

Lower bound on the representation complexity of antisymmetric tensor product functions

Yuyang Wang^{1,2}, Yukuan Hu³ & Xin Liu^{1,2,*}

¹State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;

²School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;

³Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique, École nationale des ponts et chaussées, Institut Polytechnique de Paris, Marne-la-Vallée 77455, France

Email: wangyuyang@lsec.cc.ac.cn, ykhu@lsec.cc.ac.cn, liuxin@lsec.cc.ac.cn

Received May 19, 2025; accepted January 25, 2026; published online March 12, 2026

Abstract Tensor product function (TPF) approximations have been widely adopted in solving high-dimensional problems, such as partial differential equations and eigenvalue problems, achieving desirable accuracy with computational overhead that scales linearly with problem dimensions. However, recent studies have underscored the extraordinarily high computational cost of TPFs on quantum many-body problems, even for systems with as few as three particles. A key distinction in these problems is the antisymmetry requirement on the unknown functions. In this paper, we rigorously establish that the minimum number of involved terms for a class of TPFs to be exactly antisymmetric increases exponentially with the problem dimension. This class encompasses both traditionally discretized TPFs and the recent ones parameterized by neural networks. Our proof exploits the link between the antisymmetric TPFs in this class and the corresponding antisymmetric tensors and focuses on the canonical polyadic rank of the latter. As a result, our findings reveal that low-rank TPFs are fundamentally unsuitable for high-dimensional problems where antisymmetry is essential.

Keywords tensor product function, antisymmetric wave function, antisymmetric tensor, canonical polyadic rank

MSC(2020) 15A69, 41A29, 41A63, 46B28

Citation: Wang Y Y, Hu Y K, Liu X. Lower bound on the representation complexity of antisymmetric tensor product functions. *Sci China Math*, 2026, 69, <https://doi.org/10.1007/s11425-025-2547-3>

1 Introduction

High-dimensional partial differential equations (PDEs) and eigenvalue problems frequently arise in scientific and engineering applications. In solving these problems, traditional numerical methods, such as finite difference and finite element methods, suffer from the curse of dimensionality, in that both the storage and computational costs grow exponentially with the problem dimension. For example, discretizing an N -dimensional domain with merely two grid points per dimension can result in a dense tensor of size 2^N , requiring over 1 ZB of memory for direct storage in double precision when $N \geq 70$.

* Corresponding author

To circumvent the curse of dimensionality, tensor product functions (TPFs) have been proposed to approximate the original high-dimensional ones [1, 5, 19]. For a function f defined on a Cartesian product domain $\times_{j=1}^N \Omega_j$, with $\Omega_j \subseteq \mathbb{R}^d$ ($j = 1, \dots, N$), its TPF approximation is of the form

$$f(\mathbf{r}_1, \dots, \mathbf{r}_N) \approx \sum_{i=1}^p \psi_{i1}(\mathbf{r}_1) \cdots \psi_{iN}(\mathbf{r}_N), \quad \forall \mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N) \in \times_{j=1}^N \Omega_j,$$

where $p \in \mathbb{N}$ is called the separation rank in the literature [5, 6] (see Remark 2.4 for a rigorous definition) and ψ_{ij} is a function defined on Ω_j ($i = 1, \dots, p$ and $j = 1, \dots, N$). Combined with the traditional discretization methods for $\{\psi_{ij}\}$, the TPF approximation requires a computational cost growing linearly with respect to N . The foundation of TPFs can be traced back to the work of Schmidt [47] on the case of $N = 2$ and $d = 1$ with $\psi_{ij} \in \mathcal{L}^2(\Omega_j)$, known as the Schmidt decomposition. Up to this point, this construction has been developed to handle higher-dimensional settings and function spaces with more favorable properties [4–6, 12, 17, 18, 20, 21, 40, 42].

In recent years, neural networks have gained significant attention for their powerful approximation capabilities (see [11] and the references therein for a comprehensive review) and have provided mesh-free solutions for PDEs (see, e.g., [9, 13, 29, 35, 38, 46, 48, 49, 56, 58]). Particularly in the high-dimensional regime, Wang et al. [54] introduced the tensor neural networks (TNNs) by utilizing neural networks to construct ψ_{ij} in the TPF approximation, effectively eliminating the need for high-dimensional integrals (see Subsection 2.1 for an illustration). TNN approximations have demonstrated high accuracy across various problems [27, 33, 36, 51–54]; for example, with a separation rank of $p \leq 30$, they achieved errors of $\mathcal{O}(10^{-7})$ for a 20000-dimensional Schrödinger equation with coupled quantum harmonic oscillator potential [27].

However, recent studies have highlighted the limitations of TNN approximations in addressing the electronic Schrödinger equation in quantum mechanics [34]. This problem differs from the aforementioned applications of TNN approximations in requiring the additional antisymmetry constraint on the unknown functions. Mathematically, an antisymmetric function f defined on Ω^N ($\Omega \subseteq \mathbb{R}^d$) satisfies

$$f(\mathcal{T}_{ij}(\mathbf{r}_1, \dots, \mathbf{r}_N)) = -f(\mathbf{r}_1, \dots, \mathbf{r}_N), \quad \forall i \neq j, \quad (1.1)$$

where $\mathcal{T}_{ij} : \Omega^N \mapsto \Omega^N$ swaps the i -th and j -th coordinates while keeping all the others unchanged. In the electronic Schrödinger equation, the electronic wave function describes the behavior of electrons, whose antisymmetry is grounded in the fundamental Pauli exclusion principle for fermions [44]. Liao et al. [34] and Zhou et al. [60] approximated the electronic wave function with TNNs, handled the Coulomb interaction with Legendre polynomial expansions or sum of Gaussians decompositions, and optimized the neural network parameters through a variational principle. To achieve the desired accuracy even on systems with as few as three electrons, they had to employ TNNs with $p = 50$ and undergo extraordinarily lengthy optimization processes.

The challenge can stem from the representation of antisymmetry. We first note that TPFs are not inherently antisymmetric. For example, if we set $\psi_{i1} = \psi_{i2} = \cdots = \psi_{iN}$ for each i , the resulting function is symmetric. However, antisymmetry can be crucial for applications of interest. In quantum mechanics, the antisymmetry of the wave function ensures that identical fermions cannot occupy the same quantum state. Adopting antisymmetric ansätze thus guarantees the physical meaningfulness of results, as already observed in 1930 [15, 50]. We also perform numerical comparisons between TPFs and their antisymmetrization. Figure 1 reveals that applying antisymmetrization improves both accuracy and optimization efficiency. Detailed experimental settings are provided in Appendix A. Consequently, existing studies on representing antisymmetric wave functions usually rely on determinant-based constructions, such as Slater-type wave functions [14, 15, 30, 32, 45, 50, 59], pairwise wave functions [22, 23, 43], Pfaffian wave functions [2, 3, 16], explicit antisymmetrization [37] and some implicit constructions [8, 57]. Within the framework of TPF approximation, these constructions correspond to separation ranks on the order of $N!$, rendering the computational advantages from the separability of

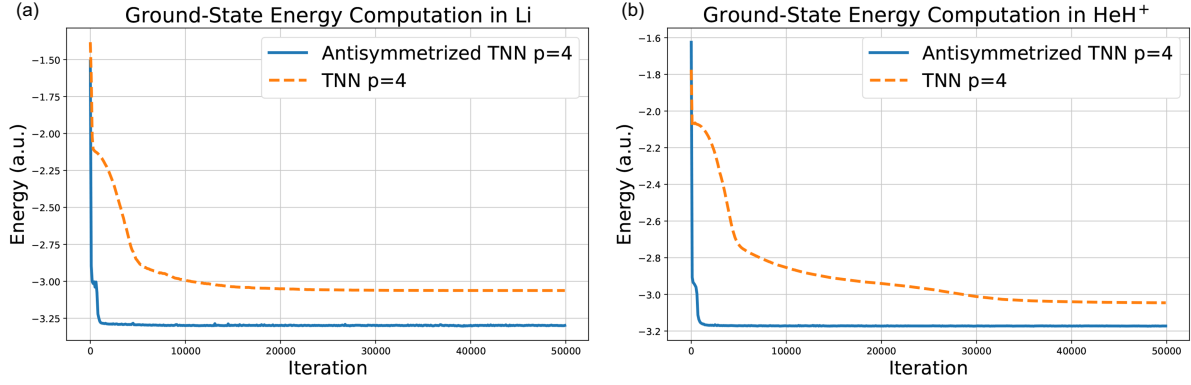


Figure 1 (Color online) Numerical comparison of the TNN approximations with and without explicit antisymmetrization on the one-dimensional systems of Li (a) and HeH⁺ (b)

TPFs inapplicable. Together, these findings motivate us to investigate the representation complexity of TPFs for antisymmetry, which refers to the minimum number of TPF terms required to represent antisymmetric functions.

Contributions. In this work, we examine a class of TPFs in which the functions $\{\psi_{ij}\}$ reside in finite-dimensional spaces. This class covers both TPFs discretized using traditional methods and the TNNs with any fixed network architecture. By linking the antisymmetric TPFs in this class to the associated antisymmetric tensors and analyzing the canonical polyadic (CP) rank of the latter, we establish an exponential lower bound of order $2^N/\sqrt{N}$ for the minimum number of terms required in the antisymmetric TPF representations. Our findings reveal that low-rank TPFs are fundamentally unsuitable for high-dimensional problems with the antisymmetry constraint, including the electronic Schrödinger equation in quantum mechanics.

