

DETERMINATION OF Q-FUNCTION OPTIMUM GRID APPLIED ON ASYNCHRONOUS ELECTRIC MOTOR CONTROL TASK

Tomáš Marada¹, Tomáš Březina², Vladislav Singule³

Summary: Asynchronous electric motor control task can be successfully solved using reinforcement learning based method called Q-learning. The main problem to solve is the convergence speed. Two-phase Q-learning can be used to speed up the learning process. Efficient prelearning phase uses mathematical model and next phase using for tutoring real system. This method can increase learning speed significantly. When the table is used as Q-function approximation, the learning speed and precision of found controllers depend highly on the Q-function table grid properties. The paper is focused on finding the optimal division of grid.

Klíčová slova: Asynchronní elektromotor, Q-učení, řízení

1. Úvod

Pro návrh metod řízení zlepšujících kvalitu řízení asynchronního elektromotoru, případně snižujících nároky na řídicí elektroniku, mohou být použity metody umělé inteligence, které využívají strojové učení v reálném čase. Vzhledem k tomu, že je poměrně obtížné získat množinu vzorů požadovaného chování regulátoru, což je nezbytná podmínka při učení s učitelem, nabízí se použití metod opakovaně posilovaného učení, které tuto množinu vzorů nevyžadují. Jednou z možných variant opakovaně posilovaného učení je takzvané Q-učení. Hlavním problémem v metodách opakovaně posilovaného učení je rychlost učení. Jednou z možností, jak urychlit proces učení, je využití dvoufázového Q-učení. Proces učení je při dvoufázovém Q-učení rozdělen na fázi předučení a fázi doučování. Fáze předučení vyžaduje výpočtový model, ale je velmi efektivní. Fáze doučování používá konvenčního Q-učení a předpokládá interakci s řízenou soustavou. Pokud je Q-funkce reprezentována tabulkou (tj. schodovitou funkcí), je rychlost a kvalita naučení značně závislá na vlastnostech tabulky. V příspěvku je studována problematika optimálního rastru této tabulky.

¹ Tomáš Marada, Ing., VUT v Brně, FSI ÚMT, Technická 2, 616 69 Brno, marada@kn.vutbr.cz

² Tomáš Březina, Doc., Ing., RNDr., CSc., VUT v Brně, FSI ÚAI, Technická 2, 616 69 Brno, brezina@uai.fme.vutbr.cz

³ Vladislav Singule, Doc., Ing., CSc., VUT v Brně, FSI ÚVSSaR, Technická 2, 616 69 Brno, singule@uvssr.fme.vutbr.cz

2. Implementace Q-funkce

Popis Q učení byl již publikován v mnoha člancích, například [1], a proto se jím již nebudeme zabývat. Stav prostředí s je chápán jako aritmetický vektor jistých veličin prostředí $\mathbf{s} = (s^1, \mathbf{K}, s^n)$, a podobně akce u jako aritmetický vektor řídicích veličin $\mathbf{u} = (u^1, \mathbf{K}, u^m)$. Necht' $\{g_j^i\}_{j=0}^{p_i}$ je posloupnost uzlových bodů rastru i -té veličiny prostředí z s^i taková, že $g_j^i < g_{j'}^i \Leftrightarrow j < j'$, $j, j' = 0, \mathbf{K}, p_i$. Hodnotou indexu, který reprezentuje hodnotu i -té veličiny prostředí vzhledem k rastru, nazveme takový index j , pro který $g_j^i \leq s^i < g_{j+1}^i$. Zcela analogicky je zavedena reprezentace akce skládající se z řídicích veličin s uzlovými body rastru $\{h_j^i\}_{j=0}^{p_i}$. Funkce $Q(\mathbf{s}, \mathbf{u})$ je potom reprezentována tabulkou jako

$$Q(\mathbf{s}, \mathbf{u}) = q_{j^1 j^2 \mathbf{K} j^{n+m}}, \quad (1)$$

kde j^i reprezentuje hodnotu i -té veličiny prostředí (resp. $i-n$ -té řídicí veličiny).

