Part III.

Nature of Ar Bonding to Small Co_n^+ Clusters and its Effect on the Structure Determination by Far-Infrared Absorption Spectroscopy

7. The Vibrational Fingerprint of Small Co_n^+ Clusters and their Ar Complexes

7.1. Introduction

The determination of atomic cluster structures constitutes a nice example of a mutual support between experiment and theory. In case of a pure computational treatment of the problem, there is never a guarantee to find the global minimum or all relevant higher-lying isomers, regardless how sophisticated the sampling scheme is designed. Furthermore, for the cluster size range of interest, where highly accurate quantum chemical methods become prohibitively expensive, there is no systematic way to converge the quality of the energetics. As a consequence, a poor description of the underlying PES might yield wrong isomers. Experimentally, it is impossible to directly measure the atomic structure but only quantities that are correlated with the underlying configuration. Theory is therefore required to interpret the measured data by simulating the experiment based upon putative cluster structures. If the computed and measured data match, the candidate structure can be considered as confirmed.

Experimentally, obtaining information on the atomic arrangement of such small clusters in the gas phase remains a challenge. Investigations of the reactivity of metal clusters with small molecules, in particular their saturation behaviour, have been used to draw conclusions about the clusters atomic arrangements [97]. Photoelectron spectroscopy provides direct information on the electronic structure of clusters which can be related to their geometries [98]. Information on the shape of cluster ions can be obtained from ion mobility measurements [99, 100, 101] and, recently, electron diffraction of trapped charged clusters has shown to be promising for revealing their structure [102, 103, 104]. Another possibility that is particularly sensitive to the internal cluster structure is the measurement of vibrational frequencies. A corresponding technique that has recently been successfully employed to determine the structure of cationic and neutral metal clusters containing between three and 20 atoms is far-infrared (vibrational) resonance enhanced multiple photon dissociation (FIR-MPD) spectroscopy [1, 2, 3, 4, 5]. The essential idea of this technique is to irradiate rare-gas complexes of the targeted clusters with infrared (IR) light in the range of the structure-specific vibrational fundamentals. When the IR light is resonant with a vibrational mode in the cluster complex, the complex can absorb several photons and subsequently evaporate off one or more rare-gas atoms. Recording the resulting abundance changes of the rare-gas complexes

as a function of the IR frequency yields the desired spectra that can then be compared to computed IR absorption spectra for different isomer structures obtained e.g. using density-functional theory (DFT). As long as different isomers exhibit distinct vibrational fingerprint patterns this then enables a unique determination of the atomic structure.

Performing the measurements on the rare-gas metal complexes, a fundamental assumption behind this approach is that the rare-gas atoms do not significantly influence the vibrational spectrum and merely act as a probe for detecting the resonant absorption of IR photons by the metal cluster [105]. A negligible influence of the employed rare-gas atoms was indeed inferred from only insignificant differences in the IR spectra of previously studied cationic V_n^+ ($n = 3 - 23$) [1, 2], Nb_n^+ ($n = 5 - 9$) [3, 4] and Ta_n^+ ($n = 6 - 20$) [5] complexes that contained one or more Ar atoms. Restricted test calculations comparing the IR spectra of bare and Ar-complexed V_3^+ and V_4^+ clusters led to the same conclusion [1, 2], so that the ensuing theoretical modeling for the mentioned systems focused exclusively on the IR spectra of the bare metal clusters.

The situation is markedly different for small cationic Co clusters. Here, the measured FIR-MPD spectra show an intriguing dependence on the number of adsorbed Ar atoms, which is strongest for clusters containing less than six cobalt atoms. The motivation for this work is therefore to investigate on the nature of the Ar-Co_n^+ bond and its implications for the vibrational spectra instancing the series Co_4^+ to Co_8^+. Among the sampling algorithms presented in chapter 3, basin-hopping was chosen in a spin-extended version to sample both the configurational and spin space on an equal footing. Additionally, the Ar probe atom was explicitly taken into account, thus identifying the Ar-binding sites in a completely unbiased way. In this chapter, the comparisons of the computed IR spectra with the measured data are presented, which allow in some cases for a unique assignment of the experimental cluster structure. The calculations reveal furthermore a characteristic increase of the Ar binding energy for the smaller clusters, rationalizing the observed increased influence of the Ar probe atoms on the IR spectra.

In the next chapter, the binding energy trend is further analysed and traced back to the predominant contribution to the Ar-Co_n^+ bond arising from the polarization of the rare-gas atom in the electrostatic field of the cationic cluster. This motivates a simple electrostatic model, that not only reproduces the binding energy trend, but also explains why the interaction of Ar is much stronger than in the previously studied systems, where little influence of the probe atoms on the vibrational spectra had been observed.

7.2. Far-Infrared Absorption Spectroscopy

7.2.1. Experimental Setup

The experiments were carried out in the group of Prof. Gerard Meijer at the Molecular Physics Department of the FHI. They employed a molecular-beam setup [2, 4, 105] that is coupled to a beam line of the Free Electron Laser for Infrared eXperiments (FELIX) at the FOM-Institute for Plasma Physics in Nieuwegein, The Netherlands [106]. Metal clusters were produced by pulsed laser ablation of a cobalt rod using the 2$^{\text{nd}}$ harmonic

output of a Nd:YAG laser and by subsequent condensation of the plasma in a mixture of Ar in He. Neutral, anionic, and cationic clusters were produced in this process and passed through a temperature controllable copper channel. Only cationic clusters were investigated in this study. Two different experimental conditions had been used. With a mixture of 0.1 % Ar in He at liquid-nitrogen-temperature, clusters containing six or more cobalt atoms bound one and to a much lesser degree two argon atoms. At the same conditions small clusters up to Co_5^+ formed complexes with up to five argon ligands. At a higher temperature of 340 K and using a mixture of 0.3 % Ar in He no attachment to bigger cobalt clusters was observed, whereas Co_4^+ and Co_5^+ added one and two rare-gas atoms which is in line with similar observations by Minemoto *et al.* [107].

The molecular beam expanded into vacuum and passed through a skimmer and an aperture before entering the extraction region of a reflectron time-of-flight mass spectrometer. A counter-propagating pulsed far-IR laser beam delivered by FELIX was overlapped with the aperture to ensure that all species in the beam that were detected in the mass spectrometer had been exposed to the IR radiation. The laser pulse consisted of a 1 GHz train of ps-duration micropulses. The duration of such a macropulse is several μs while its energy ranged typically between 15 and 30 mJ, depending on the wavelength. When the IR radiation is resonant with an IR-allowed transition of the cluster complex, sequential absorption of single photons can take place [108]. The resulting heating of the cluster may induce evaporation of the argon atoms, which leads to depletion of the complexes in the beam. IR depletion spectra were constructed by recording the ion intensities of the metal argon complexes as a function of the FELIX frequency; from these, the far-infrared absorption spectra were obtained as described in ref. [105]. As the detection is mass selective, the simultaneous measurement of far-IR spectra for different cluster sizes is possible.

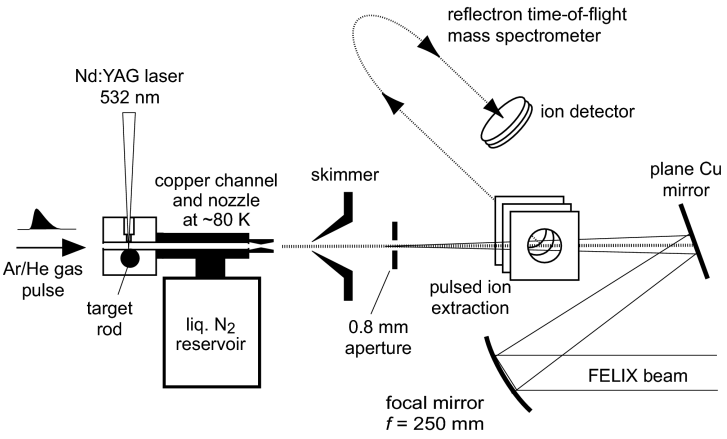

Figure 7.1.: Scheme of the experimental set-up used for IR multiple photon dissociation of metal cluster-rare gas complexes.

7.2.2. Observation of Highly Coordinated $Co_n^+Ar_m$ Complexes

For both employed complex formation conditions (77 K/0.1% Ar vs. 340 K/0.3% Ar) the recorded mass spectra of the cationic cobalt-argon complexes reveal profound differences in the binding of the rare-gas ligands to clusters containing more or less than six atoms. As shown in Fig. 7.2, at the low-temperature conditions, in which the bigger clusters

bind only one or at maximum two Ar atoms, Co_4^+ and Co_5^+ readily attach up to five ligands. The affinity of these small clusters to the rare-gas atoms is so pronounced that they still form complexes with one or two Ar atoms even at a temperature as high as 340 K. This change of the interaction with the probe atoms with cluster size is also reflected in the obtained FIR-MPD spectra, which were measured in the range 75-350 cm^{-1} with no IR absorption bands detected at higher wavenumbers. The spectra show an intriguing dependence on the number of adsorbed Ar atoms, which is strongest for the smallest clusters and is essentially lost for clusters of more than seven atoms.

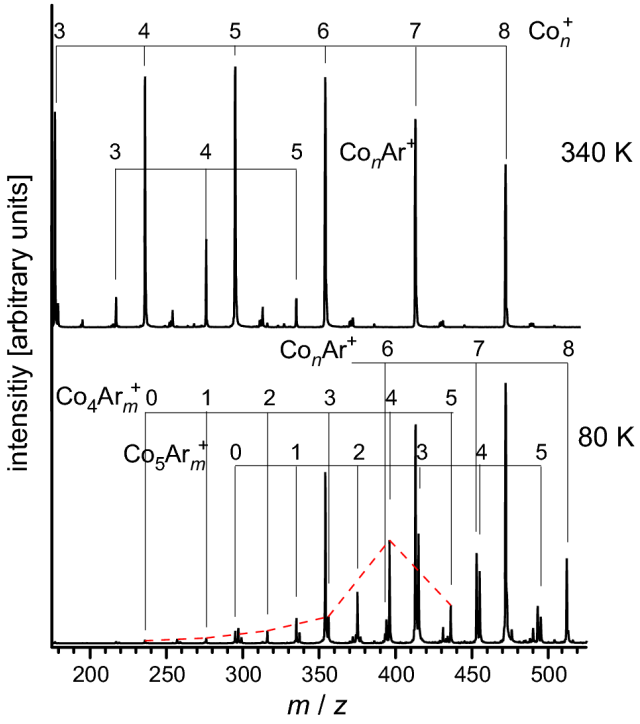

Figure 7.2: Distribution of cationic cobalt clusters and their Ar complexes obtained at source temperatures of 340 K (top) and 80 K (below). At the higher temperature rare gas complexes are only observed up to Co_5^+, whereas at low temperature also larger clusters attach Ar and the small clusters bind multiple Ar atoms. The dashed red line in the lower plot connects the maxima of the $Co_4Ar_m^+$ peaks.

7.2.3. Interpretation of FIR-MPD Spectra

Compared to linear absorption spectra, FIR-MPD spectra are known to be subject to distortions, such as an incomplete fragmentation of complexes containing more rare-gas atoms and (cross)anharmonicities in the multiple photon excitation process [105]. For the prior effect one has to keep in mind that there is no mass selection before the fragmentation of a certain complex. The spectrum of a given complex thus contains both a contribution from the depletion due to evaporation of the argon ligand as well as an increase in intensity due to incomplete fragmentation of complexes containing more argon ligands. This can lead to missing bands especially at the low-frequency end of spectra and is avoided by considering only the spectra of a certain cluster size with the maximum number of argon ligands. The effect of (cross)anharmonicities is instead rather to change the width and positions of existing lines. Corresponding effects have been observed in preceding work on cationic V_n^+ $(n = 3 - 23)$ [1, 2], Nb_n^+ $(n = 5 - 9)$

[3, 4] and Ta_n^+ ($n = 6 - 20$) [5] complexes. However, the influence of the rare-gas atoms was typically small, with only minor shifts of the frequencies of the absorption bands of the order of a few cm^{-1} and with the relative intensities of the observed bands barely affected. This situation is markedly different in the case of the small cationic Co clusters, where the spectra from different Ar containing complexes exhibit much more pronounced frequency shifts and rather strong changes in the relative intensities, cf. Figs. 7.6 and 7.7 below.

7.3. Computational Details

The DFT calculations are carried out using the all-electron full-potential code FHI-aims [11], described in part II, and the generalized gradient approximation (GGA-PBE) [70] to describe electronic exchange and correlation (xc). For comparison, some calculations are also performed using the local density approximation (PW-LDA) [69]. All calculations reported here are done with the "tier2" basis set which contains 67 and 45 basis functions for Co and Ar, respectively and are further detailed in Appendix A. As integration grids, we used $N_{\mathrm{r,div}}$= 2, corresponding to 158 and 140 radial shells for Co and Ar, respectively. The angular grid is found to be converged for $N_{\mathrm{ang,max}}$=302 with a fixed $N_{\mathrm{ang,min}}$=110. The cut-off radius was conservatively chosen with $r_{\mathrm{cut}} = 5$ Å for both species, as well as the maxium angular momentum in the multipole decomposition l_{max} with a value of 6. Convergence tests focused on the D_{2d} ground state (S=7/2) and the higher-lying D_{2h} isomer (S=9/2) of Co_4^+Ar as well as the D_{3d} ground state (S=15/2) and higher-lying C_{2v} isomer (S=13/2) of Co_6^+Ar (*vide infra*), and are given in detail in the Appendix B. Using the hierarchically constructed larger basis sets provided in FHI-aims, as well as denser integration grids, we recomputed the relaxed geometric structure and vibrational spectra of the isomers, as well as their energetic difference

$$\Delta E = E_{\mathrm{tot}}(\mathrm{isomer1}) - E_{\mathrm{tot}}(\mathrm{isomer2}) \,, \tag{7.1}$$

and the Ar binding energy

$$E_{\mathrm{b}} = E_{\mathrm{tot}}(Co_n^+Ar) - E_{\mathrm{tot}}(Co_n^+) - E_{\mathrm{tot}}(Ar) \,, \tag{7.2}$$

and found these quantities to be converged within 0.01 Å, 2 cm^{-1} and 10 meV, respectively, which is fully sufficient for the arguments and conclusions put forward below. Furthermore, no electronic smearing method as described in section 5.1.1 was required to achieve self-consistency. All calculated results are therefore based upon an undistorted energy surface.

7.3.1. Structural Relaxation and Calculation of the IR-Spectra

Local structural optimization is done using the Broyden-Fletcher-Goldfarb-Shanno method described in section 3.1.3 [27], relaxing all cartesian force components to smaller than 10^{-2} eV/Å. After relaxation, the Hessian matrix is determined numerically by finite

displacements of all atomic positions by $\Delta = 10^{-3}$ Å, and then diagonalized to get the vibrational modes. The corresponding infrared intensities are obtained by taking the derivative of the dipole moments along these modes [109]

$$I_i^{\mathrm{IR}} = \frac{N_{\mathrm{AV}}\pi}{3c}\left|\frac{d\mu}{d\mathbf{Q}_i}\right|^2 , \tag{7.3}$$

where the derivatives along the eigenmodes can easily be obtained by the derivatives with respect to cartesian coordinates, the latter being calculated along the numerical calculation of the Hessian by finite displacements.

$$\frac{d\mu}{d\mathbf{Q}_i} = \sum_\alpha \frac{d\mu}{d\mathbf{R}_\alpha}\cdot\mathbf{Q}_{\alpha,i} \tag{7.4}$$

The influence of the force convergence criterium as well as the finite displacement Δ on the calculated quantities has been carefully tested. Details on the test calculations focusing on Co_4^+Ar and Co_6^+Ar are given in the Appendix B. In order to facilitate the visual comparison to the experiment, the resulting IR spectra are finally folded with a Gaussian line shape function of half-width 1 cm^{-1}.

7.3.2. The Counterpoise Correction to the Basis Set Superposition Error

As already mentioned in section 5.2, the numerical atom-centered basis sets prevent an atomic BSSE by construction, since the total energy of the free atom is already converged at the minimum basis set level and can therefore not be variationally improved by further basis functions originating from neighbouring atoms. In the case of the Ar binding energies, calculated according to Eq. (7.2), however, the reference energy is not exclusively constituted by free atoms, but partially by the total energy of the bare Co_n^+ cluster. Hence, it is *a priori* not clear, whether the neighbouring Ar basis functions improve the total energy of the bare cluster or not. To investigate on that, the counterpoise correction of quantum chemistry [110, 111] was applied to the four test isomers of Co_4^+Ar and Co_6^+Ar as

$$\Delta^{\mathrm{CP}} = E^{\mathrm{CP}}_{\mathrm{fragment}}(Co_n^+) - E_{\mathrm{fragment}}(Co_n^+) . \tag{7.5}$$

To compensate for a potential BSSE, the total energies of the Co_n^+ fragments are hereby calculated using both the bare Co basis sets ($E_{\mathrm{fragment}}(Co_n^+)$), and as additional *ghost* basis functions the Ar basis functions originating from an empty site at the position of the Ar atom in the Co_nAr^+ complex $E^{\mathrm{CP}}_{\mathrm{fragment}}(Co_n^+)$. In this context, "fragment" means that the Co_n^+ structure is cut out from the Ar complex without further relaxing it. The difference Δ^{CP} is then a measure for the variational improvement of the total energy due to the Ar basis functions which is responsible for the overestimation of the binding energy. Hence, the corrected binding energy is obtained by

$$E_b \longrightarrow E_b - \Delta^{\mathrm{CP}} . \tag{7.6}$$

The results based upon the tier2 basis set for Co and Ar are summarized in Table 7.1. Clearly, the BSSE is below 10 meV and thus negligible at this basis set level. This conclusion is in line with ref. [11] and the counterpoise correction was therefore neglected in this work.

		Δ^{CP} [eV]
Co_4^+Ar	D_{2d}	0.005
	D_{2h}	0.005
Co_6^+Ar	D_{3d}	0.004
	C_{2v}	0.006

Table 7.1.: Corrections to the Ar binding energy due to the intermolecular BSSE.

7.3.3. Zero-Point Energy

Even at zero temperature, the total energy is increased by the zero-point energy resulting from the vibrational motion according to Eq. (2.36). In the calculation of the Ar binding energy, however, most of the zero point energy is expected to cancel in the difference $E_{\mathrm{tot}}(Co_n^+Ar) - E_{\mathrm{tot}}(Co_n^+)$, since the influence of the Ar on the vibrational spectra is rather small. The resulting differences in the zero point energies due to changes in the vibrational spectra are presented in Table 7.2 for the above introduced test systems. As already mentioned in the introduction, the influence of Ar on the vibrational spectrum increases with decreasing cluster size. Nevertheless, even the smallest isomer Co_4^+ only exhibits a correction to the binding energy of 12 meV. Therefore, zero-point corrections to the binding energies are negligible and are not considered in this work.

		ΔE_{ZPE} [eV]
Co_4^+Ar	D_{2d}	0.012
	D_{2h}	0.011
Co_6^+Ar	D_{3d}	0.007
	C_{2v}	0.008

Table 7.2.: Corrections to the Ar binding energy due to zero-point energy.

7.4. Selection of Structural Motifs by Spin-Extended Basin-Hopping

In order to identify energetically low-lying isomers all structures for the bare clusters were tested that had been discussed previously in the literature. For increasing cluster sizes this approach is likely to miss relevant geometries. This holds in particular for Co, where one has to expect several, differently Jahn-Teller distorted versions of the same structural motif. In addition, extended basin-hopping [35] runs have therefore been run to achieve a better sampling of configurational space as described in section 3.2.2.

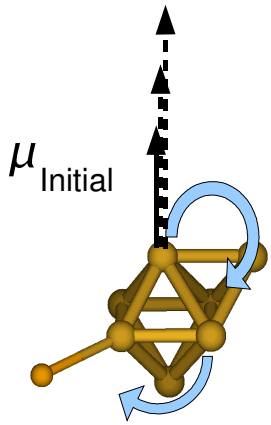

Figure 7.3.: The principle of the spin-extended basin-hopping method. Additional to a random displacement of the atomic positions, the initial spin moments sitting on each atom are randomly varied to allow the system to adopt different spin states.

