

АНАЛІЗ ДАНИХ ЕКСПЕРИМЕНТУ CMS, ЗДОБУТИХ У 2018 р., ІЗ МЕТОЮ ПОШУКУ ПОДІЙ ПРЯМОГО НАРОДЖЕННЯ ПАР НАЙЛЕГШИХ ЧАРДЖІНО

С.Т. Лук'яненко, Л.Г. Левчук

ННЦ «Харківський фізико-технічний інститут», Харків, Україна

E-mail: lukyanenko@ipt.kharkov.ua

З метою пошуку сигналу суперсиметрії (SUSY) – прямого народження пар найлегших чарджіно – проведено обробку та аналіз вибірки протон-протонних зіткнень при енергії 13 TeV, що отримано в експерименті CMS на колайдері LHC у 2018 р., у рамках спрощеного розгляду з метою освоєння та відпрацювання методології подібного аналізу даних CMS. Відібрано події з двома протилежно-зарядженими лептонами (електронами та/або мюонами) і великим дисбалансом поперечного імпульсу E_T^{miss} . За цією величиною на основі моделювання методом Монте-Карло визначено та оптимізовано шість сигнальних областей, де значимість сигналу є максимальною. Оцінено численні систематичні похибки, пов'язані з невизначеністю виміру фізичних величин в експерименті, неоднозначністю застосованих процедур відтворення подій та теоретичних розрахунків. Вони були враховані у статистичному аналізі, який виконано на основі M_{T2} -спектрів подій. Кінцевий результат свідчить, що дані CMS задовільно (у межах одного стандартного відхилення) відтворюються моделюванням методом МК. У результаті, шуканий SUSY-сигнал було виключено для широкого діапазону мас найлегших чарджіно і нейтраліно.

PACS: 13.85.-t; 07.05.Kf

Проведено обробку та аналіз вибірок протон-протонних зіткнень при енергії 13 TeV, які одержано в експерименті CMS на колайдері LHC у 2018 р., з метою пошуку сигналу суперсиметрії (SUSY) – прямого народження пар чарджіно з подальшим їх каскадним розпадом $\tilde{\chi}_1^\pm \rightarrow \tilde{l}^\pm \nu_l / l^\pm \tilde{\nu}_l \rightarrow l^\pm \tilde{\chi}_1^0 \nu_l$, тобто на два протилежно-заряджених лептони (електрони та/або мюони), нейтрино та найлегший нейтраліно. Результати отримано за допомогою “універсального” програмного комплексу фізичного аналізу даних CMS LatinoAnalysis [1]. Використовувалася остання версія (“UltraLegacy”) вибірок експериментальних даних, а також вибірок подій, згенерованих методом Монте-Карло (МК).

1. ПОПЕРЕДНЯ ОБРОБКА ДАНИХ

Перед аналізом даних CMS (їх вибірка відповідає інтегральній світимості колайдера $L_{\text{int}}=59.8 \text{ фб}^{-1}$), було проведено обробку відповідних МК-вибірок подій. Спершу до всіх вибірок застосовано “скімінг” – попередній відбір подій, в яких є хоча б один заряджений лептон, що відповідає критеріям “дуже слабко” ідентифікації, з наступним записом вибірок таких подій на сховище даних CERN. Обробка представляє собою покрокові процедури вбудовування у вхідні МК-дані значень L_{int} і поперечного перерізу процесів, що розглядаються; скейл-факторів (SF), які враховують відмінність в ефективності реєстрації b-струменів між реальним експериментом та МК-моделюванням; Rochester-поправок [2], що дозволяють додатково скоригувати 4-імпульси мюонів; SF, що враховують відмінність ефективності реєстрації (реконструкції, ідентифікації та ізоляції) заряджених лептонів в експерименті від МК-моделювання. Обчислено і вбудовано вагові коефіцієнти, за допомогою яких можна перевизначити (методом “PUreweighting” [3]) будь-які розподіли подій з урахуванням ефекту їх накладання – тобто відповідно до

експериментального розподілу подій за кількістю первинних вершин. На останньому кроці розраховано і вбудовано кінематичну величину M_{T2} [4], аналіз розподілів подій за якою є основою нашого дослідження.

МК-вибірки відтворюють умови роботи детектора CMS в 2018 р. лише до сеансу №319077, починаючи з якого відбулося знеструмлення секторів HEM15/16 переднього адронного калориметра. З цієї причини для отримання кращої згоди між моделюванням та експериментом відбраковано події, для яких спостерігається активність у напрямку зазначених секторів. Були використані спеціальні критерії ветування таких подій, які надані у розроблених у CMS рекомендаціях [5] для фізичного аналізу даних.