Přiměřenost aproximace funkce $Q(\mathbf{s}, \mathbf{u})$ tabulkou bude zřejmě ovlivněna konkrétní volbou uzlových bodů rastru. Bude-li polí rastru málo, půjde zřejmě o příliš hrubou aproximaci, bude-li rastr příliš jemný, stane se neúnosným požadavek modifikací Q-hodnot nad značně vysokým počtem polí. V prvním případě je vyloučeno dosažení funkce $Q^*(\mathbf{s}, \mathbf{u})$ s dostatečnou přesností, v druhém případě toto není dosažitelné bez zdlouhavého výpočtu.

Dalším faktorem je to, že „kvalita“ strategie řízení bude zřejmě záviset i na velikosti a poloze (tj. uzlových bodech) jednotlivých polí rastru. Pro množinu stavů s požadavkem „jemnějších“ řídicích zásahů bude třeba „jemnějšího“ rastru, než pro množinu bez této podmínky.

Běžnou strategií při použití Q-učení je využívat k učení všech stavů prostředí, které postupně nastanou v průběhu jednotlivých pokusů. Tato strategie se nazývá učení pokusem a byla použita i v našem případě.

3. Aplikace

K simulacím popsaných v dalších odstavcích byl použit známý model asynchronního elektromotoru v transformovaných souřadnicích $d, q, 0$. Model používá rovnice pro statorová napětí ve tvaru:

$$\begin{bmatrix} u_{qs} \\ u_{ds} \\ u_{0s} \end{bmatrix} = R_s \begin{bmatrix} i_{qs} \\ i_{ds} \\ i_{0s} \end{bmatrix} + \frac{d}{dt} \begin{bmatrix} \Psi_{qs} \\ \Psi_{ds} \\ \Psi_{0s} \end{bmatrix}, \quad (2)$$

rovnice pro rotorová napětí ve tvaru:

$$\begin{bmatrix} u_{qr} \\ u_{dr} \\ u_{0r} \end{bmatrix} = R_r \begin{bmatrix} i_{qr} \\ i_{dr} \\ i_{0r} \end{bmatrix} - w_r \begin{bmatrix} \Psi_{dr} \\ -\Psi_{qr} \\ 0 \end{bmatrix} + \frac{d}{dt} \begin{bmatrix} \Psi_{qr} \\ \Psi_{dr} \\ \Psi_{0r} \end{bmatrix}, \quad (3)$$

rovnice pro spřažené magnetické toky rotoru a statoru ve tvaru:

$$\begin{bmatrix} \Psi_{qs} \\ \Psi_{ds} \\ \Psi_{0s} \\ \Psi_{qr}^* \\ \Psi_{dr}^* \\ \Psi_{0r}^* \end{bmatrix} = \begin{bmatrix} L_{ls} + L_m & 0 & 0 & L_m & 0 & 0 \\ 0 & L_{ls} + L_m & 0 & 0 & L_m & 0 \\ 0 & 0 & L_{ls} & 0 & 0 & 0 \\ L_m & 0 & 0 & L_{lr}^* + L_m & 0 & 0 \\ 0 & L_m & 0 & 0 & L_{lr}^* + L_m & 0 \\ 0 & 0 & 0 & 0 & 0 & L_{lr}^* \end{bmatrix} \begin{bmatrix} i_{qs} \\ i_{ds} \\ i_{0s} \\ i_{qr}^* \\ i_{dr}^* \\ i_{0r}^* \end{bmatrix} \quad (4)$$

a rovnici pro elektromagnetický moment stroje ve tvaru:

$$M_e = \frac{3}{2} \frac{P}{2} \left[\left(\Psi_{qr}^* i_{dr}^* - \Psi_{dr}^* i_{qr}^* \right) \right] = \frac{3}{2} \frac{P}{2} (\Psi_{ds} i_{qs} - \Psi_{qs} i_{ds}). \quad (5)$$

Implementace modelu byla provedena v prostředí Borland Delphi a k řešení používá lichoběžníkovou metodu integrace. Z modelu vystupovala aktuální regulační odchylka, ze které byla vypočtena aktuální rychlost a zrychlení regulační odchylky.

4. Provedené experimenty

Jsou popsány dvě sady experimentů: s lineárním rastrem a s nelineárním rastrem, kde jsou podmínky lineárního rastru alespoň v jednom případě porušeny.

