Eur. Phys. J. C (2025) 85:1360
https://doi.org/10.1140/epjc/s10052-025-14713-w

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Computing, Software and Data Science

Development of systematic uncertainty-aware neural network trainings for binned-likelihood analyses at the LHC

CMS Collaboration*

CERN, Geneva, Switzerland

Received: 28 August 2025 / Accepted: 29 August 2025

Abstract We propose a neural network training method capable of accounting for the effects of systematic variations of the data model in the training process and describe its extension towards neural network multiclass classification. The procedure is evaluated on the realistic case of the measurement of Higgs boson production via gluon fusion and vector boson fusion in the $\tau\tau$ decay channel at the CMS experiment. The neural network output functions are used to infer the signal strengths for inclusive production of Higgs bosons as well as for their production via gluon fusion and vector boson fusion. We observe improvements of 12 and 16% in the uncertainty in the signal strengths for gluon and vector-boson fusion, respectively, compared with a conventional neural network training based on cross-entropy.

1 Introduction

A physics measurement implies the synthesis of a potentially huge data set into a single or small number of characterizing model parameters of interest (POIs) and their uncertainties. The sensitivity of a measurement is determined by its power in extracting all information relevant for the determination of the POIs from data, such that the most sensitive measurement is the one with the smallest uncertainties in the POIs.

Higgs boson (H) production at the CERN LHC [1–5] is a rare process, which, depending on the H decay, may be covered by an overwhelmingly large background from standard model (SM) processes not related to H physics. This is a particular challenge for analyses where the H decays into fermions, such as b quarks [6,7] or tau leptons ($\mathrm{H} \to \tau\tau$) [8,9]. At the same time, the complex and feature-rich nature of these decay channels makes them ideal candidates for the application of modern machine learning (ML) methods, typically exploited to separate signal from background processes. The latest measurements of differential simplified template cross sections (STXS) from $\mathrm{H} \to \tau\tau$ events of the CMS Collaboration [8] form the most precise measurements of this decay channel to date. These measurements rely heavily on the use of ML methods in the form of neural network (NN) multiclass classification using categorical cross-entropy (CE) [10] as the loss function for the NN training. In this paper we refer to this kind of NN training as (categorical) CE-based NN training (CENNT).

If the NN output function can be interpreted as a probability p_l to associate a given event to class l, the categorical CE is equivalent to the negative log-likelihood of the multinomial distribution

$$\mathbb{P}(\{k_l\},\, n,\, \{p_l\}) = n! \prod_{l=1}^{\Lambda} \frac{p_l^{k_l}}{k_l!}, \tag{1}$$

with

$$\sum_{l=1}^{\Lambda} k_l = n \quad \text{and} \quad \sum_{l=1}^{\Lambda} p_l = 1,$$

where n refers to the inclusive number of events, k_l to the number of events in class l, and Λ to the number of all classes. Hence, CENNT is a maximum likelihood estimate of a multiclass-classification model based on $\mathbb{P}(\{k_l\},\, n,\, \{p_l\})$. If successful, it provides the best possible separation of all signal and background classes based on the uncertainties related to the statistical nature of the measurement, also referred to as aleatoric or statistical uncertainties. However, CENNT is blind to any influence of imperfect knowledge of the underlying model Ω_X used to extract a POI related to a given measurement from data, where X refers to the set of observables on which the measurement is based. We usually quantify our lack of knowledge in terms of variations of external parameters of Ω_X resulting in what we call epistemic or systematic uncertainties.

The measurements of Ref. [8] have entered an era of precision where the uncertainties, even of differential measurements, are dominated by their systematic (rather than statistical) component. This statement holds even if measure-

* e-mail: cms-publication-committee-chair@cern.ch (corresponding author)

Published online: 26 November 2025

ments will at the same time become increasingly differential, since the more differential a measurement becomes the more susceptible it is to imperfect knowledge of external model parameters. In this situation, the advantage of using ML methods to maximize the sensitivity of a given measurement might be reduced or even lost. It could be preserved, however, if an NN training can be designed to take into account not only statistical, but also systematic uncertainties. In this paper, we will refer to this kind of NN training as systematic uncertainty aware NN training (SANNT). By construction, SANNT incorporates more information of Ω_X than CENNT. In this sense it is likely to require a larger investment of the analyst in characterizing Ω_X. In turn, the presented SANNT is applicable to any kind of binned likelihood analysis and should lead to at least the same or a higher sensitivity to a given POI than CENNT, when based on the same data model. Providing an objective strategy for obtaining the most sensitive result for a given Ω_X that represents a given understanding of the data, SANNT thus offers a sizable impact on many analyses in the transition from statistically to systematically dominated regions, at the LHC and beyond. A number of strategies and corresponding loss functions for SANNT have been proposed in the recent literature [11–15]. These go significantly beyond approaches that aim at only reducing the sensitivity of the NN output function to variations of input features in X, e.g. as introduced and discussed in Refs. [16,17], which are usually applied to increase the robustness of classification algorithms against differences of input features between the training and test data. This work represents the first application to an existing LHC analysis of a SANNT that includes a complete and realistic model of the systematic uncertainties. It marks an important step towards demonstrating the applicability of a SANNT in high-energy particle physics measurements.

In this paper we demonstrate the application of SANNT to the comprehensive realistic data model Ω_X presented in Ref. [8]. Our work is based on the method of Ref. [14], replacing CE by the extended binned likelihood commonly used for the extraction of a given signal in addition to a set of known background processes, based on histogram templates. From this likelihood, which can be conceptually extended to include an arbitrary number of nuisance parameters related to systematic variations in Ω_X, we derive the uncertainty in a given POI through the Fisher information [18] and the Cramér–Rao bound [19,20]. In this way, we preserve maximal congruency of the training target with the measurement target also in the presence of systematic uncertainties. Since it intrinsically relies on a predefined binning of the output node values through the binned likelihood, the SANNT described here can be viewed as a method to distribute the events optimally across these bins to obtain the minimal uncertainty in the POI, during training time. A conventional bin optimization procedure in one dimension applied to a CENNT receives fixed values of the NN output function depending on the input space vectors $\mathbf{x} \in X$, after training. Based on these output values the procedure leads to an adjustment of the bin boundaries of the histogram to be used for signal extraction. For the SANNT as described here, the binning is fixed, while the adjustment happens to the values of the NN output function, during the training. We note that this strictly holds only for tasks where the uncertainties in the POI are still dominated by the statistical part, and if a high correlation between the NN outputs after CENNT and SANNT is retained, e.g. by the choice of pre-training conditions.

During training, the minimization of the negative log-likelihood is performed with backpropagation. The issue of differentiating discrete histogram boundaries is solved through custom functions in the backward pass. As an evolution of the original method we propose an extension towards NN multiclass classification.

As a likelihood model, Ω_X comprises several hundreds of partially correlated sources of systematic uncertainty for several hundreds of individual underlying rate measurements. For practical reasons, we have restricted ourselves to the $e\tau_h$ final state of the $\tau\tau$ system, with an electron (e) and a hadronic τ lepton decay (τ_h) in the final state and the data taken in 2017. For this subset of the data we use the same model as the one used in Ref. [8]. We compare both an inclusive and a simple differential STXS measurement based on CENNT with SANNT-based measurements taking 86 sources of partially correlated uncertainties into account under the most representative data-taking conditions for the year 2017.

The paper is organized as follows: Sect. 2 describes the experimental setup of the presented studies, including a brief description of the measurement used for benchmarking. Section 3 describes the SANNT, including a discussion of the concept, related work, and adaptations made for the studies presented here. An instructive comparison of the SANNT with the original CENNT on a simplified NN model for binary classification is given in Sect. 4. A conceptual extension of the loss function for multiclass classification is introduced in Sect. 5, before we summarize the results in Sect. 6.

2 Experimental setup

2.1 The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, and a brass and scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity (η) cover-

age provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

Events of interest are selected using a two-tiered trigger system. The first level, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [21]. The second level, known as the high-level trigger, consists of a farm of processors running a version of the full event reconstruction software optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [22,23].

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [24].

2.2 Event reconstruction and selection

The reconstruction of the proton-proton (pp) collision products is based on the particle-flow (PF) algorithm [25], which combines the information from all CMS subdetectors to reconstruct a set of particle candidates, identified as charged hadrons, neutral hadrons, electrons, photons, and muons. The fully recorded detector data of a hard pp collision defines an event for further processing. The primary vertex (PV) of the interaction is taken to be the vertex corresponding to the hardest scattering in an event, evaluated using tracking information alone, as described in Section 9.4.1 of Ref. [26]. Secondary vertices, which are detached from the PV, might be associated with decays of long-lived particles emerging from the PV. Any other collision vertices in an event are associated with additional, mostly soft inelastic pp collisions called pileup.

Electron candidates are reconstructed by combining clusters of energy deposits in the electromagnetic calorimeter with hits in the tracker [27,28]. To increase their purity, reconstructed electrons are required to pass a multivariate electron identification discriminant, which combines information on track quality, shower shape, and kinematic quantities. Muons in the event are reconstructed performing a simultaneous track fit to hits in the tracker and in the muon chambers [29,30]. For further characterization of an event, all reconstructed PF candidates are clustered into jets using the anti-k_T jet clustering algorithm, as implemented in the FastJet software package [31,32], with a distance parameter of 0.4. Jets resulting from the hadronization of b quarks are identified by exploiting dedicated identification algorithms, as described in Refs. [33,34]. Jets are also used as seeds for the identification of τ_h candidates. This is done by exploiting the substructure of the jets, using the "hadrons-plus-strips" algorithm, as described in Refs. [35,36]. Decays into one or three charged hadrons with up to two neutral pions are used. To distinguish τ_h candidates from jets originating from the hadronization of quarks or gluons, as well as from electrons, or muons, the DeepTau algorithm [36] is used, which provides three corresponding discriminants D_α ($\alpha = \mathrm{e}, \mu, \mathrm{jet}$). For D_e the Tight, for D_μ the VLoose, and for D_jet the Tight working points, as defined in Ref. [36], are chosen. Eventually, the missing transverse momentum vector $\mathbf{p}_\mathrm{T}^\mathrm{miss}$ is calculated from the negative vector p_T sum of all PF candidates [37] exploiting the pileup-per-particle identification algorithm, as described in Ref. [38].

In the high-level trigger, events are selected either through the presence of both an electron and a τ_h candidate ($\mathrm{e}\tau_h$-trigger), or through the presence of only a high-p_T electron candidate (single-e trigger). In the offline selection, the e candidate is required to have $|\eta| < 2.1$ and $p_\mathrm{T} > 25\,(28)$ GeV, depending on whether the event was selected through the $\mathrm{e}\tau_h$-trigger (single-e trigger) where the offline p_T threshold is chosen to ensure high efficiency of the trigger for the selected events. Furthermore, the e candidate is required to be isolated from any hadronic activity originating from the PV. The τ_h candidate is required to have $|\eta| < 2.3$ and $p_\mathrm{T} > 35\,(30)$ GeV, depending on whether the event was selected through the $\mathrm{e}\tau_h$-trigger (single-e trigger). The e and τ_h candidates are required to be of opposite charges and separated by more than $\Delta R = 0.5$, where

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2}$$

refers to the difference between two objects in the plane defined by η and the azimuthal angle ϕ, measured in radians. The presence of e or μ candidates in addition to the selected e candidate is vetoed. A more detailed description of the selection criteria can be found in Ref. [8].

2.3 Data model

After the selection described in Sect. 2.2, referred to as the signal region (SR), the main backgrounds originate from Z boson production, also referred to as Drell–Yan (DY) production [39], in association with jets in the ττ decay channel ($\mathrm{Z} \to \tau\tau$); W boson production in association with jets (W+jets); top (t) quark pair production ($\mathrm{t\bar{t}}$) and SM events where light quark- or gluon-induced jets are produced through the strong interaction. This last background category is referred to here as quantum chromodynamics (QCD) multijet production. Minor backgrounds originate from the production of two vector bosons (diboson), single t quark, and Z boson production and decay into the ee and μμ final states (collectively denoted as $\mathrm{Z} \to \ell\ell$). Depending on the experimental signatures that these backgrounds leave in the detector, they are estimated either by the τ-embedding method [40], the ("fake factor") F_F-method [41,42], or sim-

ulation to form the data model Ω_X to be used for signal extraction.

In the τ-embedding method, events with two muons are selected in data. All energy deposits of the muons are removed from the event record and replaced by the decay products and subsequent energy deposits of simulated τ decays with the same kinematic properties as the selected muons. In this way, the method relies only on the simulation of the τ decay and its energy deposits in the detector, while all other parts of an event are obtained from data. The resulting sample is used to estimate the background from all processes that comprise two genuine τ leptons in their final states, such as $Z \to \tau\tau$, corresponding decay channels of $t\bar{t}$, and diboson production.

For the F_F-method, the selection criterion on D_{jet} is inverted, keeping all other selection criteria unchanged, to obtain a sideband close to the SR. A minimal requirement on D_{jet} is imposed to ensure kinematic proximity. Transfer functions F_F are applied to the selection to account for the differences in the D_{jet} selection. These F_F are obtained from independent determination regions in data. The F_F-method is used to estimate mostly the contributions of QCD multijet, W+jets, and $t\bar{t}$ production, where a quark- or gluon-induced jet is misidentified and selected as the τ_h candidate.

More than 80% of all expected background processes are estimated either from the τ-embedding method or the F_F-method and so are, at least partially, derived from data. Remaining backgrounds and signal processes are estimated using simulated events generated by POWHEG 2 [43–48] and MADGRAPH5_aMC@NLO 2.2.2 (2.4.2) [49,50] in configurations described in Ref. [8]. The individual processes are normalized to their theoretical predictions with the highest available accuracy in the strong coupling constant α_S. The event generator PYTHIA 8.230 [51] is used to simulate the parton shower, as well as for the simulation of additional inclusive inelastic pp collisions, according to the expected pileup profile in data. All generated events are passed through a GEANT4-based [52] simulation of the CMS detector and reconstructed using the same version of the CMS event reconstruction software as used for the data.

Control regions from event samples not included in the actual measurements are used to track how well the data can be described by Ω_X in terms of: object reconstruction; triggering, identification, and isolation efficiencies; object energy calibration scale factors; and general shape- and normalization-altering degrees of freedom of each individual background estimation method. Where needed, corrections are derived and applied as described in Ref. [8]. These corrections usually range below 10%.

2.4 Statistical inference of the signal

Events in the SR are provided as input to a fully connected feed-forward NN for multiclass classification. In Ref. [8], an NN with two hidden layers of 200 nodes and a hyperbolic tangent as the activation function for the hidden nodes was used. For this paper, we have used a simpler configuration, with one hidden layer of 100 nodes and rectified linear unit (ReLU) [53,54] activation function, following the architecture that was chosen for the analysis of Ref. [14]. We favour comparability with respect to Ref. [14] over Ref. [8], to allow for a direct comparison of the adaptations of the original method, as described in Sect. 3.2.

The 14-dimensional vector of input features $\mathbf{x}$ to be passed to the NN for each event comprises: the p_T of both τ candidates and their vector sum; the p_T of the two leading jets and their vector sum, their difference in η, and the dijet mass m_{jj}; the number of jets N_{jet}; the number of b jets N_{Btag}; a likelihood-based estimate of the mass of the $\tau\tau$ system $m_{\tau\tau}$ [55]; the mass of the visible decay products of the $\tau\tau$ system m_{vis}; and estimates of the momentum transfer of each exchanged vector boson under the vector boson fusion (VBF) hypothesis, as used in Ref. [56]. The samples used for each corresponding NN training comprise events with fewer than two jets, for which input features like the p_T of the second leading jet or m_{jj} are not defined. In these cases, these observables are filled with predefined values outside their physical value spaces. While outside, the values are chosen still close to the physical value spaces, to minimize their influence on the NN prediction. Before entering the NN, all features are standardized such that their distributions have a mean of 0 and a standard deviation of 1.

The NN architecture has seven output nodes with NN output functions $\hat{y}_l(\cdot)$, referring to the following event classes:

- Events from background processes with two genuine τ decays in the final state, estimated from the τ-embedding method (genuine τ);
- Events from processes where the τ_h candidate originates from a misidentified quark- or gluon-induced jet, estimated from the F_F-method (jet $\to \tau_h$).
- Events from $t\bar{t}$ production, which are not covered by any of the above mentioned background estimation methods.
- Events from $Z \to \ell\ell$ production.
- Events from smaller background processes that are either difficult to isolate or too small to be treated individually for signal extraction. These are combined into one residual background class (misc).
- Events from H production through gluon fusion (ggH), as defined in the STXS scheme of the LHC Higgs Working Group [57,58].

– Events from the the VBF dominated qqH process, as defined in the STXS scheme of the LHC Higgs Working Group.

The output nodes correspond to the classes of the classification task. The activation function for the output nodes is the softmax function. The evaluation of the trained NN and signal extraction are performed on independent samples of Ω_X using a two-fold cross-validation scheme.

For the signal extraction, each event is assigned to the class giving the highest value of $\hat{y}_l(\cdot)$. Histogrammed distributions of $\hat{y}_l(\cdot)$, for each corresponding class, are then used as input to an extended binned likelihood function of the form

$$\mathcal{L}\left(\{k_i\},\ \{r_s\},\ \{\theta_j\}\right) = \prod_i \mathcal{P}\big(k_i|\lambda_i(\{\theta_j\})\big) \prod_j \mathcal{C}_j(\tilde{\theta}_j|\theta_j), \tag{3}$$

with

$$\lambda_i(\{\theta_j\}) = \sum_s r_s\, S_{si}(\{\theta_j\}) + \sum_b B_{bi}(\{\theta_j\}),$$

where i labels all bins of the input distributions for each signal class (with index s) and background class (with index b) and the function $\mathcal{P}(\cdot|\cdot)$ corresponds to the Poisson probability to observe k_i events in bin i for a prediction of $\sum r_s\, S_{si}$ signal and $\sum B_{bi}$ background events in that particular bin. The scaling parameters r_s of the signal contributions S_{si} relative to the SM expectation are the POIs. Systematic uncertainties affecting the predicted yields of S_{si} and B_{bi} in each bin are incorporated in the form of penalty terms for additional nuisance parameters θ_j in the likelihood, entering as a product with predefined probability density functions $\mathcal{C}_j(\tilde{\theta}_j|\theta_j)$, where $\tilde{\theta}_j$ corresponds to the nominal value for θ_j. Eventually, the signal is obtained from a maximum likelihood fit based on Eq. (3) to the data, during which the predefined uncertainties in the $\tilde{\theta}_j$ may be further constrained. This setup is identical to the one used in Ref. [8]. In the scope of this paper, the data $\{k_i\}$ are replaced by the Asimov data set $\{k_i^{\mathrm{A}}\}$ obtained from Ω_X including all S_s and B_b under the SM hypothesis with $r_s = 1$, prior to any fit to the data. As introduced in Ref. [59] the Asimov data set is the most representative outcome of a hypothetical experiment corresponding to the expectation value according to Ω_X in each possible observable.

For the discussion in Sect. 4 we use a simplified setup, where all S_s and all B_b are combined into one single signal (S) and one single background (B) process each, resulting in a binary classification task to separate S from B. In this case, a single NN output $\hat{y}(\cdot)$ ranging from 0 (for B) to 1 (for S) is used and Eq. (3) is adapted accordingly. The final results presented in Sects. 4 and 5 are obtained using the CMS statistical analysis tool, COMBINE [60].

We note that $\hat{y}(\boldsymbol{\omega}, \mathbf{x})$ is a deterministic function of $\mathbf{x}$ and the NN weights $\boldsymbol{\omega}$. After training, the $\boldsymbol{\omega}$ are fixed so that $\mathbf{x}$ is mapped to $\hat{y}(\cdot)$. We assume no additional uncertainty from the use of the explicit NN that has been trained for the given analysis task, such that all questions of uncertainties relate to the choice and quality of Ω_X and not to the NN itself.

2.5 Systematic uncertainties

A detailed description of the original uncertainty model of Ω_X is given in Ref. [8]. The experimental uncertainties with the largest impact on the POI for an inclusive cross section measurement in the STXS setup based on the data set of the full Run 2 and all $\tau\tau$ final states are related to the τ identification efficiency $\epsilon_\tau^{\mathrm{ID}}(40, 500)$ for $40 < p_{\mathrm{T}}^{\tau_{\mathrm{h}}} < 500$ GeV and to the F_{F}-method.

When restricted to the $\mathrm{e}\tau_{\mathrm{h}}$ final state and the data-taking year 2017, the uncertainty model of the STXS cross section measurement comprises 224 nuisance parameters, of which 127 are of experimental and 50 of theoretical origins. The former relate to uncertainties in object reconstruction, triggering, and identification, as well as in specific parts of the background model. The latter relate to process normalizations including uncertainties in the signal model, which because of the fact that they might apply only to parts of the kinematic phase space, may also have shape-altering effects. The remaining 47 nuisance parameters are related to the finite population of histogrammed template distributions available for signal extraction.

In the original analysis, the last group of nuisance parameters has been incorporated for each bin of each corresponding template histogram, individually, following the approach proposed in Ref. [61]. Since they allow for independent variations of individual bins, they are also referred to as bin-by-bin uncertainties. All other uncertainties lead to correlated changes across bins taking the form of either normalization or bin-correlated, shape-altering variations. Depending on how they have been derived, correlations arise across individual signal and background samples, or individual uncertainties. The nuisance parameters most relevant for SANNT belong to the following groups of uncertainties:

– Uncertainties related to the reconstruction, triggering, and identification of the selected e or τ_{h} candidates.
– Uncertainties related to the F_{F}-method.
– Theoretical uncertainties related to the signal processes.
– Individual uncertainties, such as an uncertainty in the reweighting of the simulated sample used for estimating the background from $\mathrm{Z} \to \ell\ell$ production, in sideband regions of the data; an uncertainty in the misidentifica-

tion rate of quark- or gluon-induced jets as electron candidates; and the uncertainty in the integrated luminosity [62–65].

For practical reasons that are further described in Sect. 3.1, a total of 86 out of the original 224 nuisance parameters are considered for the statistical inference and SANNT. These comprise shape-altering variations neglecting bin-by-bin uncertainties and a small number of important normalization changing uncertainties that can lead to shape altering effects, when combining individual processes. All other uncertainties are removed from Ω_X. This setup is compared to the benchmark using CENNT with the same reduced uncertainty model and number of nuisance parameters.

