Válasz Dr. Szabó Zoltán opponensi véleményére

1. Bevezetés

Mindenekelőtt szeretném megköszönni Dr. Szabó Zoltánnak az időben és nagy gondossággal elkészített opponensi véleményét, ill. hogy a téziseimet – egy altézistől és egy összefoglaló mondattól eltekintve – elfogadja, és javasolja számomra az MTA doktora cím odaítélését.

Bár a bírálat stílusa meglepett, igyekszem annak a szakmai tartalmára koncentrálni. Véleményem szerint a félreértések és a kritikai megjegyzések nagy része visszavezethető arra, hogy eltérő tudományterületen dolgozunk: disszertációm a digitális jelfeldolgozás, és ezen belül is az audio szűrőtervezés témaköréhez kapcsolódik, míg számomra úgy tűnik, hogy a Bíráló a szabályozástechnika és az ehhez kapcsolódó rendszeridentifikáció nézőpontjából tekint az értekezésre.

A disszertációmban alkalmazott szóhasználat, jelölésrendszer és stílus, ill. az esetlegesen alapvetésnek tekintett kontextus a saját területem gyakorlatát tükrözi, és e tekintetben nem tér el más kutatók publikációitól.

Azt gondolom, saját kutatási területem irodalmát alaposan áttanulmányoztam, ezzel együtt nem zárható ki, hogy bizonyos publikációk a saját területemen is elkerülték a figyelmem. Különösen igaz ez a témájában kapcsolódó, de más alkalmazási területen megjelent publikációkra. Ugyanakkor köszönöm, hogy a Bíráló ráirányította a figyelmem az általa két fejezetben szerzőként is jegyzett, „Modelling and Identification with Rational Orthogonal Basis Functions," c. összefoglaló műre.

2. Formai észrevételek

Köszönöm a Bíráló disszertációm szerkesztését és az ábrákat dicsérő szavait. A fekete-fehér ábrák használata tudatos döntés volt, szerettem volna, ha az értekezés fekete-fehér nyomtatásban is olvasható marad.

Nagyon szerettem volna a bevezető részeket jobban kifejteni, erre azonban a doktori eljárás szabályzata által előírt, disszertáció törzsére vonatkozó 100 oldalas terjedelmi korlát nem adott lehetőséget. Sajnos erre a bevezető részek függelékbe helyezése sem jelentett volna megoldást, hiszen az elvárás az, hogy az értekezésnek a függelék elolvasása nélkül is érthetőnek kell lennie. A Bíráló jól látja, az eredeti célom valóban az volt, hogy a témából egy hosszabb monográfiát készítsek, és a megjelent könyvet használjam fel a doktori folyamatban. A megjelentetés átfutási idejét is figyelembe véve végül a hagyományos disszertáció beadása mellett döntöttem. Mindenesetre a jövőbeli tervek között szerepel, hogy a disszertáció alapján, de azt bővebb bevezető részekkel és a párhuzamos szűrők tervezésén kívül eső, egyéb eredményeimmel kiegészítve monográfiát írjak a témából. Ha ez megvalósul, akkor a Bíráló megjegyzései és kérdései is nagy segítséget nyújtanak majd abban, hogy bizonyos részeket pontosabban, ill. közérthetőbben fogalmazzak majd meg.

3. A bevezető fejezetekre irányuló észrevételek

Mivel ezen észrevételek nem az új tudományos eredményekre, hanem a bevezető fejezetekre és az irodalmi áttekintésre vonatkoznak, így egyesével csak a legfontosabb észrevételekre reagálok.

1. fejezetre irányuló észrevételek

A Bíráló több fogalom definiálását hiányolja: ebben az esetben is abból indultam ki, hogy a terület irodalmában szokásosan alkalmazott fogalmakat ismertnek tekintem. Ezzel együtt elfogadom, hogy ezen fogalmak pontos definíciója a disszertáció javára vált volna, és a terjedelmi korlátokat nélkülöző, tervezett monográfiában ez megtehető lesz.

Itt csak egy konkrét kritikai megjegyzést emelnék ki, idézem:

„Csak jelzem, hogy ezen a ponton (pp. 5), a „This work, ..., will focus on fixed-pole parallel filters, a methodology allowing the design of IIR filters at arbitrary frequency resolution profiles, while still leading to a simple filter structure with low computational complexity.” - mondat még azt is nyitva hagyja, hogy a tervezési eljárásra vagy a tervezett szűrő futására vonatkozó komplexitásról van-e szó.”

A mondat véleményem szerint pontosan kifejezi, hogy a kis számításigény a szűrőstruktúrákra vonatkozik, hiszen a „low computational complexity” a „filter structre”-el van egy tagmondatban. Azt gondolom, az értekezés többi részéből is egyértelműen kiderül, hogy a szűrési komplexitás csökkentése élvez prioritást. Ez szűrőtervezés területének alapállása is: a szűrőket jellemzően egyszer kell megtervezni (a disszertációban nem érintett adaptív szűrők kivételével), viszont folyamatosan futniuk kell, ebből pedig egyenesen következik, hogy a szűrési komplexitás sokkal kritikusabb, mint a tervezési komplexitás.

Konkrét kérdés: *„A szerző milyen mérési adatok rendelkezésre állását tételezett fel a dolgozat során. A nem egyenletes skála adatai hogyan állnak majd elő?”*

Szűrőtervezési feladat esetén a szűrő specifikációját a felhasználó adja, így ezt disszertációmban adottnak tekintettem. Ez a specifikáció gyakran (de nem minden esetben) valamilyen mérés vagy identifikáció eredményeképp áll elő. Ennek leírása azonban sem a téziseimnek, sem a disszertációmnak nem volt témája.

A disszertáció példáinak elkészítéséhez ugyanakkor magam végeztem méréseket, az akusztikai területen leggyakrabban alkalmazott logaritmikus sweep mérőjel alkalmazásával [Novak et al. 2010]. Az FFT segítségével számolt, így lineáris frekvenciaskálán adott átviteli függvényből újramintavételezéssel állítjuk elő a logaritmikus skála adatait. Amennyiben részkotávsávos simítást is alkalmazunk, ez a lépés összevonható a komplex simítás műveletével (ld. az értekezés 3. fejezetét).

