

# The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations

Zhendong Li, Junhao Li, Nikesh S. Dattani, C. J. Umrigar, and Garnet Kin-Lic Chan

Citation: *J. Chem. Phys.* **150**, 024302 (2019); doi: 10.1063/1.5063376

View online: <https://doi.org/10.1063/1.5063376>

View Table of Contents: <http://aip.scitation.org/toc/jcp/150/2>

Published by the [American Institute of Physics](#)

---

---

**PHYSICS TODAY**

WHITEPAPERS

## ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY  
 **MASTERBOND**  
ADHESIVES | SEALANTS | COATINGS

# The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations

Cite as: *J. Chem. Phys.* **150**, 024302 (2019); doi: [10.1063/1.5063376](https://doi.org/10.1063/1.5063376)

Submitted: 26 September 2018 • Accepted: 10 December 2018 •

Published Online: 8 January 2019



View Online



Export Citation



CrossMark

Zhendong Li,<sup>1</sup> Junhao Li,<sup>2</sup> Nikesh S. Dattani,<sup>3,4</sup> C. J. Umrigar,<sup>2</sup>  and Garnet Kin-Lic Chan<sup>1</sup>

## AFFILIATIONS

<sup>1</sup>Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

<sup>2</sup>Department of Physics, Laboratory of Atomic and Solid-State Physics, Cornell University, Ithaca, New York 14853, USA

<sup>3</sup>National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

<sup>4</sup>Hertford College, Oxford University, Oxford OX1 3BW, United Kingdom

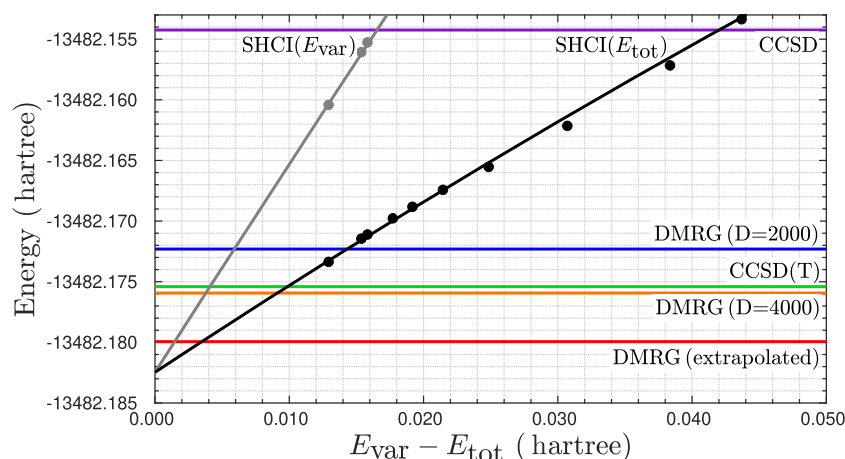
## ABSTRACT

We report that a recent active space model of the nitrogenase FeMo cofactor, proposed in the context of simulations on quantum computers, is not representative of the electronic structure of the FeMo cofactor ground-state. A more representative model does not affect much certain resource estimates for a quantum computer such as the cost of a Trotter step, while strongly affecting others such as the cost of adiabatic state preparation. Thus, conclusions should not be drawn from the complexity of quantum or classical simulations of the electronic structure of this system in this active space. We provide a different model active space for the FeMo cofactor that contains the basic open-shell qualitative character, which may be useful as a benchmark system for making resource estimates for classical and quantum computers.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5063376>

The process of nitrogen fixation, namely, that of converting atmospheric dinitrogen to a reduced form, such as ammonia, which can then be metabolized by biological species, is essential to life on this planet.<sup>1–4</sup> The industrial Haber–Bosch process to produce ammonia from the exothermic reaction  $\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3$  requires a careful balance of high temperatures and high pressures to achieve efficient catalysis. By contrast, natural bacteria and archaea carry out nitrogen fixation under ambient conditions through nitrogenases. At the molecular level, the nitrogenase enzyme, an agglomeration of a homodimeric Fe protein and the heterotetrameric MoFe protein (in the case of molybdenum nitrogenase, which is the most common form of nitrogenase), catalyzes the nitrogen bond-breaking process via a family of 3 metallic cofactors: the  $[\text{Fe}_4\text{S}_4]$  iron cubane, the  $[\text{Fe}_8\text{S}_7]$  P cluster, and the  $[\text{MoFe}_7\text{S}_9\text{C}]$  FeMo cofactor (FeMoco), with FeMoco serving as the site of nitrogen reduction.<sup>1</sup> The contrast between the conditions of biological nitrogen fixation and the Haber–Bosch process is an enduring source of fascination for chemists.

