Supplementary Material

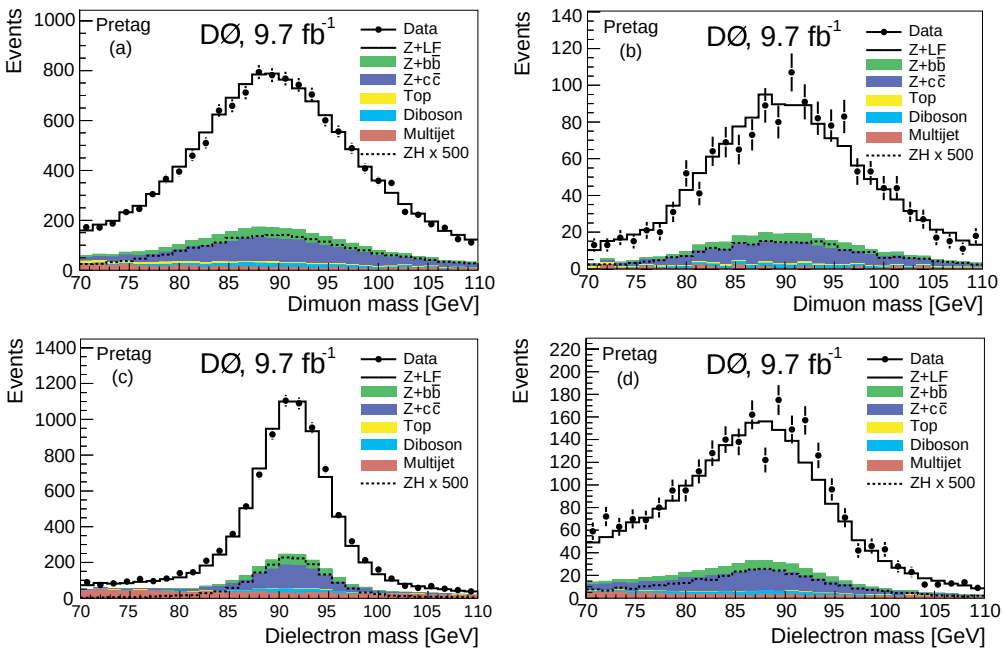

FIG. 1: The dilepton mass spectra in the (a) $\mu\mu$, (b) $\mu\mu_{\rm trk}$, (c) ee and (d) $ee_{\rm ICR}$ channels. Distributions are shown in the pretag control sample, in which all selection requirements except b-tagging are applied. Signal distributions, for $M_H = 125$ GeV, are scaled by a factor of 500.

The dimuon and dielectron mass spectra, after requiring two leptons and at least two jets are shown in Fig. 1.

Distributions of the dijet invariant mass spectra before and after adjustment by the kinematic fit, are shown in Fig. 2.

A complete list of RF input variables is shown in Table I.

Comparisons of the data and MC distributions of the $t\bar{t}$ RF output summed over all lepton channels are shown for $M_H = 125$ GeV in Figure 3.

Post-kinematic fit dijet mass distributions for ST and DT in the $t\bar{t}$ depleted region are shown in Fig. 4.

Fig. 5 displays the global RF distributions in the $t\bar{t}$ enriched region, after the fit to the background-only hypothesis.

Fig. 6 shows the observed LLR as a function of Higgs boson mass. Also shown are the expected (median) LLRs for the background-only and signal+background hypotheses, together with the one and two standard deviation bands about the background-only expectation.

Fig. 7 shows the post-fit RF distributions in the $t\bar{t}$ depleted region for the VZ search. Fig. 8 displays the post-fit distribution of the dijet invariant mass from the kinematic fit.

