Warszawa, 9 czerwca 2022

prof. dr hab. Jacek Wesołowski
Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska

Recenzja rozprawy habilitacyjnej oraz dorobku naukowego dr. Wojciecha Rejchela

Uwagi wstępne. Dr Wojciech Reichel jest absolwentem Wydziału Matematyki i Informatyki Uniwersytetu Mikołaja Kopernika w Toruniu, gdzie w 2005 roku obronił pracę magisterską dotyczącą estymacji gęstości przygotowaną pod kierunkiem prof. T. Rychlika, a w 2011 roku obronił doktorat (z wyróżnieniem) na temat statystycznych modeli regresji rangowej przygotowany pod opieką prof. W. Niemiro i od tego roku pracuje jako adiunkt na swoim macierzystym Wydziale.

Badania naukowe habilitanta koncentrują się na zastosowaniach matematyki, a konkretniej na metodach statystyki i uczenia maszynowego. Przedstawiona do oceny rozprawa habilitacyjna *M-estymatory z karą w wyborze modelu* również dotyczy zagadnień mieszczących się w tym obszarze nauki: dokładniej, chodzi o metody selekcji i estymacji parametrów w przypadku, gdy ich liczba w prawdziwym modelu jest mniejsza niż w modelu wyjściowym. Celem badań jest zaproponowanie nowych algorytmów mających służyć rozwiązaniu tego typu zadań, a twierdzenia matematyczne, czasem całkiem ciekawe, stanowią punkt wyjścia rozumowań (często heurystycznych) mających prowadzić do konkretnych konstrukcji (a czasem i do ich implementacji). O służebnej roli matematyki w badaniach prowadzonych przez habilitanta, pośrednio świadczy fakt, że dowody twierdzeń (czyli matematyka) z reguły znajdują się nie w głównych tekstach artykułów, ale w *appendixach*. To może nie podobać się matematycznym purystom, ale nie sposób w rozwoju badań naukowych przecenić roli matematyków podejmujących wyzwania, których głównym celem jest stworzenie narzędzi badawczych służącym różnorakim potrzebom innych gałęzi nauki. W mojej opinii stanowi to ważną wartość dodaną zarówno osiągnięcia habilitacyjnego jak i całego dorobku habilitanta.

Osiągniecie habilitacyjne. W ramach osiągnięcia habilitacyjnego proponowane i analizowane są różne algorytmy związane z tzw. M-estymatorami z karą typu lasso. Istnieje pokaźna literatura dotycząca analizy M-estymatorów, w tym ich asymptotyki. Poważny dorobek w tej dziedzinie ma promotor doktoratu habilitanta, W. Niemiro. Z drugiej strony, od przełomowej pracy Tibshirani (1996), estymatory z karą lasso są szeroko badane jako narzędzia redukcji modelu. Habilitant bardzo dobrze opanował te tematy i znakomicie się w nich porusza. Spośród ośmiu prac przedstawionych w ramach osiągnięcia habilitacyjnego najciekawsze są wyniki pracy A1, w której zaproponowano i przeanalizowano dwukrokową procedurę SS, ulepszającą algorytm SOS wprowadzony we wcześniejszej pracy J. Mielniczuka i P. Pokarowskiego (którzy są też współautorami pracy A1) oraz pracy A3 (wspólnej z K. Furmańczykiem), w której zaproponowano i przeanalizowano procedurę SD opartą na progowym lasso i pomysłach pochodzących z teorii wielokrotnego testowania. W obu tych pracach, ciekawych z punktu widzenia statystyki (czy uczenia maszynowego),

JW

wyzwania matematyczny znacząco wychodziły poza schematy dowodowe z pozostałych prac. Na większą uwagę zasługuje też ogólny pomysł pracy A5 polegający na zastosowaniu estymacji z karą lasso do analizy modelu Isinga, przy czym trudności techniczne powodowane niejawną postacią stałej normalizacyjnej próbuje się pokonywać metodami MCMC (podejrzewam, że w tej części pracy istotną rolę odegrał współautor B. Miasojedow).

