

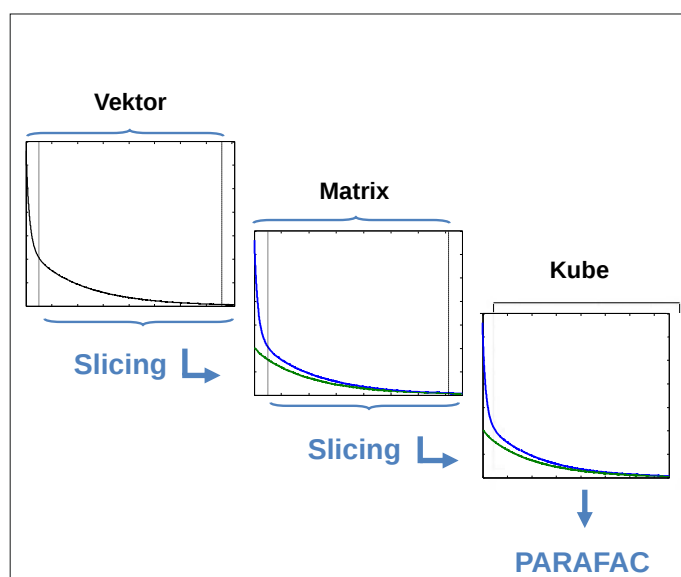
Multieksponentiel analyse: CuSO₄, relaxometri og multivejskemometri

Multivejs kemometriske modeller kan bruges til unikt at resolve multi-eksponentielle relaxationsprofiler i underliggende rene eksponentielle profiler. Metoden benytter sig af redundant information, og tricket hedder: slicing igen-igen.

Af Søren Balling Engelsen, Rasmus Bro, Institut for Fødevarevidenskab, Københavns Universitet og Lars Nørgaard, FOSS

Multieksponentiel analyse er et klassisk analytisk problem. Overlappende eksponentielle fænomener er almindeligt forekommende indenfor forskellige videnskabelige discipliner såsom dynamiske biokemiske reaktioner og relaxometriske målinger. Der er derfor fokus på robuste og effektive løsninger på dette ikke-lineære problem. I denne foreløbig sidste klumme omhandlende multivejskemometri vil vi vise, at multivejskemometri, blandt mange andre anvendelser, også kan bidrage med en effektiv alternativ løsning til resolvering af det multieksponentielle problem. Der er ikke tale om en klassisk kemometrisk anvendelse, hvor signaler fra flere prøver behandles samtidig, men om en højst overraskende anvendelse til signalbehandling af en enkelt multieksponentiel profil.

Tricket er en opgradering af den multieksponentielle profil til en datakub ved brug af en operation, som hedder dobbelt-slicing, deraf navnet på algoritmen: DOUBLESLICING [1]. Ideen bag DOUBLESLICING er at slice den eksponentielle profil to gange og derved generere en trevejs datakub, som kan analyseres ved brug af multivejs kemometriske modeller. Slicing-ideen blev præsenteret af Windig og Antalek på KODAK-laboratorierne [2]. Figur 1 viser princippet bag DOUBLESLICING. Ved den første slicing-operation (translation langs tidsaksen) skabes et antal pseudoprofiler, som alle har de samme underliggende eksponential profiler, men forskudt i tid. Det er altså samme profil, der gentages, men forskudt. Den ene profil bliver derved til en matrix af profiler, som hver repræsenterer forskudte tidsafsnit. Ved den anden slicing-operation gentages processen langs kolonneretningen, og dette resulterer i en trevejs datastruktur. Denne preprocessing kan synes som en unødvendig kompliceret proces med stærkt redundant information, men den har til formål at skabe en trevejs datastruktur fra en enkelt multieksponentiel profil. Man kan vise, at denne trevejsstruktur vil følge en PARAFAC-model, og man kan derved benytte den til at løse det diskrete multieksponentielle problem. Fordelene ved at benytte multivejskemometri er, at løsningen, givet ved PARAFAC i modsætning til



Figur 1. Princippet bag DOUBLESLICING. Det transverse NMR-signal optaget som funktion af tiden er en endimensional datastruktur (en vektor), som indeholder en sum af N eksponentielle henfald svarende til N forskellige vandpopulationer. Kurven inddeles i et antal (to er vist) af stærkt overlappende segmenter (X-slices) ved at fjerne det samme antal af de første eller sidste punkter. De to segmenter placeres i en matrix (tovejs datastruktur), hvorefter proceduren gentages for at danne Z-slices, som kan stables efter hinanden for at danne en trevejs datastruktur. Disse data kan herefter dekomponeres vha. PARAFAC.

multieksponentiel kurvefitning, bliver unik og lynhurtig under forudsætning af, at der er tale om et lavrangs (få komponenter) multieksponentielt system.

