

Deep learning topological inference-guided T_{cc}^+ pole parameter extraction

Julius B. Pagayon^{*,†}, Klarence Tomas R. Cervantes[‡], and Denny Lane B. Sombillo[‡]
National Institute of Physics, University of the Philippines Diliman, Quezon City 1101, Philippines



(Received 27 March 2026; accepted 6 July 2026; published 23 July 2026)

We perform a data-driven study of the doubly charmed tetraquark candidate T_{cc}^+ . An ensemble of deep neural network classifiers, trained on synthetic amplitudes with controlled analytic structures, identifies a dominant pole topology characterized by an isolated pole on the $[b\iota]$ Riemann sheet which is robust against left-hand cut effects. A subsequent pole parameter extraction was performed via the uniformized $\mathcal{S}$ -matrix and a complementary $\mathcal{K}$ -matrix parametrization, which respectively provides a model-independent baseline and dynamical insight on the pole position and trajectory of the resonant state. Using this two-pronged approach, we submit that the T_{cc}^+ is a shallow $D^0 D^{*+}$ bound state in the second Riemann sheet of the complex plane.

DOI: 10.1103/nr94-jzxj

I. INTRODUCTION

The observation of the doubly charmed tetraquark candidate $T_{cc}(3875)^+$ by the LHCb Collaboration in 2021 [1,2] marked a significant milestone in exotic hadron spectroscopy. This exotic state, with quark minimal content $cc\bar{u}\bar{d}$, manifests as a narrow enhancement slightly below the $D^0 D^{*+}$ threshold in the $D^0 D^0 \pi^+$ invariant mass spectrum. Despite intensive theoretical efforts, the precise nature of T_{cc}^+ remains under debate. Various studies propose it to be a compact tetraquark [3–14], a loosely bound DD^* hadronic molecule [15–36], or a virtual state [37,38]. An alternative exotic explanation invoke purely kinematical effect (e.g., triangle singularity) but such effects are numerically suppressed [39] in the $T_{cc}^+ \rightarrow D^0 D^0 \pi^+$ decay. One may refer to Refs. [40–42] for comprehensive reviews.

In this work, we utilized a two-pronged approach grounded in scattering theory. First, we adopt the coupled-channel $\mathcal{S}$ -matrix formalism, constrained by analyticity, unitarity, and hermiticity. To manage the multi-sheeted Riemann surface arising from channel thresholds, we implement a conformal uniformization scheme, which maps the four-sheeted complex s -plane into a single-sheeted complex ω -plane [43]. This enables a concise formulation of the amplitude with explicit analytic control over the placement of poles across different Riemann

sheets [44]. The uniformization scheme provides significant advantages in parametrization, numerical stability, and interpretability, especially when dealing with near-threshold phenomena where sheet ambiguities are critical to physical interpretation.

Within this framework, we generated a large ensemble of line shapes with known pole topologies and train an ensemble of deep neural network (DNN) classifiers to discriminate synthetic amplitudes into coarse-grained categories based on the number and location of nearby poles. This data-driven method, which builds upon earlier works in single-channel analysis of nucleon-nucleon scattering [45,46] and two-channel analysis [47], has been successfully applied to the study of other resonance phenomena and other application in hadron spectroscopy [48–65]. Here, we extend the method to analyze the experimental $D^0 D^0 \pi^+$ line shape associated with the T_{cc}^+ , and extract the most probable pole configuration.

A subsequent pole parameter extraction was performed via the uniformized $\mathcal{S}$ -matrix parametrization which provides a model-independent baseline. This was complemented by a dynamical fit to the experimental data using the $\mathcal{K}$ -matrix formalism in a coupled-channel setting. The latter enables the extraction of pole positions and tracing of pole trajectory, thereby providing dynamical insight on the underlying nature of the resonant state.

The primary advantage of using DNNs in our analysis lies in their ability to efficiently reduce the vast model space of 35 pole topology candidates to the most probable one for a detailed analysis. Without this model space reduction process, exhaustive parametric fitting of all pole topologies with the empirical data would be computationally exhaustive and is susceptible to subjective biases. Moreover, unlike end-to-end deep learning approaches that often treat the data as a black box, our hybrid strategy retains a strong

*Contact author: jbpagayon@up.edu.ph

†Contact author: krcervantes1@up.edu.ph

‡Contact author: dbsombillo@up.edu.ph

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

connection to underlying physical principles while leveraging the pattern-recognition power of DNNs. The main goal of the DNN is to provide an unbiased classification of the pole topology favored by the observed line shape. This work can be classified as a bottom-up treatment in hadron spectroscopy [66].

The investigative framework of this work is structured as follows. Section II discusses the coupled-channel scattering framework and the associated uniformization mapping. Section III details the construction of DNN classifiers, including the generation of the training dataset, DNN architecture, and the inference to the LHCb data. In Sec. IV, we extracted pole parameters via the uniformized $\mathcal{S}$ -matrix and coupled-channel $\mathcal{K}$ -matrix model. Section V summarizes our key findings and outlines possible directions for future work.

II. INDEPENDENT $\mathcal{S}$ -MATRIX POLES

In scattering theory, the scattering amplitude $T(s)$, expressed as a function of the Mandelstam variable s [square of the total center-of-mass (COM) energy E], exhibits an analytic structure in the complex s -plane that reflects the fundamental properties of the underlying interaction dynamics. Apart from a set of isolated and well-defined singularities, the amplitude is analytic across the complex s -plane. These singularities, which include branch points and poles, govern much of the physical content of the amplitude [67–69].

Branch points arise whenever a new scattering channel becomes kinematically accessible. This occurs when $\sqrt{s}$ exceeds the threshold energy $\varepsilon_k = m_{k1} + m_{k2}$, where m_{ki} is the mass of the i th particle in channel k . Each threshold introduces a square-root-type branch point, indicating that the function is multivalued. This multivaluedness requires the introduction of multiple Riemann sheets in the complex s -plane, with each additional channel doubling the number of sheets. These branch points are accompanied by branch cuts, which are usually defined to extend from the threshold value to positive infinity along the real axis. These cuts are commonly referred to as right-hand cuts, kinematical cuts, or unitarity cuts [67].

In addition to these kinematical singularities, the amplitude possesses dynamical singularities known as left-hand cuts (lhcs). These arise from the explicit form of the interaction potential, such as the exchange of virtual particles in the t or u channel, and are not related to production thresholds. Their location is determined by the masses of the exchanged particles, typically appearing for $s < 0$ or at unphysical values of s along the negative real axis [70,71]. While right-hand cuts are universal, the structure of the lhcs is model-dependent and encodes the forces responsible for the scattering, such as the one-pion exchange (OPE) crucial in hadronic interactions.

Moreover, the scattering amplitude also contains pole singularities, which are associated with physical phenomena

such as bound states, virtual states, and resonances [70]. Bound states and virtual states appear on the real s -axis below the threshold energy. Bound state poles reside on the physical sheet, while virtual state poles are located on the unphysical sheet. In contrast, resonance poles are located off the real axis and are found in the complex s -plane of the unphysical sheets. The analytic nature of $T(s)$ imposes the Schwarz reflection principle, which requires that if a pole exists at a complex point s_p , then a corresponding pole must also exist at its complex conjugate s_p^* [68].

Although the single-channel framework captures many essential analytic features, it becomes insufficient when handling the structure of the T_{cc}^+ state. The proximity of multiple thresholds and the interplay between open channels necessitates a coupled-channel treatment to accurately describe the pole dynamics.

A. Coupled-channel analytic structure

In a two-channel system such as the T_{cc}^+ with dominant open channels $D^0 D^0 \pi^+$ and $D^0 D^{*+}$, the analytic structure is governed by the behavior of the COM momenta for the k th channel as

$$p_k = \sqrt{\frac{[s - (m_{k1} + m_{k2})^2][s - (m_{k1} - m_{k2})^2]}{4s}}. \quad (1)$$

In this study, we consider a coupled-channel system reflecting the dominant physical process. The three-body process factorizes into a strong two-body interaction $D^0 D^{*+} \rightarrow D^0 D^{*+}$ followed by the decay $D^{*+} \rightarrow D^0 \pi^+$. Hence, the first channel represents the direct production of the observed three-body final state $D^0 D^0 \pi^+$, treated in an effective two-body formalism with asymptotic states D^0 and a composite $[D^0 \pi^+]$. This effective reduction is motivated by the argument that the D^{*+} and $D^0 \pi^+$ are physically distinct structures with different thresholds [37]. The second channel represents the intermediate resonant two-body state $D^0 D^{*+}$. This gives the corresponding nominal thresholds $\varepsilon_1 = m_{D^0} + [m_{D^0} + m_{\pi^+}] \simeq 3869.25$ [MeV] and $\varepsilon_2 = m_{D^0} + m_{D^{*+}} \simeq 3875.10$ [MeV] with the masses of D^0 , D^{*+} , and π^+ mesons obtained from PDG [72]. In the present study, the coupled-channel analysis is restricted to an effective two-channel framework involving the nearby $D^0 D^{*+}$ threshold. The treatment is already sufficient without considering the higher mass channel $D^+ D^{*0}$. In [73], the $D^0 D^{*+}$ and $D^+ D^{*0}$ are treated as isospin mixture of $I = 0$ and $I = 1$. This means that the two channels can be treated as one under isospin symmetry. In our case, we only chose the channel that is closest to the enhancement. Thus, for the purpose of identifying the dominant near-threshold pole configuration, it is sufficient in the present framework to include the $D^0 D^{*+}$ threshold that lies within its proximity.

Alternatively, one can extend the treatment of the tetraquark signal to include the $D^+ D^{*0}$ threshold explicitly.

However, such an extension would require a three-channel independent $\mathcal{S}$ -matrix pole formalism that is yet to be developed, which would enlarge the Riemann sheet structure and substantially increase the number of pole topologies to be generated and classified in the DNN workflow. This extension would also need more computational resources. With this, the present study should be understood as an analysis, focused on the dominant near-threshold structure, rather than a coupled-channel calculation including all possible thresholds.

Noting that our analysis focuses on the behavior of the amplitude in the vicinity of the relevant threshold ε_k , it is sufficient to approximate the channel momentum by the rescaled form,

$$q_k = \sqrt{s - \varepsilon_k^2}. \quad (2)$$

The structure of $q_k(s)$ introduces branch points at $s = \varepsilon_k^2$. As s crosses the branch cuts associated with the branch points, the sign of $\text{Im } q_k(s)$ flips, giving rise to two Riemann sheets per channel. Consequently, the full two-channel amplitude is defined on a four-sheeted Riemann surface. For brevity, we adopt the Pearce and Gibson notation [74] to label the Riemann surfaces. Specifically, we use the form $[\zeta_1, \zeta_2]$, where each $\zeta_k \in \{t, b\}$ indicates the sheet [top (t) indicating that $\text{Im } q_k > 0$ or bottom (b) implying that $\text{Im } q_k < 0$]. Under this convention, the four Riemann sheets are labeled as $[tt]$ (sheet I), $[bt]$ (sheet II), $[bb]$ (sheet III), and $[tb]$ (sheet IV).

In this work, a primary objective is to determine the most probable pole structure for the T_{cc}^+ state. Probing the pole structure is essential because, in principle, the $\mathcal{S}$ -matrix pole structure is a model-independent signature of a resonant state, regardless of the specific parametrization used. The physical nature of a near-threshold state is encoded in its pole structure characterized by the number of poles and the specific Riemann sheets they inhabit. This topological classification allows us to distinguish between fundamentally different physical scenarios. For instance, following the naïve pole counting criterion [75], a pole in the $[bt]$ sheet generally corresponds to a molecular state while a pole-shadow pair on the $[bt]$ and $[bb]$ sheets may be associated with a compact state picture. See the detailed discussion on this association in Appendix B. However, tracing the pole across these sheets can be technically subtle and require careful treatment.

B. Uniformization

To systematically analyze the analytic structure, we adopt a uniformizing variable ω following Kato [43] and recent implementation in hadron spectroscopy [44],

$$\omega = \frac{q_1 + q_2}{\sqrt{\varepsilon_2^2 - \varepsilon_1^2}}, \quad (3)$$

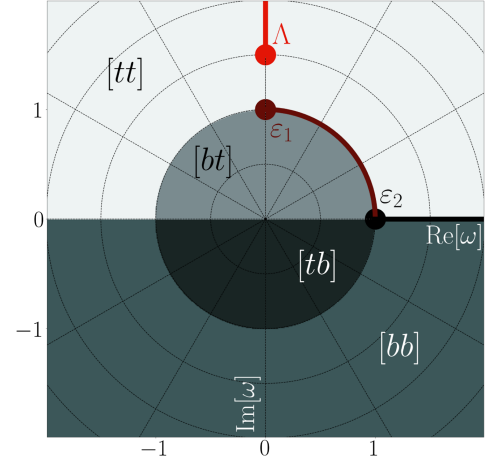


FIG. 1. Mapping of the four-sheeted complex s -plane onto the single-sheeted complex ω -plane. The red dot labeled as Λ correspond to the branch point associated with left-hand cut that arises from meson exchange. The maroon point at $\omega = i$ and the black point at $\omega = 1$ correspond to the first and second channel thresholds, ε_1 and ε_2 , respectively. The physical scattering region between the thresholds is mapped onto the upper arc of the unit circle, while the region above the second threshold is represented by the segment of the real axis extending beyond $\omega = 1$.

which ensures that every point on the ω -plane corresponds to a unique point (and sheet) in the original complex s -plane. In particular, the different Riemann sheets correspond to distinct angular sectors of the ω -plane as shown in Fig. 1, allowing one to assign poles to sheets based on the location of their corresponding ω -values.