Omówię teraz w porządku "historycznym" prace wchodzące w skład osiągnięcia habilitacyjnego.

A8: Praca dotyczy analizy esymatorów typu lasso. Zadanie polega na minimalizacji empirycznej wersji ryzyka z karą typu ℓ_1, w klasie wypukłych funkcji straty. Główny wynik to twierdzenie dotyczące asymptotyki estymatora lasso (Th. 2.1), który uogólnia/uzupełnia znane z literatury twierdzenia dla modeli liniowych i funkcji straty typu ℓ_2 lub ℓ_1. Wyniki potwierdzają dobrze rozpoznane problemy estymatorów typu lasso, tzn. trudności z spełnianiem warunków zgodności selekcyjnej lub zgodności estymacyjnej. Dzieje się tak również w ogólnej sytuacji wypukłej funkcji straty. Potwierdzenie również znajduje fakt znany dla kwadratowej funkcji straty, mówiący o tym, że tak zwane lasso adaptacyjne spełnia warunki "wyroczni", czyli zachodzi zgodność selekcyjna oraz asymptotyka gaussowska ze standardową normalizacją rzędu $\sqrt{n}$. Tego dotyczy drugi główny wynik matematyczny pracy, Th. 3.1. Rozważa się asymptotykę, wedle której liczba obserwacji dąży do nieskończoności a liczba parametrów pozostaje stała. Metodologia matematyczna w dużym stopniu wykorzystuje, odpowiednio je modyfikując, argumenty wypracowane dla M-estymatorów z kanonicznego zestawu prac Polard (1984), Niemiro (1992, 1993) i Geyer (1994).

A7: Praca dotyczy naturalnej reprezentacji ryzyka w zadaniu rangowania dwóch zmiennych. Ryzyko zależy od funkcji f z ustalonej klasy F. Optymalna funkcja f_0 to funkcja, która minimalizuje ryzyko w klasie F. Szukany jest estymator funkcji f_0 w klasie $\{f_\theta,\ \theta \in \Theta\}$, kombinacji liniowych skończonej rodziny funkcji bazowych. W odróżnieniu od pracy B7 z pozostałego dorobku habilitanta, w której podjęte jest podobne zagadnienia ale bez kary, estymator $f_{\hat{\theta}}$ otrzymywany jest na podstawie estymatora $\hat{\theta}$ minimalizującego w tej klasie ryzyko z karą typu lasso. Chodzi zatem jednocześnie o wybór niezerowych współczynników kombinacji liniowej jaki i ich estymator $\hat{\theta}$. Najciekawsze i nietrywialne twierdzenie (Th. 1) podaje nierówności probabilistyczne na tzw. nadmiarowe ryzyko (*excess risk*), czyli różnicę między ryzykiem dla funkcji $f_{\hat{\theta}}$ i ryzykiem minimalnym realizowanym przez f_0 oraz odległość w sensie ℓ_1 estymatora $\hat{\theta}$ i wartości θ^*, która poprawnie wybiera niezerowe współczynniki kombinacji liniowej. Można ja traktować jako analog Th. 1 z pracy B7. Twierdzenie to wykorzystuje ciekawą nierówność dla supremum modułu różnic ryzyka empirycznego i teoretycznego dla f_θ i f_{θ^*}. Dowód (oczywiście zawarty w dodatku) nie jest trywialny i świadczy o dobrym opanowaniu zaawansowanych metod probabilistycznych, w tym nierówności koncentracyjnych dla U-procesów, procesów empirycznych, czy dekompozycji Hoeffdinga dla U-statystyk.