2. fejezetre irányuló észrevételek

A konkrét Bírálói kérdéseken túl a főszövegben is megjelennek kérdések, ezekre is igyekszem válaszolni.

„Az előző fejezetből kiindulva és a frekvenciatartományi esetre tett nem egyenletességi megjegyzés alapján nyitva marad a kérdés: mi van az időtartományban? Ott egyenletes a mintavételezés?”

Igen, időtartományban egyenletes mintavételezést alkalmazok. Digitális jelfeldolgozás területén ez a jellemző választás, kevés speciális kivételtől eltekintve. (Nemegyenletesen mintavételezett jelek esetén pl. a szokásos FIR és IIR szűrőket sem tudnánk alkalmazni.)

„A 2.2 pontban tett megjegyzés csak SISO rendszerekre áll. Amúgy mi van a kauzalitással? Mi van a stabilitással?”

A területemen gyakorlatban előforduló, azaz kauzális és stabil rendszerekkel foglalkozom.

Konkrét kérdés: *„Jól látom, hogy az alapfeltevés ebben a pontban az, hogy SISO kauzális és stabil LTI rendszereket vizsgál (sőt, igazából véges dimenziós SISO kauzális és stabil LTI rendszereket)? ”*

Igen, valóban így van.

3. fejezetre irányuló észrevételek

„Ha jól értem, akkor az itt felsorolt simító ablakos technikák a specifikációt, vagy a használt diszkrét norma aktuális alakját érintik.”

A simítóablakos technikák a specifikációt érintik.

Konkrét kérdés: *„Hogyan biztosítható, hogy a cél (target) egy kauzális, stabil LTI rendszer legyen? ”*

Két esetet kell megkülönböztetni. Amennyiben klasszikus módon csak az amplitúdómenetet simítjuk (3.1 pont), csak a simított amplitúdómenet áll elő, fázisinformáció nélkül pedig sem a kauzalitás, sem a stabilitás nem vizsgálható. Ebben az esetben a szokásos eljárás, ahogy az értekezés 3.2 pontjának elején említem, hogy a fázist minimum-fázisú rendszernek megfelelőem írjuk elő. Minimumfázisú rendszerek esetén pedig mind a kauzalitás, mind a stabilitás biztosítva van.

Azt sejtem azonban, hogy a Bíráló kérdése a komplex simításra irányul, amikor a komplex átviteli függvényt simítjuk, és nem csak annak abszolút értékét. Itt a választ a disszertáció 13. oldalának utolsó mondata adja, miszerint a komplex simítás a frekvenciatartományban ekvivalens azzal, mintha az időtartományban az impulzusválaszt (frekvenciafüggő méretű) ablakkal szoroznánk meg. Az ekvivalencia részletei a disszertációban is hivatkozott [Hatziantoniou and Mourjopoulos 2000] irodalomban olvashatóak. Mindenesetre ezen ekvivalenciából következik, hogy ha az eredeti impulzusválasz kauzális és stabil volt, akkor az ablakfüggvénnyel megszorzott impulzusválasz is az lesz.

4. fejezetre irányuló észrevételek

Mint a többi fejezetben, a frekvenciatorzítás (warping) áttekintése során is a digitális jelfeldolgozás területén született irodalomra támaszkodtam, és a Bíráló, ill. kollégái eredményeit valóban nem ismertem. A Bíráló hiányolja a Laguerre eset és a warpolt FIR szűrők kapcsolatának tárgyalását, ez azonban a következő, 5. fejezetben megtörténik.

A Bírálóban felmerül a kérdés, hogy miért szükséges az impulzusválaszt vagy az átviteli függvényt visszatranszformálni. Amennyiben a szűrőt egy frekvenciatranszformált specifikáció alapján tervezünk meg, az átviteli függvény is a transzformált specifikációt fogja közelíteni. Mi azonban az eredeti specifikáció közelítését szeretnénk elérni. Ehhez vagy a szűrő pólusait és zérusait kell visszatranszformálnunk, vagy speciális szűrőstruktúrát kell alkalmaznunk, ahol a késleltetőelemeket elsőfokú mindentáteresztő szűrőkkel helyettesítjük. A terjedelmi korlátok miatt az eljárás viszonylag tömör leírására volt csak lehetőségem, de pl. a disszertációmban hivatkozott, [Härmä et al. 2000] irodalom kimondottan jó összefoglalást nyújt a témában.

5. fejezetre irányuló észrevételek

A Bíráló megjegyzése:

Az 5.1 szakaz utolsó paragrafusa teljesen elhibázott és talán jobb a véletlen elírások számlájára írni...

Erről a bekezdésről van szó:

„Nowadays solving a linear least squares problem is considered as one of the simplest optimization problems, thus, the orthonormality of Laguerre and Kautz basis functions has lost some of its attractiveness. However, for some cases such as adaptive filtering orthonormality is still highly beneficial since it leads to faster convergence [Salama and Cousseau 1998].”

A bírálatban azonban nincs kifejtve, hogy a Bíráló a paragrafust miért tekinti teljesen elhibázottnak. Magam részéről a fenti állításokat továbbra is igaznak tartom, még így, kontextusból kiragadva is. Disszertációmban többek között azt mutatom meg, hogy szűrőtervezési feladatok esetén az ortogonális Kautz függvények alkalmazásával ekvivalens eredményt érhetünk el egyszerűbb, nem ortogonális szűrőstruktúrával is. Bár az LS tervezés számítási igénye nagyobb a Kautz szűrők által igényelt skalárszorzat-számításnál, manapság ez már nem jelent gyakorlati problémát. Egy konkrét példa: 5000 mintapontból álló célimpulzusválasz esetén egy 1000-es fokszámú fix pólusú párhuzamos szűrő számlálóinak kiszámítása az LS feladat megoldásával egy átlagos laptopon, MATLAB környezetben kevesebb, mint egytized másodpercet vesz igénybe.