In this uniformized representation, the two-channel $\mathcal{S}$ -matrix elements take a compact and analytically controlled form,

$$\begin{aligned} \mathcal{S}_{11}(\omega) &= \frac{\mathcal{Q}(-1/\omega)}{\mathcal{Q}(\omega)}; & \mathcal{S}_{22}(\omega) &= \frac{\mathcal{Q}(1/\omega)}{\mathcal{Q}(\omega)}; \\ \det \mathcal{S} &= \frac{\mathcal{Q}(-\omega)}{\mathcal{Q}(\omega)}, \end{aligned} \quad (4)$$

where $\mathcal{Q}(\omega)$ is the Jost-like function that governs the pole structure of the amplitude which is explicitly given as

$$\begin{aligned} \mathcal{Q}(\omega) &= \prod_{j=1}^n \frac{1}{\omega^2} (\omega - \omega_{\text{pole}})(\omega + \omega_{\text{pole}}^*) \\ &\quad \times (\omega - \omega_{\text{reg}})(\omega + \omega_{\text{reg}}^*). \end{aligned} \quad (5)$$

Here, n denotes the number of physical poles ω_{pole} included in the model. The physical poles ω_{pole} represents a complex pole encoding a physical structure, ω_{reg} is a regulator pole that ensure correct asymptotic behavior $\mathcal{S}_{jk} \rightarrow 1$ as $\omega \rightarrow \infty$, as required by unitarity and Levinson's theorem [68,69]. The factor ω^{-2} prohibits highly energetic initial states from having no final states

after the interaction. Due to hermiticity of the $\mathcal{S}$ -matrix below the threshold, each pole must be paired with a symmetric pole at $-\omega_{\text{pole}}^*$. To ensure the regulator poles does not interfere with physical effects, we place them away from the scattering region, either on the $[bb]$ or $[tb]$ sheet by enforcing that $\omega_{\text{reg}} = e^{-i\pi/2}|\omega_{\text{pole}}|^{-1}$, ensuring $|\omega_{\text{pole}}\omega_{\text{reg}}| = 1$. The line shape amplitude is constructed from the T -matrix, which is extracted from the $\mathcal{S}$ -matrix via the relation,

$$\mathcal{S}_{jk} = \delta_{jk} + 2iT_{jk}, \quad (6)$$

where δ_{jk} is the Kronecker delta.

C. Left-hand cut: Phenomenological modeling

While the uniformized $\mathcal{S}$ -matrix provides a robust framework for handling right-hand cuts and pole singularities, it does not inherently contain the dynamics of the interaction potential. In hadron physics, these dynamics are dominated by cross-channel meson exchange, which introduces a critical nonanalyticity: the left-hand cut (lhc) [67].

The origin and location of lhcs are directly tied to the form of the potential. For a potential arising from meson exchange, which decays exponentially in coordinate space (Yukawa-type), the lhc appears along the imaginary axis in the complex momentum plane. They manifest as strips that start from $\pm i\Lambda$ where Λ is related to the mass of the exchange particle. This was rigorously established in nonrelativistic scattering theory, showing that for a potential $V(r) \sim e^{-\mu r}/r$, the scattering amplitude possesses a lhc along the imaginary k -axis from $k = +i\mu/2$ to $+\infty$ and a symmetric cut on the negative imaginary axis [68,69]. This structure emerges from the maximal analytic continuation of the scattering amplitude. Notably, if the interhadron potential were to decay faster than an exponential, no lhc would be present; however, the Yukawa-type potential of meson exchange guarantees its existence.

The phenomenological impact of the lhc is hierarchical and is cornerstone of modern hadronic amplitude analysis. The lightest meson, the pion (π), produces the lhc closest to the physical scattering region, making it most significant for low-energy phenomenology. Exchanges of heavier mesons, such as the ρ or ω , generate lhcs that are situated further from the threshold and thus have a comparatively subdued effect on near-threshold phenomena. Consequently, phenomenological models must carefully incorporate these singularities to ensure the analytic structure is correctly reproduced. This necessity is underscored by Ref. [76], which demonstrated that explicit inclusion of the lhc is pivotal for the correct determination of the nature of the T_{cc}^+ pole, particularly when interpreting lattice QCD data at various pion masses.

Given this imperative to include the lhc, a key challenge is to incorporate it in a manner that is both

phenomenologically effective and consistent with the principles of the $\mathcal{S}$ -matrix. A particularly efficient ansatz is to introduce a phase factor that directly modifies the diagonal $\mathcal{S}$ -matrix elements,

$$\mathcal{S}_{kk} \rightarrow \exp \left[i\zeta_k \arctan \left(\frac{q_k}{\Lambda} \right) \right] \mathcal{S}_{kk}, \quad (7)$$

where Λ sets the scale of the lhc effects and ζ_k parametrizes the strength of lhc contribution in each channel. The effectiveness of this approach is rooted in its direct connection to the analytic structure it aims to model. The choice of the $\arctan$ function is mathematically deliberate; it is analytic everywhere except for branch cuts on the imaginary axis of the complex q_k -plane starting precisely at $\pm i\Lambda$, thereby mimicking the fundamental analyticity generated by a Yukawa potential with range of $1/\Lambda$ [68,69].¹ For instance, by setting the $\Lambda \sim m_\pi$, the model explicitly accounts for long-range force scale associated with one pion exchange (OPE). Beyond OPE, this framework provides a generalized mechanism to capture further exchange physics, including two-pion exchange (TPE) and the contributions of heavier ρ and ω mesons by adjusting the scale Λ .

In its effect on the phase shift, this ansatz introduces a smooth, energy-dependent background term, $\delta_{\text{lhc}}(q_k) = (\zeta_k/2) \arctan(q_k/\Lambda)$, which represents the cumulative contribution from the t -channel forces. The parameter ζ_k serves as a phenomenological strength for the left-hand cut. A negative ζ_k , for example, yields a negative contribution to the phase shift, effectively modeling the repulsive background interactions arising from short-range t -channel dynamics. Crucially, this multiplicative modification is minimal and robust. It preserves the unitarity of the $\mathcal{S}$ -matrix, as the factor is a pure phase and leaves analytically controlled right-hand cut structure of the original $\mathcal{S}_{kk}$ intact. It therefore serves as an efficient bridge, endowing a unitary and analytic amplitude with the necessary dynamical content from cross-channel exchanges, without the need for a complicated and computationally intensive potential model.

In summary, the central innovation of this approach lies in the parametrization of the Jost-like function $\mathcal{Q}(\omega)$ via a set of known or hypothesized pole locations ω_{pole} that can be placed freely and independently in the complex plane, allowing amplitudes to be built with arbitrary combinations in four Riemann sheets. Crucially, it enables the systemic generation of amplitudes corresponding to all physically relevant pole topologies, both with or without the inclusion of left-hand cut effects. In the next section, we leverage this

¹While the $\arctan$ ansatz is numerically efficient, a rigorous treatment of left-hand cut effects would require their explicit inclusion via the N/D method or through unitarized dispersion relations. For this reason, throughout the remainder of this work, we refer to these contributions as simulated lhc effects.

theoretically controlled model space to train deep neural networks capable of inverting the scattering problem: inferring the underlying pole structure directly from the observed scattering data.

III. DEEP LEARNING FRAMEWORK

The model-independent construction of the $\mathcal{S}$ -matrix described in the previous section enables the controlled generation of a wide variety of line shapes with explicitly assigned pole structures. Each generated amplitude corresponds to a known set of pole positions on specific Riemann sheets. This capability allows us to reframe the inverse scattering problem as a supervised classification task: given an observable line shape, predict the corresponding pole topology class. In this section, we describe the generation of the synthetic training and validation dataset, the architecture of the deep neural network classifier, the overall classification strategy used to extract pole structures from experimental line shapes, and the inference framework to the T_{cc}^+ state.

A. Dataset construction

1. Synthetic dataset for training and validation

To enable supervised classification of pole structures from experimentally accessible observables, we construct a comprehensive synthetic dataset of scattering line shapes generated within the uniformized $\mathcal{S}$ -matrix formalism introduced in Sec. II. Within this framework, pole singularities are introduced explicitly through the zeros of the Jost-like function [Eq. (5)], allowing each generated amplitude to be labeled unambiguously by its pole content.

We define 35 distinct pole topology classes, each characterized by a triplet $[N_{bt}, N_{bb}, N_{tb}]$, where $N_i \in [0, 4]$ denotes the number of poles residing on the corresponding Riemann sheet. Notably, poles on the $[tt]$ sheet are excluded *a priori* to satisfy the causality requirement, ensuring that the scattered final state cannot precede the scattering interaction [77,78]. This enlarged topology basis is motivated by the analytic structure of coupled-channel scattering amplitudes, in which poles can migrate Riemann sheets under variations of interchannel coupling and background dynamics [74,75]. The resulting classes therefore span a wide spectrum of physically relevant scenarios, ranging from polefree amplitudes to configurations involving multiple nearby poles on different sheets. By construction, this complete set ensures that no analytically allowed pole configurations is excluded *a priori*, allowing the data-driven DNN analysis to discriminate among competing hypotheses without imposing model-dependent restrictions. A summary of class definitions is provided in Table II.

Crucially, to isolate the spectroscopic impact of the left-hand cut, we generated two complementary ensembles of scattering amplitudes: a baseline set of scattering amplitudes without lhc effects, and a set of scattering amplitudes

incorporating the lhc effects via the ansatz in Eq. (7). We normally sampled the cut scale $\Lambda = 280 \pm 50$ [MeV], a range physically motivated by the need to capture TPE as the dominant long-range interaction in the s -wave DD^* sector. This scale aligns with the pion masses used in lattice QCD studies, allowing for a more consistent comparison between our analysis and first-principles calculations. Meanwhile, the channel-specific strength $\zeta_k \in [-1, 1]$ was sampled uniformly to probe the wide spectrum from repulsive to attractive background interactions.

Synthetic amplitudes are constructed over the energy window [3871, 3900] MeV, aligned with the kinematic region probed by LHCb measurements of the $D^0 D^0 \pi^+$ final state [1,2]. This interval is discretized into 57 uniform bins, matching the experimental resolution. For each amplitude, we sample a single energy value per bin using a uniform distribution, simulating the finite resolution inherent in experimental detection. This binwise randomization emulates detector-level smearing and introduces variability across the dataset.

The resulting observable line shapes are computed from the differential event rate,

$$\frac{dN}{d\sqrt{s}} = \rho(\sqrt{s}) \left[|T(\sqrt{s})|^2 + b(\sqrt{s}) \right], \quad (8)$$

where $\rho(\sqrt{s})$ denotes the two-body phase space factor and $b(\sqrt{s})$ models nonresonant background contributions. The phase space factor $\rho(\sqrt{s})$ is taken as a kinematically consistent function based on the Källén triangle function,

$$\rho(\sqrt{s}) = \frac{\lambda^{1/2}(s, m_{D^0 \pi^+}^2, m_{D^0}^2)}{2\sqrt{s}}. \quad (9)$$

To enhance synthetic dataset diversity, the background term $b(\sqrt{s})$ is randomly chosen from predefined functional forms: (i) a quadratic polynomial, (ii) an exponential decay, or (iii) zero contribution. This stochastic background selection introduces nontrivial shape variations and challenges neural networks to focus on pole-driven features rather than background-specific patterns.

The full scattering amplitude $T(\sqrt{s})$ entering Eq. (8) is given by

$$T(\sqrt{s}) = \sum_{k=1}^2 \alpha_k T_{1k}(\sqrt{s}), \quad (10)$$

where α_k denotes the production coefficients modeling the relative rates for the elastic and inelastic channels. The inclusion of the off diagonal term T_{12} , which encodes interchannel coupling, is crucial for capturing interference patterns and resolving pole configurations that may be degenerate in the diagonal elements alone [44].

