double parton scattering at CMS

in the $W^{\pm}W^{\pm}$ channel

marc dünser (CERN)
on behalf of the CMS collaboration

MPI@LHC, Prague
18th of november 2019

introduction

a brief outline of the talk:

-> (very) short introduction to some theory
on DPS cross sections

-> a look at factorization and the 'pocket formula'
and sigma effective

-> a new CMS analysis: **first evidence of DPS WW!**
CMS PAS SMP-18-015: DPS in $W^{\pm}W^{\pm}$
(submitted to EPJC)

-> future analyses for DPS WW
studies done for the HL-LHC

factorized DPS cross sections

can write down a general form of a double parton scattering cross section

-> with pdf terms and partonic cross sections

$$\sigma^D_{(A,B)} = \frac{m}{2} \sum_{i,j,k,l} \int \Gamma_{ij}(x_1, x_2, b; t_1, t_2) \hat{\sigma}^A_{ik}(x_1, x_1') \hat{\sigma}^B_{jl}(x_2, x_2') \times \Gamma_{kl}(x_1', x_2', b; t_1, t_2) dx_1 dx_2 dx_1' dx_2' d^2b$$

partonic cross sections

pdf terms

distance between partons

can decompose the pdfs into longitudinal versus transverse components

$$\Gamma_{ij}(x_1, x_2, b; t_1, t_2) = D^{ij}_h(x_1, x_2; t_1, t_2) F^i_j(b)$$

-> F(b) now related to the extend of the transverse parton flux

can then also assume longitudinal factorization of the pdfs

$$D^{ij}(x_1, x_2; t_1, t_2) = f^i(x_1; t_1) f^j(x_2; t_2)$$

-> these pdf terms are now again the same as the SPS process

the 'pocket formula'

if these factorizations are assumed, the cross sections simplifies a lot

-> two SPS cross sections and some factor related to b

$$\sigma^D_{(A,B)} = \frac{m}{2}\int dx_1 dx_1' f^i(x_1;t_1) f^k(x_1';t_1)\hat{\sigma}^A(x_1,x_1')\; dx_2 dx_2' f^j(x_2;t_2) f^l(x_2';t_2)\hat{\sigma}^B(x_2,x_2')\; F^i_j(b)F^k_l(b)d^2b$$

full SPS of A

full SPS of B

transverse part

write down the transverse component as a factor, call it *'sigma effective'*

$$\sigma_{\text{eff}} = \left[\int d^2b(F(b))^2\right]^{-1}$$

can separate two integrals and write down a very simple formula

-> leading to the fully factorized cross section for DPS

$$\sigma^D_{(A,B)} = \frac{m}{2}\frac{\sigma^S_{(A)}\sigma^S_{(B)}}{\sigma_{\text{eff}}}$$

really simple to calculate cross-sections on the back of an envelope

solutions to the factorization issue

there are theoretical calculations that do *not* assume factorization

-> not yet implemented in any large-scale MC simulation (to my knowledge)

summarizing here the works of many theorists:

-> Gaunt, Stirling, arXiv:0910.4347v4, 2010
Double Parton Distributions Incorporating Perturbative QCD Evolution and Momentum and Quark Number Sum Rules

-> Ceccopieri, Rinaldi, Scopetta, arXiv:1702.05363v1, 2017
Parton correlations in same-sign W pair production via double parton scattering at the LHC

-> Bartalini, Gaunt
Multiple parton interactions at the LHC, WorldScientific, 2019

these papers introduce more complex theoretical calculations

-> the last one is a 'state of the art' summary on DPS in general (only little on $W^{\pm}W^{\pm}$ though)

implications of these (theoretical) solutions

any of the solutions presented imply *correlations*

-> especially longitudinal correlations of the partons

some of these correlations have experimental implications

-> those are subtle/small effects, difficult to test

longitudinal effects affect especially the rapidity distributions

-> e.g. relation between parton x and muon p_T/η in W production

$$x_a = e^{\eta_\mu} \frac{M_W}{\sqrt{s}} \left[\frac{M_W}{2p_T} \pm \left(\sqrt{\left(\frac{M_W}{2p_T}\right)^2 - 1} \right) \right] \qquad x_b = e^{-\eta_\mu} \frac{M_W}{\sqrt{s}} \left[\frac{M_W}{2p_T} \mp \left(\sqrt{\left(\frac{M_W}{2p_T}\right)^2 - 1} \right) \right]$$

any probe must satisfy a few criteria

-> sensitivity to the correlations

-> large enough cross section (#events)

