OPTIMÁLNÍ NÁVRH MĚŘENÍ SIGMOIDÁLNÍCH FUNKCÍ

P. Tuček[1,2], M. Tučková[2] a R. Harman[3]

[1]Regionální centrum pokročilých technologií a materiálů
Univerzita Palackého v Olomouci

[2]Katedra geoinformatiky
Univerzita Palackého v Olomouci

[3]Katedra aplikovanej matematiky a štatistiky
Univerzita Komenského, Bratislava

Robust 2012
9. – 14. září 2012, Němčičky

Obsah prezentace

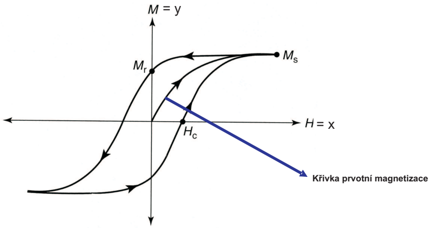
Motivace

- Nanomateriály jsou 1,2,3 - dimenzionálně vymezené prostorové útvary, vyplněné nebo obklopené hmotou, mající unikátní vlastnosti takové, které se nevyskytují u jejich makroskopických protějšků.
- Vznik těchto specifických mechanických, elektrických, optických a magnetických vlastností je spojen se zmenšováním rozměru magnetického materiálu pod hranici 100*nm*.
- Z důvodu velkého aplikačního potenciálu je analýza magnetických vlastností nanomateriálů jedním z hlavních cílů výzkumu RCPTM.
- Magnetické jevy nanomateriálů jsou řízeny dvěma faktory:
    1. kvantově-mechanické jevy $\rightarrow$ spojené s konečným rozměrem částic,
    2. povrchové jevy $\rightarrow$ fenomén rostoucího počtu atomů v povrchových vrstvách spojený s klesajícím rozměrem magnetické částice.

Fyzikální podklady

Magnetizace se měří pomocí magnetometrů dvěma způsoby:

- při různé teplotě, ale s konstantním vnějším magnetickým polem $\rightarrow$ tzv. teplotní závislost magnetizace
- při konstantní teplotě, ale s měnící se intenzitou vnějšího magnetického pole $\rightarrow$ tzv. polní závislost magnetizace $\rightarrow$ **hysterézní smyčka**



M_S - saturační magnetizace
M_r - remanentní magnetizace
H_C - koercitivní síla

Obr. 1: Hysterézní smyčka

Proces měření

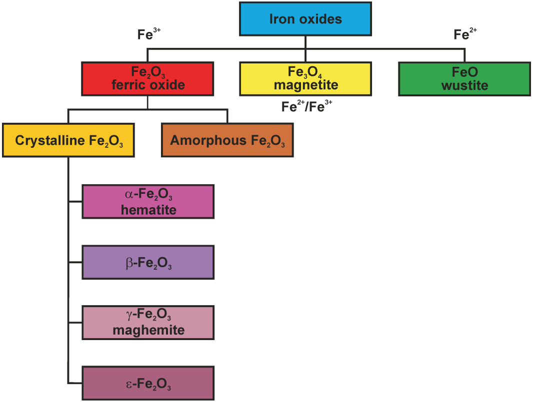
Obr. 2: Magnetometr

Měřený nanomateriál je vystaven vnějšímu magnetickému poli (VMP), kde na něj působí generující cívky elektromagnetu. Snímací cívky potom snímají magnetizaci nanomateriálu v příslušných bodech VMP.

