

Encoding electronic ground-state information with variational even-tempered basis sets

Weishi Wang,^{1, a)} Casey Dowdle,² and James D. Whitfield^{1, 3}

¹⁾*Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, USA*

²⁾*Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA*

³⁾*AWS Center for Quantum Computing, Pasadena, CA 91125, USA*

We propose a system-oriented basis-set design based on even-tempered basis functions to variationally encode electronic ground-state information into molecular orbitals. First, we introduce a reduced formalism of concentric even-tempered orbitals that achieves hydrogen energy accuracy on par with the conventional formalism, with lower optimization cost and improved scalability. Second, we propose a symmetry-adapted, even-tempered formalism specifically designed for molecular systems. It requires only primitive S-subshell Gaussian-type orbitals and uses two parameters to characterize all exponent coefficients. In the case of the diatomic hydrogen molecule, the basis set generated by this formalism produces a dissociation curve more consistent with cc-pV5Z than cc-pVTZ at the size of aug-cc-pVDZ. Finally, we test our even-tempered formalism against several types of tetra-atomic hydrogen molecules for ground-state computation and point out its current limitations and potential improvements.

I. INTRODUCTION

Orbitals are essential building blocks for the many-electron wavefunctions^{1,2}. In non-interacting systems, the eigen wavefunctions are precisely described by the Slater determinants of spin-orbitals that are the eigenfunctions of the separated one-body problems. For interacting systems, initializing the many-electron reference states with proper orbitals can also improve the accuracy or compactness of the downstream multi-configurational ground-state ansatzes.

The strategies for constructing optimal orbitals can be divided into two main types. Fitting the target orbitals (implicitly determined by an ansatz) through the linear combinations of a given basis set (e.g., orbital rotations) and directly engineering the basis set form such that it can efficiently span a function space close to where those orbitals reside. Many optimization procedures, such as the Hartree–Fock methods^{3–5} and orbital localization methods^{6–8}, fall into the first type, whereas basis-set designs belong to the second. Atomic basis sets⁹ built on contracted Gaussian-type orbitals (GTOs)¹⁰ have been a successful design for molecular electronic structure. This is due to their efficient fitting of atomic orbitals despite missing the center cusps present in Slater-type orbitals (STOs)^{11,12}. In fact, many reusable atomic Gaussian basis sets have been proposed for various systems and electronic structure methods^{9,13}. Nevertheless, there still lacks a unified framework to construct system-oriented basis sets without empirical contractions or parameterizations.

Before the rapid development of tabulated atomic Gaussian basis sets, one of the approaches to construct system-oriented basis sets was by directly optimizing GTOs with respect to the system’s ground-state energy¹⁴. To reduce the number of free exponent coefficients during the optimization, Reeves et al. proposed the ideas behind even-tempered basis sets^{15–20}. Restricted to S-subshell orbitals, even-tempered basis functions are defined as a sequence of concentric Gaussian functions

$$\phi_{mn}(\mathbf{r}|\alpha_m) \equiv \mathcal{N}(\alpha_m) \exp(-\alpha_m \|\mathbf{r} - \mathbf{R}_n\|_2^2), \quad (1)$$

where $\mathbf{R}_n$ is the center coordinate in the atomic units (a.u.), $\mathcal{N}(\alpha_m) = (2\alpha_m/\pi)^{3/4}$ is the normalization factor, and α_m are the exponent coefficients correlated by a pair of primitive parameters $\gamma \equiv (\alpha, \beta)$:

$$\alpha_m = \alpha \beta^m, \quad \alpha > 0, \beta > 0, m = 0, 1, 2, \dots \quad (2)$$

For consistency, we apply the atomic unit system to all the physical quantities presented in this paper. Particularly, we set the unit of α to be 1/a.u. (i.e., the inverse of the Bohr radius a_B) and let β remain unitless. We do not explicitly denote the unit of α throughout the rest of this paper for simplicity.

It has been empirically shown that, by optimizing γ against atomic Hartree–Fock energy, even-tempered basis sets can systematically converge to the complete basis-set (CBS) limit^{21,22}. The accuracy of the resulting basis sets does not directly transfer to excited-state computations. However, this issue can be resolved by re-optimizing the basis sets directly with respect to the excited state energy²³. There have been multiple efforts to analytically explain the effectiveness of even-tempered basis sets for atomic systems. In the case of the hydrogen atom, the construction of even-tempered basis sets is justified as a discretization of an integral transformation for hydrogen atomic orbitals^{24–26}. Alternatively, they can be seen through the lens of the Gram–Schmidt process applied to Gaussian functions²⁷. Progress has also been made on rigorously bounding the error convergence. Specifically, Kutzelnigg²⁵ showed that the ground-state energy error of the hydrogen atom asymptotically scales as $\exp(-d\sqrt{m})$ for a basis set of size m , where the positive constant d is dependent on γ . Bachmayr et al. further improved upon this study and established the error bounds on how well even-tempered basis sets approximate Slater-type functions beyond the S subshell²⁸.

Today, even-tempered basis sets are primarily used to augment atomic Gaussian basis sets for the purpose of capturing diffuse orbitals consistently^{29,30}. Optimizing even-tempered basis sets directly for molecular systems, though briefly attempted²⁰, remains an overlooked direction. As programming paradigms and hardware architectures advance, more sophisticated construction and optimization of basis sets become more feasible through modern scientific software^{31–33}. Reposition-

^{a)}Electronic mail: weishi.wang.gr@dartmouth.edu

ing even-tempered basis sets as a discretization strategy rather than an augmentation measure, we pose the following question: How efficiently can a system-oriented even-tempered basis set encode electronic ground-state information beyond atomic systems? To answer this question, we introduce a simple basis-set construction formalism based on even-tempered GTOs in Equation (1), which can be directly applied to atomic and molecular systems. We further analyze the accuracy and the stability of even-tempered basis functions with respect to their tunable parameters. We draw implications from these numerical analyses and propose new hierarchical variational optimization procedures tailored for our formalism with improved numerical stability. By combining the construction formalism and optimization strategy, we demonstrate a system-oriented and tabular data-free basis-set design.

The paper is organized as follows. In Sec. II, we first define a reduced formalism of even-tempered basis sets and formalize the concept of variational even-tempered basis sets. Second, we propose a basic one-level optimization strategy that focuses on optimizing the exponent coefficients for reduced even-tempered basis sets in Sec. III. We then apply it to atomic hydrogen to quantify the correlation between α and β , as well as their impact on the stability of ground-state computation. Next, in Sec. IV, we introduce the updated two-level optimization strategy that incorporates the optimization of basis-function centers for molecular systems. We numerically test its performance on various geometric configurations of diatomic and tetra-atomic hydrogen systems. Last, we conclude our paper in Sec. V with discussions about the limitations of our design and potential future directions.

II. VARIATIONAL EVEN-TEMPERED BASIS SETS

To study the effectiveness of constructing electronic ground states with system-oriented even-tempered basis sets, we propose a modified formalism to analyze the impact of γ on subsequent electronic-structure computations. We first define a reduced sequence of even-tempered exponent coefficients:

$$\zeta_m = \alpha \beta^m, \quad \alpha > 0, \beta > 0, m = 1, 2, \dots \quad (3)$$

Compared to the conventional form in Equation (2), index m in this new form iterates from one instead of zero, which ensures that α and β both exist for all possible ζ_m . Correspondingly, one can vary β to analyze its effect on the downstream computations while having α fixed at a global value, or vice versa. Such a subtle modification in parameterizing the exponent coefficients paves the way for an improved variational optimization strategy of even-tempered basis sets, which we shall introduce in Sec. III.

Formally, we define a **reduced even-tempered basis set** as

$$\mathcal{G}_M^r \equiv \{\phi_{mn}(\mathbf{r}|\zeta_m)| m = 1, \dots, M; n = 1, \dots, N\}, \quad (4)$$

where M is the **basis degree** that determines the multiplicity of concentric even-tempered basis functions, while N specifies the number of basis-function centers $\mathbf{R}_n$. Therefore, the basis-set size $|\mathcal{G}_M^r|$ is equal to $M \times N$. At the boundary configuration

$\mathcal{G}_M^r$, the effects of α and β converge, as the optimal value of γ is degenerate with respect to their combinations. We also note that the basis functions in $\mathcal{G}_M$, whether or not having the same centers, are all characterized by γ . This simplification further reduces the cost of basis-set parameter optimization.

In comparison to $\mathcal{G}_M^r$, we define the **non-reduced even-tempered basis set** based on the conventional formalism as

$$\mathcal{G}_M^o \equiv \{\phi_{(m-1)n}(\mathbf{r}|\alpha_{m-1})| m, n\}, \quad (5)$$

where m and n still iterates from 1 to M and N , respectively. Throughout the rest of this paper, we shall use symbol “ $\mathcal{G}_M$ ” for the statements where both the reduced and non-reduced formalisms apply.

Given a target many-electron system $\mathbb{M}$ (under specific nuclear geometry and electronic spin configuration), let the parameterized even-tempered basis set be $\mathcal{G}_M(\Theta)$, where Θ is a (vectorized) set of all its tunable parameters that are distinguished by their symbols rather than assigned values. We define a partial ground-truth variational optimization of $\mathcal{G}_M(\Theta)$ as

$$\tilde{\theta} = \arg \min_{\theta \subseteq \Theta} \hat{\mathcal{E}}(\mathbb{M})[\mathcal{G}_M(\Theta)], \quad (6)$$

where $\tilde{\theta}$ is the optimized value for a parameter subset θ of Θ ; $\hat{\mathcal{E}}(\mathbb{M})$ is ground-state energy ansatz, treated as a functional of $\mathcal{G}_M$, parameterized by $\mathbb{M}$. Moreover, we focus on the optimization of even-tempered basis sets with respect to the Hartree–Fock (HF) energy functional ($\hat{\mathcal{E}} := \hat{\mathcal{E}}_{\text{HF}}$). In this way, the accuracy of the ground-state energy approximation directly reflects the performance of the basis-set design, since the HF approximation only relies on orbital rotation. For the gradient-based implementation of the optimization procedure in Equation (6), we refer the reader to the technical literature of Quiqbox³², an open-source basis-set software library that we used to construct or optimize basis sets. Finally, we define a **variational even-tempered basis set** with respect to the optimized parameter subset θ as

$$\tilde{\mathcal{G}}_M(\Theta \setminus \theta) \equiv \mathcal{G}_M(\Theta|_{\theta=\tilde{\theta}}). \quad (7)$$

Based on Equations (4) and (5), we divide the adjustable coefficients of $\mathcal{G}_M$ into two levels. On the first level, $\{\zeta_m|m\}$ (or $\{\alpha_m|m\}$) affects how the basis functions centered at a given $\mathbf{R}_n$ span a one-body Hilbert space²⁷, which is a subspace of $L^2(\mathbb{R}^3)$. On the second level, $\{\mathbf{R}_n|n\}$ further determines the placements of the basis functions in real space $\mathbb{R}^3$. Specifically, at the limit of $N = 1$ with a fixed $\mathbf{R}_1$, $\{\zeta_m|m\}$ affects the “resolution” of $\mathcal{G}_M^r$; at the limit of $M = 1$ with a fixed $\{\zeta_m|m\}$, $\{\mathbf{R}_n|n\}$ determines the “volume” of $\mathcal{G}_M^r$. However, when M and N are both larger than one, the effects from these two levels are not necessarily orthogonal. Hence, the key to generating system-oriented $\mathcal{G}_M$ lies in efficiently optimizing the primitive parameters (e.g., γ) that control these two levels of basis-set coefficients.