As specific implementation so-called single-particle trial moves were chosen, in which a randomly picked atom is displaced in a random direction by a distance of 3.00 Å (corresponding to 1.5 times the GGA-PBE computed Co dimer bond length). In order to prevent an entropy-driven dissociation of the cluster during the BH run, trial moves as well as local relaxations were discarded that generate loosely connected or partly dissociated structures characterized by an atom having a nearest-neighbour distance larger than twice the dimer bond length. Similarly discarded were moves that place atoms at distances of less than 0.5 Å from each other.

In view of the specific electronic configuration of Co, the spin degree of freedom was furthermore included into the sampling procedure, i.e. after the trial atomic displacement the electronic self-consistency cycle was initiated with a random magnetic moment. For specific isomers this search was supplemented by fixed-spin moment calculations enforcing spin-states that had not been found in the BH runs. However, in almost all cases these states turned out not to be local minima of the spin hypersurface, suggesting that the chosen spin sampling in the BH runs yields the relevant spin minima. Equivalent BH runs were performed for the $Co_n^+Ar_m$ complexes, randomly displacing both Ar and Co atoms. The most favourable structures determined by this completely unbiased treatment of the Ar probe atoms were always Co clusters with Ar adsorbed at a top site, with Ar-Co bond lengths that range for the most stable adsorption site from 2.45 Å for the Co_4^+ ground state to 2.63 Å for the most stable isomer of Co_8^+.

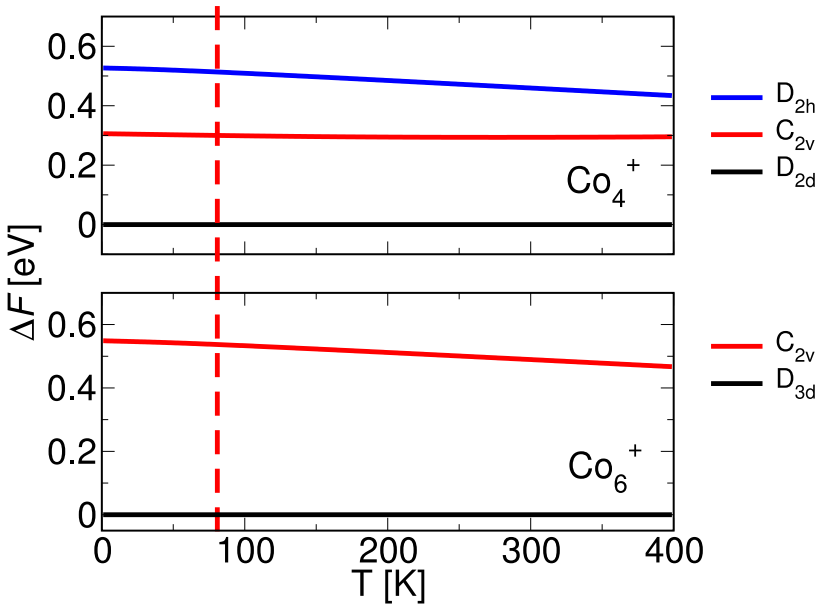

Figure 7.4.: Differences in the free energies of low-lying Co_4^+ and Co_6^+ isomers. The red dashed line depicts the temperature at which the presented IR spectra have been measured.

7.5. Thermodynamic Effects

Since the experiments were performed at finite temperatures, it can *a priori* not be ruled out that thermodynamic effects might change the energetic order of the isomers. Particularly, the vibrational frequencies and principal moments of inertia, determining both the vibrational and rotational contribution to the free energy, are sensitive to the atomic structure. The free energies according to Eq. (2.30) have therefore been calculated for the three Co_4^+ and two Co_6^+ isomers, discussed below, over a wide temperature range including the cryogenic temperature of T=77 K at which the measurements have been performed. As one can clearly see, the free energy differences do not change at all. We conclude that thermodynamic effects are insignificant for the present study and will no further be discussed.

7.6. The Influence of the Exchange-Correlation Functional

In general, there is little reason to expect that DFT with present-day local and semi-local xc functionals yields an accurate description of the intricate electronic structure of such small Co clusters. The difficulty of electronic structure calculations to properly account for the delicate balance between exchange and correlation, between magnetism and chemical bonding, and between populating d and s orbitals in these systems is al-

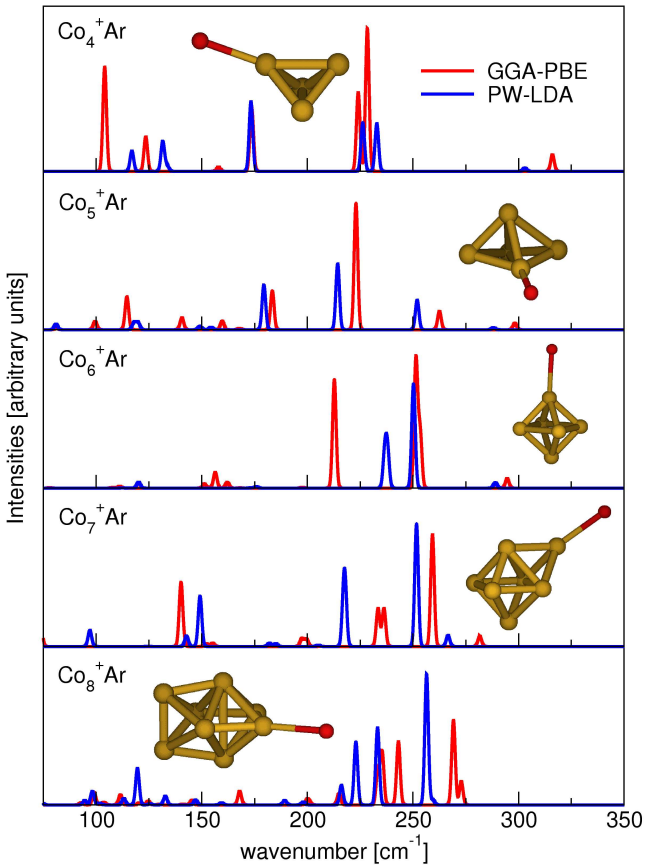

Figure 7.5.: Comparison of computed vibrational spectra of the identified most stable isomers of Co_4^+Ar to Co_8^+Ar and using the PW-LDA (blue line) or GGA-PBE (red line) functional. Applied is a common scaling factor, matching the main peak of the experimental Co_6^+Ar spectrum. See text and Figs. 7.6-7.10 below for a description of the different isomers.

ready well exemplified by the still controversial predictions of the true ground state of the neutral Co dimer (see e.g. ref. [112] and refs. therein). With respect to vibrational frequencies a common procedure to correct for some (systematic) errors due to an approximate xc treatment is to apply a general scaling factor to the calculated frequencies to bring the computed spectra in better agreement with the measured ones. The limitations of this approach for the cationic Co clusters is, however, well apparent from Fig. 7.5, which compiles the computed and scaled vibrational spectra of the identified most stable isomers of Co_n^+Ar complexes in the targeted size range (*vide infra*), and compares them to the corresponding spectra obtained using the PW-LDA functional. In light of the particularly controversial situation for the dimer, we did not follow the usual practice to determine the scaling factor from the comparison of the computed and experimental dimer vibrational frequency, but used the main peak of the simple fingerprint spectrum of Co_6^+Ar, cf. Fig. 7.8 below, instead. However, regardless of which reference is actually used, the validity of such a common scaling approach would have shown up in basically

identical spectra for all cluster sizes after scaling the PW-LDA and GGA-PBE data, which is obviously not the case.

In this situation, it makes little sense to strive for a quantitative agreement of the absolute frequencies when comparing the experimental to the theoretical spectra. Additionally considering the mentioned differences between linear IR absorption and FIR-MPD spectra, the best one can hope for is if at all a match between (relative) spacings between different peaks in the spectrum, or more modestly to identify some characteristic fingerprint pattern that is unique to a specific isomer and which would then allow for a structural assignment. It is within this perspective that we augment the detailed presentation of the low-lying isomers of Co_4^+ to Co_8^+ identified in the BH runs with a comparison of the computed and measured vibrational spectra of the corresponding Ar complexes. Despite the discussed limitations, we apply thereby a global scaling to the theoretical spectra simply to enable a better visual comparison. Using the main peak of the experimental Co_6^+Ar spectrum as reference, this yields a global scaling factor of 0.881 for GGA-PBE and 0.749 for PW-LDA, respectively.

7.7. Low-lying Isomers and Vibrational Spectra of their Ar Complexes

In the following, an overview of the putative ground-state structures and higher-lying isomers for the series $Co_4^+Ar_m$ to $Co_8^+Ar_m$ is given. For the sake of a concise overview, only a selection of the metastable structures is presented. Vibrational spectra have thereby been preferentially chosen that exhibit similar qualitative fingerprints, thus being candidate structures apart from the ground-state isomer, as well as completely different fingerprints, exemplifying identified isomers that can safely be ruled out. Shown are the spectra of the bare Co_n^+-clusters together with their Co_n^+Ar-complexes, where the most stable binding site for the Ar is depicted. For the smaller cluster structures, where the experiment yielded complexes with several Ar atoms bound to the Co_n^+ cluster, the computational counterparts are additionally presented and discussed. An exhaustive overview of all calculated structures, including all different structural motifs, spin states, Jahn-Teller distortions, and binding sites for the Ar probe atoms, can be found in the Appendix E.

7.7.1. Co_4^+

The calculations for Co_4^+ reveal as most stable isomer structure a distorted tetrahedron with D_{2d} symmetry and a magnetic moment of $7\,\mu_B$. This still quite symmetric geometry has two (opposite) bonds of 2.50 Å length, while the remaining four bonds are 2.17 Å. The next lowest structure determined is already 0.32 eV higher in energy, has a magnetic moment of $9\,\mu_B$ and corresponds to an even further distorted tetrahedron, in which one bond is elongated to 2.90 Å to yield a kind of butterfly geometry. Even further up in energy (0.53 eV) is the lowest-energy planar structure, with $9\,\mu_B$, a D_{2h} symmetry and bond lengths of 2.22 Å and 2.38 Å. The thus obtained energetic order is consistent with

the interpretation of photodissociation spectroscopy data favouring a three-dimensional structure for the cationic Co tetramer [107]. A corresponding preference for a pyramidal ground state was also obtained in earlier DFT calculations for neutral Co_4 clusters, albeit with a much smaller energetic difference to the lowest lying planar isomer [113, 114]. Theory predicts hence similar geometries for the neutral and cationic Co_4 ground state, which is in line with a corresponding interpretation derived from measured bond energy data [115].

From this quite consistent picture of the structural motif of the Co_4^+ ground state it is somewhat unexpected to achieve only a rather poor comparison of the computed vibrational spectra of its Ar complexes with the corresponding FIR-MPD measurements. Turning first to the experimental data shown in Fig. 7.6, pronounced differences are observed for the spectra derived from Co_4^+Ar, $Co_4^+Ar_2$ and $Co_4^+Ar_5$. This concerns frequency shifts and changes in the relative intensities of a group of presumably three modes in the range 160-210 cm^{-1}, as well as the appearance of a strong mode at 110 cm^{-1} and tentatively a small mode at 120 cm^{-1} only in the spectrum from the complex with five Ar atoms. As already mentioned, such apparent missing peaks at the low-frequency end of the spectrum from complexes having only one or two Ar atoms are not uncommon in FIR-MPD, and are generally ascribed to an incomplete fragmentation of complexes containing more rare-gas atoms [105]. The frequency shifts particularly of the central peak observed for the group at higher wavenumbers are, however, much stronger than those in preceding work on cationic clusters from group 5 of the Periodic Table [1, 2, 3, 4, 5]. The spectrum of the $Co_4^+Ar_5$ complex, however, cannot be affected by an incomplete fragmentation of complexes containing even more Ar ligands. From the relative abundances of $Co_4^+Ar_4$ and $Co_4^+Ar_5$ in the mass spectrum shown in Fig. 7.6 the latter complex is presumably best described as $Co_4^+Ar_4$ with an additional rather loosely bound Ar atom. Correspondingly, the spectrum derived from $Co_4^+Ar_5$ should give a most faithful representation of the vibrational frequencies of a $Co_4^+Ar_4$ complex.

The computed IR absorption spectra of the most stable D_{2d} isomer complexed with one or two Ar atoms reproduce the strong low-frequency mode observed in experiment quite well, cf. Fig. 7.6. Comparing to the also shown computed spectrum of the bare cluster it is furthermore clear that this agreement is only achieved because of the explicit consideration of the rare-gas atom(s) in the calculations. Only the unexpectedly strong binding of the Ar atom of 0.3 eV breaks the symmetry of the bare cluster and leads to the appearance of IR-active modes in the low-frequency range at all. This unusual strength of the Ar-Co_4^+ interaction prevails for up to four ligands, each time filling one of the top sites at the four corners of the pyramid. Presumably due to the lacking dispersive interactions in the semi-local xc functional, it was, however, not possible to reliably bind a fifth Ar atom to the cluster at the GGA-PBE level, which unfortunately prevents a direct comparison to the measured spectrum from the $Co_4^+Ar_5$ complex. Further complications in the experiment-theory comparison arise when turning to the higher-frequency group of bands in the experimental data. In contrast to the more or less evenly spaced three main peaks in this group, theory predicts a large spacing between one lower-energy peak and a higher-energy doublet. While the frequencies of the latter correspond quite well to the lower and higher frequency main peak of the experimental

group, either the central and dominant experimental peak is completely missing or the splitting of the doublet is significantly underestimated in the theoretical spectrum.

The calculated spectra of the second and third lowest-energy isomer also shown in Fig. 7.6 can neither achieve a better agreement with the experimental data, so that a significant population of a higher-energy isomer is not readily invoked as reason for the discrepancy. Two possible reasons for the discomforting disagreement between measured and computed data could be either an insufficient description provided by the employed GGA-PBE functional or strong (cross)anharmonicities in the experimental data, which could both particularly affect the splitting of the higher-energy doublet in the D_{2d} spectrum. The latter is very sensitive to the distortions of the pyramidal ground-state structure caused by Jahn-Teller effects [113] and the Ar bonding, both of which might not be sufficiently treated at the level of a semi-local xc functional. Simultaneously, one has to keep in mind, however, that the unusually strong Ar bonding in the complex that will be further discussed in the next chapter severely enhances the multiple photon excitation aspect of FIR-MPD, and makes the comparison to computed linear IR absorption spectra possibly less justified.

7.7.2. Co_5^+

For Co_5^+ we determine as most stable isomers three slightly differently distorted versions of a trigonal bipyramid that are all within a 0.06 eV energy range and possess all a magnetic moment of $10\mu_B$. The energetically highest of the three has still C_{2v} symmetry, while the other two are further Jahn-Teller distorted to C_1 symmetry by tilts of the apex Co atom away from the symmetry axis of the slightly folded rhombohedral base. In all three structures there are sets of shorter Co bonds of around $\sim 2.2 - 2.3$ Å and more elongated bonds to the apex atom of the order of 2.5 Å. A more symmetric C_{4v} tetragonal pyramid with $12\mu_B$ is found at 0.15 eV above the ground state, followed by the lowest-lying planar structure already at 0.81 eV. This strong preference for a distorted three-dimensional structure in a high-spin state is again similar to corresponding findings for the neutral pentamer [113, 114], further confirming that neutral and cationic Co clusters exhibit similar geometries [115]. The comparison of the computed IR absorption spectra to the FIR-MPD data is this time slightly more favourable. Figure 7.7 shows the corresponding data for one of the degenerate distorted tetragonal pyramids (the spectra of the other two differ only insignificantly), as well as for the more symmetric tetragonal pyramid and the planar structure. The measured spectra from Co_5^+Ar and $Co_5^+Ar_5$ complexes show comparable differences as in the case of the tetramer complexes: A peak at the low-frequency end of the spectrum is only observed for the complex containing more Ar atoms and a group of bands in the range 175-230 cm^{-1} exhibits pronounced frequency shifts and intensity changes. From the overall shape one would expect at least three components behind this group and this indeed is what is found in this frequency range in the computed spectrum for the lowest-energy distorted pyramids. However, even though only containing two main peaks in this frequency range, the overall spectrum of the symmetric tetragonal pyramid isomer does also fit the experimental data rather well, so that we cannot safely discriminate between these two (in any case rather similar)

structures. What is instead quite clear from Fig. 7.7 is that the planar structure can not be reconciled with the measurements.

Due to the already severely distorted geometry of the bare ground state the influence of the rare-gas atom on the computed spectrum is less dramatic than in the case of the Co tetramer. Rather than leading to the appearance of a series of additional IR-active bands, its effect is more to shift existing peaks and change their intensities. With an again rather strong computed binding energy of around 0.3 eV this concerns primarily those modes that involve displacements in the direction of the Ar adsorbate. A particularly pronounced example of such a change in intensity is the computed mode at 260 cm^{-1} for the spectra from the Co_5^+Ar and $Co_5^+Ar_5$ complexes, in nice analogy to the variations observed in the corresponding experimental spectra.

7.7.3. Co_6^+

The two most stable determined isomers of Co_6^+ correspond to slightly distorted tetragonal bipyramids, with the lowest-energy structure exhibiting a magnetic moment of $15\mu_B$ and another one 0.10 eV higher with $13\mu_B$. Both structures possess D_{3d} symmetry and bond lengths around 2.3 Å. The third lowest-energy geometry is a capped trigonal bipyramid also with $13\mu_B$, but with 0.56 eV above the ground state this is already significantly higher in energy. With the DFT calculations for Co_6 by Datta *et al.* [114] also arriving at a tetragonal bipyramid ground state, the similarity between the structural motif of cationic and neutral clusters thus continues also for the hexamer.

The clear energetic preference for the tetragonal bipyramid motif is supported by the comparison of computed and experimental vibrational spectra. The FIR-MPD data is shown in Fig. 7.8 and exhibits this time only little differences between the measurements of Co_6^+Ar and $Co_6^+Ar_2$ complexes. In both cases the spectrum is simple and consists only of two peaks around 250 cm^{-1}. This pattern is fully reproduced in the calculated spectra for the two tetragonal bipyramids and certainly inconsistent with the computed spectrum for the capped trigonal bipyramid. In the case of the cationic hexamer theory therefore achieves an unambiguous assignment of the structural motif. Furthermore, the computed spectra for the identified ground-state structure either bare or with one or two Ar atoms shown in Fig. 7.8 are virtually identical. This is distinctly different to the findings for the smaller Co_n^+ clusters, and goes hand in hand with a significant reduction of the Ar bond strength, computed as 0.12 eV for the $15\mu_B$ ground-state tetragonal bipyramid.