Při všech experimentech byly akce agenta brány z množiny $\{0Hz, 50Hz\}$, kde hodnoty „0Hz“ a „50Hz“ udávají požadovanou frekvenci na výstupu frekvenčního měniče. Této hodnoty je postupně dosaženo pomocí rampy frekvenčního měniče.

Posilovací funkce r byla definována jako:

$$r = \begin{cases} 0, & \text{pro } |e_k| < |e_{k+1}|, \text{ nebo } |e_{k+1}| \leq 2 \\ -1, & \text{jinak} \end{cases} \quad (6)$$

kde e_{k+1} je regulační odchylka následující po regulační odchylce e_k .

Pokud je hodnota aktuální regulační odchylky menší než hodnota regulační odchylky v předchozím kroku, nebo pokud je hodnota aktuální regulační odchylky menší nebo rovna 2, obdrží regulátor posílení rovné nule. Ve všech ostatních případech obdrží regulátor posílení rovné -1.

Pro experimenty bylo použito hodnot parametrů modelu uvedených v tab. 1. Tyto byly získány identifikačními technikami z asynchronního elektromotoru.

$R_r [\Omega]$	$R_s [\Omega]$	$L_r [H]$	$L_s [H]$	$L_{rs} [H]$	$J [kg.m^2]$
2.95	4.37	0.476	0.471	0.459	0.0015

Tab. 1.: Parametry asynchronního elektromotoru 4AP 90S-2

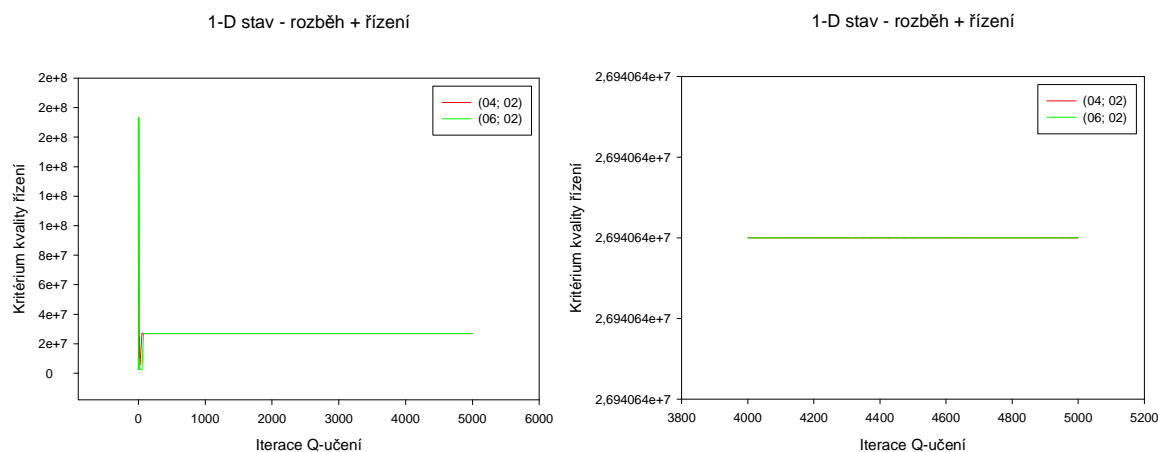
V provedených experimentech byla zjišťována kvalita regulace během 5000 kroků učení. Vždy byl 50x proveden rozběh asynchronního elektromotoru z nulových otáček na hodnotu požadovaných otáček, které byly nastaveny na hodnotu 250 min^{-1} až 2750 min^{-1} s krokem 50 min^{-1} . Simulace trvala 10 sekund a byla rozdělena na 2 fáze. Na fázi rozběhu o délce trvání 9s a na fázi řízení o délce trvání 1s. V obou fázích bylo vypočteno průměrné integrální kritérium kvality regulace.

V legendách následujících grafů udává první číslice počet dílků rastru regulační odchylky e , druhé číslo udává počet dílků rastru její rychlosti \dot{e} a třetí číslo udává počet dílků rastru jejího zrychlení \ddot{e} . Poslední číslo udává počet použitých akcí.