3 Systematic uncertainty aware neural network training

3.1 Emergence and methodology

In this section we elaborate on the abstract work mechanism of a CENNT-based measurement, as summarized in Sect. 2, and where it falls short. Based on this discussion we discuss the emergence and key difference of SANNT. Eventually, we point out the main challenges of SANNT and how they have been addressed. Measurements discussed in the scope of this paper imply the estimation of one or more POIs of a statistical model, as defined in Eq. (3), from the comparison with data. Without loss of generality, we restrict the discussion in this section to one POI that we call r_s. The measurement objective is to estimate r_s with minimal uncertainty Δr_s and no bias, corresponding to the highest sensitivity of the measurement to this POI.

The essence of statistical parameter estimation is to reduce a potentially large amount of information to a human-manageable and interpretable level. In the recent past, a sizable number of high-energy particle physics measurements have witnessed a significant increase in their sensitivity replacing traditional, human-inspired strategies for event classification by a CENNT. A typical flow chart for such an analysis is illustrated in the upper part of Fig. 1.

In a first step, the initial data set $D_X = \{\mathbf{x}_0, \ldots, \mathbf{x}_{n-1}\} \subseteq \mathbb{R}^{d\times n}$ with d characterizing observables $\mathbf{x} \in X \subseteq \mathbb{R}^d$, is reduced to a lower-dimensional summary statistic $\hat{\mathbf{y}}(\mathbf{x}, \boldsymbol{\omega}) \in Y \subseteq \mathbb{R}^l$, e.g. through NN (multiclass) classification, where l denotes the number of target classes and $\boldsymbol{\omega}$ refers to the trainable parameters of the NN model, to form a reduced data set $D_Y = \{\hat{\mathbf{y}}_0(\mathbf{x}, \boldsymbol{\omega}), \ldots, \hat{\mathbf{y}}_{l-1}(\mathbf{x}, \boldsymbol{\omega})\} \subseteq \mathbb{R}^{l\times n}$. Since the CENNT corresponds to the maximum likelihood estimate of $\mathbb{P}(\{k_l\}, n, \{p_l\})$, it is asymptotically efficient in fulfilling the separation task based on the provided d-dimensional feature space X [19,20], so that the compression from d to l happens with minimal loss of information relevant for the separation.

For the analysis strategy discussed here, in a second step, the number of n events is reduced to h bins of a histogram

$$H = \left\{ \sum_{j=0}^{n-1} H_i(\hat{\mathbf{y}}(\mathbf{x}_j, \boldsymbol{\omega})) \,\middle|\, i = 0 \ldots h-1 \right\}, \tag{4}$$

with

$$H_i\left(\hat{\mathbf{y}}(\mathbf{x}, \boldsymbol{\omega})\right) = \begin{cases} 1, & \text{if } \hat{\mathbf{y}}(\mathbf{x}, \boldsymbol{\omega}) \text{ in the boundaries of bin } i \\ 0, & \text{otherwise,} \end{cases}$$

resulting in an even further reduced data set $D_H = \{H_0(\hat{\mathbf{y}}(\mathbf{x}, \boldsymbol{\omega})), \ldots, H_{h-1}(\hat{\mathbf{y}}(\mathbf{x}, \boldsymbol{\omega}))\} \subseteq \mathbb{N}^{l\times h}$. In this nomenclature we will refer to the Asimov data set $\{k_i^{\mathrm{A}}\}$ introduced in Section 2.4 as $D_H^{\mathcal{A}}$ with $k_i^{\mathcal{A}} = H_i(\hat{\mathbf{y}}(\mathbf{x}^{\mathcal{A}}, \boldsymbol{\omega}))$, in the following.

A flaw of this strategy remains in the objective of the CENNT, i.e., the separation of S_s from B_b based on $\mathbb{P}(\{k_l\}, n, \{p_l\})$, not being the same as the objective of the measurement, which is the minimization of Δr_s. However, the ansatz remains successful, as long as the objectives of both estimates are approximately aligned, which can be assumed if the uncertainties in r_s are dominated by their statistical component $\Delta r_s^{\mathrm{stat}}$ and the yield of S_s is small relative to B_b.

The prime target of a SANNT, as proposed, e.g., in Ref. [14], is to become as congruent with the measurement objective as possible. From a given likelihood model, this can be achieved through the Fisher information matrix [18]

$$F_{\alpha_{i'}\alpha_{j'}} = \mathbb{E}\left[\frac{\partial^2}{\partial\alpha_{i'}\partial\alpha_{j'}}\Big(-\log\mathcal{L}\Big)\right]_{\hat{\alpha}_{i'}\hat{\alpha}_{j'}}, \tag{5}$$

with

$$\alpha_{i'}\,\alpha_{j'} \in \{\{r_s\}, \{\theta_j\}\},$$

where in our case $\mathcal{L}$ is defined by Eq. (3) and $\hat{\alpha}_i$ refers to the value of α_i that maximise $\mathcal{L}$. The expectation value is estimated from $D_H^{\mathcal{A}}$, which is derived from the complete training sample, to minimize statistical fluctuations. Following the Cramér–Rao inequality [19,20], in the limit where the law of large numbers applies, Δr_s can be obtained from

$$\Delta r_s = \sqrt{F^{-1}_{r_s r_s}}. \tag{6}$$

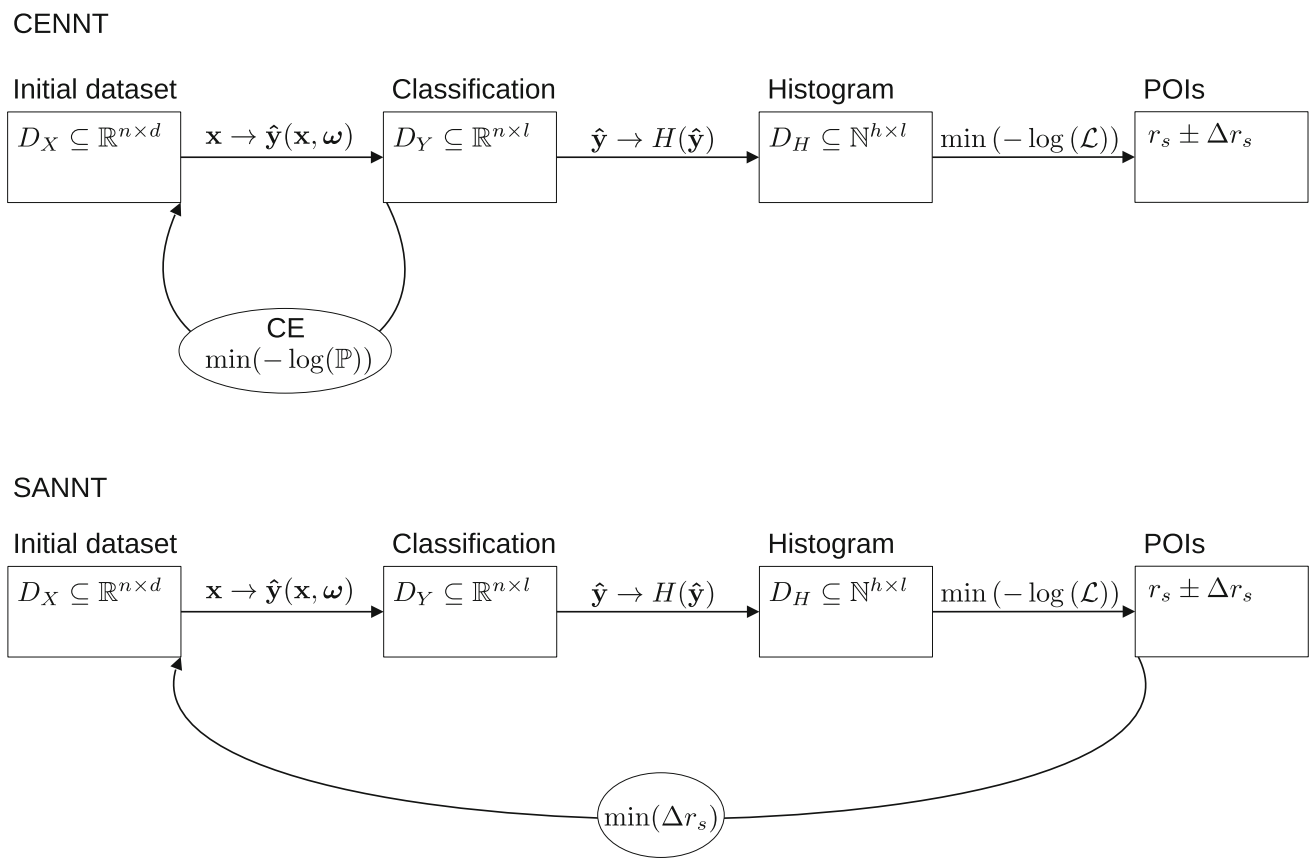

Fig. 1 Flow chart of a CENNT (upper) and SANNT (lower). In the figure D_i denotes the data set; n (d) the number of events (observables) in the initial data set D_X; l the number of classes after event classification; and h the number of histogram bins entering the statistical inference of the POIs. The function symbol $\mathbb{P}$ represents the multinomial distribution, and the symbol $\mathcal{L}$ has been defined in Eq. (3)

This estimate of Δr_s replaces CE as a loss function L_{SANNT} for SANNT. The resulting modified flow chart is illustrated in the lower part of Fig. 1.

The obvious advantage of this ansatz is that $\mathcal{L}$ may not only include the statistical, but also the systematic component of Δr_s. Thus, it also leads to the most sensitive estimate based on D_H for measurements that are not dominated by Δr_s^{stat}. On the downside, this ansatz has to address a number of challenges. It implies the computationally costly calculation and inversion of $F_{\alpha_i \alpha_j}$ during training. Moreover, it introduces an ambiguous choice of binning and implies the conceptual issue that $\mathcal{L}$, as defined in Eq. (3), is not differentiable at its bin edges, a circumstance that prevents standard backpropagation through differentiation. One solution to this conceptional issue has been proposed in Ref. [13], where the histogramming step is replaced by a modifiable softmax operator to approximate the histogram and sustain differentiability. In an ansatz followed in Ref. [15], a kernel density estimate replaces the histogram. A likelihood ratio based on this approximation is computed and subsequently used as the objective for the NN training.

At this point, we note that differentiability of L_{SANNT} is not strictly required for backpropagation, as long as the information of proximity to a decision boundary, such as the bin edges of $\mathcal{L}$ in this case, is preserved. The solution proposed in Ref. [14] is to introduce a sum of custom functions $\mathcal{B}_i$ for the backward pass of the backpropagation instead of using the gradient. In the forward pass the loss function is evaluated as usual. The role of $\mathcal{B}_i$ is to allow movements of events across bin boundaries based on residual information about their positions inside a given bin. The exact form of $\mathcal{B}_i$ is a hyperparameter choice comparable to the choice of activation function for hidden nodes of the NN architecture. The same is true for the number of bins summarizing $\hat{y}_l(\cdot)$ to form the input to $\mathcal{L}$, while we have not performed any particular optimization scan for this parameter.

The use of L_{SANNT} reveals a number of peculiarities, which are particularly important for the studies presented in Sects. 4 and 5:

– Considering l as a signal, the choice of CE as loss function implies a strict ordering in $\hat{y}_l(\cdot)$ for events being

background- or signal-like. The choice of L_{SANNT} does not imply such an ordering. Instead signal and background events obtain values of $\hat{y}_l(\cdot)$ that lead to the smallest value of Δr_s. This can most easily be appreciated from Eq. (3), which is the basis for the calculation of Δr_s. There, each bin is represented by a term i under the first product. Since the scalar outputs of $\mathcal{P}(\cdot|\cdot)$ commute under multiplication, the result of $\mathcal{L}$ does not depend on the order in which the bins are presented to $\mathcal{L}$.

– A training with Δr_s as objective is still fully supervised, but the event labels enter L_{SANNT} in a more subtle way than in the case of CE, viz. through the implicit modeling of S_s and B_b. All processes are normalized according to their expected abundances in $D_H^{\mathcal{A}}$.
– A training with Δr_s as objective only affects nuisance parameters θ_j related to shape-altering variations of $\hat{y}_l(\cdot)$. This can be best appreciated restricting the discussion to the optimization with respect to only one nuisance parameter. The parameters $\boldsymbol{\omega}$ forming the NN model induce variations of $\hat{y}_l(\cdot)$. Thus, the concrete minimization of Δr_s, which is tied to the values of $\boldsymbol{\omega}$, takes place through rearrangements of background and signal samples in the value space of $\hat{y}_l(\cdot)$ exploiting regions that are less affected by variations of θ_j. If θ_j has a pure normalization effect on $\hat{y}_l(\cdot)$, the application of $\sum \mathcal{B}_i$ has no effect on the minimization of Δr_s, since the relative size of the variation is the same across the whole value space of $\hat{y}_l(\cdot)$. We note at this point that normalization uncertainties may well be constrained during the maximum likelihood fit used for the extraction of the signal.

We conclude this section with a note of caution. A SANNT of any kind draws information from Ω_X including all the systematic variations in consideration. The additional information from the uncertainty model leads to the reduction of Δr_s compared to the outcome of a CENNT. In turn, SANNT implies stricter requirements on the quality and understanding of Ω_X. This can be appreciated by the following thoughts. We imagine a deficiency of Ω_X to describe the data in X, which is erroneously covered by a systematic variation that does not address the root cause of the deficiency. After SANNT, $\hat{y}(\cdot)$ might be derived from a region of X where the erroneously covering variation is significantly smaller. As a consequence, Ω_X might reveal a discrepancy with data that can no longer be covered by the now reduced variation. Along these lines, we anticipate that any random component in the estimation of a shape-altering systematic variation, e.g., caused by limited sizes of alternative samples used for estimating it, increases the variance of outcomes of a SANNT, namely the reduction of Δr_s. In the extreme of a purely random variation associated with θ_{rnd}, events with larger or smaller associated variation would be randomly scattered in X. After SANNT the variation associated with θ_{rnd} would be minimized to best effort, while due to the random scattering of affected events in X there is no reason why this minimization should have a positive influence on the reduction of Δr_s. Rather, decreasing or increasing the importance of random subsets in X for signal extraction should lead to a larger variance of potential outcomes of a SANNT on Δr_s. Both thoughts emphasize that special care should be taken to assure that the uncertainty model as an integral part of Ω_X be well defined and physically well motivated. While this is of obvious importance for any analysis, it is emphasized by the explicit use of the uncertainty model during a SANNT.

3.2 Adaptations to the neural network training

In this section we discuss a small number of adaptations that we have made to the SANNT, as originally proposed in Ref. [14]. For purposes of illustration, we base this discussion on the same set of pseudo-experiments as used in the original publication. This set consists of one signal and one background sample, each modeled by a multivariate Gaussian distribution centered around the points $(0, 0)$ in the case of signal and $(1, 1)$ in the case of background, in a two-dimensional feature space of a fully-connected feed-forward NN with features x_1 and x_2, one hidden layer of 100 nodes with ReLU activation function, and one single output $\hat{y}(\cdot) \in [0, 1]$ for binary classification with the sigmoid activation function. Each Gaussian distribution is chosen with an identity covariance matrix. For the background sample, a systematic variation is introduced shifting the mean of the distribution by one unit along the direction of x_2.

The original choice for $\mathcal{B}_i$ was the derivative of a Gaussian density with the expectation value μ_i at the center and standard deviation σ_i corresponding to half the width of H_i. This choice drags samples away from the edges of H_i leaving them untouched in the center. It is illustrated in Fig. 2 (left), where $\hat{y}(\cdot)$ is summarized by a histogram with eight bins, the boundaries of which are indicated by the dashed vertical lines.

The bars in the first row represent 20 randomly chosen samples of $\hat{y}(\cdot)$, in the second row, a snapshot of H, as obtained from these samples is shown. In a realistic setup, H would be filled from a corresponding batch of the training sample. The following rows display $\mathcal{B}_0$ and $\mathcal{B}_1$ for H_0 and H_1, respectively. The last row shows the collective effect of $\sum \mathcal{B}_i$ over all H_i. The intended effect per bin is clearly visible from the displays of $\mathcal{B}_0$ and $\mathcal{B}_1$. The display of $\sum \mathcal{B}_i$ however also reveals undesired effects beyond the bins that the $\mathcal{B}_i$ are supposed to act upon, visible as interference effects with neighboring bins. In total, this leads to largely reduced values of $\sum \mathcal{B}_i$ for the central bins of H whereas the original magnitude of $\mathcal{B}_i$ is only retained at the lower and upper

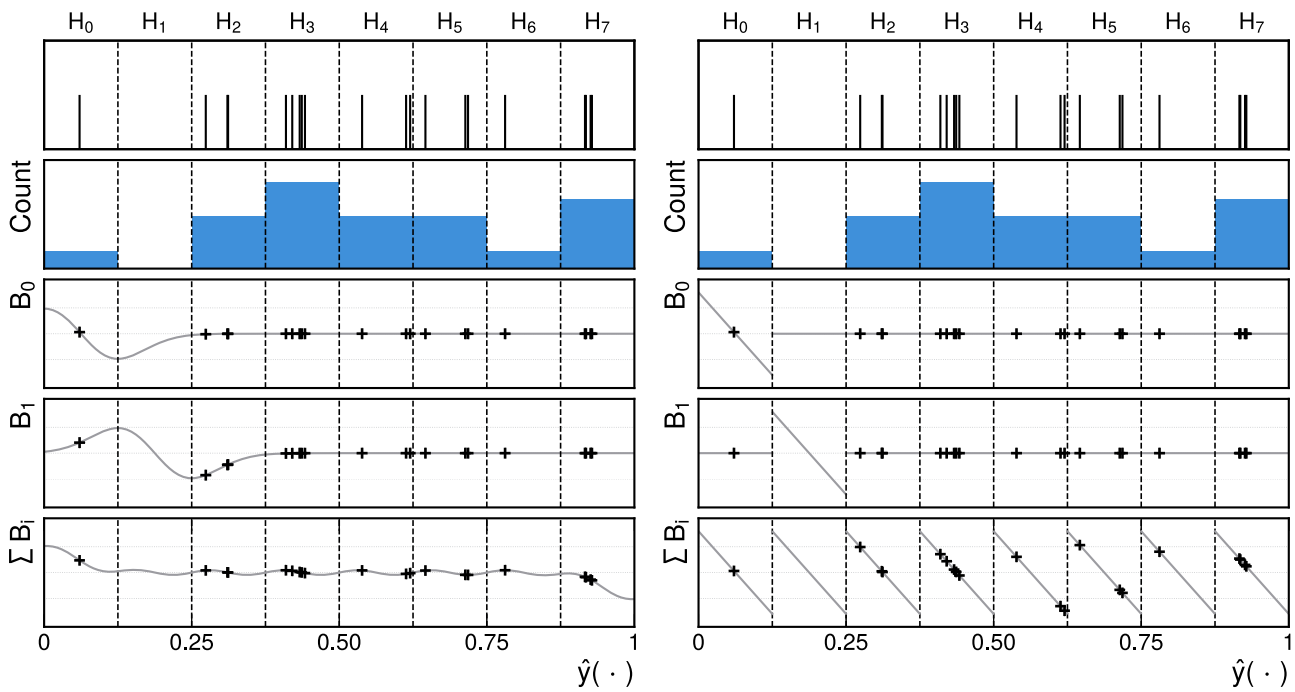

Fig. 2 Custom functions $\mathcal{B}_i$ for the backward pass of the backpropagation algorithm, as used (left) in Ref. [14] and (right) in this paper. In the first row of each sub-figure, the same 20 random samples of a simple setup of pseudo-experiments, as described in Sect. 3.2, are shown. In the second row the resulting histogram H, in the third and fourth rows the functions $\mathcal{B}_0$ and $\mathcal{B}_1$ for the individual bins H_0 and H_1, respectively, and in the last row the collective effect of $\sum \mathcal{B}_i$ are shown

bounds of H. This further leads to reduced movements across bin edges with a tendency to drag events off the lower and upper bounding bins, towards the center of H.

To make the backpropagation process more transparent and at the same time more comprehensible we have replaced the original choice of $\mathcal{B}_i$ by a simple linear function with intercept in the center of H_i. More importantly, we evaluate $\mathcal{B}_i$ only within the boundaries of H_i and set it to zero otherwise. The effect of these modifications is shown in Fig. 2 (right). The desired effect per bin is preserved, movements across bin edges are better supported, and the lower and upper bounding bins are treated as equal to any other H_i.

Another adaptation that we have made refers to the related initial observation in Ref. [14] that a training based only on Δr_s tends to collapse $\hat{y}(\cdot)$ into a usually very small number of central bins, after a few optimization steps. To mitigate this effect and improve the training stability, a pre-training of 30 optimization steps, evaluating only Δr_s^{stat} as loss function, had been applied, prior to the full use of L_{SANNT}, in the original publication.

We have replaced the loss function for pre-training by the binary CE, which all processes enter according to their expected abundance. In addition,we have extended the pre-training to 300 optimization steps to guarantee convergence. As discussed in Sect. 3.1, this modification should have an effect similar to the use only of Δr_s^{stat}, as shown in the third panel from above of Fig. 3 (right), where a reduction of Δr_s^{stat} during the CE-based pre-training is clearly visible. In addition, it introduces a separation of signal from background as an explicit pre-training objective and sustains a weak ordering of events being background- or signal-like equivalent to the CE-based pre-training. From the practical point of view, it leads to a wider spread of the samples across the available value space of $\hat{y}(\cdot)$. This strategy relies on the assumption that the result of a CENNT, which coincides with the minimization only of Δr_s^{stat}, naturally forms a good starting point for the SANNT. In the transition from a statistically to a systematically dominated measurement, the SANNT acts as a continuous extension of the CENNT.

The evolution of Δr_s^{stat} and Δr_s for a training based on the original setup and over 1000 optimization steps is shown in Fig. 3 (left). In the upper panel of the figure, the evolution of $\hat{y}(\cdot)$ for randomly selected 50 (blue) signal and 50 (orange) background samples during training is shown, which we assume to be representative of the full background and signal samples. The optimization steps of each training are shown on the x axis and the values of $\hat{y}(\cdot)$ on the y axis of the figure. The horizontal dashed lines indicate the bin edges of H, the gray shaded area indicates the pre-training. In Fig. 3 (right), the same evolution is shown for the modified setup.

