

Kraków, dnia 26 sierpnia 2020 r.

Prof. dr hab. Stanisław Drożdż
Instytut Fizyki Jądrowej PAN Kraków
oraz Wydział Informatyki i Telekomunikacji
Politechnika Krakowska

Ocena rozprawy doktorskiej
„Eigenvectors of random matrices: theory and applications ”
Pana mgr Wojciecha Tarnowskiego

Tematyka rozprawy doktorskiej Pana mgr Wojciecha Tarnowskiego wpisuje się w jedno z centralnych zagadnień współczesnej fizyki jakim jest opis, modelowanie i systematyka występujących w Naturze struktur złożonych przy użyciu koncepcji wywodzących się z teorii macierzy losowych. Stosowalność tych koncepcji okazuje się być nawet bardziej uniwersalna o czym świadczy choćby fakt, że obecny już w tytule rozprawy termin „random matrices” jest coraz częściej obecny nie tylko w fizyce ale również wszędzie tam gdzie rozważany układ składa się z dużej liczby wzajemnie oddziałujących elementów, a więc w tak rozległych dziedzinach jak ekonomia i finanse, neurologia, czy nawet nauki społeczne. Jeśli chodzi o praktyczne zastosowania to w rozważanej tu rozprawie w zasadzie wszystkie te wątki się pojawiają, a szczególnie wyeksponowany jest ten najbardziej nowatorski, skierowany na modelowanie neurologiczne.

W skład przedstawionej rozprawy wchodzi siedem współautorskich oryginalnych artykułów Pana Wojciecha Tarnowskiego opublikowanych w renomowanych światowych periodykach naukowych, z czego cztery są dwuautorskie, jeden czteroautorski a dwa pięcio-autorskie. Jest oczywiste, że udział doktoranta w tych artykułach musiał być dominujący a co najmniej znaczący. Ten fakt ma także formalne świadectwo w postaci załączonych do dokumentacji oświadczeń współautorów. Zestaw kopii tych artykułów jest poprzedzony wstępem oraz sześcioma kilkustronicowymi rozdziałami, też napisanymi po angielsku, w zasadzie stanowiącymi kilkustronicowe streszczenia poszczególnych artykułów, o których mowa powyżej. Centralnym wątkiem rozprawy jest kwestia różnorodnych zagadnień związanych z powszechnie występującymi efektami nieortogonalności wektorów własnych niehermitowskich macierzy przypadkowych.

W rozdziałach tych, po kolei, przedstawiony jest, znany wcześniej w literaturze, formalizm obliczania jednopunktowej funkcji korelacji przy użyciu metod diagramów Feynmana oraz swobodną probabilistykę (ang. *free probability*).

Rozważany jest tu w szczególności przypadek zespołów biunitarnie niezmienniczych, dla których obliczenia znacząco się upraszczają. Najważniejszy w tym kontekście wynik to rozszerzenie [A1,A2] twierdzenia Haagerupa-Larsena o spektrach macierzy przypadkowych na jednopunktową funkcję korelacji wektorów własnych.

Dyskutowana jest następnie analiza dwupunktowej funkcji korelacji dla unitarnie niezmienniczych rozkładów prawdopodobieństwa w granicy dużego rozmiaru macierzy. Przedstawiony tam [A3] formalizm jest znaczącym rozszerzeniem wcześniej znanych w literaturze wyników dla modeli gaussowskich a uzyskane równania typu Bethego-Salpetera pozwalają na znalezienie dwupunktowej funkcji korelacji dla dowolnego rozkładu macierzowego. Funkcja ta wyrażona jest przez jednopunktową funkcję Greena oraz przez funkcję generującą kumulanty. Ze względu na uproszczoną strukturę kumulant dla rozkładów biunitarnie niezmienniczych podana jest jawna formuła wyrażająca dwupunktową funkcję za pomocą funkcji jednopunktowej oraz radialnej dystrybuanty wartości własnych. W efekcie, jedno- i dwupunktowa funkcja korelacji jest związana z warunkową wartością oczekiwaną elementów macierzy przekrycia. Jest to znaczący wynik.