Ha jól értem, a Kautz szűrő pólusainak kiválasztását tekintve a Bíráló azt kifogásolja, hogy miért nem a szabályzástechnika területén igen népszerű Ho-Kalman algoritmus kerül alkalmazásra. Ez a fejezet azonban a Kautz struktúra szűrőtervezésre, és ezen belül is a nemegyenletes frekvenciafelbontás elérésére irányuló *előzmények irodalmi áttekintését* célozza. Ezen irodalom nem a Ho-Kalman algoritmust, hanem egyéb, a fejezetben említett módszereket használ a pólusok meghatározására. Ez persze nem zárja ki annak jövőbeli vizsgálatát, mennyiben előnyösebb a Ho-Kalman algoritmus alkalmazása az [Paatero and Karjalainen 2003] irodalom által használt Brandenstein-Unbehauen algoritmushoz képest az adott feladatra. (Ami számomra a legszembetűnőbb gyakorlati különbségnek tűnik, az az, hogy a Ho-Kalman algoritmus a minimális realizáció előállítását célozza, így nincs szükség fokszámbecslésre identifikáció során. Szűrőtervezés esetén azonban jellemzően nem a minimális realizációt keressük, a szűrő fokszáma adott, és tipikusan kisebb, mint amivel a specifikáció hiba nélkül közelíthető. Lehet, hogy ez az eltérő cél az oka, hogy a Ho-Kalman algoritmusra nem találtam a tágabb – tehát nem csak audio – jelfeldolgozás irodalmában semmilyen szűrőtervezési alkalmazást.)

4. Az új tudományos eredményeket tartalmazó fejezeteket érintő megjegyzések és kérdések

Ebben a pontban a saját eredményeimet taglaló fejezetekre vonatkozó kritikai észrevételekre válaszolok. Mivel ezek a tudományos eredményeimet közvetlenül érintik, a bírálat teljes szövegét bemásolom, és nem csak a konkrét kérdésekre, hanem a kritikai megjegyzésekre is reagálok.

6. fejezetre irányuló észrevételek

„A 6. fejezet írja le a javasolt fix pólusú párhuzamos szűrő tervezését és tulajdonságait. A szerző saját eredményként könyveli el a parciális tört alakban történő felírást direkt módon, vagyis az együtthatóknak a megfelelő LS feladat megoldásából származó meghatározásával.”

Természetesen az átviteli függvények parciális tört alakban történő felírása ismert eredmény. Saját eredményem a számlálók direkt módon, LS módszerrel történő meghatározása a szűrő frekvenciafelbontásának beállítására.

„A pólusrendszer rögzítésére vonatkozó megkötések nélkül, abban az általánosságban ahogy a dolgozat ezen részében szerepel, a javasolt algoritmus általában használhatatlan. Ennek magyarázatát a szerző is felírja, csak nem akarja észrevenni. A (6.25) által adott formula

$$K_{i,k} = \sqrt{1 - p_k \overline{p}_k} \prod_{j=1, j \neq i}^{k} \frac{1}{p_i - p_j} \prod_{j=1}^{k-1} (1 - \overline{p}_j p_i) \quad \text{for} \ \ i \leq k \tag{1}$$

középső tagja okozza a problémát, ha a pólusok közel vannak. Ez amúgy klasszikus ismeret és a Vandermonde mátrixok inverzének rossz numerikus kondicionáltságával van összefüggésben. Ez a probléma akkor is fennáll, ha a szerző által javasolt LS feladatot akarjuk megoldani a megoldásban szereplő Gram mátrix rendkívül rossz numerikus kondicionáltsága miatt.”

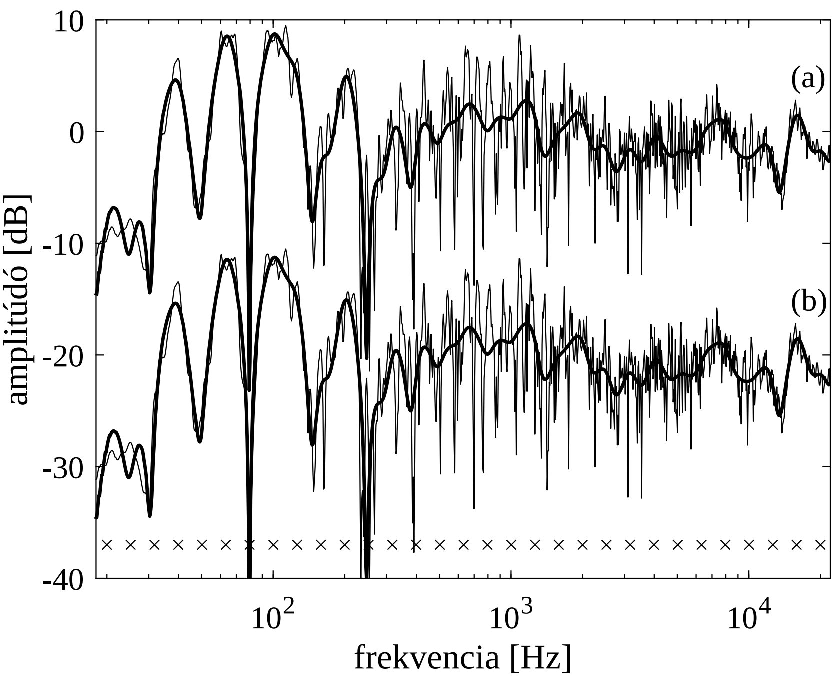
Ezt a megfogalmazást túlzónak tarom. Az egész disszertáció kontextusban, ahol a fix pólusú párhuzamos alakot használom, a Bíráló által felvetett probléma nem általános, hanem speciális eset, és inkább hibaként áll elő. Amennyiben a póluskészletet mi írjuk elő, semmi értelme egymáshoz túlságosan közel álló pólusokat megadni. Ha valahonnan mégis ilyen póluskészletet kapunk, akkor is kézben tarthatjuk a tervezést: a nagyon közeli pólusokat eldobhatjuk, vagy többszörös pólusként összevonhatjuk, amelyeket magasabb fokszámú tagokkal kezelhetünk.

Ezzel együtt köszönöm a Bíráló kérdését, ugyanis ez arra inspirált, hogy szimulációkkal is megvizsgáljam, mi történik, ha mégis előáll a közeli pólusok esete, mennyire érzékeny erre az LS tervezés, ill.

e tekintetben hogyan vizsgáznak a különböző numerikus algoritmusok. Közeli pólusok azzal a következménnyel járnak, hogy a bázisvektoraink közel azonosak lesznek. (Igazából nem is a pólusok távolsága, hanem a bázisvektorok nagy keresztkorrelációja okozza a problémát.) Legrosszabb esetben, nulla távolságú, azonos pólusok esetén – az eredeti, csak másodfokú tagokat tartalmazó struktúra kiegészítése nélkül – az eddigi egyetlen megoldás helyett végtelen sok, azonos hibájú megoldásunk lesz. Ez lineáris egyenletrendszerek megoldásánál is előforduló eset, így megoldásukra (és ennek megfelelően az LS feladat megoldására) kifejlesztett numerikus módszereknek ezt az esetet jól kell tudniuk kezelni.