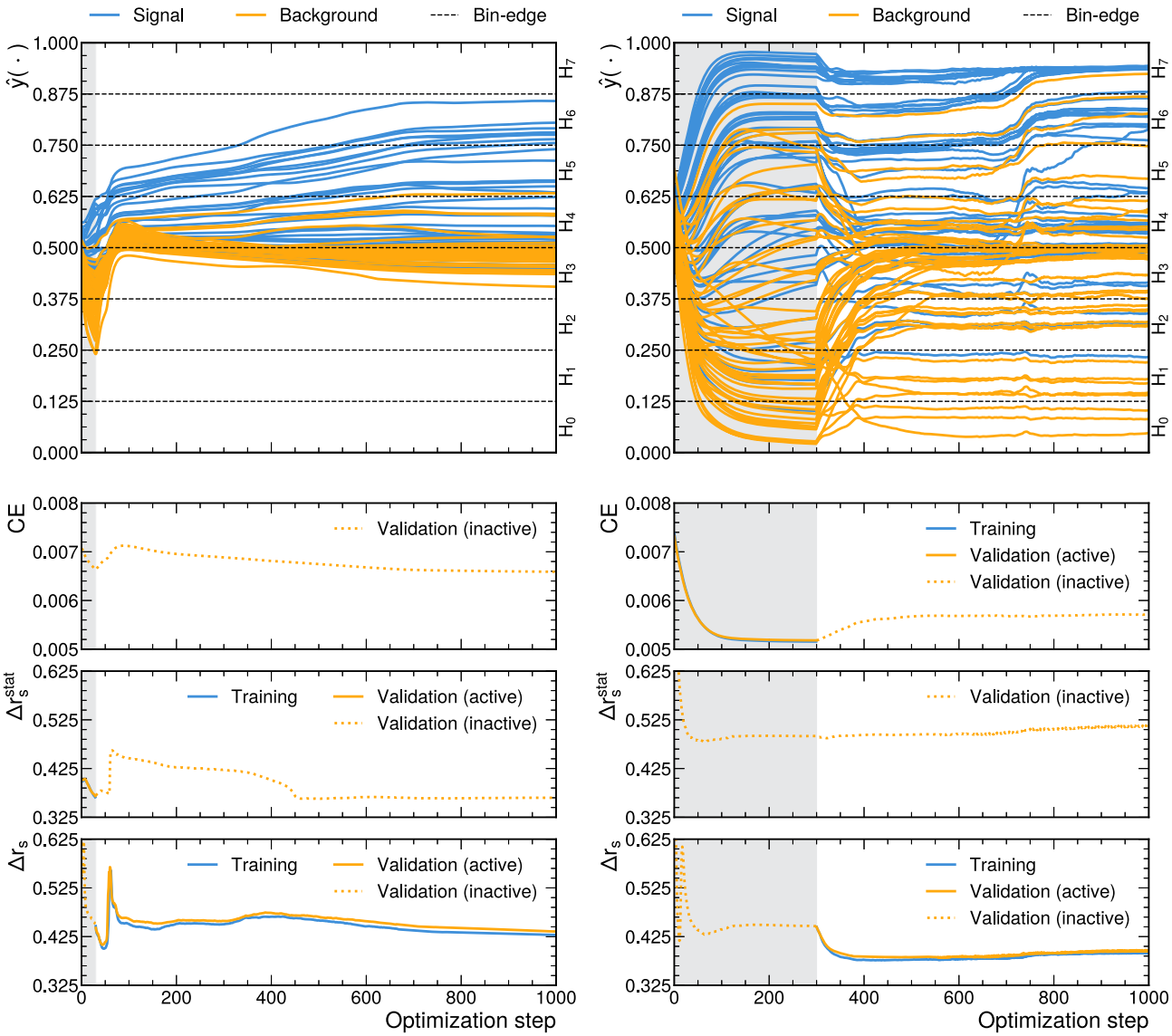

Fig. 3 Evolution of the loss functions CE, Δr_s^{stat}, and Δr_s as used (left) in Ref. [14] and (right) for this paper. In the upper panels the evolution of $\hat{y}(\cdot)$ for randomly selected 50 (blue) signal and 50 (orange) background samples during training is shown. Dashed horizontal lines indicate the boundaries of the histogram bins H_i. The gray shaded area indicates the pre-training. In the second and third panels from above the evolution of CE and Δr_s^{stat} is shown. In the lowest panels the evolution of $L_{\text{SANNT}} = \Delta r_s$ is shown. The evaluation on the training (validation) data set is indicated in blue (orange). The evaluation of the correspondingly inactive loss function, during or after pre-training, evaluated on the validation data set is indicated by the dashed orange curves

A much more pronounced spread of the samples across the value space of $\hat{y}(\cdot)$ after pre-training is observed, which is sustained despite significant rearrangements, particularly of some signal samples, across histogram bins. For both setups, the value of $\hat{y}(\cdot)$ stabilizes for each sample after roughly 500 optimization steps.

In the lower part of the figure, the evolution of the loss functions over the same period of optimization steps is shown, split by (second panel) CE, (third panel) Δr_s^{stat}, and (last panel) $L_{\text{SANNT}} = \Delta r_s$ using (left) Δr_s^{stat} and (right) CE during pre-training for SANNT. The evaluations on the training and validation samples are indicated by the blue and orange curves, respectively. Beyond the last optimization step of the pre-training, from which on Δr_s^{stat} (CE) turns inactive, the blue line is not shown. It is still instructive to follow the evolution of Δr_s^{stat} and CE on the validation sample as indicated by the orange dashed lines. The complementary case for Δr_s during pre-training and SANNT is shown in the lower panels of the figure.

As Δr_s^{stat} and CE are based on different likelihood models their magnitudes can not be directly compared. For Δr_s, however, a direct comparison of the evolution by magnitude is valid. In both setups, the pre-training coincides with a drop in Δr_s. Convergence of the pre-training is more obviously reached in the case of CE. In both cases, the decrease in Δr_s continues after pre-training, as intended. In the original setup however, a rearrangement of events across the very few bins, shown in the upper left part of the figure, leads to a significant increase in Δr_s^{stat}, visible in the third left panel from above. Shortly after, the increase in Δr_s^{stat} is outweighed by the decrease of the systematic component of Δr_s. After 500 more optimization steps the model succeeds in slowly re-adapting to a minimal value of Δr_s^{stat}, similar to the one obtained after pre-training. A qualitatively similar, but more comprehensible progression is visible for the modified setup. The value of Δr_s very quickly drops to a stable level within 100 optimization steps. The decrease in the systematic contribution of Δr_s happens at the cost of a slight increase in CE and Δr_s^{stat}. The increase of the latter is less pronounced, which can be explained by the looser requirements on Δr_s^{stat} compared to CE, as discussed in Sect. 3.1.

The $\mathcal{B}_i$ might be viewed as an extension of the so-called straight-through estimator (STE) where $\mathcal{B}_i$ is replaced by the identity operation [66]. We note that the choice of $\mathcal{B}_i$ features a number of important conceptual differences compared to the STE. Firstly, $\mathcal{B}_i$ acts only within the boundaries of bin i. Secondly, and related to the first point, $\mathcal{B}_i$ supports explicit migrations across the bins of $\mathcal{L}$, depending on the exact value of $\hat{y}(\cdot)$ within the given bin. We identify these migrations as an essential ingredient to the successful minimization of Δr_s. The evolution of a training equivalent to the one discussed here, with the only difference of replacing the $\mathcal{B}_i$ by identity operations is shown in Fig. 9 of Appendix A.

4 Comparison with the cross-entropy-based training setup

It is instructive to compare the two training setups for a binary classification task separating S from B. For this purpose, all S_s and all B_b (as discussed in Sect. 2.3) have been grouped into one sample each.

The expected distributions of $\hat{y}(\cdot)$ after CENNT, prior to any fit to $D_H^{\mathcal{A}}$, are shown in the upper panel of Fig. 4 (left). The expectations for S and B are shown by the nonstacked open blue and the filled orange histogram, respectively. The expected value of $S/B + 1$ is shown in the lower panel of the figure, where the gray band corresponds to the combined statistical and systematic uncertainty in B. A clear increase of S/B is visible with the largest value in H_7 most to the right of H, where S is expected to contribute $\approx 5\%$ to the overall yield. The largest relative effect of the uncertainty in B is expected in H_6.

Figure 4 (right) shows the same distributions after SANNT. Compared to the CENNT, both S and B are biased towards the left of the histogram and the expected largest relative effect of the uncertainties in B is shifted towards H_2 and H_3. We note that after SANNT there is no longer a strict ordering of B to the left and S to the right of H, while due to the pre-training this ordering is still weakly preserved. We have checked that the linear correlation coefficients between the output of $\hat{y}(\cdot)$ after CENNT and SANNT are 0.80 for S and 0.70 for B. This reduction in correlation relates to the impact that systematic variations have on Δr_s. We anticipate that the larger this impact is, the smaller the correlations will be.

This observation gives the following heuristic insights into the working mechanism of the SANNT implemented here. Depending on their position in X, events in D_X are subject to systematic variations. For a well-defined classification task, S and B predominantly populate distinct subspaces $\mathcal{S}_{\text{S}}, \mathcal{S}_{\text{B}} \subset X$ with some potential transition region of overlapping events. For the given example, illustrated by Fig. 4, $\hat{y}(\cdot)$ provides a mapping from X to $\hat{y} \in [0, 1]$. A fixed value of $\hat{y} = \kappa$ corresponds to the hyperplane of a discriminating boundary $\mathcal{C}_\kappa$ in X. Without loss of generality, a systematic "up"-variation might move an event lying on $\mathcal{C}_\kappa$ and belonging to B closer to $\mathcal{S}_{\text{S}}$. The "down"-variation in turn might move a similar event lying on $\mathcal{C}_\kappa$ and belonging to S closer to $\mathcal{S}_{\text{B}}$. We observe that the elements of $\mathbf{x}$ are weighted such that these events will obtain values of $\hat{y} \lesssim \kappa$, after SANNT. This explains the two observations that can be made from Fig. 4: (i) the general shift of $\hat{y}(\cdot)$ towards smaller values of $\hat{y}$ as well as; (ii) the larger values of $S/B + 1$ for the events in bins H_6 and H_7, visible from the lower panel of Fig. 4 (right). There may well be more subtle effects steering the exploitation of features $\mathbf{x}$ by the NN model playing a role into this observation.

To visualize the influence of single features and their pairwise linear correlations, obtained from the support of S and B during training, on $\hat{y}(\cdot)$ we have used the metric

$$\langle t_\alpha \rangle = \frac{1}{n} \sum_{i=1}^{n} |t_\alpha(\mathbf{x}_i)| \tag{7}$$

based on a Taylor expansion of $\hat{y}(\cdot)$, with respect to the elements of $\mathbf{x}$ up to second order, as described in Ref. [67]. In Eq. (7), α refers to a power set of the components of $\mathbf{x}$, where elements may also be repeated, t_α refers to the Taylor series coefficient for the derivative of $\hat{y}(\cdot)$ with respect to α, and n to the size of the training sample. Large (small) values of $\langle t_\alpha \rangle$ indicate that feature α has a large (small) influence on

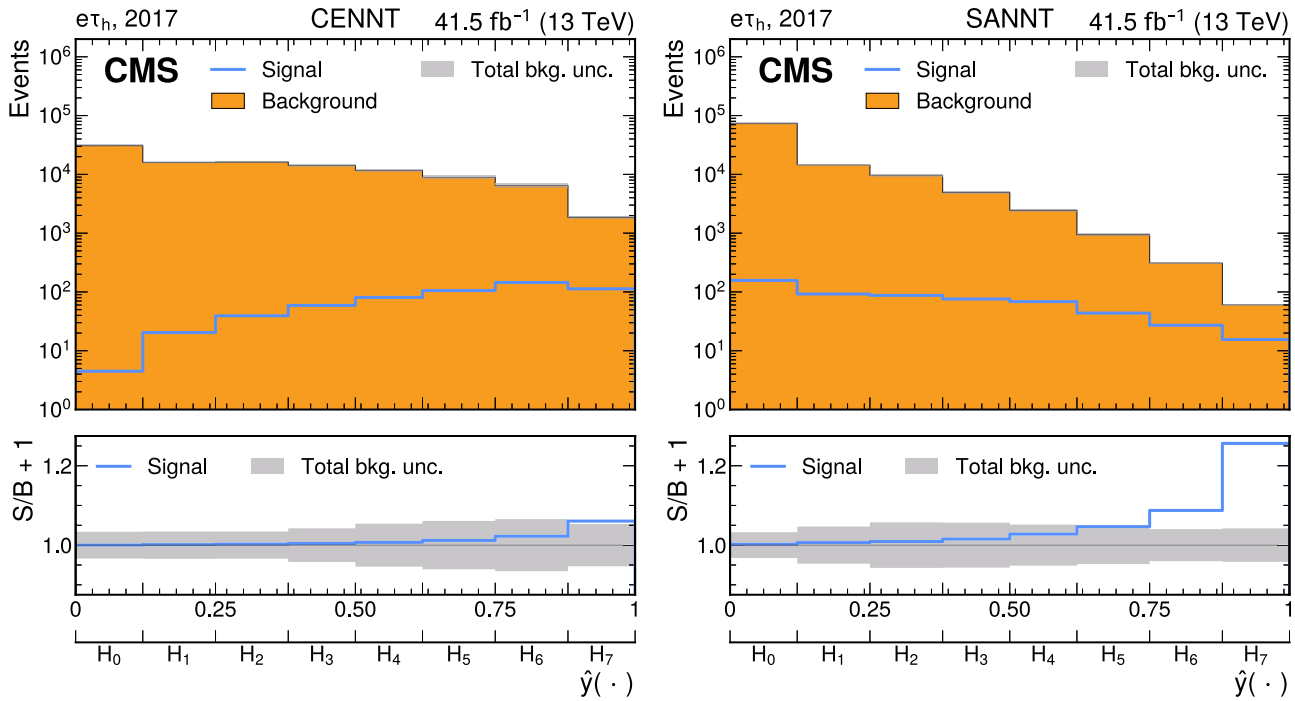

Fig. 4 Expected distributions of $\hat{y}(\cdot)$ for a binary classification task separating S from B, for a (left) CENNT and (right) SANNT, prior to any fit to $D_H^{\mathcal{A}}$. The individual distributions for S and B are shown by the nonstacked open blue and filled orange histograms, respectively. In the lower panels of the figures, the expected values of $S/B+1$ are shown. The gray bands correspond to the combined statistical and systematic uncertainty in B. The boundaries of the histogram bins H_i are also given on an extra horizontal axis

$\hat{y}(\cdot)$. For the SANNT we have replaced the derivative of $\mathcal{L}$ in the Taylor expansion by $\sum \mathcal{B}_i$.

From this study, we have identified that especially the first-order coefficients for $\alpha = m_{\tau\tau}$ and m_{vis} and the second-order coefficients for $\alpha = (m_{\text{vis}}, m_{\text{vis}})$ and $(m_{\tau\tau}, m_{\tau\tau})$ have a large influence on $\hat{y}(\cdot)$, after CENNT. The coefficients $t_{m_{\tau\tau}}$ and $t_{m_{\text{vis}}}$ indicate the location in the marginal distributions of S and B; the coefficients $t_{m_{\tau\tau},m_{\tau\tau}}$ and $t_{m_{\text{vis}},m_{\text{vis}}}$ indicate the curvatures of each distribution and thus how much it is peaking or not. Despite the simplifications of the analysis, these findings, which follow our physics intuition, are in good accord with the findings of Ref. [8]. In the transition from CENNT to SANNT, we observe that the influences of the properties associated with $t_{m_{\text{vis}},m_{\text{vis}}}$, $t_{m_{\text{vis}}}$, and $t_{m_{\tau\tau}}$ are largely reduced in favor of the curvature of $m_{\tau\tau}$, associated with $t_{m_{\tau\tau},m_{\tau\tau}}$.

The 20 nuisance parameters with the largest impacts on r_s, obtained from fitting the $S+B$-model with $r_s = 1$ to $D_H^{\mathcal{A}}$ after shifting each parameter θ_j individually by one standard deviation up and down fixing all other $\{\theta_i;\ i \neq j\}$ to their most probable values after fit, are shown in Fig. 5. The impacts can be read off from the x axis. The gray lines refer to the CENNT and the colored bars to the SANNT. Labels for each θ_j, decreasing in magnitude for the CENNT when moving from upper to lower entries in the figure, are given on the y axis, referring to the following sets of systematic uncertainties:

- Uncertainties mostly related to the τ-embedding method, viz. in ($\epsilon_\tau^{\text{trg}}$) the efficiency to trigger the $e\tau_h$ pair; ($\epsilon_\tau^{\text{ID}}(D_e)$) the efficiency to distinguish τ_h candidates from electrons; and the efficiency to distinguish τ_h candidates with ($\epsilon_\tau^{\text{ID}}(35, 40)$) $35 < p_T^{\tau_h} < 40$ GeV and ($\epsilon_\tau^{\text{ID}}(40, 500)$) $40 < p_T^{\tau_h} < 500$ GeV from quark- or gluon-induced jets. For $\epsilon_\tau^{\text{ID}}(40, 500)$ additional uncorrelated parts for τ_h decays ($\epsilon_\tau^{\text{ID}}$(3-prong)) with three charged pions, and ($\epsilon_\tau^{\text{ID}}$(1-prong*)) one charged and additional neutral pions appear in the list.
- Uncertainties related to the F_F-method, viz. in the normalizations of the estimates of events with zero, one, and two jets corresponding to F_F(0-jet), F_F(1-jet), and F_F(2-jet), respectively; the nonclosure correction to the estimate of the background contribution from QCD multijet production as a function of m_{vis}, $F_F^{\text{QCD}}(m_{\text{vis}})$; and the subtraction of W+jets events as obtained from simulation during the estimation of the same background contribution, F_F^{QCD}(W+jets).

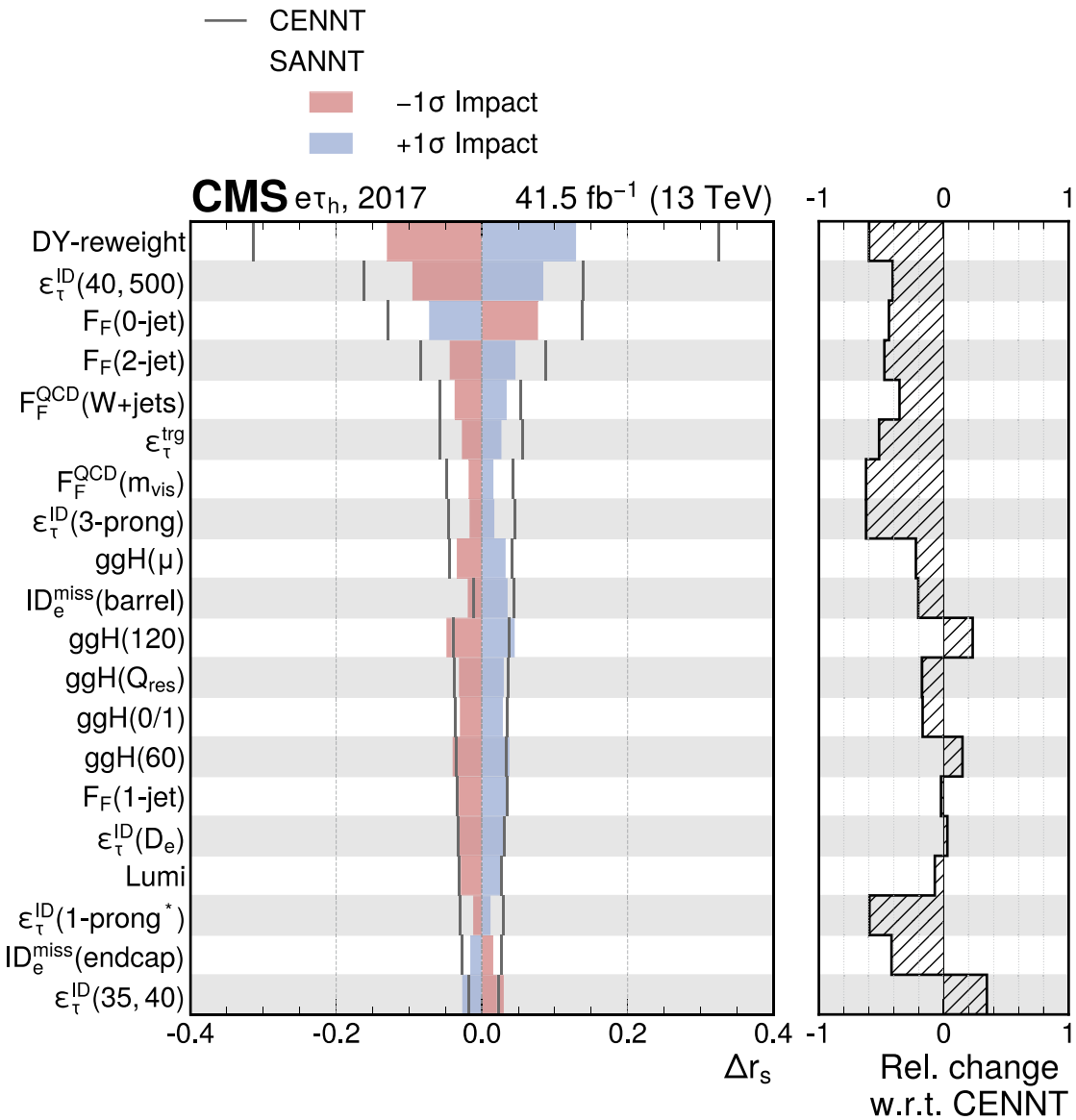

Fig. 5 Impacts for the 20 nuisance parameters θ_j with the largest impacts on r_s. The gray lines refer to the CENNT and the colored bars to the SANNT. The impacts can be read from the x axis. Labels shown on the y axis for each θ_j are defined in Table 1. The entries are ordered by decreasing magnitude for CENNT when moving from the top to the bottom of the figure. The panel on the right shows the relative change of the symmetrized impact when moving from CENNT to SANNT. A more detailed discussion is given in the text

– Uncertainties related to simulation, viz. in the reweighting of the sample used for estimating the background from $Z \rightarrow \ell\ell$ production, in sideband regions of the data (DY reweight); the misidentification rate of quark- or gluon-induced jets as electron candidates in the ($\mathrm{ID}_e^{\mathrm{miss}}$(barrel)) barrel and ($\mathrm{ID}_e^{\mathrm{miss}}$(endcap)) endcap regions of the detector; and the integrated luminosity (Lumi).
– A number of uncertainties related to the modeling of the ggH signal, viz. from variations of (ggH(μ)) the (μ_r) renormalization and (μ_f) factorization, as well as (ggH(Q_{res})) the resummation scales; migration effects across STXS bins (ggH(0/1)) with 0 and 1 jet, as well as the p_T^H STXS bin boundaries at (ggH(60)) 60 and (ggH(120)) 120 GeV.