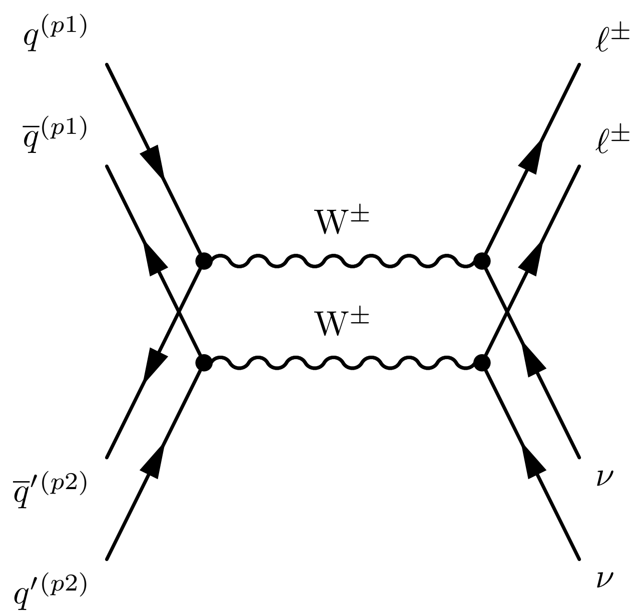
-> high purity to extract subtle correlations

the $W^\pm W^\pm$ channel is a probe that can test these correlations

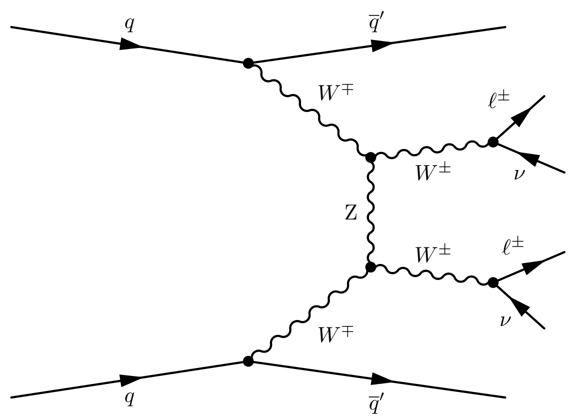
DPS in $W^{\pm}W^{\pm}$ in CMS

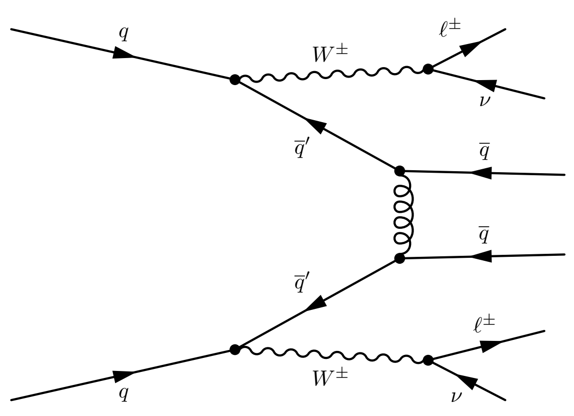
newest DPS analysis from CMS with 77.4 fb^{-1} at 13 TeV (2016+2017 datasets)

-> highly interesting channel for DPS!



versus





pro: SPS process is highly suppressed!