Organization. The rest of this paper is organized as follows. In Section 2, we introduce fundamental concepts related to TPFs, TNNs and antisymmetric functions and tensors. In Section 3, we present our main theoretical results on the antisymmetric TPFs and their application to the antisymmetric TNNs with fixed network architectures. Finally, in Section 4, we discuss the implications of these findings for relevant applications. The experimental settings are presented in Appendix A.

Notations. Throughout this paper, scalars, vectors and tensors are denoted by lowercase letters (e.g., x), bold lowercase letters (e.g., \mathbf{x}) and uppercase bold letters (e.g., \mathbf{X}), respectively. Operators are represented by uppercase script letters (e.g., \mathcal{A}). We use subscripts to indicate components or blocks of vectors and matrices. For example, x_i denotes the i -th component of \mathbf{x} , and \mathbf{r}_i denotes the i -th block of \mathbf{r} . For an N -order tensor \mathbf{X} , the entry indexed by (k_1, \dots, k_N) is written as $\mathbf{X}(k_1, \dots, k_N)$. The support of \mathbf{X} is denoted by $\text{supp}(\mathbf{X}) := \{(k_1, \dots, k_N) : \mathbf{X}(k_1, \dots, k_N) \neq 0\}$.

We use $\mathcal{B}(\Omega)$ for the Banach space of functions defined on the domain Ω . The Cartesian products of sets are denoted by either exponents (e.g., Ω^N) or multiplication symbols (e.g., $\Omega_1 \times \Omega_2$ and $\prod_{j=1}^N \Omega_j$). We use the notation \otimes or \bigotimes to represent tensor products of vectors (e.g., $\mathbf{x}_1 \otimes \mathbf{x}_2$ and $\bigotimes_{j=1}^N \mathbf{x}_j$), vector spaces (e.g., $\mathbb{C}^{K_1} \otimes \mathbb{C}^{K_2}$ and $\bigotimes_{j=1}^N \mathbb{C}^{K_j}$), functions (e.g., $\psi_1 \otimes \psi_2$ and $\bigotimes_{j=1}^N \psi_j$) and function spaces (e.g., $\mathcal{B}(\Omega_1) \otimes \mathcal{B}(\Omega_2)$ and $\bigotimes_{j=1}^N \mathcal{B}(\Omega_j)$). Note that the tensor product of two functions is defined as $(\psi_1 \otimes \psi_2)(\mathbf{r}_1, \mathbf{r}_2) := \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)$ for any \mathbf{r}_1 and \mathbf{r}_2 . The tensor product of function spaces can be obtained in an analogous way (see Definition 2.1). The notation $\binom{K}{N}$ represents the binomial coefficient, defined as $\binom{K}{N} = K!/(N!(K-N)!)$. We define the permutation group over $\{1, \dots, N\}$ as \mathcal{S}_N . The notation $\text{sgn}(\pi)$ gives the sign of the permutation π . We use $\Theta(\cdot)$ to denote quantities that are of the same asymptotic order.

2 Preliminaries

This section presents the concepts and tools necessary for our main results. We begin by detailing the structures of TPFs, including TNNs, which provide the core frameworks for approximating high-dimensional functions. Following this, we introduce the mathematical properties of antisymmetric functions and tensors. For the broadest generality and compatibility with real applications, all functions in this paper are assumed to be complex-valued and defined on real spaces.

2.1 Tensor product function and tensor neural network

We begin by defining TPFs over general Banach spaces.

Definition 2.1 (TPF). Let $\mathcal{B}(\Omega_j)$ be a Banach space of functions defined on $\Omega_j \subseteq \mathbb{R}^d$ ($j = 1, \dots, N$). The space of TPFs on $\{\mathcal{B}(\Omega_j)\}_{j=1}^N$ is defined as

$$\bigotimes_{j=1}^N \mathcal{B}(\Omega_j) := \left\{ f \in \mathcal{B} \left(\prod_{j=1}^N \Omega_j \right) : f = \sum_{i=1}^p \bigotimes_{j=1}^N \psi_{ij}, \text{ where } p \in \mathbb{N}, \psi_{ij} \in \mathcal{B}(\Omega_j), \forall i, j \right\}, \quad (2.1)$$

and each function $f \in \bigotimes_{j=1}^N \mathcal{B}(\Omega_j)$ is called a TPF (on $\{\mathcal{B}(\Omega_j)\}_{j=1}^N$).

Remark 2.2. The notation $\bigotimes_{j=1}^N \mathcal{B}(\Omega_j)$ sometimes stands for the closure of the right-hand side of (2.1) [19, Subsection 4.2.1], ensuring completeness. Since we focus on functions with explicit tensor product structures, we omit the closure and, with a slight abuse of notation, still define the resulting space by $\bigotimes_{j=1}^N \mathcal{B}(\Omega_j)$.

Unless stated, we assume that $\Omega_1 = \dots = \Omega_N = \Omega$, and we write $\bigotimes^N \mathcal{B}(\Omega) := \bigotimes_{j=1}^N \mathcal{B}(\Omega_j)$. Since for any separable Hilbert space $\mathcal{B}(\Omega)$, $\bigotimes^N \mathcal{B}(\Omega)$ is dense in $\mathcal{B}(\Omega^N)$ for any compact $\Omega \subseteq \mathbb{R}^d$ [55, Theorem 3.12], TPFs can in principle approximate any function in $\mathcal{B}(\Omega^N)$ to arbitrary accuracy.

To facilitate our analysis, we define the rank of a TPF.

Definition 2.3 (TPF rank). The TPF rank of a function $f \in \bigotimes^N \mathcal{B}(\Omega)$, denoted by $\text{rank}_{\mathcal{B}(\Omega)}(f)$, is defined as

$$\text{rank}_{\mathcal{B}(\Omega)}(f) := \min \left\{ p \in \mathbb{N} : f = \sum_{i=1}^p \bigotimes_{j=1}^N \psi_{ij}, \text{ where } \psi_{ij} \in \mathcal{B}(\Omega), \forall i, j \right\},$$

or in other words, as the minimum number of terms for the TPF representation of f .

Remark 2.4. The definitions of the TPF rank and separation rank [4–6, 17] differ fundamentally. The separation rank of a function is typically used in the context of approximation [4–6, 17]. Given a tolerance $\varepsilon > 0$, the separation rank of a function f is defined as any $p \in \mathbb{N}$ such that there exist $\psi_{ij} \in \mathcal{B}(\Omega)$ ($i = 1, \dots, p$ and $j = 1, \dots, N$) satisfying

$$\left\| f - \sum_{i=1}^p \bigotimes_{j=1}^N \psi_{ij} \right\|_{\mathcal{B}(\Omega^N)} \leq \varepsilon.$$

The optimal (or minimal) separation rank of f then refers to its minimum separation rank for the given tolerance.

In contrast, the TPF rank is defined specifically for TPFs, capturing the minimal number of terms required for an exact TPF representation. We emphasize that the exact representation of antisymmetry is meaningful in applications of interest. For example, the antisymmetry of the wave function in quantum mechanics guarantees the physical correctness of results [15, 50]. Intuitively, it is also natural to approximate antisymmetric functions with candidates in the same class.

The TNN can be viewed as a special case of TPF with $d = 1$ and $\{\psi_{ij}\}$ parameterized by neural networks [54]. Each input $r_j \in \Omega \subseteq \mathbb{R}$ is independently processed through a fully connected subnetwork that outputs a p -dimensional vector:

$$r_j \xrightarrow{\text{the } j\text{-th fully connected subnetwork}} (\psi_{1j}(r_j; \boldsymbol{\theta}_j), \dots, \psi_{pj}(r_j; \boldsymbol{\theta}_j))^\top,$$

where θ_j represents the parameters of the j -th subnetwork. Let θ collect the parameters from N subnetworks. The TNN function f_θ is then defined as

$$f_\theta := \sum_{i=1}^p \bigotimes_{j=1}^N \psi_{ij}(\cdot; \theta_j). \quad (2.2)$$

Figure 2 illustrates the architecture of the TNN.

In the context of quantum mechanics, to approximate the electronic wave function of an isolated, non-relativistic, time-independent and spinless N -electron system in $\Omega \subseteq \mathbb{R}^d$, the TNN wave function ansatz follows a similar structure. Each electronic coordinate is represented by a d -dimensional vector $\mathbf{r}_j = (r_{j1}, \dots, r_{jd})^\top \in \mathbb{R}^d$, with each one-dimensional coordinate processed by an independent subnetwork, resulting in a total of dN independent subnetworks [34]. The TNN wave function ansatz for this system is

$$f_\theta(\mathbf{r}) := \sum_{i=1}^p \prod_{j=1}^N \left(\prod_{k=1}^d \psi_{ijk}(r_{jk}; \theta_{jk}) \right), \quad \forall \mathbf{r} \in \Omega^N, \quad (2.3)$$

where θ collects the parameters of dN subnetworks.

