Guesstimation of Molecular Ensemble Electrostatics Properties Through SCERPADFT Calculation: Molecular FieldCoupled Nanocomputing as a Case Study

ADVANCED THEORY AND SIMULATIONS
www.advtheorysimul.com

Guesstimation of Molecular Ensemble Electrostatics Properties Through SCERPA-DFT Calculation: Molecular Field-Coupled Nanocomputing as a Case Study

Yuri Ardesi, Fabrizio Mo, Marco Vacca, Gianluca Piccinini, and Mariagrazia Graziano*

In the field of electronics, molecular technologies provide promising opportunities for innovators and scientists to advance technological progress. At the molecular scale, the simulation of ensembles becomes fundamental to advancing the fabrication, design, and prototyping of new technologies. This work proposes a framework leveraging the SCERPA tool and DFT calculation to efficiently evaluate the electronic properties of molecular ensembles. The Molecular Field-Coupled Nanocomputing (MolFCN) is considered as a case study to validate ab initio-comparable precision resulting from the SCERPA calculation on charge-constrained multi-molecule systems. In addition, it is demonstrated that the SCERPA results can be used as a nonrelativistic initial guess of DFT calculation, eventually reducing the ab initio computation time by 86 %. Finally, a periodic molecular FCN system is proposed, named *SelfPolarizer*. The proposed framework is employed to demonstrate that the ensemble naturally encodes QCA-like digital information, providing the first simulated proof of concept for MolFCN technology obtained with DFT precision.

1. Introduction

The advent of molecular technologies brought innovation in all technology fields, either at the research or market level.[1–3] In electronics, molecular technology gives hope to scientists to keep improving digital performance, coping with the obstacles imposed by complementary metal-oxide-semiconductor (CMOS) scaling through single-molecule devices. For this purpose, tuning the fabrication processes and choosing the correct organic compounds represent a crucial stage in the design process. Therefore, the simulation of organic materials at the nanoscale, leveraging ab initio techniques, becomes fundamental to advancing nanoelectronic fabrication, design, and prototyping.

Among the molecular-scale technologies, the Molecular Field-Coupled Nanocomputing (MolFCN), considered in this work as reference technology, encodes the information in the charge distribution of single molecules according to the quantum-dot cellular automata, and propagates information through local electrostatic interactions.[4] **Figure 1a** shows, as an example, the 1,4-diallyl butane molecule. The aggregation of charges on this molecule permits the encoding of logic information. When two molecules are encased in a MolFCN cell, see Figure 1b, the intermolecular interaction induces antiparallel dipoles which can be used to encode binary information.[4] Composing layouts of molecules makes it possible to realize digital elaboration.[5] Figure 1c shows an example of a molecular wire propagating logic '1'.[6]

Despite MolFCN promising advantages, such as highly low-power dissipation and expected very high operating frequency,[7] researchers struggle to build prototypes, challenging difficulties in correctly positioning molecules and detecting the single-molecule properties.[8] As an example, several attempts have been made to measure the molecule charge,[9,10] yet, for MolFCN, the researchers need to position molecules precisely and measure the charge once the device is on to permit integration with CMOS electronics.[8] In this context, ab initio calculation is essential to assess MolFCN and, in general, any situation where the experimental approach is challenging. The ab initio technique, such as the Density Functional Theory (DFT), is a viable solution for analyzing small molecular systems. However, it is generally limited by the high computational costs and the limitations imposed by the engineering perspective, which generally requires features not available in many DFT packages. For example, MolFCN requests simulating multi-cation ensembles and applying time-dependent non-uniform electric fields. Therefore, ab initio-based high-level models, such as the Self-Consistent Electrostatic Potential Algorithm (SCERPA), were introduced to predict the behavior of complex MolFCN circuits,[11] and

Y. Ardesi, F. Mo, M. Vacca, G. Piccinini
Department of Electronics and Telecommunications
Politecnico di Torino
Turin 10129, Italy
E-mail: yuri.ardesi@polito.it

M. Graziano
Department of Applied Science and Technology
Politecnico di Torino
Turin 10129, Italy

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adts.202500812

DOI: 10.1002/adts.202500812

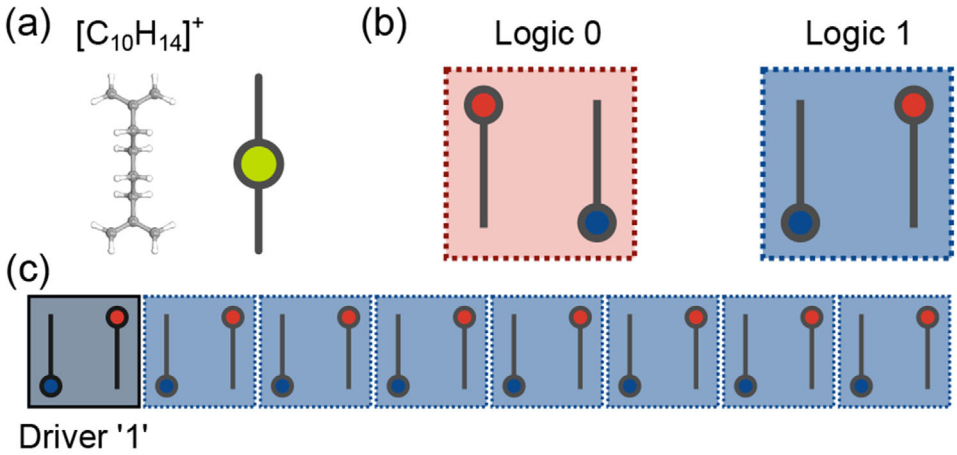

Figure 1. Molecular Field-Coupled Nanocomputing system. a) The 1,4-diallyl butane cation ($[C_{10}H_{16}]^+$) and its schematic representation; the position of the circle in the schematic represents the centre of the positive charge. b) Encoding of the binary information in diallyl butane MolFCN cells. c) MolFCN wire propagating the logic "1" imposed by a driver cell.

validated - so far - on small systems of two molecules.[12] Reliable and validated high-level simulators enable assessing beyond-CMOS technologies and, as a secondary effect, guide the realization of prototypes and experimental proofs-of-concept.

This work proposes a framework leveraging the SCERPA tool and ORCA DFT package to efficiently evaluate the electronic properties of many-molecule systems, permitting constraining the molecular charge to each molecule. As a case study, we propose a periodic molecular FCN system named *SelfPolarizer*, which naturally encodes QCA-like digital information. The charge distribution obtained with SCERPA is consistent with Quantum Mechanics/Molecular Mechanics (QM/MM) calculation, thus demonstrating the possibility of exploiting SCERPA for simulating MolFCN circuits and providing the first simulative demonstration for MolFCN technology obtained with DFT precision.

To sum up, this work validates the very high precision resulting from self-consistent calculation on multi-molecule systems for the first time. More precisely, the SCERPA results are consistent with the DFT precision, thus demonstrating the tool validity in evaluating the charge distribution of molecular ensembles. In addition, we demonstrate that the SCERPA results can be used as a nonrelativistic initial guess of DFT calculation, eventually speeding up the ab initio calculation and reducing the computation time by 86 %. This work demonstrates the validity of SCERPA, either as a DFT booster or evaluation tool, in any problem where molecules are supposed to be independent, i.e., no charge transfer among molecules is expected, making it promising for the rapid evaluation of future electronic devices based on the molecule charge distribution.

2. Results and Discussion

The proposed framework in this study simulates an ensemble of molecules, selected from a set of pre-evaluated species. The framework, schematized in **Figure 2**, includes the ab initio characterization procedure, the high-level estimation, and the DFT refinement. This section presents the proposed framework and discusses the results, whereas the computational details and models are reported in Section 4.