- vzorek je nejprve zatížen maximální kladnou hodnotou VMP, která se postupně snižuje až je dosaženo maximální záporné hodnoty VMP $\rightarrow$ **horní větev hysterézní smyčky**
- poté obdobným způsobem avšak postupem od maximální záporné hodnoty VMP k maximální kladné hodnotě VMP $\rightarrow$ **dolní větev hysterézní smyčky**

Fyzikální podklady

Jedna z výzkumných skupin RCPTM pracuje s nanočásticemi oxidů železa, resp. s nanočásticemi na bázi oxidu železitého. Oxid železitý existuje ve 4 strukturních modifikacích:

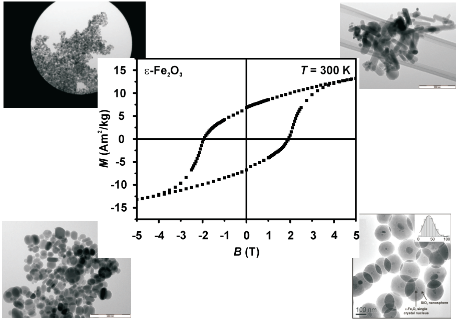


Obr. 3: Strukturní modifikace oxidu železitého.

Fyzikální podklady

Nyní se zabýváme analýzou nanočástic $\epsilon - Fe_2O_3$ označovaných jako nanomateriály nové generace záznamových médií.

Hysterézní smyčka těchto vzorků nanočástic má symetrickou horní a dolní větev $\rightarrow$ dále pracujeme pouze s horní větví smyčky.



Obr. 4: Epsilon fáze oxidu železitého.

Popis modelu

K aproximaci hysterézních smyček se užívají dvě nelineární funkce:

Langevinova funkce - horní větev smyčky

$$L(\mathbf{x}, \boldsymbol{\theta}) = \theta_1 \cdot \coth\left[\frac{\theta_2 \cdot (\mathbf{x} + \theta_3)}{k_B \cdot T}\right] - \theta_1 \cdot \frac{k_B \cdot T}{\theta_2 \cdot (\mathbf{x} + \theta_3)},$$

kde:

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\theta}) & \quad \ldots \quad \text{magnetizace,} \\ \mathbf{x} = \{x_1, \ldots, x_n\} & \quad \ldots \quad \text{množina experimentálních bodů,} \\ \boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) & \quad \ldots \quad \text{vektor neznámých parametrů,} \\ k_B = 1{,}38 \cdot 10^{-23} J/K & \quad \ldots \quad \text{známá fyzikální konstanta,} \\ T = 300K & \quad \ldots \quad \text{známá fyzikální konstanta.} \end{aligned}$$

Popis modelu

Brillouinova funkce - horní větev smyčky

$$B(\mathbf{x},\boldsymbol{\theta}) = \theta_1 \cdot \frac{2J+1}{2J} \cdot \coth\left[\frac{2J+1}{2J} \cdot \frac{g_J \cdot \theta_2 \cdot J}{k_B \cdot T}(\mathbf{x}+\theta_3)\right] - \theta_1 \cdot \frac{1}{2J} \cdot \coth\left[\frac{1}{2J} \cdot \frac{g_J \cdot \theta_2 \cdot J}{k_b \cdot T}(\mathbf{x}+\theta_3)\right],$$

kde:

$B(\mathbf{x},\boldsymbol{\theta})$ ... magnetizace,

$\mathbf{x} = \{x_1,\ldots,x_n\}$ ... množina experimentálních bodů,

$\boldsymbol{\theta} = (\theta_1,\theta_2,\theta_3)$ ... vektor neznámých parametrů,

$k_B = 1{,}38 \cdot 10^{-23} J/K$ ... známá fyzikální konstanta,

g_J ... známá fyzikální konstanta,

$J = 5/2$ ... známá fyzikální konstanta,

$T = 300K$ ... známá fyzikální konstanta.

Hlavní cíle:

1. Navrhnout optimální návrh měření za účelem získání co nejpřesnějších odhadů hodnot neznámých parametrů $\theta_1, \theta_2, \theta_3$, které jednoznačně charakterizují konkrétní vzorek nanomateriálu a rozhodují o jeho následné aplikaci v praxi.