-> need two jets to carry away some charge

-> can veto these jets in the analysis

con: pretty low cross section, very crowded phase space

-> few hundred events after all selections

-> not yet sensitive to the subtle correlation effects

the story of the DPS WW cross section

this analysis does not have a single, accurate estimation of the total cross section

-> vastly different from Higgs, W/Z, top, even SUSY cross sections
-> no (N)NLO calculations with a MC generator exist

two options to get an estimate of the inclusive cross section:

1) calculate the DPS WW cross section via the pocket formula

-> take highest order theoretical W cross section (187 ± 7 nb)
-> choose a value for σ_{eff} (say, 20.7 mb from CMS W+2jets)
-> plug it in the formula, and get: **0.87 pb**

2) ask generators what the cross section is

-> pythia is the only one with sensible results
-> the pythia cross section is *very* tune dependent (up to ~30%)
-> for the CP5 tune we get: **1.92 pb**

these numbers are very different, but reflect the uncertainty

DPS in $W^{\pm}W^{\pm}$ in CMS

this process was never measured before at a hadron collider!

goals of the analysis:

-> prove that it's there first!

-> measure a cross section and see if it agrees with pythia/factorization

-> extract a value for σ_{eff}

phase space rather crowded, no strong handle to suppress backgrounds

-> basically two W's at LO

-> no high-p_T objects

-> no (b)-jets

a fairly loose set of selection requirements implemented:

Two leptons: $e^{\pm}\mu^{\pm}$ or $\mu^{\pm}\mu^{\pm}$
$p_T^{\ell_1} > 25\,\text{GeV}$, $p_T^{\ell_2} > 20\,\text{GeV}$
$\lvert\eta_e\rvert < 2.5$, $\lvert\eta_\mu\rvert < 2.4$
$p_T^{\text{miss}} > 15\,\text{GeV}$
$N_{\text{jets}} < 2$ ($p_T^{\text{jet}} > 30\,\text{GeV}$ and $\lvert\eta_{\text{jet}}\rvert < 2.5$)
$N_{\text{b-tagged jets}} = 0$ ($p_T^{\text{b jet}} > 25\,\text{GeV}$ and $\lvert\eta_{\text{b jet}}\rvert < 2.4$)
Veto on additional e, μ, and τ_h candidates

the backgrounds very briefly

backgrounds are plentiful in this region of phase-space

-> reducible and irreducible backgrounds

two most important backgrounds:

-> irreducible **WZ->3lnu** around 40% of total backgrounds
if the wrong Z-lepton is lost, it's *very* similar

-> reducible **nonprompt leptons** around 30% of total backgrounds
estimated with standard fakerate (tight-to-loose) method

other backgrounds estimated from MC, most are pretty standard

-> $W\gamma^*$, WWW, SPS $W^\pm W^\pm$, ZZ, $W/Z\gamma$
-> charge flips for electrons very small contribution

analysis strategy - I

train two BDTs in signal versus WZ and signal versus nonprompt

-> signal and background kinematics well defined

we train on 11 kinematic input variables

-> originally chosen between signal and WZ

-> they work very well against nonprompt too

full list of variables for the training

-> full list: $p_T^{1,2}$ of the two leptons

ME_T

$eta_1 * eta_2$, $|eta_1 + eta_2|$, sensitive to correlations

M_{T2}^{ll}, has an endpoint at m_W for signal

$m_{T(l1,met)}$, $m_{T(l1,l2)}$,

$dphi_{(l1,l2)}$, $dphi_{(l2,met)}$, $dphi_{(l1l2,l2)}$

want to be able to fit a 1D distribution out of these two BDTs

-> also for plotting/presentation this is better

analysis strategy - II

combine the two BDT classifier into one discriminant variable

first make a 2D plane of BDT_{WZ} versus $BDT_{nonprompt}$

-> combine contiguous regions in this 2D plane
-> need/want some regions with:
large signal, low background
large WZ & low fakes
large fakes & low WZ
-> optimized on the expected significance

profit further from two facts:

-> larger ++ signal than - -
-> μμ much superior experimentally than eμ

perform a binned ML fit in four lepton flavor and charge channels

systematic uncertainties

largest single uncertainty by far statistical

-> at the end all systematic uncertainties are equal
to the statistical uncertainty

systematic uncertainties

-> background normalizations:
nonprompt 25(40)% for μμ(eμ) - by far dominant
uncertainties on WZ/ZZ around 10% from 3l/4/ regions
charge mis-ID 30%
other backgrounds 50% due to lack of control regions