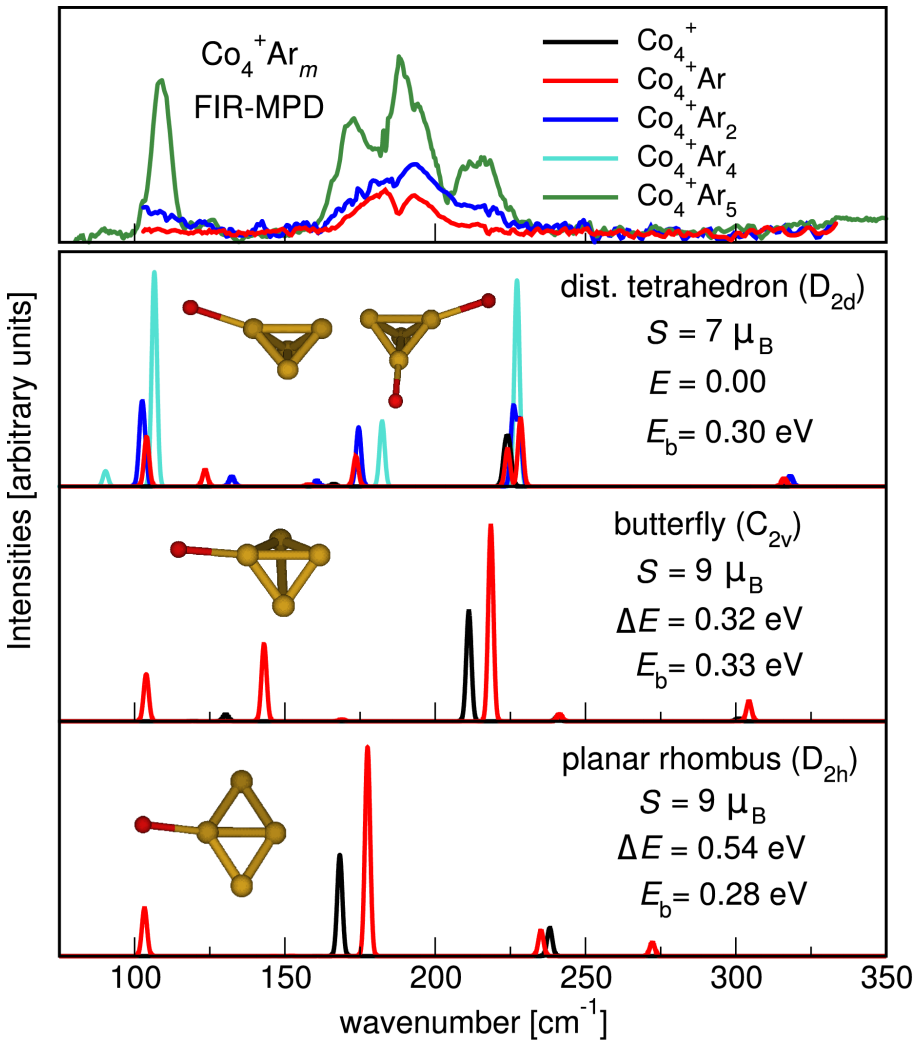

Figure 7.6.: Comparison of experimental FIR-MPD spectra (upper panel) with the computed IR absorption spectra of Ar complexes of the identified three energetically lowest isomers of Co_4^+: The most stable distorted tetrahedron (second panel from top), a butterfly-type tetrahedron (third panel from top), and a planar rhombus (bottom panel), see text. Each of the theoretical panels additionally displays a schematic representation of the cluster geometry also indicating the position of the adsorbed Ar atom in the Co_4^+Ar complex, the point group symmetry, the total energy difference ΔE of the bare cluster with respect to the most stable isomer, and the Ar binding energy E_b in the Co_4^+Ar complex. For the lowest-energy isomer, we additionally show a representation of the cluster geometry indicating the position of both adsorbed Ar atoms in the $Co_4^+Ar_2$ complex. In the $Co_4^+Ar_4$ complex each of the four Co atoms is directly coordinated to one Ar atom.

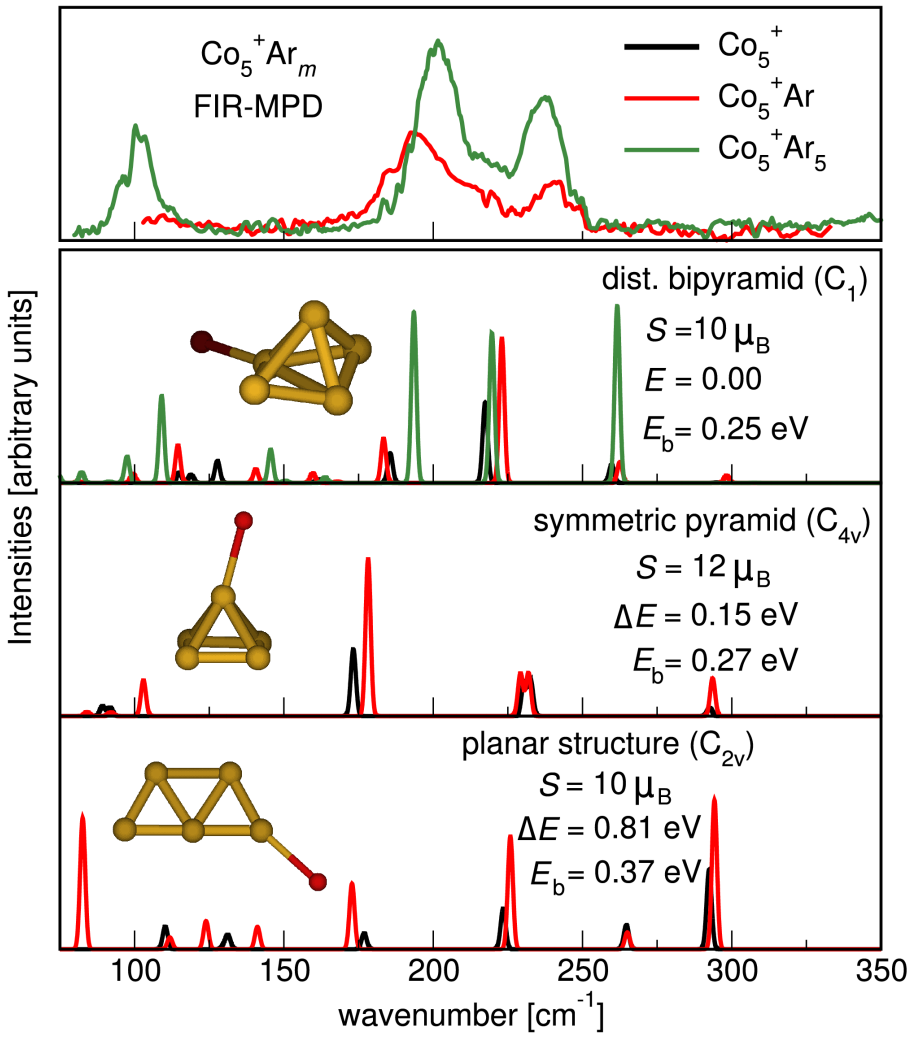

Figure 7.7.: Comparison of experimental FIR-MPD spectra (upper panel) with the computed IR absorption spectra of Ar complexes of identified energetically lowest isomers of Co_5^+: One of the three almost degenerate distorted tetragonal pyramids (second panel from top), a more symmetric C_{4v} pyramid (third panel from top), and a planar structure (bottom panel), see text. Each of the theoretical panels additionally displays a schematic representation of the cluster geometry also indicating the position of the adsorbed Ar atom in the Co_5^+-Ar complex, the point group symmetry, the total energy difference ΔE of the bare cluster with respect to the most stable isomer, and the Ar binding energy E_b in the Co_5^+Ar complex.

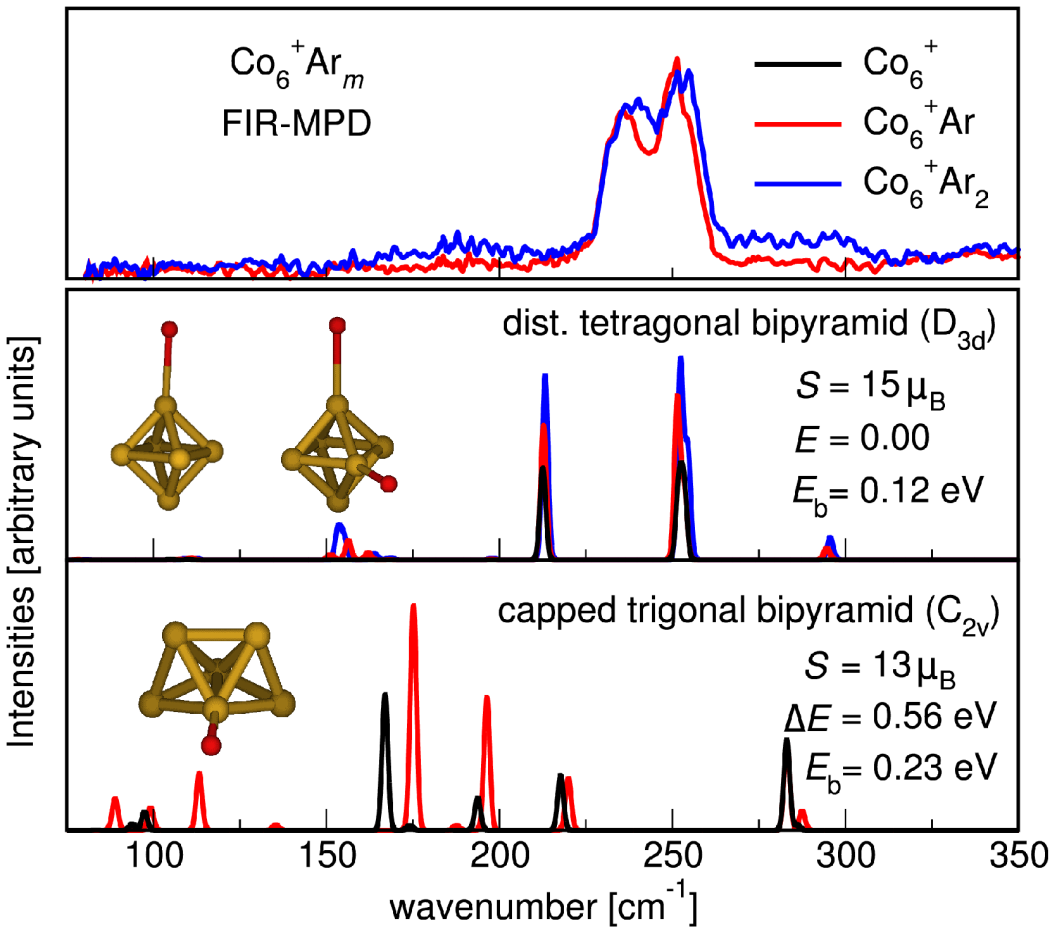

Figure 7.8.: Comparison of experimental FIR-MPD spectra (upper panel) with the computed IR absorption spectra of Ar complexes of identified energetically lowest isomers of Co_6^+: The tetragonal bipyramid with $15\mu_B$ (second panel from top) and the capped trigonal bipyramid (third panel from top). The spectrum for a second tetragonal bipyramid with $13\mu_B$ is very similar to the one of the $15\mu_B$ pyramid and omitted for brevity, see text. Each of the theoretical panels additionally displays a schematic representation of the cluster geometry also indicating the position of the adsorbed Ar atom in the Co_6^+Ar complex, the point group symmetry, the total energy difference ΔE of the bare cluster with respect to the most stable isomer, and the Ar binding energy E_b in the Co_6^+Ar complex.

7.7.4. Co_7^+

In the case of Co_7^+, a capped tetragonal bipyramid with C_{3v} symmetry and magnetic moment of $16\mu_B$ is identified as most stable structure. Again, the same structural motif for the cationic ground state is therefore found as in the preceding work for the neutral clusters [114]. The bond length in the triangle far side from the apex atom is 2.31 Å and is thus close to the computed bond length of 2.36 Å of the uncapped Co_6^+ ground-state structure, whereas the triangle near side of the apex atom has elongated bonds of 2.43 Å. The resulting average bond length of 2.34 Å for the entire cluster is thus slightly extended compared to the corresponding neutral isomer reported by Datta *et al.* [114] with an average bond length of 2.29 Å. Energetically only insignificantly higher than the capped tetragonal bipyramid are two spin-variants of a distorted pentagonal bipyramid with C_{2v} symmetry, namely one with $14\mu_B$ only 0.06 eV higher and one with $16\mu_B$ only 0.08 eV higher. The structural variety is further complemented by the energetically next lowest isomer, which corresponds to a capped trigonal prism also in a $16\mu_B$ high-spin state and 0.15 eV higher than the ground state. Still energetically quite close this is then followed by another, slightly differently distorted $16\mu_B$ C_{3v} tetragonal bipyramid (0.18 eV) and a low-symmetry $16\mu_B$ structure at 0.29 eV.

With this close energetic spacing it is particularly interesting to compare to the vibrational spectra from the FIR-MPD experiments and see if the obtained fingerprint pattern does maybe conform with one of the slightly higher-energy isomers. The corresponding measurements are shown in Fig. 7.9, exhibiting slightly larger differences between the data from the complexes with different number of Ar atoms than in the case of Co_6^+, yet still significantly less than in the case of the smaller clusters Co_4^+ and Co_5^+. The computed IR absorption spectrum for the ground-state tetragonal bipyramid agrees overall rather well with the experiments, reproducing all three higher-frequency bands. The observed influence of the Ar atom in the comparison to the bare cluster spectrum is also in line with the experimental trend, i.e. it is somewhat larger than in the case of Co_6^+, but less than in the case of Co_4^+ and Co_5^+. Defering the detailed discussion to the next chapter, particular emphasis is placed on the clear correlation with the now again slightly larger computed Ar binding energy of 0.18 eV.

A major concern in the comparison to the experimental data is the computed strong Ar-induced mode at the lower frequency end of the spectrum. For this, one cannot resort to the argument of apparent missing peaks in the spectra of complexes with one or two Ar atoms, as there are no complexes with more Ar atoms stabilized under the employed beam conditions, cf. Fig. 7.2. Comparing the theoretical spectra of the bare and Ar-complexed C_{3v} ground-state isomer it is, however, clear that the intensity of this particular mode depends sensitively on the interaction with the Ar ligand. The second most stable adsorption site at this isomer is atop one of the basal Co atoms that are on the opposite side from the capping atom, and its binding energy is only 60 meV lower. In the corresponding spectrum (see Appendix E) the mode does hardly show up. Most likely, the discrepancies between experimental and theoretical data with respect to this mode can be attributed to subtle shortcomings of the employed xc functional in describing the cluster-ligand interaction. Returning to the question of a unique structural

assignment, one would probably still conclude that the spectrum of the ground-state capped tetragonal bipyramid agrees overall best with the experimental data. Nevertheless, one also has to recognize that the fingerprint pattern of the different low-energy isomers is not sufficiently distinct to rule out small percentages of these isomers in the experimental beam as well.

7.7.5. Co_8^+

The most stable isomer found for Co_8^+ is a double-capped distorted tetragonal bipyramid with $17\mu_B$, which thus – like already the Co_7^+ ground state – also contains the very stable tetragonal bipyramid as a structural sub-unit and again coincides with the structural motif determined as most stable for the corresponding neutral cluster [114]. The geometry has nearly D_{2d} symmetry, with only bond distance differences of the order of ~ 0.01 Å reducing it to the C_s point group. The next lowest isomer found is only marginally higher in energy (0.09 eV) and corresponds to a slightly distorted double-capped trigonal prism with a magnetic moment of $17\mu_B$. This is followed by another double-capped tetragonal bipyramid of the same magnetic moment (0.21 eV), and then several differently distorted and energetically virtually degenerate capped pentagonal bipyramid structures at around 0.47 eV.

For this largest studied cluster the comparison of the computed and measured vibrational data is unfortunately again rather inconclusive with respect to a potential structural assignment. The experimental spectra are shown in Fig. 7.10 and exhibit primarily two main peaks with some suggested shoulders at the high-frequency end. There are only marginal differences between the measurements from the Co_8^+Ar and $Co_8^+Ar_2$ complexes, which is characteristic for all further FIR-MPD spectra recorded for larger clusters up to 18 Co atoms (not shown). This minor influence of the adsorbed rare-gas atom is similarly found in all computed spectra for the three most stable Co_8^+ isomers, concomitant with a moderate Ar binding energy of around 0.15 eV. In all cases the data obtained is furthermore more or less similarly consistent with the measured fingerprint pattern, with maybe a slightly better agreement for the ground-state isomer. Put more modestly there is at least no compelling evidence that it is not the latter structure that is predominantly observed, and thus there is neither a reason to question the energetic order determined at the GGA-PBE level nor to invoke the kinetic stabilization of a significant fraction of higher-energy isomers in the experiments. Part of the reason for the disappointing inconclusiveness are the rather strong Jahn-Teller deformations found for the high-spin isomers, which make it partly a matter of semantics to distinguish between e.g. a double-capped distorted trigonal prism or a double-capped distorted tetragonal bipyramid. Without clearly distinct symmetries the various geometries exhibit only quantitative differences in the IR spectra, which in turn severely limits a technique like FIR-MPD which relies on an indirect structure determination through characteristic fingerprint patterns.

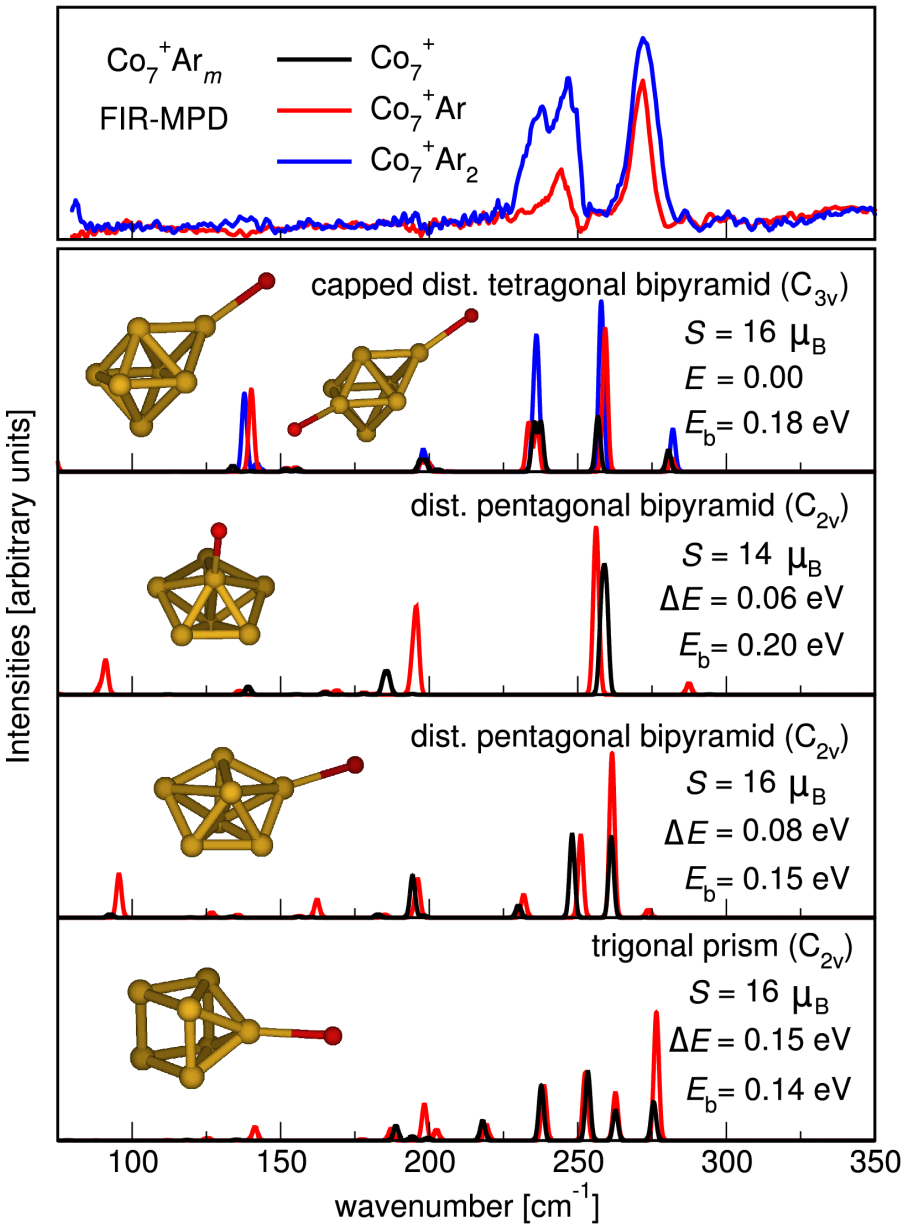

Figure 7.9.: Comparison of experimental FIR-MPD spectra (upper panel) with the computed IR absorption spectra of Ar complexes of identified energetically lowest isomers of Co_7^+: The capped tetragonal bipyramid with $16\mu_B$ (second panel from top), two spin-variants of a pentagonal bipyramid (third and fourth panel from top), as well as a capped trigonal prism (bottom panel), see text. Each of the theoretical panels additionally displays a schematic representation of the cluster geometry also indicating the position of the adsorbed Ar atom in the Co_7^+Ar complex, the point group symmetry, the total energy difference ΔE of the bare cluster with respect to the most stable isomer, and the Ar binding energy E_b in the Co_7^+Ar complex.