In the ab initio characterization procedure, each molecular species undergoes ab initio geometry optimization to determine the exact position of the molecule atoms. Secondly, DFT single-point calculation evaluates the equilibrium atomic charges of the molecules, describing their electrostatic characteristics. Therefore, each species is associated with a corresponding Aggregated Charge (AC) model. The position of the AC is optimized to provide the best intermolecular interaction evaluation, which is based on the evaluation of molecule electrostatic potential, in the SCERPA algorithm. Section 4 reports the exact procedure for deriving the position of the ACs. Finally, the aggregate charge distribution is evaluated for different input voltages using DFT calculation, thus composing the so-called Vin-AC Transcharacteristics (VACT). The characterization procedure also stores ORCA orbitals and atomic charges in each input voltage.

The **high-level estimation** section starts from the layout of the molecular ensemble given by the user. The molecular ensemble is written as a set of molecules with specific positions, rotations, and charges. The layout is passed to the SCERPA tool, which substitutes each molecule with the AC model. To provide charge constraint capabilities, SCERPA considers cations as different molecular species from their neutral counterparts. Finally, the SCERPA tool outputs the *SCERPA Aggregated Charge (SAC)* solution, i.e. each molecule input voltage and the corresponding aggregated charge distribution, fulfilling our aim to obtain the electrostatic properties of the molecular ensemble. The obtained solution can be refined by exploiting the stored atomic charges to compose the *SCERPA Interpolated (SINT)* solution.

The **DFT refinement** performs the ab initio analysis of the ensemble. The nonrelativistic initial guess is performed through the orbital estimation procedure. It associates each molecule with a first orbital configuration deriving from SCERPA results. In this work, the DFT calculation is first calculated without the self-consistent field procedure, *DFT no SCF (DnSCF)*, to determine the quality of SCERPA results. Secondly, a standard single-point calculation, *DFT with SCF (DSCF)*, determines

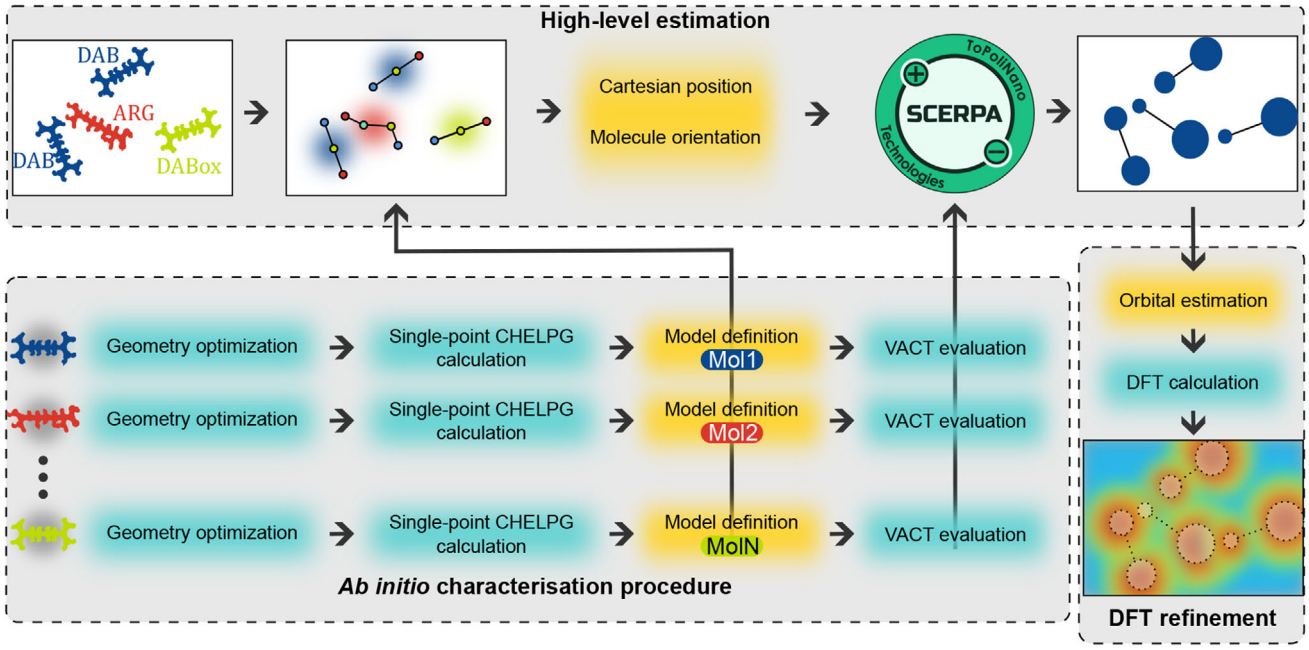

Figure 2. Proposed framework.

how far the SCERPA solution is from the lowest-energy calculation. Finally, an independent single-point calculation, named *DFT Random Guess (DRG)*, is performed with default ORCA initial guess to compare the obtained results with the unbiased calculation.

2.1. Characterization of Molecules

This work considers molecular ensembles composed of molecules from three molecular species: the neutral 1,4-diallyl butane, the arginine, and the 1,4-diallyl butane cation.

Figure 3a shows the characterization of the 1,4-diallyl butane ($C_{10}H_{14}$) in its neutral form, denoted as **(1)**. The molecule, composed of a butane (C_4H_4) and two allyl groups (C_3H_4), is a reference in the modeling of molFCN technology in its cationic form,[4] sometimes studied also as a neutral molecule.[13] This work exploits molecule **(1)** to study ensembles of neutral molecules, which can polarize due to intermolecular electrostatic forces. The molecule **(1)** is approximated with three ACs (Q1, Q2, and Q3) arranged in a line. Q1 and Q2 represent the charge aggregation in the allyl groups, whereas Q3 comprises the charge in the butane. The molecule is neutral. Thus, the Molecule Electrostatic Potential (MEP) resembles the molecule shape and vanishes for long distances. The VACT of the molecule shows the molecule polarization when exposed to electric fields. The field favors charge movement between the two diallyl groups (Q1 and Q2), whereas the charge is constant in the butane group (Q3).

Figure 3b shows the characterization of the neutral **arginine** ($C_6H_{14}N_4O_2$), denoted as **(2)**. The molecule is a polar amino acid exploited in this work to introduce electrostatic intermolecular interaction, i.e. *dipole-induced dipole forces*, in the molecular ensembles. The model uses four ACs to approximate the amino acid electrostatic behavior: Q1 and Q2 represent the external atoms of the molecule, CN_3H_3 and $C_2H_4NO_2$, respectively, whereas Q3 and Q4 model the molecule internal atoms, C_2H_5N and C_2H_4, respectively. The non-linear geometry of the arginine is reflected in its AC model, which presents a non-linear distribution of the ACs. The MEP clearly shows the polar nature of the amino acid. Indeed, the MEP demonstrates geometrical potential oscillation causing the polar nature. In particular, a low potential region can be noted on the region $y \in (0, -5)$ near Q1. The polar nature is also reflected in the VACT, which shows the low attitude of the molecule to polarize as a consequence of the electric field. In contrast to **(1)**, the arginine ACs have different values, demonstrating the non-uniform distribution of the molecule charge on **(2)**. Q1 demonstrates a fixed negative charge consistent with the shown MEP.

Finally, Figure 3c shows the characterization of the **1,4-diallyl butane cation** ($[C_{10}H_{14}]^+$), denoted as **(3)**. This molecule is generally also used in the modeling of molFCN technology.[4,14–16] Molecule **(3)** is exploited in this work to introduce *ion-induced dipole forces* in the molecular ensembles. The geometry and the AC model are similar to the neutral counterpart, depicted in Figure 3a. Instead, Figure 3c shows radially decreasing positive MEP, which is caused by the cationic nature of the molecule. The VACT of the molecule, as already highlighted in,[14] shows a clear polarization of the molecule when subjected to an electric field, which generally makes encoding information according to the molFCN paradigm possible. Notably, the MEP or the VACT shows a null net dipole moment in the molecule for no applied voltage, consistent with literature results.[16]

Overall, the characterization of the three molecular species permits modeling the electrostatic characteristics of the molecules. The obtained model geometrical information and transcharacteristics can be inserted into the SCERPA tool to provide computation capabilities.