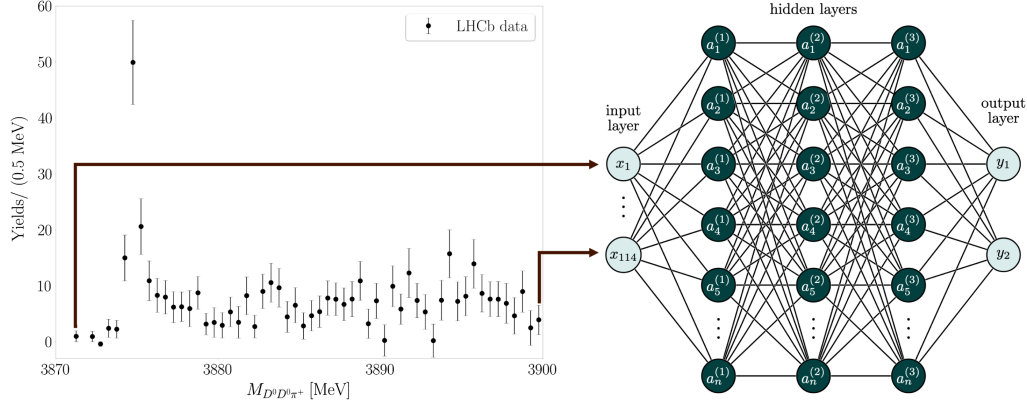


FIG. 2. Schematic workflow for inference using trained DNN classifiers on empirical data. (Left) The experimental invariant mass distribution for the $D^0 D^0 \pi^+$ final state as reported by the LHCb Collaboration [1,2], including statistical and detector-induced uncertainties. (Right) A diagram of trained DNN model with an input layer of 114 nodes corresponding to concatenated energy-amplitude pairs. The illustrated model contains three hidden layers, though architectures with up to five hidden layers were also explored. The output layer contains two nodes, the dimensionality was reduced from 35 as a result of preliminary coarse-grained classification and finalist selection strategy. The LHCb data was extracted from [79], and the TikZ code for the DNN figure was taken from [80].

Each pole is generated as a zero of a Jost-like function defined in Eq. (5) and is constrained to reside near the $D^0 D^{*+}$ threshold. Specifically, pole positions $\sqrt{s}_{\text{pole}}$ are sampled within the following bounds:

$$\begin{cases} \varepsilon_2 - 5.0 \leq \text{Re}\sqrt{s}_{\text{pole}} \leq \varepsilon_2 + 10.0, \\ 0.0 \leq |\text{Im}\sqrt{s}_{\text{pole}}| \leq 20.0. \end{cases} \quad (11)$$

This ensures the poles are located within the physically relevant region influenced by threshold effects and coupled-channel dynamics.

Specifically, we construct a 114-dimensional input vector by concatenating two components: the discretized COM energy values $\{\sqrt{s}_i\}$ and their corresponding differential event rates $\{dN/d\sqrt{s}_i\}$ across 57 bins. This dual encoding enhances the expressivity of the input representation, enabling the network to capture subtle variations across the line shape. The complete dataset is composed of labeled pairs (x^j, y^j) , where $y^j \in \{n \in \mathbb{Z} | 0 \leq n \leq 34\}$ denotes the class label corresponding to one of the 35 pole configurations, resulting in a dataset of 350,000 line shapes. For each class, we generate 10^4 synthetic samples, resulting in a well-balanced dataset suitable for supervised learning. The data are then randomly split into training and validation subsets using a 4:1 ratio, enabling consistent training and robust out-of-sample performance evaluation.

2. Empirical dataset for inference

To apply the trained DNN classifiers to the real-world data, we extract the experimental line shape from the $D^0 D^0 \pi^+$ invariant mass distribution reported by the LHCb Collaboration [1,2]. Each spectrum point includes a central energy with uncertainties on the x -axis and a

central amplitude with uncertainties on the y -axis. These uncertainties arise from both statistical fluctuations and detector resolution effects. To address this, we construct inference-ready samples that reflect the statistical and systematic uncertainties of the LHCb spectrum, as shown in the left panel of Fig. 2.

To emulate realistic detector-level effects, we generate independent realizations of the experimental spectrum. For each pseudoempirical realization, one energy-amplitude pair is sampled per reported data point. Energy values are uniformly sampled from the per-event uncertainty range. To simulate amplitude fluctuations, we explore two uncertainty models. First, the amplitude values are drawn from a Gaussian distribution centered at the reported value with a standard deviation equal to one-third of the quoted upper uncertainty. To complement this, we also implement a uniform sampling strategy over the full error band of the amplitude values. This assumption is more conservative and agnostic, avoiding any bias to the central value.

For each realization, the corresponding differential event rates $dN/d\sqrt{s}$ are taken directly from the experimental bin counts with no further smoothing or rescaling. The sampled energy values and bin heights are then concatenated to form an input vector $x^{\text{exp}} \in \mathbb{R}^{114}$, consistent with the input format used during training. Crucially, these sampling procedures are carried out independently for each trained DNN. That is, each DNN receives its own unique set of 1,000 empirically derived line shapes, ensuring that the final class probabilities reflect both model uncertainty from training variations across DNNs and data uncertainty from fluctuations in the empirical spectrum.

With both the synthetic datasets for supervised learning and the protocol for LHCb-based inference inputs outlined,

we now proceed to the design of the deep neural network architecture.

B. Neural network architecture and design

To map synthetic line shapes into their corresponding pole configuration classes, we employ supervised learning using deep neural networks (DNNs). Each input vector $x^j \in \mathbb{R}^{114}$ is fed into a fully connected feed-forward network f_θ , parametrized by a set of weights and biases θ . The network outputs a probability vector over $\mathcal{Z}$ number of classes via a softmax activation function [81–83],

$$\hat{y}^j = \text{Softmax}[f_\theta(x^j)] \in \mathbb{R}^{\mathcal{Z}}, \quad (12)$$

with the predicted class given by $\arg \max_l \hat{y}_l^j$. Training the models are performed by minimizing the categorical cross-entropy loss function [83,84], defined as

$$\mathcal{L} = - \sum_{l=1}^{\mathcal{Z}} y_l \log(\hat{y}_l), \quad (13)$$

where $y_l \in \{0, 1\}$ is the true one-hot encoded true class label and $\hat{y}_l \in [0, 1]$ is the predicted class probability. To improve convergence and expressiveness, each hidden layer employs the rectified linear unit (ReLU) activation function [85],

$$\text{ReLU}(z) = \max(0, z), \quad (14)$$

which has been shown to promote sparsity and mitigate the vanishing gradient problem [86].

To ensure robust and interpretable classification of pole configurations, we implemented two complementary approaches to neural network design and tuning: (1) a manual grid search, where model architectures and hyperparameters were explicitly specified, and (2) an automated, adaptive search using the Optuna optimization framework [87]. These approaches enabled a broad and deep exploration of the hyperparameter and architecture space, providing a diverse ensemble of models for downstream inference.

All models were implemented using PyTorch [88] and evaluated on a held-out validation set using standard classification metrics. Each model was trained using the Adam optimizer [89] with appropriate regularization imposed via dropout [90] and weight decay [91].

In manual grid search strategy, each model was selected by specifying the number of hidden layers $h_1 \in [2, 5]$, neurons per layer $h_2 \in \{100x_0 | x_0 = 2, 3, \dots, 8\}$, dropout $h_3 \in [0.025, 0.200]$. The learning rate was chosen as $h_4 = 10^{-5}$ and the weight decay $h_5 = 10^{-5}$ all throughout. This grid search allowed us to establish baseline architectures that are both interpretable and tunable.

To complement the manual design process and explore a broader space of architectures and hyperparameters, we

employed Optuna [87], a modern hyperparameter optimization framework based on Bayesian sampling and pruning strategies [92,93]. For each trial, Optuna searched the architecture and training parameter space: $h_1 \in [3, 5]$, sampled uniformly $h_2 \in [200, 600]$, sampled from the continuous range $h_3 \in [0.0, 0.3]$, sampled log-uniformly $h_4 \in [10^{-5}, 10^{-3}]$, and $h_5 \in [10^{-6}, 10^{-3}]$. The activation function was fixed to ReLU [85] for compatibility and comparability. Each Optuna trial involved training a candidate architecture for 75 epochs on the training set. The median pruner was enabled to automatically terminate underperforming trials early. This pruning strategy saved compute time and focused optimization on promising regions of architecture space. Importantly, Optuna trials are required to achieve $\geq 45\%$ validation accuracy within the 75-epoch window to be retained for downstream ensemble use.

This two-pronged strategy, manual tuning for interpretability and Optuna for adaptive exploration, yielded a robust ensemble of 216 trained DNN models. Their combined insights enabled a deeper understanding of the pole structure of the T_{cc}^+ state and formed the basis for the experimental inference. All codes used in this analysis are openly available for reproducibility in [94].

C. Coarse-grained classification and finalist class selection

Our initial classification task involved distinguishing among 35 distinct pole configurations, each representing a unique arrangement of singularities across the $[bt]$, $[bb]$, and $[tb]$ Riemann sheets. However, the attempts to train three DNN classifiers with amplitudes (sans lhc effects) directly on this 35-output node architecture showed limited success. In particular, the training accuracy plateaued below 60% and validation performance stagnated. This indicated that the classification scheme was too fine-grained relative to the degeneracies present in the line shapes. Specifically, several pole configurations led to morphologically indistinguishable patterns and consequently confounded the DNN model's ability to learn meaningful distinctions. Moreover, the computational burden of training DNNs on such large output space proved to be inefficient. Hence, the DNN models are expected to be incapable of reliably inferring the pole structure of the T_{cc}^+ from the LHCb data.

To address this, we implemented a coarse-graining strategy (Fig. 3) that grouped the 35 pole configurations into seven broader structural families based on shared topological features including the number and type of poles on each Riemann sheet. This structural grouping was organized as follows: SF1: classes 00–04, corresponding to the simplest pole topologies, SF2: classes 05–09, involving up to two poles distributed across different sheets. Moreover, SF3: classes 10–14 and SF4: classes 15–19, features three-pole configurations with different sheet

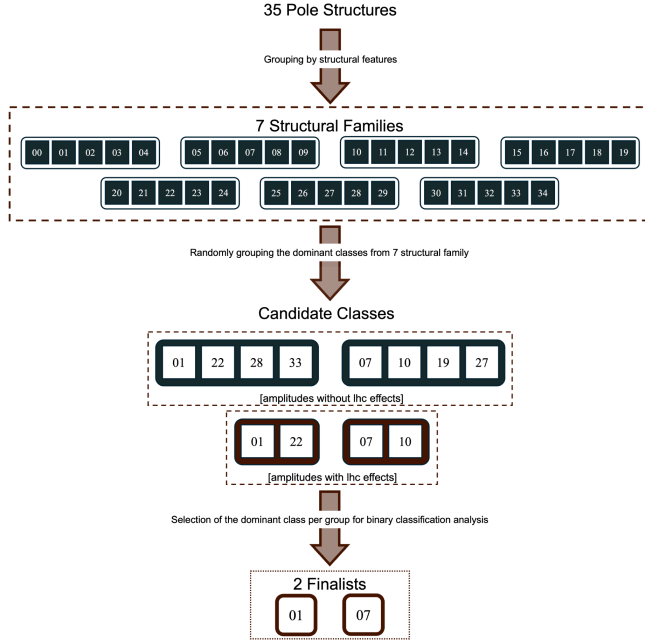


FIG. 3. Schematic overview of the coarse-grained classification and finalist selection strategy. The original 35 pole classes (top row) were grouped into seven structural families (second row) to address degeneracies in the line shapes. Dominant classes from each family (third row) formed an 8-class candidate set for amplitudes without lhc effects and 4-class candidate set for amplitudes with lhc effects. These classes were randomly grouped into two sets to select the dominant class from each set (last row) for an extensive binary classification analysis using an ensemble of 50 DNNs.

distributions. Finally, SF5: classes 20–24, SF6: classes 25–29, and SF7: classes 30–34 contains the most complex arrangements with up to four poles. One may refer to Table II for a summary of this structural categories.

For the amplitudes without lhc effects, we trained six independent DNNs (three are manually tuned and the other three are Optuna-optimized) for each of the seven structural families on a 5-output node classification task. To explore the effects of model complexity, we employed architectures with increasing depth and width. This ensemble strategy, from simple to more expressive networks, enables us to robustly generalize the results and mitigate potential overfitting associated with any single architecture. Building upon this, Fig. 4(a) compares the mean validation accuracy achieved by the manually tuned and Optuna-optimized DNN models for each of the seven structural families. The results show that hyperparameter and architecture optimization via Optuna yields modest but consistent improvement in accuracy for most families, particularly in SF1, SF3, and SF6, where gains of about 1–2% are evident. These families also exhibit the highest overall accuracies, reflecting a clearer separation between candidate configurations within these structural categories. In contrast, SF2 and SF5 show minimal differences between manual and Optuna tuning, suggesting that model performance in these families is less sensitive to hyperparameter adjustments. The lowest accuracies occur for SF4 and SF7, regardless of DNN design strategy, due to pronounced ambiguities between specific pole configurations. In SF4, the primary confusion arises between $[0_{bt}, 2_{bb}, 1_{tb}]$ and $[0_{bt}, 1_{bb}, 2_{tb}]$ pole topologies, as well as between $[1_{bt}, 0_{bb}, 2_{tb}]$ and $[1_{bt}, 1_{bb}, 1_{tb}]$, which hampers the DNN classifier’s ability to confidently separate these cases. Similarly, in SF7, misclassifications are largely driven by overlapping features between $[2_{bt}, 0_{bb}, 2_{tb}]$ and $[2_{bt}, 1_{bb}, 1_{tb}]$, as well as between $[1_{bt}, 2_{bb}, 1_{tb}]$ and $[1_{bt}, 1_{bb}, 2_{tb}]$ pole topologies. These structural families thus represent inherently more challenging classification tasks, where subtle differences in the line shapes reduces DNNs discriminative power.