III. ONE-LEVEL VARIATIONAL OPTIMIZATION

In this section, we only consider the effect of γ . Therefore, we focus on atomic systems for which even-tempered basis functions were originally considered. In this context, the potential benefit of placing the basis functions on locations other than the nucleus can be eliminated. Furthermore, we fix the values of α to preset global values, and only optimize β in the reduced formalism. In addition to applying the parameter optimization routine from Quiqbox, we propose a bootstrap strategy that iteratively uses the results of $\tilde{\mathcal{G}}_{M-1}^r$ to jump-start the optimization for $\tilde{\mathcal{G}}_M^r$. The pseudo-code of this strategy is presented in Algorithm 1.

Algorithm 1 β -adapted even-tempered basis optimization

```

1:  $\tilde{\beta} := 1$ 
2: Initialize  $\hat{\mathcal{E}}(\mathbb{M})$ ,  $\alpha$ ,  $M_{\max}$ 
3: for  $M$  in  $(1, \dots, M_{\max})$  do
4:   Initialize  $\mathcal{G}_M^r(\{\gamma\})$  with  $\gamma := (\alpha, \tilde{\beta})$ 
5:    $E_{\text{HF}} := \hat{\mathcal{E}}[\mathcal{G}_M^r]$ ,  $i := 0$ ,  $\Delta := \text{inf}$ 
6:   while  $i < i_{\max}$  and  $|\Delta| > \delta$  do
7.1)     $i := i + 1$ 
7.2)     $\tilde{\beta} := \text{optimizer}[\hat{\mathcal{E}}, \partial\hat{\mathcal{E}}/\partial\beta, \mathcal{G}_M^r]$ 
7.3)     $\gamma := (\alpha, \tilde{\beta})$ 
7.4)     $\mathcal{G}_M^r := \mathcal{G}_M^r(\{\gamma\})$ 
7.5)     $\Delta := E_{\text{HF}} - \hat{\mathcal{E}}[\mathcal{G}_M^r]$ 
7.6)     $E_{\text{HF}} := E_{\text{HF}} - \Delta$ 
8:   end while
9:   Store  $(M, \tilde{\beta}, E_{\text{HF}})$ 
10: end for
  
```

The advantage of Algorithm 1's bootstrap initialization is that, by construction, the initial guess of $\tilde{\mathcal{G}}_{M+1}^r$ is a proper superset of $\tilde{\mathcal{G}}_M^r$:

$$\mathcal{G}_{M+1}^r|_{i=0} = \tilde{\mathcal{G}}_M^r \cup \left\{ \phi_{(M+1)n}(\mathbf{r}; \gamma|_{\beta=\tilde{\beta}(\alpha; M)}, \mathbf{R}_n) \mid n \right\}. \quad (8)$$

Hence, by the variational principle, it imposes the ground-state energy bound

$$\hat{\mathcal{E}}(\mathbb{M})[\mathcal{G}_{M+1}^r|_{i=0}] \leq \hat{\mathcal{E}}(\mathbb{M})[\tilde{\mathcal{G}}_M^r] \quad (9)$$

such that extending a reduced variational even-tempered basis from degree M to $M+1$ never worsens its performance.

A. Ground-state energy of hydrogen atom

For the probing system, we choose the simplest atom: hydrogen. As an analytically solvable system, it has been tested in several literature^{22,25,34,35} to study the scaling performance

of even-tempered basis sets. Moreover, since the hydrogen atom is free of electron-electron interaction, its HF energy E_{HF} converges to the true ground-state energy E_{GT} (-0.5 Ha) at the CBS limit.

We use the optimized values of γ from reference²² to construct the conventional counterpart of $\tilde{\mathcal{G}}_M^r$, $\tilde{\mathcal{G}}_M^o$, due to its high-precision numerical computation (70-digit precision according to the authors). To verify the results from this reference, we recomputed the ground-state energies (from $M = 2$ to $M = 10$) using its parameter settings in a 128-bit floating number system and obtained a mean absolute error (MAE) $\sim 2 \times 10^{-10}$ with respect to its reported data. We generated a series of $\tilde{\mathcal{G}}_M^r$ from the optimized $\tilde{\beta}$ using Algorithm 1 under multiple values of α and M . The comparison of the ground-state energies between $\tilde{\mathcal{G}}_M^r$ and $\tilde{\mathcal{G}}_M^o$ is plotted in FIG. 1. To check the values of the representative energies and the corresponding $\tilde{\beta}$, we refer the reader to TABLE II in Appx. A.

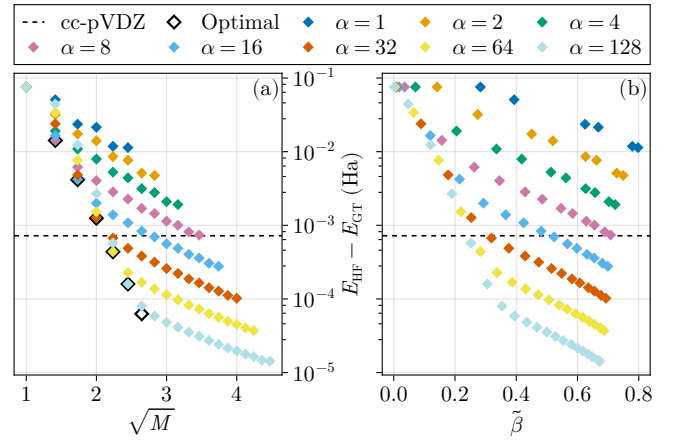


FIG. 1: The atomic hydrogen ground-state energy error of the reduced variational even-tempered basis set $\tilde{\mathcal{G}}_M^r$ (the solid colored diamonds) with respect to M (equivalent to basis set size in this case) and α . (a) shows the logarithmic errors, where the open diamonds with black strokes represent the conventionally optimized even-tempered basis set $\tilde{\mathcal{G}}_M^o$ at different M (incremented by one). (b) shows the relation between the error and optimized parameter $\tilde{\beta}$ for each configuration of $\tilde{\mathcal{G}}_M^r$. For each color group, the data points in (b) follow the same correspondence to M as the ones in (a), respectively.

As provided by previous studies^{25,35}, the asymptotic energy error of $\tilde{\mathcal{G}}_M^o$ follows an exponential decay with respect to the square root of its basis-set size:

$$\ln(|E_{\text{HF}} - E_{\text{GT}}|) \propto \sqrt{|\tilde{\mathcal{G}}_M^o|}. \quad (10)$$

This relation was verified by works^{22,28} and is reproduced again in FIG. 1(a) by the open diamonds (data points marked as “Optimal” for $\tilde{\mathcal{G}}_M^o$), as M is equal to the basis-set size $|\mathcal{G}_M^o|$ for atomic systems. Unsurprisingly, the energy errors of $\tilde{\mathcal{G}}_M^o$ form lower bounds for $\tilde{\mathcal{G}}_M^r$. Additionally, we found that there exists a numerically consistent match between the errors of

$\tilde{\mathcal{G}}_M^o$ and $\tilde{\mathcal{G}}_M^r$ with exponentially increased α in FIG. 1(a): a sequence of colored diamonds in which each one is enclosed by an open diamond.

To verify whether this matching pattern is also reflected on the exponent coefficients, we computed the mean absolute errors (MAE) of $\{\zeta_m | m = 1, \dots, M\}$ for different $\tilde{\mathcal{G}}_M^r(\{\alpha\})$ against $\{\alpha_{m'-1} | m' = 1, \dots, M\}$ from the respective $\tilde{\mathcal{G}}_M^o$. For each computed MAE, $\{\zeta_m | m\}$ is rearranged to form the same ascending order as $\{\alpha_{m'} | m'\}$. The results are shown in FIG. 2(a), where each colored diamond represents the MAE for one configuration (M, α). As M increases, the optimal value of α grows exponentially. More specifically, $\tilde{\mathcal{G}}_M^o$ can be well approximated by $\tilde{\mathcal{G}}_M^r(\alpha)$ through following mapping:

$$\tilde{\mathcal{G}}_M^o \rightarrow \tilde{\mathcal{G}}_M^r(\{\alpha = 2^{M+1}\}) \quad \text{for } M \geq 2. \quad (11)$$

This mapping relation indicates that the optimal values of α and β are inherently coupled through M . Aside from relation (10), we discovered that the scaling of the energy error with respect to β also transits to an exponential decay after the value of β surpasses the optimal value for a given α . This relation is shown in FIG. 1(b). To examine how α affects the optimization of β , we plot the rearranged $\{\zeta_m | m\}$ with respect to different values of α in the case of $M=6$ in FIG. 2(b).

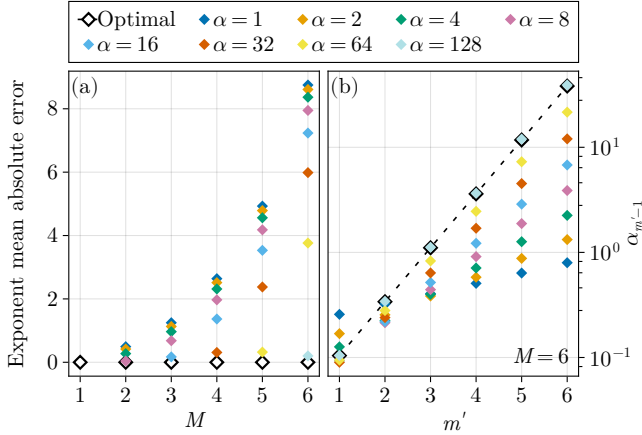


FIG. 2: The errors of the exponents $\{\zeta_m | m\}$ from $\tilde{\mathcal{G}}_M^r$ with respect to the optimal exponents $\{\alpha_{m'} | m'\}$ from $\tilde{\mathcal{G}}_M^o$. In (a), each colored diamond represents the mean absolute error of $\{\zeta_m | m\}$ (rearranged in ascending order to match $\{\alpha_{m'} | m'\}$) from $\tilde{\mathcal{G}}_M^r(\alpha)$. (b) shows that for $M=6$, the rearranged exponents of $\tilde{\mathcal{G}}_M^r(\alpha)$ converge to their respective optimal values (the open diamonds chained by a dotted line) as α increases exponentially from 1 to 128.