A disszertációban belinkelt MATLAB példakódjaimban alkalmazott `\` (azaz `mldivide`) utasítás az azonos bázisvektorok kérdését az elvártaknak megfelelően jól kezeli (jelzi az alacsony rangszámot, de helyes megoldást ad). Kíváncsiságból a szinguláris érték felbontást alkalmazó pszeudoinverz (`pinv`) függvénnyel is teszteltem a rosszul kondicionált LS feladat megoldását, és az `mldivide`-tól eltérő együtthatókat, de azonos szűrőátvitelt kaptam. A konjugált gradiens algoritmust alkalmazó `pcg` utasítás használatával is gyakorlatilag azonos szűrőátvitelt kapok. Ezen három algoritmus közül rosszul kondicionált esetek kezelése szempontjából a `pinv` alkalmazása a legelőnyösebb, mert az a számlálókra minimális normájú megoldást ad.

Illusztrációként a 1. (b) ábrán egy rendkívül rosszul kondicionált póluselrendezés látható. Az (a) görbe logaritmikus frekvenciaskálán elhelyezett pólusfrekvenciákat alkalmaz, a (b) görbe esetén pedig az összes pólust megdupláztam, de úgy, hogy továbbra is csak másodfokú tagok szerepelnek a tervezésben (a kétszeres komplex póluspárok helyes figyelembevételéhez negyedfokú tagokra is szükség lenne). Ennek eredményeképpen az $\mathbf{M}$ mátrixban *minden* bázisvektor kétszer szerepel – nehezen tudok ennél numerikus szempontból kedvezőtlenebb helyzetet elképzelni. Ennek megfelelően bár az $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ mátrix rendje 124, rangja csak 62. Ennek ellenére a szűrőtervezés mégis működőképes, és az egyszeres pólusokat tartalmazó (a) esettel gyakorlatilag azonos eredményre vezet.



1. ábra. Fix pólusú párhuzamos szűrő tervezése. Logaritmikus póluselrendezés 31 pólusfrekvenciával, mindegyik komplex póluspár egyszeres (a), ill. mindegyik póluspár kétszeres pólus (b). Vékony vonal: célátvitel, vastag vonal: szűrőátvitel. A pólusfrekvenciákat keresztek jelölik.

„Van a dolgozatban még egy hely, ahol ezt tetten lehet érni: a 9.2 táblázat a 92. oldalon."

Az a sejtésem, hogy a Bíráló a „NaN" (not a number) érékekből indult ki. A 9.2 táblázat bal oldala, ahol ezek az értékek láthatók, nem az én eljárásom eredményeit tartalmazza, hanem a részlettörtekre bontással kapott átviteli függvényeket. Továbbá, ezen érékek ott sem a pólusok közelsége, hanem a gyökhelykeresés numerikus hibáiból adódó instabil pólusok miatt állnak elő. A 9.2 táblázat jobb oldali oszlopa mutatja az általam javasolt LS eljárás eredményeit: látható, hogy még 8000-es fokszám esetén is elhanyagolható (1.1×10^{-5} dB) a közelítés hibája.

„Ezért azután a 6.4 alfejezetben levezetett, amúgy triviális, kapcsolat sem aknázható ki, általában, a kívánt reprezentáció előállítására.

A disszertáció (6.20) egyenlete után kikötöm, hogy a többszörös pólusokat kizárom, tehát leírom, milyen korlátok között értelmezem a megoldást. Azzal egyetértek, hogy érdemes lett volna leírni, hogy nem csak a többszörös (nulla távolságú), hanem nagyon közeli pólusok is okozhatnak numerikus problémát a konverzió után előálló túlságosan nagy szűrőegyütthatók miatt. Ez azért maradhatott el, mert a túlságosan közeli pólusok esetét nem tartottam gyakorlati szempontból relevánsnak.

„Még akkor sem, ha nagyvonalúan eltekintek attól, hogy a rendszer nem diszkréten ortogonális."

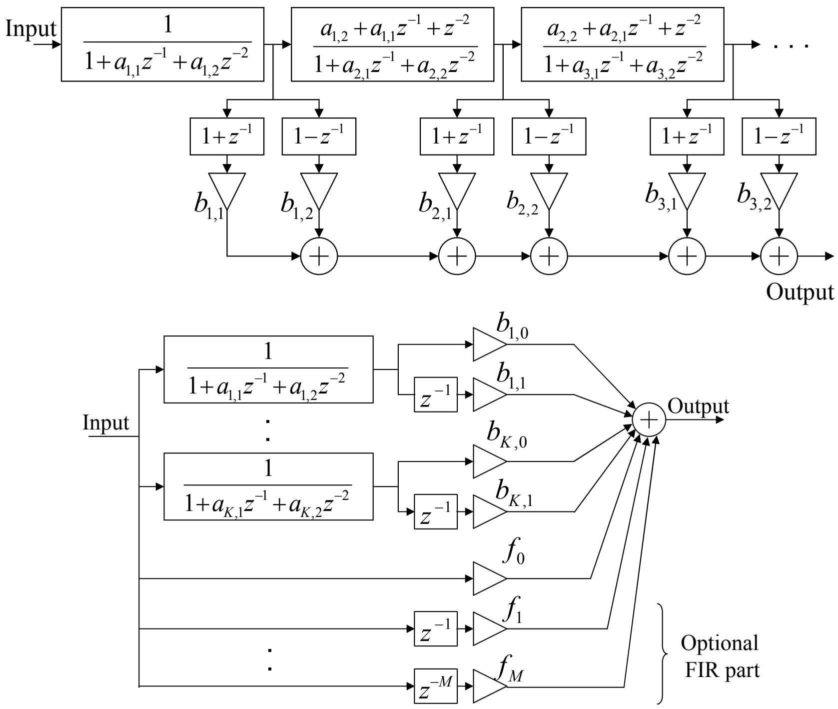
A 6.4 pontban alkalmazott részlettörtekre bontás független az ortogonalitástól. Ezzel együtt a pontosság kedvéért érdemes megjegyezni, hogy a Kautz szűrőben [Paatero and Karjalainen 2003] alkalmazott bázisfüggvények ún. diszkrét Kautz függvények, amelyek a mintavételi időköztől függetlenül diszkrét ortogonálisak (ld. az értekezés Kautz szűrőről szóló fejezetében is hivatkozott [Broome 1965] cikket a részletekért).