2.2 Antisymmetric functions

Given a domain $\Omega \subseteq \mathbb{R}^d$, the antisymmetric function f defined on Ω^N changes sign when any two inputs are exchanged, as defined in (1.1). Equivalently,

$$f(\mathbf{r}_1, \dots, \mathbf{r}_N) = \text{sgn}(\pi) f(\mathbf{r}_{\pi(1)}, \dots, \mathbf{r}_{\pi(N)}), \quad \forall \pi \in \mathcal{S}_N, \quad \mathbf{r} \in \Omega^N, \quad (2.4)$$

where \mathcal{S}_N is the permutation group on $\{1, \dots, N\}$, and $\text{sgn}(\pi)$ is the sign of permutation π . The set of all antisymmetric functions in $\bigotimes^N \mathcal{B}(\Omega)$ is denoted by

$$\mathcal{A} \left(\bigotimes^N \mathcal{B}(\Omega) \right) := \left\{ f \in \bigotimes^N \mathcal{B}(\Omega) : f \text{ satisfies (2.4)} \right\}.$$

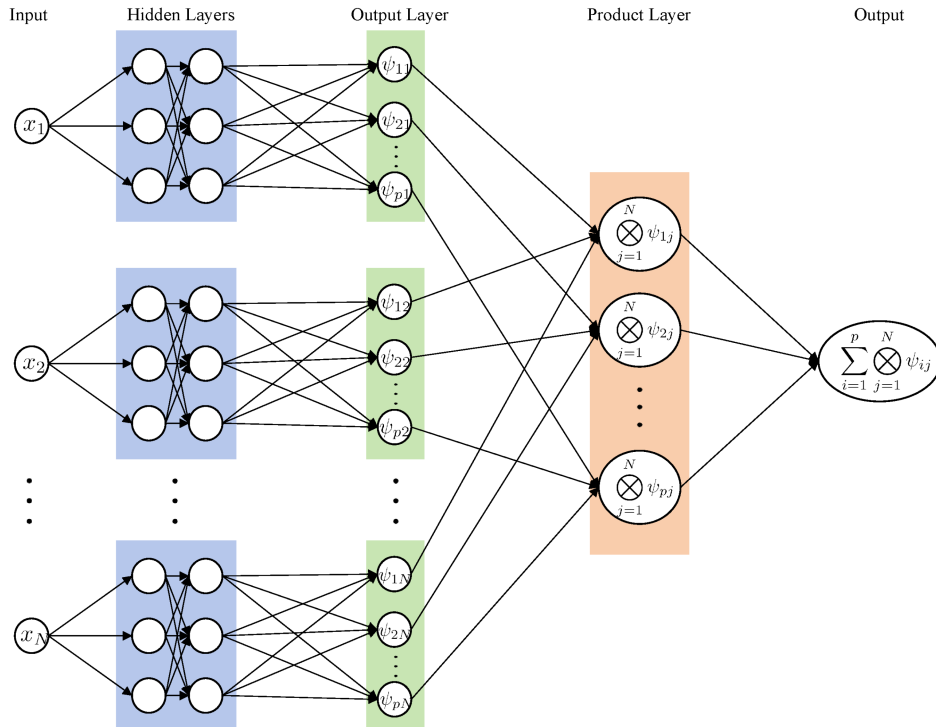


Figure 2 (Color online) Architecture of the TNN

The antisymmetrizer (also known as the antisymmetrizing operator) [39] on Ω^N , denoted by \mathcal{A} , is a linear operator that antisymmetrizes functions. For any f defined on Ω^N , the action of \mathcal{A} on f is defined as

$$\mathcal{A}[f](\mathbf{r}_1, \dots, \mathbf{r}_N) := \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) f(\mathbf{r}_{\pi(1)}, \dots, \mathbf{r}_{\pi(N)}), \quad \forall \mathbf{r} \in \Omega^N. \quad (2.5)$$

For a function $f \in \otimes^N \mathcal{B}(\Omega)$, it belongs to $\mathcal{A}(\otimes^N \mathcal{B}(\Omega))$ if and only if $f = \mathcal{A}[f]$. When $\otimes^N \mathcal{B}(\Omega)$ is a Hilbert space, the antisymmetrizer \mathcal{A} acts as the orthogonal projection operator onto $\mathcal{A}(\otimes^N \mathcal{B}(\Omega))$.

Remark 2.5. We revisit the TPF in (2.1). When $p = 1$, it reduces to the Hartree product [24], an early wave function ansatz in the quantum mechanics literature that lacks antisymmetry. When $p = N!$ with $\psi_{i1} = \text{sign}(\pi_i) \psi_{\pi_i(1)}$ and $\psi_{ij} = \psi_{\pi_i(j)}$ for $1 < j \leq N$, where $\pi_i \in \mathcal{S}_N$ for $i = 1, \dots, N!$, the TPF becomes the Slater determinant [15, 50], a well-known wave function ansatz or ingredient in various Hartree-Fock-based approximations for electronic structure calculations [25, Chapter 5]. If antisymmetrized as defined in (2.5), the TPF becomes a sum of p Slater determinants, whose expansion contains at most $pN!$ terms, and thus the TPF rank should satisfy $1 < \text{rank}_{\mathcal{B}(\Omega)}(\mathcal{A}[f]) \leq pN!$. Note that this seemingly daunting exponential rank does not necessarily imply an unacceptably high computational cost; the Slater determinant can be implemented in polynomial time due to its algebraic structure.

2.3 Antisymmetric tensors

The space of N -order (K_1, \dots, K_N) -dimensional tensors, denoted by $\otimes_{j=1}^N \mathbb{C}^{K_j}$, is defined as

$$\otimes_{j=1}^N \mathbb{C}^{K_j} := \mathbb{C}^{K_1} \otimes \dots \otimes \mathbb{C}^{K_N}.$$

In the special case where $K_1 = \dots = K_N = K$, we define this space as $\otimes^N \mathbb{C}^K$. Since the TPF rank is highly relevant to the CP rank of a tensor (as will be shown later), we provide the formal definition of the latter as follows.

Definition 2.6 (CP rank [26]). The CP rank of a tensor $\mathbf{X} \in \otimes_{j=1}^N \mathbb{C}^{K_j}$, denoted by $\text{rank}_{\text{cp}}(\mathbf{X})$, is defined as

$$\text{rank}_{\text{cp}}(\mathbf{X}) := \min \left\{ p \in \mathbb{N} : \mathbf{X} = \sum_{i=1}^p \mathbf{x}_{i1} \otimes \dots \otimes \mathbf{x}_{iN}, \mathbf{x}_{ij} \in \mathbb{C}^{K_j}, 1 \leq i \leq p, 1 \leq j \leq N \right\}.$$

In analogy to the antisymmetric function, an antisymmetric tensor changes sign with the exchange of any two indices. Following this, we define the closed subspace of $\otimes^N \mathbb{C}^K$ of all antisymmetric tensors as

$$\mathcal{A}\left(\otimes^N \mathbb{C}^K\right) := \left\{ \mathbf{X} \in \otimes^N \mathbb{C}^K : \mathbf{X}(k_1, \dots, k_N) = \text{sgn}(\pi) \mathbf{X}(k_{\pi(1)}, \dots, k_{\pi(N)}), \forall \pi \in \mathcal{S}_N \right\}.$$

The unique orthogonal projection operator onto $\mathcal{A}(\otimes^N \mathbb{C}^K)$, with a slight abuse of notation, is also denoted by \mathcal{A} : For any N -order tensor \mathbf{X} ,

$$\mathcal{A}[\mathbf{X}](k_1, \dots, k_N) := \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \mathbf{X}(k_{\pi(1)}, \dots, k_{\pi(N)}).$$

If a tensor \mathbf{X} is of a tensor product form $\mathbf{X} = \sum_{i=1}^p \mathbf{x}_{i1} \otimes \dots \otimes \mathbf{x}_{iN}$, then its projection onto $\mathcal{A}(\otimes^N \mathbb{C}^K)$ can be directly calculated by

$$\mathcal{A}[\mathbf{X}] = \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \sum_{i=1}^p \text{sgn}(\pi) \mathbf{x}_{i\pi(1)} \otimes \dots \otimes \mathbf{x}_{i\pi(N)}.$$

The following proposition gathers some key properties of antisymmetric tensors.

Proposition 2.7 (See [19, Subsection 3.5.1]). *Let $\mathbf{X} \in \mathcal{A}(\otimes^N \mathbb{C}^K)$.*

- *If there exist $i \neq j$ such that $k_i = k_j$, then $\mathbf{X}(k_1, \dots, k_N) = 0$.*
- *If $\mathcal{A}(\otimes^N \mathbb{C}^K) \neq \{0\}$, then $K \geq N$.*
- *If $K \geq N$, then*

$$\dim \mathcal{A}\left(\otimes^N \mathbb{C}^K\right) = \binom{K}{N}.$$

To write out the basis of $\mathcal{A}(\otimes^N \mathbb{C}^K)$, we introduce the multi-index set

$$\Lambda := \{\mathbf{k} = (k_1, \dots, k_N) : 1 \leq k_1 < \dots < k_N \leq K\}. \quad (2.6)$$

Notice that $|\Lambda| = \binom{K}{N}$, and the entries of an antisymmetric tensor \mathbf{X} indexed by the multi-indices in Λ are independent. For each $\mathbf{k} = (k_1, \dots, k_N) \in \Lambda$, we define the corresponding basis tensor $\mathbf{E}_{\mathbf{k}} \in \mathcal{A}(\otimes^N \mathbb{C}^K)$ as

$$\mathbf{E}_{\mathbf{k}} := \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \mathbf{e}_{k_{\pi(1)}} \otimes \dots \otimes \mathbf{e}_{k_{\pi(N)}}, \quad (2.7)$$

where $\mathbf{e}_j \in \mathbb{C}^K$ is the j -th standard unit vector in \mathbb{C}^K . By the definition of $\mathbf{E}_{\mathbf{k}}$, we have

$$\mathbf{E}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = \begin{cases} 1, & \text{if } \{\ell_1, \dots, \ell_N\} \text{ is an even permutation of } \{k_1, \dots, k_N\}, \\ -1, & \text{if } \{\ell_1, \dots, \ell_N\} \text{ is an odd permutation of } \{k_1, \dots, k_N\}, \\ 0, & \text{otherwise.} \end{cases} \quad (2.8)$$

Let $c_{\mathbf{k}} := \mathbf{X}(k_1, \dots, k_N)$ for any $\mathbf{k} \in \Lambda$. The expansion of $\mathbf{X} \in \mathcal{A}(\otimes^N \mathbb{C}^K)$ in the basis $\{\mathbf{E}_{\mathbf{k}}\}_{\mathbf{k} \in \Lambda}$ is

$$\mathbf{X} = \sum_{\mathbf{k} \in \Lambda} c_{\mathbf{k}} \mathbf{E}_{\mathbf{k}}. \quad (2.9)$$

When $K = N$, the multi-index set Λ contains only one element, and thus there is only one basis tensor, denoted by

$$\mathbf{E} := \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \mathbf{e}_{\pi(1)} \otimes \dots \otimes \mathbf{e}_{\pi(N)}. \quad (2.10)$$

This tensor is also known as the determinant tensor [10].