A potential challenge for applying even-tempered basis sets, which has been seldom discussed in the previous literature, is the numerical sensitivity of the corresponding overlap matrix. In general, the overlap matrix of a basis set becomes ill-conditioned when there are near-linearly dependent basis functions. For Gaussian basis sets, this issue can occur as more concentric (or near concentric but diffuse) GTOs with the same angular momentum are added. An ill-conditioned

overlap matrix can lead to numerical instability in solving the generalized eigenvalue problem for the free-fermion (or mean-field approximated) eigenstates:

$$HC = SC\varepsilon, \quad (12)$$

where C is the orbital coefficient matrix, ε is the orbital energies as a diagonal matrix, and H is the one-body Hamiltonian (or the Fock matrix) discretized by a basis set with its overlap matrix S . This issue typically can be mitigated by specific orthonormalization strategies^{36,37} that produce a smaller basis set at the cost of truncating the function space that the original basis sets span. However, within the scope of this paper, which focuses solely on the stability and accuracy of the basis set itself, we refrained from applying these techniques during the HF computation. In the case of the hydrogen atom, we plot the scaling of the overlap-matrix condition numbers for $\tilde{\mathcal{G}}_M^o$ and $\tilde{\mathcal{G}}_M^r$ in FIG. 3, both of which scale exponentially with respect to M . In contrast, for $\tilde{\mathcal{G}}_M^r(\{\alpha = 2^{M+1}\})$, which is highlighted by the green stroke, its overlap-matrix condition number scales sub-exponentially. This indicates that applying the mapping of Equation (11) can provide a better numerical stability. For a given basis set size, one can still increase α to estimate an asymptotic lower bound of the condition for $\tilde{\mathcal{G}}_M^r$, as shown in FIG. 3(b).

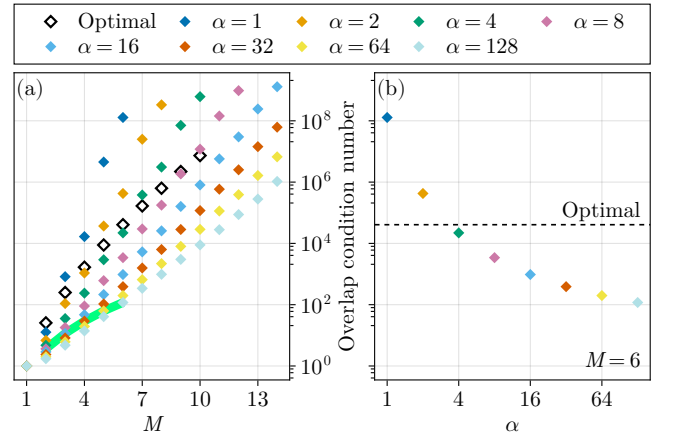


FIG. 3: The scaling of the overlap-matrix (S) condition number for various even-tempered basis set configurations tested against the hydrogen atom. In (a), the colored diamonds represent the results for $\tilde{\mathcal{G}}_M^r(\alpha)$, and the open diamonds (marked as “Optimal”) represent the results for $\tilde{\mathcal{G}}_M^o$. The light green stroke passes through the data points of $\tilde{\mathcal{G}}_M^r(\{\alpha = 2^{M+1}\})$. (b) shows how the S condition number of $\tilde{\mathcal{G}}_6^r(\alpha)$ is suppressed as α increases, where the dashed horizontal line represents the condition number for $\tilde{\mathcal{G}}_6^o$.

Although the $\tilde{\mathcal{G}}_M^r$ is generated with respect to HF energy, in the case of atomic hydrogen, the energy error directly reflects the capability of using even-tempered Gaussian functions to approximate the ground-state atomic orbital. Hence, Equation (11) indicates that one may simplify such orbital fitting procedure into the optimization of a single parameter (β). This simplified optimization strategy is useful for generalizing

approximating arbitrary spherical (S-subshell) orbitals with even-tempered Gaussian functions.

IV. TWO-LEVEL VARIATIONAL OPTIMIZATION

In the last section, we quantify the internal correlation between α and β in the case of atomic hydrogen and propose a one-level optimization strategy (for β) in Algorithm 1. For molecular systems, we need to consider the optimization of not only the exponent parameters (first-level), but also the centers of the even-tempered basis functions (second-level). This is because the distortion of the specific geometry of the centers can help incomplete basis sets fit the target molecular orbitals^{32,38–40}.

To control the shifting (floating) of the basis-function centers, we use implicit parameters that correlate the center coordinates while preserving a desirable symmetry based on the target system. Formally, we define the **correlated basis centers** with respect to a molecular system $\mathbb{M}$ by

$$\mathcal{C}_N(\boldsymbol{\nu}) \equiv \{\mathbf{R}_n(\boldsymbol{\nu}) | n = 1, \dots, N\}, \quad (13)$$

where each $\mathbf{R}_n(\boldsymbol{\nu})$ set to be the basis-function center of ϕ_{mn} for all m , parameterized by $\boldsymbol{\nu}$ such that the geometric symmetry formed by $\mathbf{R}_n$ is invariant under the optimization of $\boldsymbol{\nu}$. Furthermore, the implicit parameterization of $\mathbf{R}_n$ reduces the dimension of the basis-set parameter space compared to the direct optimization of all the coordinate components. For instance, in homonuclear diatomic systems, $\boldsymbol{\nu}$ can contain only a single parameter: the distance between the two basis-function centers.

Before incorporating the second-level optimization of basis-function centers, we shall first test the transferability of the one-level optimization strategy (Algorithm 1) alone for molecular systems. Specifically, we applied it to the diatomic hydrogen (H_2) at a bond length of 1.4 a.u. The results of the α - β relation compared to the atomic hydrogen are shown in FIG. 4.

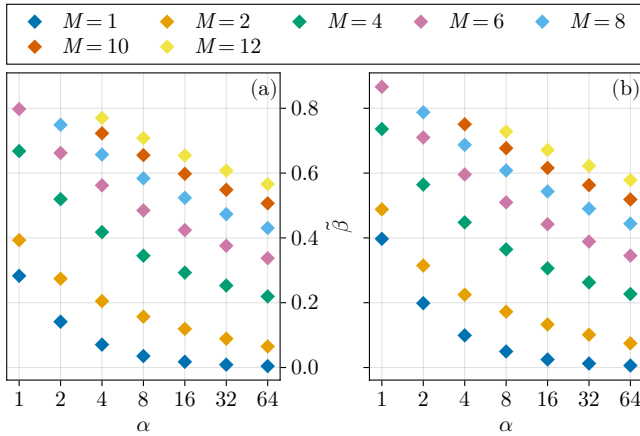


FIG. 4: The numerically-stable values of $(\alpha, \tilde{\beta})$ with respect to basis degree M of the even-tempered basis sets for: (a) the atomic hydrogen; (b) the diatomic hydrogen (H_2) molecule at the bond length of 1.4 a.u.

Comparing FIG. 4(a) and FIG. 4(b), we can see that despite being qualitatively similar, the quantitative relation between $\tilde{\beta}$ and α for H_2 is not the same as that for atomic hydrogen. Subsequently, Equation (11) no longer applies. Nevertheless, this divergence is not surprising. In the case of atomic hydrogen, the error of E_{HF} directly represents the discretization error of the basis set for the ground-truth (Slater-type) atomic orbital; whereas for molecular systems, the generated molecular orbitals carry the errors from both the basis discretization and HF approximation of the electron-electron interactions. Admittedly, it is possible to numerically re-derive a new fitting function for α . Such an effort would not be an efficient solution when the interpolated coefficients are determined system by system. Therefore, for molecular systems, we instead propose an optimization strategy that imposes an iterative tuning procedure of α on top of the simultaneous optimization of β and $\boldsymbol{\nu}$. Then, the only remaining hyperparameter is the maximal basis degree M_{max} . The pseudo-code of this new strategy is shown in Algorithm 2, where $\boldsymbol{\theta} := \boldsymbol{\nu} \frown \tilde{\beta}$ are the vectorized composite parameters concatenated from $\boldsymbol{\nu}$ and $\tilde{\beta}$. The initial value of $\boldsymbol{\nu}$ is set such that $\mathcal{C}_N(\boldsymbol{\nu})$ match the exact nuclear geometry of the target system.

Algorithm 2 α -bootstrap even-tempered basis optimization

- 1: Initialize $\alpha, \hat{\mathcal{E}}(\mathbb{M}), M_{\text{max}}, \boldsymbol{\nu}$
 - 2: $\tilde{\beta} := 1/\alpha$
 - 3: **for** M in $(1, \dots, M_{\text{max}})$ **do**
 - 4: **if** $\tilde{\beta} < 1$ **then**
 - 5: $\alpha := \alpha \tilde{\beta}^M$
 - 6: $\tilde{\beta} := 1/\tilde{\beta}$
 - 7: **end if**
 - 8: **if** $\exists n \in \mathbb{Z}_{>0}$ such that $M = 2^n$ **then**
 - 9: $\alpha := \alpha/\tilde{\beta}$
 - 10: **end if**
 - 11: Initialize $\mathcal{G}_M^r(\{\boldsymbol{\nu}, \boldsymbol{\gamma}\})$ with $\boldsymbol{\gamma} := (\alpha, \tilde{\beta})$
 - 12: $E_{\text{HF}} := \hat{\mathcal{E}}[\mathcal{G}_M^r], i := 0, \Delta := \text{inf}$
 - 13: **while** $i < i_{\text{max}}$ and $|\Delta| > \delta$ **do**
 - 14.1) $i := i + 1$
 - 14.2) $\boldsymbol{\nu} \frown \tilde{\beta} := \text{optimizer}[\hat{\mathcal{E}}, \partial \hat{\mathcal{E}} / \partial \boldsymbol{\theta}, \mathcal{G}_M^r]$
 - 14.3) $\boldsymbol{\gamma} := (\alpha, \tilde{\beta})$
 - 14.4) $\mathcal{G}_M^r := \mathcal{G}_M^r(\{\boldsymbol{\nu}, \boldsymbol{\gamma}\})$
 - 14.5) $\Delta := E_{\text{HF}} - \hat{\mathcal{E}}[\mathcal{G}_M^r]$
 - 14.6) $E_{\text{HF}} := E_{\text{HF}} - \Delta$
 - 15: **end while**
 - 16: Store $(M, \boldsymbol{\nu}, \boldsymbol{\gamma}, E_{\text{HF}})$
 - 17: **end for**
-