A summary of these variations is given in Table 1. The largest impact of a single uncertainty on r_s for CENNT is observed for DY reweight, followed by $\epsilon_\tau^{\mathrm{ID}}(40, 500)$ and F_{F}(0-jet). With a few exceptions, all uncertainties are sig-

Table 1 Association of nuisance parameters θ_j with the systematic variations they refer to for the 20 parameters with the largest impacts on r_s, as shown in Fig. 5. The label of each corresponding uncertainty is given in the first column, the type of uncertainty, process that it applies to, and rank in Fig. 5 are given in the second, third, and fourth columns, respectively. The "*" in $\epsilon_\tau^{\mathrm{ID}}$(1-prong*) refers to the fact that this is the decay channel with neutral pions in addition to the charged prong. The symbol EMB refers to τ-embedded event samples. A more detailed discussion is given in the text

Label	Type	Process	Rank	Norm	Shape	Comment
$\epsilon_\tau^{\mathrm{trg}}$	τ trigger	EMB	6	---	✓	---
$\epsilon_\tau^{\mathrm{ID}}(D_{\mathrm{e}})$	τ ID	MC, EMB	16	✓	---	Discr. against e
$\epsilon_\tau^{\mathrm{ID}}(35, 40)$	τ ID	EMB	20	---	✓	$35 < p_{\mathrm{T}}^{\tau_{\mathrm{h}}} < 40\,\mathrm{GeV}$
$\epsilon_\tau^{\mathrm{ID}}(40, 500)$	τ ID	EMB	2	---	✓	$40 < p_{\mathrm{T}}^{\tau_{\mathrm{h}}} < 500\,\mathrm{GeV}$
$\epsilon_\tau^{\mathrm{ID}}$(1-prong*)	τ ID	EMB	18	---	✓	One π^++π^0's
$\epsilon_\tau^{\mathrm{ID}}$(3-prong)	τ ID	EMB	8	---	✓	Three π^+'s
F_{F}(0-jet)	Norm.	F_{F}	3	---	✓	$N_{\mathrm{jet}} = 0$
F_{F}(1-jet)	Norm.	F_{F}	15	---	✓	$N_{\mathrm{jet}} = 1$
F_{F}(2-jet)	Norm.	F_{F}	4	---	✓	$N_{\mathrm{jet}} = 2$
$F_{\mathrm{F}}^{\mathrm{QCD}}(m_{\mathrm{vis}})$	Nonclosure	F_{F}	7	---	✓	In m_{vis}
$F_{\mathrm{F}}^{\mathrm{QCD}}$(W+jets)	Subtr.	F_{F}	5	---	✓	Subtr. of sim.
ggH(μ)	Theory	ggH	9	---	✓	μ_{r} and μ_{f}
ggH(Q_{res})	Theory	ggH	12	---	✓	Resummation
ggH(0/1)	Theory	ggH	13	---	✓	$0 \rightarrow 1$ jet migr.
ggH(60)	Theory	ggH	14	---	✓	$p_{\mathrm{T}}^{\mathrm{H}}$ migr.
ggH(120)	Theory	ggH	11	---	✓	$p_{\mathrm{T}}^{\mathrm{H}}$ migr.
$\mathrm{ID}_{\mathrm{e}}^{\mathrm{miss}}$(barrel)	e misID	MC	10	---	✓	Barrel
$\mathrm{ID}_{\mathrm{e}}^{\mathrm{miss}}$(endcap)	e misID	MC	19	---	✓	Endcap
DY reweight	Reweight	MC	1	---	✓	In $p_{\mathrm{T}}^{\mu\mu}$ and $m_{\mu\mu}$
Lumi	Int. luminosity	MC	17	✓	---	---

nificantly reduced by the SANNT demonstrating the success of the ansatz. The uncertainty due to DY reweight, with the largest impact on r_s, is reduced by nearly 60%. The $\mathrm{Z} \rightarrow \mathrm{ee}$ process, related to this uncertainty, is closely located to S in m_{vis}. The reduction of the DY reweight uncertainty is related to the reduced influence of the location and curvature of m_{vis} on $\hat{y}(\cdot)$ previously discussed in this section.

Figure 6 (upper) shows distributions of the negative log of the profile likelihood $-2\Delta \log \mathcal{L}$ as a function of r_s, based on the input distributions as depicted in Fig. 4. For the red (blue) curves all (only the statistical) uncertainties in r_s have been taken into account.

For the latter, all nuisance parameters have been fixed to their most probable values after fit, before performing a second fit to $D_H^{\mathcal{A}}$.

The most probable value of r_s is observed at 1, as expected from a fit to $D_H^{\mathcal{A}}$ demonstrating that the estimate is unbiased. The values of Δr_s and $\Delta r_s^{\mathrm{stat}}$ are obtained from the intervals in r_s with $-2\Delta \log \mathcal{L} < 1$. The results as obtained from the CENNT are indicated by the dashed lines. The continuous lines indicate the median results of ensembles of 100 independent repetitions each of the SANNT varying random initializations. The red and blue shaded bands surrounding the median expectations indicate the 68% central intervals of these ensembles, representing the variances of the obtained results due to random choices for training. In the lower panels, the underlying distributions to these central intervals are shown in red and blue, correspondingly. Based on SANNT (CENNT) values of $\Delta r_s = ^{+0.47}_{-0.44}$ ($^{+0.62}_{-0.60}$) are obtained, corresponding to a reduction of 25% in magnitude for the result obtained from SANNT with respect to the CENNT. Taking only the statistical component of Δr_s into account, a value of $\Delta r_s^{\mathrm{stat}} = ^{+0.31}_{-0.30}$ (± 0.29) for SANNT (CENNT) is obtained. This indicates that the SANNT is a significant improvement to the CENNT. It is the first time that the success of a SANNT has been demonstrated for an analysis at this level of complexity.

5 Extension for multiclass classification

In this section we report the application of SANNT to the multiclass-classification task, based on the original STXS cross section measurement described in Ref. [8] exploiting the five background and two signal classes, as described in Sect. 2.4. Each class l is represented by an output node of the NN with output function $\hat{y}_l(\cdot)$. The effect of multiclass classification is achieved through the use of a softmax operator in the output layer of the NN providing a probability estimate for an event with a feature vector $\mathbf{x}$ to originate from l. The CENNT is obtained from a simple extension of binary to categorical CE.

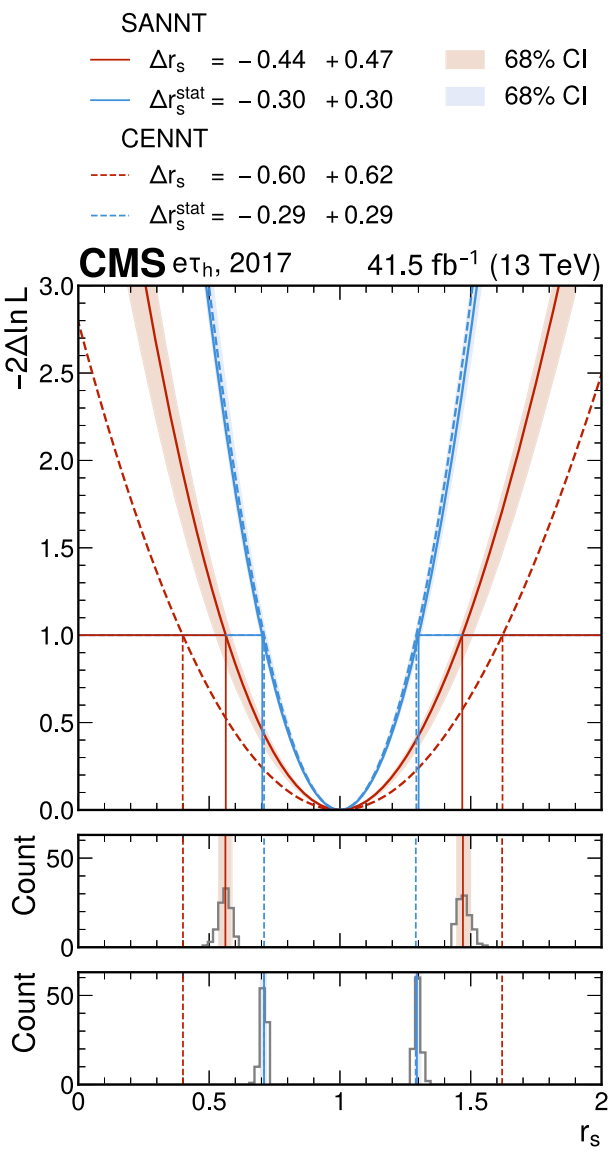

Fig. 6 Negative log of the profile likelihood $-2\Delta \log \mathcal{L}$ as a function of r_s, taking into account (red) all and (blue) only the statistical uncertainties in Δr_s. The results obtained from the CENNT are indicated by the dashed lines, and the median expected resultS of ensembles of 100 repetitions of the SANNT varying random initializations are indicated by the continuous lines. The red and blue shaded bands surrounding the median expectations indicate 68% confidence intervals (CI) from these ensembles. The lower panels show the distributions underlying these CI

Naively the same replacement of binary by categorical CE would apply for the SANNT pre-training. Instead, we keep the sigmoid activation function, applying it to all output nodes and exploit the custom function

$$\mathrm{CE}' = -\sum_i \left(w_i \sum_l \mathrm{CE}_l \right), \tag{8}$$

where CE_l corresponds to the binary CE evaluated separately for each l, i runs over all events in $D_H^{\mathcal{A}}$, and w_i normalizes each event to the expectation for the process it belongs to and the integrated luminosity of the assumed data set. We have made this modification as we observed a significant improvement of the subsequent SANNT, which as in the previous case of binary classification, we attribute to the practical effect of a wider spread of samples across the value spaces of $\hat{y}_l(\cdot)$.

As discussed at the end of Sect. 3.1, the fact that any binned likelihood function, e.g., as defined in Eq. (3), lacks immanent information about the relations between bins, necessarily implies the loss of an ordering principle for the association of events of class l with a given value of $\hat{y}_l(\cdot)$. More importantly in the case of multiclass classification, this even prevents a fixed association of events to l. While this circumstance does not pose an issue of principle for a multiclass-classification SANNT, it renders the conventional concept of output classes for multiclass classification meaningless, since, for example, there is no rule in the training to prevent events of class "genuine τ" from being associated with the output node meant to collect events of class "misc" and vice versa. Even worse, both event classes could be associated with the same output node.

In Sect. 3.2 we have introduced the result of a CENNT pre-training as natural starting point for the SANNT, preserving a weak relation between the class that an event belongs to and the value of $\hat{y}(\cdot)$, in the transition from statistically towards systematically dominated measurements. To sustain the concept of output classes we further enforce this ansatz through an explicit constraint in L_{SANNT}, retaining the result of the pre-training $\mathrm{CE}'_{\mathrm{min}}$ resembling a good separation between classes. For the subsequent SANNT, we modify L_{SANNT} to read

$$L_{\mathrm{SANNT}}^{\mathrm{mult.}} = \sum_s \Delta r_s + \omega_\lambda \, g(\cdot), \tag{9}$$

with

$$g(\cdot) = \mathrm{CE}' - \mathrm{CE}'_{\mathrm{min}},$$

following an approach as proposed in Ref. [68], with ω_λ as an additional trainable parameter of the NN model. Formally, this turns the SANNT into an optimization task with the achieved separation between classes after pre-training, expressed by $\mathrm{CE}'_{\mathrm{min}}$, as a constraint equation, where ω_λ takes the role of a Lagrange multiplier. In the nonexcluded case of $g(\cdot) < 0$, we set $\omega_\lambda = 0$ to allow improvements of the process separation during SANNT. By the sum in Eq. (9), all signal processes s are treated equally, which is an arbitrary choice.

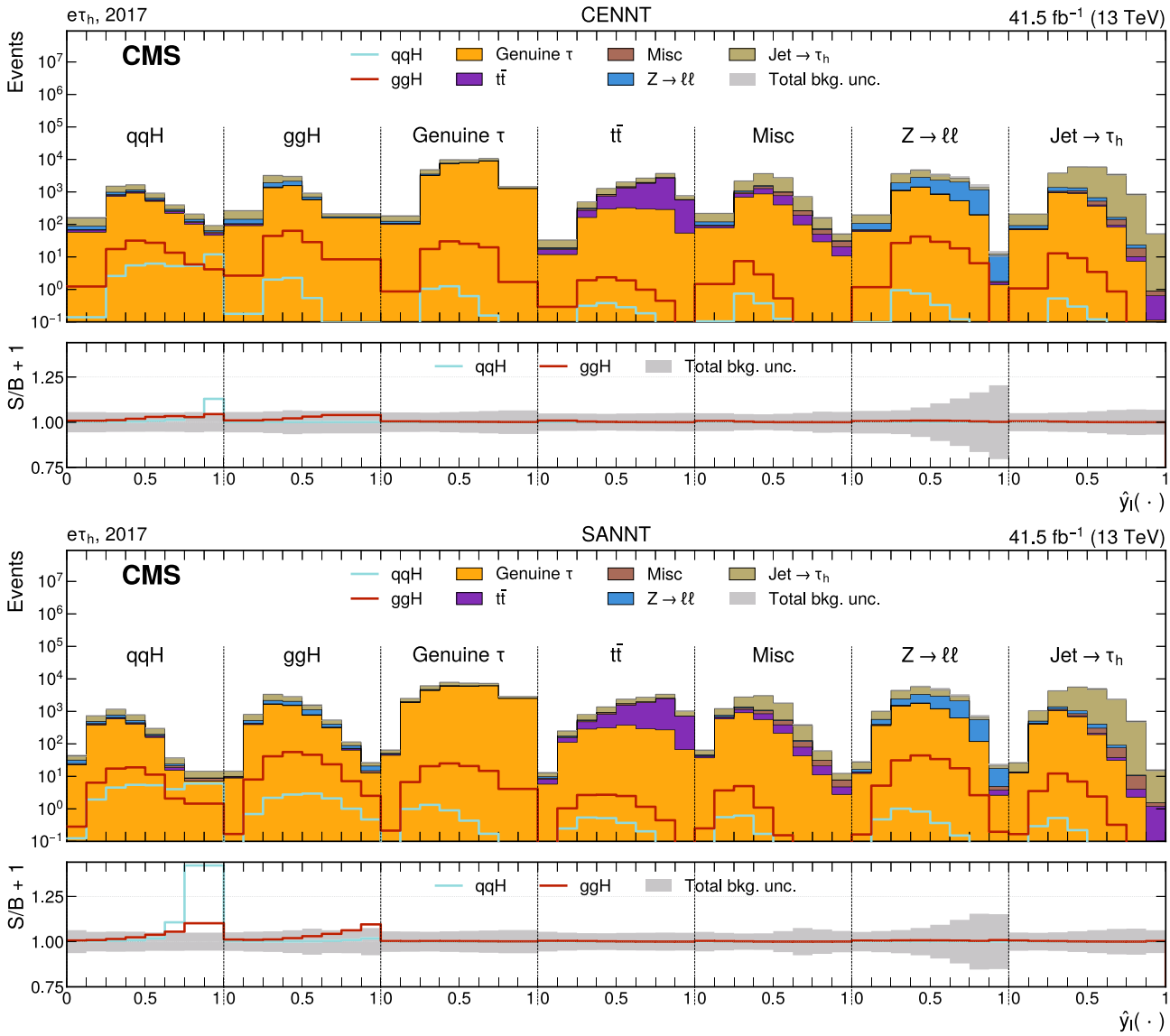

Fig. 7 Expected distributions of $\hat{y}_l(\cdot)$ for multiclass classification, based on seven event classes, as used for a differential STXS cross section measurement of H production in Ref. [8], prior to any fit to $D_H^{\mathcal{A}}$. In the upper (lower) part of the figure the results obtained after CENNT (SANNT) are shown. The background processes of Ω_X are indicated by stacked, colored, filled histograms. The expected ggH and qqH contributions are indicated by the nonstacked open histograms. In the lower panels of the figure the expected values of $S/B+1$ are shown. The gray bands correspond to the combined statistical and systematic uncertainty in the background model

Figure 7 (upper) shows the expected distributions of $\hat{y}_l(\cdot)$ for each background and signal class, and the expected values of $S/B+1$ after CENNT prior to any fit to $D_H^{\mathcal{A}}$. Individual background processes of Ω_X are indicated by stacked, differently colored, filled histograms. The expected ggH and qqH contributions are indicated by the nonstacked, red- and cyan-colored, open histograms. As for the previous discussion in Sect. 4, the gray bands in the lower panels of the figure correspond to the combined statistical and systematic uncertainty in the background model. The same distributions after SANNT are shown in Fig. 7 (lower). The change of distributions and more prominently the improved signal separation, with reduced uncertainty are visible from the comparison of the corresponding distributions, indicating the success of the training.

Figure 8 shows distributions of $-2\Delta\log\mathcal{L}$, based on the input distributions as depicted in Fig. 7, as a function of r_s. As in the previous discussion, in Sect. 4, for the red (blue) curves all (only the statistical) uncertainties in Δr_s have been taken into account. In the left part of the figure r_{inc}, for an inclusive measurement of the H production cross section in the $\text{H}\to\tau\tau$ decay channel is shown, in the middle and right

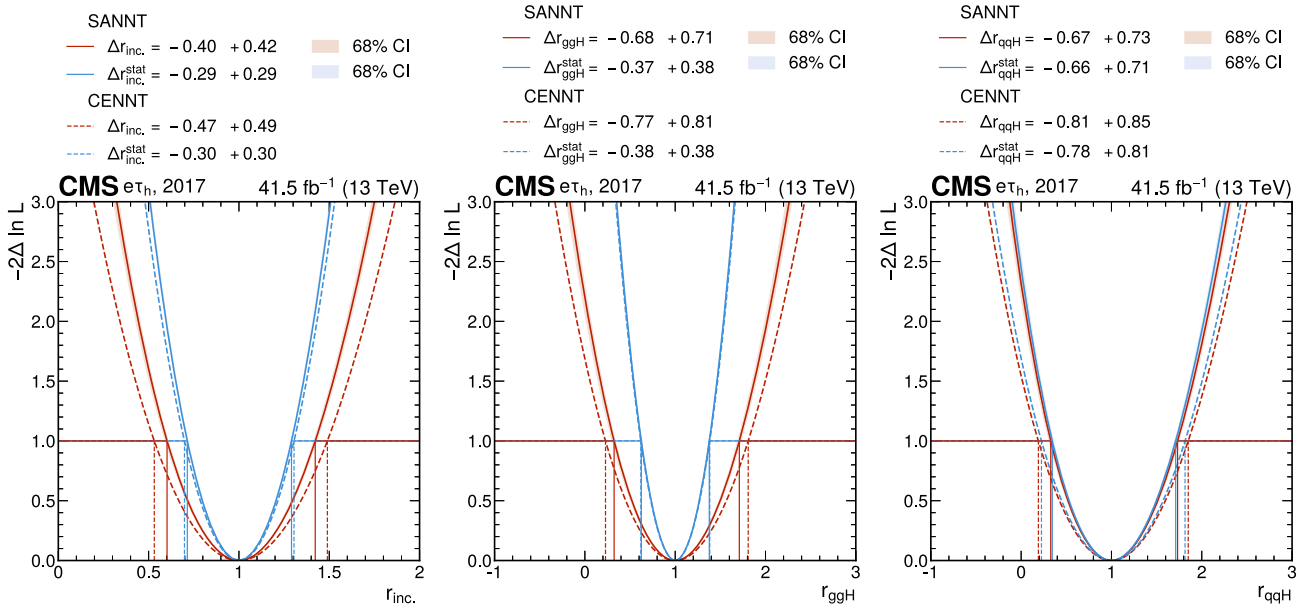

Fig. 8 Negative log of the profile likelihood $-2\Delta \log \mathcal{L}$ as a function of r_s, for a differential STXS cross section measurement of H production in the H $\rightarrow \tau\tau$ decay channel, taking (red) all and (blue) only the statistical uncertainties in Δr_s into account. In the left plot r_{inc} for an inclusive measurement is shown, and in the middle and right plots r_{ggH} and r_{qqH} for a combined differential STXS measurement of these two contributions to the signal in two bins are shown. The results as obtained from the CENNT are indicated by the dashed lines, and the median expected results of an ensembles of 100 repetitions of the SANNT varying random initializations are indicated by the continuous lines. The red and blue shaded bands surrounding the median expectations indicate 68% confidence intervals (CI) from these ensembles

Table 2 Expected combined statistical and systematic uncertainties Δr_s and statistical uncertainties Δr_s^{stat} in the parameters r_{inc} for an inclusive, and r_{ggH} and r_{qqH} for a differential STXS, cross section measurement of H production in the H $\rightarrow \tau\tau$ decay channel, as obtained from fits to D_H^A. In the second (third) column the results after SANNT (CENNT) are shown

r_s	Δr_s (Δr_s^{stat})	
	SANNT	CENNT
r_{inc}	$^{+0.42}_{-0.40}$ (± 0.29)	$^{+0.49}_{-0.47}$ (± 0.30)
r_{ggH}	$^{+0.71}_{-0.68}$ ($^{+0.38}_{-0.37}$)	$^{+0.81}_{-0.77}$ (± 0.38)
r_{qqH}	$^{+0.73}_{-0.67}$ ($^{+0.71}_{-0.66}$)	$^{+0.85}_{-0.81}$ ($^{+0.81}_{-0.78}$)

parts of the figure r_{ggH} and r_{qqH} for a combined differential STXS measurement of these contributions to the signal in two bins, are shown. The numerical results of the fits are summarized in Table 2.

As observed previously, the method is unbiased. After SANNT, an improvement of 15% relative to the result as obtained after CENNT is observed for r_{inc}. For r_{ggH} and r_{qqH}, improvements of 12 and 16% are observed, respectively. This constitutes a significant improvement of the SANNT over the CENNT. It is the first time that a multiclass-classification SANNT has been successfully demonstrated.