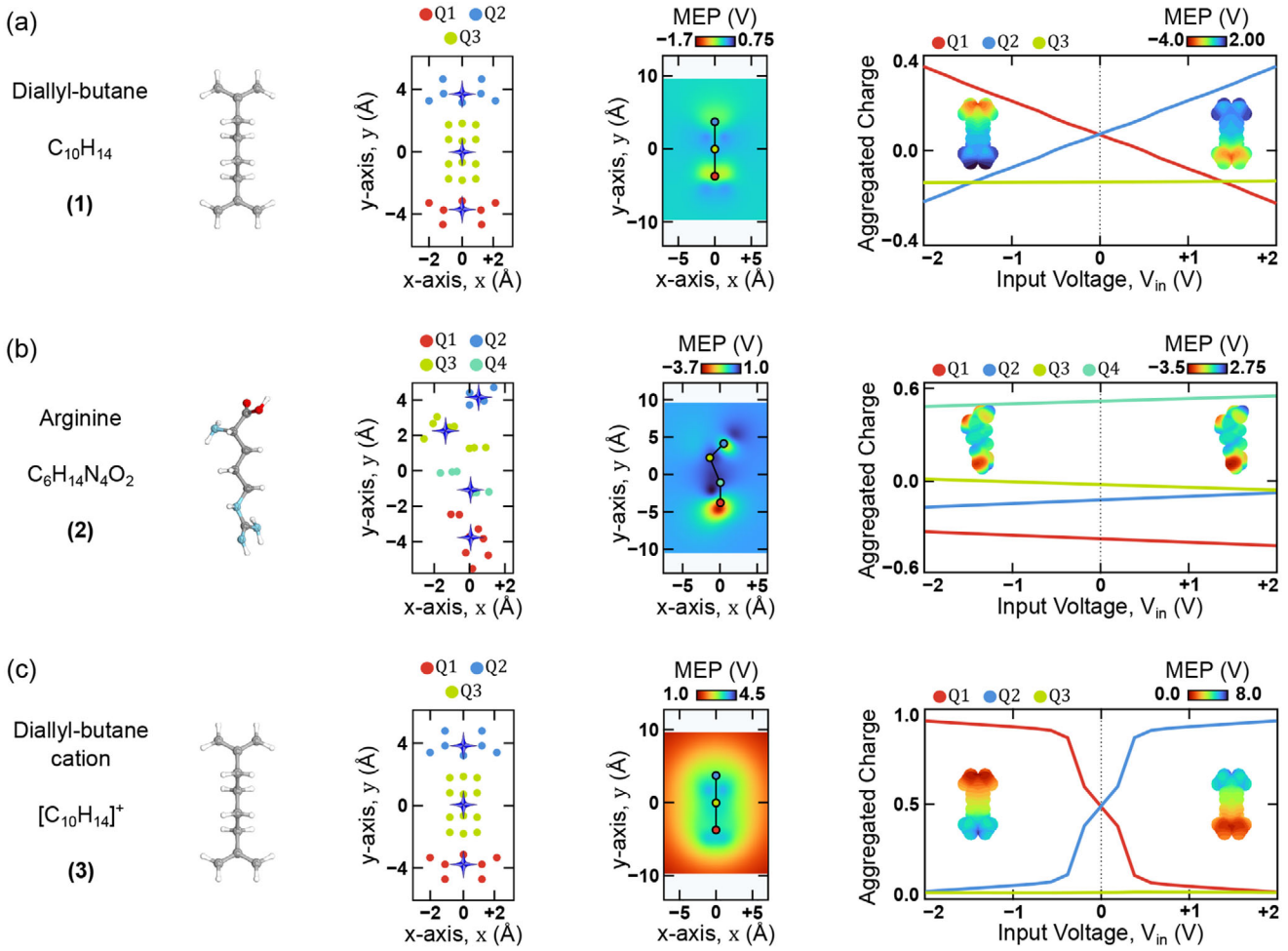

Figure 3. Characterization of molecules: structure, atomic cartesian coordinates, molecule electrostatic potential (MEP) and Vin-Aggregated Charge transcharacteristics (VACT). a) The neutral 1,4-diallyl butane ($[C_{10}H_{16}]^+$), Molecule (1). b) The 1,4-diallyl butane cation ($[C_{10}H_{16}]^+$), Molecule (2). c) The neutral arginine ($C_6H_{14}N_4O_2$), Molecule (3).

2.2. Characterisation of a Neutral Molecular Ensemble

In the first analysis, we evaluate the MEP of a simple ensemble composed of 1,4-diallyl butane and arginine molecules. **Figure** 4a shows the simulated ensemble consisting of five neutral molecules. Three neutral 1,4-diallyl butanes **(1)**, indicated as *Mol1*,*Mol2*, and *Mol3* might polarize due to intermolecular interactions with each other, and with the two polar arginines **(2)**, indicated as *Mol4* and *Mol5*.

At first, we evaluate the ensemble DRG solution, which will be the reference for the following evaluations. **Table** 1 shows the details of the DRG simulation. In particular, the charge on each molecule, obtained by summing all the atom charges in each molecule, denotes that some charge may be moved between the molecules. It is important to highlight that the atomic charges are obtained by fitting the MEP; thus, they depict the molecule electrostatic behavior rather than the spatial distribution of electrons in the molecules. The very low values obtained for the five molecules indicate that no electron transfer is present between the molecules, making this ensemble great for validating the proposed method. Figure 4b shows the obtained MEP evaluated at 0.2 nm above the molecule plane. The produced MEP depicts no discernible evidence of polarization in the 1,4-diallylbutane, instead, it highlights the non-uniform charge distribution of the arginine, which leads to negative and positive potential regions. The intuition is also confirmed by the dipole moment analysis, reported in Figure 4c, which shows a shallow dipole moment for the two 1,4-diallyl butane and a clear polarization of the two arginines. The obtained DRG solution is used as a reference to evaluate the quality of other solutions. For this purpose, Table 1 reports the maximum (E_V) and average (Δ_V) errors in the evaluated MEP, obtained by comparing the DRG solution with the following analyses.

Therefore, we perform the SCERPA simulation based on the model derived in Section 2.1 for **(1)** and **(2)** to derive the SAC solution. Figure 4d shows the obtained MEP. Again, no particular polarization can be noticed in the 1,4-diallyl butanes, whereas a non-uniform charge distribution can be appreciated in the arginines. The potential resembles the DRG solution, even though an error is clearly present ($E_V = 0.943$ V, $\Delta_V = 0.025$ V) in the evaluation. Concerning the dipole moment, reported in Figure 4e, the SAC solution provides a good estimation of the dipole moment on the