3 Main results

We state and rigorously prove our main theoretical results in this section. In particular, we focus on the finite-dimensional $\otimes^N \mathcal{B}(\Omega)$ and establish an exponential lower bound on the TPF rank of any nonzero antisymmetric function therein. This result is further applied to antisymmetric TNNs.

3.1 Representation complexity of antisymmetric tensor product functions in finite-dimensional spaces

We consider the situation $\mathcal{B}(\Omega) = \mathcal{F}_K(\Omega)$, where $\mathcal{F}_K(\Omega)$ is a K -dimensional space spanned by the linearly independent basis functions $\{\varphi_k\}_{k=1}^K$ defined on Ω , i.e.,

$$\mathcal{F}_K(\Omega) := \text{span}\{\varphi_1, \dots, \varphi_K\}.$$

Remark 3.1. It should be remarked that the finite-dimensional setting described above aligns with common practice. Typically, the functions $\{\psi_{ij}\}$ in TPFs are first parameterized using a set of basis functions, followed by the implementation of a finite-dimensional discretized version [1]. Moreover, the results established in this finite-dimensional setting are also applicable to the TNNs parameterized by neural networks, as we demonstrate in the next subsection.

Before giving the main results, we provide the following lemma on the linear independence of TPFs.

Lemma 3.2. For any $1 \leq k_1, \dots, k_N \leq K$, let the function $\Phi_{N, \mathbf{k}} (\mathbf{k} := (k_1, \dots, k_N))$ be defined as $\Phi_{N, \mathbf{k}} := \bigotimes_{j=1}^N \varphi_{k_j}$. Then $\{\Phi_{N, \mathbf{k}}\}_{\mathbf{k}}$ are linearly independent.

Proof. We prove by mathematical induction on N . For the case of $N = 1$, the conclusion holds trivially since $\{\varphi_k\}_{k=1}^K$ are already linearly independent. Now, assume that the conclusion holds for the case of $N = m - 1$, i.e., the functions $\{\Phi_{m-1, \mathbf{k}}\}_{\mathbf{k}}$ are linearly independent. Consider the case of $N = m$.

Suppose that there exist constants $\{c_{k_1, \dots, k_m}\}$ such that

$$\sum_{k_1=1}^K \cdots \sum_{k_m=1}^K c_{k_1, \dots, k_m} \bigotimes_{j=1}^m \varphi_{k_j} = 0.$$

By rearranging terms, we obtain

$$\sum_{k_1=1}^K \cdots \sum_{k_{m-1}=1}^K \left(\bigotimes_{j=1}^{m-1} \varphi_{k_j} \right) \otimes \left(\sum_{k_m=1}^K c_{k_1, \dots, k_m} \varphi_{k_m} \right) = 0.$$

Therefore, for any $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_m) \in \Omega^m$, we have

$$\sum_{k_1=1}^K \cdots \sum_{k_{m-1}=1}^K \left(\sum_{k_m=1}^K c_{k_1, \dots, k_m} \varphi_{k_m}(\mathbf{x}_m) \right) \prod_{j=1}^{m-1} \varphi_{k_j}(\mathbf{x}_j) = 0.$$

By induction, we conclude that $\sum_{k_m=1}^K c_{k_1, \dots, k_m} \varphi_{k_m}(\mathbf{x}_m) = 0, \forall 1 \leq k_1, \dots, k_{m-1} \leq K, \mathbf{x}_m \in \Omega$. Since $\{\varphi_k\}_{k=1}^K$ are linearly independent, it follows that $c_{k_1, \dots, k_m} = 0$ for any $1 \leq k_1, \dots, k_m \leq K$, which implies that the conclusion also holds for the case of $N = m$. The proof is complete. \square

For any $f \in \bigotimes^N \mathcal{F}_K(\Omega)$, we can identify its TPF rank as the CP rank of an N -order tensor, as established by the following theorem.

Theorem 3.3. For any $f \in \bigotimes^N \mathcal{F}_K(\Omega)$, there exists a tensor $\mathbf{X} \in \bigotimes^N \mathbb{C}^K$ such that

$$\text{rank}_{\mathcal{F}_K(\Omega)}(f) = \text{rank}_{\text{cp}}(\mathbf{X}).$$

Proof. Let $p := \text{rank}_{\mathcal{F}_K(\Omega)}(f)$. Then there exist functions $\{\psi_{ij}\} \subseteq \mathcal{F}_K(\Omega)$ such that f can be written in the form $f = \sum_{i=1}^p \bigotimes_{j=1}^N \psi_{ij}$. Expanding each ψ_{ij} in terms of the basis functions in $\mathcal{F}_K(\Omega)$, we obtain

$$\psi_{ij} = \sum_{k=1}^K c_{ijk} \varphi_k,$$

where $\{c_{ijk}\} \subseteq \mathbb{C}$ are the expansion coefficients. Thus, the TPF f satisfies

$$\begin{aligned} f(\mathbf{r}) &= \sum_{i=1}^p \prod_{j=1}^N \sum_{k=1}^K c_{ijk} \varphi_k(\mathbf{r}_j) = \sum_{i=1}^p \sum_{1 \leq k_1, \dots, k_N \leq K} \prod_{j=1}^N c_{ijk_j} \varphi_{k_j}(\mathbf{r}_j) \\ &= \sum_{1 \leq k_1, \dots, k_N \leq K} \left(\sum_{i=1}^p \prod_{j=1}^N c_{ijk_j} \right) \varphi_{k_1}(\mathbf{r}_1) \cdots \varphi_{k_N}(\mathbf{r}_N) \end{aligned} \quad (3.1)$$

for any $\mathbf{r} \in \Omega^N$.

Define a tensor $\mathbf{X} \in \bigotimes^N \mathbb{C}^K$ by $\mathbf{X}(k_1, \dots, k_N) := \sum_{i=1}^p \prod_{j=1}^N c_{ijk_j}$, and let $\mathbf{x}_{ij} = (c_{ij1}, \dots, c_{ijK})^\top \in \mathbb{C}^K$ for $1 \leq i \leq p$ and $1 \leq j \leq N$. Then

$$\mathbf{X} = \sum_{i=1}^p \mathbf{x}_{i1} \otimes \cdots \otimes \mathbf{x}_{iN}. \quad (3.2)$$

Therefore, the CP rank of \mathbf{X} is at most p , and it follows that $\text{rank}_{\mathcal{F}_K(\Omega)}(f) = p \geq \text{rank}_{\text{cp}}(\mathbf{X})$.

On the other hand, let $r := \text{rank}_{\text{cp}}(\mathbf{X})$, and define the CP decomposition of \mathbf{X} as

$$\mathbf{X} = \sum_{i=1}^r \tilde{\mathbf{x}}_{i1} \otimes \cdots \otimes \tilde{\mathbf{x}}_{iN},$$

where $\tilde{\mathbf{x}}_{ij} = (\tilde{c}_{ij1}, \dots, \tilde{c}_{ijK})^\top \in \mathbb{C}^K$ for $1 \leq i \leq p$ and $1 \leq j \leq N$. Similar to (3.1), we have

$$f(\mathbf{r}) = \sum_{i=1}^r \prod_{j=1}^N \sum_{k=1}^K \tilde{c}_{ijk} \varphi_k(\mathbf{r}_j).$$

Let $\tilde{\psi}_{ij} := \sum_{k=1}^K \tilde{c}_{ijk} \varphi_k$; we can write $f = \sum_{i=1}^r \bigotimes_{j=1}^N \tilde{\psi}_{ij}$, implying that $r = \text{rank}_{\text{cp}}(\mathbf{X}) \geq \text{rank}_{\mathcal{F}_K(\Omega)}(f)$, which completes the proof. \square

We then concentrate on antisymmetric TPFs.