A6: Praca dotyczy zagadnienia wyboru modelu gdy ryzyko empiryczne jest U-statystyką (od dwóch zmiennych), która zarazem jest wypukłą funkcją estymowanego parametru, przy czym stosowana jest kara typu lasso. Praca powtarza w znaczącym stopniu schemat pracy A8. Dowody głównych twierdzeń (Th. 4.1 i Th. 4.3) są w znaczącym stopniu powtórzeniem, często dokładnym, dowodów z A8; oczywiście centralne twierdzenie graniczne dla zmiennych losowych niezależnych zastępuje się jego odpowiednikami dla U-statystyk. To prowadzi jednak do pewnych niejasności: W Th. 4.1 implicite przyjmuje się założenie, że $\mathbb{V}\mathrm{ar}\, \mathbb{E}(\phi(\theta, Z_1, Z_2)|Z_1) > 0$. Wiadomo, że gdy $\mathbb{V}\mathrm{ar}(\mathbb{E}(\phi(\theta, Z_1, Z_2)|Z_1) =$

1.4

0 asymptotyka U-statystyk (z normowaniem n zamiast $\sqrt{n}$) jest typu chi-kwadrat, a nie typu gaussowskiego. To założenie nie jest niezbędne w przypadku Th. 4.3, ale w jego dowodzie mówi się wprost jedynie o asymptotyce (gaussowskiej) z normowaniem $\sqrt{n}$, jakby znowu implicite przyjmowane było wspomniane wyżej założenie. Wydaje się, że przy asymptotykce chi-kwadrat można nawet nieco wzmocnić wynik zakładając $n\lambda_n \to \infty$ zamiast $\sqrt{n}\lambda_n$.

Nowe pomysły matematyczne pojawiąją sie w dowodzie dwukrotnej różniczkowalości funkcji ryzyka dla problemu rangowania, który mieści się w rozważanym schemacie U-statystyk, ale spełnianie standardowych warunków regularnościowych (warunki (1), (2) i (3) z Sec.2) wymaga dodatkowych założeń (Prop. 4.4) i właśnie wspomnianej gładkości funkcji ryzyka odpowiedniej postaci. Wyniki otrzymane w pracy są ciekawe z punktu widzenia statystyki jako takiej. Badanie U-statystyk z karami typu lasso nie jest zagadnieniem często podejmowanym w literaturze, a otrzymane wyniki w szczególności znacząco ulepszają znane wyniki z pracy Song i Ma (2010).

A5: Punktem wyjścia pracy jest ciekawy pomysł wykorzystania estymatorów z karą lasso do analizy modelu Isinga. Pomysł polega na tym, że identyfikację krawędzi grafu w modelu Isinga można rozumieć jako wybór "istotnych" parametrów modelu, czyli zgodność selekcyjną estymatora typu lasso. Jednak nie udaje się w ten sposób ominąć powszechnie znanego problemu technicznego w modelu Isinga, braku zamkniętego wzoru na stałą normującą. Wzorując się na wcześniejszych pracach (wspólnych z B. Miasojedowem, W. Niemiro i J. Palczewskim), stałą tę estymuje się za pomocą niezależnej procedury MCMC z odpowiednio dobranym rozkładem referencyjnym. Tu z pomocą przyszedł współautor, B. Miasojedow, który jest ekspertem w dziedzinie metod MCMC. Główny matematyczny wynik pracy, Th. 2, podaje probabilistyczne oszacowanie normy $|\cdot|_\infty$ różnicy estymatora największej wiarygodności wykorzystującego MCMC do estymacji stałej normującej z karą typu lasso od prawdziwej wartości parametru, w terminach wielkości związanych z procedurą MCMC oraz z prawdziwym modelem, przy czym współczynnik λ kary lasso jest wybierany w bardzo szczególny sposób. Dowód tego twierdzenia, rozłożony na kilka lematów, jest żmudny, i rozwija metodologię pochodzącą z pracy Huang, Sun, Ying, Yu i Zhang z 2013. W części dotyczącej MCMC wkorzystuje nierówność typu Hoeffdinga pochodzącą z wcześniejszej pracy współautora. Najważniejszy wynik pracy to nowa procedura (i jej implementacja), w której autorzy proponują "optymalny" dobór stałej λ, jak i rozkładów referencyjnych w kolejnych krokach procedury MCMC (są też rozkładami typu Isinga o odpowiednio wybieranych parametrach). Problem polega na "niekonstruktywności" ograniczeń na liczbę obserwacji oraz na liczbę iteracji w próbniku Gibbsa (z losowym wyborem aktualizowanej zmiennej w każdym kroku, ang. *random scan Gibbs sampler*). Poprawienie na poziomie teoretycznym tego niedomagania wydaje się zadaniem beznadziejnym. Oczywiście, ograniczenie na liczbę iteracji w próbniku Gibbsa jest w praktyce daleko mniej groźne niż zastrzeżenie na minimalną liczebność próbki.