Hlavní cíle:

1. Navrhnout optimální návrh měření za účelem získání co nejpřesnějších odhadů hodnot neznámých parametrů $\theta_1, \theta_2, \theta_3$, které jednoznačně charakterizují konkrétní vzorek nanomateriálu a rozhodují o jeho následné aplikaci v praxi.
2. Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace.

Hlavní cíle:

1. Navrhnout optimální návrh měření za účelem získání co nejpřesnějších odhadů hodnot neznámých parametrů $\theta_1, \theta_2, \theta_3$, které jednoznačně charakterizují konkrétní vzorek nanomateriálu a rozhodují o jeho následné aplikaci v praxi.
2. Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace.
3. Odhadnout obsah plochy, která vzniká mezi horní a dolní smyčkou.

Sigmoidální funkce

Z praktických důvodů zavádíme univerzální model

$$\phi(\mathbf{x},\boldsymbol{\theta}) = \theta_1 \cdot h(\mathbf{z}) = \theta_1 \cdot h\left[\theta_2 \cdot (\mathbf{x} + \theta_3)\right],$$

kde pro Langevinovu funkci má sigmoidální funkce $h(\mathbf{z})$ tvar

$$h(\mathbf{z}) = \coth\left[\frac{\mathbf{z}}{k_B \cdot T}\right] - \frac{k_B \cdot T}{\mathbf{z}}$$

a pro Brillouinovu funkci je sigmoidální funkce $h(\mathbf{z})$ tvaru

$$h(\mathbf{z}) = \frac{2J+1}{2J} \cdot \coth\left[\frac{(2J+1)\cdot g_J \cdot J \cdot \mathbf{z}}{2J \cdot k_B \cdot T}\right] - \frac{1}{2J} \cdot \coth\left[\frac{g_J \cdot J \cdot \mathbf{z}}{2J \cdot k_B \cdot T}\right].$$

Linearizace modelu

Máme tedy nelineární regresní model měření magnetizace:

$$\mathbf{Y} = \phi(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\epsilon},$$

kde

$\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^T \in \Theta$ ... vektor neznámých parametrů,

$\Theta = (\Theta_1 \times \Theta_2 \times \Theta_3)$ ... množina "možných" hodnot parametrů,

$\boldsymbol{\epsilon}$ ... vektor chyb měření: $\mathrm{E}(\boldsymbol{\epsilon}) = 0,\ \mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2 \cdot \mathbf{I}$.

Užitím Taylorova rozvoje se zanedbáním členů 2. a vyšších řádů:

$$\mathbf{Y} \sim \phi(\mathbf{x}, \boldsymbol{\theta}^0) + \underbrace{\frac{\partial \phi(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}}_{\mathbf{F}} \cdot \delta\boldsymbol{\theta} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} + \boldsymbol{\epsilon}$$

$$\mathbf{Y} - \phi(\mathbf{x}, \boldsymbol{\theta}^0) = \mathbf{F} \cdot \delta\boldsymbol{\theta} + \boldsymbol{\epsilon},$$

dostáváme lineární regresní model měření magnetizace.

Základní pojmy

Definice

Normovanou informační maticí pro parametr $\boldsymbol{\theta}$ nazveme matici

$$\mathbf{M}_{\boldsymbol{\theta}}(\xi) = \sum_{x_i;\xi(x_i)>0} \xi(x_i) \left(\frac{\partial \phi(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \phi(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} ; \ \boldsymbol{\theta} \in \Theta.$$

kde

$\xi(x_i)$ ... návrh, tj. pravděpodobnostní míra, na množině experimentálních bodů $\mathbf{x}$.

Definice

Kritérium optimality je funkce $\Phi : \mathbf{M}_{\boldsymbol{\theta}}(\xi) \mapsto \Phi\left[\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right] \in R$ *taková, že jestliže je* $\xi \prec \tau$, *potom* $\Phi\left[\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right] \leq \Phi\left[\mathbf{M}_{\boldsymbol{\theta}}(\tau)\right]$.

