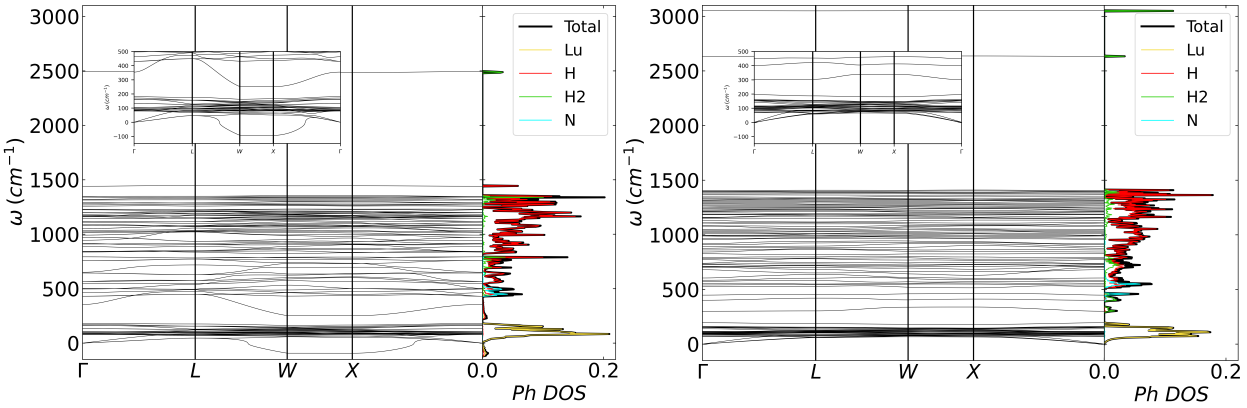

FIG. SF8. Left panel: The phonon spectra obtained for $LuH_{2.875}N_{0.125}$ with one H_2 molecule. Inset: focus on low frequencies region showing the Lu modes (below 200 cm^{-1}) and the first softened modes around 250 cm^{-1}. Right panel: $LuH_{2.875}N_{0.125}$ in its metastable phase (2 molecules) phonon spectra. Inset: focus on low frequencies region showing the Lu modes (below 250 cm^{-1}) and the molecular *librational* modes (around $300 - 500$ cm^{-1}). Each plot is accompanied by its density of states. The hydrogen character of eigenvalues is highlighted (red the total H character, in green the molecular contribution, in cyan the nitrogen one, and in yellow the total Lu character).

As shown in Fig.SF8, the representative sub-unit containing just one H_2 molecule results to be dynamically unstable at the harmonic level. The instability is not so deep and extended in the whole Brillouin zone, as in the case of the pristine $Fm\bar{3}m$ LuH_3 system, and is due to a widespread softening involving hydrogen atoms (see projections in Fig.SF8). The softening also affects other modes, at higher energies, still involving hydrogen atoms.

On the other hand, the representative sub-system containing two H_2 molecules result metallic (Fig.SF7) and dynamically stable.

Despite the instability observed for the first structure, the phonon dispersion and related

DOS appear overall similar. In particular, the Lu-derived modes are found at low frequencies (<250 cm^{-1}), nearly separated from the other branches; N-related branches are located around 500 cm^{-1} and the H and H_2 modes extend up to $\simeq$1500 cm^{-1}. The characteristic molecular vibrational modes are well separated from the rest of the spectrum at very high frequencies.

VII. CALCULATION OF THE SUPERCONDUCTING CRITICAL TEMPERATURE

We used the $ab-initio$ Superconducting density functional theory (SCDFT)[1–3], an extension of DFT to the superconducting phase, to predict the superconducting critical temperature of the studied systems.

We solved the SCDFT gap equation using the functional form proposed in Ref. 4, in the isotropic approximation using the calculated α^2F and the Sham-Kohn approximation[5] for the electron self-energy for the normal state repulsion. The result are reported in Fig.SF9.