Corollary 3.4. Assume that $K \geq N$ and $f \in \mathcal{A}(\bigotimes^N \mathcal{F}_K(\Omega)) \setminus \{0\}$. Then the TPF rank of f satisfies

$$\text{rank}_{\mathcal{F}_K(\Omega)}(f) \geq \min_{\mathbf{X} \in \mathcal{A}(\bigotimes^N \mathbb{C}^K) \setminus \{\mathbf{0}\}} \text{rank}_{\text{cp}}(\mathbf{X}).$$

Proof. Assume f has the expansion as in (3.1), and define the corresponding tensor \mathbf{X} as in (3.2). Applying the antisymmetrizer to f , we obtain

$$\begin{aligned} \mathcal{A}[f](\mathbf{r}) &= \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) \sum_{i=1}^p \prod_{j=1}^N \sum_{k=1}^K c_{ijk} \varphi_k(\mathbf{r}_{\tau(j)}) \\ &\quad (\text{let } \pi = \tau^{-1} \text{ and } \ell = \tau(j), \text{ and thus } j = \pi(\ell)) \\ &= \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \sum_{i=1}^p \prod_{\ell=1}^N \sum_{k=1}^K c_{i\pi(\ell)k} \varphi_k(\mathbf{r}_\ell) \\ &= \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \sum_{1 \leq k_1, \dots, k_N \leq K} \left(\sum_{i=1}^p \prod_{j=1}^N c_{i\pi(j)k_j} \right) \varphi_{k_1}(\mathbf{r}_1) \cdots \varphi_{k_N}(\mathbf{r}_N) \\ &= \sum_{1 \leq k_1, \dots, k_N \leq K} \left(\frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \text{sgn}(\pi) \left(\sum_{i=1}^p \prod_{j=1}^N c_{i\pi(j)k_j} \right) \right) \varphi_{k_1}(\mathbf{r}_1) \cdots \varphi_{k_N}(\mathbf{r}_N) \\ &\quad (\text{let } \tau = \pi^{-1} \text{ and } \ell = \pi(j), \text{ and thus } j = \tau(\ell)) \\ &= \sum_{1 \leq k_1, \dots, k_N \leq K} \left(\frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) \left(\sum_{i=1}^p \prod_{\ell=1}^N c_{i\ell k_{\tau(\ell)}} \right) \right) \varphi_{k_1}(\mathbf{r}_1) \cdots \varphi_{k_N}(\mathbf{r}_N) \\ &= \sum_{1 \leq k_1, \dots, k_N \leq K} \left(\frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) \mathbf{X}(k_{\tau(1)}, \dots, k_{\tau(N)}) \right) \varphi_{k_1}(\mathbf{r}_1) \cdots \varphi_{k_N}(\mathbf{r}_N) \end{aligned} \quad (3.3)$$

for any $\mathbf{r} \in \Omega^N$. Since f is antisymmetric, we have $f = \mathcal{A}[f]$. Noticing that $\{\varphi_k\}_{k=1}^K$ are linearly independent, we conclude after comparing (3.1) and (3.3) and using Lemma 3.2 that the tensor \mathbf{X} satisfies

$$\begin{cases} \mathbf{X}(k_1, \dots, k_N) = \frac{1}{N!} \sum_{\tau \in \mathcal{S}_N} \text{sgn}(\tau) \mathbf{X}(k_{\tau(1)}, \dots, k_{\tau(N)}), & \forall 1 \leq k_1, \dots, k_N \leq K, \\ \exists k_1, \dots, k_N, \text{ s.t. } \mathbf{X}(k_1, \dots, k_N) \neq 0. \end{cases} \quad (3.4)$$

Thus, (3.4) implies that $\mathbf{X} \in \mathcal{A}(\bigotimes^N \mathbb{C}^K)$ and $\mathbf{X} \neq 0$. By Theorem 3.3, it follows that

$$\text{rank}_{\mathcal{F}_K(\Omega)}(f) = \text{rank}_{\text{cp}}(\mathbf{X}) \geq \min_{\mathbf{Y} \in \mathcal{A}(\bigotimes^N \mathbb{C}^K) \setminus \{\mathbf{0}\}} \text{rank}_{\text{cp}}(\mathbf{Y}),$$

which completes the proof. \square

Remark 3.5. The condition $K \geq N$ is necessary to avoid the trivial case. If $K < N$, by the basic property of antisymmetric tensors (see Proposition 2.7), it must hold that $f \equiv 0$.

This corollary establishes that the TPF rank of any $f \in \mathcal{A}(\otimes^N \mathcal{F}_K(\Omega)) \setminus \{0\}$ is not smaller than the lowest CP rank of nonzero tensors in $\mathcal{A}(\otimes^N \mathbb{C}^K)$. To estimate this lowest CP rank, we turn to analyze the CP ranks of the basis tensors $\{\mathbf{E}_{\mathbf{k}}\}_{\mathbf{k} \in \Lambda}$, as suggested by (2.9). For this purpose, we introduce an existing lower bound of $\text{rank}_{\text{cp}}(\mathbf{E})$, where \mathbf{E} is the determinant tensor in (2.10).

Lemma 3.6 (See [10]). *Let $\mathbf{E} \in \mathcal{A}(\otimes^N \mathbb{C}^N)$ be the determinant tensor defined in (2.10). Then*

$$\binom{N}{\lfloor \frac{N}{2} \rfloor} \leq \text{rank}_{\text{cp}}(\mathbf{E}) \leq N! \cdot \left(\frac{5}{6}\right)^{\lfloor \frac{N}{3} \rfloor}.$$

The antisymmetric basis tensors $\{\mathbf{E}_{\mathbf{k}}\}_{\mathbf{k} \in \Lambda}$ defined in (2.7) generalize the determinant tensor \mathbf{E} to higher dimensions. We now demonstrate that their CP ranks are identical.

Theorem 3.7. *Assume that $K \geq N$. Let Λ be the multi-index set defined in (2.6). For any multi-index $\mathbf{k} \in \Lambda$, $\text{rank}_{\text{cp}}(\mathbf{E}_{\mathbf{k}}) = \text{rank}_{\text{cp}}(\mathbf{E})$.*

Proof. Let $\mathbf{k} := (k_1, \dots, k_N)$, $p := \text{rank}_{\text{cp}}(\mathbf{E})$ and $p_{\mathbf{k}} := \text{rank}_{\text{cp}}(\mathbf{E}_{\mathbf{k}})$. By the definition of the CP rank, there exist vectors $\mathbf{a}_{ij} \in \mathbb{C}^N$, $j = 1, \dots, N$, $i = 1, \dots, p$, such that $\mathbf{E} = \sum_{i=1}^p \mathbf{a}_{i1} \otimes \dots \otimes \mathbf{a}_{iN}$. We extend each \mathbf{a}_{ij} to \mathbb{C}^K by adding zero entries and define the resulting vector by $\tilde{\mathbf{a}}_{ij} \in \mathbb{C}^K$, which satisfies that the ℓ -th entry of \mathbf{a}_{ij} is the k_ℓ -th entry in $\tilde{\mathbf{a}}_{ij}$. Based on $\{\tilde{\mathbf{a}}_{ij}\}$, we define

$$\mathbf{B}_{\mathbf{k}} := \sum_{i=1}^p \tilde{\mathbf{a}}_{i1} \otimes \dots \otimes \tilde{\mathbf{a}}_{iN} \in \otimes^N \mathbb{C}^K. \quad (3.5)$$

In the following, we discuss the value of $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N)$ with $1 \leq \ell_1, \dots, \ell_N \leq K$ and finally conclude that $\mathbf{B}_{\mathbf{k}} = \mathbf{E}_{\mathbf{k}}$.

Case I. There exists an $\ell_j \notin \{k_1, \dots, k_N\}$. By the definition of $\tilde{\mathbf{a}}_{ij}$, we have $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = 0$.

Case II. For any $j \in \{1, \dots, N\}$, $\ell_j \in \{k_1, \dots, k_N\}$.

- If there exist $j \neq m$ such that $\ell_j = \ell_m$, by the definitions of $\tilde{\mathbf{a}}_{ij}$ and \mathbf{E} , $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = 0$.
- If $\ell_j \neq \ell_m$ for any $j \neq m$, then there exists a permutation $\pi \in \mathcal{S}_N$ such that $\ell_j = k_{\pi(j)}$ for $j = 1, \dots, N$. We thus have $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = \mathbf{E}(\pi(1), \dots, \pi(N))$. By the definition of \mathbf{E} , if π is even, $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = 1$; otherwise, $\mathbf{B}_{\mathbf{k}}(\ell_1, \dots, \ell_N) = -1$.

By the above two cases and the definition of $\mathbf{E}_{\mathbf{k}}$ in (2.8), we have $\mathbf{E}_{\mathbf{k}} = \mathbf{B}_{\mathbf{k}}$. From (3.5) and the definition of the CP rank, $p_{\mathbf{k}} \leq p$.

Conversely, there exist vectors $\tilde{\mathbf{b}}_{ij} \in \mathbb{C}^K$, $j = 1, \dots, N$, $i = 1, \dots, p$, such that

$$\mathbf{E}_{\mathbf{k}} = \sum_{i=1}^{p_{\mathbf{k}}} \tilde{\mathbf{b}}_{i1} \otimes \dots \otimes \tilde{\mathbf{b}}_{iN}.$$

Let $\mathbf{b}_{ij} \in \mathbb{C}^N$ be the restriction of $\tilde{\mathbf{b}}_{ij}$ to the entries indexed by \mathbf{k} , i.e., the ℓ -th entry of \mathbf{b}_{ij} is the k_ℓ -th entry of $\tilde{\mathbf{b}}_{ij}$ ($\ell = 1, \dots, N$). Similar arguments then yield

$$\mathbf{E} = \sum_{i=1}^{p_{\mathbf{k}}} \mathbf{b}_{i1} \otimes \dots \otimes \mathbf{b}_{iN},$$

implying $p \leq p_{\mathbf{k}}$. Therefore, $p = p_{\mathbf{k}}$ as desired. \square

With the above theorem, we obtain the following inequalities for the CP rank of any nonzero antisymmetric tensor.