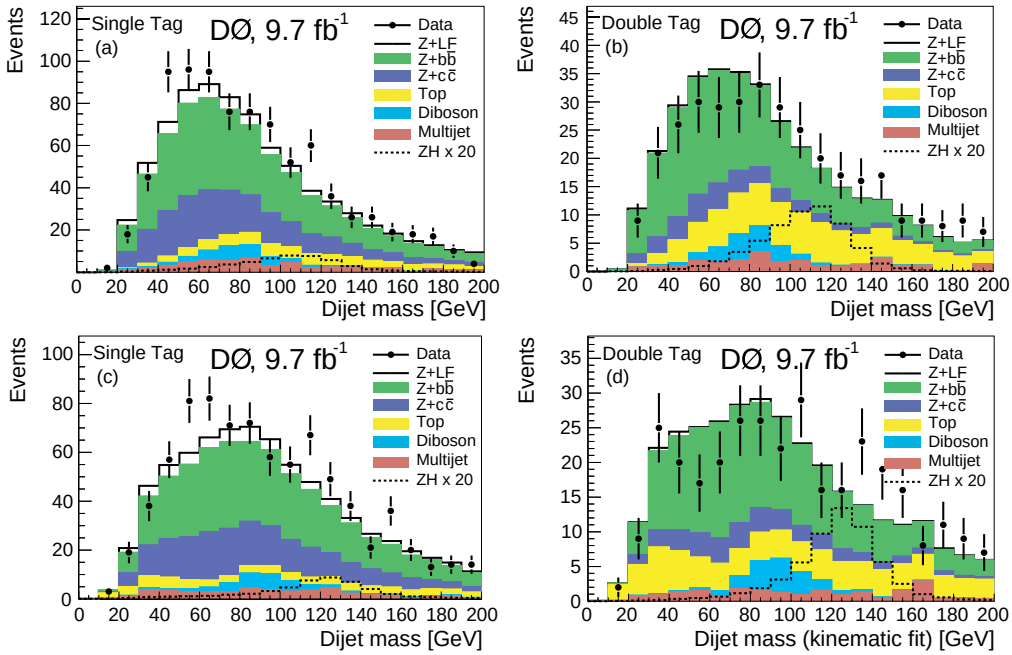

FIG. 2: Dijet invariant mass distributions before the kinematic fit in (a) ST events and (b) DT events; and after the kinematic fit in (c) ST events and (d) DT events, combined for all lepton channels. Signal distributions ($M_H = 125$ GeV) are shown with the SM cross section multiplied by 20.

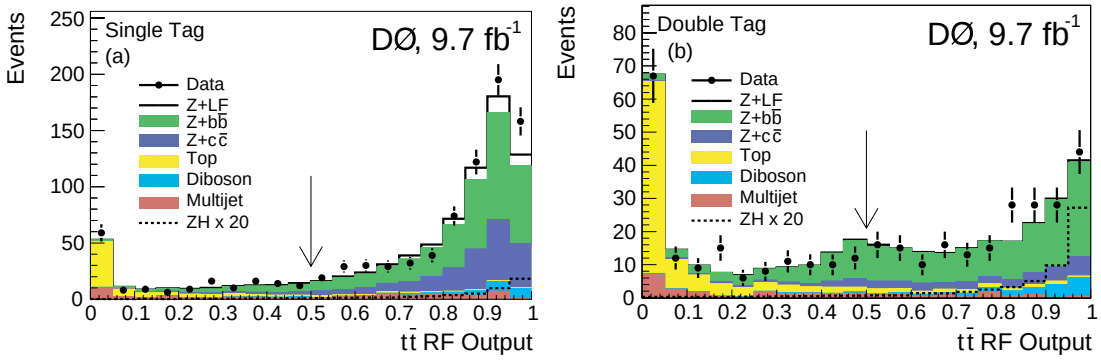

FIG. 3: The $t\bar{t}$ RF output ($M_H = 125$ GeV) for all lepton channels combined (a) ST and (b) DT events. Signal distributions are shown with the SM cross section multiplied by 20. The vertical arrows indicate the cut at $t\bar{t}$ RF $= 0.5$ that defines the $t\bar{t}$ enriched and depleted regions.