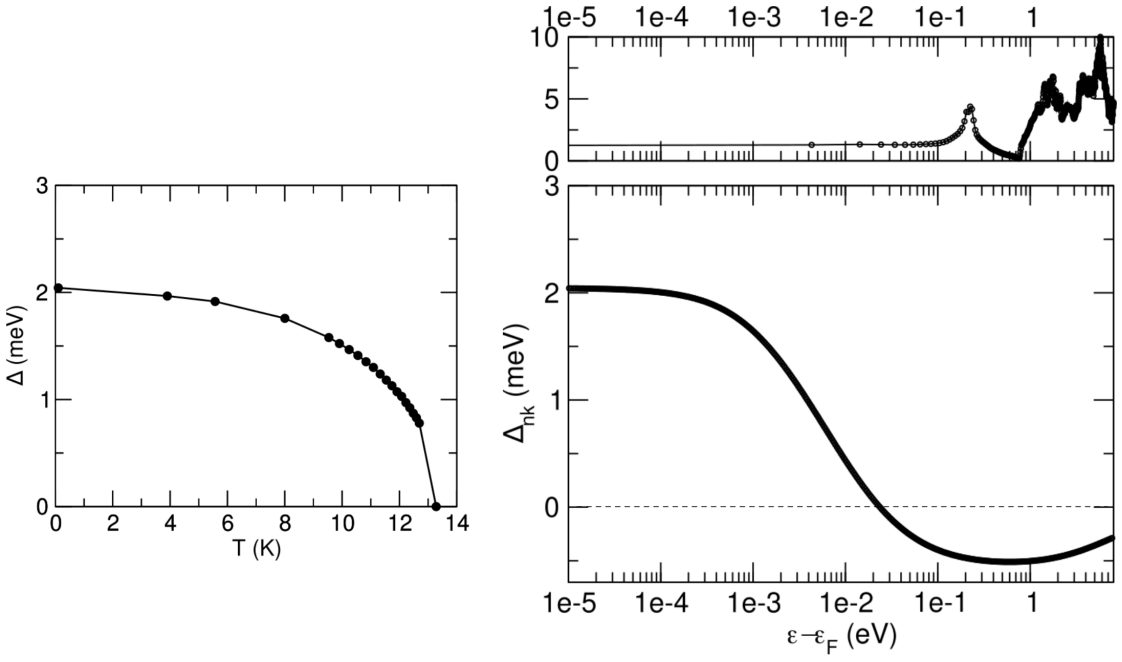

FIG. SF9. Left panel: The superconducting energy gap as a function of temperature for the metallic phase with two H_2 molecules. Right panel: The superconducting energy gap at "zero" temperature as a function of the energy from the Fermi level and the electronic DOS in (states/spin/eV) on the same energy scale (on top).

The calculated superconducting gap at zero temperature is $\Delta(0) \simeq 2$ meV and it goes to zero at a critical temperature of $T_C \simeq 13$ K, with a ratio $\Delta(0)/K_B T_C = 1.78$, in line with the BCS universal ratio of 1.764 for weak coupling superconductors.

The density of states close to the Fermi energy is nearly constant and the superconducting gap results positive up to $\sim 10^{-2}$ meV, an energy distance of the order of the phononic energy range ($\omega_{log} \simeq 24$meV), then it changes sign at higher energies. Interestingly, the van Hove singularity falls into the negative region of the gap, contributing to the gap renormalization.

We have estimated the superconducting critical temperature also using Allen and Dynes formula resulting in $T_C \simeq 8$ K, with $\mu^* = 0.1$, slightly lower than the SCDFT value probably due to the different treatment of renormalization effects due to the Coulomb interaction.

VIII. ACCURACY OF MLFF-MD CALCULATIONS

During the molecular MLFF-MD dynamics simulations, at each time step, the machine-learned force field predicts energy, forces, and the corresponding Bayesian error estimations [6, 7]. In our setup, if the Bayesian error is above a certain threshold, an *ab-initio* calculation is performed, and the reference energy and forces are added to the training data set. If the error is below the threshold, the ab-initio step is omitted, and the system is propagated via machine learning force-field predictions. The threshold is dynamically updated. Figure SF10 shows the Bayesian error estimation on the forces (BEEF) as a function of time, and the corresponding threshold for the DFT calculations.

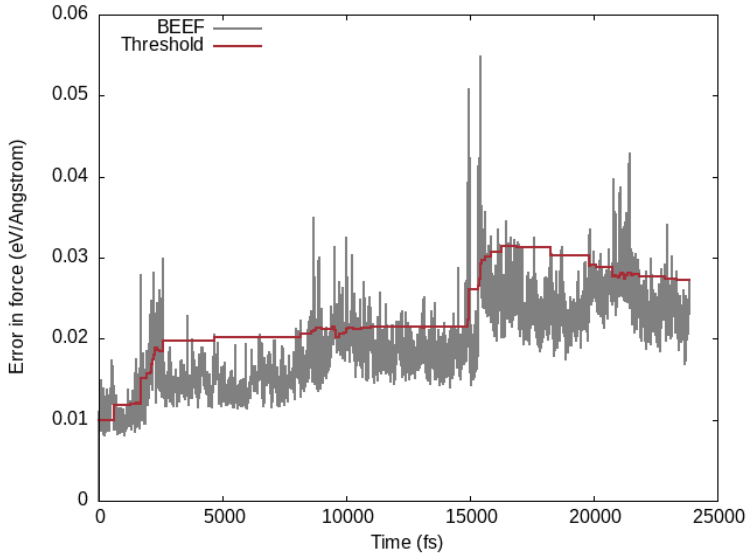

FIG. SF10. Estimated errors in the MLFF-MD run at 300 K. The red line represents the threshold activating the DFT calculations; the gray line represents the Bayesian error estimate of forces (BEEF).

[1] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Density-functional theory for superconductors, Phys. Rev. Lett. **60**, 2430 (1988).

[2] M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross, Ab initio theory of superconductivity. i. density functional formalism and approximate functionals, Phys. Rev. B **72**, 024545 (2005).

[3] M. A. L. Marques, M. Lüders, N. N. Lathiotakis, G. Profeta, A. Floris, L. Fast, A. Continenza, E. K. U. Gross, and S. Massidda, Ab initio theory of superconductivity. ii. application to elemental metals, Phys. Rev. B **72**, 024546 (2005).

[4] A. Sanna, C. Pellegrini, and E. K. U. Gross, Combining eliashberg theory with density functional theory for the accurate prediction of superconducting transition temperatures and gap functions, Phys. Rev. Lett. **125**, 057001 (2020).

[5] L. J. Sham and W. Kohn, One-particle properties of an inhomogeneous interacting electron gas, Phys. Rev. **145**, 561 (1966).

[6] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with bayesian inference, Physical Review Letters **122**, 225701 (2019).

[7] R. Jinnouchi, F. Karsai, and G. Kresse, On-the-fly machine learning force field generation: Application to melting points, Physical Review B **100**, 014105 (2019).