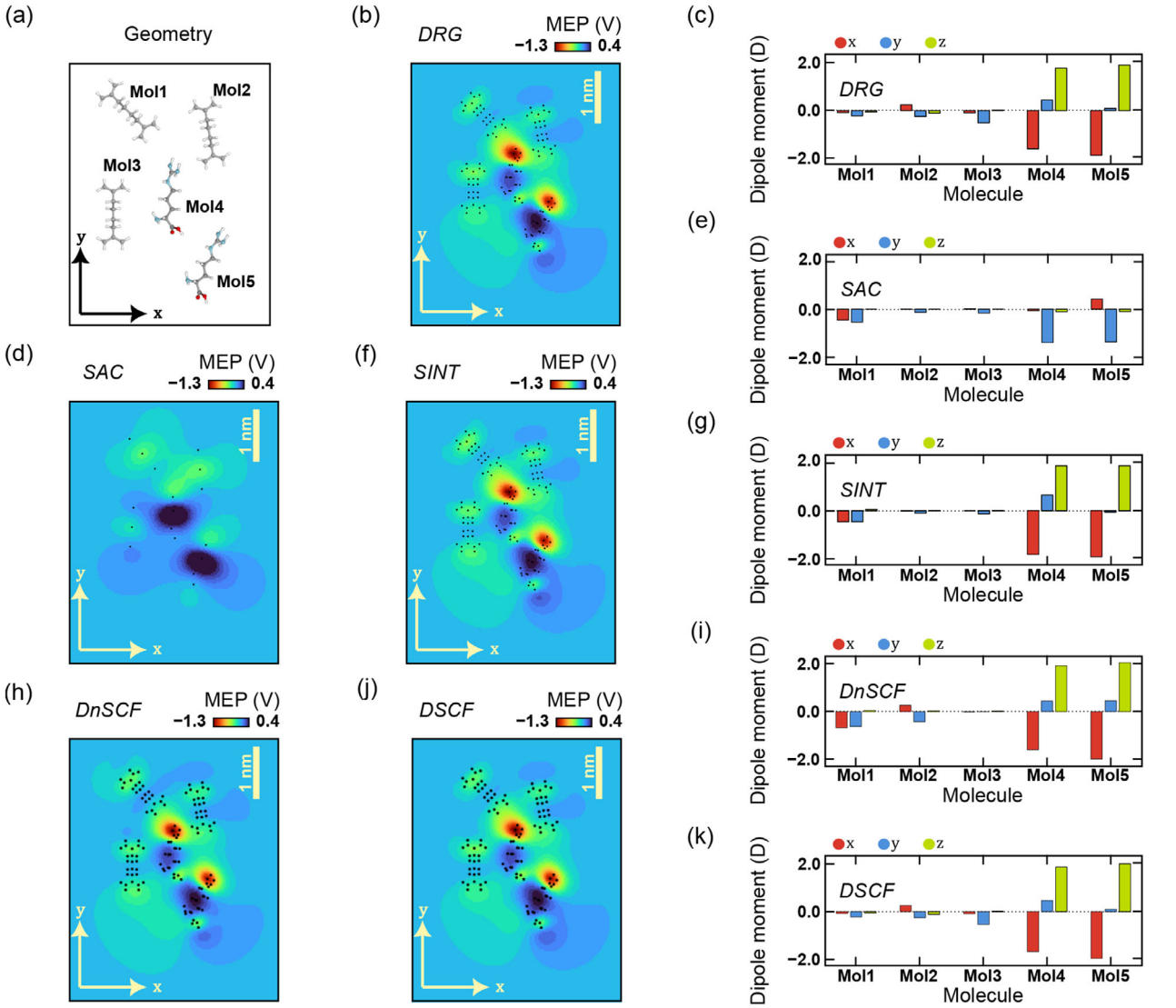

Figure 4. Neutral ensemble. a) Geometry of the analyzed neutral ensemble composed of three neutral (1) and two polar (2) molecules; b) MEP evaluated from the DRG solution; c) Dipole analysis evaluated from the DRG solution; d) MEP evaluated from the SAC solution; e) Dipole analysis evaluated from the SAC solution; f) MEP evaluated from the SINT solution; g) Dipole analysis evaluated from the SINT solution; h) MEP evaluated from the DnSCF solution; i) Dipole analysis evaluated from the DnSCF solution; j) MEP evaluated from the DSCF solution; k) Dipole analysis evaluated from the DSCF solution.

1,4-diallylbutane molecules; in addition, the result on the y-axis is always consistent with the DRG solution, denoting SCERPA capabilities to provide precise simulation results on molecular ensembles.

To obtain the SINT solution, we interpolate the ab initio data. Figure 4f shows the obtained results, which will approximate the DRG solution ($E_V = 0.025$ V, $\Delta_V = 0.001$ V). In this case, the MEP well resembles the DRG solution, demonstrating the non-uniform charge distribution of the arginine molecules. The resulting quality is also confirmed by the dipole moment analysis shown in Figure 4g, denoting excellent agreement with the DRG solution, shown in Figure 4c. It is worth noticing that the very high increment in precision is obtained with a minimal computational time overhead. Indeed, the SINT solution only requires an interpolation, whose computational cost can be made $O(1)$.[12] The obtained results, either SAC or SINT, have been obtained in 0.22 s, which is a minimum time if compared to the DRG calculation (6927.477 s).

So far, the evaluation permitted assessing SCERPA as a valid and fast estimator of the electrostatic characteristics of neutral molecular ensembles. As a second possibility, this work demonstrates that the SCERPA tool can be exploited to determine the ab initio initial guess and speed up the first-principle calculation. Figure 4h reports the result of a DFT calculaton with a disabled SCF loop (DnSCF). This result permits analysing the quality of the initial guess, since no energy minimization procedure is performed in the calculation. The calculation, performed in 778.283 s, well resembles the DRG solution even though the obtained error ($E_V = 0.090$ V, $\Delta_V = 0.001$ V) is more significant with respect to the SINT solution. The DnSCF initial guess is ob-

Table 1. Neutral 3 neutral dab + 2 arginine simulation.

Simulation	QM1	QM2	QM3	QM4	QM5	Time (s)	Energy (Eh)	E_V (V)	Δ_V (V)
Mol1: (1), Mol2: (1), Mol3 (1), Mol4 (2), Mol5 (2)									
DRG	−0.0003	0.0168	−0.0026	−0.0166	0.0027	6927.477	−2384.19493933	—	—
SAC	10^{-6}	10^{-6}	$2{\cdot}10^{-}6$	10^{-6}	10^{-6}	0.220324	—	0.943399	0.024927
SINT	10^{-6}	-10^{-6}	0.0	$2{\cdot}10^{-6}$	0.0	—	—	0.024760	0.001431
DnSCF	0.0008	0.0169	−0.0016	−0.0157	−0.0004	778.283	—	0.089272	0.000865
DSCF	−0.0006	0.0170	−0.0026	−0.0166	0.0027	924.595	−2384.19494349	0.000426	0.000004
Mol1: (3), Mol2: (3), Mol3 (3), Mol4 (2), Mol5 (2)									
DRG	0.9984	1.0070	0.9924	−0.0069	0.0090	19138.349	−2383.52462274	—	—
SAC	1.0000	1.0000	1.0000	0.0000	0.0000	0.175509	—	0.917489	0.024303
SINT	1.0000	1.0000	1.0000	0.0000	0.0000	—	—	0.133219	0.003293
DnSCF	0.9996	1.0139	0.9945	−0.0121	0.0042	781.198	—	0.150148	0.003057
DSCF	0.9986	1.0068	0.9925	−0.0069	0.0090	2652.39	−2383.52449168	0.100879	0.002064

tained by associating each molecule input voltage, obtained with SCERPA, to the orbitals in the characterization library with the closest input voltage. Therefore, no interpolation is performed in the initial guess determination, which is the primary error source. Therefore, the final result is less precise than the SINT solution. Figure 4i shows the DnSCF dipole moment analysis, which again resembles the DRG solution, yet provides more difference to all the other solutions: e.g., Mol3 polarization is almost negligible. Table 1 reports the molecular charges, which are non-null in all the molecules. The molecule charges are relatively small values resembling the DRG solution. Notwithstanding the low precision of the obtained results, the DnSCF solution is attractive for evaluating possible properties of the molecular ensemble, e.g., the orbital shape or the electron density, directly from the SCERPA calculation without the need for a complete DFT calculation. Indeed, the computational time is significantly lowered with respect to the DRG solution. Finally, the same calculation can be performed by enabling the SCF loop. In this last result, the obtained MEP is precise ($E_V = 4 \times 10^{-5}$ V, $\Delta_V = 4 \times 10^{-6}$ V). Figure 4j,k shows the obtained MEP and dipole moment, respectively, which greatly resemble the DRG solution. It is fundamental to notice that the computational time is 924.595 s, which represents a reduction of 87 % computational time to the pure DFT DRG calculation. In addition, it is interesting to notice that the obtained result represents a more stable solution than the DRG calculation, indeed, the DSCF solution has lower energy (-4.16×10^{-6} Eh).