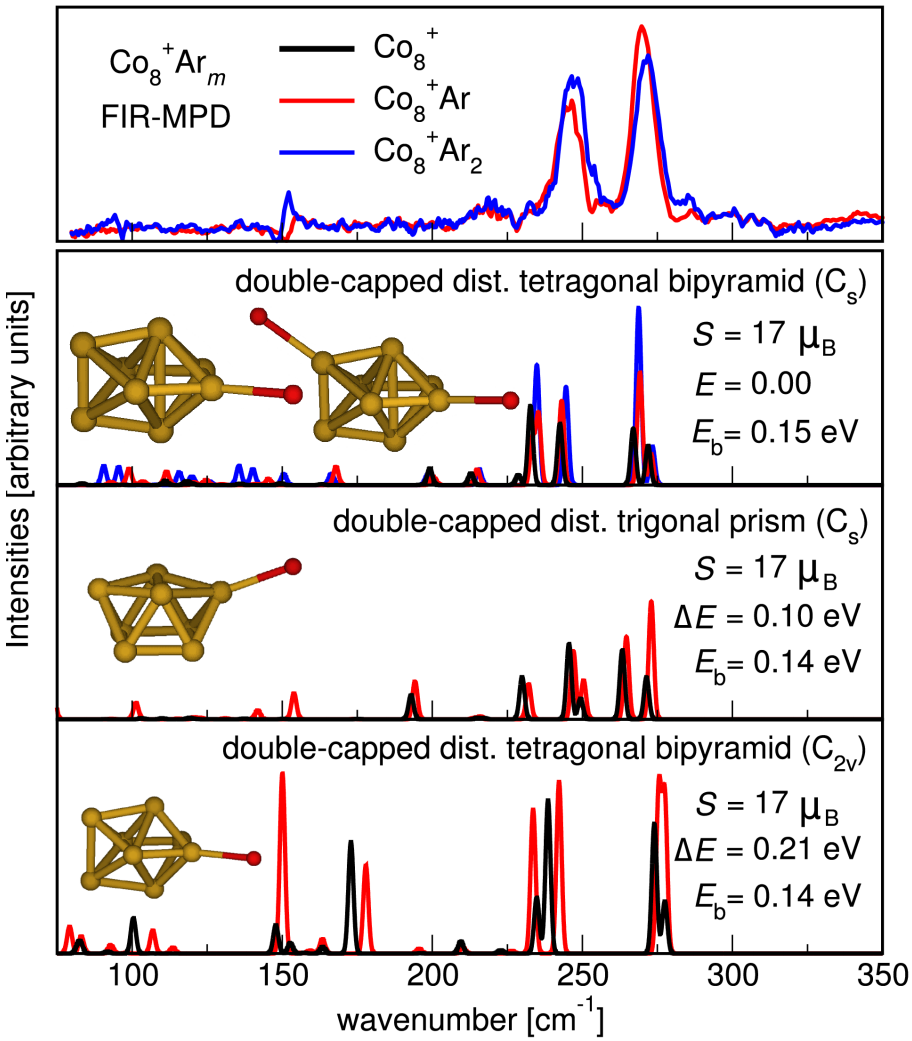

Figure 7.10.: Comparison of experimental FIR-MPD spectra (upper panel) with the computed IR absorption spectra of Ar complexes of identified energetically lowest isomers of Co_8^+: The double-capped distorted tetragonal bipyramid with $17\mu_B$ (second panel from top), the double-capped trigonal prism (third panel from top), as well as another double-capped tetragonal bipyramid (bottom panel), see text. Each of the theoretical panels additionally displays a schematic representation of the cluster geometry also indicating the position of the adsorbed Ar atom in the Co_8^+Ar complex, the point group symmetry, the total energy difference ΔE of the bare cluster with respect to the most stable isomer, and the Ar binding energy E_b in the Co_8^+Ar complex.

7.8. Summary

Summarizing this chapter, density-functional theory has been used to study energetically low-lying isomers of cationic Co clusters containing from four to eight atoms. Supplementing the search for structural motifs with first-principles basin-hopping sampling the identified ground-state structures agree for each cluster size with the geometry determined in preceding DFT studies for the corresponding neutrals. All structures are in high-spin states and exhibit sizable Jahn-Teller distortions. The specific motivation has been a detailed comparison of the vibrational properties of the clusters with experimental far-infrared multiple photon dissociation data. While this comparison allows only in few cases for a unique structural assignment and leaves some open questions and doubts in particular for Co_4^+, there is at least no compelling evidence that would question the energetic order of the isomers obtained at the GGA-PBE level or reciprocally indicate a kinetic stabilization of a significant fraction of higher-energy isomers in the experiments.

Reproduced quite well by the calculations is the general trend of the influence of the rare-gas probe atoms used in the measurements, with Ar-induced changes of the IR spectra becoming smaller for larger clusters. This correlates nicely with an overall reduction of the calculated Ar-Co_n^+ bond strength from 0.3 eV for the lowest lying isomers of Co_4^+ to around 0.1 eV for the largest studied clusters. This clear correlation therefore motivates a deeper analysis of the nature of the Ar-Co_n^+ bond, presented in the next chapter, with the expectation of understanding the behaviour in the previously reported studies on cationic vanadium, niobium and tantalum clusters [1, 2, 3, 4, 5].

For the investigated Co_n^+ vibrational spectra, complexation with Ar leads in general to the appearance of new IR-active modes, hand in hand with frequency shifts and splittings of existing peaks and often an increase in the absorption intensity. These effects can be primarily traced back to Ar-induced symmetry breakings in the geometric structure of the cluster. The latter become obviously stronger with increasing Ar binding energy, but can never be excluded. For a cluster geometry on the verge of a Jahn-Teller distortion even the tiniest disturbance by a rare-gas ligand may induce symmetry breakings that can then substantially affect the spectrum. In this respect, the present work casts some doubts on the hitherto employed practice of comparing experimental data to computed IR spectra of bare clusters. The influence of the probe atom can not be judged from a comparison of experimental spectra from different complexes, as the latter reveal only the differences between adsorption of one or more rare-gas atoms, and not with respect to the bare cluster. A computed negligible effect of Ar on the spectrum of one cluster isomer does neither justify to dismiss the ligand in the calculations, since this tells nothing about the liability toward symmetry breaking of other isomers of the same or other cluster sizes. In this situation, explicit consideration of the rare-gas atoms in the modelling is always advisable at little additional computational cost.

The appearance of Jahn-Teller distortions and therefore qualitatively similar vibrational fingerprints clearly limits FIR-MPD in connection with DFT as a method for structure determination. Since a unique assignment of a computed spectrum is impossible in these cases, it is recommended to additionally resort to further spectroscopic data, in the hope of gaining a more reliable picture of the cluster structure under consideration.

8. The Nature of the Ar-Co_n^+ Bonding

8.1. Introduction

While the comparison of computed and measured vibrational spectra in particular of the smallest studied Co clusters leaves a number of open questions and doubts with respect to the detailed interpretation of the FIR-MPD data, the general trend of the influence of the rare-gas atoms with cluster size is reproduced quite well by the calculations presented in the previous chapter. In general, the Ar-induced changes of the IR spectra become smaller for larger clusters, correlating nicely with an overall reduction of the Ar-Co_n^+ bond strength from an "unusually" high value of 0.3 eV for the lowest lying isomers of Co_4^+ to the more "intuitive" value of around 0.1 eV for the largest studied clusters. Particularly for the smallest clusters, the complexation with Ar leads to the appearance of new IR-active modes in the spectra, hand in hand with frequency shifts and splittings of existing peaks and often an increase in the absorption intensity. These effects can be primarily traced back to Ar-induced symmetry breakings in the geometric structure of the cluster, together with an enhanced polarization of the cluster electron density. Both effects depend obviously on the strength of the Ar-Co_n^+ interaction, rationalizing the correlation with the computed Ar binding energy. The explanation for the observation of a much more pronounced dependence of the FIR-MPD spectra on the number of Ar atoms compared to the preceding work on the cationic group 5 clusters [1, 2, 3, 4, 5] is therefore coupled to the understanding of the larger Ar binding energy at the smaller Co clusters compared to the e.g. ~ 0.1 eV computed also with GGA-PBE for the previously studied V_n^+ clusters [1, 2]. Hence, the motivation for this part of the work is to analyze the origin of the Ar-Co_n^+ binding in order to deduce an explanation of the different behaviour of the Ar probe atom in the previous studies.

8.2. Electrostatic Model of the Ar-Co_n^+ Bonding

8.2.1. Multipole Decomposition via Hirshfeld Analysis

As an important key to understand the particularities of the Ar-Co_n^+ binding energy, Fig. 8.1 shows a plot of the electron density redistribution when adsorbing an Ar atom at the ground-state isomer of Co_4^+. Observable is a strong polarization of the Ar atom, which is equivalently obtained in a corresponding analysis of Ar adsorption at all larger cluster sizes studied. This suggests as the predominant contribution to the Ar-Co_n^+ bond a mere polarization of the rare-gas atom in the electrostatic field of the cationic cluster. Figure 8.2 provides further insights into the charge density distribution of the bare Co_4^+ cluster

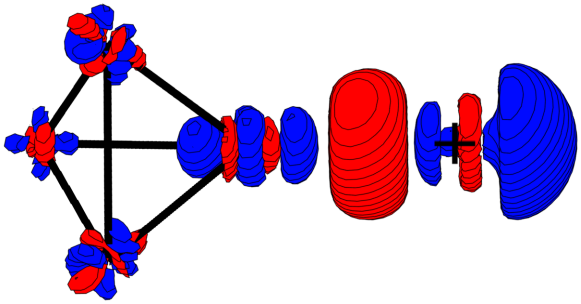

Figure 8.1.: Electron density difference induced by the adsorption of Ar at the ground-state isomer of Co_4^+, see text. For the optimized geometry of the Co_4^+Ar complex, the plot shows the isosurface corresponding to 0.015 $Å^{-3}$ electron accumulation (red) and depletion (blue) upon Ar addition, obtained by subtracting from the full self-consistent electron density of the Co_4^+Ar complex the self-consistent electron densities of the bare Co_4^+ cluster and an isolated Ar atom at the corresponding atomic positions. For better visualization, the drawn skeleton of black "bonds" at the left of the figure represents the cluster geometry, while the black cross marks the position of the Ar atom.

and clearly indicates that the positive charge seems to be smeared out over the surface thus leading to local dipoles pointing outward of the atomic cluster. Further support for this view on the nature of the Ar-Co_n^+ interaction comes from a multipole decomposition of the charge density of the determined ground-state isomers of the bare Co_n^+ clusters. Via Hirshfeld analysis [84] the point charge and dipole moment on each atom in the bare Co_n^+ cluster is computed, yet at the relaxed geometry of the corresponding Co_n^+Ar complex. The qualitatively expected results of having radially outward pointing local dipoles is herewith confirmed and depicted in Fig. 8.3, instancing both spin states of the ground-state structure of Co_6^+.

8.2.2. Ar Binding Energy in the Electrostatic Model

In a second step the Hartree-type interaction energy of an Ar atom is then computed, which is placed at the position in the Co_n^+-Ar complex into the frozen electrostatic field generated by the computed point charges and dipole moments of the bare cluster . The correspondingly obtained binding energies, which are henceforth termed $E_b(q$ +all $p)$, are finally compared to the fully self-consistently computed Ar binding energies at the different cluster sizes in the upper panel of Fig. 8.6. The rather good agreement suggests that the polarization of the Ar atom describes indeed the predominant character of the bond, with larger multipole moments and the back-reaction of the Co charge density on the polarized Ar atom contained in the true self-consistent calculation forming only a small correction. These findings hold similarly for all of the studied cluster sizes, and even carry over to less stable Ar adsorption sites and other low-energy isomers than the ground state (see Table 8.1), for all of which we obtain a similarly good agreement of the Ar

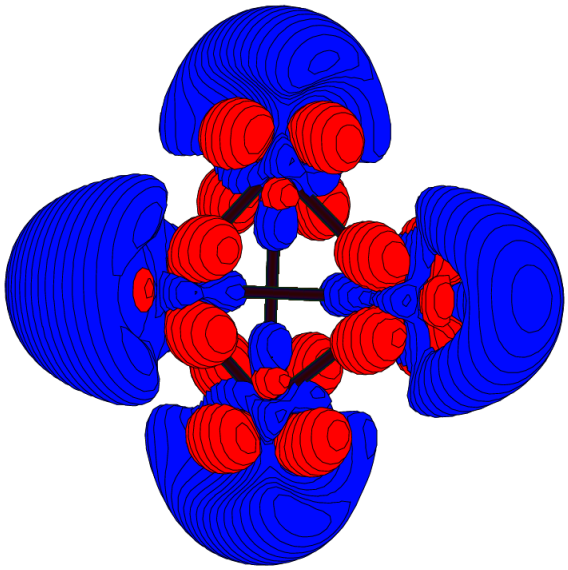

Figure 8.2.: Electron density difference between the ground-state isomer of Co_4^+ and the correspondong neutral isomer to illustrate the distribution of the positive charge. The plot shows the isosurface corresponding to $0.02\,\text{Å}^{-3}$ electron accumulation (red) and depletion (blue) obtained by subtracting from the full self-consistent electron density of the Co_4^+ the self-consistent electron densities of the neutral Co_4^+ cluster at exactly the same geometry. For better visualization, the drawn skeleton of black "bonds" represents the cluster geometry.

binding energy computed in the electrostatic model with the true self-consistent value. In order to better understand the tendency of the Ar binding energy, the individual energy contributions have further been disentangled. The binding energy that arises in the electrostatic model when the local monopoles are exclusively considered is termed as $E_\text{b}(q)$. Additionally considering the dipole to which the Ar atom is directly correlated (therefore denoted as the direct dipole p_d) yields a binding energy that is termed $E_\text{b}(q+p_\text{d})$. The contribution due to the direct dipole can then be defined as

$$E_\text{b}(p_\text{d}) \equiv E_\text{b}(q+p_\text{d}) - E_\text{b}(q)\ . \tag{8.1}$$

Correspondingly, the contribution arising from the other dipoles can conveniently be defined as

$$E_\text{b}(p_\text{other}) \equiv E_\text{b}(q+\text{all}\ p) - E_\text{b}(q+p_\text{d})\ . \tag{8.2}$$

Figure 8.6 clearly shows that taking the local monopoles exclusively into account yields only a low binding energy below 100 meV in the electrostatic model. Most of the binding energy instead results from the direct dipole field which is further offset by the other dipoles, obviously being unfavourably oriented. Both dipole contributions show

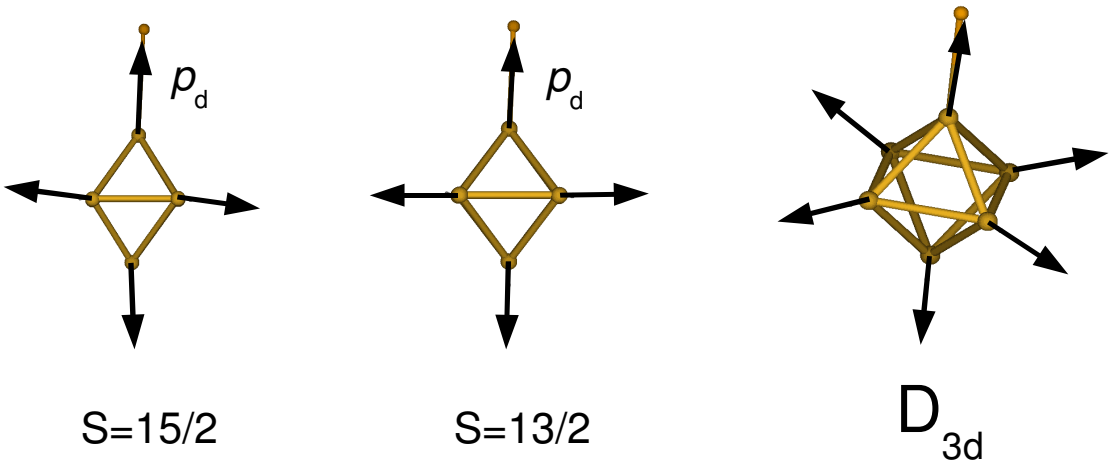

Figure 8.3.: Electrostatic model of the Ar-Co_n^+ bonding instancing both spin states of the D_{3d} structure of Co_6^+Ar. The local dipole moments obtained by a Hirshfeld analysis are pointing outwards. The two different spin states exhibit slightly different orientations of the dipoles. The dipole to which the Ar atom is directly coordinated is depicted as p_d.

the overall tendency to get smaller with increasing cluster size, with the variation of the binding energy predominantly resulting from the attractive direct dipole field. Figure 8.5 furthermore reveals the striking correlation between p_d and the resulting binding energy, once more underlining the picture of the direct dipole field playing the dominant role in the electrostatic interaction of the Ar probe atom with the Co cluster. With that picture in mind, the behaviour of the binding energy is easy to understand. The local dipole moments are obviously correlated with the cluster structure which discontinuously varies, thus giving rise to some kind of zig-zag trend, being typical for properties in the context of cluster physics as already mentioned in the introduction of the thesis. Superimposed is the general tendency towards a lower binding energy, which can be rationalized by the picture of the positive charge being smeared over an increasing cluster surface. The larger the cluster surface on which the single positive charge is distributed, the smaller the local charge density, thus giving rise to smaller local dipole moments. The clear correlation between the direct dipole and the overall binding energy again carries over to less stable adsorption sites and different spin states, as computed in Table 8.1.