Corollary 3.8. *Assume that $K \geq N$. Let $\mathbf{X} \in \mathcal{A}(\otimes^N \mathbb{C}^K) \setminus \{0\}$. Then*

$$\binom{N}{\lfloor \frac{N}{2} \rfloor} \leq \text{rank}_{\text{cp}}(\mathbf{X}) \leq N! \cdot \binom{K}{N} \cdot \left(\frac{5}{6}\right)^{\lfloor \frac{N}{3} \rfloor}.$$

Proof. By using the expansion of the antisymmetric tensor \mathbf{X} in (2.9), the definition of the CP rank, Lemma 3.6 and Theorem 3.7, we obtain

$$\text{rank}_{\text{cp}}(\mathbf{X}) \leq |\Lambda| \cdot \text{rank}_{\text{cp}}(\mathbf{E}_k) \leq N! \cdot \binom{K}{N} \cdot \left(\frac{5}{6}\right)^{\lfloor \frac{N}{3} \rfloor}.$$

Now, consider the other side. Choose $\mathbf{u} := (k_1, \dots, k_N) \in \Lambda$ that satisfies $c_{\mathbf{u}} := \mathbf{X}(k_1, \dots, k_N) \neq 0$. Define the truncated tensor $\tilde{\mathbf{X}}$ of \mathbf{X} by

$$\tilde{\mathbf{X}}(\ell_1, \dots, \ell_N) := \begin{cases} \mathbf{X}(\ell_1, \dots, \ell_N), & \text{if } \ell_i \in \{k_1, \dots, k_N\} \text{ for } 1 \leq i \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

It is straightforward to verify from (2.8) that

$$\text{supp}(\tilde{\mathbf{X}}) = \text{supp}(\mathbf{E}_{\mathbf{u}}) \quad \text{and} \quad \text{supp}(\mathbf{E}_{\mathbf{u}'}) \cap \text{supp}(\mathbf{E}_{\mathbf{u}''}) = \emptyset, \quad \forall \mathbf{u}', \mathbf{u}'' \in \Lambda, \quad \mathbf{u}' \neq \mathbf{u}''.$$

Therefore, from the expansion in (2.9), we see that $\tilde{\mathbf{X}} = c_{\mathbf{u}} \mathbf{E}_{\mathbf{u}}$. Let $p := \text{rank}_{\text{cp}}(\mathbf{X})$. The CP decomposition of \mathbf{X} then takes the form

$$\mathbf{X} = \sum_{i=1}^p \mathbf{x}_{i1} \otimes \cdots \otimes \mathbf{x}_{iN},$$

where $\mathbf{x}_{ij} \in \mathbb{C}^K$ for $i = 1, \dots, p$ and $j = 1, \dots, N$. By using similar arguments as in the proof of Theorem 3.7, we have that

$$\tilde{\mathbf{X}} = \sum_{i=1}^p \tilde{\mathbf{x}}_{i1} \otimes \cdots \otimes \tilde{\mathbf{x}}_{iN},$$

where $\tilde{\mathbf{x}}_{ij} \in \mathbb{C}^K$ shares the (k_1, \dots, k_N) -th entry with \mathbf{x}_{ij} while letting others be zero. By the definition of the CP rank, Lemma 3.6, and Theorem 3.7, it follows that

$$p = \text{rank}_{\text{cp}}(\mathbf{X}) \geq \text{rank}_{\text{cp}}(\tilde{\mathbf{X}}) = \text{rank}_{\text{cp}}(\mathbf{E}_{\mathbf{u}}) \geq \binom{N}{\lfloor \frac{N}{2} \rfloor}.$$

The proof is complete. \square

Notice that the minimum CP rank of any antisymmetric tensor depends only on N . Thus, after taking the minimum over $\mathcal{A}(\otimes^N \mathbb{C}^K) \setminus \{\mathbf{0}\}$, the lower bound remains unchanged. Combining Corollaries 3.4 and 3.8 and applying the Stirling approximation $N! = \Theta(\sqrt{2\pi N}(N/e)^N)$, we obtain the following result.

Corollary 3.9. *Assume that $K \geq N$ and $f \in \mathcal{A}(\otimes^N \mathcal{F}_K(\Omega)) \setminus \{0\}$. Then the TPF rank of f satisfies*

$$\text{rank}_{\mathcal{F}_K(\Omega)}(f) \geq \Theta\left(\frac{2^N}{\sqrt{N}}\right).$$

Corollary 3.9 demonstrates that the TPF rank of any antisymmetric TPF in a finite-dimensional space grows at least exponentially with the dimension. Moreover, this exponential lower bound is independent of the basis size $K \geq N$. This independence originates from the intrinsic structure of the determinant tensor in (2.10). For $K < N$, Proposition 2.7 implies that the only antisymmetric TPF is the trivial one $f \equiv 0$.

Remark 3.10. Huang et al. [28] took the so-called backflow ansätze—a generalized Slater determinant—as the basis function and showed that, in the worst case, approximating a nonzero antisymmetric polynomial may require exponentially many determinants. In comparison, we take the TPF as the basis function and adopt a best-case perspective. We establish that even the *minimum* number of terms in representing a nonzero antisymmetric function grows exponentially with the problem dimension.

3.2 Representation complexity of antisymmetric tensor neural networks

Consider the TNN defined in (2.3). We begin with the cases where each subnetwork has only one hidden layer with a width of $m \in \mathbb{N}$. In this case, each univariate function takes the form

$$\psi_{ijk}(r_{jk}; \boldsymbol{\theta}_{jk}) = \sum_{\ell=1}^m a_{ijkl} \cdot \sigma(\omega_{jkl} r_{jk} + b_{jkl}) + c_{ijk}, \quad (3.6)$$

where $\boldsymbol{\theta}_{jk}$ collects $\{a_{ijkl}\}_{i\ell}$, $\{c_{ijk}\}_i$, $\{\omega_{jkl}\}_\ell$, $\{b_{jkl}\}_\ell \subseteq \mathbb{C}$, and $\sigma(\cdot)$ is an activation function. We aim to analyze the TPF rank of antisymmetric TNNs.

To apply the theoretical results in the previous subsection, we proceed to formulate the TNNs into TPFs in finite-dimensional spaces. To this end, for any $i \in \{1, \dots, p\}$ and $j \in \{1, \dots, N\}$, define

$$\psi_{ij}(\mathbf{r}_j; \boldsymbol{\theta}_j) := \prod_{k=1}^d \psi_{ijk}(r_{jk}; \boldsymbol{\theta}_{jk}), \quad \forall \mathbf{r}_j = (r_{j1}, \dots, r_{jd})^\top \in \Omega,$$

where $\boldsymbol{\theta}_j := (\boldsymbol{\theta}_{j1}, \dots, \boldsymbol{\theta}_{jd})^\top$. Recalling the definition of ψ_{ijk} in (3.6), one can thus verify that $\psi_{ij}(\cdot; \boldsymbol{\theta}_j) \in \tilde{\mathcal{F}}_{N,m,1}(\Omega) := \text{span}\{\tilde{\varphi}_{t,\mathcal{I},\boldsymbol{\ell}}\}$, where

$$\tilde{\varphi}_{t,\mathcal{I},\boldsymbol{\ell}}(\hat{\mathbf{r}}) := \prod_{k \in \mathcal{I}} \sigma(\omega_{tk\ell_k} \hat{r}_k + b_{tk\ell_k}), \quad \forall \hat{\mathbf{r}} = (\hat{r}_1, \dots, \hat{r}_d)^\top \in \Omega$$

for $t \in \{1, \dots, N\}$, $\mathcal{I} \subseteq \{1, \dots, d\}$ and $\boldsymbol{\ell} = (\ell_k)_{k \in \mathcal{I}}$ with each $\ell_k \in \{1, \dots, m\}$. Note that when $\mathcal{I} = \emptyset$, the function is defined to be identically equal to 1. Consequently, the TNN $f_{\boldsymbol{\theta}} \in \bigotimes^N \tilde{\mathcal{F}}_{N,m,1}(\Omega)$. Although the functions $\{\tilde{\varphi}_{t,\mathcal{I},\boldsymbol{\ell}}\}$ are not necessarily linearly independent, the space $\tilde{\mathcal{F}}_{N,m,1}(\Omega)$ is finite-dimensional for fixed parameters.

Using Corollary 3.9, we obtain the following bound for any antisymmetric TNN with one hidden layer and a width of m .

Corollary 3.11. *For a non-zero antisymmetric TNN with one hidden layer, a width of m , and any activation function, for any network parameters, its TPF rank p satisfies*

$$p \geq \Theta(2^N / \sqrt{N}).$$

Proof. From the discussions above, the TNN $f_{\boldsymbol{\theta}}$ with any fixed set of network parameters always belongs to the finite-dimensional function space $\tilde{\mathcal{F}}_{N,m,1}(\Omega)$. By Corollary 3.9, for any fixed network parameters, we have

$$p \geq \Theta(2^N / \sqrt{N}).$$

Since varying the parameters only alters the basis functions without changing the intrinsic finite dimensionality, this lower bound holds uniformly for all network parameters. \square

Notice that the above result is applicable under any fixed network architecture. The only modification required is to adjust the finite-dimensional space.