Aside from including the optimization of $\boldsymbol{\nu}$, the most significant addition in Algorithm 2 compared to Algorithm 1 is the bootstrap-based tuning procedure of α by lines 3–10. This

procedure can be divided into two subroutines that are intermittently executed. The first subroutine (lines 3–7) effectively relabels the basis functions by restricting the domains of α and $\tilde{\beta}$ to be $(0, 1]$ and $(1, \infty)$, respectively. This subroutine preserves the single-particle function space that $\tilde{\mathcal{G}}_M^r$ spans at the end of last optimization cycle, while resuming the optimization on a submanifold of $\hat{\mathcal{E}}(\mathbb{M})$ to avoid degenerate local minima. Moreover, it helps compensate for the exponential growth of α by the polynomial growth of $\tilde{\beta}$. When the second subroutine (lines 8–10) is executed, α is divided by $\tilde{\beta}$, thus still bounded by $(0, 1]$. To avoid an exponential decay of α with respect to M , the update rate of α also follows an exponential decay. The two cases in which α is updated or remains the same as in the last cycle correspond to two directions for extending the basis set: appending a more diffuse basis function or a more localized one. In this way, the bootstrap initialization of the even-tempered basis set is more balanced compared to Algorithm 1. To verify the numerical stability of Algorithm 2, we test it against the diatomic hydrogen at a typical bond length of 1.4 a.u. Specifically, the exponent parameters γ , the overlap condition number, and the HF energy for $\tilde{\mathcal{G}}_M^r$ of different M are plotted in FIG. 5.

First, FIG. 5(a) shows that the HF energy with $\tilde{\mathcal{G}}_M^r$ generated by Algorithm 2 converges as M increases. Particularly, each jump of the overlap condition number of the optimized basis set reflects the iteration cycle where α is updated at the beginning, which helps suppress the subsequent growth of the condition number. This update has a similar effect on the convergence of $\alpha\tilde{\beta}$ and $\tilde{\beta}$ in FIG. 5(b). Specifically, when α is updated at the beginning of an optimization cycle, the resulting $\tilde{\beta}$ jumps (while $\alpha\tilde{\beta}$ drops), leading to a damped value decay in the subsequent cycles until the next update. Due to the consistent behavior across the performance indicators (energy and condition number) and the basis set parameters, one may choose one over the others as the criterion for obtaining the optimal minimum of M , depending on the required strictness for convergence.

A. Dissociation of diatomic hydrogen

We have demonstrated the performance scalability of Algorithm 2 with respect to the basis degree M for a fixed H_2 geometry. In this subsection, we test out its performance consistency at different H_2 bond lengths. Specifically, we compare $\tilde{\mathcal{G}}_9^r$ generated by Algorithm 2 against several correlation-consistent basis sets^{41,42} (in the Cartesian representation) for producing the dissociation energy curve of H_2 on the HF level. The optimized parameters of the generated $\tilde{\mathcal{G}}_9^r$ are listed in TABLE III in Appx. A, and results are shown in FIG. 6. cc-pV5Z (with optimized general contraction⁴³) is also applied to the full configuration-interaction (FCI) computation (realized by PySCF⁴⁴) to estimate the ground-truth values near the CBS limit. The justification for choosing such a basis set is backed by numerical observations. For tested bond lengths, the energy differences between FCI/cc-pV5Z and FCI/aug-cc-pV5Z^{41,42} are always below $\sim 5 \times 10^{-4}$ Ha, approximately one third of the energy resolution at chemical accuracy.

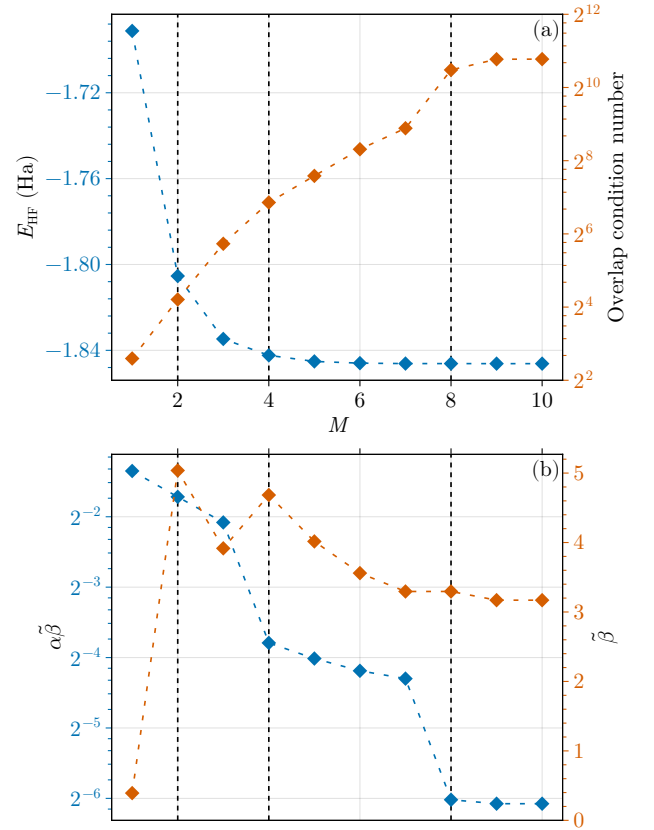


FIG. 5: The convergence characteristics of a variationally optimized (through Algorithm 2) reduced even-tempered basis set $\tilde{\mathcal{G}}_M^r$ with respect to the basis degree M in the case of H_2 at a bond length of 1.4 a.u. (a) shows the restricted closed-shell Hartree-Fock energy (blue dashed line) for $\tilde{\mathcal{G}}_M^r$ and the condition number of the corresponding overlap matrix (orange line). (b) shows the optimized values of the smallest basis-function exponent $\alpha\tilde{\beta}$ (blue dashed line) and the growth rate $\tilde{\beta}$ (orange dashed line). For (a) and (b), the vertical dashed lines cross the data points corresponding to the steps when α is updated.

Overall, the correlation energies for different basis sets exhibit the same monotonic trend in the medium-to-long bond length region $[2, 5]$ a.u. However, in $(0, 1.8]$ a.u., only the correlation energies of $\tilde{\mathcal{G}}_9^r$ and the near-complete cc-pV5Z still decrease, while the remaining tested basis sets produce qualitatively incorrect curves that increase the correlation energies. As a result, the HF energy of $\tilde{\mathcal{G}}_9^r$ is lower than that of aug-cc-pVDZ and cc-pVTZ in this lower bond-length region. Notably, $\tilde{\mathcal{G}}_9^r$ is of the same size as aug-cc-pVDZ (2×9) and only contains S-subshell orbitals.

To better estimate the discretization error of $\tilde{\mathcal{G}}_9^r$ in the low bond-length region, we compare its Hartree-Fock charge density (HF density) against that of aug-cc-pVDZ and cc-pVTZ at the bond length of 0.6 a.u., where the maximal performance divergence appears in FIG. 6. By taking the difference between the HF density of each compared basis set and the FCI/cc-

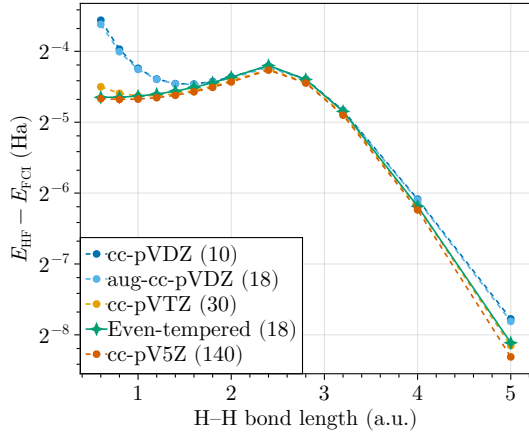


FIG. 6: The unrestricted open-shell Hartree-Fock (UHF) energy errors of the H_2 molecule at different bond lengths with respect to the reduced variational even-tempered basis set ($\tilde{\mathcal{G}}_9^r$) versus correlation-consistent basis sets. $\tilde{\mathcal{G}}_9^r$ is generated by Algorithm 2 also targeting the UHF functional. The size of each tested basis set is included in the parentheses after its name in the legend box.

pV5Z charge density (FCI density), we measure the density errors and visualize them in FIG. 7. The HF density errors of cc-pV5Z are also included as a baseline reference for the HF approximation. It is possible to further integrate over the density difference and use the scalar result as a metric for the difference⁴⁵. However, the distinct shapes of those density differences provide more information about the distribution of density errors, revealing inhomogeneous biases imposed by the incomplete basis sets. For instance, as shown by FIG. 7(a-d), the density difference for each basis set forms multiple filled contours of distinct shapes in the x - y plane. This geometric discrepancy indicates that the specific construction of incomplete basis sets imposes a non-uniformly distributed, yet spatially structured, overestimation (in red) or underestimation (in blue) in approximating the HF density at the CBS limit.

Focusing on the density errors along the z axis, we also observe wavelet-like curves in FIG. 7(e), which all have their respective main lobes centered at the origin, corresponding to the maximal amplitude of the single-point error. Restricting to the magnitude of the density errors (indicated by the flipped dashed lines), we can analyze the magnitude of the density error for each basis set more quantitatively. Compared to aug-cc-pVDZ, the error magnitude of $\tilde{\mathcal{G}}_9^r$ decays more consistently and is especially lower near the origin and in the far region $(-\infty, -6) \cup (6, \infty)$. Compared to cc-pVTZ, $\tilde{\mathcal{G}}_9^r$ becomes less advantageous. It produces lower errors in the region $(-4, 4)$ a.u., higher errors in the region $(-6.2, -4) \cup (4, 6.2)$ a.u., and asymptotically match the error of cc-pVTZ in the region $(-\infty, -6.2) \cup (6.2, \infty)$.