We observe that SANNT leads to an improvement of the statistical component of r_{qqH}, where the contribution of systematic variations to the overall uncertainty has been small, already from the beginning. This can be explained by the fact that the targets of both trainings are not in sync. While L_{SANNT} has been constructed to be as close to the measurement target as possible, the target of CENNT is the best possible separation of all processes comprising two signal and five background processes, in this case. While for the case of binary classification discussed in Sect. 4, this particular difference between the training methods was covered by the experimental setup, it becomes more obvious here and emphasizes the conceptual advantage of SANNT over CENNT.

6 Summary

We have proposed a neural network training method capable of accounting for the effects of systematic variations of the data model used in the training process and described its extension to neural network multiclass classification. The procedure has been evaluated on the realistic case of the measurement of Higgs boson production via gluon fusion and vector boson fusion in the $\tau\tau$ decay channel at the CMS experiment. The neural network output functions are used to

infer the signal strengths for inclusive production of Higgs bosons as well as for their production via gluon fusion and vector boson fusion. We observe improvements of 12 and 16% in the uncertainty in the signal strengths for gluon and vector-boson fusion, respectively, compared with a conventional neural network training based on cross-entropy. This is the first time that a neural network training capable of accounting for the effects of systematic variations has been demonstrated on a data model of this complexity and the first time that such a training has been applied to multiclass classification.

Acknowledgements We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid and other centers for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: SC (Armenia), BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENESCYT (Ecuador); ERC PRG, RVTT3 and MoER TK202 (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); SRNSF (Georgia); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LMTLT (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MES and NSC (Poland); FCT (Portugal); MESTD (Serbia); MICIU/AEI and PCTI (Spain); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); MHESI and NSTDA (Thailand); TUBITAK and TENMAK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie program and the European Research Council and Horizon 2020 Grant, contract Nos. 675440, 724704, 752730, 758316, 765710, 824093, 101115353, 101002207, and COST Action CA16108 (European Union); the Leventis Foundation; the Alfred P. Sloan Foundation; the Alexander von Humboldt Foundation; the Science Committee, project no. 22rl-037 (Armenia); the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Beijing Municipal Science & Technology Commission, No. Z191100007219010 and Fundamental Research Funds for the Central Universities (China); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Shota Rustaveli National Science Foundation, grant FR-22-985 (Georgia); the Deutsche Forschungsgemeinschaft (DFG), among others, under Germany's Excellence Strategy -- EXC 2121 "Quantum Universe" -- 390833306, and under project number 400140256 - GRK2497; the Hellenic Foundation for Research and Innovation (HFRI), Project Number 2288 (Greece); the Hungarian Academy of Sciences, the New National Excellence Program - ÚNKP, the NKFIH research grants K 131991, K 133046, K 138136, K 143460, K 143477, K 146913, K 146914, K 147048, 2020-2.2.1-ED-2021-00181, TKP2021-NKTA-64, and 2021-4.1.2-NEMZ_KI-2024-00036 (Hungary); the Council of Science and Industrial Research, India; ICSC -- National Research Center for High Performance Computing, Big Data and Quantum Computing and FAIR -- Future Artificial Intelligence Research, funded by the NextGenerationEU program (Italy); the Latvian Council of Science; the Ministry of Education and Science, project no. 2022/WK/14, and the National Science Center, contracts Opus 2021/41/B/ST2/01369 and 2021/43/B/ST2/01552 (Poland); the Fundação para a Ciência e a Tecnologia, grant CEECIND/01334/2018 (Portugal); the National Priorities Research Program by Qatar National Research Fund; MICIU/AEI/10.13039/501100011033, ERDF/EU, "European Union NextGenerationEU/PRTR", and Programa Severo Ochoa del Principado de Asturias (Spain); the Chulalongkorn Academic into Its 2nd Century Project Advancement Project, and the National Science, Research and Innovation Fund via the Program Management Unit for Human Resources & Institutional Development, Research and Innovation, grant B39G670016 (Thailand); the Kavli Foundation; the Nvidia Corporation; the SuperMicro Corporation; the Welch Foundation, contract C-1845; and the Weston Havens Foundation (USA).

Data availability Release and preservation of data used by the CMS Collaboration as the basis for publications is guided by the CMS data preservation, re-use and open access policy.

Code Availability Statement The CMS core software is publicly available on GitHub (https://github.com/cms-sw/cmssw).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A: Custom backpropagation based on identity operations

Figure 9 shows the evolution of $\hat{y}(\cdot)$, CE, Δr_s^{stat}, and Δr_s, as used for this paper, for a SANNT where the custom functions $\mathcal{B}_i$ have been replaced by an identity operation, also referred to as STE in the literature [66].

The upper panel of the figure shows the evolution of $\hat{y}(\cdot)$ during training, for randomly selected 50 (blue) signal and 50 (orange) background samples, which we assume to be representative of the full background and signal samples. The optimization steps of the training are shown on the x axis and the values of $\hat{y}(\cdot)$ on the y axis of the figure. The horizontal dashed lines indicate the bin edges of H, the gray shaded area indicates the pre-training based on CE. We observe an even more pronounced collapse of $\hat{y}(\cdot)$, into bins H_6 and H_7

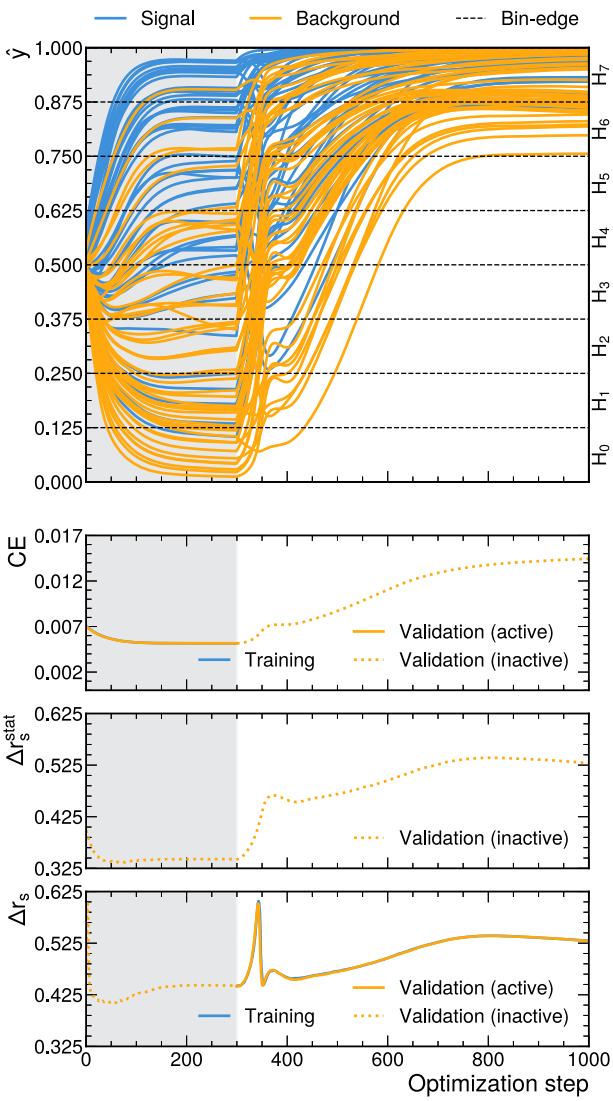

Fig. 9 Evolution of the loss functions CE, Δr_s^{stat}, and Δr_s, as used for this paper. Instead of the custom functions $\mathcal{B}_i$ the identity operation (the so-called straight-through estimator) is used for SANNT. In the upper panel, the evolution of $\hat{y}$ for randomly selected 50 (blue) signal and 50 (orange) background samples during training is shown. The gray shaded area indicates the pre-training. In the second panel from above the evolution of CE is shown. Though not actively used for the SANNT Δr_s^{stat} is also shown, in the third panel from above. In the lower panel, the evolution of Δr_s is shown. The evaluation on the training (validation) data set is indicated in blue (orange). The evolution of inactive loss functions, evaluated on the validation data set, is indicated by the dashed orange curves

in that case, as for the original choice of $\mathcal{B}_i$, shown in Fig. 3 (left).

The lower part of the figure shows the evolution of (second panel) CE, (third panel) Δr_s^{stat}, and (last panel) Δr_s over the same period of optimization steps. The evaluations on the training and validation samples are indicated by the blue and orange curves, respectively. The continuous lines indicate the active loss function during each corresponding optimization step. The evolution of inactive loss functions on the validation sample is indicated by orange dashed lines. The progression of $\hat{y}(\cdot)$ and CE during pre-training is identical to Fig. 3 (right), while we have adapted the y axis range here to be able to display the full development of CE up to optimization step 1000. After pre-training we observe that both Δr_s and Δr_s^{stat} increase with a pronounced structure in the course of Δr_s around optimization step 350, coinciding with the period of a few optimization steps, during which $\hat{y}(\cdot)$ collapses into H_6 and H_7. Even after 1000 optimization steps, Δr_s has not improved beyond its value immediately after pre-training, which demonstrates that $\mathcal{B}_i$ can not be replaced by a naive STE and that this change rather destabilizes the training, compromising its success.

References

1. ATLAS Collaboration, Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. Phys. Lett. B **716**, 1 (2012). https://doi.org/10.1016/j.physletb.2012.08.020. arXiv:1207.7214
2. CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys. Lett. B **716**, 30 (2012). https://doi.org/10.1016/j.physletb.2012.08.021. arXiv:1207.7235
3. CMS Collaboration, Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV. JHEP **06**, 081 (2013). https://doi.org/10.1007/JHEP06(2013)081. arXiv:1303.4571
4. CMS Collaboration, A portrait of the higgs boson by the CMS experiment ten years after the discovery. Nature **607**, 60 (2022). https://doi.org/10.1038/s41586-022-04892-x. arXiv:2207.00043. [Erratum: 10.1038/s41586-023-06164-8]
5. ATLAS Collaboration, Characterising the Higgs boson with ATLAS data from Run 2 of the LHC. Phys. Rept. **11**, 001 (2024). https://doi.org/10.1016/j.physrep.2024.11.001. arXiv:2404.05498
6. ATLAS Collaboration, Measurements of Higgs bosons decaying to bottom quarks from vector boson fusion production with the ATLAS experiment at $\sqrt{s}$ = 13TeV. Eur. Phys. J. C **81**, 537 (2021). https://doi.org/10.1140/epjc/s10052-021-09192-8. arXiv:2011.08280
7. CMS Collaboration, Measurement of simplified template cross sections of the Higgs boson produced in association with W or Z bosons in the H → bb decay channel in proton-proton collisions at $\sqrt{s}$ = 13TeV. Phys. Rev. D **109**, 092011 (2024). https://doi.org/10.1103/PhysRevD.109.092011. arXiv:2312.07562
8. CMS Collaboration, Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton-proton collisions at $\sqrt{s}$ = 13TeV. Eur. Phys. J. C **83**, 562 (2023). https://doi.org/10.1140/epjc/s10052-023-11452-8. arXiv:2204.12957
9. ATLAS Collaboration, "Differential cross-section measurements of Higgs boson production in the H → $\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s}$ = 13TeV with the ATLAS detector", 2024. arXiv:2407.16320. Submitted to *JHEP*
10. I.J. Good, Rational decisions. J. Royal Stat. Soc. B **14**, 107 (1952). https://doi.org/10.1111/j.2517-6161.1952.tb00104.x
11. R. M. Neal, "Computing likelihood functions for high-energy physics experiments when distributions are defined by simulators with nuisance parameters", in *PHYSTAT-LHC Workshop on Sta-*

tistical Issues for LHC Physics, p. 111. 2008. https://doi.org/10.5170/CERN-2008-001.111
12. K. Cranmer, J. Pavez, and G. Louppe, "Approximating likelihood ratios with calibrated discriminative classifiers", 2015. arXiv:1506.02169. https://doi.org/10.48550/arXiv.1506.02169
13. P. Castro, T. Dorigo, INFERNO: Inference-aware neural optimisation. Comput. Phys. Commun. **244**, 170 (2019). https://doi.org/10.1016/j.cpc.2019.06.007. arXiv:1806.04743
14. S. Wunsch, S. Jörger, R. Wolf, G. Quast, Optimal statistical inference in the presence of systematic uncertainties using neural network optimization based on binned Poisson likelihoods with nuisance parameters. Comput. Softw. Big Sci. **5**, 4 (2021). https://doi.org/10.1007/s41781-020-00049-5. arXiv:2003.07186
15. N. Simpson, L. Heinrich, neos: End-to-end-optimised summary statistics for high energy physics. J. Phys: Conf. Ser. **2438**, 012105 (2023). https://doi.org/10.1088/1742-6596/2438/1/012105. arXiv:2203.05570
16. G. Louppe, M. Kagan, and K. Cranmer, "Learning to pivot with adversarial networks", in *Advances in Neural Information Processing Systems*, p. 982. 2017. arXiv:1611.01046
17. S. Wunsch, S. Jörger, R. Wolf, G. Quast, Reducing the dependence of the neural network function to systematic uncertainties in the input space. Comput. Softw. Big Sci. **4**, 5 (2020). https://doi.org/10.1007/s41781-020-00037-9. arXiv:1907.11674
18. R.A. Fisher, On the mathematical foundations of theoretical statistics. J. R. Stat. Soc. **82**, 87 (1922). https://doi.org/10.1098/rsta.1922.0009
19. H. Cramér, "Mathematical methods of statistics". Princeton University Press, 1946
20. C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters", in *Breakthroughs in statistics*, p. 235. Springer, 1992
21. CMS Collaboration, Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}$ = 13 TeV. JINST **15**, P10017 (2020). https://doi.org/10.1088/1748-0221/15/10/P10017. arXiv:2006.10165
22. C.M.S. Collaboration, The CMS trigger system. JINST **12**, P01020 (2017). https://doi.org/10.1088/1748-0221/12/01/P01020. arXiv:1609.02366
23. CMS Collaboration, Performance of the CMS high-level trigger during LHC Run 2. JINST **19**, P11021 (2024). https://doi.org/10.1088/1748-0221/19/11/P11021. arXiv:2410.17038
24. C.M.S. Collaboration, The CMS experiment at the CERN LHC. JINST **3**, S08004 (2008). https://doi.org/10.1088/1748-0221/3/08/S08004
25. CMS Collaboration, Particle-flow reconstruction and global event description with the CMS detector. JINST **12**, P10003 (2017). https://doi.org/10.1088/1748-0221/12/10/P10003. arXiv:1706.04965
26. CMS Collaboration, "Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid", CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015. http://cds.cern.ch/record/2020886
27. CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}$ = 8 TeV. JINST **10**, P06005 (2015). https://doi.org/10.1088/1748-0221/10/06/P06005. arXiv:1502.02701
28. CMS Collaboration, Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. JINST **16**, P05014 (2021). https://doi.org/10.1088/1748-0221/16/05/P05014. arXiv:2012.06888
29. C.M.S. Collaboration, Performance of CMS muon reconstruction in pp collision events at $\sqrt{s}$ = 7 TeV. JINST **7**, P10002 (2012). https://doi.org/10.1088/1748-0221/7/10/P10002. arXiv:1206.4071
30. CMS Collaboration, Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ = 13 TeV. JINST **13**, P06015 (2018). https://doi.org/10.1088/1748-0221/13/06/P06015. arXiv:1804.04528
31. M. Cacciari, G.P. Salam, G. Soyez, The anti-k_T jet clustering algorithm. JHEP **04**, 063 (2008). https://doi.org/10.1088/1126-6708/2008/04/063. arXiv:0802.1189
32. M. Cacciari, G.P. Salam, G. Soyez, FastJet user manual. Eur. Phys. J. C **72**, 1896 (2012). https://doi.org/10.1140/epjc/s10052-012-1896-2. arXiv:1111.6097
33. CMS Collaboration, Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. JINST **13**, P05011 (2018). https://doi.org/10.1088/1748-0221/13/05/P05011. arXiv:1712.07158
34. E. Bols et al., Jet flavour classification using DeepJet. JINST **15**, P12012 (2020). https://doi.org/10.1088/1748-0221/15/12/P12012. arXiv:2008.10519
35. CMS Collaboration, Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$ = 13 TeV. JINST **13**, P10005 (2018). https://doi.org/10.1088/1748-0221/13/10/P10005. arXiv:1809.02816
36. CMS Collaboration, Identification of hadronic tau lepton decays using a deep neural network. JINST **17**, P07023 (2022). https://doi.org/10.1088/1748-0221/17/07/P07023. arXiv:2201.08458
37. CMS Collaboration, Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV. JINST **10**, P02006 (2015). https://doi.org/10.1088/1748-0221/10/02/P02006. arXiv:1411.0511
38. D. Bertolini, P. Harris, M. Low, N. Tran, Pileup per particle identification. JHEP **10**, 059 (2014). https://doi.org/10.1007/JHEP10(2014)059. arXiv:1407.6013
39. S.D. Drell, T.-M. Yan, Massive lepton-pair production in hadron-hadron collisions at high energies. Phys. Rev. Lett. **25**, 316 (1970). https://doi.org/10.1103/PhysRevLett.25.316
40. CMS Collaboration, An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data. JINST **14**, P06032 (2019). https://doi.org/10.1088/1748-0221/14/06/P06032. arXiv:1903.01216
41. CMS Collaboration, Measurement of the $Z\gamma^* \to \tau\tau$ cross section in pp collisions at $\sqrt{s}$ = 13 TeV and validation of τ lepton analysis techniques. Eur. Phys. J. C **78**, 708 (2018). https://doi.org/10.1140/epjc/s10052-018-6146-9. arXiv:1801.03535
42. CMS Collaboration, Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV. JHEP **09**, 007 (2018). https://doi.org/10.1007/JHEP09(2018)007. arXiv:1803.06553
43. P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms. JHEP **11**, 040 (2004). https://doi.org/10.1088/1126-6708/2004/11/040. arXiv:hep-ph/0409146
44. S. Frixione, P. Nason, C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method. JHEP **11**, 070 (2007). https://doi.org/10.1088/1126-6708/2007/11/070. arXiv:0709.2092
45. S. Alioli, P. Nason, C. Oleari, E. Re, NLO Higgs boson production via gluon fusion matched with shower in POWHEG. JHEP **04**, 002 (2009). https://doi.org/10.1088/1126-6708/2009/04/002. arXiv:0812.0578
46. S. Alioli, P. Nason, C. Oleari, E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. JHEP **06**, 043 (2010). https://doi.org/10.1007/JHEP06(2010)043. arXiv:1002.2581
47. S. Alioli et al., Jet pair production in POWHEG. JHEP **04**, 081 (2011). https://doi.org/10.1007/JHEP04(2011)081. arXiv:1012.3380
48. E. Bagnaschi, G. Degrassi, P. Slavich, A. Vicini, Higgs production via gluon fusion in the POWHEG approach in the SM

and in the MSSM. JHEP **02**, 088 (2012). https://doi.org/10.1007/JHEP02(2012)088. arXiv:1111.2854
49. J. Alwall et al., MadGraph 5: Going beyond. JHEP **06**, 128 (2011). https://doi.org/10.1007/JHEP06(2011)128. arXiv:1106.0522
50. J. Alwall et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. JHEP **07**, 079 (2014). https://doi.org/10.1007/JHEP07(2014)079. arXiv:1405.0301
51. T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, https://doi.org/10.1016/j.cpc.2015.01.024, arXiv:1410.3012
52. S. Agostinelli et al., Geant4-a simulation toolkit. Nucl. Instrum. Meth. A **506**, 250 (2003). https://doi.org/10.1016/S0168-9002(03)01368-8
53. K. Fukushima, Cognitron: A self-organizing multilayered neural network. Biol. Cybernetics **20**, 121 (1975). https://doi.org/10.1007/BF00342633
54. V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines", in *Proc. 27th Int. Conf. on Machine Learning*, ICML'10, p. 807. Omnipress, Madison, WI, USA, 2010
55. L. Bianchini et al., Reconstruction of the Higgs mass in events with Higgs bosons decaying into a pair of τ leptons using matrix element techniques. Nucl. Instrum. Meth. A **862**, 54 (2017). https://doi.org/10.1016/j.nima.2017.05.001. arXiv:1603.05910
56. A.V. Gritsan, R. Röntsch, M. Schulze, M. Xiao, Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques. Phys. Rev. D **94**, 055023 (2016). https://doi.org/10.1103/PhysRevD.94.055023. arXiv:1606.03107
57. LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", *CERN Yellow Reports: Mongraphs* **2** (2017) https://doi.org/10.23731/CYRM-2017-002, arXiv:1610.07922
58. N. Berger et al., "Simplified template cross sections - stage 1.1", LHC Higgs Cross Section Working Group Report LHCHXSWG-2019-003, DESY-19-070, 2019. https://doi.org/10.48550/arXiv.1906.02754, arXiv:1906.02754
59. G. Cowan, K. Cranmer, E. Gross, O. Vitells, Asymptotic formulae for likelihood-based tests of new physics. Eur. Phys. J. C **71**, 1554 (2011). https://doi.org/10.1140/epjc/s10052-011-1554-0. arXiv:1007.1727. [Erratum: 10.1140/epjc/s10052-013-2501-z]
60. CMS Collaboration, The CMS statistical analysis and combination tool: Combine. Comput. Softw. Big Sci. **8**, 19 (2024). https://doi.org/10.1007/s41781-024-00121-4. arXiv:2404.06614
61. R.J. Barlow, C. Beeston, Fitting using finite Monte Carlo samples. Comput. Phys. Commun. **77**, 219 (1993). https://doi.org/10.1016/0010-4655(93)90005-W
62. CMS Collaboration, "CMS luminosity measurements for the 2016 data-taking period", Technical Report CMS-PAS-LUM-17-001, CERN, 2017. https://cds.cern.ch/record/2257069
63. CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$TeV", CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018. https://cds.cern.ch/record/2621960
64. CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019. https://cds.cern.ch/record/2676164
65. CMS Collaboration, Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. Eur. Phys. J. C **81**, 800 (2021). https://doi.org/10.1140/epjc/s10052-021-09538-2. arXiv:2104.01927
66. Y. Bengio, N. Léonard, and A. C. Courville, "Estimating or propagating gradients through stochastic neurons for conditional computation", 2013. arXiv:1308.3432. https://doi.org/10.48550/arXiv.1308.3432
67. S. Wunsch, R. Friese, R. Wolf, G, Quast, Identifying the relevant dependencies of the neural network response on characteristics of the input space. Comput. Softw. Big Sci. **2**, 5 (2018). https://doi.org/10.1007/s41781-018-0012-1. arXiv:1803.08782
68. J. C. Platt and A. H. Barr, "Constrained differential optimization", in *NIPS*, p. 612. American Institue of Physics, 1987