2. АНАЛІЗ ДАНИХ

2.1. СИГНАЛЬНІ ОБЛАСТІ

На цьому етапі визначаються сигнальні (SR) та контрольні області (CR); для кінематичних величин, що нас цікавлять, будуються розподіли подій у вигляді ROOT-гістограм, на основі яких формуються набори графіків для візуалізації результатів. SR визначають кінематичні області, де бажано мати велику значущість сигналу, тобто максимальне співвідношення сигнал/ $\sqrt{\text{фон}+\text{сигнал}}$. Оптимізацію SR було проведено за дисбалансом поперечного імпульсу $E_T^{\text{miss}}=|\mathbf{P}_T^{\text{miss}}|$ у події і кількістю адронних струменів. Для цього використовувалися МК-вибірки, які відповідають “базовій” сигнальній області: відібрано події з двома ізольованими протилежно-зарядженими лептонами (електронами та/або мюонами) з поперечними імпульсами $p_T > 25 \text{ GeV}$ у разі сублідируючого лептону і псевдорapidіті $|\eta| < 2.4$; на інваріантну масу M_{ll} двох лептонів однакового аромату накладено обмеження $|M_{ll}-M_Z| > 15 \text{ GeV}$ (M_Z – маса Z-бозону), а для довільного аромату $M_{ll} > 20 \text{ GeV}$; відбраковано події, якщо вони містять адронні струмені, мічені b-

кварком; $E_T^{\text{miss}} > 140$ GeV. Такий відбір подій відповідає сигнатурі шуканого сигналу із суттєвою дискримінацією основних фонових внесків. У результаті оптимізації ми отримали розбиття базової сигнальної області на шість SR, як це показано в Табл. 1.

Таблиця 1

Інтервали значень E_T^{miss} і число адронних струменів N_{jets} , що відповідають оптимізованим SR

	SR1	SR2	SR3	SR4	SR5	SR6
E_T^{miss} , GeV	160... 220	160... 220	220... 280	220... 280	280... 380	>380
N_{jets}	≥ 1	0	≥ 1	0	≥ 0	≥ 0

Зауважимо, що з наміром не збільшувати кількість SR, ми не проводили додаткову оптимізацію SR за величиною M_{T2} .

2.2. КОНТРОЛЬНІ ОБЛАСТІ

CR необхідні для перевірки якості відтворення МК-моделюванням експериментальних розподілів фонових подій, які обумовлюються партонними процесами в рамках Стандартної Моделі (CM). Задовільне відтворення дає підґрунтя для оцінки фонових внесків у SR виключно через моделювання і без додаткової корекції M_{T2} -спектрів.

Основні фонові внески обумовлені процесами парного (tt) та поодинокого (tW) утворення t-кварків, двох слабких векторних бозонів (WW, WZ, ZZ) та процесами Дрела-Яна (DY). Також розглянуто рідкісні ("Rare") події від народження трьох бозонів, частинки Хіггса тощо.

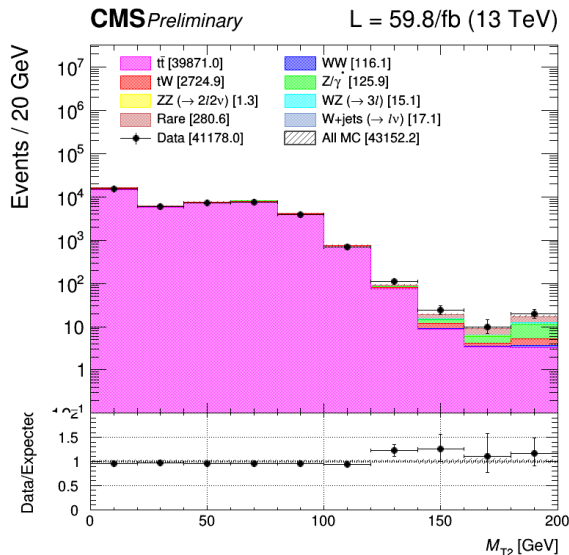


Рис. 1. M_{T2} -спектр подій для контрольної області, яка відрізняється від базової сигнальної області лише зверненням вимоги на відсутність b-струменів у події; заштрихована вузька смуга невизначеності включає статистичну похибку та систематичні похибки, пов'язані лише з визначенням інтегральної світимості та ефективності тригерів (це стосується також Рис. 2 нижче)

У CR, де домінує утворення t-кварків, МК-моделювання задовільно відтворює отриманий в експерименті M_{T2} -спектр (Рис. 1).