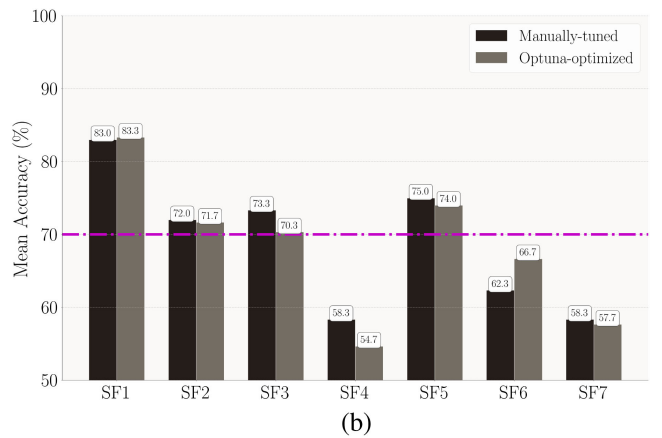
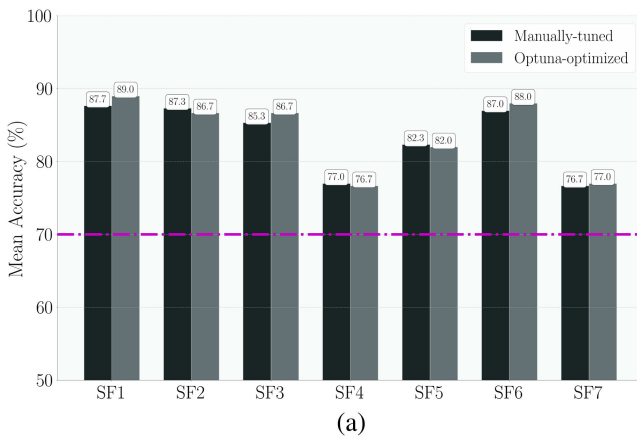


FIG. 4. Mean validation accuracy of DNN models across seven structural families. Each bar represents the average over three independently trained manually tuned and Optuna-optimized DNN models per family. (a) shows the classification performance of DNNs trained on amplitudes without lhc effects and the (b) displays the classification performance of DNNs trained on amplitudes with lhc effects. The fuchsia horizontal dashed-dotted line marks the 70% validation accuracy threshold; i.e., predictions below this threshold are insufficiently reliable for subsequent inference using LHCb data.

It is worth noting however, that despite these average gains, many of the individual Optuna-optimized DNN models we tested were not ultimately helpful because they tended to overfit. In our Optuna setup, candidate architectures are selected based on achieving high validation accuracy at relatively early epochs, a criterion that can favor models which quickly memorize the training data but fail to generalize when trained to convergence. This reflects a well-documented issue in neural architecture search (NAS), where reliance on low-fidelity evaluations or short-horizon validation signals often misranks candidate architectures [95]. More broadly, this challenge reflects the inherent tension between model expressivity and generalization: highly expressive architectures discovered through automated search can easily overfit unless balanced by appropriate regularization or multiobjective selection criteria. Nonetheless, the searched models that failed to generalize were systematically rejected from our analysis.

When the amplitude samples from the LHCb spectrum is fed to the DNNs trained on amplitudes without lhc effects, this strategy yielded eight dominant classes. In particular, class 01 ($[1_{bt}, 0_{bb}, 0_{tb}]$) consistently dominated in the first structural family SF1. Meanwhile, class 07 ($[1_{bt}, 1_{bb}, 0_{tb}]$) was strongly favored in SF2. In SF3 and SF4, the leading configurations were class 10 ($[3_{bt}, 0_{bb}, 0_{tb}]$) and class 19 ($[1_{bt}, 1_{bb}, 1_{tb}]$), respectively. The pole configuration $[1_{bt}, 1_{bb}, 1_{tb}]$ had previously been proposed by [49] as a leading pole structure for the $P_{\psi}^N(4312)^+$, hinting at a recurring structural pattern in near-threshold exotic states. From SF5, class 22 ($[0_{bt}, 0_{bb}, 4_{tb}]$) was identified, suggesting that in some cases, multiple poles on a single sheet can approximate the amplitude structure. In SF6, despite achieving a relatively high training and validation accuracy of approximately 87% across six DNN models, the DNNs could not decisively select between class 27 ($[1_{bt}, 0_{bb}, 3_{tb}]$) and class 28 ($[0_{bt}, 1_{bb}, 3_{tb}]$). This ambiguity likely stems from a structural mismatch between the classes in SF6 and the empirical features of the LHCb data, compelling the DNNs to distribute their predictions across suboptimal class candidates. Nevertheless, both classes were retained in the subsequent stage to ensure that no potentially viable configuration was prematurely ruled out. Finally, in SF7, class 33 ($[1_{bt}, 2_{bb}, 1_{tb}]$) was favored, a configuration also suggested in [52] for the $P_{\psi}^N(4312)^+$ state, further strengthening the hypothesis of structural analogies among exotic hadrons.

When the same analysis is repeated using amplitudes that include lhc effects, the overall classification performance follows the same qualitative trends but with significantly reduced accuracies across all structural families. This is clearly visible in Fig. 4(b), where the mean validation accuracies fall by approximately 5%–20%. The decrease is most pronounced in SF4, SF6, and SF7,

whose accuracies drop below 70% regardless of whether the models are manually tuned or Optuna-optimized. Even families that perform comparatively well in the case where lhc effects are not incorporated, such as SF3 and SF6, exhibit a substantial loss when lhc effects are present, reflecting the increased overlap in the line shapes. In contrast, SF1 and SF5 remain the most robust even under lhc contamination, with accuracies that stay marginally above 70% accuracy. Consequently, the choice to set the 70% reliability threshold was made after a thorough inspection of the confusion matrices from the trained classifiers, and thereby rejecting the pole topology classes involved in these structural families. Thus, for the DNN analysis involving amplitudes with lhc effects, only the pole topologies $[1_{bt}, 0_{bb}, 0_{tb}]$, $[1_{bt}, 1_{bb}, 0_{tb}]$, $[3_{bt}, 0_{bb}, 0_{tb}]$, and $[0_{bt}, 0_{bb}, 4_{tb}]$ was considered for the subsequent class-selection stage.

To further refine the analysis, we randomly partitioned the dominant candidate configurations into two groups for comparative evaluation. In particular, for the case involving amplitudes without lhc effects, we have G1: classes {01, 22, 28, 33} and G2: classes {07, 10, 19, 27}. Meanwhile, for the dominant candidate classes associated with amplitudes contaminated by lhc effects, we performed two separate binary classification analysis such that G1: classes {01, 22} and G2: classes {07, 10}. Each group was treated as a standalone classification task, where the objective was to identify the most likely configuration within the group that best explains the LHCb data. For each group, we trained eight independent DNN models (four are manually tuned and the remaining four models are Optuna-optimized), each operating on a four-output and binary-output node architecture corresponding to the configurations within its assigned set.

Upon inference using the LHCb data, the classifiers trained both on amplitudes with and without lhc effects consistently identified class 01 ($[1_{bt}, 0_{bb}, 0_{tb}]$) as the most probable configuration within G1, and class 07 ($[1_{bt}, 1_{bb}, 0_{tb}]$) as dominant in G2. These two configurations thus advanced as the finalist candidates, representing two structurally distinct hypotheses for the nature of the T_{cc}^+ state.

D. Binary classification analysis: Molecular vs compact interpretations

In the final stage of the classification pipeline, we conducted an extensive analysis comparing the two finalist configurations: class 01 ($[1_{bt}, 0_{bb}, 0_{tb}]$) and class 07 ($[1_{bt}, 1_{bb}, 0_{tb}]$). To ensure robust and statistically significant results, we trained an ensemble of 50 independently initialized DNN models, each tasked with a binary classification between the two pole topologies. Consistent with the earlier subgroup training results, where manual tuning performed on par with Optuna-optimized models, we

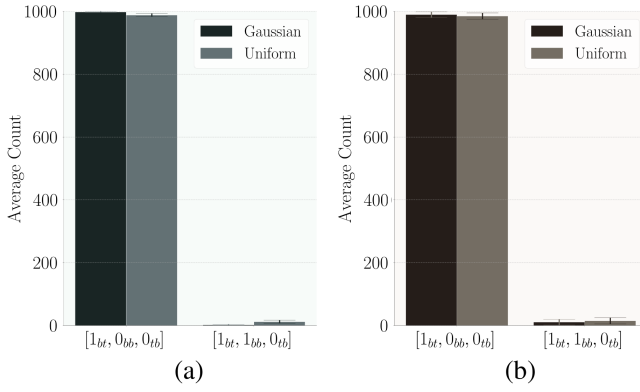


FIG. 5. Molecular vs compact tetraquark picture. Ensemble-averaged predicted class distributions from 100 DNN models trained on (a) amplitudes without lhc effects and (b) on amplitudes with lhc effects. Predictions are made on LHCb data using two amplitude sampling assumptions: Gaussian and uniform distribution. The mean classification counts are obtained from 50,000 LHCb amplitude realizations, with predictions aggregated across an ensemble of 50 DNNs. Error bars represent standard errors of the ensemble means.

adopted only the manual tuning strategy at this stage. In the binary classification setting, the Optuna optimization framework loses its practical advantage since the DNN models trained on amplitudes (sans lhc effects) typically converge to near-maximum validation performance after ~ 300 – 500 epochs, rendering the long search cycles required by Optuna computationally prohibitive. The manually tuned ensemble was deliberately diversified to explore the model space thoroughly, with architectures comprising 2–5 hidden layers and the number of neurons per hidden layer are selected in the range [200, 900] (see Appendix E for the complete list of the DNN models, their classification accuracy, and predicted class distributions). Across this ensemble, validation accuracies hovered in the 93%–95% range for DNNs trained on amplitudes without lhc effects and 81%–87% classification accuracy for DNNs trained on amplitudes with lhc effects, underscoring both the robustness of the manual tuning approach and the high separability of the two pole topologies in the reduced hypothesis space.

The final ensemble-averaged class distribution with the corresponding standard error of the mean are shown in Fig. 5. Under Gaussian and uniform sampling strategies in the intensity axis of the LHCb data, the pole configuration $[1_{bt}, 0_{bb}, 0_{tb}]$ is overwhelmingly favored, the class which correspond to a single pole on the $[bt]$ sheet. These results exhibit both sharp dominance and low ensemble variance, indicating strong agreement across models and sampling assumptions. Notably, in contrast to conventional fitting procedures, such as the $\Lambda(1405)$ analysis in [96], where increasing model complexity (e.g., by introducing more poles and free parameters) naturally improves the fit quality

as measured by the reduced chi-square χ^2_ν , our trained DNNs consistently favor the simplest adequate explanation of the LHCb data. It avoids overfitting by rejecting more complex pole topologies that might artificially improve agreement with fine features of the line shape. For instance, the pole configuration $[1_{bt}, 1_{bb}, 0_{tb}]$ which includes additional pole on the $[bb]$ sheet, could in principle refine the line shape and absorb residual binning or resolution effects. However, this more ambiguous and flexible topology is not selected by the classifiers, demonstrating the model’s preference for parsimony and its insensitivity to artifacts that might otherwise bias a traditional fit.

This minimal inference strengthens the robustness of the result and supports a topologically grounded interpretation of the T_{cc}^+ as a hadronic molecule. The interpretation is made within the $\mathcal{S}$ -matrix pole topology formalism. In this framework, the nature of the near-threshold state is inferred from the analytic structure of the observable scattering amplitude. The interpretation is also consistent with the possibility of meson-exchange due to the presence of light quarks inside the tetraquark state. However, since the present framework does not contain explicit color or gluon degrees of freedom, it cannot directly rule out the compact tetraquark interpretation at the QCD level.

The interpretation that the T_{cc}^+ is a loosely bound state requires careful consideration at this point, since a virtual state pole of the $D^0 D^{*+}$ channel may traverse onto the $[bt]$ sheet due to strong interchannel coupling [70,74,97]. Despite this distinction, bound and virtual states are encompassed within the broader classification of hadronic molecules [98]. In any case, a reliable interpretation must be substantiated by the corresponding pole trajectory in the complex plane, which can only be established using model-dependent methods [99–101].

Moreover, we emphasize that the role of the DNN is not limited to handling left-hand cut effects. If the relevant pole is sufficiently far from the lhc, then its effect may not be that significant. However, the primary purpose of the DNN in the present work is to reduce the large space of possible pole configurations into a smaller set of physically relevant candidates using a model-independent approach. The inclusion of simulated lhc, therefore, serves as a robustness test rather than being the sole motivation for the DNN usage.

To close this section, while the classification results unambiguously point to an isolated pole in the $[bt]$ sheet, the neural network framework is inherently agnostic to underlying dynamics that generate such a pole. In particular, it does not provide direct access to the pole position and trajectory of the pole in the complex plane. Hence, we now transition to pole parameter extraction through the uniformized $\mathcal{S}$ -matrix parametrization complemented by the model-dependent $\mathcal{K}$ -matrix analysis.