In the search to unravel the secrets of biological nitrogen fixation, the first stage is to understand the structure of the enzyme itself. After many decades, we now possess atomic scale resolution structures of nitrogenase, including all cofactors.<sup>5,6</sup> However, the electronic structure of the cofactors, and, in particular, the large P cluster and FeMo cofactor, remains poorly understood. This is due to the complexity of tackling the multiple transition metal ions with their multiple charge states and complicated spin-couplings. Even though we believe the qualitative electronic structure to be captured using only the valence active space of the metals and bridging S ligands which provides a great reduction of the problem size (to, for example, 103 electrons in 71 orbitals in the case of FeMoco, considering the Fe 3d, S 3p, Mo 4d, and the interstitial C 2s2p for the  $[\text{MoFe}_7\text{S}_9\text{C}]$  core, assuming a total charge of  $-1$ ), no satisfactory classical many-electron simulation within this valence active space has yet been performed. Because of the need for tangible objectives for quantum simulations of the electronic structure,



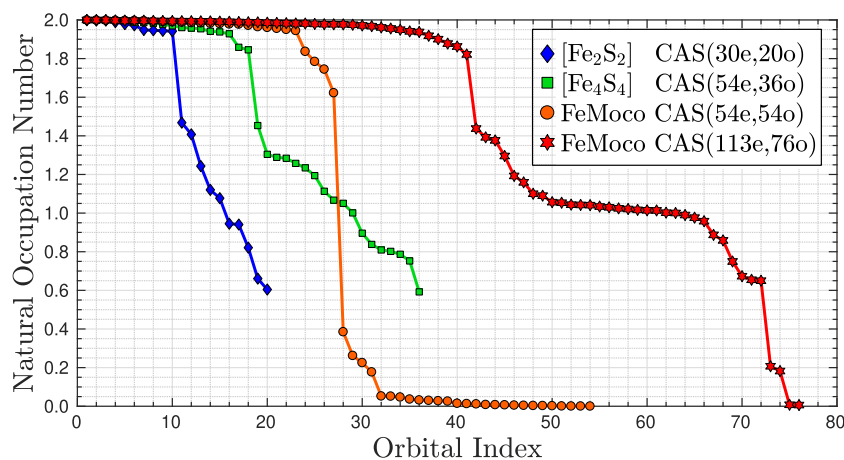
**FIG. 1.** SHCI variational and total energies for progressively decreasing cutoffs (circles) along with weighted quadratic fits (curves through the circles) of the (54e, 54o) model of FeMoco in Ref. 7, variational DMRG energies at bond dimensions of  $D = 2000$  and  $D = 4000$  as well as the extrapolated DMRG energy (obtained from a three-point extrapolation of the  $D = 4000, 3500, 3000$  data in a reverse schedule), and CCSD and CCSD(T) energies. The weighted quadratic fit for the SHCI calculations used weights of  $1/(E_{\text{var}} - E_{\text{tot}})^2$ . All calculations are for the  $S = 0$  state.

these metallic cofactors have thus been suggested as an interesting target for future quantum simulators.<sup>7,8</sup> Reference 7 provides a pedagogical discussion of the chemical questions that must be considered when elucidating a complex reaction such as nitrogen fixation, as well as concrete resource estimates resulting from 54 electron in 54 orbital (54e, 54o) and 65 electron in 57 orbital (65e, 57o) models of the FeMoco cluster.

Although the focus of Ref. 7 was the resource estimates for quantum computers, it is natural to ask whether a calculation of the electronic structure of FeMoco for the (54e, 54o) model of Ref. 7 is feasible on a classical computer. For this reason, we report that this active space does not actually contain the representative features of the electronic structure of the low-energy states of FeMoco that make its classical (and potentially quantum) simulation difficult. Consequently, if taken out of context, it provides a misleading characterization of the classical and quantum complexity of obtaining the low-energy states. In fact, as shown in Fig. 1 for the (54e, 54o) model of Ref. 7, we can obtain accurate energies of the lowest  $S = 0$  state studied in Ref. 7 using standard classical algorithms such as coupled cluster theory,<sup>9,10</sup> variational density matrix

renormalization group (DMRG),<sup>11–16</sup> and the semistochastic<sup>17,18</sup> heatbath configuration interaction (SHCI) method.<sup>19</sup> The extrapolated DMRG variational energy and the extrapolated SHCI energy agree to within  $2 \text{ mE}_h$  or about  $0.25 \text{ mE}_h$  per metal center. Note that higher accuracy calculations are feasible (for example, by including higher-order excitation operators in coupled cluster, more variational determinants in selected CI, or larger bond dimensions in DMRG).