variables	definition	$t\bar{t}$ RF	global RF
m_{bb} ($m_{bb}{}^{fit}$)	invariant mass of the dijet system before (after) the kinematic fit	$\surd$	$\surd$
p_T^{b1} ($p_T^{b1\,fit}$)	transverse momentum of the first jet before (after) kinematic fit	$\surd$	$\surd$
p_T^{b2} ($p_T^{b2\,fit}$)	transverse momentum of the second jet before (after) kinematic fit	$\surd$	$\surd$
p_T^{bb}	transverse momentum of the dijet system before the kinematic fit	$\surd$	$\surd$
$\Delta\phi(b_1, b_2)$	$\Delta\phi$ between the two jets in the dijet system	–	$\surd$
$\Delta\eta(b_1, b_2)$	$\Delta\eta$ between the two jets in the dijet system	–	$\surd$
$m(\sum j_i)$	invariant mass of all jets in the event (the multijet mass)	$\surd$	$\surd$
$p_T(\sum j_i)$	transverse momentum of all jets in the event	$\surd$	$\surd$
$H_T(\sum j_i)$	scalar sum of the transverse momenta of all jets in the event	$\surd$	–
$p_T^{bb}/(\lvert p_T^{b1}\rvert + \lvert p_T^{b2}\rvert)$	ratio of dijet system p_T over the scalar sum of the p_T of the two jets	$\surd$	–
$m_{\ell\ell}$	invariant mass of the dilepton system	$\surd$	–
p_T^Z	transverse momentum of the dilepton system	$\surd$	$\surd$
$\Delta\phi(\ell_1, \ell_2)$	$\Delta\phi$ between the two leptons	$\surd$	$\surd$
colinearity(ℓ_1, ℓ_2)	cosine of the angle between the two leptons (colinearity)	$\surd$	$\surd$
$\Delta\phi(\ell\ell, bb)$	$\Delta\phi$ between the dilepton and dijet systems	$\surd$	$\surd$
$\cos\theta^*$	cosine of the angle between the incoming proton and the Z in the zero momentum frame [34]	–	$\surd$
$m(\ell\ell bb)$	Invariant mass of dilepton plus dijet system	–	$\surd$
$H_T(\ell\ell bb)$	Scalar sum of the transverse momenta of the leptons and jets	–	$\surd$
$\not{E}_T$	missing transverse energy of the event	$\surd$	–
$\not{E}_T^{sig}$	the $\not{E}_T$ significance [35]	$\surd$	$\surd$
$-\ln\ L_{fit}$	negative log likelihood from the kinematic fit	$\surd$	$\surd$
$t\bar{t}$ `RF`	$t\bar{t}$ RF output	–	$\surd$

TABLE I: Variables used for the $t\bar{t}$ and global RF training. The jets that form the Higgs boson candidate are referred to as b_1 and b_2.