Kritérium lokální D-optimality

Označíme-li $\widehat{\boldsymbol{\theta}}$ jako NLNO parametru $\boldsymbol{\theta}$, pak je snadné dokázat platnost vztahu:

$$\mathrm{Var}(\widehat{\boldsymbol{\theta}}) = \left[\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right]^{-1}.$$

S cílem získat co nejpřesnější odhad vektorového parametru $\widehat{\boldsymbol{\theta}} \to$ volíme kriteriální funkci lokální D-optimality, jejíž základní vlastností je minimalizace objemu konfidenčního elipsoidu. Definováno jako

$$\Phi\left[\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right] = \det\left[\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right].$$

Protože uvažuje kritérium v jeho konkávní podobě, musí lokálně D-optimální návrh ξ^* splňovat podmínku:

$$\det\left(\mathbf{M}_{\boldsymbol{\theta}}(\xi^*)\right) = \arg\max_{\xi\in\Xi} \det\left(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right),$$

kde Ξ je množina všech návrhů.

Kritérium lokální D-optimality

Je snadné ověřit, že

$$\det\left(\mathbf{M}_{\theta_1,\theta_2,\theta_3}(\xi)\right) = \theta_1^4 \det\left(\mathbf{M}_{1,\theta_2,\theta_3}(\xi)\right),$$

pro jakékoliv $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) \in \Theta$ a jakýkoliv návrh $\xi \in \Xi$.

Současně platí

$$\arg\max_{\xi\in\Xi} \det\left(\mathbf{M}_{\theta_1,\theta_2,\theta_3}(\xi)\right) = \arg\max_{\xi\in\Xi} \det\left(\mathbf{M}_{1,\theta_2,\theta_3}(\xi)\right)$$

a proto D-optimální návrh nezávisí na hodnotě parametru $\theta_1 \in \Theta_1$, ale závisí pouze na hodnotách parametrů $\theta_2 \in \Theta_2$ a $\theta_3 \in \Theta_3$.

$$\Downarrow$$

PROBLÉM !!!

Závislost lokálně D-optimálního návrhu měření na bodech linearizace (θ_2^0 a θ_3^0) uvažovaného nelineárního modelu.

Parametrické prostory

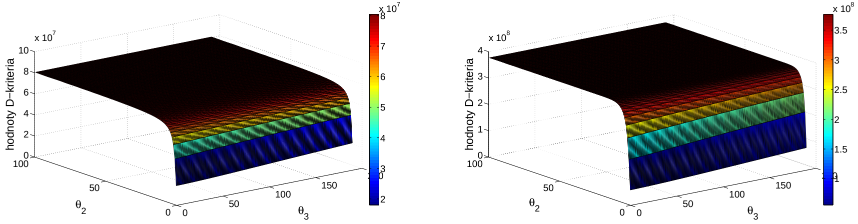
Vymezení přípustného parametrického prostoru

Langevinovy funkce:

$$\theta_2 = \left[1 \cdot 10^{-18} : 1 \cdot 10^{-18} : 1 \cdot 10^{-16}\right] \quad ; \quad \theta_3 = [1000 : 100 : 20000]$$

Brillouinovy funkce:

$$\theta_2 = \left[1 \cdot 10^{-19} : 1 \cdot 10^{-19} : 1 \cdot 10^{-17}\right] \quad ; \quad \theta_3 = [1000 : 100 : 20000]$$



Obr. 5: Hodnoty lokálně D-optimálních návrhů v bodech přípustných parametrických prostorů Langevinovy a Brillouinovy funkce.