As we have mentioned, the reason for the simplicity of the classical simulations is not the intrinsic electronic structure of the FeMo cofactor but is due to the choice of the active space in Ref. 7. In FeMoco, the Fe and Mo ions are expected to be in the (II), (III), or (IV) formal oxidation states,<sup>20–25</sup> which leads to approximately 35 open shells (singly filled orbitals) depending on the charge state of the cluster. The prevalence of Fe(II) and Fe(III) oxidation states is supported experimentally by the Fe Mössbauer spectrum<sup>20–22</sup> and can be seen in direct theoretical calculations of smaller pieces of the FeMo cofactor, such as the  $[\text{Fe}_2\text{S}_2]$  or  $[\text{Fe}_4\text{S}_4]$  clusters.<sup>26,27</sup> However, the one-body density matrix in the FeMo cofactor model of Ref. 7 has no open-shell orbitals (orbitals with an occupation number close to 1), as seen from the eigenvalues of the one-body density matrix



**FIG. 2.** Natural occupations obtained with DMRG for the  $S = 0$  state of a  $[\text{Fe}_2\text{S}_2]$  complex with CAS(30e, 20o) and  $D = 8000$ , the  $S = 0$  state of a  $[\text{Fe}_4\text{S}_4]$  complex with CAS(54e, 36o) and  $D = 4000$ , the  $S = 0$  state of FeMoco with CAS(54e, 54o) reported in Ref. 7 and  $D = 2000$ , and the  $S = 3/2$  state of FeMoco with CAS(113e, 76o) constructed in this work and  $D = 2000$ . In contrast to the other models, the CAS(54e, 54o) ground-state has no open shells.

(Fig. 2). A related point is that the coefficient of the dominant (natural orbital) determinant in SHCI is very large (0.67), indicating that the wavefunction has mainly single or few determinantal characters, which is not possible for a low-spin system with many open shells. A large determinant weight has also been observed in Ref. 28 (in fact, they observed an even larger determinant weight, probably because of using a smaller number of variational determinants). As shown in Fig. 1, the coupled-cluster singles and doubles with perturbative triples [CCSD(T)] energy is also within 5–7  $mE_h$ , or less than 1  $mE_h$  per metal center, of our extrapolated DMRG and SHCI energies, providing further evidence for the single reference nature of this problem.

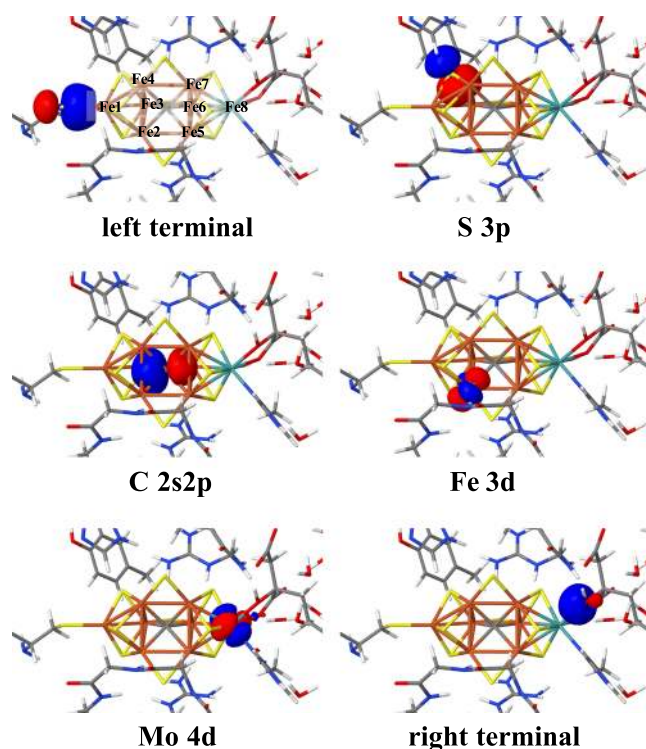
While the electronic structure of the low-energy states within the active space of Ref. 7 is qualitatively incorrect, the effect on estimates of resources for quantum computers depends on the computational step under consideration. For example, the cost of a Trotter step (the primary focus in Ref. 7) depends primarily on the number of Hamiltonian matrix elements and their magnitudes. Thus the estimates in Ref. 7 are probably reasonable even if one were to use a different, more appropriate, set of active orbitals, assuming the same overall active space size. However, the character of the ground-state greatly impacts the efficiency of adiabatic state preparation. With respect to FeMoco, this was left as an open problem in Ref. 7 but was considered, for example, in Ref. 28. In this case, we expect a significant increase in the cost of adiabatic state preparation as one can no longer simply use a single determinant state.