-> shape uncertainties
5-10% on all backgrounds from MC
nonprompt uncertainties from the variation of
the 'fake-rate' in p_T/|eta|
signal shape free between pythia and herwig (they are similar)

other smaller uncertainties

-> luminosity/trigger/reco/ID/ISO/pileup/

results

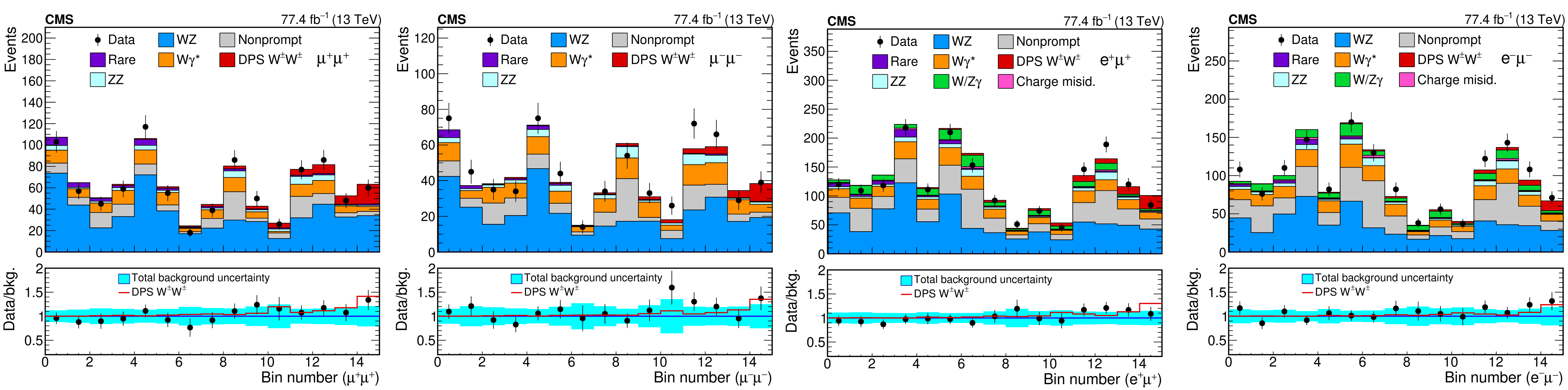
showing *postfit* plots of the final 1D classifier

-> somewhat of an under fluctuation
w/r/t to the prediction from pythia8

found a total of 4921 events in data

-> most of them to constrain backgrounds

-> fit extracts ~209 ± 59 signal events



$\mu^+\mu^+$ $\mu^-\mu^-$ $e^+\mu^+$ $e^-\mu^-$

decreasing sensitivity

first evidence of DPS WW

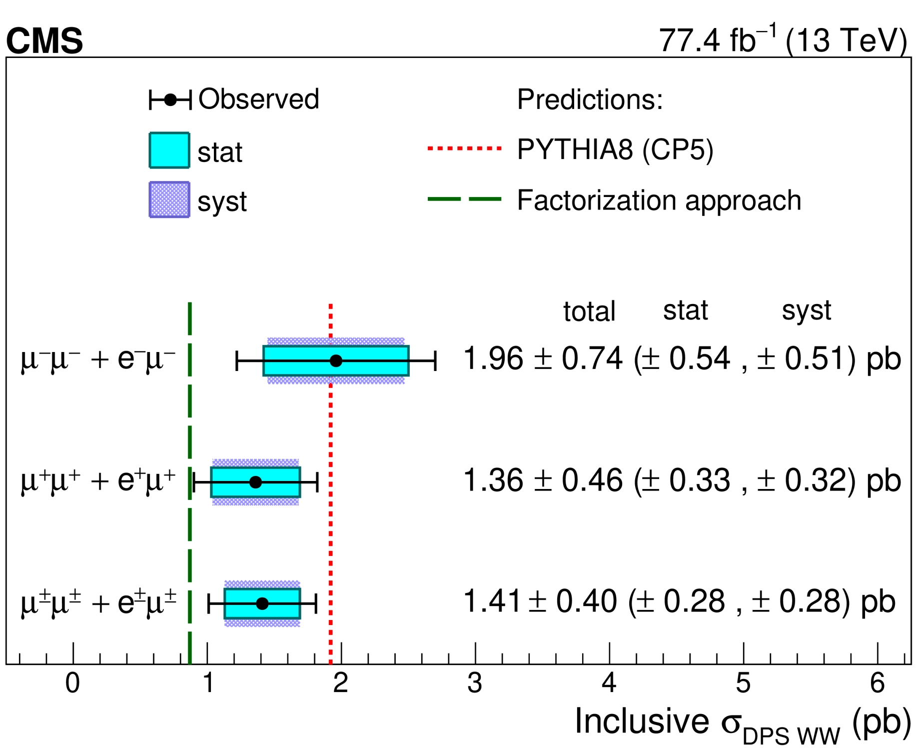
sensitivity large enough to claim first evidence