Побудова низки CR не може здійснюватися на пряму. Тому виділення деяких подій ґрунтується на методі "штучної вибірки". Він застосовується для імітації народження, наприклад пар $ZZ \rightarrow 2l2\nu$ за допомогою використання додаткової 4-лептонної МК-вибірки подій, що відповідають процесам $ZZ \rightarrow 4l$. Вибираючи в таких подіях випадковим чином Z-бозон, який задовольняє умові $|M_{ll} - M_Z| < 15$ GeV, була побудована дволептонна вибірка шляхом видалення з події двох лептонів від такого Z-бозону та додавання їх сумарного поперечного імпульсу до P_{T}^{miss} події. Далі накладається обмеження $E_T^{\text{miss}} > 140$ GeV. На Рис. 2 наведено M_{T2} -спектри для контрольної області, де домінують події з народженням пар Z-бозонів і розпадом $ZZ \rightarrow 2l2\nu$.

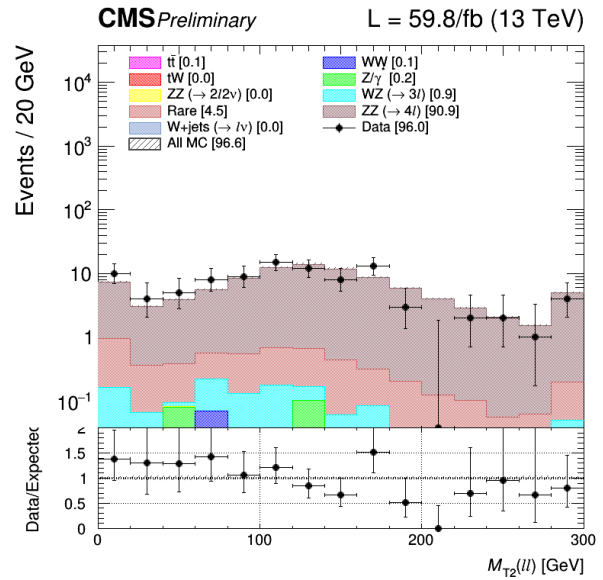


Рис. 2. Спостережувані та очікувані розподіли подій за величиною M_{T2} , отримані методом побудови «штучної» вибірки подій з народженням $ZZ \rightarrow 2l2\nu$ на основі вибірки $ZZ \rightarrow 4l$; отриманій таким чином МК-вибірки відповідає гістограма сірого кольору

Маємо задовільне відтворення методом МК отриманих в експерименті даних, навіть у межах статистичної похибки. Загалом це спостерігається і для інших CR.

2.3. СИСТЕМАТИЧНІ ПОХИБКИ

На етапі аналізу обчислено систематичні похибки для всіх SR. Для цього були сформовані, з урахуванням особливостей набору даних у 2018 р., «альтернативні» МК-вибірки подій, які, на відміну від «номінальних» вибірок, містять фізичні величини, перераховані в рамках невизначеності (звичай, в межах одного стандартного відхилення $\pm\sigma$) вимірювання базових характеристик фізичних об'єктів. Зокрема, були «згенеровані» вибірки з перерахованими значеннями величини M_{T2} .

Основні джерела похибок пов'язані з кінцевим дозволом по енергії адронних струменів (JER), що

реєструються, і неоднозначністю масштабу енергії, як для самих струменів (JES), так і некластеризованих у струмені частинок, що мають малий поперечний імпульс p_T (UnclEn). Значний внесок роблять також теоретичні невизначеності, пов'язані з вибором щільності розподілу партонів за імпульсами (PDF), а також відповідних параметрів перенормування μ_R і факторизації μ_F . Відносно велику похибку дає також застосування методики «PURewighting». Крім того, ми врахували деякі

інші, менш значущі, джерела систематичних похибок. До таких належать неоднозначності визначення інтегральної світимості колайдера, ефективності тригерів та багатьох SF, що враховують різницю між експериментом та МК-модельюванням. Обчислено також похибку, пов'язану з перевизначенням p_T -спектра t-кварків для подій з народженням t-кваркових пар [6]. У Табл. 2 наведено значення основних систематичних похибок.