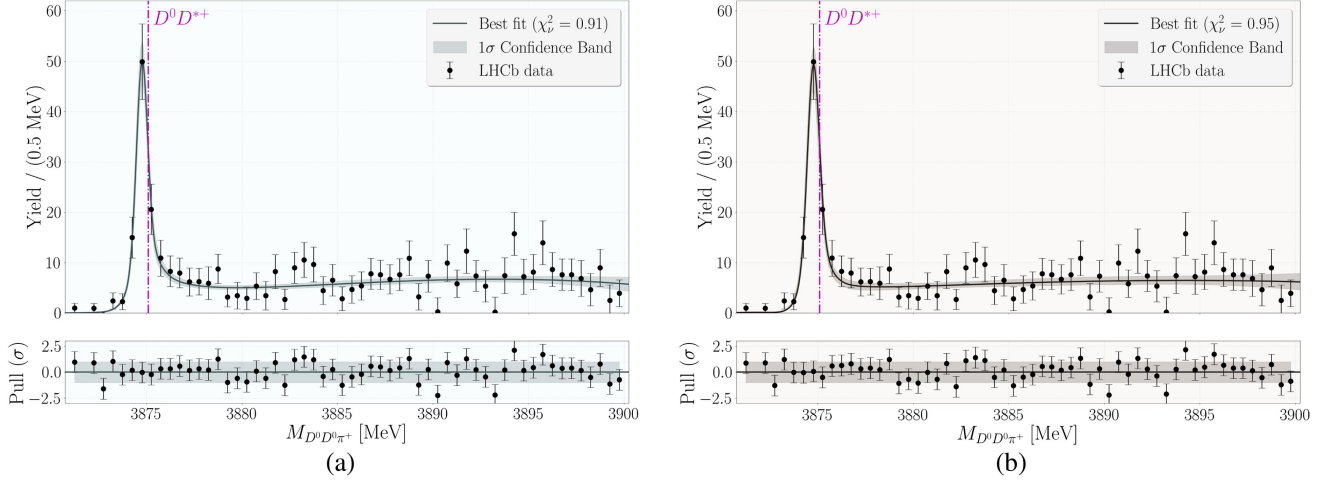


FIG. 6. Best fit line shape and statistical diagnostics for the isolated pole in the $[bt]$ sheet model via (a) the uniformized $\mathcal{S}$ -matrix and (b) $\mathcal{K}$ -matrix framework. The experimental data points used are from the LHCb [2,79].

IV. POLE PARAMETER EXTRACTION

In this section, we extract the pole position of the T_{cc}^+ by performing an extensive parametric fit to the LHCb invariant mass spectrum using the uniformized coupled-channel $\mathcal{S}$ -matrix parametrization discussed in Sec. II. To complement this, we also extract the pole parameters via a model-dependent approach using the two-channel $\mathcal{K}$ -matrix formalism. Both fitting procedure is implemented using the Minuit minimization framework [102] and is designed to determine the resonance mass and width from the analytic structure of the scattering amplitude.

From the minimal and robust classification results of the DNN workflow in Sec. III, the fitting procedure only consider the minimal scenario of a single isolated pole located on the $[bt]$ sheet. The pole position is parametrized directly as

$$\sqrt{s}_{T_{cc}^+}^{(\text{pole})} = M_{T_{cc}^+} - i \frac{\Gamma_{T_{cc}^+}}{2}, \quad (15)$$

where $M_{T_{cc}^+}$ is the resonance mass and $\Gamma_{T_{cc}^+}$ is the resonance decay width. The predicted invariant mass spectrum is given by Eq. (8), and we modeled the nonresonant background as a smooth quadratic polynomial $b(\sqrt{s}) = c_0 + c_1\sqrt{s} + c_2s$ with coefficient c_i determined in the fit.

To eliminate binning bias and accurately capture the sharp threshold dynamics of the T_{cc}^+ , the predicted yield is calculated by integrating the probability density function [103] over the 0.5 [MeV] bin width. In our numerical implementation, this is approximated via a multipoint grid average with a fine spacing of 0.005 [MeV]. Crucially, the theoretical model is convolved with a double-Gaussian resolution kernel to match the LHCb detector resolution

following their spectroscopic paper on T_{cc}^+ [2].² The convolution is performed on a high-resolution grid extended by 5.0 [MeV] to prevent edge artifacts.

The fit parameters are extracted via a multistart χ^2 cost function minimization using the MIGRAD algorithm provided by Minuit [102]. To mitigate sensitivity to local minima in the high-dimensional parameter space, we performed 10^3 independent fits with randomized initial conditions within physically motivated parameter bounds.

A. Uniformized $\mathcal{S}$ -matrix

To provide a model-independent baseline for the pole extraction, we first carry out a parametric fit via the uniformized $\mathcal{S}$ -matrix. Here, the mass and width of the T_{cc}^+ are directly treated as fitting parameters. Figure 6(a) presents the invariant mass distribution of $D^0 D^0 \pi^+$ candidates with the best-fit result. The solid aqua green curve accurately reproduces the sharp, near-threshold enhancement corresponding to the T_{cc}^+ state, capturing the rapid rise at the $D^0 D^{*+}$ threshold and the subsequent $D^0 D^0 \pi^+$ decay. The surrounding gray shaded region represents the 1σ confidence band, generated via Monte Carlo propagation of the fit parameter uncertainties. The confidence band is

²The double-Gaussian convolution kernel reads

$$\mathcal{R}_{\text{LHCb}}(\sqrt{s}, \sqrt{s}_i) = \kappa_{\text{core}} \mathcal{G}(\sqrt{s}, \sqrt{s}_i, \sigma_{\text{core}}) + (1 - \kappa_{\text{core}}) \mathcal{G}(\sqrt{s}, \sqrt{s}_i, \sigma_{\text{tail}}).$$

In particular, the core component is a narrow Gaussian with $\sigma_{\text{core}} \approx 0.25$ [MeV], represents the primary detector resolution. The tail component is a broader Gaussian with $\sigma_{\text{tail}} \approx 0.65$ [MeV] which captures the extended resolution effects. These components are combined with a core fraction $\kappa_{\text{core}} = 0.70$, yielding a total effective root mean square (RMS) resolution of approximately 0.40 [MeV] [2].

notably narrow, which indicates that the model parameters are constrained with high precision by the data.

The lower panel displays the pull distribution showing that the pulls are randomly distributed around zero without systematic structures, and the majority of data points fall within the expected $\pm 1\sigma$ band (shaded gray rectangle). This statistical behavior, corroborated by the reduced chi-square value of $\chi^2_\nu = 0.91$, confirms that an isolated pole in the $[bt]$ sheet model from the uniformized $\mathcal{S}$ -matrix framework provides a robust description of the experimental spectrum.

B. $\mathcal{K}$ -matrix

A complementary fit was also performed using a $\mathcal{K}$ -matrix parametrization to extract the numerical values of the pole parameters. By construction, the $\mathcal{K}$ -matrix formalism preserves the unitarity of the $\mathcal{S}$ -matrix, thus ensuring conservation of probability [104]. In the two-channel case relevant to the analysis, the first and second channels incorporated here are the $D^0 D^0 \pi^+$ and $D^0 D^{*+}$ channels. The $\mathcal{K}$ -matrix is then constructed as a real symmetric 2×2 matrix,

$$\mathcal{K} = \begin{pmatrix} \mathcal{K}_{11} & \mathcal{K}_{12} \\ \mathcal{K}_{21} & \mathcal{K}_{22} \end{pmatrix}, \quad (16)$$

where the corresponding elements are treated as free parameters in the fitting scheme. In general, each element of the $\mathcal{K}$ -matrix can be written as the sum of the resonant and nonresonant contributions,

$$\hat{\mathcal{K}}_{ij} = \sum_a \frac{\gamma_{ai} \gamma_{aj} m_a \Gamma_a^0}{m_a^2 - m^2} + \hat{c}_{ij}. \quad (17)$$

Here the γ_{ai} denotes the coupling strength of resonance a to the i^{th} channel, while m_a and Γ_a^0 represent the mass and width of the resonance. The term $\hat{c}_{ij}$ accounts for the nonresonant contribution.

We formalize and modify the general representation of the $\mathcal{K}$ -matrix elements using the result of an isolated pole from Sec. III as motivation. The presence of the isolated pole structure seen in our DNN result allows for the elimination of the resonance mass/width dependence in the usual $\mathcal{K}$ -matrix element equation as seen in Eq. (17), effectively reducing the $\mathcal{K}$ -matrix elements into energy-independent terms. This also reduces the number of free parameters present during the fitting scheme, allowing for a more accurate extraction of pole parameter values in a fitting perspective while also capturing the relevant resonance behavior near the peak and threshold.

The obtained $\mathcal{K}$ -matrix parameters are used to build the scattering T -matrix using the relation defined by

$$T = \rho^{1/2} \mathcal{K} (I - i\mathcal{K})^{-1} \rho^{1/2}, \quad (18)$$

TABLE I. Extracted best fit pole parameters for an isolated pole on the $[bt]$ sheet model.

Pole parameters	Uniformized $\mathcal{S}$ -matrix	$\mathcal{K}$ -matrix
$M_{T_{cc}^+}$ [MeV]	3874.783 ± 0.042	3874.749 ± 0.072
$\Gamma_{T_{cc}^+}$ [MeV]	0.097 ± 0.010	0.232 ± 0.031

where I is the identity matrix and ρ is the phase-space matrix which is a 2×2 diagonal matrix. The T -matrix is then used to reconstruct the invariant mass distribution using the differential event rate given by Eq. (8) that was introduced in Sec. III.

We take advantage of the relation between the $\mathcal{S}$ - and T -matrices mentioned in Sec. II and use it together with the obtained best-fit values to numerically compute the poles of the $\mathcal{S}$ -matrix, which correspond to peaks seen in the scattering data. The resulting fit, as shown in Fig. 6(b), reproduces the narrow enhancement across the concerned energy region. The reduced chi-square value of the best fit is $\chi^2_\nu = 0.95$, which suggests a good agreement between the model and data.

C. Pole position and trajectory of T_{cc}^+

In the fitting scheme, the global minimum was identified by the smallest χ^2 value among all valid converged trials. Table I summarizes the best-fit pole parameters with their associated uncertainties. One may also refer to Appendix D for the complete list of the extracted fitting parameters. The extracted mass is remarkably consistent across both parameterizations, reflecting the clear threshold enhancement in the LHCb data.

However, we observed a rather large deviation with the extracted width across both parametric framework. The discrepancy between the uniformized $\mathcal{S}$ -matrix ($\Gamma_{T_{cc}^+} \sim 98$ [keV]) and $\mathcal{K}$ -matrix ($\Gamma_{T_{cc}^+} \sim 232$ [keV]) highlights the inherent sensitivity of the T_{cc}^+ width to the choice of unitarization scheme. Given the detector resolution ($\sigma \approx 400$ [keV]) significantly exceeds both extracted widths, these two values define the envelope of our systematic uncertainty, consistently placing the T_{cc}^+ as a narrow, near-threshold state regardless of the specific mathematical framework employed.

Nonetheless, our result is physically consistent and sits between the LHCb's discovery paper [1] which employed a Breit-Wigner parametrization extracting an apparent width of $\Gamma_{\text{B.W.}} = 410.0 \pm 165.0$ [keV] and the rigorous spectroscopic LHCb analysis [2], which employed a unitarized Breit-Wigner parametrization with momentum-dependent width, extracting a pole width $\Gamma_{\text{U.B.W.}} = 48.0 \pm 2.0$ [keV].

We must point out however, that these extracted values should be viewed as effective pole parameters within a two-body, two-channel approximation. In particular, our model employs an effective two-body treatment of the

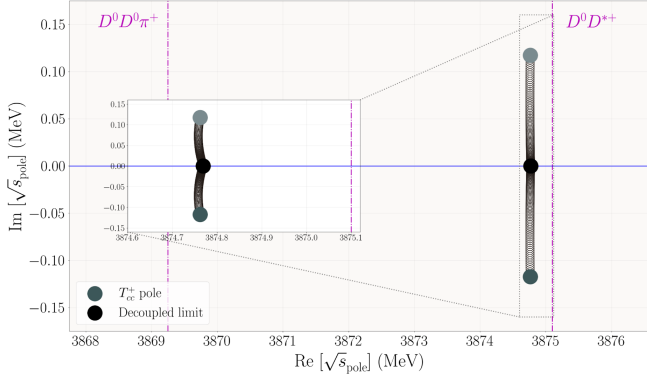


FIG. 7. Trajectory of the pole found in the $[bt]$ sheet. Turning off the coupling to the $D^0 D^{*+}$ channel shifts the pole towards the real energy axis, which is a signature of an unstable bound state.

$D^0 D^0 \pi^+$ system; Ref. [23] argued that neglecting explicit three-body $DD\pi$ dynamics with its associated branch cut structure and the energy-dependent pion exchange dynamics inherently leads to an overestimation of the pole width. Moreover, our analysis is restricted to a two-channel framework, neglecting the $D^+ D^{*0}$ threshold. Since this third mass threshold lies only ~ 1.4 [MeV] above the $D^0 D^{*+}$ channel, the omission likely forces the primary channels to absorb the spectral density that would otherwise be distributed across the isospin-multiplet, further deviating the extracted pole width.

In addition to the pole parameter extraction, we resolve the loosely-bound state ambiguity within the $[bt]$ pole topology. As mentioned in Sec. III, an isolated pole in the $[bt]$ sheet cannot be definitively used to support the loosely bound state interpretation for the T_{cc}^+ because such a pole topology can still be reached by a virtual state pole of the $D^0 D^{*+}$ channel due to strong interchannel coupling [70,74,97].