8.2.3. Electrostatic Driving Forces

On the basis of the established electrostatic picture a rather consistent explanation for the observed trend of the Ar binding energy with Co cluster size can be achieved, as well as for the stronger interaction compared to the previously studied cationic clusters from group 5 of the Periodic Table. Taking the ground-state isomer of Co_4^+ and Co_8^+ as well as both spin states of the Co_6^+ ground-state structure as examples for stronger and weaker Ar bonding, respectively, the upper panel of Fig. 8.7 shows the variation of the interaction energy in the electrostatic model when changing the Ar-Co_n^+ bond length by

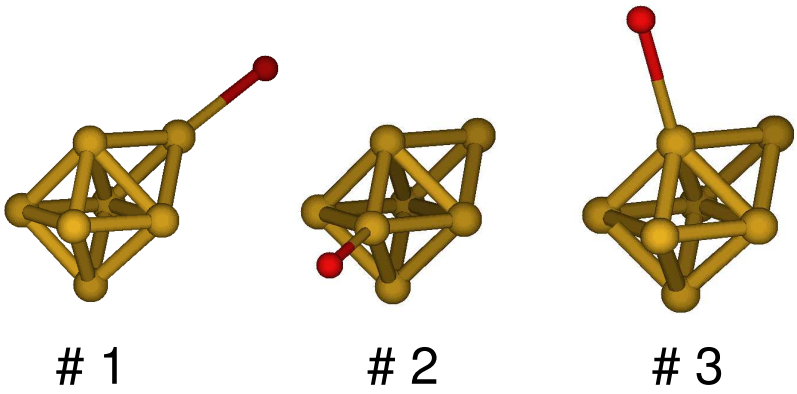

Figure 8.4.: The three symmetry-inequivalent on-top binding sites of Co_7^+Ar.

		p_d[eÅ]	E_b(PBE)	$E_b(q$ +all $p)$	$E_b(q)$	$E_b(p_d)$	$E_b(p_{other})$
Co_6^+Ar	S=15/2	0.46	0.12	0.10	0.04	0.13	-0.07
	S=13/2	0.56	0.30	0.27	0.07	0.37	-0.17
Co_7^+Ar	# 1	0.50	0.18	0.20	0.05	0.20	-0.05
	# 2	0.46	0.13	0.11	0.04	0.14	-0.07
	# 3	0.39	0.07	0.05	0.03	0.07	-0.05

Table 8.1.: Binding energies in the electrostatic model for different spin states and Ar binding sites. The different binding sites of Co_7^+Ar, labeled as # 1, # 2, and # 3, are illustrated in Fig. 8.4

moving the Ar atom radially away from the cluster. Immediately apparent is the smaller attraction in the case of Co_8^+ and the $15\mu_B$ state of Co_6^+ which can be rationalized with the weaker direct dipole moment compared to Co_4^+ and Co_6^+ being in the $13\mu_B$ state. The lower panels shows the contributions to the interaction energy due to the attractive field constituted by the monopoles and the direct dipole on the one hand (left lower panel) and the repulsive field due to the unfavourably oriented other dipoles (right lower panel) on the other hand. Perfectly in line with the observations of the previous section, the differences in the attraction predominantly arise from the differences in the direct dipole part, with the other dipoles only playing a minor role. Neglecting electronic self-consistency effects, the equilibrium Ar-Co_n^+ bond length is determined by the interplay of this attractive electrostatic contribution and a strongly repulsive short-range component arising from Pauli repulsion. Since the latter is primarily due to the inner shell electrons of the directly coordinated Co atom, one can assume this component to be rather similar for the four test cases to a first approximation. With the stronger electrostatic attraction this then leads to a shorter Ar-Co_4^+ and Ar-$Co_6^+(13\mu_B)$ bond length compared to Ar-Co_8^+ and Ar-$Co_6^+(15\mu_B)$, and in turn to a stronger Ar binding energy. In the same spirit one can also attribute the comparably stronger interaction of Ar with Co_n^+ than with the previously studied V_n^+, Nb_n^+ and Ta_n^+ clusters to this interplay of electrostatic attraction and Pauli repulsion. The distance at which the latter sets in depends of course

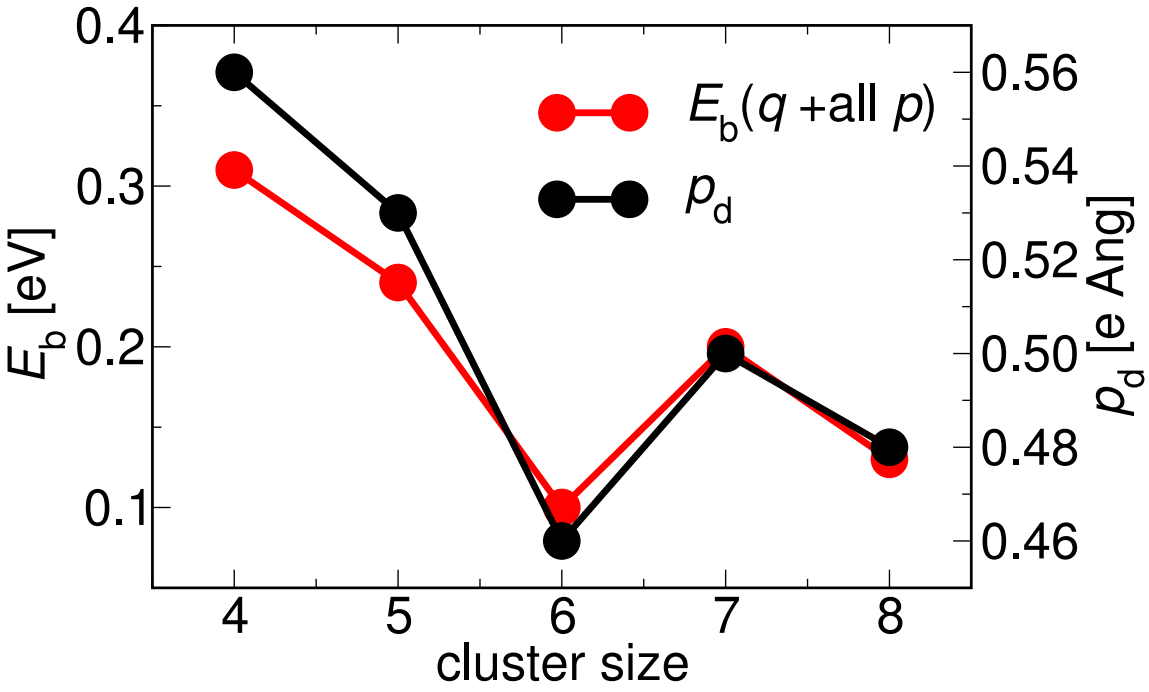

Figure 8.5.: Correlation between Ar binding energy and direct dipole moment p_d.

on the size of the atomic radius, which is smaller for Co than for any of the group 5 elements. With the steep Pauli repulsion setting in at already larger distances, one can in general expect longer equilibrium Ar-metal cluster bond lengths and therewith smaller Ar binding energies for the group 5 elements. As suggested by Fig. 8.7, the differences in the electrostatic driving forces is also smaller at such increased interaction distances, so that one would furthermore conclude on a smaller variation of the Ar binding energy with cluster size as in the case of Co.

Simplistically equating the Ar bond strength with the influence of the rare-gas atom on the measured FIR-MPD spectra as discussed above, the electrostatic picture can thus fully account for the general trends observed in the corresponding experiments on these materials. Tacitly assuming a similar character of the rare-gas−metal cluster bond, one would even predict that equally pronounced effects of the probe atoms as in the case of Co will be obtained whenever studying cationic clusters formed of other elements with small atomic radii or for more polarizable heavier rare-gas atoms.

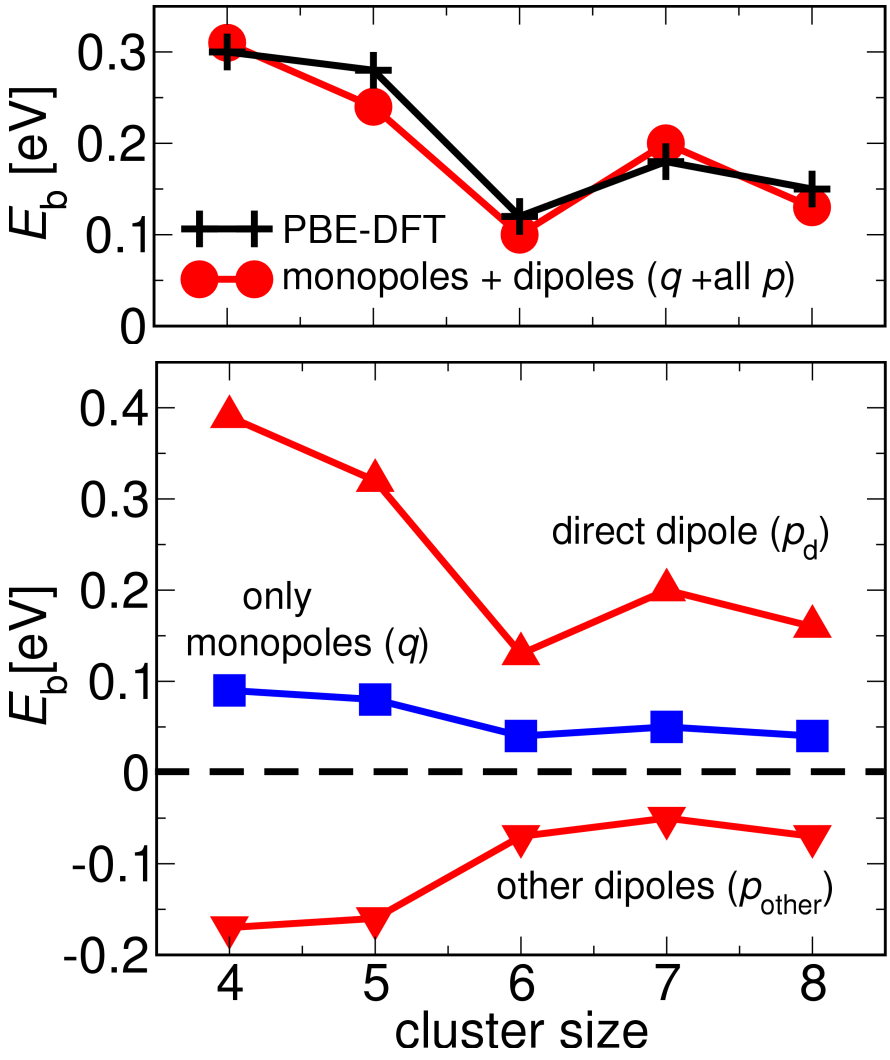

Figure 8.6.: Ar binding energy at the most stable adsorption site offered by the determined ground-state isomers of Co_4^+ to Co_8^+. In the upper panel, values obtained from the real self-consistent calculation (black crosses) are compared to the Hartree-type interaction energy obtained when placing Ar into the frozen electrostatic field formed by the point charges and dipole moments of the bare cluster (red circles), see text. The lower panel shows the separation of this interaction energy into contributions due to the monopoles (blue squares), the direct dipole (upper triangles) and the other dipoles (lower triangles).

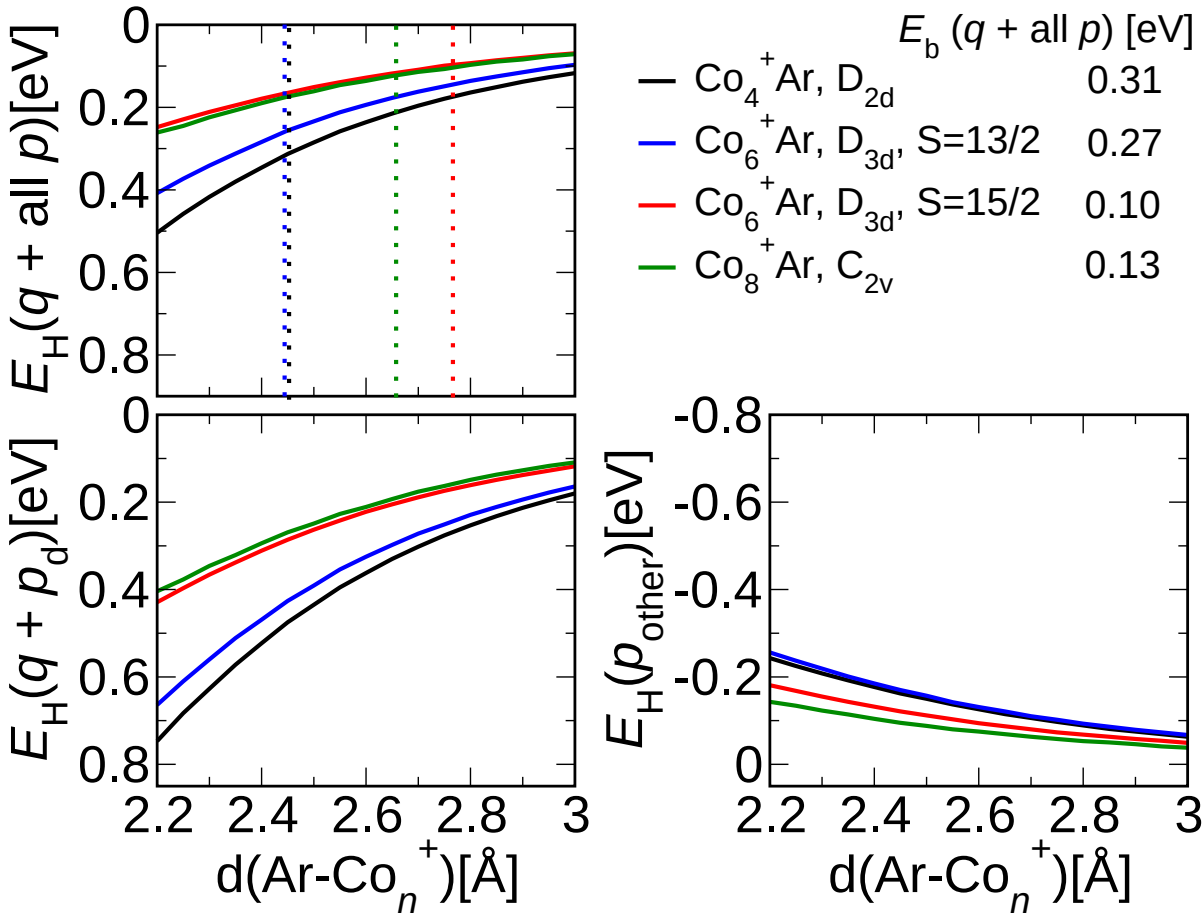

Figure 8.7.: Attractive Hartree-type Ar-Co_n^+ interaction energy computed in the electrostatic model, see text. Shown is the variation with bond length when moving the Ar atom radially outwards in the four complexes Co_4^+Ar (D_{2d}, black line), Co_8^+Ar (C_{2v}, green line), Co_6^+Ar (D_{3d}, S=15/2, red line) and Co_6^+Ar (D_{3d}, S=13/2, blue line). The four dotted vertical lines indicate the corresponding computed equilibrium bond lengths.

8.3. Summary and Conclusion

In this chapter, an electrostatic model was established that can capture most of the Ar-Co_n^+ binding energy, in particular it can fully account for the variation of the binding energy with the cluster size. For this, the Ar atom is self-consistently calculated in the multipole field set up by the local monopoles and dipoles in the bare Co_n^+ cluster obtained via Hirshfeld analysis. Disentangling the individual energy contributions reveals that the binding energy predominantly results from the dipole moment to which the Ar atom is directly coordinated, being offset by the unfavourably oriented dipole fields originating from the other Co atoms. Supported is this picture by the striking correlation between the full binding energy in the electrostatic model with the direct dipole strength, the latter being highly sensitive to the cluster structure. Superimposed to the thus resulting

zig-zag behaviour is the general trend towards lower binding energies with increasing cluster sizes since the larger surface area on which the positive charge is distributed yields smaller local dipoles. In the real cluster, the bond distance results from the interplay between the attractive Hartree energy and the Pauli repulsion, where the latter acts locally between the Ar atom and the Co atom to which the former is directly coordinated, thus being more or less constant for all isomers under consideration. The variation in the bond distances and the corresponding binding energies can therefore be traced back to the variation in the electrostatic driving forces resulting from the variation of the direct dipole strength. The interatomic distance at which the Pauli repulsion sets in is clearly controlled by the atomic radii of the constituent atoms. Hence, the larger binding energy for Ar-Co_n^+ can be attributed to the smaller radius of Co compared to V, Nb and Ta in the previous studies. As for the binding site, the preference for on-top sites becomes trivial to understand since this maximizes the polarization of the charge density in the radially outwards pointing direct dipole and in turn the resulting binding energy. As further conclusion, a stronger influence on the IR spectra is to be expected for heavier, more polarizable probe atoms.

Part IV.

Assessing the Efficiency of First-Principles Basin-Hopping Sampling

9. Method Optimization Based upon Quantitative Efficiency Criteria

9.1. Introduction

In the previous study with systems containing only few atoms (Co_4^+Ar to Co_8^+Ar), the sampling efficiency is not critical, reasonable settings for sampling parameters provided. In the context of the applied basin-hopping scheme, this means large enough move distances, enabling the system to find new isomers. The average move distance for displacing single randomly chosen atoms has therefore been empirically set to 1.5 times the corresponding dimer bond length for the whole sampling run, which worked satisfactorily. With increasing cluster size, however, the exponential growth of the number of different isomers clearly requires a more careful analysis of the involved technical parameters. Taking the well studied Lennard-Jones clusters as an example, the number of different isomers grows from 64 to 1505, when going from ten to only thirteen atoms [6]. To reveal inefficient settings, such an analysis necessitates a proper definition of efficiency, with which runs under different settings can quantitatively be compared. Such a quantity must obviously be coupled to the desired sampling goal. A common scheme is to investigate the mean first encounter, i.e. one counts the number of moves, or local structural relaxations, respectively, until the global minimum has been identified for the first time (see chapter 3). If it were exclusively the ground state being of interest, such a quantity would clearly be correlated with the overall computational time, thus being an appropriate efficiency indicator. However, as the previous study on $Co_n^+Ar_m$ has illustrated, also energetically higher-lying isomers are to be identified and discussed. Hence, the move number should be somehow coupled to the relevant isomers, where one has to consider the actual problem to decide which isomers are relevant. Moreover, rather than revealing inefficient settings only *a posteriori*, this optimization would best be carried out by monitoring on-the-fly analyzable performance indicators that allow to adapt an ongoing run. Unfortunately, there are few to none general prescriptions of how to set technical move parameters that do not require detailed system-specific insight. With respect to on-the-fly performance indicators there exists at best the empirical rule-of-thumb to aim at an overall acceptance of new trial structures of roughly one half [45, 116]. However, this rule emerges from the empirical observation that a factor one half ensures an efficient sampling of canonic ensemble averages and must not necessarily carry over to the intended goal of searching for the energetically lowest lying isomers with the least possible number of energy and force evaluations. A second complication arises from the stochastic nature of the BH algorithm. Any analysis mea-

suring the efficiency of technical BH settings or the reliability of suggested on-the-fly performance indicators therefore necessarily needs to involve an averaging over a sufficiently large number of different BH runs starting from different initial structures and using different random number seeds. This would not be too much of a problem when using numerically undemanding model potentials, but then it would be unclear whether the obtained findings are meaningful for proper quantum-mechanical PESs. A straightforward evaluation based on first-principles energetics, on the other hand, is hitherto computationally prohibitive even when only considering smaller clusters up to say 10 atoms. In this situation, the aim of the present study is to establish a corresponding framework for a systematic performance analysis of first-principles BH sampling runs. An important ingredient herein is the use of a hopping matrix type concept that significantly brings down the computational cost for the manifold of first-principles BH runs required in the averaging procedure and thus allows to arrive at statistically reliable performance values. Using DFT, this scheme is illustrated for Si clusters as a system with more directional, covalent type of bonding and for Cu clusters as representative of a metallic system. With the obtained insights, a simple adaptive scheme based upon on-the-fly efficiency indicators is finally suggested and illustrated.

9.2. Computational Details

9.2.1. Density-Functional Theory

The underlying PES is again obtained by DFT in the generalized-gradient approximation (GGA-PBE) [70] in order to transfer the gained insights to practical applications that require a quantum-mechanical treatment of the energetics. In order to suppress a potential complication in the performance analysis due to the spin degrees of freedom all calculations were consistently carried out in a non-polarized way. All calculations reported here were conducted with the so-called "minimal+*spd*" basis set (see Appendix A). For each considered system we recomputed all stable cluster isomers within an energy range up to 1 eV above the ground state, namely those listed in Figs. 9.5-9.7 below, also with hierarchically constructed larger basis sets available in FHI-aims. These calculations clearly show that the relative energies between these isomers are converged to within 10 meV at the "minimal+*spd*" basis set level, which is fully sufficient for the arguments and conclusions put forward below since it is the sampling efficiency that is in the focus of interest, not tightly converged physical quantities. To be on the safe side, several further basin-hopping runs with larger basis sets have been run, but no isomers other than those already revealed at the "minimal+*spd*" level have been obtained. This suggests that not only the local minima, but also the other parts of the PES are sufficiently described with the employed "minimal+*spd*" basis, in order to pursue the aspired efficiency study. The integration grids are found to be converged at $N_{\mathrm{ang,max}} = 302$ with $N_{\mathrm{r,div}} = 2$. The cut-off radius is slightly reduced to 4 Å which is again fully sufficient for the energy differences and thus saves computational time for the sampling. As l_{max} a conservative value of 6 was chosen. Furthermore, a slight Gaussian smearing of width 0.1

eV was employed to ensure scf-convergence for the distorted structures resulting from the trial move. Details of the convergence tests can be found in Appendix C.

Local structural optimization is done using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [27], relaxing all force components to smaller than $10^{-2}\,\mathrm{eV/Å}$. In order to check whether the thus identified different isomers are true local minima and not saddle points, they were subjected after the BH run to a vibrational analysis based upon a Hessian matrix obtained by finite differences of the analytical atomic forces when displacing all atoms by 10^{-3} Å.

9.2.2. Basin-Hopping

In this section, the intricate technical details of the basin-hopping scheme that influence the efficiency of the sampling, are further analyzed. The key ingredients are the way new structures are generated, the trial move scheme, and the acceptance criterium, which decides, whether a new structure is accepted or not. While in the previously presented study on $Co_n^+Ar_m$-complexes these parameters have been adjusted intuitively, they are now subject to a critical discussion.