CMS Collaboration

Yerevan Physics Institute, Yerevan, Armenia
V. Chekhovsky, A. Hayrapetyan, V. Makarenko, A. Tumasyan[1]

Institut für Hochenergiephysik, Vienna, Austria
W. Adam, J. W. Andrejkovic, L. Benato, T. Bergauer, S. Chatterjee, K. Damanakis, M. Dragicevic, P. S. Hussain, M. Jeitler[2], N. Krammer, A. Li, D. Liko, I. Mikulec, J. Schieck[2], R. Schöfbeck[2], D. Schwarz, M. Sonawane, W. Waltenberger, C.-E. Wulz[2]

Universiteit Antwerpen, Antwerpen, Belgium
T. Janssen, H. Kwon, T. Van Laer, P. Van Mechelen

Vrije Universiteit Brussel, Brussels, Belgium
N. Breugelmans, J. D'Hondt, S. Dansana, A. De Moor, M. Delcourt, F. Heyen, Y. Hong, S. Lowette, I. Makarenko, D. Müller, S. Tavernier, M. Tytgat[3], G. P. Van Onsem, S. Van Putte, D. Vannerom

Université Libre de Bruxelles, Brussels, Belgium
B. Bilin, B. Clerbaux, A. K. Das, I. De Bruyn, G. De Lentdecker, H. Evard, L. Favart, P. Gianneios, A. Khalilzadeh, F. A. Khan, K. Lee, A. Malara, M. A. Shahzad, L. Thomas, M. Vanden Bemden, C. Vander Velde, P. Vanlaer

Ghent University, Ghent, Belgium
M. De Coen, D. Dobur, G. Gokbulut, J. Knolle, L. Lambrecht, D. Marckx, K. Skovpen, N. Van Den Bossche, J. van der Linden, J. Vandenbroeck, L. Wezenbeek

Université Catholique de Louvain, Louvain-la-Neuve, Belgium
S. Bein, A. Benecke, A. Bethani, G. Bruno, C. Caputo, J. De Favereau De Jeneret, C. Delaere, I. S. Donertas, A. Giammanco, A. O. Guzel, S a. Jain, V. Lemaitre, J. Lidrych, P. Mastrapasqua, T. T. Tran, S. Turkcapar

Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
G. A. Alves, E. Coelho, G. Correia Silva, C. Hensel, T. Menezes De Oliveira, C. Mora Herrera[4], P. Rebello Teles, M. Soeiro, E. J. Tonelli Manganote[5], A. Vilela Pereira[4]

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
W. L. Aldá Júnior, M. Barroso Ferreira Filho, H. Brandao Malbouisson, W. Carvalho, J. Chinellato[6], E. M. Da Costa, G. G. Da Silveira[7], D. De Jesus Damiao, S. Fonseca De Souza, R. Gomes De Souza, T. Laux Kuhn[7], M. Macedo, J. Martins, K. Mota Amarilo, L. Mundim, H. Nogima, J. P. Pinheiro, A. Santoro, A. Sznajder, M. Thiel

Universidade Estadual Paulista, Universidade Federal do ABC, São Paulo, Brazil
C. A. Bernardes[7], L. Calligaris, T. R. Fernandez Perez Tomei, E. M. Gregores, I. Maietto Silverio, P. G. Mercadante, S. F. Novaes, B. Orzari, Sandra S. Padula, V. Scheurer

Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria
A. Aleksandrov, G. Antchev, R. Hadjiiska, P. Iaydjiev, M. Misheva, M. Shopova, G. Sultanov

University of Sofia, Sofia, Bulgaria
A. Dimitrov, L. Litov, B. Pavlov, P. Petkov, A. Petrov, E. Shumka

Instituto De Alta Investigación, Universidad de Tarapacá, Casilla 7 D, Arica, Chile
S. Keshri, D. Laroze, S. Thakur

Beihang University, Beijing, China
T. Cheng, T. Javaid, L. Yuan

Department of Physics, Tsinghua University, Beijing, China
Z. Hu, Z. Liang, J. Liu

Institute of High Energy Physics, Beijing, China
G. M. Chen[8], H. S. Chen[8], M. Chen[8], F. Iemmi, C. H. Jiang, A. Kapoor[9], H. Liao, Z.-A. Liu[10], R. Sharma[11], J. N. Song[10], J. Tao, C. Wang[8], J. Wang, Z. Wang[8], H. Zhang, J. Zhao

State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
A. Agapitos, Y. Ban, A. Carvalho Antunes De Oliveira, S. Deng, B. Guo, C. Jiang, A. Levin, C. Li, Q. Li, Y. Mao, S. Qian, S. J. Qian, X. Qin, X. Sun, D. Wang, H. Yang, Y. Zhao, C. Zhou

Guangdong Provincial Key Laboratory of Nuclear Science and Guangdong-Hong Kong Joint Laboratory of Quantum Matter, South China Normal University, Guangzhou, China
S. Yang

Sun Yat-Sen University, Guangzhou, China
Z. You

University of Science and Technology of China, Hefei, China
K. Jaffel, N. Lu

Nanjing Normal University, Nanjing, China
G. Bauer[12], B. Li[13], H. Wang, K. Yi[14], J. Zhang

Institute of Modern Physics and Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) - Fudan University, Shanghai, China
Y. Li

Zhejiang University, Hangzhou, Zhejiang, China
Z. Lin, C. Lu, M. Xiao

Universidad de Los Andes, Bogotá, Colombia
C. Avila, D. A. Barbosa Trujillo, A. Cabrera, C. Florez, J. Fraga, J. A. Reyes Vega

Universidad de Antioquia, Medellin, Colombia
J. Jaramillo, C. Rendón, M. Rodriguez, A. A. Ruales Barbosa, J. D. Ruiz Alvarez

Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia
D. Giljanovic, N. Godinovic, D. Lelas, A. Sculac

Faculty of Science, University of Split, Split, Croatia
M. Kovac, A. Petkovic, T. Sculac

Institute Rudjer Boskovic, Zagreb, Croatia
P. Bargassa, V. Brigljevic, B. K. Chitroda, D. Ferencek, K. Jakovcic, A. Starodumov[15], T. Susa

University of Cyprus, Nicosia, Cyprus
A. Attikis, K. Christoforou, A. Hadjiagapiou, C. Leonidou, J. Mousa, C. Nicolaou, L. Paizanos, F. Ptochos, P. A. Razis, H. Rykaczewski, H. Saka, A. Stepennov

Charles University, Prague, Czech Republic
M. Finger, M. Finger Jr., A. Kveton

Escuela Politecnica Nacional, Quito, Ecuador
E. Ayala

Universidad San Francisco de Quito, Quito, Ecuador
E. Carrera Jarrin

Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
H. Abdalla[16], Y. Assran[17,18], B. El-mahdy

Center for High Energy Physics (CHEP-FU), Fayoum University, El-Fayoum, Egypt
M. Abdullah Al-Mashad, M. A. Mahmoud

National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
K. Ehataht, M. Kadastik, T. Lange, C. Nielsen, J. Pata, M. Raidal, L. Tani, C. Veelken

Department of Physics, University of Helsinki, Helsinki, Finland
K. Osterberg, M. Voutilainen

Helsinki Institute of Physics, Helsinki, Finland
N. Bin Norjoharuddeen, E. Brücken, F. Garcia, P. Inkaew, K. T. S. Kallonen, T. Lampén, K. Lassila-Perini, S. Lehti, T. Lindén, M. Myllymäki, M. m. Rantanen, J. Tuominiemi

Lappeenranta-Lahti University of Technology, Lappeenranta, Finland
H. Kirschenmann, P. Luukka, H. Petrow

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
M. Besancon, F. Couderc, M. Dejardin, D. Denegri, J. L. Faure, F. Ferri, S. Ganjour, P. Gras, G. Hamel de Monchenault, M. Kumar, V. Lohezic, J. Malcles, F. Orlandi, L. Portales, A. Rosowsky, M. Ö. Sahin, A. Savoy-Navarro[19], P. Simkina, M. Titov, M. Tornago

Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
F. Beaudette, G. Boldrini, P. Busson, A. Cappati, C. Charlot, M. Chiusi, T. D. Cuisset, F. Damas, O. Davignon, A. De Wit, I. T. Ehle, B. A. Fontana Santos Alves, S. Ghosh, A. Gilbert, R. Granier de Cassagnac, A. Hakimi, B. Harikrishnan, L. Kalipoliti, G. Liu, M. Nguyen, S. Obraztsov, C. Ochando, R. Salerno, J. B. Sauvan, Y. Sirois, G. Sokmen, L. Urda Gómez, E. Vernazza, A. Zabi, A. Zghiche

CNRS, IPHC UMR 7178, Université de Strasbourg, Strasbourg, France
J.-L. Agram[20], J. Andrea, D. Apparu, D. Bloch, J.-M. Brom, E. C. Chabert, C. Collard, S. Falke, U. Goerlach, R. Haeberle, A.-C. Le Bihan, M. Meena, O. Poncet, G. Saha, M. A. Sessini, P. Van Hove, P. Vaucelle

Centre de Calcul de l'Institut National de Physique Nucleaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France
A. Di Florio

Institut de Physique des 2 Infinis de Lyon (IP2I), Villeurbanne, France
D. Amram, S. Beauceron, B. Blancon, G. Boudoul, N. Chanon, D. Contardo, P. Depasse, C. Dozen[21], H. El Mamouni, J. Fay, S. Gascon, M. Gouzevitch, C. Greenberg, G. Grenier, B. Ille, E. Jourd'huy, I. B. Laktineh, M. Lethuillier, L. Mirabito, S. Perries, A. Purohit, M. Vander Donckt, P. Verdier, J. Xiao

Georgian Technical University, Tbilisi, Georgia
G. Adamov, I. Lomidze, Z. Tsamalaidze[22]

RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany
V. Botta, S. Consuegra Rodríguez, L. Feld, K. Klein, M. Lipinski, D. Meuser, A. Pauls, D. Pérez Adán, N. Röwert, M. Teroerde

RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
S. Diekmann, A. Dodonova, N. Eich, D. Eliseev, F. Engelke, J. Erdmann, M. Erdmann, B. Fischer, T. Hebbeker, K. Hoepfner, F. Ivone, A. Jung, M. y. Lee, F. Mausolf, M. Merschmeyer, A. Meyer, S. Mukherjee, D. Noll, F. Nowotny, A. Pozdnyakov, Y. Rath, W. Redjeb, F. Rehm, H. Reithler, V. Sarkisovi, A. Schmidt, C. Seth, A. Sharma, J. L. Spah, F. Torres Da Silva De Araujo[23], S. Wiedenbeck, S. Zaleski

RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany
C. Dziwok, G. Flügge, T. Kress, A. Nowack, O. Pooth, A. Stahl, T. Ziemons, A. Zotz

Deutsches Elektronen-Synchrotron, Hamburg, Germany
H. Aarup Petersen, M. Aldaya Martin, J. Alimena, S. Amoroso, Y. An, J. Bach, S. Baxter, M. Bayatmakou, H. Becerril Gonzalez, O. Behnke, A. Belvedere, F. Blekman[24], K. Borras[25],

A. Campbell, A. Cardini, F. Colombina, M. De Silva, G. Eckerlin, D. Eckstein, L. I. Estevez Banos, E. Gallo[24], A. Geiser, V. Guglielmi, M. Guthoff, A. Hinzmann, L. Jeppe, B. Kaech, M. Kasemann, C. Kleinwort, R. Kogler, M. Komm, D. Krücker, W. Lange, D. Leyva Pernia, K. Lipka[26], W. Lohmann[27], F. Lorkowski, R. Mankel, I.-A. Melzer-Pellmann, M. Mendizabal Morentin, A. B. Meyer, G. Milella, K. Moral Figueroa, A. Mussgiller, L. P. Nair, J. Niedziela, A. Nürnberg, J. Park, E. Ranken, A. Raspereza, D. Rastorguev, J. Rübenach, L. Rygaard, M. Scham[28,25], S. Schnake[25], P. Schütze, C. Schwanenberger[24], D. Selivanova, K. Sharko, M. Shchedrolosiev, D. Stafford, F. Vazzoler, A. Ventura Barroso, R. Walsh, D. Wang, Q. Wang, K. Wichmann, L. Wiens[25], C. Wissing, Y. Yang, S. Zakharov, A. Zimermmane Castro Santos

University of Hamburg, Hamburg, Germany
A. Albrecht, S. Albrecht, M. Antonello, S. Bollweg, M. Bonanomi, P. Connor, K. El Morabit, Y. Fischer, E. Garutti, A. Grohsjean, J. Haller, D. Hundhausen, H. R. Jabusch, G. Kasieczka, P. Keicher, R. Klanner, W. Korcari, T. Kramer, C. c. Kuo, V. Kutzner, F. Labe, J. Lange, A. Lobanov, C. Matthies, L. Moureaux, M. Mrowietz, A. Nigamova, Y. Nissan, A. Paasch, K. J. Pena Rodriguez, T. Quadfasel, B. Raciti, M. Rieger, D. Savoiu, J. Schindler, P. Schleper, M. Schröder, J. Schwandt, M. Sommerhalder, H. Stadie, G. Steinbrück, A. Tews, B. Wiederspan, M. Wolf

Karlsruher Institut fuer Technologie, Karlsruhe, Germany
S. Brommer, E. Butz, T. Chwalek, A. Dierlamm, G. G. Dincer, U. Elicabuk, N. Faltermann, M. Giffels, A. Gottmann, F. Hartmann[29], R. Hofsaess, M. Horzela, U. Husemann, J. Kieseler, M. Klute, O. Lavoryk, J. M. Lawhorn, M. Link, A. Lintuluoto, S. Maier, S. Mitra, A. A. Monsch, M. Mormile, T h. Müller, M. Neukum, M. Oh, E. Pfeffer, M. Presilla, G. Quast, K. Rabbertz, B. Regnery, R. Schmieder, N. Shadskiy, I. Shvetsov, H. J. Simonis, L. Sowa, L. Stockmeier, K. Tauqeer, M. Toms, B. Topko, N. Trevisani, T. Voigtländer, R. F. Von Cube, J. Von Den Driesch, M. Wassmer, S. Wieland, F. Wittig, R. Wolf, X. Zuo

Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
G. Anagnostou, G. Daskalakis, A. Kyriakis, A. Papadopoulos[29], A. Stakia

National and Kapodistrian University of Athens, Athens, Greece
G. Melachroinos, Z. Painesis, I. Paraskevas, N. Saoulidou, K. Theofilatos, E. Tziaferi, K. Vellidis, I. Zisopoulos

National Technical University of Athens, Athens, Greece
G. Bakas, T. Chatzistavrou, G. Karapostoli, K. Kousouris, I. Papakrivopoulos, E. Siamarkou, G. Tsipolitis, A. Zacharopoulou

University of Ioánnina, Ioánnina, Greece
I. Bestintzanos, I. Evangelou, C. Foudas, C. Kamtsikis, P. Katsoulis, P. Kokkas, P. G. Kosmoglou Kioseoglou, N. Manthos, I. Papadopoulos, J. Strologas

HUN-REN Wigner Research Centre for Physics, Budapest, Hungary
C. Hajdu, D. Horvath[30,31], K. Márton, A. J. Rádl[32], F. Sikler, V. Veszpremi

MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
M. Csanád, K. Farkas, A. Fehérkuti[33], M. M. A. Gadallah[34], Á. Kadlecsik, P. Major, G. Pásztor, G. I. Veres

Faculty of Informatics, University of Debrecen, Debrecen, Hungary
B. Ujvari, G. Zilizi

HUN-REN ATOMKI-Institute of Nuclear Research, Debrecen, Hungary
G. Bencze, S. Czellar, J. Molnar, Z. Szillasi

Karoly Robert Campus, MATE Institute of Technology, Gyongyos, Hungary
T. Csorgo[33], F. Nemes[33], T. Novak

Panjab University, Chandigarh, India
S. Bansal, S. B. Beri, V. Bhatnagar, G. Chaudhary, S. Chauhan, N. Dhingra[35], A. Kaur, A. Kaur, H. Kaur, M. Kaur, S. Kumar, T. Sheokand, J. B. Singh, A. Singla

University of Delhi, Delhi, India
A. Bhardwaj, A. Chhetri, B. C. Choudhary, A. Kumar, A. Kumar, M. Naimuddin, K. Ranjan, M. K. Saini, S. Saumya

Saha Institute of Nuclear Physics, HBNI, Kolkata, India
S. Baradia, S. Barman[36], S. Bhattacharya, S. Das Gupta, S. Dutta, S. Dutta, S. Sarkar

Indian Institute of Technology Madras, Madras, India
M. M. Ameen, P. K. Behera, S. C. Behera, S. Chatterjee, G. Dash, P. Jana, P. Kalbhor, S. Kamble, J. R. Komaragiri[37], D. Kumar[37], T. Mishra, B. Parida[38], P. R. Pujahari, N. R. Saha, A. Sharma, A. K. Sikdar, R. K. Singh, P. Verma, S. Verma, A. Vijay

Tata Institute of Fundamental Research-A, Mumbai, India
S. Dugad, G. B. Mohanty, M. Shelake, P. Suryadevara

Tata Institute of Fundamental Research-B, Mumbai, India
A. Bala, S. Banerjee, S. Bhowmik, R. M. Chatterjee, M. Guchait, S h. Jain, A. Jaiswal, B. M. Joshi, S. Kumar, G. Majumder, K. Mazumdar, S. Parolia, A. Thachayath

National Institute of Science Education and Research, An OCC of Homi Bhabha National Institute, Bhubaneswar, Odisha, India
S. Bahinipati[39], C. Kar, D. Maity[40], P. Mal, K. Naskar[40], A. Nayak[40], S. Nayak, K. Pal, P. Sadangi, S. K. Swain, S. Varghese[40], D. Vats[40]

Indian Institute of Science Education and Research (IISER), Pune, India
S. Acharya[41], A. Alpana, S. Dube, B. Gomber[41], P. Hazarika, B. Kansal, A. Laha, B. Sahu[41], S. Sharma, K. Y. Vaish

Isfahan University of Technology, Isfahan, Iran
H. Bakhshiansohi[42], A. Jafari[43], M. Zeinali[44]

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
S. Bashiri, S. Chenarani[45], S. M. Etesami, Y. Hosseini, M. Khakzad, E. Khazaie, M. Mohammadi Najafabadi, S. Tizchang[46]

University College Dublin, Dublin, Ireland
M. Felcini, M. Grunewald

INFN Sezione di Bari[a], Università di Bari[b], Politecnico di Bari[c], Bari, Italy
M. Abbrescia[a,b], A. Colaleo[a,b], D. Creanza[a,c], B. D'Anzi[a,b], N. De Filippis[a,c], M. De Palma[a,b], W. Elmetenawee[a,b,47], N. Ferrara[a,b], L. Fiore[a], G. Iaselli[a,c], L. Longo[a], M. Louka[a,b], G. Maggi[a,c], M. Maggi[a], I. Margjeka[a], V. Mastrapasqua[a,b], S. My[a,b], S. Nuzzo[a,b], A. Pellecchia[a,b], A. Pompili[a,b], G. Pugliese[a,c], R. Radogna[a,b], D. Ramos[a], A. Ranieri[a], L. Silvestris[a], F. M. Simone[a,c], Ü. Sözbilir[a], A. Stamerra[a,b], D. Troiano[a,b], R. Venditti[a,b], P. Verwilligen[a], A. Zaza[a,b]

INFN Sezione di Bologna[a], Università di Bologna[b], Bologna, Italy
G. Abbiendi[a], C. Battilana[a,b], D. Bonacorsi[a,b], P. Capiluppi[a,b], A. Castro[a,b], F. R. Cavallo[a], M. Cuffiani[a,b], G. M. Dallavalle[a], T. Diotalevi[a,b], F. Fabbri[a], A. Fanfani[a,b], D. Fasanella[a], P. Giacomelli[a], L. Giommi[a,b], C. Grandi[a], L. Guiducci[a,b], S. Lo Meo[a,48], M. Lorusso[a,b], L. Lunerti[a], S. Marcellini[a], G. Masetti[a], F. L. Navarria[a,b], G. Paggi[a,b], A. Perrotta[a], F. Primavera[a,b], A. M. Rossi[a,b], S. Rossi Tisbeni[a,b], T. Rovelli[a,b], G. P. Siroli[a,b]

INFN Sezione di Catania[a], Università di Catania[b], Catania, Italy
S. Costa[a,b,49], A. Di Mattia[a], A. Lapertosa[a], R. Potenza[a,b], A. Tricomi[a,b,49]

INFN Sezione di Firenze[a], Università di Firenze[b], Firenze, Italy
P. Assiouras[a], G. Barbagli[a], G. Bardelli[a,b], B. Camaiani[a,b], A. Cassese[a], R. Ceccarelli[a], V. Ciulli[a,b], C. Civinini[a], R. D'Alessandro[a,b], E. Focardi[a,b], T. Kello[a], G. Latino[a,b], P. Lenzi[a,b], M. Lizzo[a], M. Meschini[a], S. Paoletti[a], A. Papanastassiou[a,b], G. Sguazzoni[a], L. Viliani[a]

INFN Laboratori Nazionali di Frascati, Frascati, Italy
L. Benussi, S. Bianco, S. Meola[50], D. Piccolo

INFN Sezione di Genova[a], Università di Genova[b], Genoa, Italy
M. Alves Gallo Pereira[a], F. Ferro[a], E. Robutti[a], S. Tosi[a,b]

INFN Sezione di Milano-Bicocca[a], Università di Milano-Bicocca[b], Milan, Italy
A. Benaglia[a], F. Brivio[a], F. Cetorelli[a,b], F. De Guio[a,b], M. E. Dinardo[a,b], P. Dini[a], S. Gennai[a], R. Gerosa[a,b], A. Ghezzi[a,b], P. Govoni[a,b], L. Guzzi[a], G. Lavizzari[a,b], M. T. Lucchini[a,b], M. Malberti[a], S. Malvezzi[a], A. Massironi[a], D. Menasce[a], L. Moroni[a], M. Paganoni[a,b], S. Palluotto[a,b], D. Pedrini[a], A. Perego[a,b], B. S. Pinolini[a], G. Pizzati[a,b], S. Ragazzi[a,b], T. Tabarelli de Fatis[a,b]