A4: W pracy zaproponowany jest nowy algorytm estymacyjny typu lasso dla modelu liniowego. Polega on na wykorzystaniu progowego lasso do identyfikcaji wstępnego zbioru parametrów (wiadomo, że takie ograniczenie z dużym prawdopodobieństwem jest nadzbiorem zbioru paramterów istotnych). Następnie konstruowana jest sekwencja nierosnących statystyk i progów (wykorzystujących to wstępne ograniczenie), a ostateczny zbiór zmiennych istotnych polega na wyborze tych, dla których odpowiednie progi są przekraczane. Główny wynik matematyczny (Th. 3.1) podaje ograniczenie na prawdopodobieństwo niepoprawnego wyboru istotnych parametrów. Istotnym założeniem w tym twierdzeniu jest dwustronna nierówność na współczynnik λ kary lasso. Autorzy, jako przewagę proponowanej metody, wskazują na jej "konstruktywność", mającą polegać na tym, że dolne ograniczenie na λ za-

1. 4

leży jedynie od znanych parametrów modelu σ, p, n. Pomijając fakt, że znajomość σ jest założeniem optymistycznym, anonsowana "konstruktywność" budzi wątpliwości (choć nie są to wątpliwości natury matematycznej). Mianowicie, żeby można było przyjąć "kostruktywnie", jak autorzy sugerują, λ równe lewej stronie nierówności (10) i korzystać z wyniku Th. 3.1, trzeba wiedzieć, że prawa strona w (10) jest nie większa od lewej strony, a tego nie znając zbioru I_0 indeksów parametrów istotnych, nie sposób stwierdzić. Podobne zastrzeżenie dotyczy założenia (11). Tym niemniej, Th. 3.1 wskazuje jak można próbować działać w konkretnych sytuacjach i przedstawione w pracy eskeprymenty numeryczne pokazują, że procedura może dawać dobre wyniki dla konkretnych zbiorów danych, w szczególności w porównaniu z innymi metodami. Dowód Th. 3.1, przedstawiony w dodatku, składa się z dwóch części. Pierwsza, łatwiejsza, dotycząca ograniczenia wynikającego z zastosowania progowego lasso w dwóch pierwszych krokach algorytmu, wykorzystuje w istotny sposób nierówność z pracy Ye i Zhang (2010). Druga, trudniejsza, dotyczy trzeciego kroku algorytmu. Tu dowód wymagał sporego wysiłku i pomysłowości. W szczególności, ciekawym pomysłem było pozbycie się zależności od losowego wstępnego zbioru parametrów istotnych S, co pozwoliło na wykorzystanie założenia o podgaussowskości obserwacji. Niestety, w tej części dowodu autorzy w wielu miejscach stosują mylącą notację odwołując się do zmiennej $\hat{\beta}_j^{LS}$, wprowadzonej we wzorze (4), która zależy od S, podczas gdy chodzi im o zmienną losową o podobnej konstrukcji, ale z losowym S zamienionym na nielosowe $I_0 \cup J$.