B. Ground-state approximation of tetra-atomic hydrogen

In the last subsection, we have demonstrated that, through Algorithm 2, we can generate variational even-tempered basis sets consisting solely of multi-center S-subshell orbitals that exhibit consistent performance over a range of bond lengths for the diatomic hydrogen molecule. In this subsection, we further test out the potential of Algorithm 2 on molecular systems with relatively more complex geometries. Specifically, we compute the restricted closed-shell Hartree-Fock (RHF) energies with variational even-tempered basis set configurations against three types of tetra-atomic hydrogen (H_4) molecules: linear chain, square planar, and rhombus. The geometry configurations of these target systems are shown in FIG. 8.

We first directly applied Algorithm 2 to generate variational even-tempered basis sets $\tilde{\mathcal{G}}_M^r$ of three basis degrees with the correlated basis centers specified in FIG. 8. Their respective optimized parameters are listed in TABLE IV in Appx. A, and the corresponding RHF energies are shown in TABLE I (marked as “Direct variational even-tempered”). Compared to aug-cc-pVDZ, which has the same number of basis functions, $\tilde{\mathcal{G}}_9^r$ only produces lower RHF energy for the linear chain H_4 at bond lengths of 1.2 a.u. and 1.6 a.u. To verify if this underwhelming performance is caused by the lack of GTOs of higher angular momenta in $\tilde{\mathcal{G}}_9^r$, we computed the RHF using subsets of aug-cc-pVDZ and cc-pVTZ that contain only S subshell orbitals (and S plus P subshell orbitals). Compared to the results of these subsets, which are also included in TABLE I, the directly generated variational even-tempered basis set does produce lower RHF energy when utilizing the same number of primitive S-subshell GTOs (e.g., $\tilde{\mathcal{G}}_6^r$ vs. the S-subshell part of aug-cc-pVDZ). However, it does not necessarily offer any advantage in terms of basis-set size. This is because in aug-cc-pVDZ and cc-pVTZ, the primitive GTOs can be combined to form contracted GTOs. Moreover, the inclusion of P-subshell GTOs indeed significantly lowers the energy, which cannot be achieved with such a direct even-tempered construction.

To address the limitations of S-subshell orbitals centered on correlated basis centers within the current construction and optimization framework, we introduce the notion of **nested variational even-tempered basis sets**. They are generated based on the already-optimized $\tilde{\mathcal{G}}_M^r$, which are re-categorized as the **direct variational even-tempered basis sets**. We extend $\tilde{\mathcal{G}}_M^r$ by adding even-tempered basis functions parameterized by a new γ on the augmented centers that are at midpoints between its basis-function centers, then optimize γ by a procedure similar to Algorithm 2. We can characterize such a nested construction by a pair of basis degrees. The first basis degree inherits from the original $\tilde{\mathcal{G}}_M^r$, and the second basis degree specifies the number of even-tempered basis functions located on each augmented center. For the previously tested H_4 geometries, we define the augmented centers in FIG. 8 by gray markers shaped like “+”. We used these augmented centers to generate nested variational even-tempered basis sets based on $\tilde{\mathcal{G}}_6^r$. The optimized parameters of the appended basis subsets with different augmented basis degrees are listed in TABLE V in Appx. A, and the RHF energies of the corresponding nested basis sets are shown in rows 3–6 of TABLE I.

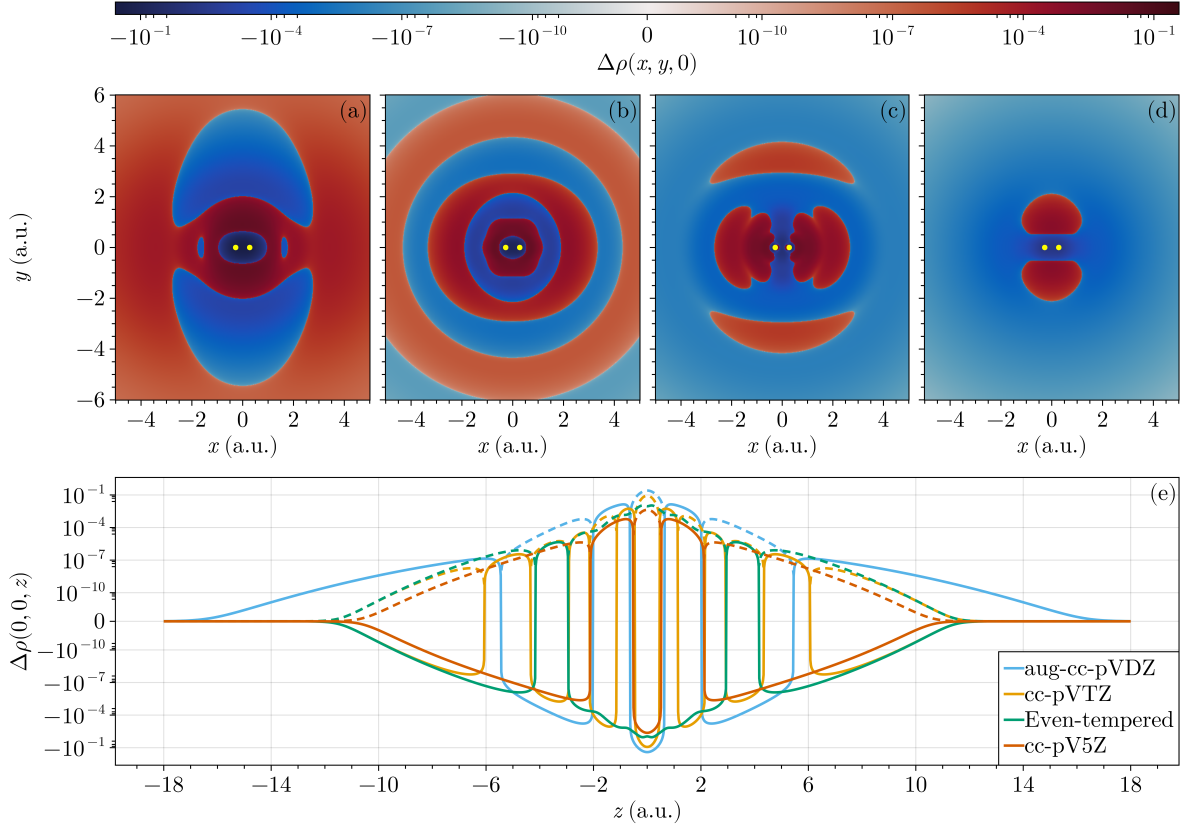


FIG. 7: The differences between the charge densities on the HF level in various basis sets and the ground-truth charge density (FCI/cc-pV5Z) for the H_2 molecule at the bond length of 0.6 a.u. Sub-figures (a–d) correspond to the density differences on the Cartesian x – y plane ($z = 0$) of using aug-cc-pVDZ, cc-pVTZ, the even-tempered basis ($\tilde{\mathcal{G}}_9^r$), and cc-pV5Z, respectively. The red areas represent overestimated regions, the blue areas represent underestimated regions, and the two yellow dots represent the nuclear positions. Sub-figure (e) shows the density differences along the z axis ($x = y = 0$) for the compared basis sets, where the solid lines represent the signed differences and the dashed lines represent the absolute differences.

Even just adding even-tempered basis functions with basis degree $M = 1$, the performance of the even-tempered basis set is already significantly improved. At a basis-degree pair of (6, 2), the nested variational even-tempered basis set contains fewer basis functions than aug-cc-pVDZ for all H_4 geometries. Yet, this configuration produces lower RHF energy for the linear chain H_4 and comparable energies (difference below 2.5×10^{-3} Ha) for the square H_4 . To confirm that for each tested H_4 geometry, all the HF states corresponding to the compared RHF energies have the same spatial symmetry, we also computed their respective charge densities. The visualized comparisons of these charge densities are included in Appx. B.

V. CONCLUSIONS AND DISCUSSIONS

In this paper, we revisit the even-tempered basis functions and use them as a direct approach to construct molecular orbitals for electronic ground-state wavefunctions. Based on the conventional even-tempered formalism, we propose system-oriented basis set designs with variational and numerically stable optimization procedures.

In Sec. II, we first propose a reduced formalism of even-

tempered basis sets and define the (reduced) variational even-tempered basis sets $\tilde{\mathcal{G}}_M^r$ with tunable parameters optimized³² against the electronic ground-state energy. Next, in Sec. III, we apply it to the atomic hydrogen to quantify the correlation between the two parameters α and β that determine the optimal exponent coefficients of the even-tempered basis functions. We show that even when these two parameters are not optimized simultaneously (through Algorithm 1), one can still systematically achieve the exponential decay of the ground-state energy error (see FIG. 1). This numerical observation rationalizes the hierarchical separation of tuning α and $\tilde{\beta}$. We further note the numerical instability of conventionally optimized even-tempered basis sets, which arises from the exponential scaling of the overlap matrix's condition number (see Fig. 3). We propose a simplified formula for generating $\tilde{\mathcal{G}}_M^r$ in Equation (11), where only $\tilde{\beta}$ is actively optimized for the hydrogen atom. Not only does the accuracy of this simplified version of $\tilde{\mathcal{G}}_M^r$ match the conventional approach across different basis degree M , but it also presents better numerical scalability.