2.3. Characterization of a Cationic Molecular Ensemble

In this section, we evaluate the MEP of the ensemble analyzed in Section 4 by substituting the neutral 1,4-diallyl butane with the oxidized counterpart, denoted as **(3)**. **Figure** 5a shows the geometry of the simulated ensemble consisting of five molecules. Three oxidized 1,4-diallyl butanes **(3)**, designated as *Mol1*, *Mol2*, and *Mol3*, may undergo polarization due to intermolecular interactions with each other and with the two polar arginines **(2)**, i.e. *Mol4* and *Mol5*.

Again, we evaluate the reference DRG solution, whose result details are reported in Table 1. The almost unitary charge obtained for *Mol1*, *Mol2*, and *Mol3* and the almost null charge obtained for *Mol4* and *Mol5* suggest molecule charge is well-constrained, thereby supporting the validity of the ensemble in validating the proposed method. Figure 5b shows the obtained MEP evaluated at 0.2 nm above the molecule plane. The produced MEP depicts clear polarization in the oxidized 1,4-diallyl butane, especially in *Mol1* and *Mol3*, whereas *Mol2* MEP appears non-polarized. Concerning the arginine, *Mol4* and *Mol5* show clear polarization, which is expected due to the polar nature of **(2)**. The intuition is confirmed by the dipole moment analysis, reported in Figure 5c, which clearly shows polarization variations among the five molecules, especially in the **(3)** species.

Figure 5d shows the MEP resulting from the SCERPA calculation in the SAC solution. SCERPA well assesses the polarization of the oxidized 1,4-diallyl butanes. The potential closely resembles the DRG solution, leading to errors consistent with the neutral calculation ($E_V = 0.917$ V, $\Delta_V = 0.024$ V). Regarding the dipole moment, as shown in Figure 5(e), the SAC solution provides an accurate estimation of the dipole moment for the oxidized 1,4-diallyl butane molecules, especially on the x and y-axis, which remain consistent with the DRG solution. Once more, this result demonstrates SCERPA capability to deliver precise simulation outcomes for molecular ensembles, also when ions are present.

Figure 5f shows the obtained SINT MEP ($E_V = 0.133$ V, $\Delta_V = 0.003$ V). For the SINT calculation, the dipole moment analysis shown in Figure 5g shows excellent agreement with the DRG solution since the dipole distribution resembles Figure 5c, despite obtained with minimal computational overhead.

Finally, Figure 5h reports the result of the DnSCF. The calculation, performed in 781.198 s, well resembles the DRG solution with an error slightly larger than the SINT solution ($E_V = 0.150$ V, $\Delta_V = 0.003$ V). The reason for the error increase have been already discussed in Section 2.2. Figure 5i shows the DnSCF dipole moment analysis. Finally, Figure 5j,k shows the obtained MEP and dipole moment, respectively, for the DSCF calculation. The DSCF calculation was obtained in 2652.39 s, which represents a reduction of 86 % to the pure DFT DRG calculation, with error ($E_V = 0.101$ V, $\Delta_V = 0.002$ V)

Overall, this analysis demonstrates that SCERPA correctly predicts the electrostatic properties of charged molecular ensem-

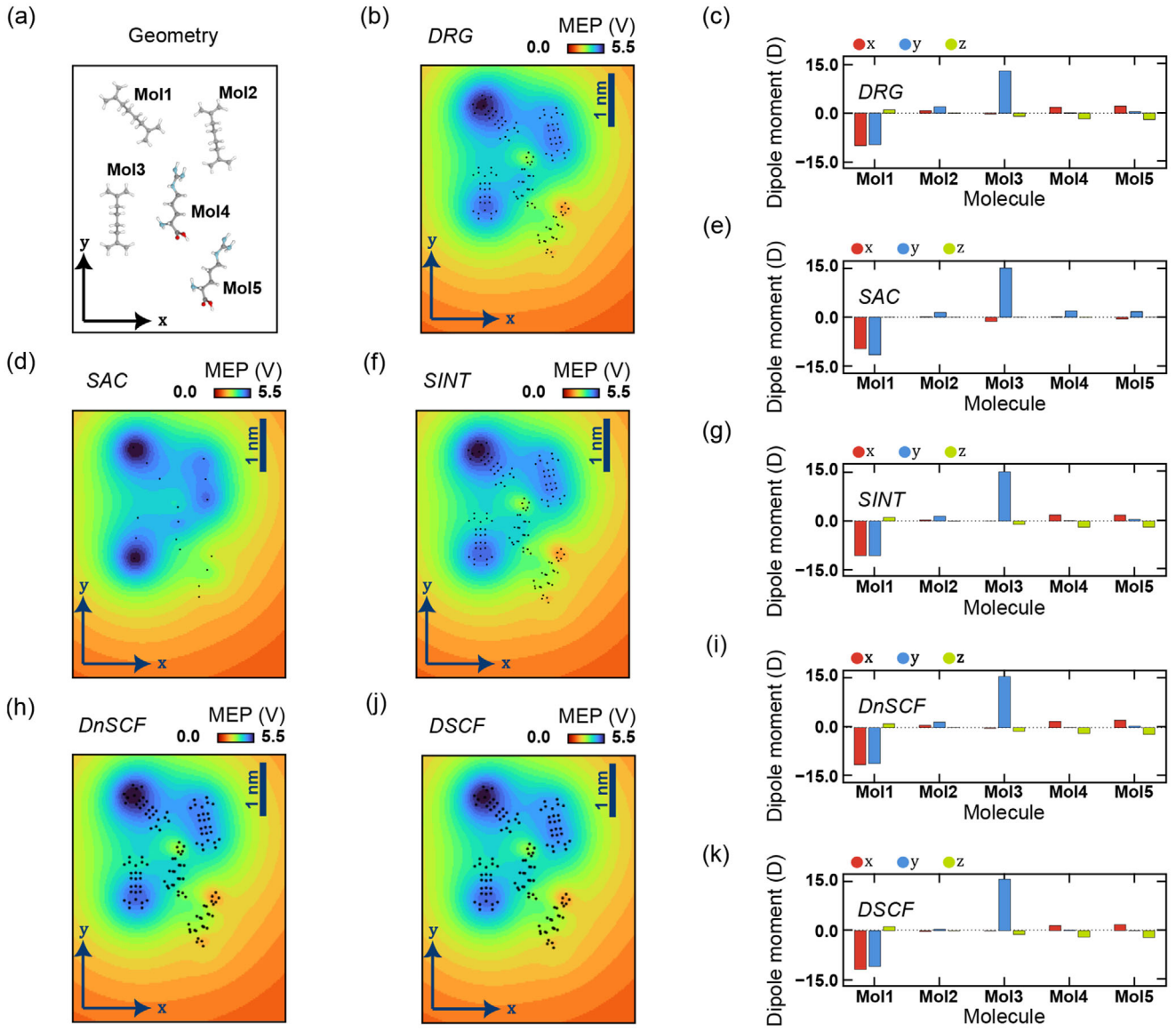

Figure 5. Charged ensemble. a) Geometry of the analyzed cationic ensemble composed of three cations (3) and two polar (2) molecules; b) MEP evaluated from the DRG solution; c) Dipole analysis evaluated from the DRG solution; d) MEP evaluated from the SAC solution; e) Dipole analysis evaluated from the SAC solution; f) MEP evaluated from the SINT solution; g) Dipole analysis evaluated from the SINT solution; h) MEP evaluated from the DnSCF solution; i) Dipole analysis evaluated from the DnSCF solution; j) MEP evaluated from the DSCF solution; k) Dipole analysis evaluated from the DSCF solution.

bles. In particular, Table 1 reports the molecular charges in all the molecules, demonstrating consistent charges and no intermolecular charge transfer. The charge is well-constrained in each molecule.