A3: Praca dotyczy statystyk liniowych w zastosowaniu do klasyfikacji binarnej. Badane są dwa zagadnienia: (1) oszacowanie ryzyka nadmiarowego, tzn. różnicy między ryzykiem estymatora z karą lasso i minimalnym ryzykiem wybijanym przez tzw. estymator bayesowski; (2) problem selekcji modelu, w którym $\eta(x) = \mathbb{P}(Y = 1|X = x)$ zależy (w nieznany sposób) od statystyki liniowej. Ograniczenie górne na ryzyko nadmiarowe w przypadku kwadratowej funkcji straty podane jest w Th. 2(a), a w przypadku straty logistycznej w Th. 2(b). Twierdzenie dotyczy sytuacji gdy predyktory są podgaussowskie i mają ciągłą gęstość na pewnym przedziale ograniczonym. W obu wzorach istotną rolę pełni $\mathbb{P}(|\hat{b} - b_*|_\infty > c)$ odpowiadające za "stopień" zgodności estymatora $\hat{b}$. Oszacowaniu $|\hat{b} - b_*|_q$ z dużym prawdopodobieństwem dla dowolnego $q \geq 1$ i kwadratowej funkcji straty poświęcone jest Th. 3. Drugie zagadnienie, czyli zgodność selekcyjna badana jest w zasadzie przy założeniu eliptyczności rozkładu predyktorów, które ukryte jest w liniowości regresji form liniowych. W tym kontekście, nie rozumiem, co autorzy mają na myśli, gdy przytaczając aprobująco cytat z literatury, piszą, że *it is a non-restrictive condition*. Główny wynik tej części pracy to selekcyjna zgodność estymatora wykorzystującego progowe lasso z niekonstruktywnymi ograniczeniami na próg. Dowody, choć nie wnoszą specjalnie nowych pomysłów matematycznych, są świadectwem bardzo dobrego opanowania probabilistyczno-analitycznych metod wykorzystywanych w analizie M-estymatorów i etymatorów typu lasso.

A2: Praca dotyczy procedury lasso dla rang obserwacji w modelu z monotoniczną funkcją wiążącą i predykcją liniową. Na uwagę zasługuje brak założeń o postaci funkcji wiążącej oraz o momentach błędów losowych. Zasadniczo badana jest sytuacja gdy liczba predyktorów jest większa od liczby obserwacji. Przy dość standardowych założeniach o rozkładzie predyktorów pokazano (Th. 1), że oba zbiory istotnych predyktorów: w modelu liniowym i modelu lasso dla rang są identyczne, co uzasadnia rozważanie takiej procedury. Kolejny wynik dotyczy już samej estymacji i pokazuje zgodność estymacyjną rangowego lasso. Sam wynik i dowód są bardzo podobne do omówionego w poprzednim punkcie Th. 3. Dowód prawie dosłownie przytacza rozumowania z A3. W zasadzie jedyna poważniejsza różnica polega na szacowaniu wyrażeń we wzorze (39); zob. odpowiadający temu wyrażeniu wzór (A15) w pracy A3. Więc z punktu widzenia atrakcyjności matematycznej nieco za dużo tu powtórzeń. Znacznie bardziej interesujący jest dowód Th. 7 dotyczącego zgodności selekcyjnej i estymacyjnej ważonej rangowej procedury lasso z wagami zależnymi od estymatorów otrzymanych za pomocą czystej rangowej procedury lasso. Appendix A poświęcony jest sytuacji nisko-wymiarowej tzn. takiej gdy liczba predyktorów jest mała w porównaniu z liczbą obserwacji. Wyniki asymptotyczne dla estymatora rangowego lasso i jego modyfikacji są bardzo bliskie analogicznym wynikom dla standardowego estymatora lasso: Th. 8 jest analogiem Cor. 2.3 z A8, Lem. 9 jest analogiem Th. 2.2 z A8, Th. 10(b) jest analogiem Th. 3.1(b) z A8. Można je uznać za ciekawe z punktu widzenia statystyki, ale matematycznie dowody są niemal powtórzeniem odpowiednich dowodów z pracy A8; szczególnie drastyczny jest przykład dowodu Th. 10 (b): Część dowodu zaczynająca się od słów "Let us start...", str. 36, a kończąca frazą "it finishes the proof of the second claim." pod koniec str. 37, jest prawie dosłownym powtórzeniem (z oczywistymi drobnymi modyfikacjami) identycznie się zaczynającego ("Let us start...") i kończącego ("it finishes...") dowodu Theorem 3.1(b) z pracy A8 (str. 1997-1998). Dowód w pracy A2 ma jednak tę przewagę nad dowodem z A8, że poprzedza go wyjaśnienie terminów związanych z ważnym i użytecznym w badanym kontekście analitycznym pojęciem epi-zbieżności (zob. Geyer, 1994).