After improving the optimization of atomic even-tempered basis sets, we consider the direct generation of multi-center even-tempered basis sets for molecular systems. In Sec. IV,

Basis set		H ₄ RHF energy excluding nuclear repulsions (Ha)					
		Linear chain			Square planar		Rhombus
Type	Basis degree	1.2 a.u.	1.6 a.u.	2.0 a.u.	2.0 a.u.	2.4 a.u.	2.2 a.u.
Direct variational even-tempered	3	−5.61693	−4.86449	−4.30367	−4.63276	−4.17920	−4.54378
	6	−5.67733	−4.88849	−4.31855	−4.65320	−4.20025	−4.56171
	9	−5.67814	−4.88905	−4.31922	−4.65369	−4.20099	−4.56200
Nested variational even-tempered	(6, 1)	−5.67905	−4.89083	−4.32096	−4.65734	−4.20421	−4.56993
	(6, 2)	−5.67918	−4.89126	−4.32138	−4.65816	−4.20474	−4.57030
	(6, 3)	−5.67931	−4.89141	−4.32231	−4.65837	−4.20538	−4.57116
aug-cc-pVDZ	S subshell	6 → 3	−5.65354	−4.88051	−4.31547	−4.64378	−4.19302
	Total (S+P)	12 → 9	−5.66054	−4.88549	−4.32032	−4.65981	−4.20704
cc-pVTZ	S subshell	7 → 3	−5.67545	−4.88672	−4.31724	−4.64549	−4.19539
	S+P subshells	13 → 9	−5.67935	−4.89210	−4.32297	−4.66207	−4.20831
	Total (S+P+D)	19 → 15	−5.67945	−4.89218	−4.32331	−4.66251	−4.20859

TABLE I: Comparison between even-tempered basis sets generated by Algorithm 2 and several atomic (Cartesian) Gaussian basis sets in being applied to computing the RHF energy for various geometries of H₄ molecules. For the linear chain type, three interatomic distances (i.e., bond lengths) are tested: 1.2 a.u., 1.6 a.u., and 2.0 a.u. For the square planar type, two edge lengths are tested: 2.0 a.u. and 2.4 a.u. Additionally, a rhombus type with a 60° angle and an edge length of 2.2 a.u. is tested. For each “nested variational even-tempered basis set” configuration, its basis degrees are specified by a pair of integers. The first represents the basis degree of its based “direct variational even-tempered” basis set, and the second represents the number of even-tempered basis functions added at each augmented center. For the atomic Gaussian basis sets (and respective subsets divided by angular momentum subshells), “basis degree” shows the number of orbitals at each nuclear position before and after (indicated by “→”) the contraction of primitive GTOs.

we propose a two-level optimization strategy (in Algorithm 2) that optimizes both the exponent coefficients and the basis-function centers. In particular, the basis-function centers are specified by the “correlated basis centers” defined in Equation (13), which are correlated through primitive parameters ν . By such construction, these centers remain floating (non-fixed) while preserving a desired geometry when ν is being optimized. To test the performance of Algorithm 2, we first apply it to the diatomic hydrogen (H₂) molecule and compare its dissociation curve with those of correlation-consistent basis sets. We observe a consistent error across the region where the bond length exceeds 1.8 a.u. As for the region where the bond length is below 1.8 a.u., the variational even-tempered basis set of degree nine ($\tilde{G}_9^r$) produces lower errors than aug-cc-pVDZ (which has the same basis-set size) and cc-pVTZ. To verify the discretization advantage of $\tilde{G}_9^r$ over aug-cc-pVDZ on constructing proper molecular orbitals, we further compare their corresponding electron charge densities at the bond length of 0.6 a.u., and the results are shown in FIG. 7. Particularly, we show that $\tilde{G}_9^r$ more efficiently describes the charge density in the compressed and stretched bond regions than the traditional atomic Gaussian basis sets. Next, we test the performance of variational even-tempered basis sets generated by Algorithm 2 on various tetra-atomic hydrogen molecules (H₄). Based on the test results in TABLE I, we conclude

that for systems more complicated than homonuclear diatomic molecules, simply placing on even-tempered basis functions on centers that maintain the spatial symmetry of the nuclear positions is not enough to produce energy with the level of accuracy from atomic Gaussian basis sets. However, as we add more even-tempered basis functions on secondary basis-function centers (i.e., the augmented centers), the performance of the even-tempered basis set can be significantly improved through further partial variational optimization.

Now, we look back at the question in Sec. I: How efficiently can a system-oriented even-tempered basis set encode the electronic ground-state information beyond atomic systems? We have demonstrated a few cases in which even-tempered basis sets can provide a more efficient representation of molecular orbitals at the mean-field level. However, based on the results for the H₄ systems, we believe there is room for further improvement. For instance, we can try including even-tempered basis functions of angular momentum higher than the S subshell. Moreover, we would like to compare the benefits of such additions to those of the nested approaches based on S-subshell GTOs at augmented centers. Specifically, we can implement more sophisticated correlated basis centers for the augmented S-subshell GTOs and compare their contribution against directly adding even-tempered GTOs in the higher angular momentum subshells. Testing these potential modifi-

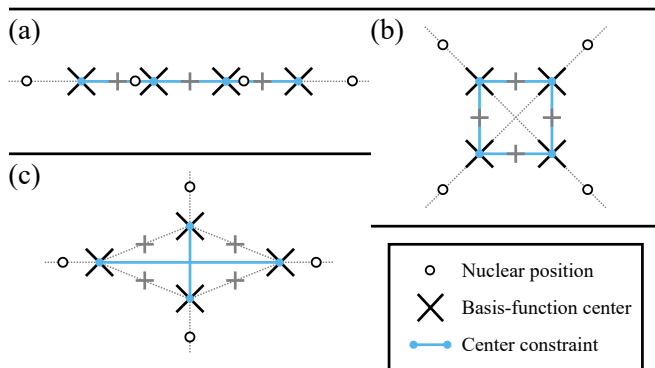


FIG. 8: Three types of molecular H_4 used to test the variational even-tempered basis sets $\tilde{G}_M^r$ generated by Algorithm 2. For each system type, the respective correlated basis center C_N , characterized by the “center constraints” (light blue edges), controls the basis-function centers of applied $\tilde{G}_M^r$. Sub-figure (a) represents the linear chain type, where the correlated basis center is constrained such that it stays in the same axis (marked as the dotted gray line) as the evenly spaced nuclear chain, shares a coincident midpoint, and has a tunable equal spacing parameter that determines the distance between adjacent basis-function centers. Sub-figure (b) represents the square planar type, where the correlated basis center is constrained such that it shares the same diagonal axes of symmetry (marked as the dotted gray lines) with the square formed by the nuclei, but has a tunable edge length parameter. Sub-figure (c) represents the rhombus type, where the correlated basis center is constrained such that the vertical (horizontal) center constraint stays in the same axis as the two vertically (horizontally) aligned nuclei and shares a midpoint, but has a separate tunable length parameter. The augmented centers used for the nested variational even-tempered basis sets are also marked by “+” in gray.

cations across a wider range of molecules can help us better understand the limitations of the current framework and improve it.

On the one hand, our work provides a minimal starting point for developing more holistic electronic-structure methods without relying on tabulated empirical data. This system-oriented generation of basis sets on the Hartree-Fock level provides a more direct connection between basis-set discretizations and the formation of molecular orbitals for many-electron systems than direct linear combinations of atomic basis sets. On the other hand, more in-depth questions also need to be answered in order to assess the practicality of using system-oriented even-tempered basis sets for electronic structure computation. Is there a more systematic way to nest multiple even-tempered basis sets based on the geometry of the target system? How should one determine the basis degree for each subset of the even-tempered basis functions before the variational optimization procedure? What is the average optimization complexity of the parameter space of the even-tempered basis sets? Admittedly, the conventional atomic Gaussian basis sets share some similar concerns, such as determining the

cardinal number (n in cc-pVnZ) of the correlation-consistent basis sets. However, compared to the system-oriented approach, they have little to no overhead of reconfiguring the construction of the final basis set. Therefore, for data-free but not parameter-free basis sets like variational even-tempered basis sets, the main challenge for their applicability within the computational chemistry community does not only depend on whether they can efficiently encode the electronic ground-state information, but also depends on how easily and consistently they can do it. In addition to further improving the upper-bound performance of system-oriented even-tempered basis sets, we hope that future work will also explore the practicality of incorporating this basis-set design framework into a more realistic working pipeline for electronic structure computation.

ACKNOWLEDGMENTS

W.W. thanks Brent Harrison for the helpful suggestions on visualizing the density differences in FIG.7. W.W. and J.D.W. acknowledge support by the U.S. Department of Energy, Office of Basic Energy Sciences, under Award DE-SC0019374. W.W. and J.D.W. were also partially supported by the ARO grant W911NF2410043. J.D.W. holds concurrent appointments at Dartmouth College and as an Amazon Visiting Academic. This paper describes work performed at Dartmouth College and is not associated with Amazon.

- ¹A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation, 2012).
- ²T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, 2013).
- ³J. C. Slater, *Physical Review* **35**, 210 (1930).
- ⁴V. Fock, *Zeitschrift für Physik* **61**, 126 (1930).
- ⁵P. Echenique and J. L. Alonso, *Molecular Physics* **105**, 3057 (2007).
- ⁶C. Edmiston and K. Ruedenberg, *Reviews of Modern Physics* **35**, 457 (1963).
- ⁷V. Magnasco and A. Perico, *The Journal of Chemical Physics* **47**, 971 (1967).
- ⁸Z. Li, H. Li, B. Suo, and W. Liu, *Accounts of chemical research* **47**, 2758 (2014).
- ⁹F. Jensen, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **3**, 273 (2013).
- ¹⁰S. F. Boys, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **200**, 542 (1950).
- ¹¹J. C. Slater, *Physical review* **36**, 57 (1930).
- ¹²W. J. Hehre, R. F. Stewart, and J. A. Pople, *The Journal of Chemical Physics* **51**, 2657 (1969).
- ¹³K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, *Journal of chemical information and modeling* **47**, 1045 (2007).
- ¹⁴C. Reeves, *The Journal of Chemical Physics* **39**, 1 (1963).
- ¹⁵C. Reeves and M. Harrison, *Journal of Chemical Physics* **39**, 11 (1963).
- ¹⁶R. C. Raffanetti, *The Journal of Chemical Physics* **59**, 5936 (1973).
- ¹⁷R. D. Bardo and K. Ruedenberg, *The Journal of Chemical Physics* **59**, 5956 (1973).
- ¹⁸R. D. Bardo and K. Ruedenberg, *The Journal of Chemical Physics* **59**, 5966 (1973).
- ¹⁹R. C. Raffanetti and K. Ruedenberg, *The Journal of Chemical Physics* **59**, 5978 (1973).
- ²⁰R. D. Bardo and K. Ruedenberg, *The Journal of Chemical Physics* **60**, 918 (1974).
- ²¹E. S. Kryachko and S. Wilson, *International journal of quantum chemistry* **93**, 112 (2003).
- ²²V. Bakken and T. Helgaker, *Theoretical Chemistry Accounts* **112**, 124 (2004).