Таблиця 2

Основні внески в систематичну похибку обчислення очікуваного виходу фонових подій і форми M_{T2} -спектрів для перших чотирьох SR

Джерело невизначеності	SR1		SR2		SR3		SR4	
	Вихід	Форма	Вихід	Форма	Вихід	Форма	Вихід	Форма
JES	4%	4%	8%	14%	7%	5%	4%	6%
JER	2%	9%	2%	5%	2%	4%	1%	5%
UnclEn	1%	6%	3%	10%	2%	5%	2%	6%
PDF	8%	5%	5%	4%	4%	2%	7%	5%
μ_R/μ_F	9%	5%	4%	3%	3%	2%	10%	6%
PURewighting	1%	16%	10%	11%	3%	5%	2%	12%

У разі шуканого сигналу маємо додаткові джерела невизначеності. Вони пов'язані з тим, що сигнальні вибірки формуються за допомогою спрощеного (FASTSIM) моделювання подій, на відміну від повного (FULLSIM) моделювання для «фонових» вибірок. Тут найбільшу похибку може вносити незадовільне відтворення величини E_T^{miss} у разі FASTSIM.

Враховання систематичних похибок дозволяє адекватніше інтерпретувати результати експерименту. Це наочно демонструє Рис. 3.

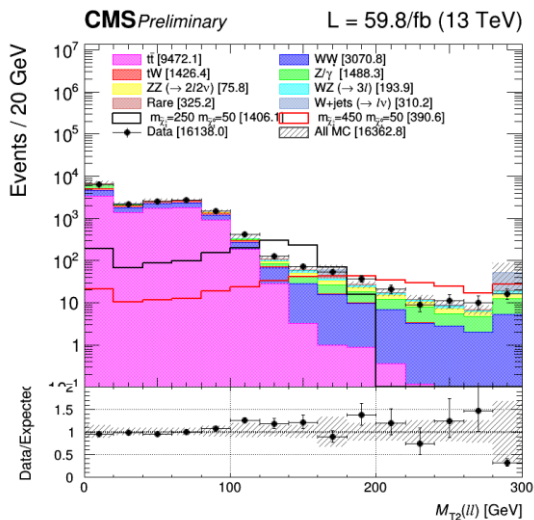


Рис. 3. M_{T2} -спектри подій для базової SR. Червона та чорна лінії показують можливі внески шуканого сигналу при двох різних значеннях (450 та 250 GeV) маси найлегшого чарджіно, смуга невизначеності не враховує похибки для сигналу

Систематична похибка може впливати на висновки щодо наявності надлишку подій в

експерименті (наприклад, в інтервалі $100 \text{ GeV} < M_{T2} < 120 \text{ GeV}$) або, навпаки, переоцінки виходу подій моделюванням методом МК ($M_{T2} \approx 300 \text{ GeV}$).

3. СТАТИСТИЧНИЙ АНАЛІЗ

Після побудови усіх необхідних M_{T2} -спектрів для фонових і сигнальних подій було виконано статистичний аналіз. Для цього використовувався програмний пакет HiggsAnalysis-CombinedLimit [7] та попередньо сформовані для всіх шести SR «карти даних», які містять оцінки систематичних похибок, обумовлених різноманітними їх джерелами. Паке [7] дозволив розрахувати із заданим рівнем достовірності (у нашому випадку $CL=95\%$) верхні межі на силу сигналу $\mu = \sigma / \sigma_{SUSY}$, яка є відношенням отриманого в експерименті поперечного перерізу прямого народження пари найлегших чарджіно до перерізу, обчисленого в рамках розглянутої SUSY-моделі. Застосовано метод AsymptoticLimits, головною перевагою якого є те, що розподіл тестової статистики визначається аналітично в асимптотичному наближенні. Метод обчислює рівень достовірності $CL_s(\mu)$, використовуючи відношення функцій правдоподібності як тестову статистику [8]. Далі були побудовані одновимірні ROOT-діаграми спостережуваних та очікуваних кривих виключення (на них верхні межі на μ дорівнюють одиниці з $CL=95\%$). Нижче кривих лежать області виключення можливих комбінацій мас найлегших чарджіно та нейтраліно (тут верхні межі на μ менше одиниці). Нарешті, формується двовимірна діаграма зі спостережуваними та очікуваними кривими виключення, де для кожної масової точки представлено значення верхньої межі поперечного перерізу шуканого процесу.