„Ami ebből a fejezetből hiányzik: annak igazolása (pld. hibabecslések), hogy a javasolt póluselrendezéssel és a javasolt algoritmussal (milyennel is?) valóban előállítható a feladat megoldása a feladatosztályra."

A disszertációmban négyzetes hibakritériumot alkalmazok, mivel a feladat paramétereiben lineáris, az optimális megoldás egyértelmű és zárt alakban megkapható. (Ranghiányos esetben végtelen sok azonos hibájú megoldás van, ami közül egy – pl. a minimális normájú – a fentiekben említett numerikus módszerekkel szintén előállítható.) Más hibakritériumok alkalmazása esetén valóban érdemes lehet annak diszkussziója, hogy előállítható-e a megoldás, de ez a disszertációmnak nem volt témája.

„Ahhoz képest, hogy milyen fontosnak tartja a szerző, hogy a parciális tört felbontás az implementáció szempontjából előnyös, egy rövid fél oldallal elintézi a témát a 6.4.2 alfejezetben. Véleményem szerint a kérdéskör ennél jóval összetettebb és több figyelmet és magyarázatot igényel. Lásd még az idevágó kanonikus alakra alapozott példám. Arról nem is beszélve, hogy az adatmozgatás és memóriához való hozzáférés ugyan olyan tényező lehet egy adott implementáció során, mint a műveletszám."

Azt gondolom, hogy az értekezés 5.1 és a 6.1 ábrájának összehasonlításából látható, hogy a Kautz szűrő a párhuzamos szűrőhöz képest lényegesen nagyobb számításigénnyel rendelkezik. A blokkvázlatokat válaszomban is megismétlem a 2. ábrán. Disszertációmban a szorzások és összeadások számát elemzem, ill. megemlítem, hogy digitális jelfeldolgozás területén alkalmazott DSP-k esetén az adatmozgatás nem igényel plusz órajelciklust, mert az a MAC (multiply-and-accumulate) utasítással párhuzamosan történik. Egyéb processzorarchitektúrák esetén is igaz, hogy az összeadások és különösen a szorzások száma az, ami az órajelszámot jelentősen meghatározza. Ugyanakkor a blokkvázlatokból az is látható, hogy a Kautz szűrő az adatmozgatás tekintetében is műveletigényesebb a párhuzamos szűrőnél. Tekintve, hogy a bonyolultabb Kautz szűrőstruktúra a ciklusszervezés tekintetében is nagyobb terhet jelent, a disszertációmban megadott számitásigény-csökkenés inkább konzervatív, alsó becslés arra vonatkozóan, mekkora a nyereség a párhuzamos szűrők alkalmazásával. A párhuzamos struktúra előnyét tovább növeli, hogy a párhuzamos szűrő másodfokú tagjainak számítása teljes mértékben párhuzamosítható, míg a Kautz szűrőnél ez csak a kicsatoló tagok esetén, tehát a műveletek egyharmadára tehető meg.



2. ábra. Kautz (fent) és párhuzamos (lent) szűrőstruktúra.

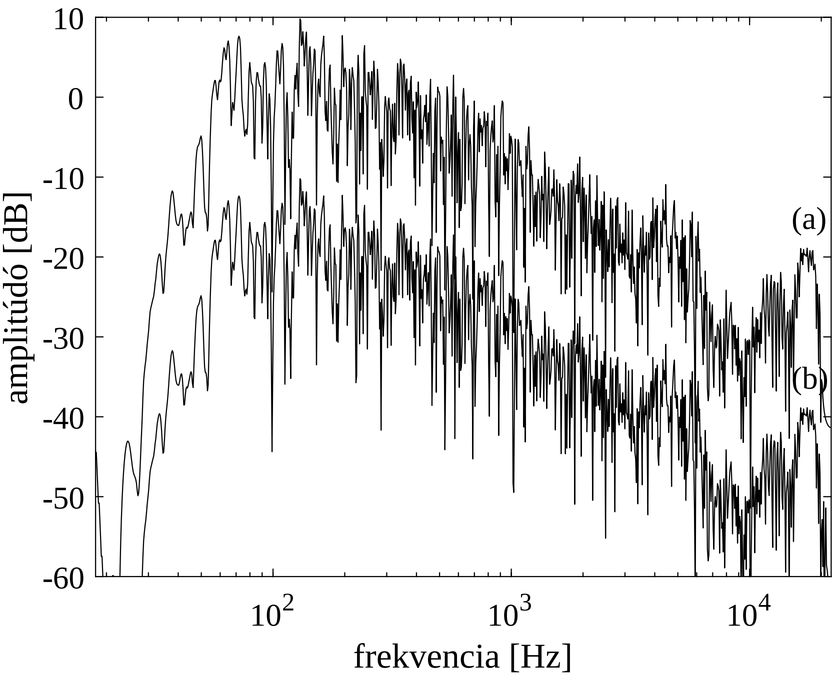
„A 6.5. alfejezet meglátásai érdekesek, hasznosak és újak. A tézis megfogalmazásakor én inkább erre a részre koncentráltam volna.”

Köszönöm a Bíráló elismerő szavait! Az első tézis harmadik altézise ezt tartalmazza, ugyanakkor a tézist összefoglaló főszövegben ez sajnos nem került bele. E tekintetben valóban szerencsés lett volna a „I have also shown the connection between fixed-pole parallel filter design and complex smoothing.” mondatot az összefoglaló szöveghez is hozzátenni.