It seems desirable to have a more qualitatively reasonable active space for future studies. For this purpose, we attach a valence active space Hamiltonian<sup>29</sup> of the FeMo cofactor constructed from all Fe 3d, S 3p, Mo 4d, and C 2s2p orbitals in the [MoFe<sub>7</sub>S<sub>9</sub>C] core, as well as some bonding ligand orbitals (for details, see Table I). The active orbitals were obtained by

**TABLE I.** Composition of the active space with 76 orbitals for FeMoco. For the labels of Fe atoms, refer to the labels in Fig. 3.

Group	Orbital	Orbital index
Left cubane		
1	Left terminal	1, 2
2	Fe1 3d	3, 4, 5, 6, 7
3	S 3p	8, 9, 10, 11, 12, 13, 14, 15, 16
4	Fe2 3d	17, 18, 19, 20, 21
5	Fe3 3d	22, 23, 24, 25, 26
6	Fe4 3d	27, 28, 29, 30, 31
Central part		
7	S 3p, C 2s2p	32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44
Right cubane		
8	Fe5 3d	45, 46, 47, 48, 49
9	Fe6 3d	50, 51, 52, 53, 54
10	Fe7 3d	55, 56, 57, 58, 59
11	S 3p	60, 61, 62, 63, 64, 65, 66, 67, 68
12	Mo8 4d	69, 70, 71, 72, 73
13	Right terminal	74, 75, 76

first performing high-spin unrestricted Kohn-Sham calculations with the spin-free exact two-component (sf-X2C) Hamiltonian,<sup>30–32</sup> the B3LYP functional,<sup>33–35</sup> and the TZP-DKH<sup>36</sup> basis for Fe, S, and Mo and the def2-SVP basis<sup>37</sup> for the other atoms (C, H, O, and N) using a structure<sup>38</sup> in Ref. 25 [with the rest of the protein environment mimicked by the conductor-like screening model (COSMO)<sup>39</sup> with  $\epsilon = 4.0$ ] and then split-localizing the unrestricted natural orbitals. This results in an active space model with 113 electrons in 76 orbitals. The detailed composition is shown in Table I, and some selected localized orbitals are shown in Fig. 3. The dimension of the full configuration interaction (FCI) space is on the order of  $10^{35}$  (specifically,  $\binom{76}{58} \times \binom{76}{55} = 3.6 \times 10^{35}$ ) for the spin  $S = 3/2$  ground state<sup>20,21</sup> in this FeMoco active space. We have performed preliminary DMRG calculations to check the qualitative features of the active space. As shown in Fig. 2, the natural occupation numbers obtained with a DMRG solution ( $D = 2000$ ) for  $S = 3/2$  show a large number of singly occupied orbitals, which demonstrates that this active space captures the open-shell character of FeMoco in contrast to the previous model.<sup>7</sup> While we emphasize that a detailed and chemically meaningful study on FeMoco should consider many other factors, such as the convergence of the environment representation,<sup>40,41</sup> different protonations,<sup>42</sup> etc., we conclude that the active space Hamiltonian we provide contains at least a



**FIG. 3.** Illustration of some selected active orbitals for FeMoco in the active space model CAS(113e, 76o) constructed in this work. Labels for the Fe atoms in the top-left subfigure explain the labels given in Table I.

qualitative model of the open-shell character and low-energy states of the cofactor. We hope that this will be useful in future quantum or classical studies of the complexity of the FeMo cofactor electronic structure.

Z.L. and G.K.-L.C. were supported by the U.S. National Science Foundation via No. CHE-1665333. J.L. and C.J.U. were supported by AFOSR, Grant No. FA9550-18-1-0095. One of the SHCI computations was performed on the Bridges computer at the Pittsburgh Supercomputing Center supported by NSF, Grant No. ACI-1445606, as part of the XSEDE program supported by NSF, Grant No. ACI-1548562.