Następny przedstawiony ważny wynik dotyczy łącznej gęstości prawdopodobieństwa dla wartości własnej i stowarzyszonego z nią diagonalnego elementu macierzy przekrycia. Obliczenia przeprowadzone są dla tak zwanego rzeczywistego zespołu eliptycznego, który pozwala na ciągłą interpolację pomiędzy macierzami symetrycznymi a całkowicie asymetrycznymi. Głównym wynikiem jest wzór na to prawdopodobieństwo otrzymany dla skończonego rozkładu macierzy. Jego asymptotyczna analiza pozwoliła na znalezienie trzech obszarów, charakteryzujących się nieco innym zachowaniem wektorów własnych. W tym kontekście najciekawszym wynikiem pracy [A4] jest odkrycie obszaru, w którym macierz jest prawie symetryczna a jej stopień asymetrii maleje wraz ze wzrostem rozmiaru macierzy. Zachowanie wartości własnych w takim obszarze zostało udokumentowane w literaturze ponad dwie dekady temu i ten obszar został nazwany słabą niehermitowskością, ale dopiero w pracach Tarnowskiego pokazane zostało, że parametr asymetrii wpływa także na wektory własne i dlatego ten obszar został nazwany słabą nienormalnością (ang. *weak nonnormality*).

Ostatnie z zaprezentowanych zagadnień teoretycznych dotyczy techniki zaproponowanej przez Borodina i Olshanskiego, opierającej się na spostrzeżeniu, że łączna gęstość prawdopodobieństwa wszystkich wartości własnych w unitarnie niezmienniczych rozkładach hermitowskich macierzy jest dana przez kwadrat wyznacznika Slatera, co od razu sugeruje analogię kwantowo-mechaniczną. Stąd narzucenie mikroskopowego skalowania znacznie upraszcza równanie Schroedingera, pozwalając nieraz szybko odczytać jak powinna wyglądać struktura jądra korelacji. Jednym z wyników pracy [A5] omawianej w tym zestawieniu jest obserwacja, że trzy klasyczne mikroskopowe uniwersalności związane są transformatami całkowitymi (Fouriera, Airy'ego i Hankela), a jądro dane jest przez obcięcie zakresu całkowania. Zauważono także, że ten obraz przenosi się łatwo na tzw. zespoły biortogonalne, w których łączne prawdopodobieństwo jest iloczynem

dwóch wyznaczników Slatera ale zbudowanych z różnych funkcji. Struktura biortogonalna tych funkcji jest podobna do biortogonalnych wektorów własnych macierzy niehermitowskich i okazuje się, że te funkcje można interpretować jako 'lewe' i 'prawe' funkcje własne niehermitowskiego 'hamiltonianu'.

W ostatnich dwóch, bardziej aplikacyjnych rozdziałach, prezentowane są zastosowania elementów powyższego formalizmu w kontekście sieci neuronowych. Rozważany jest więc model biologicznej sieci neuronowej, który uwzględnia dwa rodzaje neuronów – wzmacniające i wygaszające – oraz odzwierciedla regułę Dale'a mówiącą, że wszystkie synapsy wychodzące z neuronu albo wygaszają albo wzmacniają a nie ma mieszanego działania. O dynamice takiego modelu decydują wartości własne macierzy połączeń synaptycznych, które w tym przypadku układają się wewnątrz dysku z kilkoma znacznie poza nim. Dodanie warunku balansu między wzmacnianiem i wygaszaniem neuronów ściąga wszystkie wartości własne do dysku, sprawiając, że dynamika przestaje być chaotyczna. Jako nowatorski wynik w pracy [A6] zauważono, że struktura z neuronami wygaszającymi i wzmacniającymi uwzględniająca regułę Dale'a, w połączeniu z warunkiem równowagi między wzmocnieniem a wygaszaniem, jest źródłem silnej nieortogonalności wektorów własnych tej macierzy. Jest to związane z dużą czułością spektrum macierzy na zmiany wag synaps co ma miejsce w procesie uczenia sieci oraz jest źródłem silnych efektów przejściowych w dynamice zlinearyzowanej.

Jako ostatnie z dyskutowanych zagadnień przedstawiony jest problem znikających gradientów w uczeniu głębokich sztucznych sieci neuronowych (ang. *deep learning*). Problem ten jest analizowany poprzez jacobian wejście-wyjście i rozkład jego wartości osobliwych (ang. *singular values*). W idealnym przypadku wartości własne tego jacobianu powinny być wszystkie równe jedności, natomiast problem zanikających gradientów występuje gdy wartości osobliwe zbliżają się do zera wraz z dodawaniem kolejnych warstw. Problem tzw. wybuchających gradientów pojawia się natomiast gdy wartości osobliwe rosną szybko razem z głębokością takiej sieci. W pracy [A7], przy użyciu technik analitycznych wypracowanych w [A1], pokazane zostało, że dla architektury rezydualnej wartości osobliwe koncentrują się w pobliżu jedności w sposób uniwersalny – ich gęstość jest opisywana przez jeden parametr odpowiedzialny za całą zależność od inicjalizacji wag sieci oraz od funkcji aktywacji. Ten wynik w sposób ilościowy pokazuje dlaczego sieci o architekturze rezydualnej są łatwiejsze do trenowania. Wyniki te są także weryfikowane w eksperymentach numerycznych w problemie klasyfikacji obrazów.