Konkrét kérdés: „*Mekkora modellméret és mintaszám esetén gondolja még alkalmazhatónak a javasolt eljárást?*

Tapasztalataim szerint az alkalmazásnak a tervezést végző számítógép memóriája szab határt. Az értekezés 9.2 táblázatának jobb oldali oszlopához pl. 8000-es maximális fokszámú szűrőt terveztem. A mintaszám (akár az impulzusválasz hossza, akár a frekvenciapontok tekintetében) megválasztásánál nyilvánvalóan törekedni kell arra, hogy az egyenletek száma magasabb legyen a szabad paraméterek számánál. A fenti 8000-es fokszámú szűrőket 32000 minta hosszú impulzusválasz alapján terveztem.

Példaként álljon itt egy még nagyobb méretű, 20000-es fokszámú szűrő, logaritmikus póluselrendezés mellett, 40000 minta hosszú koncerterem-impulzusválasz alapján tervezve (3. ábra). Ilyen nagy fokszámú szűrőkre nagy zengési idejű termek impulzusválaszának modellezésén kívül nem igazán látok más gyakorlati alkalmazást. (Hangszertestek átvitelének modellezésére ezres nagyságrendű fokszám elegendő, míg pl. hangszórók frekvenciamenetének kiegyenlítésére száz alatti fokszámok alkalmazása jellemző). Mindenesetre a példa jól mutatja a tervezési eljárás robusztusságát, ami – a szabad paraméterbeli linearitás miatt – a FIR szűrők tervezésének robusztusságával állítható párhuzamba.



3. ábra. Fix pólusú párhuzamos szűrő tervezése 20000-es fokszám esetén: szűrőátvitel (a) és specifikáció (b). A két görbe a jobb áttekinthetőség kedvéért 20 dB-el el van tolva egymástól.

7. fejezetre irányuló észrevételek

„Új póluselrendezési eljárásokat javasol. Jó lett volna, ha az olyan utalások, mint: „the parallel filter is more advantageous from the implementation point of view in terms of computational complexity and quantization noise" alá lennének támasztva a dolgozatban. Ez ugyan is valóban tézis értékű állítás lehetne, ellentétben sok mással, ami végül is bekerült a dolgozatba."

A kontextusából kiemelt félmondat önmagában valóban nem áll meg, ugyanakkor ugyanazon bekezdés korábbi mondataiban (ld. 49. oldal) olvasható az indoklás, ill. a hivatkozás is az ezzel foglalkozó publikációmra [Bank and Horváth 2017].

A terjedelmi korlát miatt mérlegelnem kellet, hogy mi kerüljön be a disszertációba, és én úgy döntöttem, hogy a numerikus kérdések, így a párhuzamos szűrő és a warpolt szűrő numerikus szempontból történő összehasonlítása csak hivatkozásként és a kapcsolódó eredmények összefoglalásaként (ld. 10.1.1 pont a disszertációban) kap teret. Ennek többek között az az oka, hogy arra törekedtem, hogy a téziseim alapjait egyszerzős cikkek képezzék, a numerikus kérdések vizsgálatát pedig doktorandusz hallgatómmal együttműködésben végeztem.

„Itt is vannak olyan részek, megfogalmazások, amit nem egészen értek: pld. 7.2 alfejezetben „It has been observed for Kautz filters that a logarithmic pole set produces logarithmic frequency resolution ..." állítás alatt mit kell érteni? Ha jól értelmezem a fogalmat a „logarithmic frequency resolution" az egy ábrázolási döntés következménye, és azt a rácsot határozza meg, amin kiértékelem/definiálom, pld., a normákat. Gondolom nem ebben az értelemben használta a szerző az adott szövegkörnyezetben. Ide kellett volna egy magyarázó utalás, hogy hogyan is kell érteni az adott mondatot."

Területemen azon szűrőket tekintjük logaritmikus frekvenciafelbontásúnak, ahol a szűrő által megvalósított átviteli függvény változási képessége logaritmikus skálán ábrázolva lesz egyenletes.

„Itt is visszaköszön az, hogy nincs pontosan megfogalmazva, hogy mit is akarok. Ezért nem világos, pld., hogy miért is baj az, ha kézzel kell valamit beállítani (pld. a pólusok sugarát)”

Egy tíz alatti fokszámú szűrőnél lehet, hogy a csillapítási tényezők beállítása gyorsan megoldható, de egy századfokú szűrőnél a paraméterek manuális állítgatása – úgy, hogy az egyes tagok átvitelei kb. a -3 dB-es pontjaiknál találkozzanak – igencsak időigényes. Ha ez a felesleges munka könnyen elkerülhető, és a csillapításra zárt alakú (közelítő) formula is megadható, akkor érdemes így eljárni, arról a járulékos előnyről nem is beszélve, hogy az automatizálás a reprodukálhatóságot is biztosítja.

„A fejezet erénye, hogy összehasonlít számos pólusválasztási stratégiát és ad egy intuitív képet az egyes eljárásokról. Elkerülhetetlen azonban, hogy számos, potenciálisan releváns stratégia kimaradjon a vizsgálatokból.”

Köszönöm a pozitív visszajelzést. Én is egyetértek azzal, hogy bármely kutató számára lehetetlen lenne az összes potenciális stratégia vizsgálata.

Konkrét kérdés: *„Miért is jobb a javasolt pólusválasztási stratégia a kitűzött feladatosztályra, mint más egyéb, például lásd az idézett szekvenciális stratégiát, pólusválasztási eljárás?”*

Soumelidis, Alexandros and Bokor, József and Schipp, Ferenc (2013) An iterative identification of pole-stucture in dynamic systems based on hyperbolic metrics and Malmquist-Takenaka representation. In: Proceedings of the 52nd IEEE Conference on Decision and Control, 2013-12-10 – 2013-12-13, Florence, Italy.

Köszönöm, hogy a Bíráló felhívta a figyelmem a fenti publikációra. Amennyiben jól értem, a cikkben javasolt módszernek nem célja a logaritmikus frekvenciafelbontás elérése. A kérdés tehát inkább olyan szempontból releváns, hogy a fenti eljárás alkalmazható-e a kétsávos és a custom warping módszerek belső szűrőtervezési lépéseként.