A1: Praca (mająca pięcioro autorów), jak już wspomniano, wydaje się najważniejsza w zestawie prac składającym się na osiągnięcie habilitacyjne. Proponuje się w niej dwukrokowy algorytm: w pierwszym kroku za pomocą klasycznego lasso wyodrębnia się wstępny zbiór istotnych predyktorów a w drugim minimalizuje się kryterium GIC (generalized information criterion) z odpowiednio dobraną karą we wstępującej rozdzinie zbiorów indeksów uporządkowanych zgodnie z malejącymi modułami estymatorów otrzymanych w pierwszym kroku. Procedura, nazwana SS, jest znaczącym ulepszeniem algorytmu SOS zaproponowanego w pracy Mielniczuka i Pokarowskiego (2015), którzy są współautorami A1. Zgodnie z oświadczeniami współautorów, pomysł algorytmu SS pochodzi od J. Mielniczuka i P. Pokarowskiego. Matematycznie najciekawszy w pracy jest dowód Th. 2 dotyczącego zgodności selekcyjnej proponowanej procedury w modelu GLM. Kluczowe momenty dowodu są subtelne, a dotyczą zawierania się zbiorów, na które dekomponowane jest zdarzenie niepoprawnej selekcji, w odpowiednio skonstruowanych zbiorach krytycznych (Lem. 4 i Lem. 5). Szczególnie trudny do analizy jest przypadek gdy procedura daje w wyniku istotny nadzbiór prawdziwego zbioru istotnych predyktorów. Th. 4 dotyczy analogicznego zagadnienia dla wypukłej funkcji ryzyka, a jego dowód, mimo że częściowo wykorzystuje pomysły z dowodu Th. 2, jest też daleki od trywialności. W szczególności dotyczy to analogii do wspomnianych Lem. 4 i Lem. 5 (przy czym daleko tu od prostego automatyzmu), czy zastępowania nierówności wykładniczych z Lem. 2 założeniami (26) i (27). Zgodnie z deklaracją współautorów (niestety w dokumentacji brak deklaracji dwojga z nich) habilitant był jedną z trzech osób odpowiedzialnych za w dowód Th. 2, a dowód Th. 4 jest jego samodzielnym dziełem. Taki udział należy uznać za znaczący.

Pozostałe osiągnięcia naukowo-badawcze. Pozostały dorobek habilitanta można zakwalifikować do trzech grup tematycznych, z których dwie pierwsze pojawiają się (w wersji "lasso") w rozprawie habilitacyjnej. Są to: estymatory rangowe - asymptotyka i nierówności probabilistyczne dla różnicy między ryzykiem estymatora a ryzykiem najlepszej reguły w odpowiedniej klasie reguł; metody Monte Carlo w zastosowaniu do estymatorów największej wiarogodności gdy stała normująca zależy w skomplikowany sposób od estymowanych parametrów; wykorzystanie metod statystycznych w kartografii.

1. Estymacja rangowa:

   B1: Praca (wspólna z promotorem doktoratu) dotyczy silnej zgodności i asymptotycznej normalności estymatora rangowego.

   B2: W pracy wykorzystano metodologię procesów empirycznych i U-procesów do badania szybkości zbieżności w procedurach estymacji rangowej opartych na wypukłej funkcji straty.