Sampling Parameter

As a representative and widely used move class this work focuses on single-particle and collective moves, in which either a single randomly chosen atom or all atoms in the cluster are randomly displaced, respectively. The corresponding displacement vector of atom α is suitably expressed in spherical coordinates as

$$\Delta\mathbf{R}_\alpha = r\vec{e}_r(\theta,\phi) \quad , \tag{9.1}$$

where $\vec{e}_r(\theta,\phi)$ is a unit vector in the displacement direction defined by the angles θ and ϕ with respect to an arbitrary, but fixed axis. In order to obtain an unbiased sampling, $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$ must obviously be obtained as uniformly distributed random numbers. On the contrary, the move distance r is *a priori* not specified, but will sensitively determine the jumps in configuration space and therewith the algorithmic performance. It provides therefore a nice example of a technical parameter that one would like to optimize for a first-principles sampling run, yet without introducing bias or system-specific insight. As any one specifically chosen value for r (suitably measured in units of the dimer bond length a) could potentially already be too restrictive and therewith lead to a biased result, it appears advantageous to use a distribution of r values around a specified average value r_0. As two natural distributions, both a normal distribution (width $0.07\sqrt{r_0}$) around r_0 and a uniform distribution (width r_0) around r_0 are therefore investigated.

Apart from the move class the second important ingredient that needs to be specified in a BH run is the acceptance criterion according to which a generated trial structure is accepted and replaces the current cluster structure as starting point for the following trial move. In order to introduce a downhill driving force towards the energetically low-lying (and ultimately ground-state) isomers it is clear that a more stable trial structure should

always be unconditionally accepted. In its classical form, the BH scheme also accepts less stable trial structures according to a Metropolis rule, $\exp(-\Delta E/k_{\mathrm{B}}T_{\mathrm{eff}})$, where k_{B} is the Boltzmann constant and introducing another unspecified technical parameter which crucially affects the algorithmic performance, the effective temperature T_{eff}. The original motivation behind this Metropolis rule is that the finite possibility to climb uphill enables the algorithm to effectively surmount high-energy barrier regions on multiple-funnel type transformed PESs $\tilde{E}(\{\mathbf{R}_\alpha\})$. However, as long as the employed move class enables efficient jumps between all parts of configuration space, this acceptance criterion is only of subordinate importance. As it will become clear below this is still the case for the small cluster sizes studied here, and therefore in this work simply all generated cluster structures within a predefined energy range of interest are accepted. This allows to separate the analysis of the performance variation with the move distance from the specifically employed acceptance criterion.

In order to prevent an entropy-driven dissociation of the cluster during the BH run, trial moves as well as local relaxations are discarded that generate loosely connected or partly dissociated structures characterized by an atom having a nearest-neighbour distance larger than twice the dimer bond length. Similarly discarded were moves that place atoms at distances of less than 0.5 Å from each other.

Structure Distinction

For a performance analysis, it is of particular importance that the algorithm achieves to uniquely distinguish structures. In case of a simple analytical model potential, the total energy can be converged up to any desired accuracy, so that two isomers can be distinguished from each other exclusively by their different energetics. Using *ab initio* methods, however, the energy is noisy, thus prohibiting the structure distinction exclusively based upon the energy to uniquely distinguish different isomers from each other, in particular when Jahn-Teller distortions show up that are close in energy or in case of accidental degeneracies in energy space. Hence, the difference norm of all interatomic distances in the cluster is used as additional tool for the comparison of isomer structures. Two isomers a and b are hereby considered to be equivalent if

$$\frac{\sum_i \left(d_{a,\{i\}} - d_{b,\{i\}}\right)^2}{\sum_i \left(d_{a,\{i\}}^2 + d_{b,\{i\}}^2\right)} < \Delta \,, \tag{9.2}$$

where $d_{a,\{i\}}$ and $d_{b,\{i\}}$ are the sorted interatomic distances of the two isomers to compare. The denominator serves as normalization which yields a dimensionless quantity that is furthermore species- and cluster-size independent. Δ can be tuned such that all isomers are unambiguously distinguished and was taken as 10^{-4} as is illustrated in Fig. 9.1. Shown are the number of different isomers that are considered to be different with respect to the tolerance Δ. Additionally, the maximum energy interval ΔE associated with an isomer is given, which is the range of the total energies of all structures that are found to be identical according to the difference norm. A large value in the range of eV hereby indicates that different isomers could obviously be not distinguished. Hence, a small

value of ΔE is a necessary criterium for a sufficiently small tolerance value Δ. In case of a too large threshold, isomers are considered to be identical, though being different, and the number of different structures that are identified is too small and the energy range is way above the typical noise of DFT calculations in the meV range. Taking an extremely large value of 10^{-1}, for instance, structures cannot be distinguished at all and thus there is only one isomer yielded by the structure distinction scheme. Reducing Δ consequently increases the number of isomers until it saturates at 10^{-4} when all different structures can uniquely be distinguished from each other. This perfectly coincides with a drop of the maximum energy range down to few meV, indicating that only energy noise due to the finite self-consistency criterium and force convergence criterium of the DFT calculations is left. Decreasing the threshold further leads again to a slight increase of the number of isomers. This can be rationalized by a second effect. In this case, the noise in the bond distances is too large and two structures are considered to be different, though corresponding to the same isomer. This picture is supported by the fact that the new structures that pop up must be energetically close to the already identified structures at a larger tolerance value since the maximum energy range ΔE is already at the order of meV. As Fig. 9.1 clearly shows, 10^{-4} is a reasonable value for all three cases. Hence, Eq. (9.2) indeed leads to a size- and species-independent norm for the structural difference between two isomers.

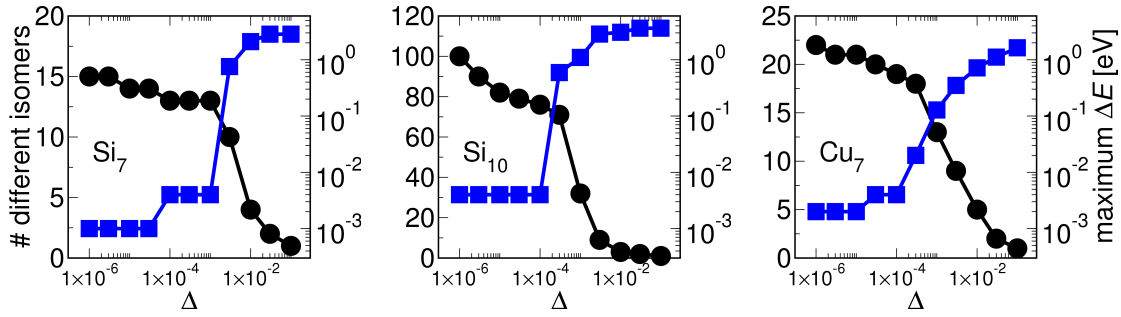

Figure 9.1.: Adjusting the difference tolerance Δ according to Eq. (9.2). Shown are the three cases Si_7, Si_{10} and Cu_7 discussed in this work. The sets of isomers result from sampling runs, using single-particle moves and a normal distribution of the move distance around 1.5 times the corresponding dimer bond length.

9.2.3. Measuring Sampling Efficiency

The intended performance analysis requires a well-defined measure for the success of a sampling run. A common choice for this in the literature is the number of moves until the global minimum has been found for the first time. Here, we adapt this criterion to the stated goal of identifying not only the global minimum, but also all relevant energetically lowest-lying isomers. Correspondingly, the considered indicator of sampling efficiency which we aim to optimize is the number of moves N until all relevant isomers have been found at least once, where, of course, one needs to define what a relevant isomer

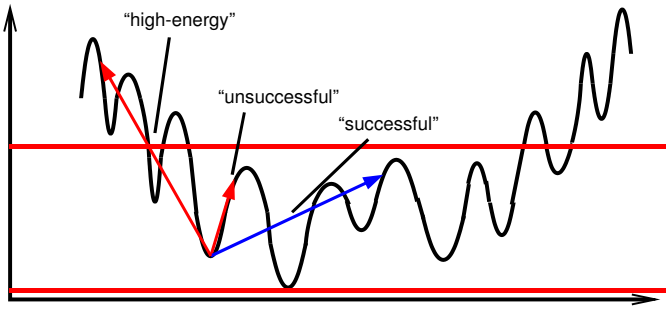

Figure 9.2.: Schematic illustration of successful, unsuccessful and high-energy trial moves in the BH scheme. The horizontal red lines indicate an energy window as acceptance criterium according to which moves within the target window are accepted.

is. While certainly a useful measure for the performance of the employed BH moves, it should still be stressed that due to the slightly varying number of geometry steps for the local relaxation of each trial structure, N is only roughly proportional to the total computational cost of the first-principles basin-hopping run.

Due to the stochastic nature of the BH method, both with respect to the generation of the initial starting structure and the generation of trial structures, N is only a statistically meaningful quantity after averaging over sufficiently many runs. Even for the small cluster sizes considered here, this implies having to run hundreds of different first-principles BH runs to obtain an N_{av} that is converged to within ± 1, and this for each BH setting (e.g. move distance) one wants to analyze. Since this straightforward approach would quickly lead to an unfeasible computational cost, we instead resort to the concept of a "hopping matrix" h_{ij}, which summarizes the transition probabilities between all isomers under the chosen BH settings. Specifically, the matrix element h_{ij} is then the probability to jump from the local minimum i to the local minimum j. If all local minima are explicitly accounted for, one obviously has the condition

$$\sum_j h_{ij} = 1 \quad . \tag{9.3}$$

Assuming that the matrix h_{ij} is completely known, a sufficiently large number of sampling runs starting in random isomers can be quickly simulated entirely on the basis of these transition probabilities without the need for further first-principles calculations. Notwithstanding, with an exponentially growing number of isomers with cluster size this approach merely shifts the computationally unfeasible number of direct BH runs to the equally unfeasible computation of an exploding number of hopping matrix elements, i.e. converged transition probabilities. Yet, below we will show that an approximate, but for our purposes sufficient determination of N_{av} is possible by restricting the explicit calculations to a limited number of hopping matrix elements.

Even then, N_{av} is only a quantity that reflects the efficiency of the sampling run *a posteriori*. As potential on-the-fly indicators of the sort as discussed in the introduction, we therefore disentangle the different possible outcomes of a trial move. First of all, the

system might relax back into the structure from which the trial move has been performed so that in terms of isomer information nothing has been gained. Correspondingly, we denote such a move as unsuccessful (see Fig. 9.2) and define at the nth trial move during the run the fraction of hitherto unsuccessful moves $\gamma_{\text{unsucc.}}$ as

$$\gamma_{\text{unsucc.}} = \frac{n_{\text{unsucc.}}}{n} \quad , \tag{9.4}$$

where $n_{\text{unsucc.}} < n$ is the number of unsuccessful moves up to that point. Even if the trial move leads to a different local minimum, the move might still be rejected due to the acceptance criterion, if the new minimum is higher up in energy. The fraction of moves rejected on this basis is defined as

$$\gamma_{\text{high}E} = \frac{n_{\text{high}E}}{n} \quad , \tag{9.5}$$

where $n_{\text{high}E} < n$ is the corresponding number of rejected moves up to that point. Only the remaining fraction

$$\gamma_{\text{succ.}} = 1 - \gamma_{\text{unsucc.}} - \gamma_{\text{high}E} \tag{9.6}$$

are successful moves at least in the sense that they bring the algorithm to a new minimum out of which the next trial move is performed, albeit not necessarily leading to a minimum that had hitherto not yet been sampled. Just as in the case of N_{av}, it only makes sense to analyze the fractions $\gamma_{\text{unsucc.,av}}$, $\gamma_{\text{high}E\text{,av}}$ and $\gamma_{\text{succ.,av}}$ once averaged over sufficiently many different BH runs, and for the small cluster sizes we will also evaluate only the fractions for the entire run until all relevant isomers are found. Yet, it is obvious from the construction that one can also evaluate these fractions on-the-fly at any trial move and furthermore possibly also only considering the outcomes of a certain number of the most recently performed trial moves, and we will further discuss this possibility in section 9.4.

9.3. Performance Analysis For Small Cluster Sizes

Our performance analysis concentrates on small clusters formed of Si and small clusters formed of Cu atoms. Both systems have already been subject to extensive theoretical studies and are therefore natural choices for the intended benchmarking. Extensive works on small silicon clusters have been done with Hartree-Fock methods and beyond [117, 118]. Further calculations based upon DFT are e.g. [119, 120]. A database of small silicon isomers based upon the PBE functional and an exhaustive sampling based upon minima hopping can be found in the PhD-thesis by Hellmann [121]. Recent works on small copper clusters based upon *ab initio* methods are e.g. [122, 123, 124, 125, 126]. The choice of these two materials is further motivated by their different chemistry, which can be characterized as more covalent and directional in the case of Si, and more metallic in the case of Cu. The direct comparison of the results obtained for Si_7 and Cu_7 will therefore reflect a possible material-specificity of the findings, while an additional

comparison of the results obtained for Si_7 and Si_{10} aims at assessing the variation with cluster size in the range where due to the limited dimensionality of the configuration space the BH acceptance criterion does not yet play a role (*vide infra*).

9.3.1. Existence of Dominant Isomers

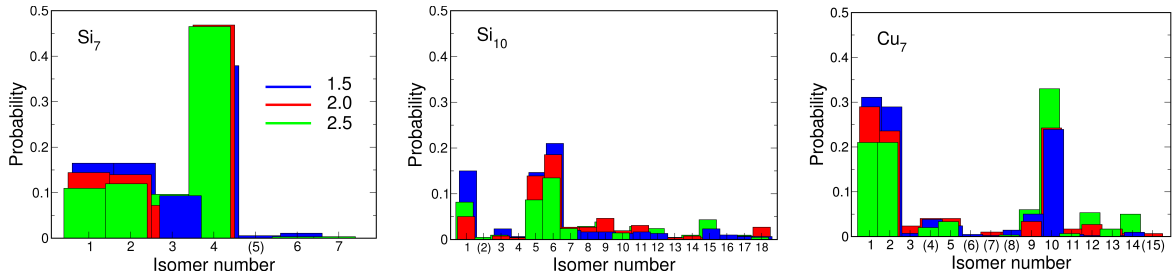

Figure 9.3.: Histograms of the probability with which trial moves end up in the lowest-energy isomers of Si_7, Si_{10} and Cu_7. The identified isomers are numbered with decreasing stability, with isomer #1 corresponding to the ground state and those isomers shown with bracketed numbers in fact revealed as unstable by an *a posteriori* vibrational analysis (see text). The histograms comprise all isomers found in an energy range up to 2 eV above the ground-state isomer, as obtained from long BH hopping runs using single-particle moves and normally distributed move distances around the average values $r_0 = 1.5$ a, 2.0 a and 2.5 a (with a the computed dimer bond length). The geometric structures behind the truly stable isomers in an energy range up to 1 eV above the identified ground state are summarized in Figs. 9.5-9.7.

As a prelude to the actual performance analysis we present in Fig. 9.3 the histograms of the number of times with which low-energy isomers were identified in long BH runs for the three systems addressed, i.e. Si_7, Si_{10}, and Cu_7. Each run consisted of several hundred moves and was carried out until the shape of the histogram, i.e. the normalized probability with which the different low-energy isomers are identified, was fully converged. In all cases the evolution towards convergence was rather uniform as demonstrated by Fig. 9.4 for Si_7, which presents the histogram entries binned over consecutive sampling periods containing 50 moves each. Apparently, the ratios of the histogram entries for each sampling period are roughly the same. In view of the overall rather limited system dimensionality and concomitant small number of low-energy isomers, this is most naturally interpreted in terms of the initially made assumption that the employed move types enable jumps between any parts of the PES, thereby suggesting that for these systems the algorithm efficiency is dominated by the actual moves and not the acceptance criterion.

Even though Fig. 9.3 comprises the data obtained using single-particle moves with three quite different move distances it is interesting to observe that some isomers are always sampled much more often than others. For Si_7 for example, more than one third of

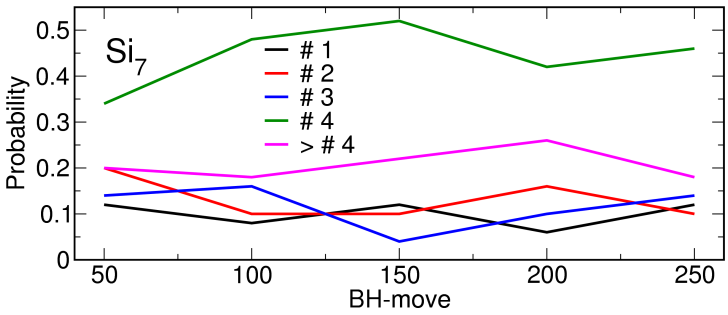

Figure 9.4.: Normalized histogram entries for the lowest-energy isomers of Si_7 as in Fig. 9.3. Shown is the evolution when binning the histogram entries over consecutive sampling periods containing 50 moves each using single-particle moves and normally distributed move distances around the average values $r_o = 2.5$. Entries for all isomers higher in energy than isomer #4 are bundled into one entry labeled $>\#4$.

all executed moves in the BH runs ended up in the isomer structure labeled #4, regardless of the actual move distance employed. In the case of collective moves, the corresponding histograms look qualitatively the same so that the existence of such "preferred" isomers, which we will henceforth term dominant isomers, is even independent of the specific move class employed. In this respect, one should mention that some of the isomers listed in Fig. 9.3 turned out to be unstable in the concluding vibrational analysis and are correspondingly not further considered below. Distinguishing and discarding these structures which correspond either to flat or saddle-point PES regions directly in the BH run is unfortunately impossible as it would imply a prohibitive computational cost when performing a vibrational analysis directly after each trial move. As apparent from Fig. 9.3 the total number of times in which the BH runs end up in such unstable structures is at least not too large, so that the actual computational time wasted is small. The one notable exception is isomer #4 of Cu_7, which exhibits small imaginary eigenmodes, but is sampled about as frequently as the truly stable isomer #5. Since the algorithm thus spends some appreciable time in this basin, we retained isomer #4 in the ensuing performance analysis despite its instability.

One immediate rationalization for the existence of dominant isomers is simply that their corresponding basins of attraction on the PES is huge and thus hit by the trial moves many times. Inspection of the geometric structures of the lowest-energy isomers as summarized in Figs. 9.5 - 9.7 for the three systems points, however, at a second potential reason. Many of the dominant isomers correspond to rather low-symmetry structures, e.g. isomer #4 for Si_7, isomer #6 for Si_{10} or isomer #10 for Cu_7. More clearly stated, Si_7 can e.g. be regarded as composed of a Si_6 subunit with one capped facet. In the case of its low-symmetric isomer #4 this capping can be done in four different ways whereas in the case of the high-symmetry C_{3v}-structure #2 there are only two possibilities. Quantitatively, this can be expressed in terms of the number of versions of an isomer that correspond to the multiplicity of the basins of attractions leading to this isomer, which is given by Eq. (2.24). Table 9.1 summarizes the number of versions

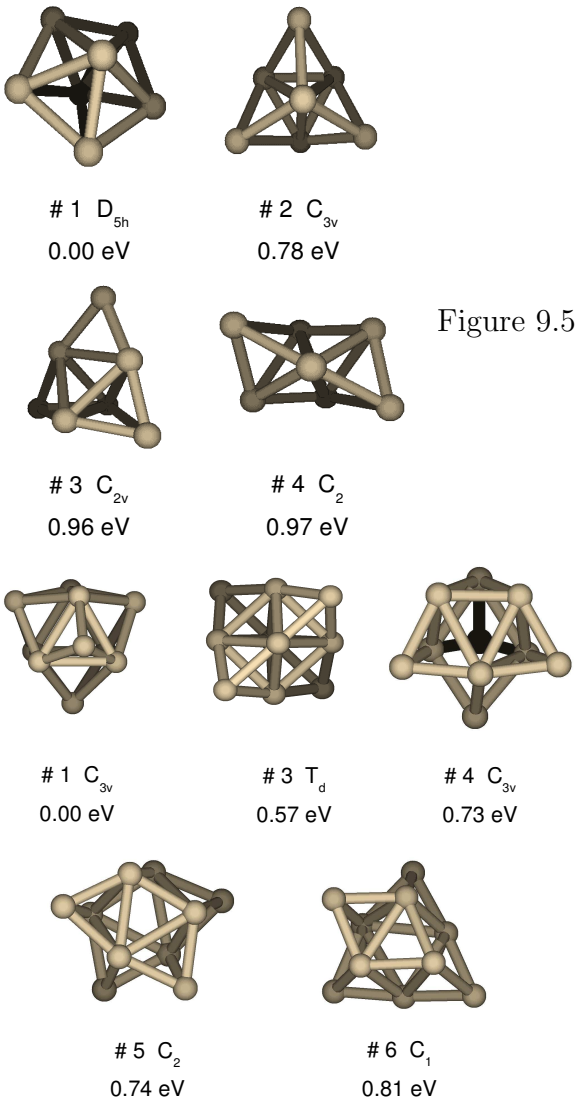

Figure 9.5: Identified stable Si_7-isomers in the energy range up to 1 eV above the ground state. The isomer numbering follows the one of Fig. 9.3 and reflects the decreasing cluster stability as indicated by the stated energies relative to the ground-state isomer #1.