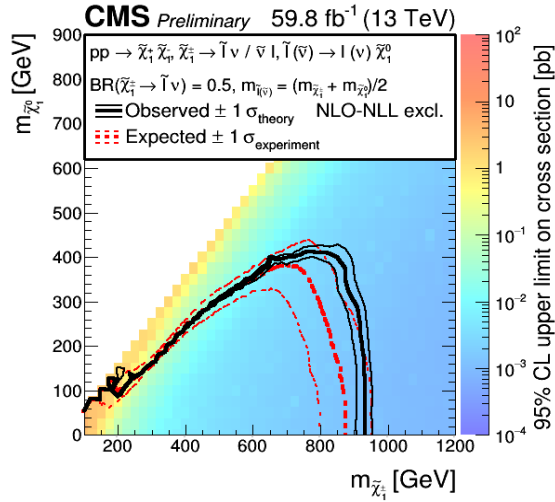


Рис. 4. Области виключення масових точок, а також відповідні їм верхні межі на поперечні перерізи шуканого сигналу: товста штрихова червона лінія – очікувана медіанна крива виключення; тонкі штрихові червоні лінії відповідають невизначеності (статистичній та систематичній) в одне стандартне відхилення ($\pm\sigma$); товста суцільна чорна лінія – спостережувана крива виключення; тонкі суцільні чорні лінії вказують на невизначеність у теоретичних розрахунках перерізу прямого народження пари найлегших чарджіно; на горизонтальній /вертикальній осі відкладено масу чарджіно/нейтраліно

Рис. 4 показує результати виконаного нами статистичного аналізу даних CMS, здобутих у 2018 р. Спостережувана та очікувана криві виключення збігаються з точністю до розрахункової похибки, що вказує на задовільне відтворення даних експерименту CMS розрахунками в рамках CM, а також на відсутність (з рівнем достовірності $CL=95\%$) SUSY-ефектів, обумовлених прямим народженням пар найлегших чарджіно та подальшим утворенням нейтраліно, для широкого діапазону мас цих частинок – $100 \text{ GeV} \leq M_{\text{ch}} \leq 930 \text{ GeV}$ і $1 \text{ GeV} \leq M_{\text{neu}} \leq 410 \text{ GeV}$ відповідно.

ANALYSIS OF 2018 CMS DATA AIMED AT SEARCH FOR EVENTS OF DIRECT CHARGINO PAIR PRODUCTION

S.T. Lukyanenko, L.G. Levchuk

A sample of 13 TeV proton-proton collisions recorded in the CMS experiment at the LHC for 2018 was processed and analyzed to search for a supersymmetry (SUSY) signal, the direct lightest chargino pair production, within a simplified consideration aimed at gaining an experience in usage of CMS analysis tools. Events with two oppositely charged leptons (electrons and/or muons) and a large missing transverse momentum E_T^{miss} were selected. Based on E_T^{miss} , six signal regions with maximum signal significance have been determined and optimized with Monte Carlo (MC) simulation. Estimates have been obtained for numerous systematic uncertainties due to degree of precision in measurements of physical quantities in the experiment, ambiguities of procedures applied for event reconstruction and of theoretical calculations. They were taken into account in a statistical analysis, which was performed through examining M_{T2} spectra for the events. The final result suggests that the CMS data are satisfactorily (within one standard deviation) reproduced by the MC simulation. As a result, the SUSY signal searched for has been excluded across a wide range of the lightest chargino and neutralino masses.

Насамкінець, підкреслимо, що головними цілями цієї роботи були освоєння та відпрацювання методів і засобів фізичного аналізу даних експерименту CMS (зокрема, комплексу [1]), які використовувалися при пошуку сигналу прямого народження пар чарджіно на основі аналізу даних CMS, деталі та результати якого представлено в роботі [9]. При цьому ми обмежилися спрощеним (порівняно з аналізом [9]) розглядом із значно меншою кількістю оптимізованих SR та скороченою оцінкою систематичних невизначеностей.

БІБЛІОГРАФІЧНИЙ СПИСОК

1. LatinoAnalysis, <https://github.com/latinos/LatinoAnalysis>.
2. RochesterCorrections, <https://twiki.cern.ch/twiki/bin/view/CMS/RochcorMuon>
3. PileupReweighting, <https://twiki.cern.ch/twiki/bin/view/CMS/PileupMCReweightingUtilities>.
4. C.G. Lester and D.J. Summers. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders // *Phys. Lett. B.* 463 (1999) 99.
5. SUSYRecommendations, <https://twiki.cern.ch/twiki/bin/view/CMS/SUSRecommendationsRun2UltraLegacy>.
6. TopPtReweighting, <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting>
7. HiggsAnalysis-CombinedLimit, <https://github.com/cms-/HiggsAnalysis-CombinedLimit>.
8. ATLAS and CMS Collaborations, LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in Summer 2011 // *CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11* (2011).
9. CMS Collaboration, Search for supersymmetric particle pair production in final states with two oppositely charged leptons and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ // *CMS Physics Analysis Summary CMS-PAS-SUS-23-002* (2024).