   B10: Praca zawiera sporo interesującej matematyki dotyczącej nierówności probabilistycznych, a najważniejsze twierdzenie dotyczy ograniczeń probabilistycznych na nadmiarowe ryzyko dla estymatora rangowego.

   B8: Praca dotyczy tego samego problemu co praca B10. Otrzymane ograniczenia na nadmiarowe ryzyko są lepsze niż w B10. Matematyka jest nietrywialna. Częściowo pomysły dowodowe z tej pracy wykprzystane były w pracy A6.

   B4: W pracy o charakterze, który trzeba chyba uznać za popularyzatorski, przedstawiono wybrane fragmenty pracy B8.

   B9: Krótka notka dotyczy błędu znalezionego przez habilitanta w pracy Li, Ren, Li (2014). Habilitant jest autorem kontrprzykładów, natomiast część notki dotycząca poprawy dowodu odpowiedniego lematu ze wspomnianej pracy, pochodzi od współautorów: Li, Ren i Li.

2. Metody Monte Carlo w estymatorach największej wiarogodności, MCML (zestaw trzech prac wspólnych m.in. z B. Miasojedowem i W. Niemiro). Metody wypracowane w toku tych badań zostały częściowo wykorzystane w pracy A5.

   B7: Praca dotyczy asymptotyki normalnej estymatorów MCML, w których niejawną stałą normalizującą rozkład przybliża się za pomocą metod MCMC. Dowody są zaawansowane matematycznie.

   B5: Praca rozszerza wyniki z B7 na adaptacyjną procedurę Monte Carlo. Dowodzi się zgodności i asymptotycznej normalności wykorzystując twierdzenie graniczne dla różnic martyngałowych.

   B3: Praca dotyczy asymptotycznej normalności MCML, w sytuacji gdy zarówno liczebność próbki jak i liczba kroków MCMC dążą do nieskończoności. Praca zawiera znaczącą dawkę zaawansowanej asymptotycznej probabilistyki. Jest jedną z ciekawszychw dorobku habilitanta.

3. Zastosowania do oceny jakości map: B6 i B11. Są to prace czysto zastosowaniowe i polegają na wykorzystaniu znanych metod statystycznych, w tym metod rangowych.

Całościowo dorobek naukowy zawarty w pracach B1-B11 należy uznać za solidny. Odnoszę wrażenie jakby pytania stricte matematyczne miały w tej części dorobku habilitanta (oczywiście, poza dwiema pracami poświęconymi kartografii) nieco większą wagę, niż w pracach stanowiących osiągnięcie habilitacyjne. W tym zestawie prac zarówno atrakcyjnością otrzymanych wyników jak i poziomem matematycznym wyróżniają się prace B3 (wspólna z B. Miasojedowem i W. Niemiro, opublikowana w *Annales Institute H. Poincaré*), B8 (publikacja w *Neurocomputing*) oraz B10 (publikacja w *Journal of Machine Learning Research*). Praca B4, jak napisałem wyżej, jest w zasadzie fragmentem wyjętym z B8, dlatego odwołanie się do tej pracy w zdaniu "Problem ten jest badany w [B4, B8], w których skonstruowano nierówności probabilistyczne (4.6.1) z rzędem zbieżności $1/n^{\beta}$, $\beta \in (1/2, 1)$." pochodzącym z autoreferatu, uznaję za nieadekwatne.