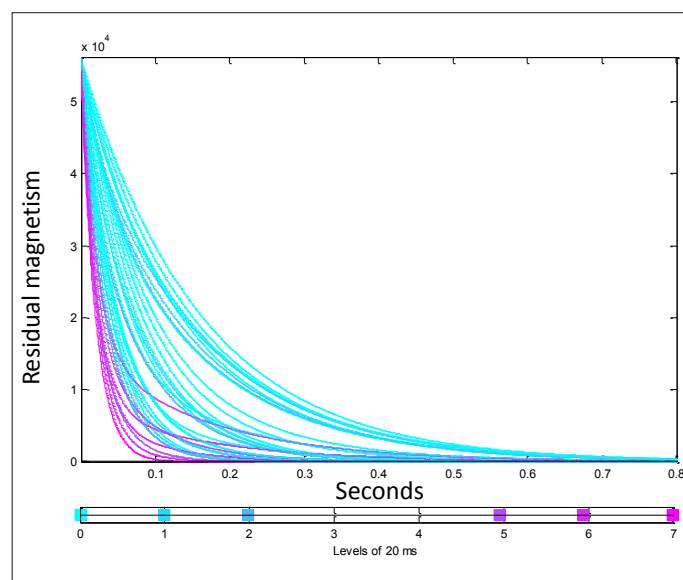
Vi vil i denne klumme demonstrere algoritmen til brug for multieksponentielle profiler målt med et benchtop kernemagnetisk resonans (NMR) instrument også kaldet NMR-relaxometri. Denne type målinger er udbredt indenfor fødevareforskningen, hvor de bl.a. bruges til at bestemme "solid fat" i spiseolier, hvor profilen deles op i en hurtigt relaxerende del (fast fedt) og en langsomt relaxerende del (flydende fedt).

Sample no.	1 ($T_2 = 19.6\text{ms}$)	2 ($T_2 = 40.6\text{ms}$)	3 ($T_2 = 81.4\text{ms}$)	4 ($T_2 = 159.4\text{ms}$)	Sample label
1	0	0	0	7	0007
2	1	0	1	5	1015
3	0	6	0	1	0601
4	1	0	6	0	1060
5	1	6	0	0	1600
6	0	0	6	1	0061
7	0	2	0	5	0205
8	2	5	0	0	2500
9	0	5	2	0	0520
10	0	0	7	0	0070
11	0	6	1	0	0610
12	0	5	0	2	0502
13	0	0	2	5	0025
14	6	0	1	0	6010
15	0	0	5	2	0052
16	0	1	0	6	0106
17	0	2	5	0	0250
18	1	5	1	0	1510
19	1	0	0	6	1006
20	2	0	5	0	2050
21	6	1	0	0	6100
22	0	0	1	6	0016
23	5	2	0	0	5200
24	7	0	0	0	7000
25	6	0	0	1	6001
26	0	7	0	0	0700
27	5	0	2	0	5020
28	0	2	0	5	0205
29	0	1	6	0	0160
30	5	0	0	2	5002



Figur 2. Det eksperimentelle design. Koncentrationen af de 30 prøver er i enhederne 1-7, hvor hver enhed repræsenterer en syvendedel af prøvevolumet. Prøverne er mærket iht. deres koncentration.

For at kunne illustrere den nye algoritmes effektivitet har vi designet et prøvesæt, som eliminerer vekselvirkninger mellem de forskellige eksponentielle komponenter. NMR er baseret på diamagnetisme, men det er velkendt, at tilstedeværelse af paramagnetiske kerner vil quenche NMR-signalerne. Tilsætning af forskellige koncentrationer af CuSO_4 til vand kan således designes til en given tilsyneladende T_2 -relaxation. Vi har til formålet lavet fire moderopløsninger med forskellige CuSO_4 -koncentrationer på hhv. 55, 26, 13, og 6.5 mM. Det svarer til tilsyneladende transverse T_2 -relaxationstider på 20, 40, 80 og 160 ms. Mono-eksponentiel kurvefitning af de målte relaxationsprofiler for de fire moderopløsninger resulterede i følgende T_2 -værdier: 19.6, 40.6, 81.4, and 159.4 ms.



Figur 3. De 30 relaxationsprofiler farvet efter deres indhold af den hurtigst relaxerende komponent (20 ms).