Figure 9.6: Identified stable Si_{10}-isomers in the energy range up to 1 eV above the ground state. The isomer numbering follows the one of Fig. 9.3 and reflects the decreasing cluster stability as indicated by the stated energies relative to the ground-state isomer #1.

for the dominant isomers. In the case of Si_7 and Cu_7, for instance, the dominant isomers with low-symmetry have ten times more local minima on the PES than the symmetric ground-state structures and are therefore identified in the sampling much more often. Hence, the histogram can be regarded as a convolution of the size of the basin with the number of versions.

This relation to the underlying PES shape also motivates why certain isomers are dominant irrespective of the employed move class. Any general purpose move class that is intended to achieve unbiased jumps on the PES will be similarly affected by a varying size or multiplicity of the different basins of attraction. This is an important point as an at first glance appealing approach to improve the efficiency of BH sampling would be to reduce the number of times that the algorithm gets stuck in always the same dominant isomers and instead aim to increase jumps into the rare minima. Within the understanding of the relation to the PES shape it seems unlikely that this can be realized without either resorting to moves that are specifically tailored to the system at hand or

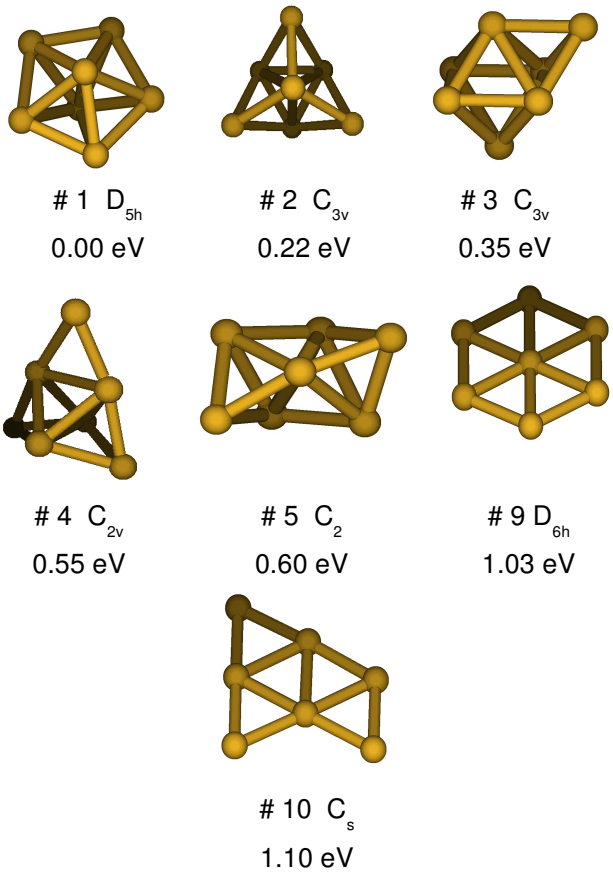

Figure 9.7: Identified stable Cu_7-isomers in the energy range up to 1.1 eV above the ground state. The isomer numbering follows the one of Fig. 9.3 and reflects the decreasing cluster stability as indicated by the stated energies relative to the ground-state isomer #1. Note that isomer #4 exhibits small imaginary eigenmodes but is nevertheless retained in the performance analysis, see text.

use other local PES information. From the perspective of an unbiased sampling ansatz with purely stochastic moves it correspondingly seems inevitable that the algorithm will (at least for the limited isomer number of the small cluster sizes studied here) often revisit the same dominant isomers. The overall performance is then obviously dictated by the way the algorithm can deal with these dominant isomers, e.g. how efficiently it can hop out of them.

9.3.2. Approximate Hopping Matrix

On the basis of the histograms presented in Fig. 9.3 we can now specify which of the energetically lowest-lying isomers are the target of the sampling runs. In the general case, this would be dictated by the physics of the problem at hand, e.g. prescribing that the sampling should yield the ground-state isomer, as well as all isomers in a certain energy range above it. However in view of the discussion above, it is clear that the overall sampling performance is governed by the dominant isomers, since the algorithm spends most of its time jumping out of these few minima. For the intended performance analysis we therefore choose as the sampling target the identification of all dominant isomers. As indicator of the sampling efficiency we correspondingly focus on the number

		point group	N_{versions}
Si_7	# 1	D_{5h}	504
	# 2	C_{3v}	1680
	# 3	C_{2v}	2520
	# 4	C_2	5040
Cu_7	# 1	D_{5h}	504
	# 9	D_{6h}	420
	# 10	C_s	5040
Si_{10}	# 1	C_{3v}	1209600
	# 5	C_2	3628800
	# 6	C_1	7257600

Table 9.1.: The number of different versions for the dominant isomers.

of moves N until all dominant isomers are found at least once. In the case of Si_7 and Si_{10} all dominant isomers are included in an energy range up to 1 eV above the ground state as apparent from Figs. 9.3, 9.5 and 9.7. In the case of Cu_7, this energy range is slightly extended to 1.1 eV above the ground state to also include the dominant isomer #10, cf. Figs. 9.3 and 9.7.

With a thus defined sampling target the natural BH acceptance criterion is to unconditionally accept trial moves that lead into any isomer in the corresponding energy window, and to unconditionally reject any trial move that leads into an isomer that is higher in energy. It would only be necessary to change the latter to some, e.g. Boltzmann weighted, conditional acceptance rule, if a multiple-funnel type PES would necessitate passages via such higher energy isomers. However, as discussed above this is not the case for the systems studied here. In terms of a hopping matrix corresponding energy-window BH runs require only the knowledge of a limited number of hopping matrix elements. Definitely required are the transition probabilities between any of the targeted low-energy isomers. Since trial moves into higher energy isomers are rejected, it suffices to know the overall probability to jump from each one of the low-energy isomers into any of the higher energy ones, without the need to further resolve the latter. For the example of Si_7 the targeted energy window comprises 4 different isomers, and energy-window BH runs can therefore be simulated on the basis of 20 hopping matrix elements: 16 transition probabilities between any of the 4 different low-energy isomers, as well as one hopping matrix element per low-energy isomer that describes the sub-summed transition probability to jump out of the isomer into any of the higher energy ones.

For a specified BH setting (i.e. fixed move type and fixed technical move parameters) we obtain the required hopping matrix elements by simply performing a fixed number of trial moves out of each of the low-energy isomers, recording the probabilities with which the moves led into each of the other low-energy isomers or any of the higher-energy ones. After 100 moves these probabilities are converged to within ± 0.1 at a confidence interval

at the level of 95%, which we found to be sufficient for the conclusions put forward below. With the thus determined hopping matrix, a large number of energy-window BH runs from different starting isomers and with different random number sequences can be quickly simulated without the need for further first-principles calculations. This allows to arrive at a properly averaged number N_{av} of moves required to determine all low-energy isomers at least once, albeit with the disadvantage that the transition probabilities are only known within the confidence interval of ± 0.1. To account for the latter, we therefore randomly varied the individual hopping matrix elements within this uncertainty range and under the constraint of Eq. (9.3). Determining the N_{av} for several hundreds of correspondingly created hopping matrices, we finally quote below the average value together with error bars given by the standard deviation.

This remaining uncertainty incurred from the approximate hopping matrix procedure does not affect any of the trend conclusions made below, yet on the other hand leads to a substantial reduction in the computational effort: In order to determine a well converged N_{av} for the systems studied here, required is typically an averaging over several hundreds BH runs starting from different initial isomers and with different random number sequences. As apparent below N_{av} for e.g. Cu_7 is of the order of 50, so that a straightforward determination of N_{av} by averaging over individual first-principles BH runs would require of the order of 5000 or more trial moves, with a corresponding number of first-principles energy and force evaluations. For the described hopping matrix based approach, however, only 100 moves out of each of the six low-energy isomers need to be done on the basis of first-principles calculations, while the ensuing hopping matrix based simulations are computationally undemanding. This reduces the overall computational cost by about an order of magnitude.

9.3.3. CPU Reduction by Using Optimum Sampling Parameters

For the ensuing move analysis, we first focus on a normal distribution which one would intuitively consider to be better, since it centers on a potential optimum value for the move distance. Further below, we then elaborate more on the random number distribution, taking additionally the uniform distribution into account.

Si_7 and Cu_7

In Fig. 9.8 we plotted the average move number along with the different fractions of moves introduced in section 9.2.3. It is apparent that for all cases a too small value of r_0 leads to a large move number N_{av} which goes down with increasing move distance until an optimum value for r_0 is reached. This rather obvious behaviour is reflected in the fraction of unsuccessful moves $\gamma_{\mathrm{unsucc.}}$ which approaches one for too small move distances and goes down with increasing values for r_0. Intriguingly, $\gamma_{\mathrm{unsucc.}}$ never reaches zero but seems to asymptotically approach a finite value. To understand this behaviour we take a closer look at the diagonal elements of the hopping matrix h_{ii}, which describe the probability of the system to relax back in the previous structure after a trial move.

Obviously, $\gamma_{\text{unsucc.}}$ is just the average of $\{h_{ii}\}$ weighted by the corresponding histogram entries H_i:

$$\gamma_{\text{unsucc.}} = \frac{1}{N} \sum_{i=1}^{N} H_i h_{ii} \tag{9.7}$$

Consequently, if all diagonal elements vanished, $\gamma_{\text{unsucc.}}$ would reach zero. The upper panel of Fig. 9.9 presents the diagonal elements of all dominant isomers of Si_7 and Cu_7 with respect to the move distance for single-particle moves. For the case of Si_7, the diagonal elements of the rather symmetric isomers # 1 - # 3 vanish with large enough distances as one would expect since the trial move then brings the system out of the current basin of attraction. For the unsymmetric Si_7 isomer with pointgroup C_2, however, the hopping matrix element seems to stay constant. The same tendency can be observed for the case of Cu_7, yet less pronounced. Clearly, the diagonal element of the unsymmetric C_s isomer does not further decrease when increasing the move distance beyond 2.0 a while the corresponding matrix elements of the more symmetric isomers seem to decrease monotonically. This is a further consequence of the comparably large number of versions of the low-symmetry structures of Si_7 and Cu_7 the system can relax into. In other words, by choosing a sufficiently large move distance, the system can be prevented from relaxing back into the previous structure but not from jumping between symmetry-equivalent basins of attraction. Consequently, $\gamma_{\text{unsucc.}}$ saturates at some point and N_{av} does not decrease any further. A similar behaviour can be observed for collective moves shown in the lower panel of Fig. 9.9.

In almost all cases, the move number starts to grow again, albeit insignificantly for practical purposes. A rational for this behaviour can be obtained by having a closer look at the fraction of rejected moves $\gamma_{\text{high}E}$ which monotonically increases with the move distance r_0. In other words, the larger r_0, the more often the system jumps out of the energy region of interest. The remaining fraction of accepted moves $\gamma_{\text{succ.}}$ is finally the kind of moves that yield new isomers within the energy window of interest, so that intuitively it should be correlated with the overall move number N_{av}. This is indeed true for all four cases, in which $\gamma_{\text{succ.}}$ reaches a maximum at the optimum move distance that minimizes N_{av}.

For Si_7, using single-particle moves, the optimum value for r_0 lies somewhere between $r_0 = 1.5\,a$ and $2.5\,a$ with only minor differences in the move number. Using collective moves, the best choice of r_0 is 0.75 a. It is clear that displacing all atoms at once, the absolute value of r_0 can be smaller compared to single-particle moves, since the more atoms are involved the less each of them needs to be disturbed to change the configuration significantly. In the case of Cu_7, the optimum move distance seems to be slightly extended to 2.0 a and 1.0 a using single-particle and collective moves, respectively. Comparing the move numbers N_{av} at the optimum value of r_0 for both move schemes, we note that there is basically no performance difference.

We further note that the fraction of accepted moves $\gamma_{\text{succ.}}$ is different for Si_7 and Cu_7 at their maximum. In the first case, roughly half of all the moves need to be accepted to optimize the efficiency, while for the latter case, the optimum fraction is

significantly higher (roughly 70 %). Hence, this example illustrates that aiming at a certain and system-independent absolute value for the acceptance ratio for all different kinds of systems is clearly not the best strategy to obtain an optimized sampling. For instance, in the case of Cu_7 using collective-moves, the move number and thus the overall efficiency is more than twice as large for an acceptance fraction of roughly a half ($N_{\mathrm{av}} > 20$) compared to the optimum value of roughly 70 % ($N \approx 10$).

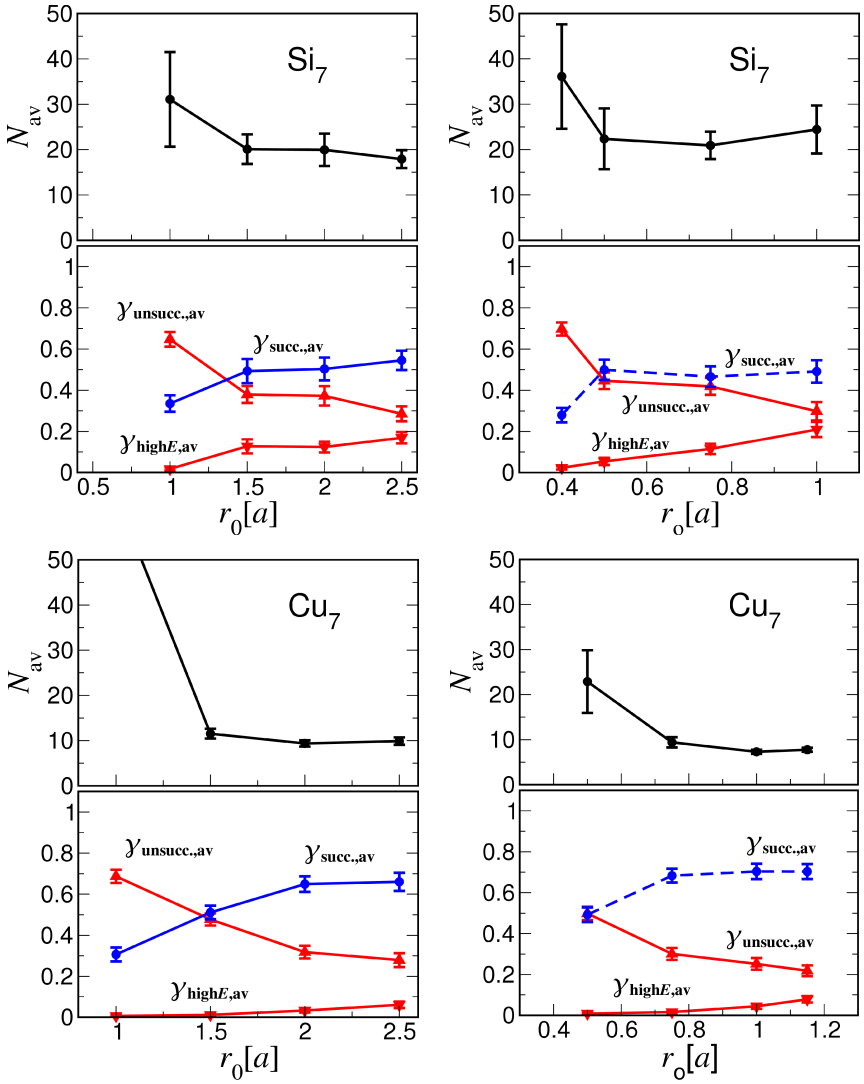

Figure 9.8.: Performance analysis for Si_7 and Cu_7 for both single-particle (left) and collective (right) moves. Given are the move number N_{av} including error-bar (upper panels) along with the different fraction of moves including error-bar (lower panels) $\gamma_{\mathrm{unsucc.}}$ (red curves with triangles pointing up), $\gamma_{\mathrm{high}E}$ (red curves with triangles pointing down) and $\gamma_{\mathrm{succ.}}$ (blue curve with circles).

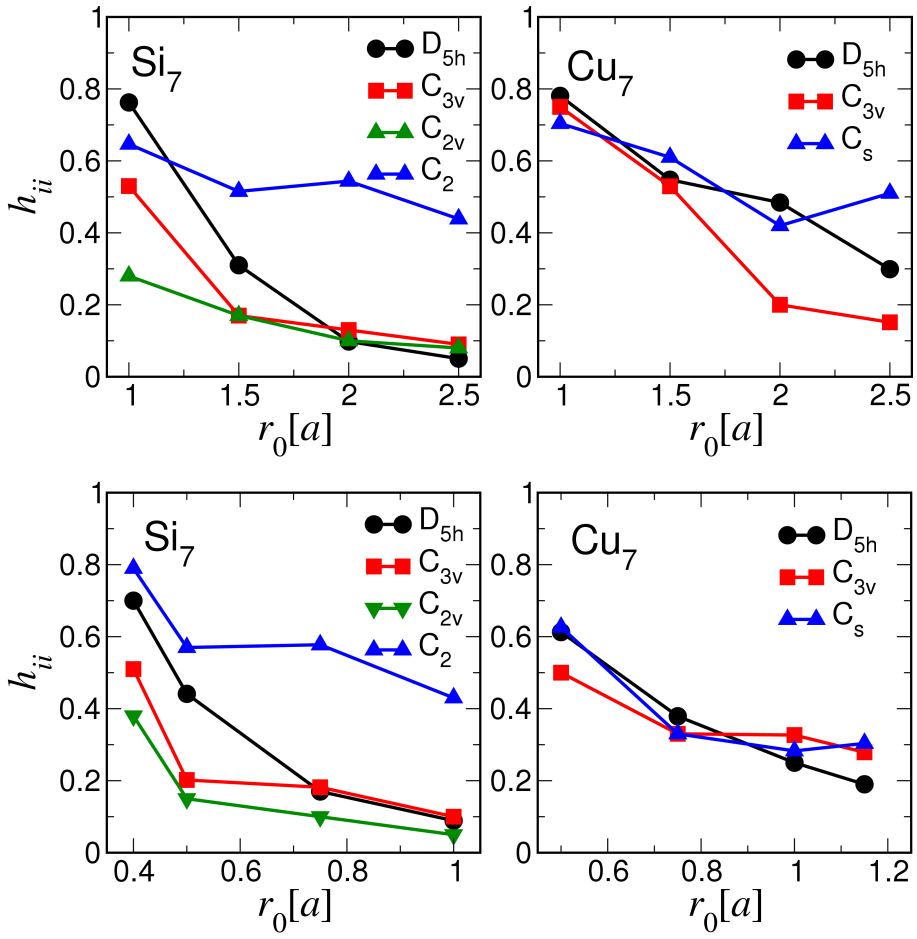

Figure 9.9.: Diagonal elements of the hopping matrix of the dominant isomers of Si_7 (# 1 (D_{5h}), # 2 (C_{3v}), # 3 (C_{2v}) and # 4 (C_2)) and Cu_7 (# 1 (D_{5h}), # 2 (C_{3v}), # 10 (C_s)) using single-particle (left) and collective (right) moves.