A válasz igen, ahogy az értekezés 48. oldalának első bekezdésében írom, a belső lépésként alkalmazott Steiglitz-McBride algoritmus tetszés szerint lecserélhető bármilyen IIR szűrőtervezési vagy identifikációs algoritmusra, így igény szerint a fenti algoritmusra is. Az általam javasolt módszereknek tehát nem lényegi eleme, hogy az eljáráson belül milyen IIR szűrőtervezési (vagy identifikációs) algoritmust alkalmazunk, ahogy az irodalomból korábban ismert, egyszerű (egy λ-t alkalmazó) warpolásnak sem. Az irodalomhoz képesti új eredményeim szintén nem az alkalmazott szűrőtervezési algoritmusban, hanem a warpolás olyan új módozataiban rejlenek, melyek az egyszerű warpoláshoz képest logaritmikus skálán egyenletesebb frekvenicafelbontást eredményeznek.

„Érdemes lenne összehasonlítani a javasolt „custom warping” eljárást a pólusok által meghatározott frekvenciatorzítási függvénnyel kapott megfontolásokkal (lásd β függvény az idézett cikkben, vagy lásd a monográfiát a részletekért).”

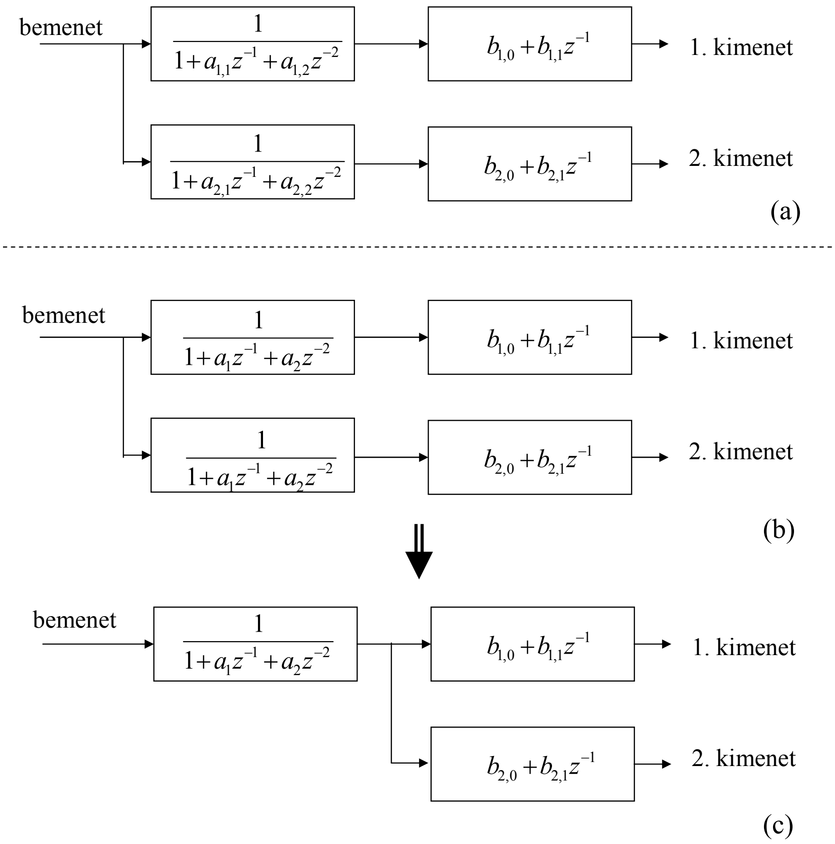
Véleményem szerint a β függvény a hagyományos (egy λ-t alkalmazó) warpolásnak feleltethető meg. A custom warping és a hagyományos warpolás között pedig a 7.4 fejezet ábrái és a 7.2 táblázat nyújtanak összehasonlítást.

8. fejezetre irányuló észrevételek

„MIMO rendszerekre terjeszti ki a javasolt eljárást. Javaslom a szerzőnek átgondolni a megtakarításokra vonatkozó állítást, ami nyilvánvaló tévedés. Ennek belátására vegyen egy diagonális rendszert egyszerűség végett egy kétszer kettes MIMO rendszert. Triviális, hogy a komplexitás csak a rendtől függ, és nem a konkrét pólusoktól. Az egy más kérdés, hogy más egyéb szempontból előnyös/célszerű közös pólusrendszerrel dolgozni.”

Azzal természetesen egyetértek, hogy általános esetben – ha nem használjuk ki a közös pólusok adta lehetőségeket, és minden egyes átvitelt független szűrővel valósítunk meg – a komplexitás nem függ a pólusoktól.

Disszertációmban azt állítom, hogy amennyiben az átviteli függvény mátrix elemeinek azonosak a pólusai, az lehetőséget biztosít arra, hogy a szűrők nevezőit a különböző ágakban összevontan, csak egyszer valósítsuk meg, és ennek módját az értekezés 8.1.1. pontjában részletesen kifejtem. Ahogy a disszertációban is írom, a közös pólusú elrendezés alacsonyabb számításigénye nemcsak MIMO, de SIMO vagy MISO esetben is megnyilvánul. A 4. ábrán ezt egy egy-bemenetű, két-kimenetű (SIMO) IIR szűrőn mutatom be, az átláthatóság kedvéért egy-egy másodfokú taggal.



4. ábra. Egy-bemenetű, két-kimenetű IIR szűrő: (a) különböző pólusok, ill. közös pólusok összevonás előtt (b) és után (c).

Látható, hogy míg a különböző pólusokat és így különböző nevezőket tartalmazó (a) szűrő megvalósítása nyolc szorzást igényel, a közös pólusok lehetővé teszik a nevezők összevont megvalósítását (c), így csak hat szorzásra van szükség. A számításigény-csökkenést egyéb bemenet- és kimenetszámokra az értekezés 8.1 táblázata tartalmazza.

9. fejezetre irányuló észrevételek

„9. fejezet az úgynevezett késleltetett párhuzamos szűrők tervezésével foglalkozik. Ez egy érdemi fejezet, ami a javasolt reprezentáció praktikus elő- állítására javasol numerikusan használható módszert.”

Itt is köszönöm a pozitív visszajelzést.

„Magam részéről nem tartom meggyőzőnek a fejezetben használt érvelést. Nem látom be, hogy a késleltetés miért segítene a módszer alapproblémáján (potenciálisan rosszul kondicionált feladat). Az idő és

frekvenciatartományi érvelés keveredése sem segíti a megértést.”