D-eficience

V konkrétním bodě parametrického prostoru je možné výkonnost libovolného návrhu ξ vzhledem k lokálně optimálnímu návrhu ξ^* vyjádřit pomocí jeho D-eficience *(Pukelsheim F. (1993))*:

$$\mathrm{eff}_{\boldsymbol{\theta}}(\xi) = \frac{\det\left(\mathbf{M}_{\boldsymbol{\theta}}(\xi)\right)^{1/k}}{\det\left(\mathbf{M}_{\boldsymbol{\theta}}(\xi^*)\right)^{1/k}},$$

kde k je počet neznámých parametrů modelu.

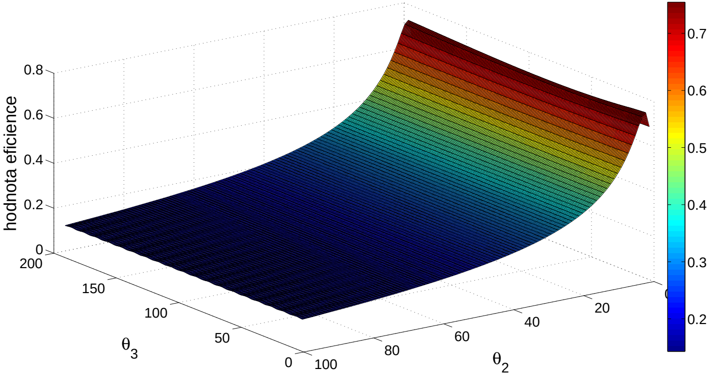
Tímto postupem můžeme například stanovit eficienci rovnoměrného návrhu (tj. měření magnetizace v každém bodě množiny experimentálních bodů) vzhledem k lokálně D-optimálnímu návrhu.

$$\downarrow$$

Popis současného stavu procesu měření magnetizace nejen $\epsilon - Fe_2O_3$. Takto prováděné měření je velice nákladné a to jak časově, tak finančně.

Eficience rovnoměrných návrhů měření Langevinovy funkce

Maximální hodnota eficience je 75,4% avšak minimální hodnota eficience je 14,3%.



Zavedení pojmu

Hledáním ”ucházející” hodnoty eficience libovolně volených návrhů ξ v přípustných parametrických prostorech Langevinovy a Brillouinovy funkce jsme narazili na zvláštní fenomén.

Jedná se o **PRAVIDELNÝ** výskyt hodnoty minimální eficience na okrajích přípustného parametrického prostoru.

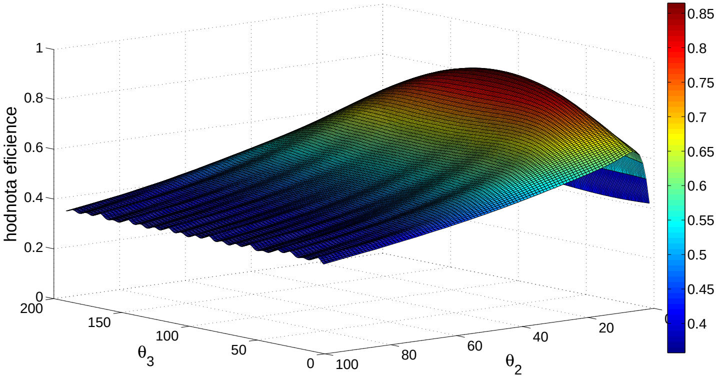
Z tohoto důvodu jsme přistoupili k možnosti nalezení maximin eficentního návrhu, který je definován vztahem *(Müller CH., Pázman A. (1998))*

$$\xi^*_{ME} \in \arg\max_{\xi \in \Xi} \min_{\theta \in \Theta} \mathrm{eff}_\theta(\xi).$$

Pomocí výpočetních algoritmů hill climbing a simulovaného žíhání jsme prohledávali okraje obou přípustných parametrických prostorů abychom tak nalezli návrhy, které jsou stabilní vzhledem k uvažovaným parametrickým prostorům a dosahují přijatelných hodnot minimální eficience.