## REFERENCES

- J. B. Howard and D. C. Rees, *Chem. Rev.* **96**, 2965 (1996).
- H. Beinert, R. H. Holm, and E. Münck, *Science* **277**, 653 (1997).
- D. C. Rees and J. B. Howard, *Science* **300**, 929 (2003).
- B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean, and L. C. Seefeldt, *Chem. Rev.* **114**, 4041 (2014).
- T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. Andrade, E. Schleicher, S. Weber, D. C. Rees, and O. Einsle, *Science* **334**, 940 (2011).
- K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. Hu, M. W. Ribbe, F. Neese, U. Bergmann, and S. DeBeer, *Science* **334**, 974 (2011).
- M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Proc. Natl. Acad. Sci. U. S. A.* **114**, 7555 (2017).
- M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush, and G. K. Chan, preprint [arXiv:1808.02625](https://arxiv.org/abs/1808.02625) (2018).
- I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, 2009).
- R. J. Bartlett and M. Musial, *Rev. Mod. Phys.* **79**, 291 (2007).
- S. R. White and R. L. Martin, *J. Chem. Phys.* **110**, 4127 (1999).
- G. K.-L. Chan and M. Head-Gordon, *J. Chem. Phys.* **116**, 4462 (2002).
- Ó. Legeza and J. Sólyom, *Phys. Rev. B* **68**, 195116 (2003).
- G. K.-L. Chan and S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 (2011).
- S. Sharma and G. K.-L. Chan, *J. Chem. Phys.* **136**, 124121 (2012).
- U. Schollwöck, *Rev. Mod. Phys.* **77**, 259 (2005).
- S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, *J. Chem. Theory Comput.* **13**, 1595 (2017).
- J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar, *J. Chem. Phys.* **149**, 214110 (2018); e-print [arXiv:1809.04600](https://arxiv.org/abs/1809.04600).
- A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- E. Münck, H. Rhodes, W. Orme-Johnson, L. Davis, W. Brill, and V. Shah, *Biochim. Biophys. Acta, Protein Struct.* **400**, 32 (1975).
- R. Zimmermann, E. Münck, W. J. Brill, V. K. Shah, M. T. Henzl, J. Rawlings, and W. H. Orme-Johnson, *Biochim. Biophys. Acta, Protein Struct.* **537**, 185 (1978).
- S. J. Yoo, H. C. Angove, V. Papaefthymiou, B. K. Burgess, and E. Münck, *J. Am. Chem. Soc.* **122**, 4926 (2000).
- R. Björnsson, F. A. Lima, T. Spatzal, T. Weyhermüller, P. Glatzel, E. Bill, O. Einsle, F. Neese, and S. DeBeer, *Chem. Sci.* **5**, 3096 (2014).
- T. Spatzal, J. Schlesier, E.-M. Burger, D. Sippel, L. Zhang, S. L. Andrade, D. C. Rees, and O. Einsle, *Nat. Commun.* **7**, 10902 (2016).
- R. Björnsson, F. Neese, and S. DeBeer, *Inorg. Chem.* **56**, 1470 (2017).
- S. Sharma, K. Sivalingam, F. Neese, and G. K.-L. Chan, *Nat. Chem.* **6**, 927 (2014).
- Z. Li and G. K.-L. Chan, *J. Chem. Theory Comput.* **13**, 2681 (2017).
- N. M. Tubman, C. Mejuto-Zaera, J. M. Epstein, D. Hait, D. S. Levine, W. Huggins, Z. Jiang, J. R. McClean, R. Babbush, M. Head-Gordon, and K. B. Whaley, e-print [arXiv:1809.05523](https://arxiv.org/abs/1809.05523).
- See <https://github.com/zhendongli2008/Active-space-model-for-FeMoco> for information about the new active space Hamiltonian.
- W. Liu, *Mol. Phys.* **108**, 1679 (2010).
- T. Saue, *ChemPhysChem* **12**, 3077 (2011).
- D. Peng and M. Reiher, *Theor. Chem. Acc.* **131**, 1081 (2012).
- A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- P. Stephens, F. Devlin, C. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- F. Jorge, N. A. Canal, G. Camiletti, and S. Machado, *J. Chem. Phys.* **130**, 064108 (2009).
- F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- This is the truncated model based on X-ray crystal structures starting on p. 20 of the supplementary material of Ref. 25.
- A. Klant and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.
- B. Benediktsson and R. Björnsson, *Inorg. Chem.* **56**, 13417 (2017).
- L. Cao and U. Ryde, *Int. J. Quantum Chem.* **118**, e25627 (2018).
- L. Cao, O. Caldararu, and U. Ryde, *J. Phys. Chem. B* **121**, 8242 (2017).