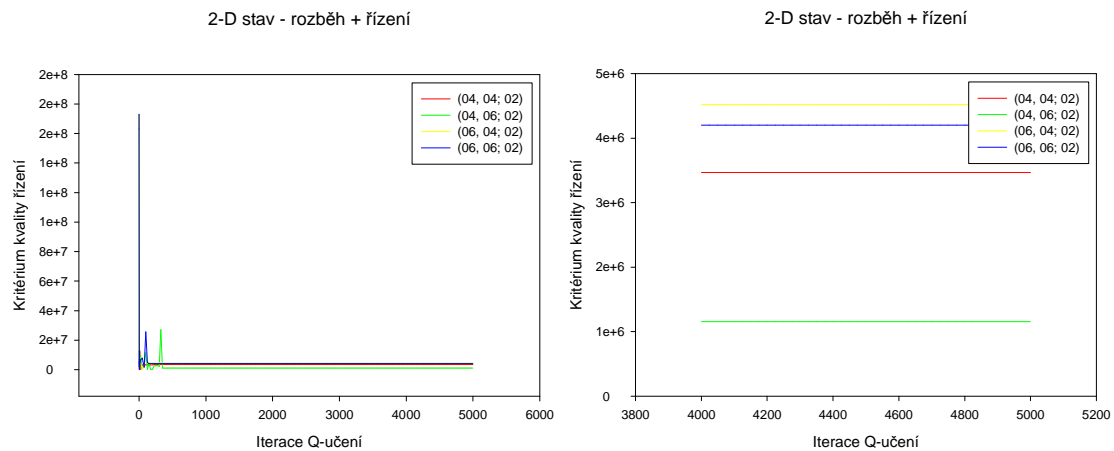
4.1. Určení stavů prostředí

Počáteční testy měly za úkol zjistit, jeví-li se jako vhodnější z hlediska úspěšnosti učení stav prostředí definovaný jako 1D (uvažující pouze regulační odchylku), 2D (uvažující regulační odchylku a její rychlost), nebo jako 3D (uvažující regulační odchylku, její rychlost a zrychlení).

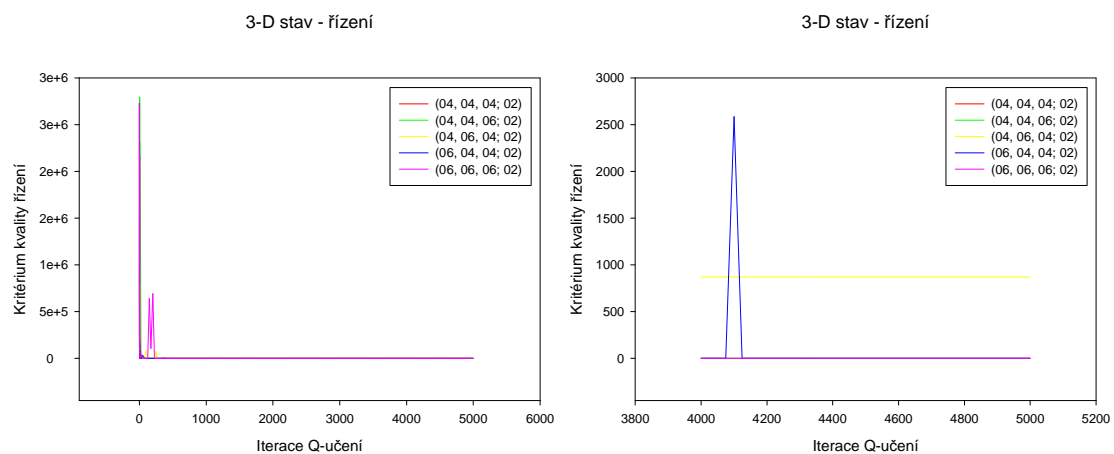
Z výsledků simulačních experimentů shrnutých na obr. 1 až 4 vyplynulo, že pro další experimenty se jeví jako nejnadějnější 3D stav prostředí. Z obr. 3 je zřejmé, že průměrná hodnota integrálního kritéria kvality regulace při řízení (9-10 sekunda simulace) je téměř ve všech případech rovna nule.