Maximin eficientní návrh - Langevinova funkce

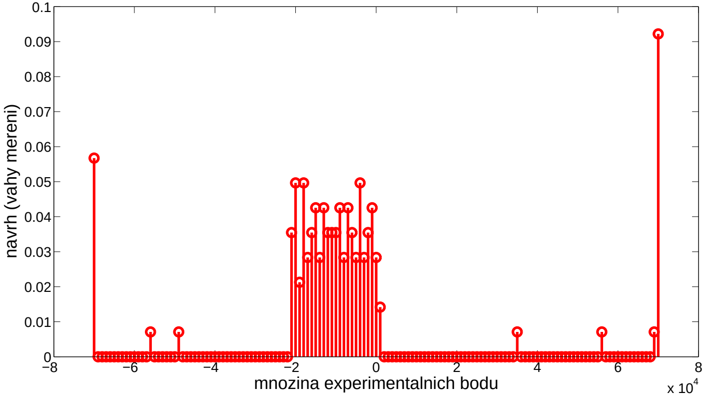
Doposud nejlepším dosaženým výsledkem je minimální hodnota eficience 35,80%.



Obr.7: Maximin eficientní návrh měření Langevinovy funkce.

Maximin eficientní návrh - Langevinova funkce

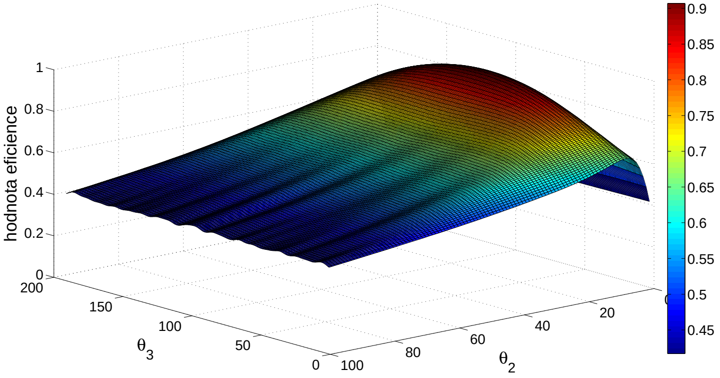
Pro dosažení zmíněné minimální hodnoty eficience je nutné provědět měření v 30 optimálních bodech množiny **x**.



Obr.8: Váhy v optimálních bodech měření Langevinovy funkce.

Maximin eficientní návrh - Brillouinova funkce

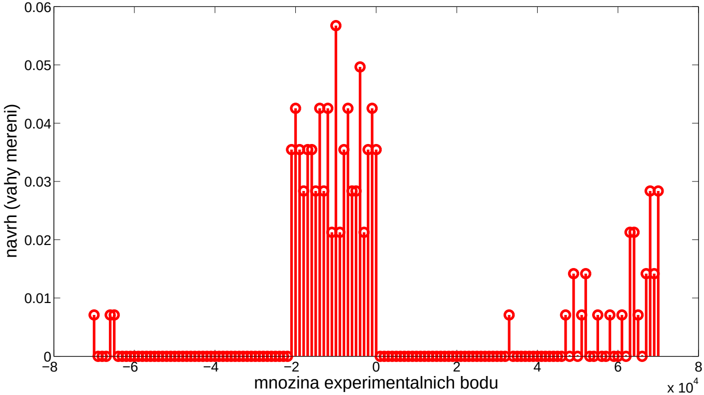
Doposud nejlepším dosaženým výsledkem je minimální hodnota eficience 41,70%.



Obr.9: Maximin eficientní návrh měření Brillouinovy funkce.

Maximin eficientní návrh - Brillouinova funkce

Pro dosažení zmíněné minimální hodnoty eficience je nutné provědět měření v 40 optimálních bodech množiny **x**.



Obr.10: Váhy v optimálních bodech měření Brillouinovy funkce.

Postup dalšího zpracování

Jedním z hlavních cílů analýzy je: ”Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace.”