- ²³V. Glushkov and S. Wilson, in *Recent Advances in the Theory of Chemical and Physical Systems* (Springer, 2006) pp. 107–126.
- ²⁴D. F. Feller and K. Ruedenberg, *Theoretica chimica acta* **52**, 231 (1979).
- ²⁵W. Kutzelnigg, *International journal of quantum chemistry* **51**, 447 (1994).
- ²⁶W. Kutzelnigg, in *Strategies and Applications in Quantum Chemistry: From Molecular Astrophysics to Molecular Engineering* (Springer, 1996) pp. 79–101.
- ²⁷I. Cherkes, S. Klaiman, and N. Moiseyev, *International Journal of Quantum Chemistry* **109**, 2996 (2009).
- ²⁸M. Bachmayr, H. Chen, and R. Schneider, *Numerische Mathematik* **128**, 137 (2014).
- ²⁹W. J. Morgan and R. C. Fortenberry, *Theoretical Chemistry Accounts* **134**, 47 (2015).
- ³⁰J. Koput, *Journal of Computational Chemistry* **36**, 1286 (2015).
- ³¹G. Schoendorff, in *Rare Earth Elements and Actinides: Progress in Computational Science Applications* (ACS Publications, 2021) pp. 329–346.
- ³²W. Wang and J. D. Whitfield, *Journal of Chemical Theory and Computation* **19**, 8032 (2023).
- ³³R. A. Shaw and J. G. Hill, *The Journal of Chemical Physics* **159** (2023).
- ³⁴M. Schmidt and K. Ruedenberg, *The Journal of Chemical Physics* **71**, 3951 (1979).
- ³⁵W. Klopper and W. Kutzelnigg, *Journal of Molecular Structure: THEOCHEM* **135**, 339 (1986).
- ³⁶P.-O. Löwdin, in *Advances in quantum chemistry*, Vol. 5 (Elsevier, 1970) pp. 185–199.
- ³⁷S. Lehtola, *The Journal of Chemical Physics* **151** (2019).
- ³⁸A. A. Frost, *Journal of Chemical Physics* **47**, 3707 (1967).
- ³⁹H. Huber, *Chemical Physics Letters* **62**, 95 (1979).
- ⁴⁰A. Hurley, *Journal of computational chemistry* **9**, 75 (1988).
- ⁴¹T. H. Dunning Jr, *The Journal of chemical physics* **90**, 1007 (1989).
- ⁴²R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, *The Journal of chemical physics* **96**, 6796 (1992).
- ⁴³B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus, *Journal of chemical information and modeling* **59**, 4814 (2019).
- ⁴⁴Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, *et al.*, *The Journal of chemical physics* **153** (2020).
- ⁴⁵R. Carbó, L. Leyda, and M. Arnau, *International journal of quantum chemistry* **17**, 1185 (1980).

Appendix A: Optimized basis set parameters

α	$\tilde{\beta}$	M	Energy (Ha)
1	0.393140	2	-0.44916
	0.667944	4	-0.47852
	0.797569	6	-0.48864
2	0.274151	2	-0.46806
	0.519582	4	-0.48597
	0.662166	6	-0.49235
	0.748984	8	-0.49524
4	0.204947	2	-0.48093
	0.417854	4	-0.49206
	0.562148	6	-0.49560
	0.657208	8	-0.49722
	0.722829	10	-0.49809
8	0.157000	2	-0.48574
	0.345140	4	-0.49596
	0.484849	6	-0.49773
	0.583781	8	-0.49855
	0.655322	10	-0.49900
	0.708195	12	-0.49926
16	0.119268	2	-0.48345
	0.292807	4	-0.49800
	0.424110	6	-0.49892
	0.523709	8	-0.49930
	0.597673	10	-0.49951
	0.654368	12	-0.49964
	0.698894	14	-0.49972

α	$\tilde{\beta}$	M	Energy (Ha)
32	0.088965	2	-0.47608
	0.253055	4	-0.49873
	0.375734	6	-0.49951
	0.473296	8	-0.49968
	0.548382	10	-0.49978
	0.607553	12	-0.49983
64	0.655120	14	-0.49987
	0.693002	16	-0.49990
	0.065135	2	-0.46606
	0.219475	4	-0.49847
	0.337393	6	-0.49977
	0.430413	8	-0.49986
128	0.506717	10	-0.49990
	0.565911	12	-0.49993
	0.616830	14	-0.49994
	0.655119	16	-0.49995
	0.686586	18	-0.49996
	0.046955	2	-0.45571
256	0.189911	4	-0.49732
	0.305748	6	-0.49984
	0.394849	8	-0.49994
	0.467925	10	-0.49996
	0.527660	12	-0.49997
	0.579845	14	-0.49998
	0.616841	16	-0.49998
	0.648062	18	-0.49998
	0.672647	20	-0.49999

TABLE II: The values of $\tilde{\beta}$ output by Algorithm 1 under a given α to generate the reduced variational even-tempered basis sets $\tilde{\mathcal{G}}_M^r$ for the hydrogen atom. The rightmost column lists the respective ground-state energies.

H ₂ bond length (a.u.)	α	$\tilde{\beta}$	ν (a.u.)	Energy (Ha)
0.6	0.054307	2.644041	0.552297	-2.39608
0.8	0.026985	2.788918	0.737008	-2.22986
1.0	0.013222	2.865272	0.924141	-2.08395
1.2	0.007695	3.022383	1.114314	-1.95690
1.4	0.004678	3.170136	1.307021	-1.84620
1.6	0.003054	3.081738	1.501414	-1.74944
1.8	0.002135	3.113571	1.698202	-1.66445
2.0	0.001594	3.206404	1.897016	-1.58941
2.4	0.000955	3.936617	2.299774	-1.46442
2.8	0.000745	4.052290	2.761710	-1.38011
3.2	0.000581	4.206682	3.181930	-1.32327
4.0	0.000461	3.458980	3.994340	-1.25240
5.0	0.000303	3.675462	4.998815	-1.20001

TABLE III: The values of α , $\tilde{\beta}$, and the correlated-cluster parameters (ν) output by Algorithm 2 to generate the reduced variational even-tempered basis sets $\tilde{\mathcal{G}}_9^r$ for the molecular H₂ at different bond lengths. The rightmost column lists the respective UHF energies (excluding nuclear repulsion), as the UHF functional was used as the objective in Algorithm 2.

Type	Geometry	α	$\tilde{\beta}$	ν (a.u.)	M	Energy (Ha)
Linear chain	1.2 a.u.	0.117587	2.715992	1.161747	3	−5.61693
		0.043294	3.154529	1.179023	6	−5.67733
		0.014507	3.010633	1.180780	9	−5.67814
	1.6 a.u.	0.059646	3.991194	1.563404	3	−4.86449
		0.014944	3.630591	1.576631	6	−4.88849
		0.004282	3.478584	1.578644	9	−4.88905
	2.0 a.u.	0.035272	4.865904	1.961840	3	−4.30367
		0.007249	3.903687	1.976556	6	−4.31855
		0.001929	3.771001	1.978518	9	−4.31922
Square planar	2.0 a.u.	0.032225	4.632960	1.890077	3	−4.63276
		0.006956	3.796694	1.924379	6	−4.65320
		0.001869	3.732178	1.927855	9	−4.65369
	2.4 a.u.	0.022950	4.988391	2.292533	3	−4.17920
		0.004601	3.949170	2.335201	6	−4.20025
		0.001194	3.860293	2.339247	9	−4.20099
Rhombus	(60°, 2.2 a.u.)	0.028346	4.774232	(3.589486, 1.996967)	3	−4.54378
		0.005937	3.815942	(3.655913, 2.045451)	6	−4.56171
		0.001583	3.771706	(3.661470, 2.047756)	9	−4.56200

TABLE IV: The values of α , $\tilde{\beta}$, and the correlated-cluster parameters (ν) output by Algorithm 2 to generate the direct variational even-tempered basis sets $\tilde{\mathcal{G}}_M^r$ for the molecular H_4 geometries listed in Table I. The rightmost column lists the respective RHF energies (excluding nuclear repulsions), as the RHF functional was used as the objective in Algorithm 2.

Type	Geometry	α	$\tilde{\beta}$	M	Energy (Ha)
Linear chain	1.2 a.u.	1	1.522179	1	−5.67905
		0.105112	4.235839	2	−5.67918
		0.105112	2.866763	3	−5.67931
	1.6 a.u.	1	0.695538	1	−4.89083
		0.230038	2.377820	2	−4.89126
		0.230038	1.839277	3	−4.89141
	2.0 a.u.	1	0.741479	1	−4.32096
		0.215785	1.768776	2	−4.32138
		0.215785	1.370233	3	−4.32231
Square planar	2.0 a.u.	1	0.372993	1	−4.65734
		0.107241	2.396058	2	−4.65816
		0.107241	1.764328	3	−4.65837
	2.4 a.u.	1	0.280993	1	−4.20421
		0.138664	1.639747	2	−4.20474
		0.138664	1.332234	3	−4.20538
Rhombus	(60°, 2.2 a.u.)	1	0.312303	1	−4.56993
		0.190375	1.469011	2	−4.57030
		0.190375	1.213282	3	−4.57116

TABLE V: The values of α and $\tilde{\beta}$ output by Algorithm 2 to generate the additional even-tempered basis functions located on the augmented centers (shown in FIG. 8) to construct the nested variational even-tempered basis sets based on $\tilde{\mathcal{G}}_6^r$ for the molecular H_4 geometries listed in Table I. Particularly, for basis degree $M = 1$, the value of α is always normalized to one so that the corresponding $\tilde{\beta}$ directly represents the exponent coefficient of the single S-subshell GTO at each augmented center. The rightmost column lists the respective RHF energies (excluding nuclear repulsions), as the RHF functional was used as the objective in Algorithm 2.

Appendix B: Charge densities of tetra-atomic hydrogen

This appendix includes the figures (FIGs. 9–14) of HF charge densities for all tested H_4 geometries in Sec. IV B, computed with the basis sets (and subsets) compared in TABLE I. Within each figure, the panels (i.e., the sub-figures) appear in the same order as the basis sets listed in TABLE I, and each panel is labeled by the corresponding basis abbreviation. For example, “ET-6-1” in FIG. 9 denotes the HF charge density obtained with the variational even-tempered basis set having basis-degree pair (6, 1). In each panel, the nuclear positions of the system are marked by the yellow dots.