Corollary 3.12. *For a non-zero antisymmetric TNN with any fixed finite depths and widths and any activation function, for any network parameters, its TPF rank p satisfies*

$$p \geq \Theta(2^N / \sqrt{N}).$$

Remark 3.13. During the training process, the TNN transfers among different finite-dimensional function spaces. Corollary 3.12 states that only when $p \geq \Theta(2^N / \sqrt{N})$ can it possibly reach a space that contains a nonzero antisymmetric function.

Remark 3.14. According to our theoretical results, for a three-electron system (i.e., $N = 3$), the TPF rank of the electronic wave function in finite-dimensional spaces is at least $\binom{3}{1} = 3$. On the other hand, Iiten and Teitler [31] demonstrated that the CP rank of the third-order determinant tensor is exactly five.

These indicate that a TPF rank of $p = 5$ is sufficient to represent antisymmetry in finite-dimensional spaces. Notably, this does not contradict the experimental results in [34], where a rank of $p = 50$ was employed in TNNs. Our results establish a theoretical condition required to ensure the existence of nonzero antisymmetric solutions in finite-dimensional spaces, whereas the TNNs in [34] are specific members within the associated spaces.

4 Conclusion and discussions

Recent studies reveal that the TPF approximations (specifically, the TNNs) can incur extraordinarily high computational costs to achieve desirable accuracy on quantum many-body problems, even for systems with as few as three electrons. This stands in sharp contrast to their success on other high-dimensional applications. Preliminary numerical results point to their representation complexity in representing antisymmetry, a fundamental property that the electronic wave function should always satisfy. In this paper, we rigorously establish an exponential lower bound on the TPF rank for the antisymmetric TPFs in finite-dimensional spaces. In other words, the minimum number of involved terms for the TPFs in finite-dimensional spaces to ensure antisymmetry increases exponentially with the problem dimension. Notably, this finite-dimensional setting is compatible with the traditional discretization methods and also includes the TNNs (the TPFs parameterized by neural networks) as special cases. In our proof, we mainly leverage the connection between the antisymmetric TPFs and antisymmetric tensors and dig into the CP rank of the latter.

Our theoretical results offer new insights and perspectives for applications with antisymmetry requirements. First, low-rank TPFs, while favorable due to the separability of integrals, cannot be antisymmetric in high-dimensional settings. For example, in the case of $N = 20$, representing antisymmetry requires a rank of at least 1.8×10^5 , making such representations impractical. However, this does not necessarily imply an unacceptably high computational cost in implementations, especially when certain algebraic structures are available. In contrast, our findings partly explain why determinant-based constructions have become the standard choice in modern practice, and some work has investigated the rank with Slater determinants as basis functions [28]. Alternatively, one can resort to approximations based on other tensor formats, such as tensor train formats [41], whose capability to represent antisymmetry have not been studied. Second, quantitatively investigating the impact of errors in representing antisymmetry on relevant applications would be valuable. As an analogy, the principal component analysis in statistics essentially approximates a covariance matrix with a low-rank one, while retaining most of the critical information. Understanding the trade-off between computational efficiency and accuracy in antisymmetry representation can guide further developments.

Acknowledgements Yukuan Hu was supported by European Research Council under the European Union’s Horizon 2020 Research and Innovation Programme (Grant No. 810367). Xin Liu was supported by National Natural Science Foundation of China (Grant Nos. 12125108, 12021001 and 12288201) and Research Grants Council for the Chinese Academy of Sciences, Academy of Mathematics and Systems Science-Hong Kong Polytechnic University Joint Laboratory in Applied Mathematics (Grant No. JLFS/P-501/24).

References

- 1 Bachmayr M. Low-rank tensor methods for partial differential equations. *Acta Numer*, 2023, 32: 1–121
- 2 Bajdich M, Mitas L, Drobný G, et al. Pfaffian pairing wave functions in electronic-structure quantum Monte Carlo simulations. *Phys Rev Lett*, 2006, 96: 130201
- 3 Bajdich M, Mitas L, Wagner L K, et al. Pfaffian pairing and backflow wavefunctions for electronic structure quantum Monte Carlo methods. *Phys Rev B*, 2008, 77: 115112
- 4 Bazarkhanov D, Temlyakov V. Nonlinear tensor product approximation of functions. *J Complexity*, 2015, 31: 867–884
- 5 Beylkin G, Mohlenkamp M J. Numerical operator calculus in higher dimensions. *Proc Natl Acad Sci USA*, 2002, 99: 10246–10251

- 6 Beylkin G, Mohlenkamp M J. Algorithms for numerical analysis in high dimensions. *SIAM J Sci Comput*, 2005, 26: 2133–2159
- 7 Bradbury J, Frostig R, Hawkins P, et al. JAX: Composable transformations of Python+NumPy programs. [Http://github.com/jax-ml/jax](http://github.com/jax-ml/jax), 2018
- 8 Chen Z, Lu J. Exact and efficient representation of totally anti-symmetric functions. [arXiv:2311.05064](https://arxiv.org/abs/2311.05064), 2023
- 9 Dai X, Fan Y, Sheng Z. Subspace method based on neural networks for eigenvalue problems. [arXiv:2410.13358](https://arxiv.org/abs/2410.13358), 2024
- 10 Derksen H. On the nuclear norm and the singular value decomposition of tensors. *Found Comput Math*, 2016, 16: 779–811
- 11 DeVore R, Hanin B, Petrova G. Neural network approximation. *Acta Numer*, 2021, 30: 327–444
- 12 Dolgov S, Kressner D, Strössner C. Functional Tucker approximation using Chebyshev interpolation. *SIAM J Sci Comput*, 2021, 43: A2190–A2210
- 13 E W, Yu B. The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems. *Commun Math Stat*, 2018, 6: 1–12
- 14 Feynman R P, Cohen M. Energy spectrum of the excitations in liquid helium. *Phys Rev*, 1956, 102: 1189–1204
- 15 Fock V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z Physik*, 1930, 61: 126–148
- 16 Gao N, Günemann S. Neural Pfaffians: solving many many-electron Schrödinger equations. *Adv Neural Inform Process Systems*, 2024, 37: 125336–125369
- 17 Griebel M, Harbrecht H. Analysis of tensor approximation schemes for continuous functions. *Found Comput Math*, 2023, 23: 219–240
- 18 Griebel M, Harbrecht H, Schneider R. Low-rank approximation of continuous functions in Sobolev spaces with dominating mixed smoothness. *Math Comp*, 2023, 92: 1729–1746
- 19 Hackbusch W. *Tensor Spaces and Numerical Tensor Calculus*. Berlin: Springer, 2012
- 20 Hackbusch W, Khoromskij B N. Tensor-product approximation to operators and functions in high dimensions. *J Complexity*, 2007, 23: 697–714
- 21 Hackbusch W, Khoromskij B N. Tensor-product approximation to multidimensional integral operators and Green's functions. *SIAM J Matrix Anal Appl*, 2008, 30: 1233–1253
- 22 Han J, Li Y, Lin L, et al. Universal approximation of symmetric and anti-symmetric functions. *Commun Math Sci*, 2022, 20: 1397–1408
- 23 Han J, Zhang L, E W. Solving many-electron Schrödinger equation using deep neural networks. *J Comput Phys*, 2019, 399: 108929
- 24 Hartree D R. The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods. *Math Proc Cambridge Philos Soc*, 1928, 24: 89–110
- 25 Helgaker T, Jørgensen P, Olsen J. *Molecular Electronic-Structure Theory*. New York: John Wiley & Sons, 2000
- 26 Hitchcock F L. The expression of a tensor or a polyadic as a sum of products. *J Math Phys*, 1927, 6: 164–189
- 27 Hu Z, Shukla K, Karniadakis G E, et al. Tackling the curse of dimensionality with physics-informed neural networks. *Neural Networks*, 2024, 176: 106369
- 28 Huang H, Landsberg J M, Lu J. Geometry of backflow transformation ansatz for quantum many-body fermionic wavefunctions. *Commun Math Sci*, 2023, 21: 1447–1453
- 29 Huang J, Wu H, Zhou T. Adaptive neural network basis methods for partial differential equations with low-regular solutions. *Commun Comput Phys*, 2026, 39: 553–577
- 30 Hutter M. On representing (anti)symmetric functions. [arXiv:2007.15298](https://arxiv.org/abs/2007.15298), 2020
- 31 Ilten N, Teitler Z. Product ranks of the 3×3 determinant and permanent. *Canad Math Bull*, 2016, 59: 311–319
- 32 Jastrow R. Many-body problem with strong forces. *Phys Rev*, 1955, 98: 1479–1484
- 33 Kao T, Zhang H, Zhang L, et al. pETNNs: Partial evolutionary tensor neural networks for solving time-dependent partial differential equations. *CSIAM Trans Appl Math*, 2025, 6: 862–891
- 34 Liao Y, Lin Z, Liu J, et al. Solving Schrödinger equation using tensor neural network. [arXiv:2209.12572](https://arxiv.org/abs/2209.12572), 2024
- 35 Liao Y, Ming P. Deep Nitsche method: Deep Ritz method with essential boundary conditions. *Commun Comput Phys*, 2021, 29: 1365–1384
- 36 Liao Y, Wang Y, Xie H. Solving high dimensional partial differential equations using tensor type discretization and optimization process. [arXiv:2211.16548](https://arxiv.org/abs/2211.16548), 2022
- 37 Lin J, Goldshlager G, Lin L. Explicitly antisymmetrized neural network layers for variational Monte Carlo simulation. *J Comput Phys*, 2023, 474: 111765
- 38 Lin Z, Wang Y, Xie H. Adaptive neural network subspace method for solving partial differential equations with high accuracy. [arXiv:2412.02586](https://arxiv.org/abs/2412.02586), 2024
- 39 Löwdin P-O. Correlation problem in many-electron quantum mechanics I. Review of different approaches and discussion of some current ideas. *Adv Chem Phys*, 1958, 2: 207–322
- 40 Mohlenkamp M J, Monzón L. Trigonometric identities and sums of separable functions. *Math Intelligencer*, 2005, 27: 65–69