2.4. Case Studio: Self-Polarizer

Section 2.2 and Section 2.3 validated the method on ensembles composed of a few molecules or ions. This section presents a case study application related to molecular electronics. We propose a MolFCN circuit, i.e., a molecular ensemble named "self-polarizer," ideally composed of infinite molecules with a known a priori charge distribution. This molecular system naturally encodes MolFCN information without external driver stimuli. It can be exploited for validation purposes in both simulation and, in principle, experimental measurements when forcing information encoding with external fields is difficult. For this work, this molecular ensemble demonstrates the effectiveness and precision of the SCERPA algorithm in evaluating the information encoding and propagation in the MolFCN technology on an ensemble composed of many molecules.

Figure 6a shows the structure of the self-polarizer. It is composed of infinite replications of the here-named "basic element." Figure 6b shows the basic element, which is composed of four molecules aligned with fixed intermolecular distance, arranged

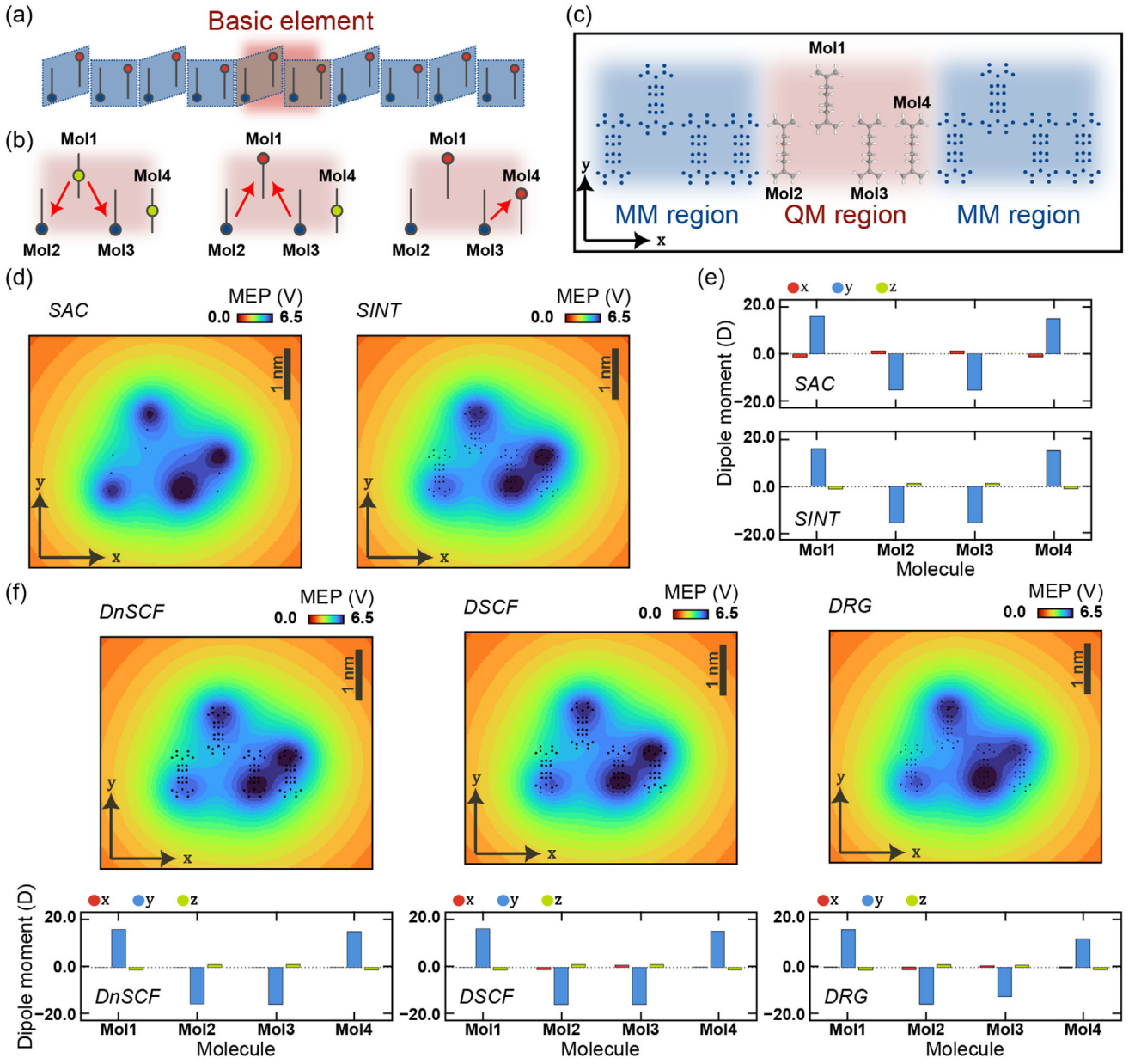

Figure 6. Self-polarizer. a) Geometry of the self-polarizer, composed of the (infinite) replication of the basic element; b) Working principle of the self-polarizer; c) Structure of the self-polarizer simulated with QM/MM; d) MEP evaluated from the SAC and SINT solutions; e) Dipole analysis evaluated for the SAC and SINT solutions; f) MEP and dipole analysis evaluated from the DnSCF, DSCF, and DRG solutions.

so that *Mol1* is slightly shifted up by a quantity Δy on the y-axis. The *Mol1* shift is responsible for the self-polarization of all the molecules. We here refer to molecules as "up-polarized" when the molecule dipole has a positive y-component and "down-polarized" when the molecule dipole has a negative y-component. If the molecules are all oxidized, *Mol1* produces a radial electric field that down-polarizes *Mol2* and *Mol3*. In turn, *Mol2* and *Mol3* up-polarize *Mol1*. Finally, *Mol4* is up-polarized due to the intermolecular interaction with *Mol3*. Notice that the system is periodical; therefore, *Mol4* is up-polarized due to the influence of *Mol2*, considered adjacent. Figure 6b shows the expected charge distribution of the self-polarizer.

Studying this system with DFT requires simulating an ensemble with many molecules and constrained charge distribution. The DFT tool used in this work, i.e. ORCA, does not allow exploiting periodic boundary conditions nor charge constraints in the single molecules. Therefore, we make use of QM/MM technique to simulate the self-polarizer. Figure 6c shows the structure of the QM/MM system composed of molecules **(3)** for the simulation of the self-polarizer. Section QM/MM Calculation describes the details of the model.

The system is first analyzed with SCERPA to validate the device and to determine the charges requested for the MM region. The simulated circuit consists of 11 replications of the basic element.

Table 2. ab initio results for the self-polarizer.

Simulation	QM1	QM2	QM3	QM4	Time (s)	Energy (Eh)
Mol1: (3), Mol2: (3), Mol3 (3), Mol4 (3)						
DRG	1.0083	0.9933	1.1930	0.8054	15840.929	−1559.17719894
SAC	1.0000	1.0000	1.0000	1.0000	0.782411	—
SINT	1.0000	1.0000	1.0000	1.0000	—	—
DnSCF	1.0007	0.9990	1.0008	0.9995	468.93	—
DSCF	1.0023	0.9984	1.0009	0.9984	4500.894	−1559.17580761

Figure 6d shows the result obtained with the SCERPA calculation, either SINT or SAC. The figure shows only the molecules corresponding to the QM region for consistency with the following simulations. The MEP clearly shows the polarization of all the molecules in perfect accordance with the expectations: *Mol1* and *Mol4* up-polarized, whereas *Mol2* and *Mol3* are down-polarized. The result is also confirmed by the dipole analysis reported in Figure 6e. Figure 6f reports the ab initio based calculation: DnSCF, DSCF and DRG. All the ab initio calculations are consistent with the expectations and the results obtained with SCERPA.