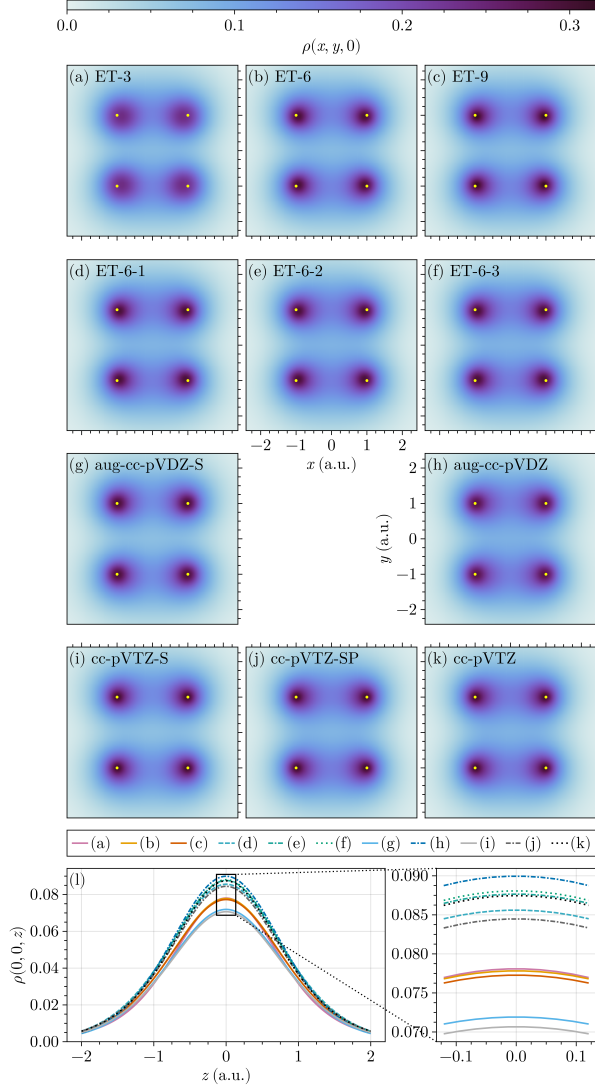


FIG. 9: Square planar (2.0 a.u.).

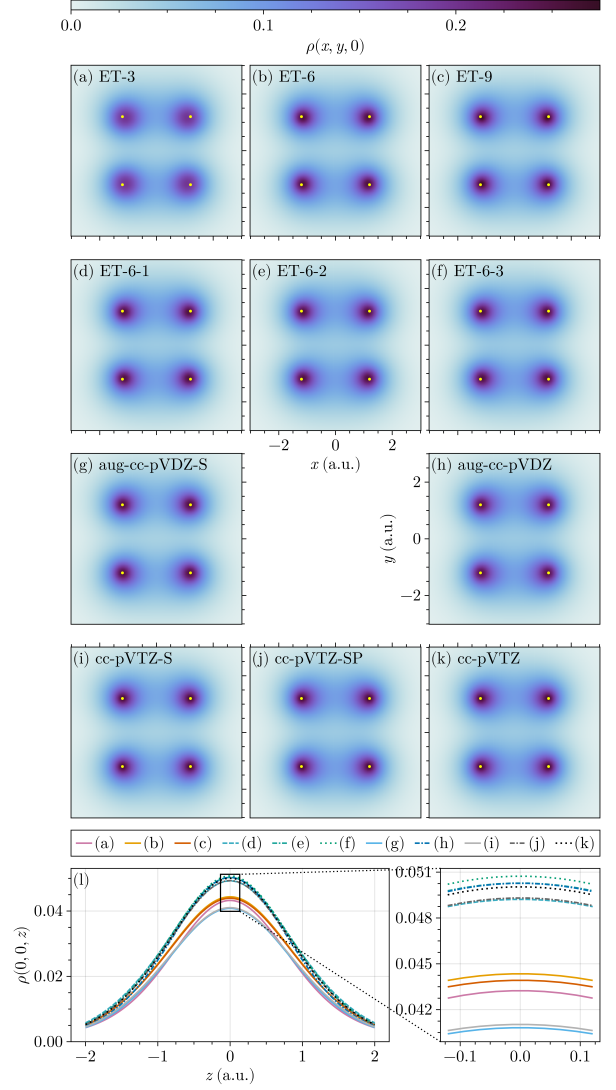


FIG. 10: Square planar (2.4 a.u.).

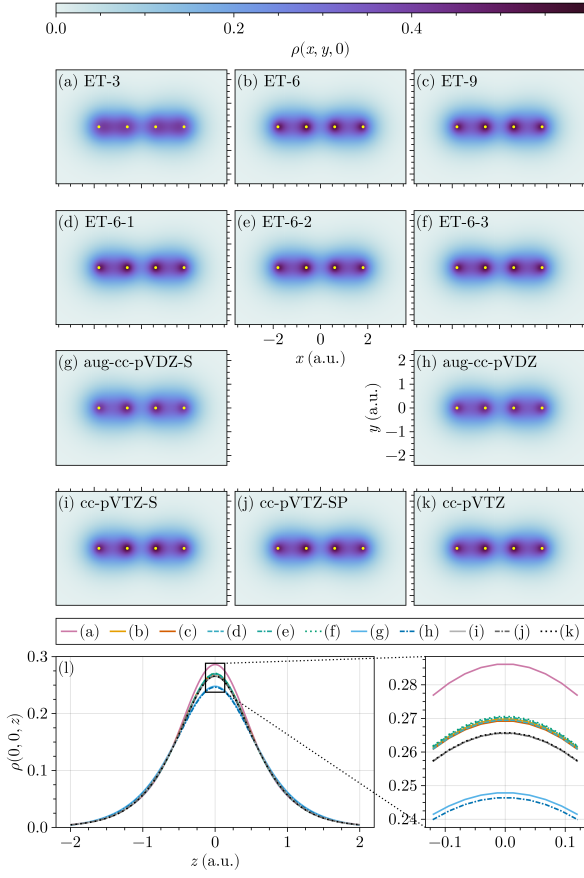


FIG. 11: Linear chain (1.2 a.u.).

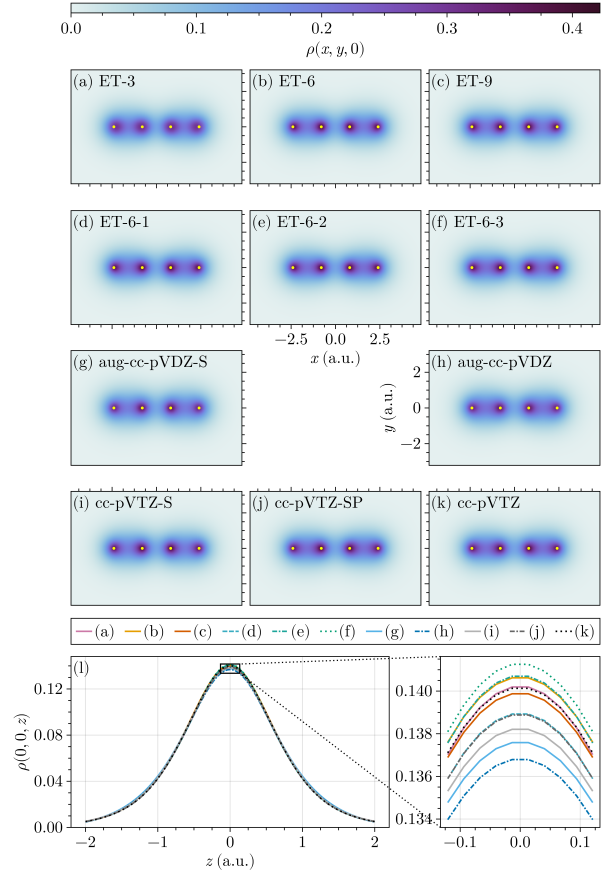


FIG. 12: Linear chain (1.6 a.u.).

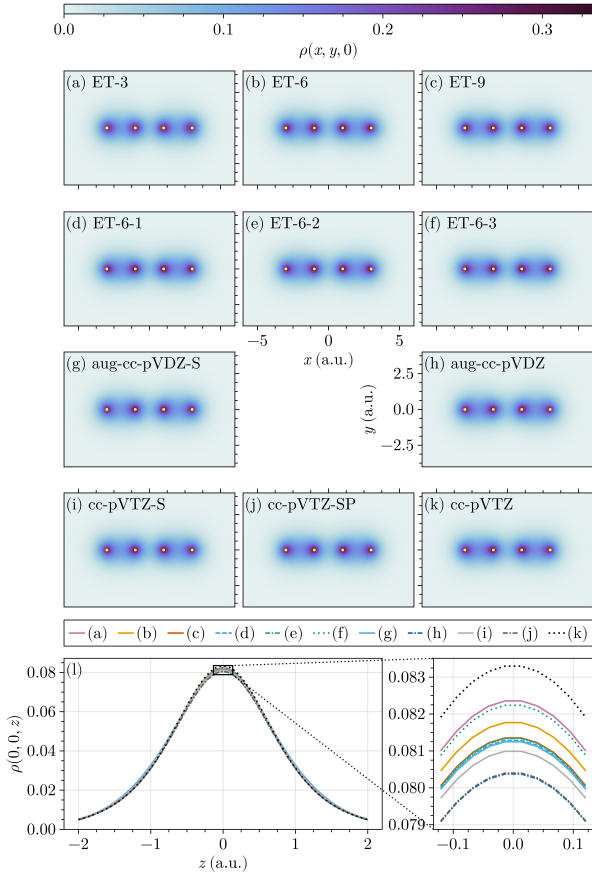


FIG. 13: Linear chain (2.0 a.u.).

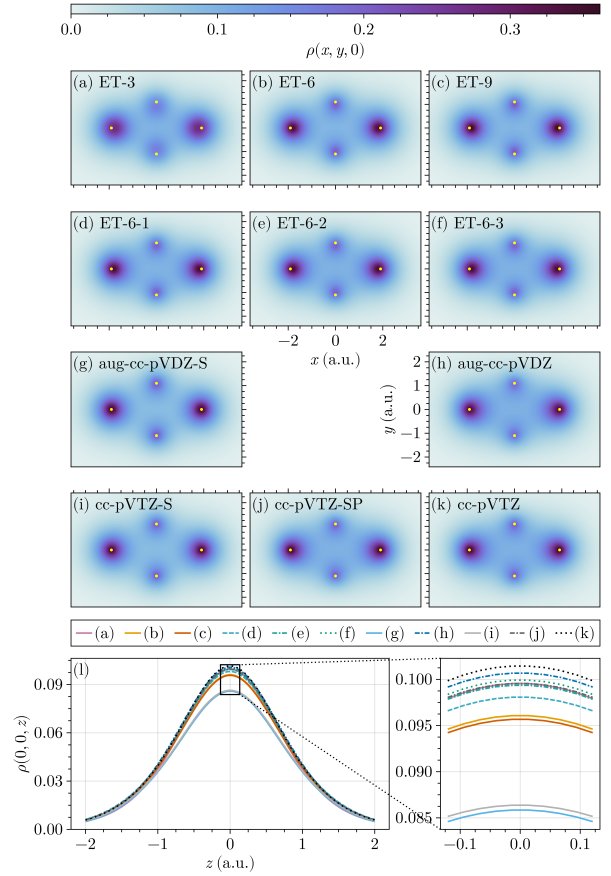


FIG. 14: Rhombus (60°, 2.2 a.u.).