Vha. de fire moderopløsninger har vi konstrueret et 4-komponent (D-optimal) design med syv koncentrationsniveauer. Designet på i alt 30 prøver blev etableret ved at placere syv små NMR-rør (5mm) fyldt med de fire moderopløsninger indeni i et 18mm NMR-rør (figur 2). Koncentrationsværdier, som svarer til antallet af de små rør med en given moderopløsning (værdier fra 1 til 7), er vist i figur 2. Hver enhed svarer til en syvendedel af det totale NMR-volumen. Skemaet (figur 2) angiver, hvor mange af hver af de fire moderopløsninger, der er til stede ved en given måling (i prøve nummer 1 er alle de små NMR-rør fyldt med moderopløsning 4). Ved at bruge dette eksperimentelle design er det muligt at kontrollere det eksakte bidrag fra hver af relaxationskomponenterne til det totale NMR-signal samtidig med, at vi kan undgå vekselvirkninger, som typisk forekommer, når koncentrationerne varieres. Man skal dog være opmærksom på, at opdeling af NMR-røret i mindre rum kan føre til artefakter, som stammer fra mellemrummene mellem de små NMR-rør (NMR susceptibility). Desuden kan der opstå mindre artefakter, pga. at forskellige dele af den samme prøve kan opleve en smule anderledes magnetisk felt.

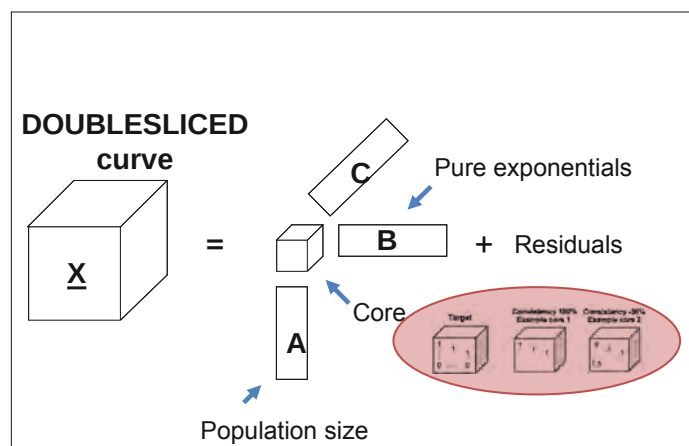
For hver af de 30 prøver målt det transverse relaxations-signal af vandkomponenten vha. en CPMG (Carr-Purcell-Meiboom-Gill) pulssekvens (figur 3). Dette unikke datasæt er tilgængeligt i MATLAB-format fra www.models.life.ku.dk sammen med MATLAB-koden til DOUBLES LICING.

De resulterede 30 NMR-relaxationsprofiler er vist i figur 3, hvor de er farvet efter deres indhold af den hurtigt relaxerende komponent (20 ms). Figur 5, side 24 viser dispersionen af de

■ DET KEMOMETRISKE RUM

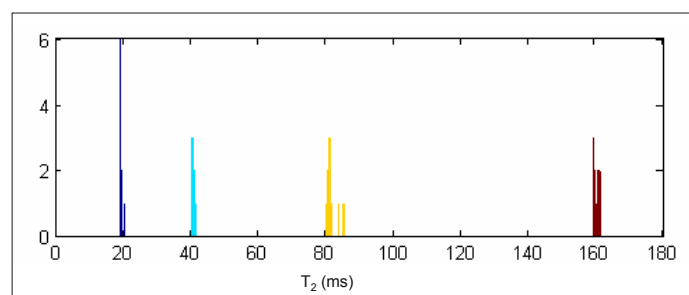
ekstraherede T_2 -værdier ved applikation af DOUBLESILING. Algoritmen (se skitse i figur 4) har ingen problemer med at ekstrahere de fire underliggende rene eksponentielle komponenter fra disse relativt komplekse NMR-henfaldskurver.

Vi har gennemført en række test af DOUBLESLICING, og den viser sig i det store hele at give nogenlunde de samme resultater som en effektiv klassisk kurvefitningsalgoritme. På to afgørende punkter er der imidlertid en forskel.



Figur 4. PARAFAC- modellen som den benyttes i DOUBLESLICING. Den dobbelt-slicede relaxationskurve resolveres til rene mono-eksponentielle komponenter og deres tilsvarende koncentrationer.