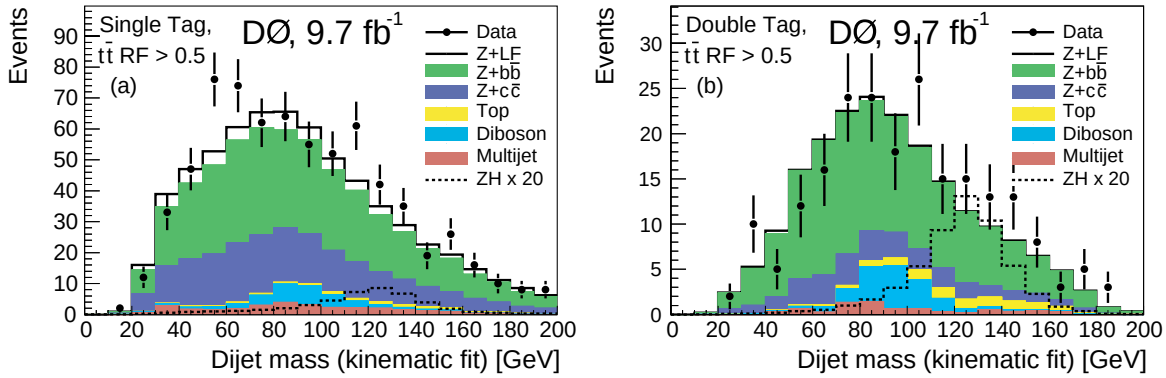

FIG. 4: Post-kinematic fit dijet mass distributions in the $t\bar{t}$ depleted region for all lepton channels combined assuming $M_H = 125$ GeV for (a) ST events and (b) DT events. Signal distributions are shown with the SM cross section multiplied by 20.

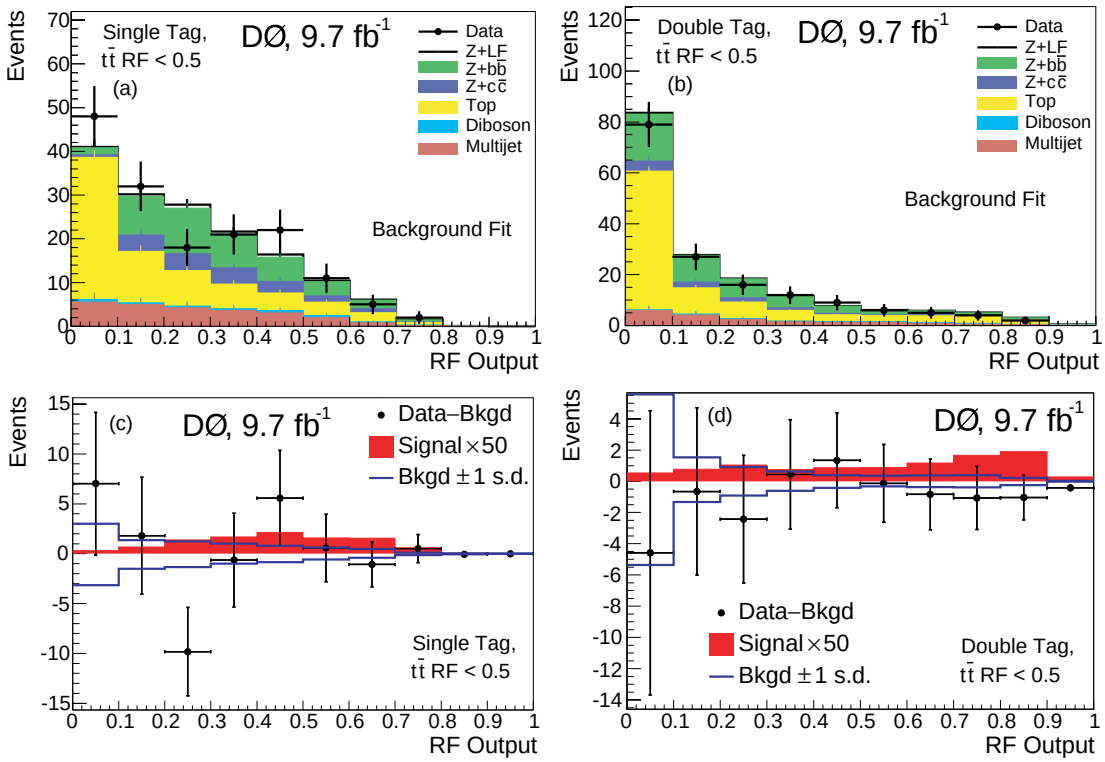

FIG. 5: Post-fit RF output distributions in the $t\bar{t}$ enriched region, assuming $M_H = 125$ GeV, after the fit to the background-only model for (a) ST events and (b) DT events. Background-subtracted distributions for (a) and (b) are shown in (c) and (d), respectively. Signal distributions are shown with the SM cross section scaled to 50 × SM prediction in (c) and (d). The blue lines indicate the uncertainty from the fit.