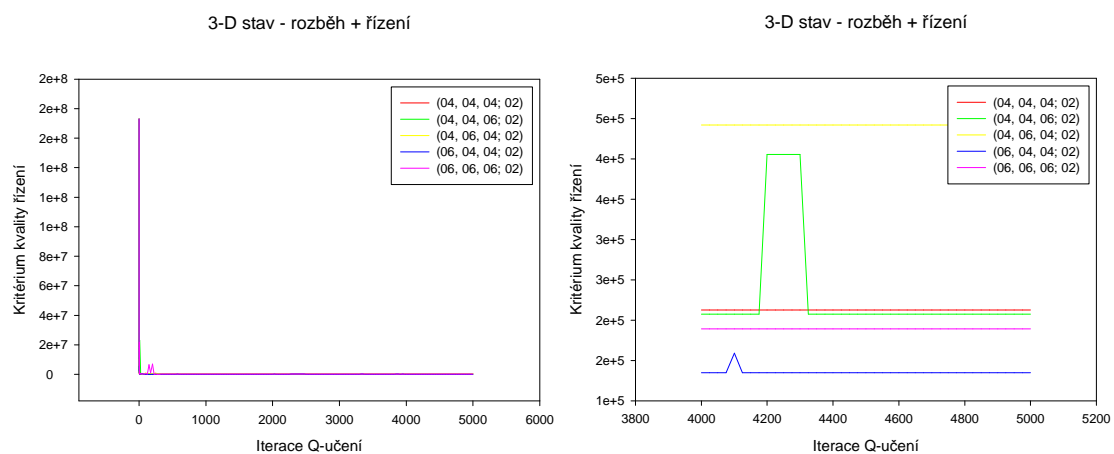
Obr. 1.: 1D stav - Rozběh + řízení AM



Obr. 2.: 2D stav - Rozběh + řízení AM



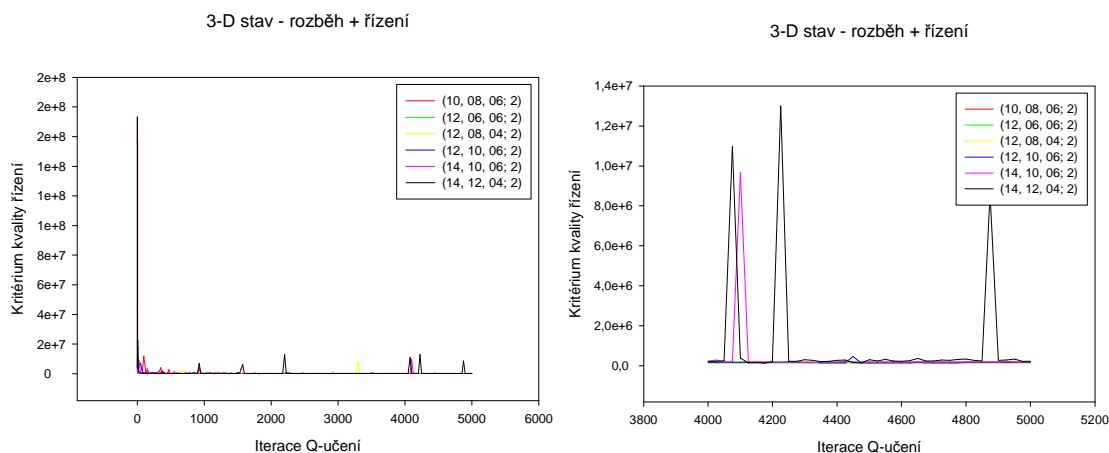
Obr. 3.: 3D stav - Řízení AM



Obr. 4.: 3D stav - Rozběh + řízení AM

4.2. Lineární rastr

Experimenty prováděné v této etapě se týkaly optimalizace počtu intervalů rastru jednotlivých stavových veličin. Jako stavový prostor byl na základě předchozích experimentů, brán 3D stav (e, \dot{e}, \ddot{e}) . Na základě následujících experimentů byl vybrán lineární rastr $(12,08,04)$, který dával nejlepší výsledky (dosahoval nejnižší hodnoty integrálního kritéria kvality regulace). Výsledky jsou shrnuty na obr. 5.