INFN Sezione di Napoli[a], Università di Napoli 'Federico II'[b], Napoli, Italy; Università della Basilicata[c], Potenza, Italy; Scuola Superiore Meridionale (SSM)[d], Naples, Italy
S. Buontempo[a], A. Cagnotta[a,b], F. Carnevali[a,b], N. Cavallo[a,c], F. Fabozzi[a,c], A. O. M. Iorio[a,b], L. Lista[a,b,51], P. Paolucci[a,29], B. Rossi[a]

INFN Sezione di Padova[a], Università di Padova[b], Padova, Italy; Università di Trento[c], Trento, Italy
R. Ardino[a], P. Azzi[a], N. Bacchetta[a,52], D. Bisello[a,b], P. Bortignon[a], G. Bortolato[a,b], A. C. M. Bulla[a], P. Checchia[a], T. Dorigo[a,53], F. Gasparini[a,b], U. Gasparini[a,b], S. Giorgetti[a], M. Gulmini[a,54], E. Lusiani[a], M. Margoni[a,b], A. T. Meneguzzo[a,b], M. Migliorini[a,b], J. Pazzini[a,b], P. Ronchese[a,b], R. Rossin[a,b], F. Simonetto[a,b], M. Tosi[a,b], A. Triossi[a,b], S. Ventura[a], M. Zanetti[a,b], P. Zotto[a,b], A. Zucchetta[a,b], G. Zumerle[a,b]

INFN Sezione di Pavia[a], Università di Pavia[b], Pavia, Italy
A. Braghieri[a], S. Calzaferri[a], D. Fiorina[a], P. Montagna[a,b], V. Re[a], C. Riccardi[a,b], P. Salvini[a], I. Vai[a,b], P. Vitulo[a,b]

INFN Sezione di Perugia[a], Università di Perugia[b], Perugia, Italy
S. Ajmal[a,b], M. E. Ascioti[a,b], G. M. Bilei[a], C. Carrivale[a,b], D. Ciangottini[a,b], L. Fanò[a,b], V. Mariani[a,b], M. Menichelli[a], F. Moscatelli[a,55], A. Rossi[a,b], A. Santocchia[a,b], D. Spiga[a], T. Tedeschi[a,b]

INFN Sezione di Pisa[a], Università di Pisa[b], Scuola Normale Superiore di Pisa[c], Pisa, Italy; Università di Siena[d], Siena, Italy
C. Aimè[a], C. A. Alexe[a,c], P. Asenov[a,b], P. Azzurri[a], G. Bagliesi[a], R. Bhattacharya[a], L. Bianchini[a,b], T. Boccali[a], E. Bossini[a], D. Bruschini[a,c], R. Castaldi[a], M. A. Ciocci[a,b], M. Cipriani[a,b], V. D'Amante[a,d], R. Dell'Orso[a], S. Donato[a], A. Giassi[a], F. Ligabue[a,c], A. C. Marini[a], D. Matos Figueiredo[a], A. Messineo[a,b], S. Mishra[a], V. K. Muraleedharan Nair Bindhu[a,b,40], M. Musich[a,b], S. Nandan[a], F. Palla[a], A. Rizzi[a,b], G. Rolandi[a,c], S. Roy Chowdhury[a], T. Sarkar[a], A. Scribano[a], P. Spagnolo[a], R. Tenchini[a], G. Tonelli[a,b], N. Turini[a,d], F. Vaselli[a,c], A. Venturi[a], P. G. Verdini[a]

INFN Sezione di Roma[a], Sapienza Università di Roma[b], Rome, Italy
P. Barria[a], C. Basile[a,b], F. Cavallari[a], L. Cunqueiro Mendez[a,b], D. Del Re[a,b], E. Di Marco[a,b], M. Diemoz[a], F. Errico[a,b], R. Gargiulo[a,b], E. Longo[a,b], L. Martikainen[a,b], J. Mijuskovic[a,b], G. Organtini[a,b], F. Pandolfi[a], R. Paramatti[a,b], C. Quaranta[a,b], S. Rahatlou[a,b], C. Rovelli[a], F. Santanastasio[a,b], L. Soffi[a], V. Vladimirov[a,b]

INFN Sezione di Torino[a], Università di Torino[b], Torino, Italy; Università del Piemonte Orientale[c], Novara, Italy
N. Amapane[a,b], R. Arcidiacono[a,c], S. Argiro[a,b], M. Arneodo[a,c], N. Bartosik[a], R. Bellan[a,b], C. Biino[a], C. Borca[a,b], N. Cartiglia[a], M. Costa[a,b], R. Covarelli[a,b], N. Demaria[a], L. Finco[a], M. Grippo[a,b],

B. Kiani[a,b], F. Legger[a], F. Luongo[a,b], C. Mariotti[a], L. Markovic[a,b], S. Maselli[a], A. Mecca[a,b], L. Menzio[a,b], P. Meridiani[a], E. Migliore[a,b], M. Monteno[a], R. Mulargia[a], M. M. Obertino[a,b], G. Ortona[a], L. Pacher[a,b], N. Pastrone[a], M. Pelliccioni[a], M. Ruspa[a,c], F. Siviero[a,b], V. Sola[a,b], A. Solano[a,b], A. Staiano[a], C. Tarricone[a,b], D. Trocino[a], G. Umoret[a,b], R. White[a,b]

INFN Sezione di Trieste[a], Università di Trieste[b], Trieste, Italy
J. Babbar[a,b], S. Belforte[a], V. Candelise[a,b], M. Casarsa[a], F. Cossutti[a], K. De Leo[a], G. Della Ricca[a,b]

Kyungpook National University, Daegu, Korea
S. Dogra, J. Hong, J. Kim, D. Lee, H. Lee, S. W. Lee, C. S. Moon, Y. D. Oh, M. S. Ryu, S. Sekmen, B. Tae, Y. C. Yang

Department of Mathematics and Physics - GWNU, Gangneung, Korea
M. S. Kim

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea
G. Bak, P. Gwak, H. Kim, D. H. Moon

Hanyang University, Seoul, Korea
E. Asilar, J. Choi[56], D. Kim, T. J. Kim, J. A. Merlin, Y. Ryou

Korea University, Seoul, Korea
S. Choi, S. Han, B. Hong, K. Lee, K. S. Lee, S. Lee, J. Yoo

Department of Physics, Kyung Hee University, Seoul, Korea
J. Goh, S. Yang

Sejong University, Seoul, Korea
Y. Kang, H. S. Kim, Y. Kim, S. Lee

Seoul National University, Seoul, Korea
J. Almond, J. H. Bhyun, J. Choi, J. Choi, W. Jun, J. Kim, Y. W. Kim, S. Ko, H. Lee, J. Lee, J. Lee, B. H. Oh, S. B. Oh, H. Seo, U. K. Yang, I. Yoon

University of Seoul, Seoul, Korea
W. Jang, D. Y. Kang, S. Kim, B. Ko, J. S. H. Lee, Y. Lee, I. C. Park, Y. Roh, I. J. Watson

Department of Physics, Yonsei University, Seoul, Korea
S. Ha, K. Hwang, B. Kim, H. D. Yoo

Sungkyunkwan University, Suwon, Korea
M. Choi, M. R. Kim, H. Lee, Y. Lee, I. Yu

College of Engineering and Technology, American University of the Middle East (AUM), Dasman, Kuwait
T. Beyrouthy, Y. Gharbia

Kuwait University-College of Science-Department of Physics, Safat, Kuwait
F. Alazemi

Riga Technical University, Riga, Latvia
K. Dreimanis, A. Gaile, C. Munoz Diaz, D. Osite, G. Pikurs, A. Potrebko, M. Seidel, D. Sidiropoulos Kontos

University of Latvia (LU), Riga, Latvia
N. R. Strautnieks

Vilnius University, Vilnius, Lithuania
M. Ambrozas, A. Juodagalvis, A. Rinkevicius, G. Tamulaitis

National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia
I. Yusuff[57], Z. Zolkapli

Universidad de Sonora (UNISON), Hermosillo, Mexico
J. F. Benitez, A. Castaneda Hernandez, H. A. Encinas Acosta, L. G. Gallegos Marí nez, M. León Coello, J. A. Murillo Quijada, A. Sehrawat, L. Valencia Palomo

Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico
G. Ayala, H. Castilla-Valdez, H. Crotte Ledesma, E. De La Cruz-Burelo, I. Heredia-De La Cruz[58], R. Lopez-Fernandez, J. Mejia Guisao, C. A. Mondragon Herrera, A. Sánchez Hernández

Universidad Iberoamericana, Mexico City, Mexico
C. Oropeza Barrera, D. L. Ramirez Guadarrama, M. Ramírez García

Benemerita Universidad Autonoma de Puebla, Puebla, Mexico
I. Bautista, F. E. Neri Huerta, I. Pedraza, H. A. Salazar Ibarguen, C. Uribe Estrada

University of Montenegro, Podgorica, Montenegro
I. Bubanja, N. Raicevic

University of Canterbury, Christchurch, New Zealand
P. H. Butler

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
A. Ahmad, M. I. Asghar, A. Awais, M. I. M. Awan, H. R. Hoorani, W. A. Khan

AGH University of Krakow, Kraków, Poland
V. Avati, A. Bellora, L. Forthomme, L. Grzanka, M. Malawski, K. Piotrzkowski

National Centre for Nuclear Research, Swierk, Poland
H. Bialkowska, M. Bluj, M. Górski, M. Kazana, M. Szleper, P. Zalewski

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland
K. Bunkowski, K. Doroba, A. Kalinowski, M. Konecki, J. Krolikowski, A. Muhammad

Warsaw University of Technology, Warsaw, Poland
P. Fokow, K. Pozniak, W. Zabolotny

Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal
M. Araujo, D. Bastos, C. Beirão Da Cruz E Silva, A. Boletti, M. Bozzo, T. Camporesi, G. Da Molin, P. Faccioli, M. Gallinaro, J. Hollar, N. Leonardo, G. B. Marozzo, A. Petrilli, M. Pisano, J. Seixas, J. Varela, J. W. Wulff

Faculty of Physics, University of Belgrade, Belgrade, Serbia
P. Adzic, P. Milenovic

VINCA Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia
D. Devetak, M. Dordevic, J. Milosevic, L. Nadderd, V. Rekovic, M. Stojanovic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
J. Alcaraz Maestre, Cristina F. Bedoya, J. A. Brochero Cifuentes, Oliver M. Carretero, M. Cepeda, M. Cerrada, N. Colino, B. De La Cruz, A. Delgado Peris, A. Escalante Del Valle, D. Fernández Del Val, J. P. Fernández Ramos, J. Flix, M. C. Fouz, O. Gonzalez Lopez, S. Goy Lopez, J. M. Hernandez, M. I. Josa, J. Llorente Merino, C. Martin Perez, E. Martin Viscasillas, D. Moran, C. M. Morcillo Perez, Á. Navarro Tobar, C. Perez Dengra, A. Pérez-Calero Yzquierdo, J. Puerta Pelayo, I. Redondo, J. Sastre, J. Vazquez Escobar

Universidad Autónoma de Madrid, Madrid, Spain
J. F. de Trocóniz

Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias (ICTEA), Universidad de Oviedo, Oviedo, Spain
B. Alvarez Gonzalez, J. Cuevas, J. Fernandez Menendez, S. Folgueras, I. Gonzalez Caballero, P. Leguina,

E. Palencia Cortezon, J. Prado Pico, V. Rodríguez Bouza, A. Soto Rodríguez, A. Trapote, C. Vico Villalba, P. Vischia

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain
S. Blanco Fernández, I. J. Cabrillo, A. Calderon, J. Duarte Campderros, M. Fernandez, G. Gomez, C. Lasaosa García, R. Lopez Ruiz, C. Martinez Rivero, P. Martinez Ruiz del Arbol, F. Matorras, P. Matorras Cuevas, E. Navarrete Ramos, J. Piedra Gomez, L. Scodellaro, I. Vila, J. M. Vizan Garcia

University of Colombo, Colombo, Sri Lanka
B. Kailasapathy[59], D. D. C. Wickramarathna

Department of Physics, University of Ruhuna, Matara, Sri Lanka
W. G. D. Dharmaratna[60], K. Liyanage, N. Perera

CERN, European Organization for Nuclear Research, Geneva, Switzerland
D. Abbaneo, C. Amendola, E. Auffray, G. Auzinger, J. Baechler, D. Barney, A. Bermúdez Martínez, M. Bianco, A. A. Bin Anuar, A. Bocci, L. Borgonovi, C. Botta, A. Bragagnolo, E. Brondolin, C. E. Brown, C. Caillol, G. Cerminara, N. Chernyavskaya, D. d'Enterria, A. Dabrowski, A. David, A. De Roeck, M. M. Defranchis, M. Deile, M. Dobson, G. Franzoni, W. Funk, S. Giani, D. Gigi, K. Gill, F. Glege, M. Glowacki, J. Hegeman, J. K. Heikkilä, B. Huber, V. Innocente, T. James, P. Janot, O. Kaluzinska, O. Karacheban[27], G. Karathanasis, S. Laurila, P. Lecoq, E. Leutgeb, C. Lourenço, M. Magherini, L. Malgeri, M. Mannelli, M. Matthewman, A. Mehta, F. Meijers, S. Mersi, E. Meschi, V. Milosevic, F. Monti, F. Moortgat, M. Mulders, I. Neutelings, S. Orfanelli, F. Pantaleo, G. Petrucciani, A. Pfeiffer, M. Pierini, H. Qu, D. Rabady, B. Ribeiro Lopes, F. Riti, M. Rovere, H. Sakulin, R. Salvatico, S. Sanchez Cruz, S. Scarfi, C. Schwick, M. Selvaggi, A. Sharma, K. Shchelina, P. Silva, P. Sphicas[61], A. G. Stahl Leiton, A. Steen, S. Summers, D. Treille, P. Tropea, D. Walter, J. Wanczyk[62], J. Wang, S. Wuchterl, P. Zehetner, P. Zejdl, W. D. Zeuner

PSI Center for Neutron and Muon Sciences, Villigen, Switzerland
T. Bevilacqua[63], L. Caminada[63], A. Ebrahimi, W. Erdmann, R. Horisberger, Q. Ingram, H. C. Kaestli, D. Kotlinski, C. Lange, M. Missiroli[63], L. Noehte[63], T. Rohe, A. Samalan

ETH Zurich-Institute for Particle Physics and Astrophysics (IPA), Zurich, Switzerland
T. K. Aarrestad, M. Backhaus, G. Bonomelli, A. Calandri, C. Cazzaniga, K. Datta, P. De Bryas Dexmiers D'archiac[62], A. De Cosa, G. Dissertori, M. Dittmar, M. Donegà, F. Eble, M. Galli, K. Gedia, F. Glessgen, C. Grab, N. Härringer, T. G. Harte, D. Hits, W. Lustermann, A.-M. Lyon, R. A. Manzoni, M. Marchegiani, L. Marchese, A. Mascellani[62], F. Nessi-Tedaldi, F. Pauss, V. Perovic, S. Pigazzini, B. Ristic, R. Seidita, J. Steggemann[62], A. Tarabini, D. Valsecchi, R. Wallny

Universität Zürich, Zurich, Switzerland
C. Amsler[64], P. Bärtschi, M. F. Canelli, K. Cormier, M. Huwiler, W. Jin, A. Jofrehei, B. Kilminster, S. Leontsinis, S. P. Liechti, A. Macchiolo, P. Meiring, F. Meng, J. Motta, A. Reimers, P. Robmann, M. Senger, E. Shokr, F. Stäger, R. Tramontano

National Central University, Chung-Li, Taiwan
C. Adloff[65], D. Bhowmik, C. M. Kuo, W. Lin, P. K. Rout, P. C. Tiwari[37]

National Taiwan University (NTU), Taipei, Taiwan
L. Ceard, K. F. Chen, Z. g. Chen, A. De Iorio, W.-S. Hou, T. h. Hsu, Y. w. Kao, S. Karmakar, G. Kole, Y. y. Li, R.-S. Lu, E. Paganis, X. f. Su, J. Thomas-Wilsker, L. s. Tsai, D. Tsionou, H. y. Wu, E. Yazgan

High Energy Physics Research Unit, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand
C. Asawatangtrakuldee, N. Srimanobhas, V. Wachirapusitanand

Tunis El Manar University, Tunis, Tunisia
Y. Maghrbi

Physics Department, Science and Art Faculty, Çukurova University, Adana, Turkey
D. Agyel, F. Boran, F. Dolek, I. Dumanoglu[66], E. Eskut, Y. Guler[67], E. Gurpinar Guler[67], C. Isik, O. Kara, A. Kayis Topaksu, Y. Komurcu, G. Onengut, K. Ozdemir[68], A. Polatoz, B. Tali[69], U. G. Tok, E. Uslan, I. S. Zorbakir

Physics Department, Middle East Technical University, Ankara, Turkey
M. Yalvac[70]

Bogazici University, Istanbul, Turkey
B. Akgun, I. O. Atakisi, E. Gülmez, M. Kaya[71], O. Kaya[72], S. Tekten[73]

Istanbul Technical University, Istanbul, Turkey
A. Cakir, K. Cankocak[66,74], S. Sen[75]

Istanbul University, Istanbul, Turkey
O. Aydilek[76], B. Hacisahinoglu, I. Hos[77], B. Kaynak, S. Ozkorucuklu, O. Potok, H. Sert, C. Simsek, C. Zorbilmez

Yildiz Technical University, Istanbul, Turkey
S. Cerci, B. Isildak[78], D. Sunar Cerci, T. Yetkin

Institute for Scintillation Materials of National Academy of Science of Ukraine, Kharkiv, Ukraine
A. Boyaryntsev, B. Grynyov

National Science Centre, Kharkiv Institute of Physics and Technology, Kharkiv, Ukraine
L. Levchuk

University of Bristol, Bristol, UK
D. Anthony, J. J. Brooke, A. Bundock, F. Bury, E. Clement, D. Cussans, H. Flacher, J. Goldstein, H. F. Heath, M.-L. Holmberg, L. Kreczko, S. Paramesvaran, L. Robertshaw, V. J. Smith, K. Walkingshaw Pass

Rutherford Appleton Laboratory, Didcot, UK
A. H. Ball, K. W. Bell, A. Belyaev[79], C. Brew, R. M. Brown, D. J. A. Cockerill, C. Cooke, A. Elliot, K. V. Ellis, K. Harder, S. Harper, J. Linacre, K. Manolopoulos, D. M. Newbold, E. Olaiya, D. Petyt, T. Reis, A. R. Sahasransu, G. Salvi, T. Schuh, C. H. Shepherd-Themistocleous, I. R. Tomalin, K. C. Whalen, T. Williams

Imperial College, London, UK
I. Andreou, R. Bainbridge, P. Bloch, O. Buchmuller, C. A. Carrillo Montoya, G. S. Chahal[80], D. Colling, J. S. Dancu, I. Das, P. Dauncey, G. Davies, M. Della Negra, S. Fayer, G. Fedi, G. Hall, A. Howard, G. Iles, C. R. Knight, P. Krueper, J. Langford, K. H. Law, J. León Holgado, L. Lyons, A.-M. Magnan, B. Maier, S. Mallios, M. Mieskolainen, J. Nash[81], M. Pesaresi, P. B. Pradeep, B. C. Radburn-Smith, A. Richards, A. Rose, K. Savva, C. Seez, R. Shukla, A. Tapper, K. Uchida, G. P. Uttley, T. Virdee[29], M. Vojinovic, N. Wardle, D. Winterbottom

Brunel University, Uxbridge, UK
J. E. Cole, A. Khan, P. Kyberd, I. D. Reid

Baylor University, Waco, TX, USA
S. Abdullin, A. Brinkerhoff, E. Collins, M. R. Darwish, J. Dittmann, K. Hatakeyama, V. Hegde, J. Hiltbrand, B. McMaster, J. Samudio, S. Sawant, C. Sutantawibul, J. Wilson

Catholic University of America, Washington, DC, USA
R. Bartek, A. Dominguez, A. E. Simsek, S. S. Yu

The University of Alabama, Tuscaloosa, AL, USA
B. Bam, A. Buchot Perraguin, R. Chudasama, S. I. Cooper, C. Crovella, S. V. Gleyzer, E. Pearson, C. U. Perez, P. Rumerio[82], E. Usai, R. Yi

Boston University, Boston, MA, USA
A. Akpinar, C. Cosby, G. De Castro, Z. Demiragli, C. Erice, C. Fangmeier, C. Fernandez Madrazo, E. Fontanesi, D. Gastler, F. Golf, S. Jeon, J. O'cain, I. Reed, J. Rohlf, K. Salyer, D. Sperka, D. Spitzbart, I. Suarez, A. Tsatsos, A. G. Zecchinelli

Brown University, Providence, RI, USA
G. Barone, G. Benelli, D. Cutts, L. Gouskos, M. Hadley, U. Heintz, K. W. Ho, J. M. Hogan[83], T. Kwon, G. Landsberg, K. T. Lau, J. Luo, S. Mondal, T. Russell, S. Sagir[84], X. Shen, M. Stamenkovic, N. Venkatasubramanian

University of California, Davis, Davis, CA, USA
S. Abbott, B. Barton, C. Brainerd, R. Breedon, H. Cai, M. Calderon De La Barca Sanchez, M. Chertok, M. Citron, J. Conway, P. T. Cox, R. Erbacher, F. Jensen, O. Kukral, G. Mocellin, M. Mulhearn, S. Ostrom, W. Wei, S. Yoo, F. Zhang

University of California, Los Angeles, CA, USA
K. Adamidis, M. Bachtis, D. Campos, R. Cousins, A. Datta, G. Flores Avila, J. Hauser, M. Ignatenko, M. A. Iqbal, T. Lam, Y. f. Lo, E. Manca, A. Nunez Del Prado, D. Saltzberg, V. Valuev

University of California, Riverside, Riverside, CA, USA
R. Clare, J. W. Gary, G. Hanson