For det første er DOUBLESLICING ikke-iterativ, da den har en analytisk løsning og er for disse data minimum en faktor fire hurtigere i beregningstid. Den hurtigere beregningstid giver DOUBLESLICING et stort potentiale indenfor applikationer, hvor mange oligo-eksponentialfunktioner skal analyseres eller i andre tilfælde, hvor omfattende resampling er nødvendig for et sikkert analyseresultat. Den lynhurtige DOUBLESLICING kan også benyttes som et superkvalificeret begyndelsesgæt for en traditionel numerisk kurvefitningsalgoritme, og indenfor Magnetic Resonance Imaging (MRI) kan DOUBLESLICING finde anvendelse som matematisk kontrast af MRI-billeder [3].

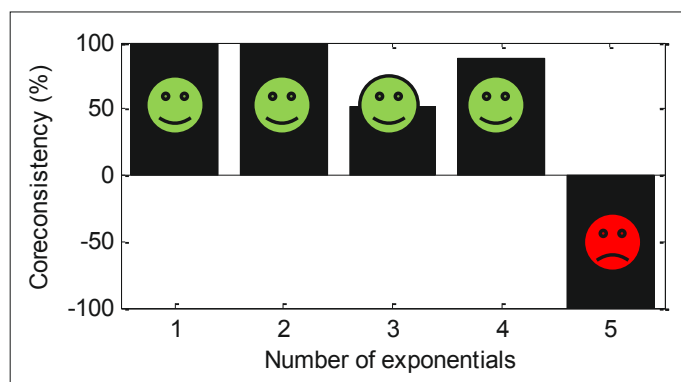


Figur 5. Dispersion af T_2 -værdier fundet ved brug af DOUBLESLICING. Vi ser, at der for de 30 prøver er en overordentlig fin opløsning af de fire underliggende T_2 -komponenter ved hhv. 20, 40, 80 og 160 ms.

For det andet giver brugen af multivejskemometri til analyse af multieksponentielle kurver nogle gode diagnostiske redskaber. Et af hovedproblemerne ved anvendelse af traditionel kurvefitning er at estimere antallet af komponenter. Dette foretages som regel ved at undersøge RMSE (root mean square error) for kurvefittet som en funktion af antallet af komponenter. Men ofte er det subtile forskelle, der afgør, om man skal bruge to eller tre komponenter på trods af, at de resulterende modeller kan være væsentligt forskellige. Ved brug af DOUBLESLICING kan man anvende diverse diagnostiske redskaber som f.eks.

den såkaldte core consistency (figur 6), der samlet kan give et bedre fingerpeg om, hvor mange komponenter der er passende [4]. Normalt ses kun to- eller trekomponent-løsninger indenfor NMR-relaxometri, men som figur 6 antyder, er dette formentlig et firekomponent-system.

Brugen af DOUBLESLICING til multieksponentiel analyse giver således en væsentlig bedre diagnosticering med hensyn til det vigtige valg af antallet af komponenter.



Figur 6. Eksempel på evaluering af PARAFAC core-consistency for en firekomponents-blanding. Værdier tæt på eller under nul, antyder, at for mange komponenter er brugt. Eksemplet viser, at en firekomponents PARAFAC-model beskriver data godt

Outro

Multivejskemometri kan på overraskende vis bidrage med en ny vinkel på multi-eksponentiel analyse. DOUBLESLICING har vist sig at være tilstrækkeligt nøjagtig til at estimere relaxationstider (T_2 -værdier) og deres relaterede koncentrationer. Hovedfordelen ved DOUBLESLICING er imidlertid dens meget kortere beregningstider sammenlignet med de traditionelle kurvefitningsalgoritmer, samt de forbedrede muligheder for det kritiske valg af antal af eksponentialkomponenter i en given profil.

E-mail

Søren Balling Engelsen: se@life.ku.dk

Rasmus Bro: rb@life.ku.dk

Lars Nørgaard: lno@foss.dk

Referencer

1. L. Andrade, E. Micklander, I.A. Farhat, R. Bro & S.B. Engelsen, DoubleSlicing: a non-iterative single profile multi-exponential curve resolution procedure. Application to time-domain NMR transverse relaxation data, *Journal of Magnetic Resonance* (2007), **189**(2), 286-292.
2. W. Windig and B. Antalek, Direct exponential curve resolution algorithm (DECRA): A novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles, *Chemometrics and Intelligent Laboratory Systems* (1997) **37**, 241-254
3. H.F. Seefeldt, F. van den Berg, W. Köckenberger, S.B. Engelsen and B. Wollenweber, Water mobility in the endosperm of high beta-glucan barley mutants as studied by Nuclear Magnetic Resonance Imaging, *Magnetic Resonance Imaging* (2007), **25**(3), 425-432.
4. R. Bro and H.A.L. Kiers, A new efficient method for determining the number of components in PARAFAC models, *Journal of Chemometrics*, **17**(5), 274-286.