- 41 Oseledets I-V. Tensor-train decomposition. SIAM J Sci Comput, 2011, 33: 2295–2317
- 42 Oseledets I-V. Constructive representation of functions in low-rank tensor formats. Constr Approx, 2013, 37: 1–18
- 43 Pang T, Yan S, Lin M. $\mathcal{O}(N^2)$ universal antisymmetry in fermionic neural networks. arXiv:2205.13205, 2022
- 44 Pauli W. Über den zusammenhang des abschlusses der elektronengruppen im atom mit der komplexstruktur der spektren. Z Physik, 1925, 31: 765–783
- 45 Pfau D, Spencer J S, Matthews A G D G, et al. *Ab initio* solution of the many-electron Schrödinger equation with deep neural networks. Phys Rev Res, 2020, 2: 033429
- 46 Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys, 2019, 378: 686–707
- 47 Schmidt E. Zur theorie der linearen und nichtlinearen integralgleichungen. Math Ann, 1907, 63: 433–476
- 48 Siegel J W, Hong Q, Jin X, et al. Greedy training algorithms for neural networks and applications to PDEs. J Comput Phys, 2023, 484: 112084
- 49 Sirignano J, Spiliopoulos K. DGM: A deep learning algorithm for solving partial differential equations. J Comput Phys, 2018, 375: 1339–1364
- 50 Slater J-C. Note on Hartree’s method. Phys Rev, 1930, 35: 210–211
- 51 Wang T, Hu Z, Kawaguchi K, et al. Tensor neural networks for high-dimensional Fokker-Planck equations. Neural Networks, 2025, 185: 107165
- 52 Wang Y, Lin Z, Liao Y, et al. Solving high-dimensional partial differential equations using tensor neural network and *a posteriori* error estimators. J Sci Comput, 2024, 101: 1–29
- 53 Wang Y, Xie H. Computing multi-eigenpairs of high-dimensional eigenvalue problems using tensor neural networks. J Comput Phys, 2024, 506: 112928
- 54 Wang Y, Xie H, Jin P. Tensor neural network and its numerical integration. J Comput Math, 2024, 42: 1714–1742
- 55 Weidmann J. Linear Operators in Hilbert Spaces. New York: Springer, 1980
- 56 Xiao S, Jin P, Tang Y. Learning solution operators of PDEs defined on varying domains via MIONet. arXiv:2402.15097, 2024
- 57 Ye H, Li R, Gu Y, et al. $\tilde{\mathcal{O}}(N^2)$ representation of general continuous anti-symmetric function. arXiv:2402.15167, 2024
- 58 Yu H, Zhang S. A natural deep Ritz method for essential boundary value problems. J Comput Phys, 2025, 537: 114133
- 59 Zhou D, Chen H, Ho C H, et al. A multilevel method for many-electron Schrödinger equations based on the atomic cluster expansion. SIAM J Sci Comput, 2024, 46: A105–A129
- 60 Zhou Q, Wu T, Liu J, et al. Sum-of-Gaussians tensor neural networks for high-dimensional Schrödinger equation. arXiv:2508.10454, 2025

Appendix A Experimental settings

For the one-dimensional system, we consider the Hamiltonian with soft-Coulomb interaction [59]:

$$\mathcal{H} := -\frac{1}{2} \sum_{i=1}^N \Delta_{r_i} - \sum_{i=1}^N \sum_{I=1}^M \frac{Z_I}{\sqrt{1 + (r_i - R_I)^2}} + \sum_{1 \leq i < j \leq N} \frac{1}{\sqrt{1 + (r_i - r_j)^2}},$$

where $\{r_i\}_{i=1}^N$ and $\{R_I\}_{I=1}^M$ represent electron and nucleus positions, respectively, and Z_I is the charge of the I -th nucleus.

We consider a spinless system. For the non-antisymmetric TNN wave function f_{θ} defined in (2.2), the loss function is chosen as in [34]:

$$L^{\text{TNN}}(\theta) := \frac{\langle f_{\theta}, \mathcal{H} f_{\theta} \rangle}{\langle f_{\theta}, f_{\theta} \rangle} + \beta \sum_{1 \leq i < j \leq N} \frac{\langle f_{\theta}, f_{\theta} \circ \mathcal{I}_{ij} \rangle}{\langle f_{\theta}, f_{\theta} \rangle},$$

where $\langle \cdot, \cdot \rangle$ is the inner product in \mathcal{L}^2 , and $\beta > 0$ is the penalty parameter.

For the antisymmetrized TNN wave function, we apply the antisymmetrizer \mathcal{A} defined in (2.5) to f_{θ} , yielding

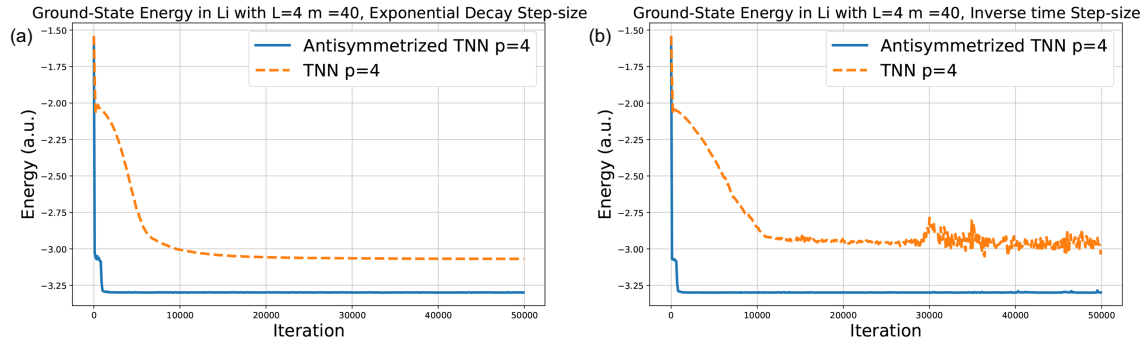
$$f_{\theta}^{\text{anti}}(r_1, \dots, r_N) := \frac{1}{N!} \sum_{\pi \in \mathcal{S}_N} \sum_{i=1}^p \text{sgn}(\pi) \prod_{j=1}^N \psi_{ij}(r_{\pi(j)}; \theta_j).$$

The corresponding loss function is

$$L^{\text{anti}}(\theta) := \frac{\langle f_{\theta}^{\text{anti}}, \mathcal{H} f_{\theta}^{\text{anti}} \rangle}{\langle f_{\theta}^{\text{anti}}, f_{\theta}^{\text{anti}} \rangle}.$$

Table A1 Neural network and optimization parameters

Parameter		Value
Neural network	Hidden layers	$L = 2$
	Hidden dimension	$m = 20$
	Activation function	$\sigma(x) = \tanh(x)$
	Initialization	Default in JAX
Optimization	Optimizer	Adam
	Initial learning rate	$\eta_0 = 1.00 \times 10^{-3}$
	Learning rate decay	$\eta_k = 0.7^{\lfloor k/3000 \rfloor} \eta_0$
	Iterations	50000
	Penalty parameter	$\beta = 200$

**Figure A1** (Color online) Comparison of TNN approximations with and without antisymmetrization for the one-dimensional Li system using $L = 4$ and $m = 40$, under the exponential decay learning rate schedule (a) and the inverse time schedule (b)

Our experiments are conducted on one-dimensional Lithium (Li) and Helium hydride ion (HeH^+). For Li, the nucleus is positioned at 0.0. For HeH^+ , the nuclei are located at 0.0 (He) and 1.463 (H^+). Since the loss functions only involve one- and two-dimensional integrals, we compute them using Gauss-Legendre quadrature. The integration interval is truncated to $[-10, 10]$, divided uniformly into 30 subintervals, with 30 quadrature points applied in each subinterval.

The neural network architecture and optimization settings are detailed in Table A1, closely following [34]. All experiments were implemented in JAX [7] on an Ubuntu 20.04.1 operating system and trained on GPUs ($8 \times \text{NVIDIA GeForce RTX 4090}$) with CUDA 12.0. Results for $p = 4$ are shown in Figure 1.

The numerical experiments aim to show that the lack of antisymmetry can substantially degrade accuracy and efficiency. Such a deficiency cannot be resolved easily by tuning hyperparameters. This argument is supported by additional experiments using either more complicated neural networks ($L = 4$, $m = 40$) or different learning rate schedules ($\eta_k = \eta_0 / (1 + \alpha k)$ and $\alpha = 1 \times 10^{-3}$) (see Figure A1). The obtained results are qualitatively consistent with those in Figure 1.