To break this degeneracy, we trace the pole trajectory across the Riemann sheets by switching off the coupling to the second channel (Fig. 7). Within our $\mathcal{K}$ -matrix formalism, this is equivalent to setting the off diagonal elements to zero to indicate the type of the resulting one-channel pole. In this case, the pole moves toward the real energy axis below the $D^0 D^{*+}$ threshold. Using the pole-counting argument, the observed pole in the $[bt]$ sheet is classified as a loosely-bound state of the $D^0 D^{*+}$ system.

Ultimately, while the extracted pole parameters for the T_{cc}^+ are subject to the inherent limitations of a simplified two-body, two-channel parametrization, this analysis serves as a successful proof-of-concept for our broader methodology. We have demonstrated that a *two-pronged strategy* (utilizing DNNs for topological model selection followed by multitrial Minuit-based pole extraction) provides

a robust framework for identifying the nature of exotic hadron candidates.

The simplifications employed within the uniformized $\mathcal{S}$ -matrix model were tactical necessities. In the present work, we utilized a two-channel uniformized $\mathcal{S}$ -matrix framework [43] to leverage its capacity for independent and controlled pole position formulation [44] for a rigorous inference-assisted pole topology DNN classification. While alternative parametrizations, such as the $\mathcal{K}$ -matrix, could in principle accommodate more channels or three-body effects, they lack direct mapping between $\mathcal{S}$ -matrix analyticity and pole locations required to generate the high-fidelity, labeled training datasets essential for our data-driven approach.

V. CONCLUSIONS

This work implemented a two-pronged strategy combining inference-assisted pole topology classification and parametric fitting to extract pole parameters to analyze the doubly charmed tetraquark candidate T_{cc}^+ . A hierarchical DNN classification scheme overcame intrinsic degeneracies among line shapes of viable pole topologies. While lhc effects introduce additional analytic structures and moderately reduce classification performance, the dominant pole topology remains robust: DNN classifiers trained on amplitudes with and without lhc effects consistently feature an isolated pole on the $[bt]$ sheet.

Quantitative pole parameter extraction demonstrate that while the mass of the T_{cc}^+ remains remarkably stable across disparate parametric schemes, the observed sensitivity in the pole width highlights systematic trade-offs of a simplified coupled-channel framework. Finally, pole trajectory analysis supports a quasibound $D^0 D^{*+}$ molecular interpretation for the T_{cc}^+ .

Moving forward, regression-based deep neural networks offer a promising complementary avenue for pole parameter determination, enabling a direct, data-driven mapping from line shapes to pole positions beyond conventional parametric fitting schemes.

ACKNOWLEDGMENTS

JBP acknowledges the scholarship support through DOST-ASTHRDP. K. T. R. C. acknowledges the support from the Pisayan graduate scholarship.

DATA AVAILABILITY

The data and codes that support the findings of this article are openly available in [94].

APPENDIX A: POLE-TOPOLOGY CLASSES IN A COUPLED TWO-CHANNEL SYSTEM

We define 35 distinct pole topology classes, each characterized by a triplet $[N_{bt}, N_{bb}, N_{tb}]$ with $N_i \in [0, 4]$ counting poles on the corresponding Riemann sheet (Table II). Poles on the $[tt]$ sheet are excluded to satisfy causality [77,78]. This basis captures the possible analytic structures of coupled-channel amplitudes, where poles migrate between sheets under varying

TABLE II. Pole-topology classification in a coupled two-channel system. The 35 configurations (00–34) are labeled by $[N_{bt}, N_{bb}, N_{tb}]$, indicating pole multiplicities on the $[bt]$, $[bb]$, and $[tb]$ sheets. They are organized into seven structural families (SF1–SF7) based on shared analytic structure.

Label	Pole structure	Physical interpretation
SF1		
00	$[0_{bt}, 0_{bb}, 0_{tb}]$	Pure nonresonant background
01	$[1_{bt}, 0_{bb}, 0_{tb}]$	Hadronic molecule (loosely bound state or a virtual state) [75]
02	$[0_{bt}, 1_{bb}, 0_{tb}]$	Not possible for ERE [70]
03	$[0_{bt}, 0_{bb}, 1_{tb}]$	Virtual state
04	$[2_{bt}, 0_{bb}, 0_{tb}]$	Two molecular states
SF2		
05	$[0_{bt}, 2_{bb}, 0_{tb}]$	...
06	$[0_{bt}, 0_{bb}, 2_{tb}]$	...
07	$[1_{bt}, 1_{bb}, 0_{tb}]$	May be interpreted as a compact state coupled to the lower channel
08	$[1_{bt}, 0_{bb}, 1_{tb}]$	Hadronic molecule + virtual state
09	$[0_{bt}, 1_{bb}, 1_{tb}]$	May be interpreted as a compact state coupled to the higher channel
SF3		
10	$[3_{bt}, 0_{bb}, 0_{tb}]$	...
11	$[0_{bt}, 3_{bb}, 0_{tb}]$	...
12	$[0_{bt}, 0_{bb}, 3_{tb}]$	...
13	$[2_{bt}, 1_{bb}, 0_{tb}]$	Hadronic molecule + compact state
14	$[2_{bt}, 0_{bb}, 1_{tb}]$	...
SF4		
15	$[1_{bt}, 2_{bb}, 0_{tb}]$	...
16	$[0_{bt}, 2_{bb}, 1_{tb}]$	...
17	$[1_{bt}, 0_{bb}, 2_{tb}]$	...
18	$[0_{bt}, 1_{bb}, 2_{tb}]$	Compact state coupled to the higher channel + virtual state
19	$[1_{bt}, 1_{bb}, 1_{tb}]$	Compact state coupled to the lower channel + virtual state [49]
SF5		
20	$[4_{bt}, 0_{bb}, 0_{tb}]$	...
21	$[0_{bt}, 4_{bb}, 0_{tb}]$	...
22	$[0_{bt}, 0_{bb}, 4_{tb}]$	...
23	$[3_{bt}, 1_{bb}, 0_{tb}]$	...
24	$[3_{bt}, 0_{bb}, 1_{tb}]$	...
SF6		
25	$[1_{bt}, 3_{bb}, 0_{tb}]$	...
26	$[0_{bt}, 3_{bb}, 1_{tb}]$	...
27	$[1_{bt}, 0_{bb}, 3_{tb}]$	...
28	$[0_{bt}, 1_{bb}, 3_{tb}]$	...
29	$[2_{bt}, 2_{bb}, 0_{tb}]$	Possible two compact states
SF7		
30	$[2_{bt}, 0_{bb}, 2_{tb}]$	...
31	$[0_{bt}, 2_{bb}, 2_{tb}]$	...
32	$[2_{bt}, 1_{bb}, 1_{tb}]$	...
33	$[1_{bt}, 2_{bb}, 1_{tb}]$	Compact state + virtual state pole with nonvanishing width [52]
34	$[1_{bt}, 1_{bb}, 2_{tb}]$	...

coupling and background conditions [74,75]. The complete set spans pole-free amplitudes to configurations with multiple nearby poles, ensuring that no analytically allowed topology is excluded *a priori* in the subsequent DNN analysis.

APPENDIX B: POLE TOPOLOGY FOR COMPACT AND MOLECULAR STATES

The association of the pole-shadow pair to a compact state interpretation can be understood from the analytic continuation of a bare state coupled to two scattering channels. In a single-channel problem, a compact or bare state coupled to the continuum appears as a resonance pole generated by the zero of its corresponding Jost function. In a two-channel system, however, the same state can couple to the second channel. The analytic continuation across the channel cuts then produces more than one nearby solution of the pole equation. Schematically, the two-channel Jost condition may be written as

$$[(q_1 - i\beta_1)^2 - \alpha_1^2] + \lambda[(q_2 - i\beta_2)^2 - \alpha_2^2] = 0, \quad (\text{B1})$$

where λ parametrizes the coupling to the second channel, while α and β denote the real and imaginary parts of the corresponding channel momentum. This condition can be manipulated to obtain the form,

$$\begin{aligned} & [(q_1 - i\beta_1)^2 - \alpha_1^2] \left\{ \left[q_1 - i\beta_1 \left(\frac{1-\lambda}{1+\lambda} \right) \right]^2 \right. \\ & \left. - \left[\alpha_1^2 + \frac{4\lambda}{(1+\lambda)^2} \beta_2^2 \right] \right\} = 0, \end{aligned} \quad (\text{B2})$$

or equivalently,

$$\begin{aligned} & [(q_2 - i\beta_2)^2 - \alpha_2^2] \left\{ \left[q_2 + i\beta_2 \left(\frac{1-\lambda}{1+\lambda} \right) \right]^2 \right. \\ & \left. - \left[\alpha_2^2 + \frac{4\lambda}{(1+\lambda)^2} \beta_1^2 \right] \right\} = 0. \end{aligned} \quad (\text{B3})$$

The first factor corresponds to the main pole, while the second factor inside the curly brackets gives the shadow-pole solution generated by the coupling to the other channel. The sheet assignment is controlled by the signs of the imaginary parts of the channel momenta. Thus, changing the sign of the imaginary part of q_2 moves the associated shadow pole between adjacent sheets, such as $[bt]$ and $[bb]$. Therefore, a compact state coupled to two channels may appear not only as a pole on the $[bt]$ sheet, but also as a nearby shadow pole on the $[bb]$ sheet. This is why a nearby $[bt] - [bb]$ pole pair is interpreted as the compact state

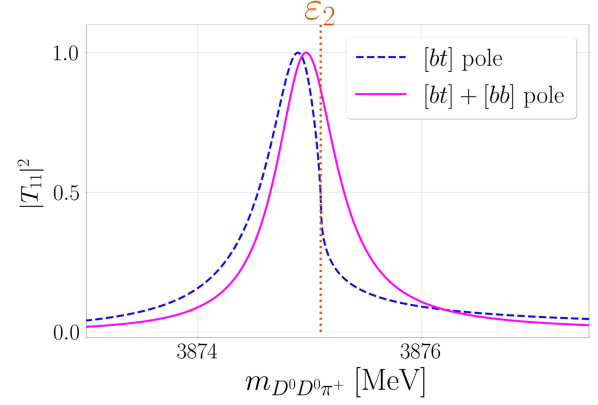


FIG. 8. Schematic comparison of near-threshold line shapes from different pole configurations. The isolated $[bt]$ -sheet pole gives a threshold-dominated enhancement near the $D^0 D^{*+}$ threshold (ε_2) marked by the vertical dotted line, while the compact state $[bt] - [bb]$ pole-shadow pair produces a comparatively more resonancelike structure.

topology. This interpretation should be understood as a pole-counting criterion within the $\mathcal{S}$ -matrix framework [75].

In principle, the isolated $[bt]$ -sheet pole and the $[bt] - [bb]$ pole-shadow pair topology can generate different near-threshold line shapes (Fig. 8).

The isolated $[bt]$ -sheet pole, associated here with the hadronic molecule interpretation, produces a threshold-dominated enhancement, characterized by a sharper and steeper behavior near the second threshold. This reflects the strong influence of the nearby threshold on the amplitude. By contrast, the compact state of the $[bt] - [bb]$ pole-shadow pair topology can produce a more resonancelike line shape, which is comparatively more symmetric around the peak. However, this distinction is not always easy to identify by eye in the experimentally measured invariant-mass distribution, as the observed spectrum is not simply the scattering amplitude, but also includes background contributions and is affected by detector resolution and statistical uncertainties. These effects can partially obscure the visual differences between compactlike and molecular-like line shapes.

APPENDIX C: SUMMARY OF DNN-BASED POLE TOPOLOGY INFERENCE

The original 35 pole-topology configurations were grouped into seven structural families according to pole multiplicity and Riemann-sheet distribution. Ensemble of DNN classifiers were trained for each family, and inference was performed on pseudoempirical amplitudes generated from the LHCb line shape.

For amplitudes without lhc effects, dominant pole topology emerged in each structural family. The leading candidates were class 01: $[1_{bt}, 0_{bb}, 0_{tb}]$, class 07: $[1_{bt}, 1_{bb}, 0_{tb}]$,

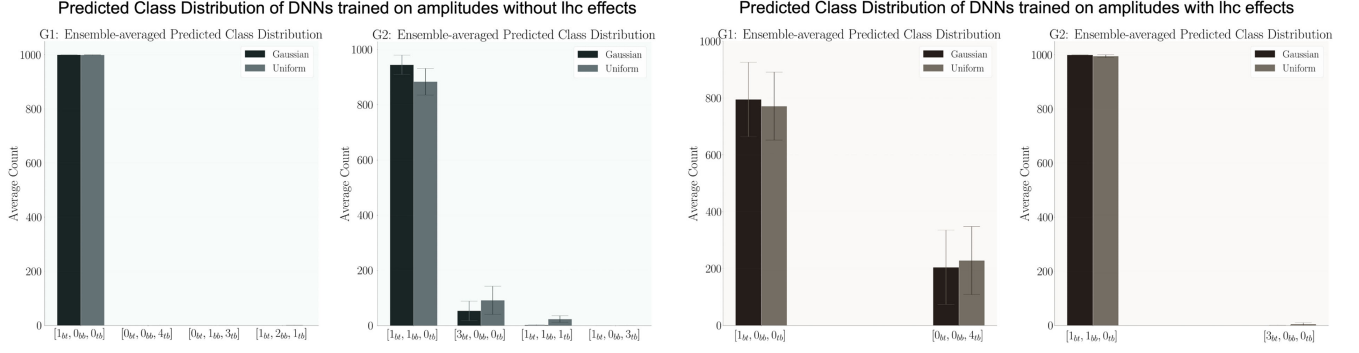


FIG. 10. Ensemble mean with standard error of the predicted class distribution for the LHCb data for two groups of candidate classes.