A félreértést az okozhatja, hogy ebben az esetben a rossz kondicionáltságot nem a Bíráló által a 6. fejezet kapcsán felvetett, túlságosan közeli pólusok megléte okozza. Pontosabban, itt nem a tervezési feladat kondicionáltsága okozza a problémát, több ezres fokszámú párhuzamos szűrőket is lehet tervezni az eredeti, nem késleltetett alakban is MATLAB segítségével. A gond az, hogy a FIR és az IIR részek átvitele sokkal nagyobb lesz, mint az eredeti átvitel, ahogy pl. az értekezés 9.2 ábrájának bal oldalán látható. Ez MATLAB-ban, duplapontos lebegőpontos számábrázolás esetén nem okoz problémát. A problémával a szűrő gyakorlati, kisebb bitszámú (pl. 24 bites) implementációjakor szembesülünk, mivel kénytelenek leszünk a bemenőjelet leskálázni, hogy elkerüljük az egyes ágakban előálló túlvezérlést, és ezt kompenzálandó, a kimenetet a szűrés végeztével vissza kell skáláznunk. A skálázás pedig az effektív bitszám csökkenéshez, azaz a kvantálási zaj növekedéséhez vezet. Ezen nemkívánatos jelenség kerülhető el azzal, hogy az IIR részt megkésleltetjük, az értekezésben részletezett okokból kifolyólag.

Konkrét kérdés: *„Instabil szűrő esetén honnan lesz impulzus válasz információnk vagy honnan származik* $A(z^{-1})$*?”*

A fejezetben stabil szűrők felbontásával foglalkozom. Az instabilitás akkor jelenhet meg, amikor a felbontani kívánt, racionális törtfüggvény alakban adott átvitel pólusait meghatározom. A gyökhelykeresés ugyanis numerikus szempontból rendkívül rosszul kondicionált feladat, így eredetileg stabil rendszernek megfelelő átviteli függvény esetén is adhat egységkörön kívüli pólust. Pontosan ez az oka, ami miatt a résszlettörtekre bontás nem vezet eredményre nagy fokszámok esetén (ld. 9.2 táblázat bal oldali oszlopa). Én a javasolt eljárásban ezt úgy oldom meg, hogy a numerikus hibák által eredményezett instabil pólusokat az egységkörön belülre tükrözöm, és az így korrigált, bár az eredetitől némileg eltérő, de stabil pólusok felhasználásával határozom meg a másodfokú tagok számlálóit a legkisebb négyzetek módszerének alkalmazásával. Mivel eljárásomban az impulzusválaszt, amire a tervezést végzem, az eredeti, stabil szűrő futtatásával számítom (tehát ahhoz nem használom fel az esetleges numerikus hibákat tartalmazó pólusokat), nem merül fel stabilitási probléma.

5. A tézisek értékelése

A Bíráló alapvetően elfogadja a téziseket, mint új tudományos eredményeket, amit ezúton is szeretnék megköszönni.

Ez alól az 1.2 pont képez kivételt. Ezt a Bíráló nem fogadja el tézis értékű állításnak, ill. új tudományos eredménynek. Én az 1.2 pontot nem önálló tézisként, hanem az 1. tézis altéziseként fogalmaztam meg, és egyetértek azzal, hogy önálló tézisként nem állná meg a helyét, ill. azzal is, hogy az altézisek közül ez igényelte a legkevesebb kutatómunkát. Ugyanakkor ez az eredmény abból a szempontból fontos, hogy megmutatja, hogy az 1.3 altézis és a 2. tézis által összefoglalt eredményeim nem csak a párhuzamos, hanem a Kautz szűrő esetén is alkalmazhatóak. Arról a gyakorlati alkalmazásról nem is beszélve, hogy lehetőséget biztosít arra, hogy a tervezést Kautz alakban, az implementációt pedig a hatékonyabb párhuzamos alakban tegyük meg. Erre én korábban nem láttam példát (hozzátéve, hogy a digitális szűrőtervezés irodalmát ismerem), így ez véleményem szerint tudományos újdonságnak tekinthető. Remélem, hogy az 1.2 pont altézisként (tehát nem különálló tézisként) a Bíráló számára is elfogadható.

Szintén kivételt képez a 3. tézis főszövegének első mondata. Reményeim szerint a közös pólusú szűrők kisebb számításigénynek kérdését a 8. fejezetre irányuló észrevételek kapcsán a 4. ábrával tisztáztam, így a Bíráló számára a tézis ezen mondata is elfogadható lesz.

6. Összefoglalás

Összefoglalásként ismét szeretném megköszönni a Bíráló munkáját. Kérdéseivel számos olyan pontra világított rá, melyeket a későbbi publikációimban érdemes lesz pontosabban, ill. érthetőbben megfogalmazni, ill. a javasolt irodalom, és az azok alapján megtalált egyéb publikációk újabb vizsgálatok, és remélhetőleg új kutatási eredmények alapjait képezhetik.

A végső értékelést tekintve örömmel olvastam, hogy a kritikai megjegyzések ellenére a Bíráló új tudományos eredményeimet jelentősnek tartja és javasolja számomra az MTA doktora cím odaítélését, amit ezúton is köszönök.

Hivatkozások

Bank, B. and Horváth, K. (2017). Quantization noise of warped and parallel filters using floating point arithmetic. In *Proc. 142$^{\text{nd}}$ AES Conv., eBrief No. 337*.

Broome, P. W. (1965). Discrete orthonormal sequences. *J. of the Association for Computing Machinery*, 12(2):151–168.

Hatziantoniou, P. D. and Mourjopoulos, J. N. (2000). Generalized fractional-octave smoothing of audio and acoustic responses. *J. Audio Eng. Soc.*, 48(4):259–280.

Härmä, A., Karjalainen, M., Savioja, L., Välimäki, V., Laine, U. K., and Huopaniemi, J. (2000). Frequency-warped signal processing for audio applications. *J. Audio Eng. Soc.*, 48(11):1011–1031.

Novak, A., Simon, L., Kadlec, F., and Lotton, P. (2010). Nonlinear system identification using exponential swept-sine signal. *IEEE Trans. Instrum. Meas.*, 59(8):2220–2229.

Paatero, T. and Karjalainen, M. (2003). Kautz filters and generalized frequency resolution: Theory and audio applications. *J. Audio Eng. Soc.*, 51(1–2):27–44.

Budapest, 2022. december 1.

Bank Balázs