- S využitím T-kriteriální funkce nebo DT- kriteriální funkce chceme navrhnout vhodný plán měření, díky němuž bude možné toto rozhodnutí učinit.

Postup dalšího zpracování

Jedním z hlavních cílů analýzy je: "Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace."

- S využitím T-kriteriální funkce nebo DT- kriteriální funkce chceme navrhnout vhodný plán měření, díky němuž bude možné toto rozhodnutí učinit.
- Atkinson A.C., Donev A.N. (1992): *Optimum Experimental Designs.* Oxford University Press.

Postup dalšího zpracování

Jedním z hlavních cílů analýzy je: ”Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace.”

- S využitím T-kriteriální funkce nebo DT- kriteriální funkce chceme navrhnout vhodný plán měření, díky němuž bude možné toto rozhodnutí učinit.
- Atkinson A.C., Donev A.N. (1992): *Optimum Experimental Designs.* Oxford University Press.
- Atkinson A.C. (2008): *DT-optimum designs for model discrimination and parameter estimation.* Journal of Statistical Planning and Inference.

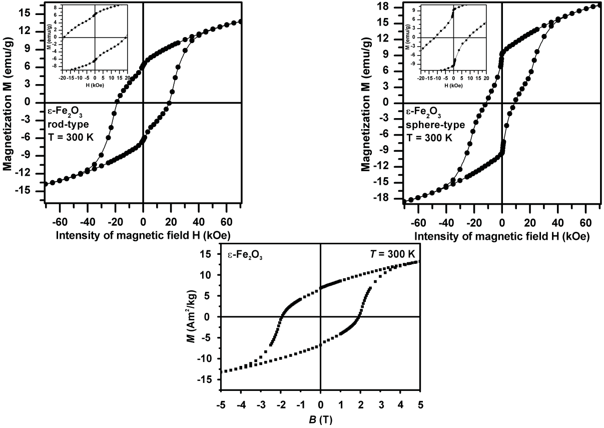
Postup dalšího zpracování

Jedním z hlavních cílů analýzy je: "Navrhnout optimální návrh měření tak aby bylo možné jednoznačně rozhodnout, kterou funkci (Langevinovu nebo Brillouinovu) použít k aproximaci naměřených dat - naměřené magnetizace."

- S využitím T-kriteriální funkce nebo DT- kriteriální funkce chceme navrhnout vhodný plán měření, díky němuž bude možné toto rozhodnutí učinit.
- Atkinson A.C., Donev A.N. (1992): *Optimum Experimental Designs.* Oxford University Press.
- Atkinson A.C. (2008): *DT-optimum designs for model discrimination and parameter estimation.* Journal of Statistical Planning and Inference.
- Jednou z možností je provádět měření horní větve smyčky dle T-optimálního návrhu a následně dle odpovídajícího modelu (Langevin x Brillouin) provádět měření dolní větve hysterézní smyčky dle D-optimálního návrhu.

Postup dalšího zpracování

Analýza "smíšené" fáze $\epsilon - Fe_2O_3$, kde se v neznámém množství vyskytuje fáze $\gamma - Fe_2O_3$, což vede ke vzniku nežádoucích zubů.



Obr.11: Nanomateriály smíšených fází.

Literatura

1. Atkinson A.C., Donev A.N. (1992). *Optimum Experimental Designs.* Oxford University Press, Oxford.
2. Darby M. (1967). *Tables of the Brillouin function and of the related function for the spontaneous magnetization.* British Journal of Applied Physics **18**, 1415 – 1417.
3. Müller CH., Pázman A. (1998). *Application of necessary and sufficient conditions for maximin efficient designs*. Metrika **48**, 1 – 19.
4. Pukelsheim F. (1993). *Optimal design of experiments*. Wiley and Sons, New York.

DĚKUJI VÁM
ZA VAŠI POZORNOST.

Pavel Tuček

pavel.tucek@upol.cz