Table 2 reports the simulation data. The analysis of the self-polarizer shows, for the first time, a limitation of the DRG calculation. Indeed, assuming the work aims to simulate four oxidized molecules, the charge should be unitary in all the molecules. In the DRG calculation, the charge moves among molecules, e.g., $QM3 = 1.1939 a.u.$ and $QM4 = 0.8054 a.u.$, clearly demonstrating charge transfer between *Mol4* and *Mol3*. All the other calculations show no charge transfer between the molecules. In particular, it is worth highlighting the DSCF calculation result. This calculation is obtained by minimizing the energy of the system with a given initial guess, determined by SCERPA. Therefore, DSCF is a correct solution to the quantum problem. This result shows that the molecular ensembles allow at least two possible solutions. The DRG solution is the most stable, whereas the DSCF is a local minimum preserving charge constraint on each molecule. This analysis demonstrates that SCERPA can find solutions alternative to the normal ab initio calculation, favoring the charge-constrained solutions. For completeness, we also highlight the time calculation improvement by comparing the calculation time with the DRG calculation (15840.1772 s). The SAC/SINT, DnSCF and DSCF permits saving 99.99%, 97.04%, and 71.59% of the computational time, respectively.

Table 3 reports the errors between the calculations. For this analysis, the DRG calculation error is slightly higher than those reported in Section 4 and Section 5. It is important to consider that the SCERPA simulation led to a solution different from the DRG calculation, yet ab initio valid. The DSCF calculation convergence guarantees the validity of SCERPA in providing a possible solution to the quantum problem, which is, in this case, a local energy minimum. Therefore, the quality of the SCERPA-based calculation is performed by comparing the calculations with the DSCF. For the SINT and DnSCF calculation, the maximum error is reduced by one order of magnitude, demonstrating the validity and precision of the SCERPA calculation.

Table 3. Calculation errors for the self-polarizer.

	DRG		DSCF	
Simulation	E_V (V)	Δ_V (V)	E_V (V)	Δ_V (V)
DRG	—	—	0.455647	0.005138
SAC	0.644550	0.006464	0.678883	0.001326
SINT	0.473326	0.006801	0.070755	0.001664
DnSCF	0.464411	0.006528	0.075794	0.001390
DSCF	0.455647	0.005138	—	—

3. Conclusion

This work proposes and validates efficient self-consistent procedures for evaluating electrostatic properties in molecular ensembles. The method leverages intermolecular interaction models which consider molecules as black boxes characterized by first principle calculation. The results show that self-consistent procedure, specifically the SCERPA tool, guarantees ab initio precision with minimal calculation time. When precision is crucial, the self-consistent procedure can be used as initial guess estimator for facilitating ab initio calculation, finally reducing computation time. In addition, the results show that this approach can favor the ab initio calculation to converge to possible energy minima with constrained charge on single-molecules, overcoming one of the problems of state-of-the-art DFT packages, which often makes challenging charge constraining in molecular fragments. As a case studio, this work focuses on Molecular Field-Coupled Nanocomputing, which makes use of molecular ensembles for realizing electronics devices based on intermolecular electrostatic properties. The validity of the SCERPA results pose an important milestone on this technology, which still misses a validated working prototype. Indeed, provided correct ab initio characterisation of molecules, SCERPA results can be used to demonstrate the feasibility of information propagation and elaboration at the molecular scale, without the need for electric current. Experimental validation is still fundamental, yet the obtained results bode well for the assessment of the technology, and motivate further experimental analyses for physically detecting charges in molecular FCN circuits. In addition, to validate the molecular FCN paradigm, this work proposes a molecular system which shows fixed polarization and can provide fixed input logic to molecular FCN devices in realizing majority-voter-based AND/OR gates. In addition, the proposed self-polarizer facilitates the experimental characterization of molecular devices without the need for external electrodes that polarize single-molecules and makes the characterization difficult with scanning probe microscopy techniques.

4. Experimental Section

Ab Initio Characterization: Each molecule of the *N* species used in this work undergoes ab initio characterization. The procedure includes geometry optimization to determine the exact position of the atoms and a DFT single-point calculation to evaluate the equilibrium electrostatic properties represented by the atomic charges $\{q_i\}$. Molecules have been characterized with the ORCA package version 4.2.0.[17,18] The electric field used

to polarize the molecule in the ab initio characterization has been inserted using point charges, following the procedure reported in [14]. All the calculations are performed with the Unrestricted Kohn–Shann (UKS) method using CAM-B3LYP functional with def2-TZVP basis set.[19] The calculation utilizes the atom-pairwise dispersion correction with the Becke–Johnson damping scheme (D3BJ).[20,21] Atomic charges are obtained by fitting the MEP using the CHELPG (CHarges from ELectrostatic Potentials using a Grid-based method) method.[22]

Figures of Merit Extraction: Each molecule is modeled with the Aggregated Charge (AC) distribution within the MoSQuiTo methodology to permits the SCERPA simulation.[6,12,13,23] In SCERPA, the electrostatic interaction between a Molecule Under Test (MUT) and the rest of the circuit molecules, is evaluated through the concept of *Input Voltage* (V_in) of the MUT, which is obtained by integrating the electric field generated by all the molecules of the circuit between the so-called *active dots* of the MUT.[6] The active dots are generally the two dots (i.e. two ACs) of the molecule providing logic encoding in molecular FCN. In general, the input voltage of the MUT depends on the other molecule aggregated charges and corresponding positions. The value of the charges is generally obtained through DFT characterization. Instead, the AC position is generally associated with an atom considered effective by the designer without any optimization related to the final electric field generated by the molecule.[13] In this work, we optimize the AC position to improve the electrostatic description of the molecule. With the optimized position, we expect the MEP evaluated with electrostatic equations based on the AC values to re-establish the MEP evaluated with DFT precision. This optimization improves the interaction evaluation in the SCERPA algorithm. The position of the generic j-th AC, denoted as $\mathbf{R}_j$ and grouping a set of atomic charges denoted as ACG-j is evaluated as:

$$\mathbf{R}_j = \frac{\sum_{\text{ACG-}j} \mathbf{r}_i |q_i|}{\sum_{\text{ACG-}j} |q_i|} \tag{1}$$

Where $\{\mathbf{r}_i\}$ is the position of the molecule atom and $\{q_i\}$ is the associated CHELPG charge. Finally, the aggregate charge distribution is evaluated on a set of different input voltages ($\mathbf{V}_{char}$) using DFT calculation, thus composing the VACT. The characterization procedure also keeps track of the orbitals, thus storing the `GBW` file (used for the DnSCF and SCF calculations, see Section Orbital Estimation Procedure) and the CHELPG charges (used for the SINT calculation) in each input voltage.