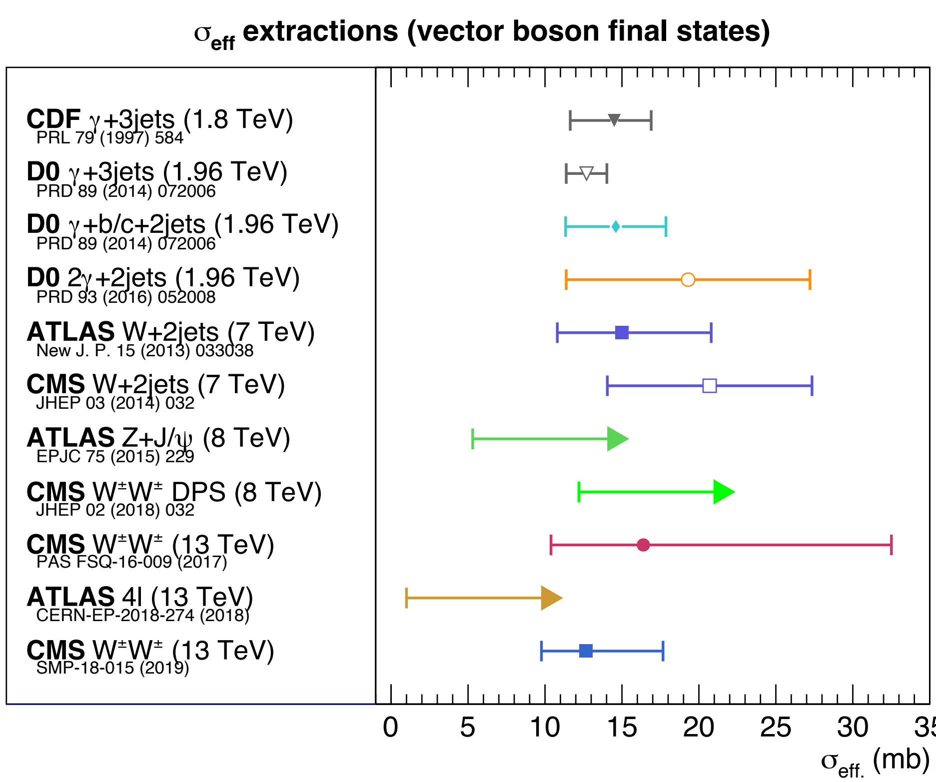
-> including 2018 should be enough to get to observation

	Value	Significance (standard deviations)
$\sigma^{\text{PYTHIA}}_{\text{DPS WW, exp}}$	1.92 pb	5.4
$\sigma^{\text{factorized}}_{\text{DPS WW, exp}}$	0.87 pb	2.5
$\sigma_{\text{DPS WW, obs}}$	$1.41 \pm 0.28\,(\text{stat}) \pm 0.28\,(\text{syst})$ pb	3.9
σ_{eff}	$12.7^{+5.0}_{-2.9}$ mb	—

also extract:

-> signal strength as function of charge





quo vadis, DPS?

the LHC is only at the beginning of data-taking

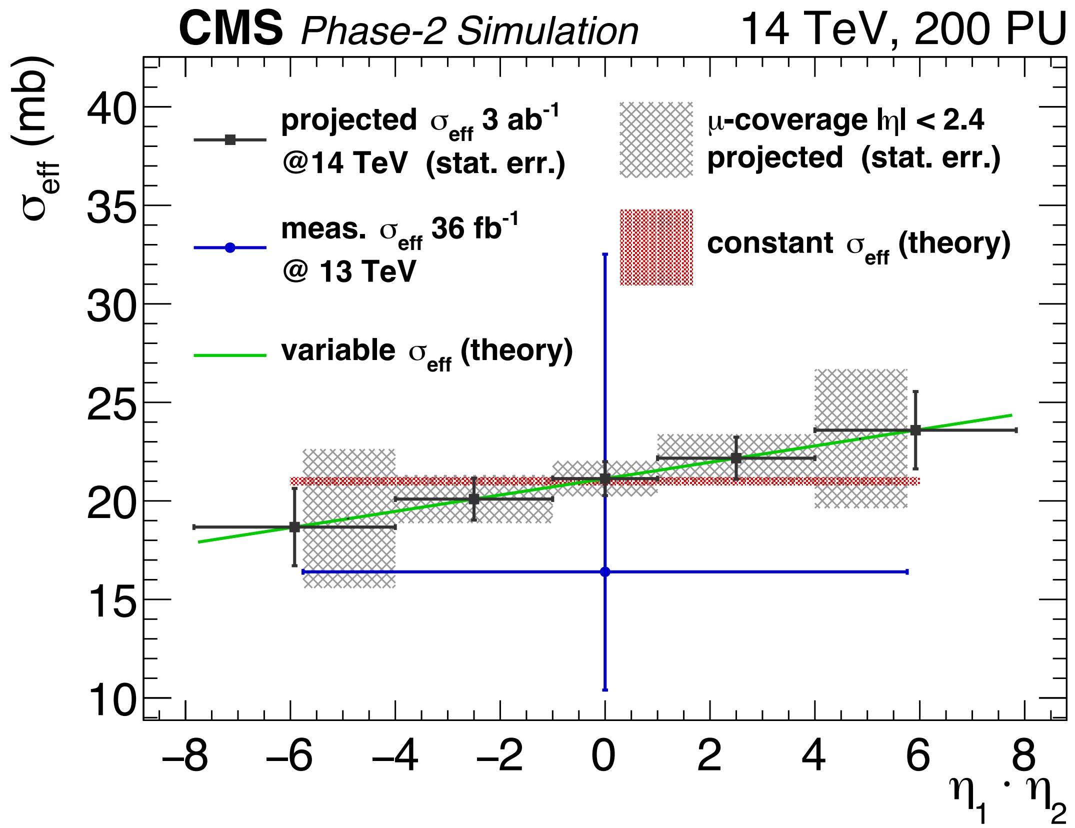
-> roughly 150 fb^{-1} taken out of >3000 fb^{-1}

studied the effect of extended muon coverage of the CMS detector

-> this analysis especially would profit from
CMS trigger upgrade to larger eta values!

reminder: if partons are correlated σ_{eff} will vary with $\eta_1\eta_2$

-> we will be sensitive to this at the latest with HL-LHC!



summary

CMS made public the first evidence for DPS WW production

-> public pages: CMS SMP-18-015 and CDS record and arXiv:1909.06265

-> the paper is submitted to EPJC

we find a sigma effective value of $12.7^{+5.0}_{-2.9}$ mb

-> compatible with past boson induced processes

this is only a first step in using $W^{\pm}W^{\pm}$ production for DPS studies

-> the future holds interesting physics to be extracted

it would be good to have theorists input as to what and how we can test certain more complex theoretical models

-> we're looking forward to talking this week!

the end

marc dünser (CERN)

extras