TABLE III. Extracted best fit parameters for an isolated pole on the $[bt]$ sheet model.

Parameter	Uniformized $\mathcal{S}$ -matrix	$\mathcal{K}$ -matrix
$M_{T_{cc}^+}$ [MeV]	3874.78 ± 0.04	...
$\Gamma_{T_{cc}^+}$ [keV]	97.20 ± 9.40	...
$\mathcal{K}_{11}$	...	-5.84 ± 0.74
$\mathcal{K}_{12}$	...	-15.93 ± 2.66
$\mathcal{K}_{22}$	...	-76.57 ± 7.11
$\mathcal{A}^a$	1.00 ± 0.00	1.00 ± 0.00
α_1^a	1.68 ± 0.17	52.02 ± 0.52
α_2^b	1.00 ± 0.00	1.00 ± 0.00
c_0^b	0.00 ± 0.00	0.00 ± 0.00
c_1	$(2.78 \pm 0.69) \times 10^{-3}$	$(-0.72 \pm 0.50) \times 10^1$
c_2	$(-8.30 \pm 0.33) \times 10^{-5}$	$(-0.18 \pm 1.06) \times 10^{-1}$

^aFixed to its best fit values from prior runs.^bFixed.

APPENDIX E: DNN PERFORMANCE AND ENSEMBLE INFERENCE FOR BINARY CLASSIFICATION

The performance evaluation for the DNN architectures trained on amplitudes without and with lhc effects are

comprehensively detailed in Tables IV and V, respectively. The ensemble classification results consistently favor the isolated $[bt]$ pole structure which underscores the stability of the physical extraction independent of DNN model depth, capacity, or variations in the empirical data sampling methodology.

TABLE IV. Comprehensive performance summary of the 50 DNN models trained on amplitudes without lhc effects. The table includes the model architecture, validation accuracy, and the ensemble-averaged predicted class distribution from inference on LHCb data using both Gaussian and uniform sampling. The final row shows the ensemble mean for the inference counts.

DNN	Architecture	Accuracy	Gaussian sampling		Uniform sampling	
			$[1_{bt}, 0_{bb}, 0_{tb}]$	$[1_{bt}, 1_{bb}, 0_{tb}]$	$[1_{bt}, 0_{bb}, 0_{tb}]$	$[1_{bt}, 1_{bb}, 0_{tb}]$
1	[200,200]	93%	1000.0	0.0	1000.0	0.0
2	[300,300]	94%	1000.0	0.0	1000.0	0.0
3	[400,400]	94%	1000.0	0.0	1000.0	0.0
4	[500,500]	95%	1000.0	0.0	997.0	3.0
5	[600,600]	94%	1000.0	0.0	999.0	1.0
6	[700,700]	94%	1000.0	0.0	1000.0	0.0
7	[800,800]	94%	996.0	4.0	927.0	73.0
8	[900,900]	95%	1000.0	0.0	1000.0	0.0
9	[200,200,200]	95%	1000.0	0.0	1000.0	0.0
10	[200,300,200]	95%	1000.0	0.0	1000.0	0.0
11	[300,300,300]	95%	1000.0	0.0	999.0	1.0
12	[300,400,300]	95%	1000.0	0.0	1000.0	0.0
13	[400,400,400]	95%	1000.0	0.0	1000.0	0.0
14	[400,500,400]	95%	1000.0	0.0	999.0	1.0
15	[500,500,500]	95%	982.0	18.0	872.0	128.0
16	[500,600,500]	94%	1000.0	0.0	1000.0	0.0
17	[600,600,600]	94%	1000.0	0.0	1000.0	0.0
18	[600,700,600]	95%	1000.0	0.0	1000.0	0.0
19	[700,700,700]	95%	1000.0	0.0	985.0	15.0
20	[700,800,700]	94%	1000.0	0.0	997.0	3.0
21	[800,800,800]	95%	1000.0	0.0	983.0	17.0
22	[200,200,200,200]	94%	1000.0	0.0	1000.0	0.0
23	[200,300,300,200]	94%	1000.0	0.0	998.0	2.0
24	[300,300,300,300]	94%	1000.0	0.0	999.0	1.0
25	[300,400,400,300]	95%	1000.0	0.0	1000.0	0.0
26	[400,400,400,400]	95%	1000.0	0.0	1000.0	0.0
27	[400,500,500,400]	95%	1000.0	0.0	1000.0	0.0
28	[500,500,500,500]	95%	1000.0	0.0	1000.0	0.0
29	[500,600,600,500]	94%	1000.0	0.0	998.0	2.0
30	[600,600,600,600]	94%	1000.0	0.0	1000.0	0.0
31	[600,700,700,600]	94%	1000.0	0.0	1000.0	0.0
32	[700,700,700,700]	95%	993.0	7.0	948.0	52.0
33	[700,800,800,700]	94%	1000.0	0.0	1000.0	0.0
34	[800,800,800,800]	94%	1000.0	0.0	1000.0	0.0
35	[200,200,200,200,200]	95%	1000.0	0.0	1000.0	0.0
36	[200,300,300,300,200]	95%	937.0	63.0	809.0	191.0
37	[200,300,400,300,200]	95%	1000.0	0.0	1000.0	0.0
38	[300,300,300,300,300]	95%	1000.0	0.0	999.0	1.0
39	[300,400,400,400,300]	95%	1000.0	0.0	1000.0	0.0
40	[300,400,500,400,300]	95%	1000.0	0.0	998.0	2.0
41	[400,400,400,400,400]	95%	1000.0	0.0	974.0	26.0
42	[400,500,500,500,400]	95%	990.0	10.0	927.0	73.0
43	[400,500,600,500,400]	94%	1000.0	0.0	999.0	1.0
44	[500,500,500,500,500]	95%	1000.0	0.0	1000.0	0.0
45	[500,600,600,600,500]	95%	1000.0	0.0	997.0	3.0
46	[500,600,700,600,500]	95%	1000.0	0.0	1000.0	0.0
47	[600,600,600,600,600]	95%	1000.0	0.0	1000.0	0.0
48	[600,700,700,700,600]	95%	1000.0	0.0	1000.0	0.0
49	[600,700,800,700,600]	95%	1000.0	0.0	995.0	5.0
50	[700,700,700,700,700]	94%	1000.0	0.0	1000.0	0.0
			997.96 ± 1.32	2.04 ± 1.32	987.98 ± 4.96	12.02 ± 4.96

TABLE V. Comprehensive performance summary of the 50 DNN models trained on amplitudes with lhc effects from two-pion exchange. The table includes the model architecture, validation accuracy, and the ensemble-averaged predicted class distribution from inference on LHCb data using both Gaussian and uniform sampling. The final row shows the ensemble mean for the inference counts.

DNN	Architecture	Accuracy	Gaussian sampling		Uniform sampling	
			$[1_{bt}, 0_{bb}, 0_{tb}]$	$[1_{bt}, 1_{bb}, 0_{tb}]$	$[1_{bt}, 0_{bb}, 0_{tb}]$	$[1_{bt}, 1_{bb}, 0_{tb}]$
1	[200,200]	84%	1000.0	0.0	1000.0	0.0
2	[300,300]	85%	1000.0	0.0	1000.0	0.0
3	[400,400]	83%	1000.0	0.0	1000.0	0.0
4	[500,500]	85%	1000.0	0.0	1000.0	0.0
5	[600,600]	85%	1000.0	0.0	1000.0	0.0
6	[700,700]	83%	1000.0	0.0	1000.0	0.0
7	[800,800]	83%	1000.0	0.0	1000.0	0.0
8	[900,900]	81%	1000.0	0.0	1000.0	0.0
9	[200,200,200]	87%	1000.0	0.0	1000.0	0.0
10	[200,300,200]	86%	1000.0	0.0	1000.0	0.0
11	[300,300,300]	86%	1000.0	0.0	1000.0	0.0
12	[300,400,300]	86%	1000.0	0.0	1000.0	0.0
13	[400,400,400]	85%	1000.0	0.0	994.0	0.0
14	[400,500,400]	86%	1000.0	0.0	998.0	2.0
15	[500,500,500]	87%	1000.0	0.0	999.0	1.0
16	[500,600,500]	84%	549.0	451.0	539.0	461.0
17	[600,600,600]	84%	1000.0	0.0	1000.0	0.0
18	[600,700,600]	85%	1000.0	0.0	1000.0	0.0
19	[700,700,700]	85%	1000.0	0.0	1000.0	0.0
20	[700,800,700]	86%	1000.0	0.0	1000.0	0.0
21	[800,800,800]	86%	1000.0	0.0	991.0	9.0
22	[200,200,200,200]	86%	1000.0	0.0	1000.0	0.0
23	[200,300,300,200]	86%	1000.0	0.0	987.0	13.0
24	[300,300,300,300]	85%	990.0	10.0	922.0	78.0
25	[300,400,400,300]	85%	1000.0	0.0	1000.0	0.0
26	[400,400,400,400]	86%	1000.0	0.0	1000.0	0.0
27	[400,500,500,400]	85%	1000.0	0.0	995.0	5.0
28	[500,500,500,500]	86%	1000.0	0.0	1000.0	0.0
29	[500,600,600,500]	86%	1000.0	0.0	1000.0	0.0
30	[600,600,600,600]	84%	1000.0	0.0	1000.0	0.0
31	[600,700,700,600]	85%	1000.0	0.0	1000.0	0.0
32	[700,700,700,700]	83%	1000.0	0.0	1000.0	0.0
33	[700,800,800,700]	84%	1000.0	0.0	1000.0	0.0
34	[800,800,800,800]	84%	1000.0	0.0	1000.0	0.0
35	[200,200,200,200,200]	86%	1000.0	0.0	1000.0	0.0
36	[200,300,300,300,200]	86%	1000.0	0.0	1000.0	0.0
37	[200,300,400,300,200]	85%	1000.0	0.0	1000.0	0.0
38	[300,300,300,300,300]	83%	1000.0	0.0	1000.0	0.0
39	[300,400,400,400,300]	85%	1000.0	0.0	1000.0	0.0
40	[300,400,500,400,300]	84%	1000.0	0.0	1000.0	0.0
41	[400,400,400,400,400]	85%	1000.0	0.0	1000.0	0.0
42	[400,500,500,500,400]	85%	1000.0	0.0	1000.0	0.0
43	[400,500,600,500,400]	84%	1000.0	0.0	1000.0	0.0
44	[500,500,500,500,500]	82%	1000.0	0.0	1000.0	0.0
45	[500,600,600,600,500]	86%	923.0	77.0	824.0	176.0
46	[500,600,700,600,500]	85%	1000.0	0.0	1000.0	0.0
47	[600,600,600,600,600]	83%	1000.0	0.0	1000.0	0.0
48	[600,700,700,700,600]	84%	1000.0	0.0	1000.0	0.0
49	[600,700,800,700,600]	84%	1000.0	0.0	1000.0	0.0
50	[700,700,700,700,700]	81%	1000.0	0.0	1000.0	0.0
			989.24 ± 9.12	10.76 ± 9.12	984.98 ± 9.87	15.02 ± 9.87