Molecule Electrostatic Potential Evaluation: The evaluation of the MEP depends on the solution type. In particular, SAC and SINT solutions lead to electrostatic charges. In the SAC solution, the MEP is evaluated by exploiting the ACs obtained from the SCERPA calculation. In a set of N molecules, each identified as Mol_i and described with a set of M_i aggregated charges $\{Q_{i,j}\}$ in positions $\{R_{i,j}\}$:

$$MEP(\mathbf{r}) = \sum_{i=1}^{N} \sum_{j=1}^{M_i} \frac{1}{4\pi\varepsilon_0} \frac{Q_{i,j}}{R_{i,j}} \tag{2}$$

Instead, the SINT solution exploits all the atomic CHELPG charges. The MEP is evaluated by considering each molecule Mol_i as a group of m_i atomic charges with values $\{q_{i,j}\}$ and positions $\{r_{i,j}\}$:

$$MEP(\mathbf{r}) = \sum_{i=1}^{N} \sum_{j=1}^{m_i} \frac{1}{4\pi\varepsilon_0} \frac{q_{i,j}}{r_{i,j}} \tag{3}$$

Concerning the other solutions, SINT, DnSCF and DSCF, the MEP is directly evaluated with ORCA. In general, in the quantum chemistry description, all the molecules are treated as a single group of H atoms. Each $k-th$ atom is described by the position of the nucleus R_k and the nuclear charge Z_k. Therefore, in the DFT perspective, the electrons are described with the electron density $\rho(\mathbf{r})$. The potential can be evaluated according to [24]:

$$MEP(\mathbf{r}) = \sum_{k=1}^{H} \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}|} - \iiint \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' \tag{4}$$

Circuit Simulation: The molecular ensemble is defined as a group of molecules with specific positions, rotations, and chemical species. The layout is analyzed by the SCERPA tool, which substitutes each molecule with the aggregated charge model comprised of charges in position $\{\mathbf{R}_j\}$ evaluated with (1). Among the ACs, the active dots (here denoted as Q_1 and Q_2) are used to define the input voltage of the molecule. For each molecule, denoted as the $i-th$ molecule, the SCERPA tool evaluates the potential generated by all the molecules of the circuit except the $i-th$ molecule using Equation (2). SCERPA evaluates the so-called input voltage of the Molecule as:

$$V_{in,i} = MEP(\mathbf{R}_{i,1}) - MEP(\mathbf{R}_{i,2}) \tag{5}$$

Where, $\mathbf{R}_{i,1}$ and $\mathbf{R}_{i,2}$ are the positions of the two active dots of the $i-th$ molecule. Therefore, the input voltage of the molecule is used to compute the aggregated charges of the $i-th$ molecule by exploiting the Vin-AC Transcharacteristics (VACT), which links the input voltage (V_{in}) to the DFT-derived AC value.[13] The procedure is repeated to derive the input voltage of all the molecules of the ensemble.

It is important to notice that the input voltage of the $i-th$ molecule, described with Equation (5), depends on the AC values of other molecules, which in turn depend on the AC (i.e., the input voltage) of the $i-th$ molecule.[12] Therefore, the solution of the equation is obtained through a self-consistent procedure, i.e., the input voltage evaluation is iterated until the convergence is reached. Finally, the SCERPA tool outputs the input voltage and the charge distribution of each molecule (SAC solution), fulfilling our aim to obtain the electrostatic properties of the molecular ensemble.

Orbital Estimation Procedure: The procedure constructs the nonrelativistic orbital initial guess of the DnSCF/DSCF solutions by exploiting the SCERPA results. In particular, for each molecule of the ensemble, the characterization procedure saves the ORCA `GBW` file, containing the orbitals of the molecule, with the associated input voltage V_{char}. In the characterization procedure, we create a database composed of GBW files, each identified with a specific input voltage V_{char}. Then, the ensemble is analyzed with SCERPA to obtain the input voltage of each molecule V_{in}, i.e., the SAC solution. The initial guess of the DnSCF/DSCF calculation is calculated with the `orca_mergefrag` tool, merging all the `GBW` files of the molecules in the layout. The orbital files (GBW) of each molecule are chosen as the one minimizing $|V_{char} - V_{in}|$.

QM/MM Calculation: The QM region is composed of four 1,4-diallyl butane molecules. The total charge of the QM region is set equal to four to emulate the oxidation of all the molecules. No charge constraint is applied on molecules, since this feature is unavailable in ORCA. The MM region comprises 40 molecules representing the self-polarizer periodic structure (10 replications of the basic element, five on each side). Ten replications were considered large enough to assume the QM region behaves like an infinite number of molecules. The intermolecular distance is fixed to $d_x = 0.75$ nm and the *Mol1* is up-shifted by $\Delta y = 0.45$ nm. The intermolecular distances have been chosen empirically to favor the electrostatic interaction, thus favor the polarization, and prevent charge transfer between molecules in the DRG calculation. All the molecules in the MM region are approximated with point charges evaluated with a preliminary SCERPA calculation (SAC). Once each molecule input voltage is determined with SCERPA, the charges in the MM region are set equal to the charges of the calculation performed in the Characterisation stage with the closest V_{in}.

Acknowledgements

The authors thank Dr. Giuliana Beretta for her valuable support during the initial stages of this work.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

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

Keywords

ab initio, density-functional theory, molecular electronics, nanoelectronics, organic electronics

Received: April 29, 2025
Revised: June 3, 2025
Published online:

[1] H. Zhang, Y. Zhu, P. Duan, M. Shiri, S. C. Yelishala, S. Shen, Z. Song, C. Jia, X. Guo, L. Cui, K. Wang, *Appl. Phys. Rev.* **2024**, *11*, 041312.
[2] F. Mo, Y. Ardesi, M. R. Roch, M. Graziano, G. Piccinini, *IEEE Sens. J.* **2022**, *22*, 19152.
[3] X. Wang, A. Alajmi, Z. Wei, M. Alzanbaqi, N. Wei, C. Lambert, A. Ismael, *ACS Appl. Mater. Interfaces* **2024**, *16*, 66290.
[4] C. S. Lent, B. Isaksen, M. Lieberman, *J. Am. Chem. Soc.* **2003**, *125*, 1056.
[5] G. Beretta, Y. Ardesi, M. Graziano, G. Piccinini, *IEEE Trans. Nanotechnol.* **2022**, *21*, 52.
[6] Y. Ardesi, G. Turvani, M. Graziano, G. Piccinini, *IEEE Trans. Very Large Scale Integr. VLSI Syst.* **2021**, *29*, 558.
[7] J. Timler, C. S. Lent, *J. Appl. Phys.* **2002**, *91*, 823.
[8] F. Ravera, G. Beretta, Y. Ardesi, M. Krzywiecki, M. Graziano, G. Piccinini, in IEEE, editor, *2023 IEEE Nanotechnology Materials and Devices Conference (NMDC)*, IEEE, Paestum (ITALY), **2023**, pp. 212–213.
[9] F. Mohn, L. Gross, N. Moll, G. Meyer, *Nat. Nanotechnol.* **2012**, *7*, 227.
[10] B. Mallada, M. Ondráček, M. Lamanec, A. Gallardo, A. Jiménez-Martín, B. de la Torre, P. Hobza, P. Jelínek, *Nat. Commun.* **2023**, *14*, 4954.
[11] Y. Ardesi, U. Garlando, F. Riente, G. Beretta, G. Piccinini, M. Graziano, *ACM J. Emerging Technol. Computing Systems* **2022**, *19*, 1.
[12] Y. Ardesi, R. Wang, G. Turvani, G. Piccinini, M. Graziano, *IEEE Trans. Comput.-aided Des. Integr. Circuits Syst.* **2020**, *39*, 2749.
[13] Y. Ardesi, A. Pulimeno, M. Graziano, F. Riente, G. Piccinini, *J. low power electron. appl.* **2018**, *8*, 24.
[14] Y. Ardesi, M. Graziano, G. Piccinini, *J. low power electron. appl.* **2022**, *12*, 1.
[15] Y. Ardesi, A. Gaeta, G. Beretta, G. Piccinini, M. Graziano, *JICS. J. Integr. Circuits Syst.* **2021**, *16*, 1.
[16] E. Rahimi, J. R. Reimers, *Phys. Chem. Chem. Phys.* **2018**, *20*, 17881.
[17] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73.
[18] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1327.
[19] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
[20] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
[21] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
[22] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361.
[23] G. Beretta, Y. Ardesi, G. Piccinini, M. Graziano, *vlsi-nanocomputing/scerpa: Scerpa v4.0.1*, **2022**, https://doi.org/10.5281/zenodo.7457038.
[24] C. H. Suresh, G. S. Remya, P. K. Anjalikrishna, *WIREs Computational Molecular Science* **2022**, *12*, e1601.