University of California, San Diego, La Jolla, CA, USA
A. Aportela, A. Arora, J. G. Branson, S. Cittolin, S. Cooperstein, D. Diaz, J. Duarte, L. Giannini, Y. Gu, J. Guiang, R. Kansal, V. Krutelyov, R. Lee, J. Letts, M. Masciovecchio, F. Mokhtar, S. Mukherjee, M. Pieri, D. Primosch, M. Quinnan, V. Sharma, M. Tadel, E. Vourliotis, F. Würthwein, Y. Xiang, A. Yagil

Department of Physics, University of California, Santa Barbara, Santa Barbara, CA, USA
A. Barzdukas, L. Brennan, C. Campagnari, K. Downham, C. Grieco, M. M. Hussain, J. Incandela, J. Kim, A. J. Li, P. Masterson, H. Mei, J. Richman, S. N. Santpur, U. Sarica, R. Schmitz, F. Setti, J. Sheplock, D. Stuart, T. Á. Vámi, X. Yan, D. Zhang

California Institute of Technology, Pasadena, CA, USA
S. Bhattacharya, A. Bornheim, O. Cerri, J. Mao, H. B. Newman, G. Reales Gutiérrez, M. Spiropulu, J. R. Vlimant, C. Wang, S. Xie, R. Y. Zhu

Carnegie Mellon University, Pittsburgh, PA, USA
J. Alison, S. An, P. Bryant, M. Cremonesi, V. Dutta, T. Ferguson, T. A. Gómez Espinosa, A. Harilal, A. Kallil Tharayil, M. Kanemura, C. Liu, T. Mudholkar, S. Murthy, P. Palit, K. Park, M. Paulini, A. Roberts, A. Sanchez, W. Terrill

University of Colorado Boulder, Boulder, CO, USA
J. P. Cumalat, W. T. Ford, A. Hart, A. Hassani, N. Manganelli, J. Pearkes, C. Savard, N. Schonbeck, K. Stenson, K. A. Ulmer, S. R. Wagner, N. Zipper, D. Zuolo

Cornell University, Ithaca, NY, USA
J. Alexander, X. Chen, D. J. Cranshaw, J. Dickinson, J. Fan, X. Fan, S. Hogan, P. Kotamnives, J. Monroy, M. Oshiro, J. R. Patterson, M. Reid, A. Ryd, J. Thom, P. Wittich, R. Zou

Fermi National Accelerator Laboratory, Batavia, IL, USA
M. Albrow, M. Alyari, O. Amram, G. Apollinari, A. Apresyan, L. A. T. Bauerdick, D. Berry, J. Berryhill, P. C. Bhat, K. Burkett, J. N. Butler, A. Canepa, G. B. Cerati, H. W. K. Cheung, F. Chlebana, G. Cummings, I. Dutta, V. D. Elvira, J. Freeman, A. Gandrakota, Z. Gecse, L. Gray, D. Green, A. Grummer, S. Grünendahl, D. Guerrero, O. Gutsche, R. M. Harris, T. C. Herwig, J. Hirschauer, B. Jayatilaka, S. Jindariani, M. Johnson, U. Joshi, T. Klijnsma, B. Klima, K. H. M. Kwok, S. Lammel, C. Lee, D. Lincoln, R. Lipton, T. Liu, K. Maeshima, D. Mason, P. McBride, P. Merkel, S. Mrenna, S. Nahn, J. Ngadiuba, D. Noonan, S. Norberg, V. Papadimitriou,

N. Pastika, K. Pedro, C. Pena[85], F. Ravera, A. Reinsvold Hall[86], L. Ristori, M. Safdari, E. Sexton-Kennedy, N. Smith, A. Soha, L. Spiegel, S. Stoynev, J. Strait, L. Taylor, S. Tkaczyk, N. V. Tran, L. Uplegger, E. W. Vaandering, I. Zoi

University of Florida, Gainesville, FL, USA
C. Aruta, P. Avery, D. Bourilkov, P. Chang, V. Cherepanov, R. D. Field, C. Huh, E. Koenig, M. Kolosova, J. Konigsberg, A. Korytov, K. Matchev, N. Menendez, G. Mitselmakher, K. Mohrman, A. Muthirakalayil Madhu, N. Rawal, S. Rosenzweig, Y. Takahashi, J. Wang

Florida State University, Tallahassee, FL, USA
T. Adams, A. Al Kadhim, A. Askew, S. Bower, R. Hashmi, R. S. Kim, S. Kim, T. Kolberg, G. Martinez, H. Prosper, P. R. Prova, M. Wulansatiti, R. Yohay, J. Zhang

Florida Institute of Technology, Melbourne, FL, USA
B. Alsufyani, S. Butalla, S. Das, T. Elkafrawy[87], M. Hohlmann, E. Yanes

University of Illinois Chicago, Chicago, IL, USA
M. R. Adams, A. Baty, C. Bennett, R. Cavanaugh, R. Escobar Franco, O. Evdokimov, C. E. Gerber, M. Hawksworth, A. Hingrajiya, D. J. Hofman, J. h. Lee, D. S. Lemos, C. Mills, S. Nanda, G. Oh, B. Ozek, D. Pilipovic, R. Pradhan, E. Prifti, T. Roy, S. Rudrabhatla, N. Singh, M. B. Tonjes, N. Varelas, M. A. Wadud, Z. Ye, J. Yoo

The University of Iowa, Iowa City, IA, USA
M. Alhusseini, D. Blend, K. Dilsiz[88], L. Emediato, G. Karaman, O. K. Köseyan, J.-P. Merlo, A. Mestvirishvili[89], O. Neogi, H. Ogul[90], Y. Onel, A. Penzo, C. Snyder, E. Tiras[91]

Johns Hopkins University, Baltimore, MD, USA
B. Blumenfeld, L. Corcodilos, J. Davis, A. V. Gritsan, L. Kang, S. Kyriacou, P. Maksimovic, M. Roguljic, J. Roskes, S. Sekhar, M. Swartz

The University of Kansas, Lawrence, KS, USA
A. Abreu, L. F. Alcerro Alcerro, J. Anguiano, S. Arteaga Escatel, P. Baringer, A. Bean, Z. Flowers, D. Grove, J. King, G. Krintiras, M. Lazarovits, C. Le Mahieu, J. Marquez, M. Murray, M. Nickel, M. Pitt, S. Popescu[92], C. Rogan, C. Royon, S. Sanders, C. Smith, G. Wilson

Kansas State University, Manhattan, KS, USA
B. Allmond, R. Gujju Gurunadha, A. Ivanov, K. Kaadze, Y. Maravin, J. Natoli, D. Roy, G. Sorrentino

University of Maryland, College Park, MD, USA
A. Baden, A. Belloni, J. Bistany-riebman, Y. M. Chen, S. C. Eno, N. J. Hadley, S. Jabeen, R. G. Kellogg, T. Koeth, B. Kronheim, Y. Lai, S. Lascio, A. C. Mignerey, S. Nabili, C. Palmer, C. Papageorgakis, M. M. Paranjpe, E. Popova[93], A. Shevelev, L. Wang, L. Zhang

Massachusetts Institute of Technology, Cambridge, MA, USA
C. Baldenegro Barrera, J. Bendavid, S. Bright-Thonney, I. A. Cali, P. c. Chou, M. D'Alfonso, J. Eysermans, C. Freer, G. Gomez-Ceballos, M. Goncharov, G. Grosso, P. Harris, D. Hoang, D. Kovalskyi, J. Krupa, L. Lavezzo, Y.-J. Lee, K. Long, C. Mcginn, A. Novak, M. I. Park, C. Paus, C. Reissel, C. Roland, G. Roland, S. Rothman, G. S. F. Stephans, Z. Wang, B. Wyslouch, T. J. Yang

University of Minnesota, Minneapolis, MN, USA
B. Crossman, C. Kapsiak, M. Krohn, D. Mahon, J. Mans, B. Marzocchi, M. Revering, R. Rusack, R. Saradhy, N. Strobbe

University of Nebraska-Lincoln, Lincoln, NE, USA
K. Bloom, D. R. Claes, G. Haza, J. Hossain, C. Joo, I. Kravchenko, A. Rohilla, J. E. Siado, W. Tabb, A. Vagnerini, A. Wightman, F. Yan, D. Yu

State University of New York at Buffalo, Buffalo, NY, USA
H. Bandyopadhyay, L. Hay, H. w. Hsia, I. Iashvili, A. Kalogeropoulos, A. Kharchilava, M. Morris, D. Nguyen, S. Rappoccio, H. Rejeb Sfar, A. Williams, P. Young

Northeastern University, Boston, MA, USA
G. Alverson, E. Barberis, J. Bonilla, B. Bylsma, M. Campana, J. Dervan, Y. Haddad, Y. Han, I. Israr, A. Krishna, P. Levchenko, J. Li, M. Lu, R. Mccarthy, D. M. Morse, V. Nguyen, T. Orimoto, A. Parker, L. Skinnari, E. Tsai, D. Wood

Northwestern University, Evanston, IL, USA
S. Dittmer, K. A. Hahn, D. Li, Y. Liu, M. Mcginnis, Y. Miao, D. G. Monk, M. H. Schmitt, A. Taliercio, M. Velasco

University of Notre Dame, Notre Dame, IN, USA
G. Agarwal, R. Band, R. Bucci, S. Castells, A. Das, R. Goldouzian, M. Hildreth, K. Hurtado Anampa, T. Ivanov, C. Jessop, K. Lannon, J. Lawrence, N. Loukas, L. Lutton, J. Mariano, N. Marinelli, I. Mcalister, T. McCauley, C. Mcgrady, C. Moore, Y. Musienko[22], H. Nelson, M. Osherson, A. Piccinelli, R. Ruchti, A. Townsend, Y. Wan, M. Wayne, H. Yockey, M. Zarucki, L. Zygala

The Ohio State University, Columbus, OH, USA
A. Basnet, M. Carrigan, L. S. Durkin, C. Hill, M. Joyce, M. Nunez Ornelas, K. Wei, D. A. Wenzl, B. L. Winer, B. R. Yates

Princeton University, Princeton, NJ, USA
H. Bouchamaoui, K. Coldham, P. Das, G. Dezoort, P. Elmer, P. Fackeldey, A. Frankenthal, B. Greenberg, N. Haubrich, K. Kennedy, G. Kopp, S. Kwan, D. Lange, A. Loeliger, D. Marlow, I. Ojalvo, J. Olsen, F. Simpson, D. Stickland, C. Tully, L. H. Vage

University of Puerto Rico, Mayaguez, PR, USA
S. Malik, R. Sharma

Purdue University, West Lafayette, IN, USA
A. S. Bakshi, S. Chandra, R. Chawla, A. Gu, L. Gutay, M. Jones, A. W. Jung, A. M. Koshy, M. Liu, G. Negro, N. Neumeister, G. Paspalaki, S. Piperov, J. F. Schulte, A. K. Virdi, F. Wang, A. Wildridge, W. Xie, Y. Yao

Purdue University Northwest, Hammond, IN, USA
J. Dolen, N. Parashar, A. Pathak

Rice University, Houston, TX, USA
D. Acosta, A. Agrawal, T. Carnahan, K. M. Ecklund, P. J. Fernández Manteca, S. Freed, P. Gardner, F. J. M. Geurts, I. Krommydas, W. Li, J. Lin, O. Miguel Colin, B. P. Padley, R. Redjimi, J. Rotter, E. Yigitbasi, Y. Zhang

University of Rochester, Rochester, NY, USA
A. Bodek, P. de Barbaro, R. Demina, J. L. Dulemba, A. Garcia-Bellido, O. Hindrichs, A. Khukhunaishvili, N. Parmar, P. Parygin[93], R. Taus

Rutgers, The State University of New Jersey, Piscataway, NJ, USA
B. Chiarito, J. P. Chou, S. V. Clark, D. Gadkari, Y. Gershtein, E. Halkiadakis, M. Heindl, C. Houghton, D. Jaroslawski, S. Konstantinou, I. Laflotte, A. Lath, R. Montalvo, K. Nash, J. Reichert, P. Saha, S. Salur, S. Schnetzer, S. Somalwar, R. Stone, S. A. Thayil, S. Thomas, J. Vora

University of Tennessee, Knoxville, TN, USA
D. Ally, A. G. Delannoy, S. Fiorendi, S. Higginbotham, T. Holmes, A. R. Kanuganti, N. Karunarathna, L. Lee, E. Nibigira, S. Spanier

Texas A&M University, College Station, TX, USA
D. Aebi, M. Ahmad, T. Akhter, K. Androsov[62], O. Bouhali[94], R. Eusebi, J. Gilmore, T. Huang, T. Kamon[95], H. Kim, S. Luo, R. Mueller, D. Overton, A. Safonov

Texas Tech University, Lubbock, TX, USA
N. Akchurin, J. Damgov, Y. Feng, N. Gogate, Y. Kazhykarim, K. Lamichhane, S. W. Lee, C. Madrid, A. Mankel, T. Peltola, I. Volobouev

Vanderbilt University, Nashville, TN, USA
E. Appelt, Y. Chen, S. Greene, A. Gurrola, W. Johns, R. Kunnawalkam Elayavalli, A. Melo, D. Rathjens, F. Romeo, P. Sheldon, S. Tuo, J. Velkovska, J. Viinikainen

University of Virginia, Charlottesville, VA, USA
B. Cardwell, H. Chung, B. Cox, J. Hakala, R. Hirosky, A. Ledovskoy, C. Mantilla, C. Neu, C. Ramón Álvarez

Wayne State University, Detroit, MI, USA
S. Bhattacharya, P. E. Karchin

University of Wisconsin-Madison, Madison, WI, USA
A. Aravind, S. Banerjee, K. Black, T. Bose, E. Chavez, S. Dasu, P. Everaerts, C. Galloni, H. He, M. Herndon, A. Herve, C. K. Koraka, A. Lanaro, R. Loveless, J. Madhusudanan Sreekala, A. Mallampalli, A. Mohammadi, S. Mondal, G. Parida, L. Pétré, D. Pinna, A. Savin, V. Shang, V. Sharma, W. H. Smith, D. Teague, H. F. Tsoi, W. Vetens, A. Warden

Authors Affiliated with an International Laboratory Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
S. Afanasiev, V. Alexakhin, D. Budkouski, I. Golutvin, I. Gorbunov, V. Karjavine, O. Kodolova[96,93], V. Korenkov, A. Lanev, A. Malakhov, V. Matveev[97], A. Nikitenko[98,96], V. Palichik, V. Perelygin, M. Savina, V. Shalaev, S. Shmatov, S. Shulha, V. Smirnov, O. Teryaev, N. Voytishin, B. S. Yuldashev[99], A. Zarubin, I. Zhizhin, Y u. Andreev, A. Dermenev, S. Gninenko, N. Golubev, A. Karneyeu, D. Kirpichnikov, M. Kirsanov, N. Krasnikov, I. Tlisova, A. Toropin

Authors Affiliated with an Institute Formerly Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
G. Gavrilov, V. Golovtcov, Y. Ivanov, V. Kim[100], V. Murzin, V. Oreshkin, D. Sosnov, V. Sulimov, L. Uvarov, A. Vorobyev, T. Aushev, K. Ivanov, V. Gavrilov, N. Lychkovskaya, V. Popov, A. Zhokin, M. Chadeeva[100], R. Chistov[100], S. Polikarpov[100], V. Andreev, M. Azarkin, M. Kirakosyan, A. Terkulov, E. Boos, V. Bunichev, M. Dubinin[85], L. Dudko, A. Ershov, A. Gribushin, V. Klyukhin, M. Perfilov, V. Savrin, P. Volkov, V. Blinov[100], T. Dimova[100], A. Kozyrev[100], O. Radchenko[100], Y. Skovpen[100], V. Kachanov, S. Slabospitskii, A. Uzunian, A. Babaev, V. Borshch, D. Druzhkin

1: Also at Yerevan State University, Yerevan, Armenia
2: Also at TU Wien, Vienna, Austria
3: Also at Ghent University, Ghent, Belgium
4: Also at Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
5: Also at FACAMP-Faculdades de Campinas, Sao Paulo, Brazil
6: Also at Universidade Estadual de Campinas, Campinas, Brazil
7: Also at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
8: Also at University of Chinese Academy of Sciences, Beijing, China
9: Also at China Center of Advanced Science and Technology, Beijing, China
10: Also at University of Chinese Academy of Sciences, Beijing, China
11: Also at China Spallation Neutron Source, Guangdong, China
12: Now at Henan Normal University, Xinxiang, China
13: Also at University of Shanghai for Science and Technology, Shanghai, China
14: Now at The University of Iowa, Iowa City, IA, USA
15: Also at An Institute Formerly Covered by a Cooperation Agreement with CERN, Geneva, Switzerland

16: Also at Cairo University, Cairo, Egypt
17: Also at Suez University, Suez, Egypt
18: Now at British University in Egypt, Cairo, Egypt
19: Also at Purdue University, West Lafayette, IN, USA
20: Also at Université de Haute Alsace, Mulhouse, France
21: Also at Istinye University, Istanbul, Turkey
22: Also at An International Laboratory Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
23: Also at The University of the State of Amazonas, Manaus, Brazil
24: Also at University of Hamburg, Hamburg, Germany
25: Also at RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
26: Also at Bergische University Wuppertal (BUW), Wuppertal, Germany
27: Also at Brandenburg University of Technology, Cottbus, Germany
28: Also at Forschungszentrum Jülich, Juelich, Germany
29: Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland
30: Also at HUN-REN ATOMKI-Institute of Nuclear Research, Debrecen, Hungary
31: Now at Universitatea Babes-Bolyai-Facultatea de Fizica, Cluj-Napoca, Romania
32: Also at MTA-ELTE Lendület CMS Particle and Nuclear Physics Group, Eötvös Loránd University, Budapest, Hungary
33: Also at HUN-REN Wigner Research Centre for Physics, Budapest, Hungary
34: Also at Physics Department, Faculty of Science, Assiut University, Assiut, Egypt
35: Also at Punjab Agricultural University, Ludhiana, India
36: Also at University of Visva-Bharati, Santiniketan, India
37: Also at Indian Institute of Science (IISc), Bangalore, India
38: Also at Amity University Uttar Pradesh, Noida, India
39: Also at IIT Bhubaneswar, Bhubaneswar, India
40: Also at Institute of Physics, Bhubaneswar, India
41: Also at University of Hyderabad, Hyderabad, India
42: Also at Deutsches Elektronen-Synchrotron, Hamburg, Germany
43: Also at Isfahan University of Technology, Isfahan, Iran
44: Also at Sharif University of Technology, Tehran, Iran
45: Also at Department of Physics, University of Science and Technology of Mazandaran, Behshahr, Iran
46: Also at Department of Physics, Faculty of Science, Arak University, Arak, Iran
47: Also at Helwan University, Cairo, Egypt
48: Also at Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Bologna, Italy
49: Also at Centro Siciliano di Fisica Nucleare e di Struttura Della Materia, Catania, Italy
50: Also at Università degli Studi Guglielmo Marconi, Rome, Italy
51: Also at Scuola Superiore Meridionale, Università di Napoli 'Federico II', Naples, Italy
52: Also at Fermi National Accelerator Laboratory, Batavia, IL, USA
53: Also at Lulea University of Technology, Lulea, Sweden
54: Also at Laboratori Nazionali di Legnaro dell'INFN, Legnaro, Italy
55: Also at Consiglio Nazionale delle Ricerche-Istituto Officina dei Materiali, Perugia, Italy
56: Also at Institut de Physique des 2 Infinis de Lyon (IP2I), Villeurbanne, France
57: Also at Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia
58: Also at Consejo Nacional de Ciencia y Tecnología, Mexico City, Mexico
59: Also at Trincomalee Campus, Eastern University, Nilaveli, Sri Lanka
60: Also at Saegis Campus, Nugegoda, Sri Lanka
61: Also at National and Kapodistrian University of Athens, Athens, Greece
62: Also at Ecole Polytechnique Fédérale Lausanne, Lausanne, Switzerland
63: Also at Universität Zürich, Zurich, Switzerland
64: Also at Stefan Meyer Institute for Subatomic Physics, Vienna, Austria
65: Also at Laboratoire d'Annecy-le-Vieux de Physique des Particules, IN2P3-CNRS, Annecy-le-Vieux, France
66: Also at Near East University, Research Center of Experimental Health Science, Mersin, Turkey
67: Also at Konya Technical University, Konya, Turkey

68: Also at Izmir Bakircay University, Izmir, Turkey
69: Also at Adiyaman University, Adiyaman, Turkey
70: Also at Bozok Universitetesi Rektörlügü, Yozgat, Turkey
71: Also at Marmara University, Istanbul, Turkey
72: Also at Milli Savunma University, Istanbul, Turkey
73: Also at Kafkas University, Kars, Turkey
74: Now at Istanbul Okan University, Istanbul, Turkey
75: Also at Hacettepe University, Ankara, Turkey
76: Also at Erzincan Binali Yildirim University, Erzincan, Turkey
77: Also at Faculty of Engineering, Istanbul University-Cerrahpasa, Istanbul, Turkey
78: Also at Yildiz Technical University, Istanbul, Turkey
79: Also at School of Physics and Astronomy, University of Southampton, Southampton, UK
80: Also at IPPP Durham University, Durham, UK
81: Also at Faculty of Science, Monash University, Clayton, Australia
82: Also at Università di Torino, Turin, Italy
83: Also at Bethel University, St. Paul, MN, USA
84: Also at Karamanoğlu Mehmetbey University, Karaman, Turkey
85: Also at California Institute of Technology, Pasadena, CA, USA
86: Also at United States Naval Academy, Annapolis, MD, USA
87: Also at Ain Shams University, Cairo, Egypt
88: Also at Bingol University, Bingol, Turkey
89: Also at Georgian Technical University, Tbilisi, Georgia
90: Also at Sinop University, Sinop, Turkey
91: Also at Erciyes University, Kayseri, Turkey
92: Also at Horia Hulubei National Institute of Physics and Nuclear Engineering (IFIN-HH), Bucharest, Romania
93: Now at Another Institute Formerly Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
94: Also at Texas A&M University at Qatar, Doha, Qatar
95: Also at Kyungpook National University, Daegu, Korea
96: Also at Yerevan Physics Institute, Yerevan, Armenia
97: Also at Another International Laboratory Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
98: Also at Imperial College, London, UK
99: Also at Institute of Nuclear Physics of the Uzbekistan Academy of Sciences, Tashkent, Uzbekistan
100: Also at Another Institute Formerly Covered by a Cooperation Agreement with CERN, Geneva, Switzerland
* Deceased