Si_{10}

Figure 9.10 presents the same analysis as above for Si_{10} using collective moves. There are two striking features that can be assigned to the exponential growth of the number of different isomers. Since there are obviously more isomers above the energy window of interest, the probability to jump in the high-energy region is much higher than in the case of the smaller system Si_7. This is reflected in much higher values for $\gamma_{\mathrm{high}E}$. At the maximum move distance of 1.0 a, roughly 20 % of the moves have led out of the energy window in the case of Si_7. With the same move parameter setting, four times as many moves jumped into the high-energy region in case of the larger system Si_{10}. The second effect is the stronger dependence of the unsuccessful moves $\gamma_{\mathrm{unsucc.}}$ on the move distance. Due to the higher number of local minima, the effect of the multiplicity of individual isomers that saturates the fraction of unsuccessful moves becomes less significant, thus making $\gamma_{\mathrm{unsucc.}}$ approach zero more closely. Consequently, the two competing effects of

having a too small or a too large move distance, reflected in the two different fractions of disadvantageous moves $\gamma_{\text{unsucc.}}$ and $\gamma_{\text{high}E}$, become much stronger, thus yielding a clear optimum of $\gamma_{\text{succ.}}$, which again nicely correlates with the move number N, the latter having a clear minimum in the range around 0.5 to 0.6 a. It is in particular the intuitively chosen acceptance criterium that gives rise to this clear correlation, thus making $\gamma_{\text{succ.}}$ a potential candidate for an on-the-fly indicator of efficiency. Since moves are unconditionally accepted if they lie within the energy window of interest, $\gamma_{\text{succ.}}$ corresponds exactly to the fraction of moves that explore the regions of the PES of interest.

As to the overall sampling efficiency, we further note that except for extremely unreasonable move distances, the basin-hopping scheme using unbiased collective moves is quite robust with respect to the exact value of the move distance. Around the range of 0.5 and 0.6 a, the move number varies less than a factor of two, which is in practice an insignificant uncertainty.

normal vs. uniform distribution

We now turn our attention to the random number distribution for the move distances. Intuitively, one would expect that focusing on an appropriate value r_0 as was done above is more efficient than using a uniform distribution around r_0. The rational is simple. A uniform distribution contains both a large fraction of moves with quite small distances and a fraction of moves with rather large displacements as well. While the former is inefficient since the system tends to relax back into the previous structure, the latter can decrease the efficiency since the system jumps to high in energy as was shown above. To investigate on this we explicitly compared the efficiency of both distributions. The results are summarized in Table 9.2 and clearly prove that in the case of single-particle moves using the optimum value of r_0, the normal distribution is more efficient than the uniform one. In the case of Si_7 with 1.5 a for instance, an average move number of 30.8 is obtained with the uniform distribution compared to 20.1 using a normal distribution. The same holds for Cu_7 with r_0 equals 1.5 or 2.0 a, where the uniform distribution yields an average move number about twice as large as in the case of the normal distribution. In the case of collective-moves, however, both uniform and normal distribution perform roughly the same for Si_7 and Cu_7. The rational is here simply that in the case of collective moves, each atom has an individual move distance determined by a random number around r_0. Consequently, even if some atoms are disturbed too much or too little, there are always certain atoms which are displaced with an appropriate move distance. In the case of Si_{10}, however, the uniform distribution seems to get less efficient compared to the normal distribution (e.g. 15.1 compared to 9.8 moves in the case of the optimum move distance 1.5 a). Again, this can be rationalized with the effect that a huge fraction of the moves lead to high-energy regions since the uniform distribution contains a larger ratio of move distances that are larger than the average value r_0. Summarized, the uniform distribution can only be disadvantageous for the systems under consideration and is thus no longer considered.

	move settings		normal	uniform
Si_7	single	1.0 a	31.1	34.6
		1.5 a	20.1	30.8
	collective	0.5 a	22.4	20.3
		0.75 a	20.9	17.7
Cu_7	single	1.5 a	11.5	20.3
		2.0 a	9.4	21.9
	collective	0.5 a	22.9	24.7
		0.75 a	9.4	8.0
Si_{10}	collective	0.5 a	9.8	15.1
		0.75 a	20.5	29.1

Table 9.2.: Performance comparison of normal-distributed and uniformly distributed random numbers for the move distance.

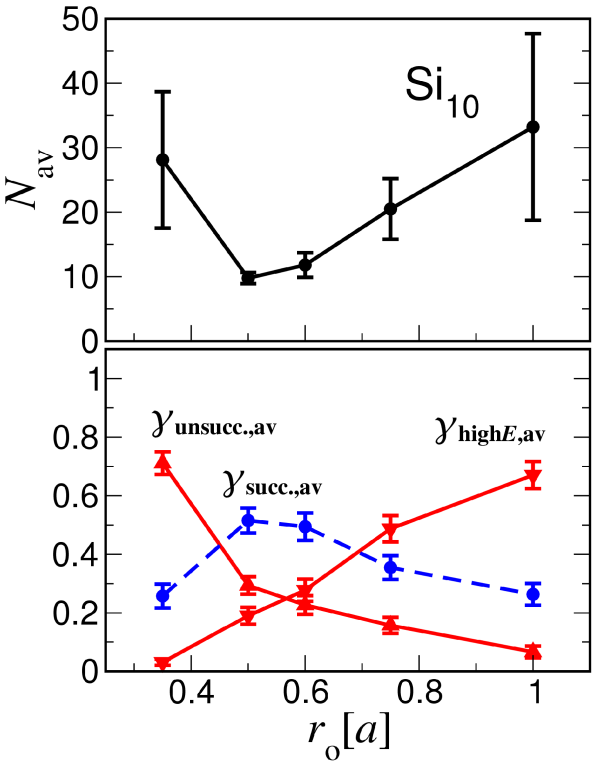

Figure 9.10.: Performance analysis for Si_{10} using collective moves.

9.4. Adaptive Move Scheme

On the basis of the above obtained insights, a simple adaptive scheme is now presented, which guides an unknown system towards reasonable move settings based upon optimizing $\gamma_{\text{succ.}}$. This efficiency indicator is therefore averaged over the last 50 moves, thus being converged up to ± 0.1 at a confidence interval at the level of 90 % (see as an example Fig. 9.11). As move scheme, collective moves are applied with a normal distribution around an average value r_0. Our suggested self-adapting scheme then updates the average move distance according to simple rules. Starting with some initial value r_0 and a step width Δr, the change in the fraction of successful moves $\Delta\gamma_{\text{succ.}}$ between two averaging periods with r_0 and $r_0 + \Delta r$ is then monitored. Since $\gamma_{\text{succ.}}$ contains noise, the sign of $\Delta\gamma_{\text{succ.}}$ cannot directly be used as an indicator whether an updated move distance has improved the efficiency or not. Hence, a tolerance value of γ_{tol} needs to be defined which was set according to the convergence rate discussed above to 0.1. After two averaging periods, three cases then need to be distinguished. If $\gamma_{\text{succ.}}$ has increased more than the prescribed tolerance value

$$\Delta\gamma_{\text{succ.}} > \gamma_{\text{tol}} \ , \tag{9.8}$$

the change in the step width Δr has obviously improved the efficiency significantly.

Consequently, the adjustment Δr points into the direction of the maximum of $\gamma_{\text{succ.}}$ and the sign is therefore kept for the next move distance adaption. Additionally, the magnitude is slightly increased to accelerate the approach to the optimum value of r_0, so $\Delta r \rightarrow \beta_1 \Delta r$, with $\beta_1 = 1.2$.

In the opposite case, if $\gamma_{\text{succ.}}$ has decreased more than the prescribed tolerance value

$$\Delta\gamma_{\text{succ.}} < -\gamma_{\text{tol}} \ , \tag{9.9}$$

the efficiency has significantly degraded and the sign of the adjustment Δr is therefore changed. To enable the system to pin down the optimum value of the move distance, the magnitude is decreased in this case, so $\Delta r \rightarrow -\beta_2 \Delta r$, with $\beta_2 = 0.3$.

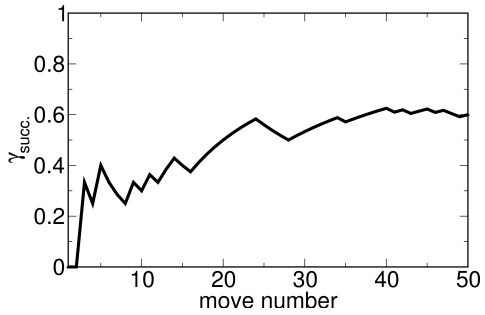

Figure 9.11.: The convergence of γ_{succ} with respect to the move number.

If $\Delta\gamma_{\text{succ.}}$ has not changed within the tolerance range,

$$-\gamma_{\text{tol}} < \Delta\gamma_{\text{succ.}} < \gamma_{\text{tol}} \ , \tag{9.10}$$

the current move distance r_0 is obviously close to the optimum range. However, in case of small "gradients" of $\gamma_{\text{succ.}}$, the system thus might get stuck in regions away from the maximum. Hence, the adjustment Δr is not set to zero, but instead only decreased in magnitude, so $\Delta r \rightarrow +\beta_2 \Delta r$.

Further fast adjustments according to $\gamma_{\text{unsucc.}}$ and $\gamma_{\text{high}E}$ additionally lead to a quick and coarse adjustment of the move distance by directly exploiting the obtained insights from the move analysis in the previous section. If the system got stock in at least 9 out of 10 moves, it is clear that the current move distance must be way too small without further averaging $\gamma_{\text{succ.}}$. So if

$$\gamma_{\text{unsucc.}} \geq 0.9 \,, \tag{9.11}$$

the move distance r_0 is immediately increased by a certain amount Δr, which was typically chosen as equal to the initial adjustment. If on the other hand, more than half of the previous 10 moves lead to the high-energy region, so

$$\gamma_{\text{high}E} \geq 0.5 \,, \tag{9.12}$$

the move distance is likely to be too large and thus decreased by Δr, again (arbitrarily) set to the initial adjustment.

This simple scheme is illustrated for Si_{10}, using collective moves, for which the optimum parameter range for r_0 is known.

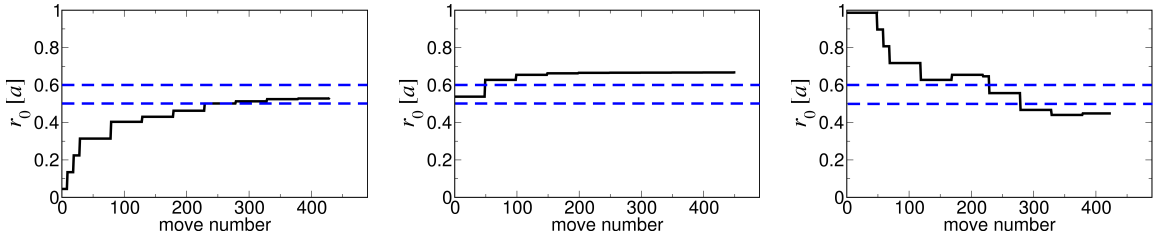

Figure 9.12.: Self-adapting move distance during a Si_{10} sampling run with window acceptance criterium and collective moves. Exemplified are three runs starting with a way too small move distance (left panel), a move distance that lies already in the optimum range (middle panel) and a way too large move distance (right panel). $\gamma_{\text{succ.}}$ is averaged over 50 moves with additional fast adjustments according to $\gamma_{\text{unsucc.}}$ and $\gamma_{\text{high}E}$. The blue dotted lines indicate the optimum range according to the performance analysis in section 9.3.3. The initial step adjustment Δr is set to 0.2 Å or 0.09 a, respectively.

Fig. 9.12 instances three cases, starting with a way too small move distance, a way too large one and with a setting that lies already within the reasonable range. Indicated are the two values 0.5 and 0.6 a that are optimum values according to sec. 9.3.3. As these three examples clearly show, the self-adapting scheme succeeds in guiding the system towards a reasonable move distance within O(100) moves. Compared to this computational demand, even one data point in the move analysis for Si_{10} required 2000 trial moves, 100 moves for each of the 20 hopping matrix elements. Hence, our suggested self-adapting scheme provides an efficient scheme which exploits the on-the-fly efficiency indicators, thus preventing the system from inefficient sampling due to unreasonable move settings. As to the converged move distances in the three example runs, we note

that the optimum range where $\gamma_{\text{succ.}}$ is maximized, obviously lies somethere between 0.45 and 0.7, thus not exhibiting a sharp peak but rather a broad plateau where the overall sampling efficiency is quite unaffected by the exact value of the move distance, which is in line with the insights obtained in section 9.3.3.

9.5. Summary and Conclusion

In this part of the work, the efficiency of an *ab initio* based basin-hopping run was quantitatively analyzed to reveal inefficient settings, thus suggesting potential ways to optimize the technical settings, in particular discussing consequences and possible strategies for larger cluster runs. To pin down the efficiency of a sampling run under a certain set of parameters, the common scheme in literature to measure the mean first encounter was extended to the number of moves until all relevant isomers within a pre-defined energy range have been identified at least once. This defined measure of efficiency is therewith coupled to the sampling goal of not only finding the ground state but also higher-lying isomers. To make the analysis computationally tractable, the concept of a so-called hopping matrix has been developed, thereby circumventing the necessity to perform hundreds of *ab initio* basin-hopping runs in order to get a statistically meaningful move number with respect to the stochastic noise and the dependence upon the starting geometry. Further on-the-fly measurable quantities have additionally been suggested that monitor the efficiency during a sampling run instead of just identifying an inefficient setting *a posteriori*, thus being candidates for possible self-adapting schemes. In order to elaborate on the question which isomers are relevant, long basin-hopping runs for Si_7, Si_{10} and Cu_7 under different move settings, using both single-particle and collective moves, have been performed, revealing dominant isomers as an intrinsic feature of the PES. This property can be traced back to either a huge size of the corresponding basin of attraction or to the huge multiplicity of different versions corresponding to the same isomer. Since the system spends most of the time jumping between the dominant isomers, they obviously govern the overall performance and have thus been chosen as sampling goal. The performance analysis of the trial move is thereby separated from the acceptance criterium by using a window acceptance scheme, according to which isomers that lie within the energy window containing the dominant isomers are unconditionally accepted. Supported is this strategy by the fact, that for these cluster sizes, the simple stochastic trial move schemes enable efficient jumps between all parts of the configuration space. Using both single-particle and collective moves, the performance analysis yields a clear correlation between the move number N_{av} and the on-the-fly indicator $\gamma_{\text{succ.}}$, which is defined as the fraction of moves that lead to new isomers that lie within the energy window of interest. Comparing single-particle and collective moves, there is basically no difference between the move numbers at the corresponding optimum move distances. Further analysis of the random number distribution for the move distance only revealed a slight disadvantage of the uniform distribution compared to the normal distribution. Hence, the sampling in connection with simple stochastic moves turns out to be quite robust with respect to the analyzed parameters. Significantly inefficient

settings result mainly from way too small move distances, so that the system relaxes essentially always back to the previous minimum, or way too large distances, which tend to bring the system into high-energy PES regions, thus being out of the energy range of interest. In particular the last tendency becomes stronger with increasing cluster size, since the number of isomers exponentially grows and with that the configurational space in the high-energy region. Such inefficient settings can be avoided by a simple self-adapting scheme that is based upon the optimization of the indicator $\gamma_{\text{succ.}}$, which has been presented and illustrated for Si_{10}. Apart from unknown systems, for which this scheme is able to quickly identify the reasonable range of move distances, the system is expected to locally adjust the move distance in case of significant differences in the topologies in multi-funnel-systems. To gain further significant improvements in the sampling efficiency, more sophisticated trial moves would have to be designed that are able to explore the properties of the PES, like in the spirit of the MD-trajectories exploiting the BEP-principle. However, this always bears the danger of biasing the sampling. A completely different strategy would therefore be not to put emphasis on the sampling of the PES, but on the evaluation of the PES itself using sophisticated analytical potentials that are able to capture at least the same structural motifs as an *ab initio* energy surface.

10. Summary and Outlook

The motivation of the present work was to develop and analyze a DFT based basin-hopping algorithm for the structure determination of atomic clusters. This methodological approach has been applied for the interpretation of FIR-MPD data of small $Co_n^+Ar_m$ (n=4-8) clusters, measured in the group of Prof. Gerard Meijer at the Molecular Physics Department of the FHI. The dataset of isomers generated from the unbiased sampling of configurational space included several cases of differently Jahn-Teller distorted versions of the same basic structural motif, which consequently would likely have been missed using the commonly employed scheme to simply try out "usual suspect" structures derived from chemical intuition. Specific conclusions concerning this part of the work, in particular on the nature of the Co-Ar binding, were already presented at the end of the corresponding chapter, which is why this final chapter is restricted to an outlook of only the methodological aspects.

The structure determination based on the comparison of vibrational spectra relies on the fundamental assumption that different isomers do exhibit distinguishable fingerprints. In the present work, however, many isomers, in particular the differently Jahn-Teller distorted ones, exhibited rather similar fingerprints with only minor quantitative differences, which in turn severly limited the approach to exclusively use FIR-MPD data to indirectly determine the atomic structure. Only in few cases the structure could be unambiguously determined, while in many cases some isomers could at best be ruled out or, put more modestly, turned out to be energetically rather unlikely. Taking additional experimental data into account, like for instance photoelectron spectroscopy, can therefore significantly facilitate the structural assignment by analyzing different independent fingerprints.

The involved exploration of the PES requires to combine two aspects. In order to obtain reliable energetics, a quantum-mechanical description of the PES is essential, which is nowadays typically achieved by density-functional theory. While single local optimizations of clusters at the size of few atoms constitute simple tasks with present day computational facilities, a basin-hopping run already at this system size may contain a few hundred optimizations, dictating already a high efficiency of both the calculation of the underlying PES and the sampling parameter. Using small silicon and copper clusters, we therefore demonstrated how the efficiency of an *ab initio* basin-hopping run can be critically assessed. For the system sizes studied, the crucial aspect is the trial move scheme, i.e. the way a new trial structure is generated from the previous one. The comparison of the efficiency for different move settings herein necessitates a quantity that pins down the computational demand of a certain run. For this, we extended the common scheme in the literature to count the number of moves until the ground state has been identified for the first time, to the mean first encounter of defined low-energy

isomers, since it is not only the ground-state structure that is of fundamental interest. The first step of the performance analysis hereby identified dominant (or frequently visisted) isomers as an intrinsic feature of the PES. For the size range investigated, they consequently govern the overall computational demand of the sampling and are therefore the relevant isomers for the performance analysis. To lower the computational cost and additionally obtain detailed analysis data we circumvented the straightforward calculation of the manifold of sampling runs by introducing the concept of a hopping matrix. By recording the transition probabilities between the individual isomers within the energy range of interest, statistically meaningful quantities can then be extracted without the need for further first-principles calculations and the converged hopping matrix elements provide the connectivity of the low-lying isomers. This led to the observation of clear correlations between the number of required moves with on-the-fly efficiency indicators, that could then be used to develop a simple self-adapting scheme.

The move analysis focused on single-particle and collective moves. At the small cluster sizes studied no clear dominance of one move class over the other could be identified. Furthermore, each individual class turned out to be quite robust with respect to the average step width with which atoms are randomly displaced. The governing factors leading to this optimum move distance are the inability to escape from the basin of attraction of the present configuration at too short distances and the increased probability to end up in high-energy isomers at too large distances. The suggested self-adapting scheme exploits these insights and its performance has been demonstrated for Si_{10}, for which the self-adapting scheme guides and keeps the system within reasonable move settings.

In this respect, one has to note that simple stochastic move schemes provide only little flexibility to significantly improve the sampling efficiency. While this motivates the design of more sophisticated trial moves at first glance, such an approach always bears the danger to introduce bias or loose the general-purpose character. Nevertheless, when assessing such more specialized move types (also in view of the much more demanding size range beyond ten atoms) the evaluation should be based on a performance analysis protocol as presented in this work. In any case, orders of magnitude in the efficiency are much more likely to be gained not by optimizing the exploration, but instead the calculation of the PES itself. Simple empirical potentials can thereby not guarantee that the same structural motifs are captured, not to speak of their energetic order, thus making them unsuitable for a fast coarse screening of the configurational space. More promising are approaches that are based upon a parameterization of the *ab initio* PES, and therefore allow for a systematic convergence of the accuracy. In a collaboration with Dr. Jörg Behler from the Ruhr-Universität Bochum, we currently pursue this direction based on a neural network approach, which has recently been successfully applied to bulk silicon and achieved *ab initio* accuracy [127]. With an equally accurate parameterizaton for finite systems, this approach would then be an ideal candidate to enable sampling runs for significantly larger cluster sizes that are presently unfeasible with straightforward DFT energetics.