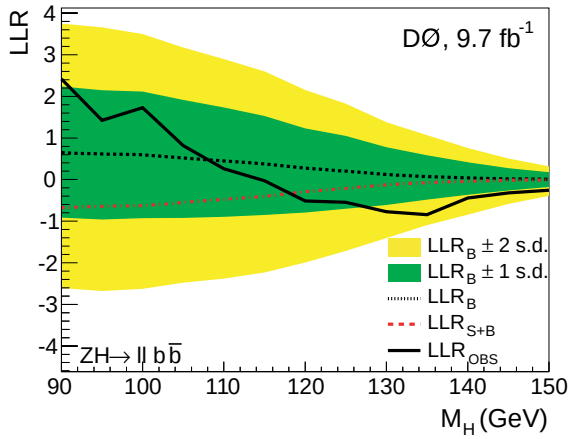

FIG. 6: Observed LLR as a function of Higgs boson mass. Also shown are the expected LLRs for the background-only (B) and signal+background (S+B) hypotheses, together with the one and two standard deviation (s.d.) bands about the background-only expectation.

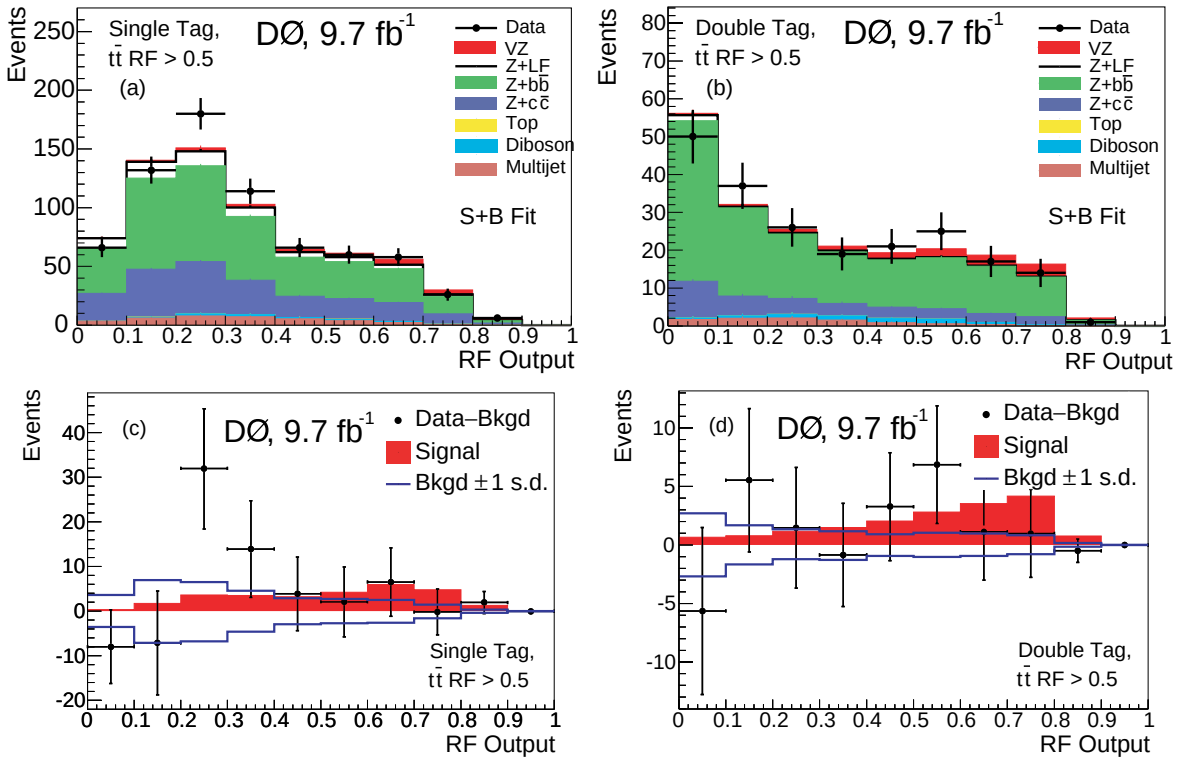

FIG. 7: Post-fit VZ RF output distributions in the $t\bar{t}$ depleted region after the fit to the S+B model for (a) ST events and (b) DT events. Background-subtracted distributions for (a) and (b) are shown in (c) and (d), respectively. Signal distributions are scaled to the fit VZ cross section. The blue lines indicate the uncertainty from the fit.

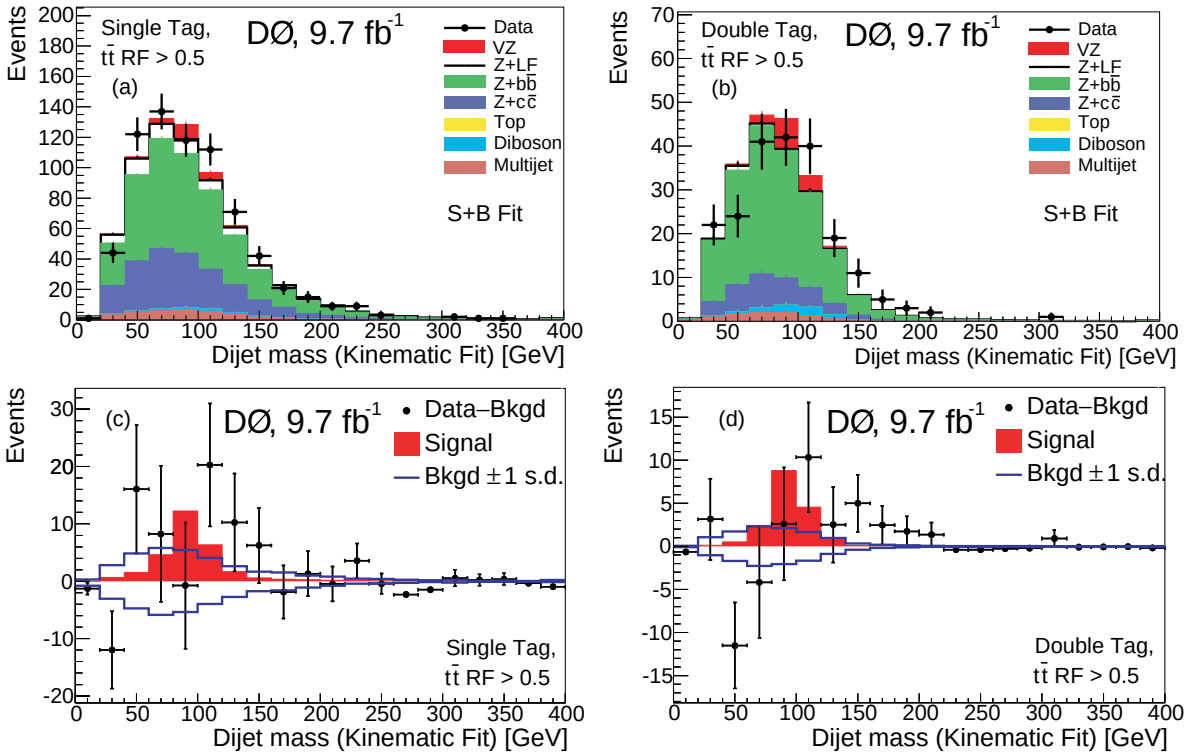

FIG. 8: Post-fit distributions of the dijet invariant mass (from the kinematic fit) in the $t\bar{t}$ depleted region after the fit to the S+B model for (a) ST events and (b) DT events. Background-subtracted distributions for (a) and (b) are shown in (c) and (d), respectively. Signal distributions are scaled to the fit VZ cross section. The blue lines indicate the uncertainty from the fit.