- [1] R. Aaij *et al.* (LHCb Collaboration), *Nat. Phys.* **18**, 751 (2022).
- [2] R. Aaij *et al.* (LHCb Collaboration), *Nat. Commun.* **13**, 3351 (2022).
- [3] J. P. Ader, J. M. Richard, and P. Taxil, *Phys. Rev. D* **25**, 2370 (1982).
- [4] L. Heller and J. A. Tjon, *Phys. Rev. D* **35**, 969 (1987).
- [5] J. Carlson, L. Heller, and J. A. Tjon, *Phys. Rev. D* **37**, 744 (1988).
- [6] B. Silvestre-Brac and C. Semay, *Z. Phys. C* **57**, 273 (1993).
- [7] C. Semay and B. Silvestre-Brac, *Z. Phys. C* **61**, 271 (1994).
- [8] S. Pepin, F. Stancu, M. Genovese, and J. M. Richard, *Phys. Lett. B* **393**, 119 (1997).
- [9] J. I. Ballot and J. M. Richard, *Phys. Lett.* **123B**, 449 (1983).
- [10] S. Zouzou, B. Silvestre-Brac, C. Gignoux, and J. M. Richard, *Z. Phys. C* **30**, 457 (1986).
- [11] T. Guo, J. Li, J. Zhao, and L. He, *Phys. Rev. D* **105**, 014021 (2022).
- [12] X.-Z. Weng, W.-Z. Deng, and S.-L. Zhu, *Chin. Phys. C* **46**, 013102 (2022).
- [13] Y. Kim, M. Oka, and K. Suzuki, *Phys. Rev. D* **105**, 074021 (2022).
- [14] W.-C. Dong and Z.-G. Wang, *Nucl. Phys. B* **1012**, 116828 (2025).
- [15] X.-Z. Ling, M.-Z. Liu, L.-S. Geng, E. Wang, and J.-J. Xie, *Phys. Lett. B* **826**, 136897 (2022).
- [16] X.-K. Dong, F.-K. Guo, and B.-S. Zou, *Commun. Theor. Phys.* **73**, 125201 (2021).
- [17] A. Feijoo, W. H. Liang, and E. Oset, *Phys. Rev. D* **104**, 114015 (2021).
- [18] S. Fleming, R. Hodges, and T. Mehen, *Phys. Rev. D* **104**, 116010 (2021).
- [19] H. Ren, F. Wu, and R. Zhu, *Adv. High Energy Phys.* **2022**, 9103031 (2022).
- [20] K. Chen, R. Chen, L. Meng, B. Wang, and S.-L. Zhu, *Eur. Phys. J. C* **82**, 581 (2022).
- [21] R. Chen, Q. Huang, X. Liu, and S.-L. Zhu, *Phys. Rev. D* **104**, 114042 (2021).
- [22] M. Albaladejo, *Phys. Lett. B* **829**, 137052 (2022).
- [23] M.-L. Du, V. Baru, X.-K. Dong, A. Filin, F.-K. Guo, C. Hanhart, A. Nefediev, J. Nieves, and Q. Wang, *Phys. Rev. D* **105**, 014024 (2022).
- [24] V. Baru, X.-K. Dong, M.-L. Du, A. Filin, F.-K. Guo, C. Hanhart, A. Nefediev, J. Nieves, and Q. Wang, *Phys. Lett. B* **833**, 137290 (2022).
- [25] N. Santowsky and C. S. Fischer, *Eur. Phys. J. C* **82**, 313 (2022).
- [26] C. Deng and S.-L. Zhu, *Phys. Rev. D* **105**, 054015 (2022).
- [27] S. Agaev, K. Azizi, and H. Sundu, *J. High Energy Phys.* **06** (2022) 057.
- [28] Y. Kamiya, T. Hyodo, and A. Ohnishi, *Eur. Phys. J. A* **58**, 131 (2022).
- [29] L. M. Abreu, *Nucl. Phys. B* **985**, 115994 (2022).
- [30] S. Chen, C. Shi, Y. Chen, M. Gong, Z. Liu, W. Sun, and R. Zhang, *Phys. Lett. B* **833**, 137391 (2022).
- [31] M. Albaladejo and J. Nieves, *Eur. Phys. J. C* **82**, 724 (2022).
- [32] F.-Z. Peng, M. Sánchez Sánchez, M.-J. Yan, and M. Pavon Valderrama, *Phys. Rev. D* **105**, 034028 (2022).
- [33] P. G. Ortega, J. Segovia, D. R. Entem, and F. Fernandez, *Phys. Lett. B* **841**, 137918 (2023); **847**, 138308(E) (2023).
- [34] S.-Y. Li, Y.-R. Liu, Z.-L. Man, Z.-G. Si, and J. Wu, *Phys. Rev. D* **110**, 094044 (2024).
- [35] L. R. Dai, J. Song, and E. Oset, *Phys. Lett. B* **846**, 138200 (2023).
- [36] M. Abolnikov, V. Baru, E. Epelbaum, A. A. Filin, C. Hanhart, and L. Meng, *Phys. Lett. B* **860**, 139188 (2025).
- [37] L.-Y. Dai, X. Sun, X.-W. Kang, A. P. Szczepaniak, and J.-S. Yu, *Phys. Rev. D* **105**, L051507 (2022).
- [38] T. Whyte, D. J. Wilson, and C. E. Thomas (Hadron Spectrum Collaboration), *Phys. Rev. D* **111**, 034511 (2025).
- [39] N. N. Achasov and G. N. Shestakov, *Phys. Rev. D* **105**, 096038 (2022).
- [40] H.-X. Chen, W. Chen, X. Liu, Y.-R. Liu, and S.-L. Zhu, *Rep. Prog. Phys.* **86**, 026201 (2023).
- [41] A. Francis, *Prog. Part. Nucl. Phys.* **140**, 104143 (2025).
- [42] M.-Z. Liu, Y.-W. Pan, Z.-W. Liu, T.-W. Wu, J.-X. Lu, and L.-S. Geng, *Phys. Rep.* **1108**, 1 (2025).
- [43] M. Kato, *Ann. Phys. (N.Y.)* **31**, 130 (1965).
- [44] L. M. Santos and D. L. B. Sombillo, *Phys. Rev. C* **108**, 045204 (2023).
- [45] D. L. B. Sombillo, Y. Ikeda, T. Sato, and A. Hosaka, *Phys. Rev. D* **102**, 016024 (2020).
- [46] D. L. B. Sombillo, Y. Ikeda, T. Sato, and A. Hosaka, *Few Body Syst.* **62**, 52 (2021).
- [47] D. L. B. Sombillo, Y. Ikeda, T. Sato, and A. Hosaka, *Phys. Rev. D* **104**, 036001 (2021).
- [48] Z. Zhang, J. Liu, J. Hu, Q. Wang, and U.-G. Meißner, *Sci. Bull.* **68**, 981 (2023).
- [49] L. M. Santos, V. A. A. Chavez, and D. L. B. Sombillo, *J. Phys. G* **52**, 015104 (2025).
- [50] D. A. O. Co, V. A. A. Chavez, and D. L. B. Sombillo, *Phys. Rev. D* **110**, 114034 (2024).
- [51] D. A. O. Co and D. L. B. Sombillo, *Proc. Sci. QNP2024* (2025) 031 [arXiv:2411.14044].
- [52] F. Frohnert, D. L. B. Sombillo, E. van Nieuwenburg, and P. Emonts, *Phys. Rev. D* **113**, 056019 (2026).
- [53] L. Ng, Ł. Bibrzycki, J. Nys, C. Fernández-Ramírez, A. Pilloni, V. Mathieu, A. J. Rasmusson, and A. P. Szczepaniak, *Nuovo Cimento C* **47**, 207 (2024).
- [54] D. Sadasivan, I. Cordero, A. Graham, C. Marsh, D. Kupcho, M. Mourad, and M. Mai, *Phys. Rev. D* **114**, 016002 (2026).
- [55] A. B. Mapa and D. L. B. Sombillo, *Proc. Sci. HA-DRON2025* (2026) 214 [arXiv:2506.04628].
- [56] A. Dersy, M. D. Schwartz, and A. Zhiboedov, *J. High Energy Phys.* **05** (2024) 200.
- [57] L. Ng, Ł. Bibrzycki, J. Nys, C. Fernandez-Ramirez, A. Pilloni, V. Mathieu, A. J. Rasmusson, and A. P. Szczepaniak (Joint Physics Analysis Center, JPAC Collaborations), *Phys. Rev. D* **105**, L091501 (2022).
- [58] D. Binosi, A. Pilloni, and R.-A. Tripolt, *Phys. Lett. B* **839**, 137809 (2023).
- [59] V. A. A. Chavez and D. L. B. Sombillo, *Proc. Sci. HA-DRON2025* (2026) 253 [arXiv:2506.04622].

- [60] J. Landay, M. Mai, M. Döring, H. Haberzettl, and K. Nakayama, *Phys. Rev. D* **99**, 016001 (2019).
- [61] R. Molina, J. Landay, B. Hu, M. Doring, and C. Fernández-Ramírez, *Proc. Sci. Hadron2017* (**2018**) 135.
- [62] T. Alghamdi *et al.*, *Phys. Rev. D* **108**, 094030 (2023).
- [63] P. Bydžovský, A. Cieplý, D. Petrellis, D. Skoupil, and N. Zachariou, *Phys. Rev. C* **104**, 065202 (2021); **110**, 029901 (E) (2024).
- [64] D. Petrellis and D. Skoupil, *Phys. Rev. C* **107**, 045206 (2023).
- [65] M. Malekshosseini, S. Rostami, A. R. Olamaei, R. Ostovar, and K. Azizi, *Phys. Rev. D* **110**, 054011 (2024).
- [66] M. Albaladejo *et al.* (JPAC Collaboration), *Prog. Part. Nucl. Phys.* **127**, 103981 (2022).
- [67] R. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix*, 2nd ed. (Cambridge University Press, London, 1966).
- [68] J. Taylor, *Scattering Theory: Quantum Theory on Non-relativistic Collisions* (Wiley, New York, 1972).
- [69] R. G. Newton, *Scattering Theory of Waves and Particles*, Theoretical and Mathematical Physics (Springer-Verlag, Berlin Heidelberg, 1982).
- [70] W. R. Frazer and A. W. Hendry, *Phys. Rev.* **134**, B1307 (1964).
- [71] J. A. Oller, *A Brief Introduction to Dispersion Relations: With Modern Applications*, 1st ed., SpringerBriefs in Physics (Springer International Publishing, Cham, Switzerland, 2019), pp. VIII, 141.
- [72] S. Navas *et al.* (Particle Data Group Collaboration), *Phys. Rev. D* **110**, 030001 (2024).
- [73] S. Ampuku, Y. Yamaguchi, and M. Harada, *Phys. Rev. D* **113**, L031505 (2026).
- [74] B. C. Pearce and B. F. Gibson, *Phys. Rev. C* **40**, 902 (1989).
- [75] D. Morgan, *Nucl. Phys.* **A543**, 632 (1992).
- [76] M.-L. Du, A. Filin, V. Baru, X.-K. Dong, E. Epelbaum, F.-K. Guo, C. Hanhart, A. Nefediev, J. Nieves, and Q. Wang, *Phys. Rev. Lett.* **131**, 131903 (2023).
- [77] G. R. Sreaton, *Phys. Rev.* **182**, 1415 (1969).
- [78] G. Minerbo, *Phys. Rev. D* **3**, 928 (1971).
- [79] LHCb Collaboration, Observation of an exotic narrow doubly charmed tetraquark, HEPData (collection) (2021), <https://doi.org/10.17182/hepdata.114869>.
- [80] I. Neutelings, TikZ.net: Neural networks (2021), <https://tikz.net/neural-networks/> (accessed: 2025-07-10).
- [81] J. Bridle, in *Advances in Neural Information Processing Systems*, edited by D. Touretzky (Morgan-Kaufmann, San Mateo, CA, 1989), Vol. 2.
- [82] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).
- [83] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
- [84] K. P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT Press, Cambridge, MA, 2012).
- [85] V. Nair and G. E. Hinton, in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10 (OmniPress, Madison, WI, USA, 2010), pp. 807–814.
- [86] X. Glorot, A. Bordes, and Y. Bengio, in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research Vol. 15 (PMLR, 2011), pp. 315–323.
- [87] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2019), arXiv:1907.10902.
- [88] A. Paszke *et al.*, arXiv:1912.01703.
- [89] D. P. Kingma and J. Ba, in *International Conference on Learning Representations (ICLR)* (2015), arXiv:1412.6980.
- [90] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, *J. Mach. Learn. Res.* **15**, 1929 (2014), <https://jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf>.
- [91] A. Krogh and J. Hertz, in *Advances in Neural Information Processing Systems* (Morgan-Kaufmann, San Mateo, CA, 1991), Vol. 4.
- [92] J. Snoek, H. Larochelle, and R. P. Adams, arXiv:1206.2944.
- [93] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., Red Hook, NY, 2011), Vol. 24.
- [94] D. L. B. Sombillo, TccML, Version 1.0, GitHub repository (2025), https://github.com/sombillo/Tcc_ML.git.
- [95] T. Elsken, J. H. Metzen, and F. Hutter, *J. Mach. Learn. Res.* **20**, 1 (2019), <https://www.jmlr.org/papers/volume20/18-598/18-598.pdf>.
- [96] W. A. Yamada and O. Morimatsu, *Phys. Rev. C* **103**, 045201 (2021).
- [97] C. Fernández-Ramírez, A. Pilloni, M. Albaladejo, A. Jackura, V. Mathieu, M. Mikhasenko, J. A. Silva-Castro, and A. P. Szczepaniak (JPAC Collaboration), *Phys. Rev. Lett.* **123**, 092001 (2019).
- [98] I. Matuschek, V. Baru, F.-K. Guo, and C. Hanhart, *Eur. Phys. J. A* **57**, 101 (2021).
- [99] C. Hanhart, J. R. Pelaez, and G. Rios, *Phys. Lett. B* **739**, 375 (2014).
- [100] V. Baru, J. Haidenbauer, C. Hanhart, Y. Kalashnikova, and A. E. Kudryavtsev, *Phys. Lett. B* **586**, 53 (2004).
- [101] A. M. Badalyan, L. P. Kok, M. I. Polikarpov, and Y. A. Simonov, *Phys. Rep.* **82**, 31 (1982).
- [102] F. James and M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975).
- [103] V. V. Gligorov, S. Hageboeck, T. Nanut, A. Sciandra, and D. Y. Tou, *J. Instrum.* **16**, T08004 (2021).
- [104] S. U. Chung, J. Brose, R. Hackmann, E. Klempt, S. Spanier, and C. Strassburger, *Ann. Phys. (Berlin)* **4**, 404 (1995).