Tradycja formułowania recenzji rozpraw doktorskich w Polsce jako nieodłączny element wskazuje również potrzebę przedstawienia uwag krytycznych. W przypadku rozprawy Pana Tarnowskiego nie jest to łatwe bo przedstawiony materiał jest wybitnie dobry a na dodatek już opublikowany w czasopiśmie recenzowanych. Moja zasadnicza uwaga typu krytycznego dotyczyłaby więc sposobu redagowania tej rozprawy. Nie bardzo widzę uzasadnienia dla pisania po angielsku rozdziałów w postaci streszczeń załączonych artykułów. Przecież cały ten materiał, jako

opublikowany po angielsku, jest już dostępny dla wszystkich zainteresowanych na świecie. Napisanie po polsku byłoby znakomitą okazją do ugruntowania polskich odpowiedników często bardzo specjalistycznych i nowych terminów używanych w tym obszarze badań. Sam podczas pisania tej opinii po polsku doświadczyłem tu pewnego tego typu zakłopotania. Na szczęście prywatnymi kanałami mogłem się skutecznie w tym temacie skonsultować z Panem Wojciechem Tarnowskim dzięki temu, że znamy się osobiście od kilku lat ponieważ jako student miesięczną praktykę odbywał w Zakładzie Teorii Systemów Złożonych, którym kieruję w IFJ PAN. Pewien niedosyt u mnie wywołuje również fakt, że klarowna ilustracja interpretacji składowych wektorów własnych jest przedstawiona tylko dla macierzy symetrycznych i to na dodatek wzięta z obcego źródła (Rys.1.1). Z przyjemnością wśród cytowanych widzę artykuł mojej grupy (Acta Physica Polonica B 37 (2006) 3039-3049), gdzie zastosowane są asymetryczne macierze korelacji do badania opóźnionych korelacji na rynkach finansowych. Niestety podany w rozprawie numer (B36) tomu jest błędny. Nie sprawdzałem na ile tego typu błędy mają miejsce w innych pozycjach. Niedopatrzaniem wydaje mi się też brak jakiegokolwiek w tej rozprawie wzmianki o artykule, jaki jeszcze znacznie wcześniej, bo już ponad dwie dekady temu, powstał w Krakowie, opublikowany w Physical Review E64 (2000) 5557-5564, w którym, w terminologii macierzy przypadkowych, badane są korelacje sygnałów elektromagnetycznych pochodzących z mózgu. Co więcej, w kontekście badania korelacji opóźnionych, skonstruowana jest tam, wydaje się po raz pierwszy w literaturze, asymetryczna macierz korelacji a jej spektrum wartości własnych zilustrowane na płaszczyźnie zespolonej. Nieprecyzyjne jest natomiast cytowanie książki Ch.E. Portera z 1965 roku ([93]) jako właściwego źródła dla kwantowego chaosu. Tamte lata to dopiero początki formalnego rozwoju teorii klasycznego chaosu a tak zwany kwantowy chaos to dopiero lata 80, wszystko wieku XX.

Podsumowując, materiał przedstawiony przez Pana mgr Wojciecha Tarnowskiego jako rozprawa doktorska zawiera wielorakość interesujących oraz bardzo wartościowych wyników o implikacjach interdyscyplinarnych, na których zgłębianie chętnie poświęciłbym znacznie więcej czasu niż pozwalają na to różne ograniczenia, a przede wszystkim ustawowe limity czasowe na przygotowanie recenzji. Osobiście jestem entuzjastą tego typu aktywności naukowej. Jestem w pełni przekonany, że rozprawa ta z wielkim powodzeniem spełnia wszelkie warunki stawiane tego typu rozprawom i wnoszę o dopuszczenie Pana magistra Wojciecha Tarnowskiego do dalszych etapów przewodu doktorskiego. Co więcej, zgodnie z załączonym dodatkowym uzasadnieniem, stawiam również wniosek o wyróżnienie tej rozprawy.



Stanisław Drożdż