Inne aspekty dorobku naukowego. Ilościowo dorobek habilitanta to 19 pozycji, w tym jedna praca, B3, jest opublikowana w prestiżowym czasopiśmie probabilistycznym *AIHP*, jedna, A1, w wysoko cenionym czasopiśmie statystycznym *Scandinavian Journal of Statistics*, trzy w *Journal of Machine Learning Research*, najważniejszym czasopiśmie w dziedzine uczenia maszynowego, dwie w *Neurocomputing*. Pozostałe czasopisma mają mniejszą rangę. Spośród współautorów warto wymienić ważne nazwiska polskiej statystyki i statystyki matematycznej: M. Bogdan, J. Mielniczuka, W. Niemiro, P. Pokarowskiego czy, z nieco młodszgo pokolenia, B. Miasojedowa. Brak publikacyjnego potwierdzenia pracy z młodzieżą naukową, ale według załączonej dokumentacji dr Rejchel jest promotorem pomocniczym dwóch przygotowywanych rozpraw doktorskich na UMK w Toruniu. Habiltant odbył dwa staże naukowe, oba na Uniwersytecie Warszawskim: jeden w ramach grantu OPUS w panelu NCN "Informatyka i Technologie Informacyjne, którego kierownikiem był P. Pokarowski, i jeden w ramach grantu FUGA, którego kierownikiem był habilitant (a opiekunem naukowym W. Niemiro). O rosnącej pozycji dr Rejchela w polskiej statystyce matematycznej świadczy wybór w 2020 roku do Komisji Statystyki Komitetu Matematyki PAN. Nieco słabiej wygląda współpraca międzynarodowa. Nieliczac drobnego epizodu "chińskiego" związanego z pracą B9, jedynym współautorem prac, który nie pracuje w Polsce jest J. Palczewski, który wchodził w skład czterosobowego zespołu autorskiego prac B5 i B7. Aktywność habilitanta na arenie międzynarodowej polega w zasadzie wyłącznie na udziale w konferencjach naukowych (16 wystąpień na konferencjach międzynarodowych, w tym 4 zaproszone).

Wskaźniki bibliometryczne według bazy AMS nie są wysokie: liczba cytowań wynosi 21, ale poza autocytowaniami, cytowana jest jedynie praca B10 (13 razy) i praca A3 (1 raz przez współautorkę). Odpowiednie wskaźniki wg baz WoS oraz Scopus są znacznie lepsze, odpowiednio 52 (44) oraz 66 (52), co może wskazywać na interdyscyplinarność badań habilitanta. Jednak indeks Hirscha wg obu tych baz pozostaje niski i wynosi 3.

Konkluzja: Rozprawa habilitacyjna, której autorem jest dr W. Rejchel, dotyczy statystyki/uczenia maszynowego. Istotną i niezbędną jej częścią są nowe wyniki matematyczne: nierówności probabilistyczne i twierdzenia graniczne. Stanowią one punkt wyjścia dla konstrukcji różnych wariantów algorytmów typu lasso służących do selekcji modelu i/lub estymacji/predykcji. Dr Rejchel znakomicie porusza się w zaawansowanej metodologii probabilistycznej dotyczącej teoretycznych aspektów tego typu zagadnień. Erudycja matematyczna habilitanta odpowiada zarówno za to, że część dowodów wygląda schematycznie, jak i za to, że w kilku z nich zaprezentowane zostały ciekawe i całkiem nowe pomysły matematyczne, które mają szansę wejść na stałe do arsenału technik teorii selekcji modeli statystycznych. Jednak dla rozwoju badań naukowych rozumianych szeroko ważniejsze od wyników matematycznych wydają się być proponowane konkretne algorytmy selekcji i estymacji/predykcji. Gdyby w polskim systemie taka możliwość istniała, w dyscyplinie statystyka omawiana rozprawa habilitacyjna oceniona byłaby z pewnością bardzo wysoko. Uważam, że środowisko matematyczne powinno mocno wspierać działalność matematyków zajmujących się tworzeniem konkretnych narzędzi (w tym algorytmów i ich implementacji) dla potrzeb statystyki, czy uczenia maszynowego. W przeciwnym razie będą to robili (i w pewnym stopniu robią) ludzie bez odpowiedniej wiedzy matematycznej, co może być szkodliwe na wiele sposobów.

Biorąc pod uwagę wszystkie wymienione aspekty, wyrażam zdecydowane poparcie dla nadania dr. Wojciechowi Rejchelowi stopnia naukowego doktora habilitowanego w dyscyplinie matematyka.

J. Wa[illegible]