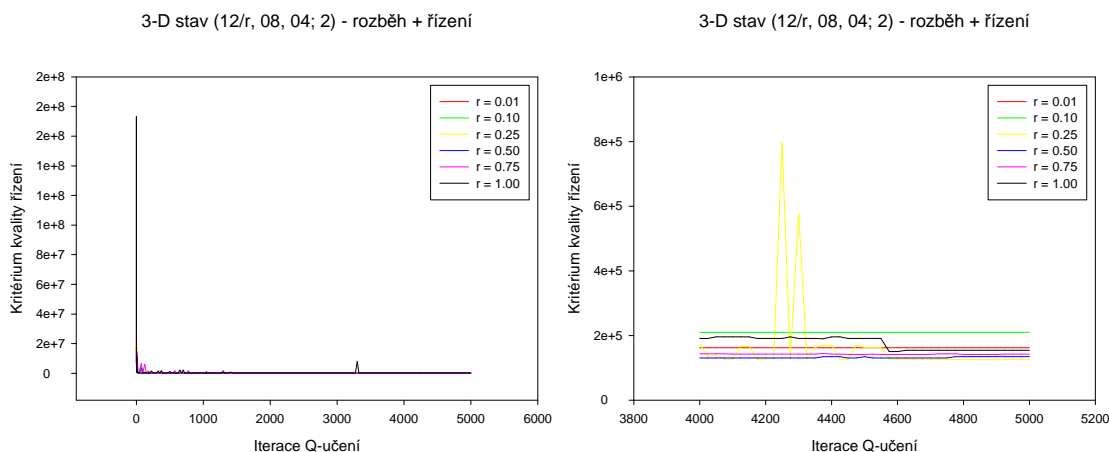


Obr. 5.: Vliv volby lineárního rastru na průběh naučení

4.3. Nelineární rastr

Další část experimentů se týkala nelineárního dělení rastrů, které by při zachování počtu dílků rastrů jednotlivých veličin mohlo přinést další zlepšení průběhu učení. Vycházelo se z nejlepšího lineárního rastru $(12,08,04)$. Nelinearita rastru byla nastavována hodnotou parametru $r \in (0,1)$ úměrně členu $r^{|\Delta i|}$, kde $|\Delta i|$ značí vzdálenost indexu i -tého dílku rastru od indexu středového dílku (tj. $r = 1$ odpovídá lineárnímu rastru).

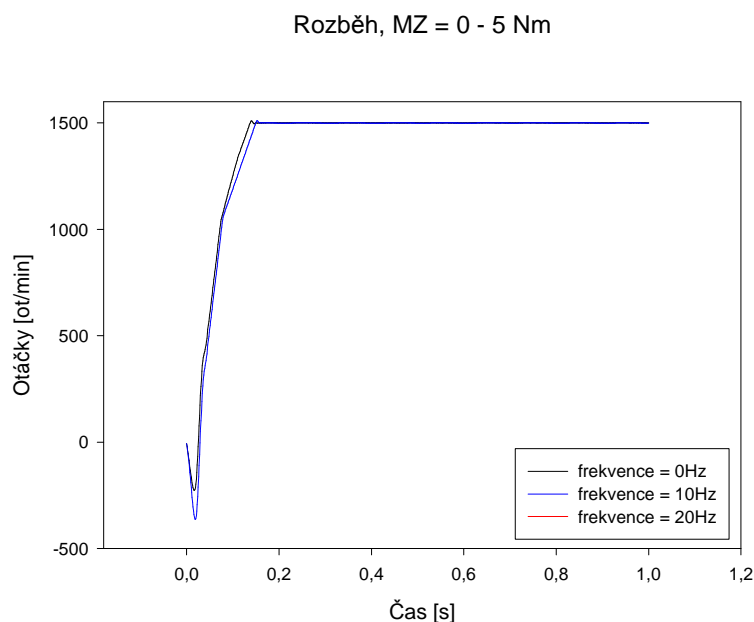
V legendách následujících grafů je původní značení rozšířeno o hodnotu tohoto parametru. Výsledky jsou shrnuty na obr. 6.



Obr. 6: Vliv volby nelineárního rastru na průběh naučení

Nelineární rast regulační odchylky přinesl zlepšení kvality regulace, které je nejmarkantnější pro hodnotu $r = 0.01$ (viz. obr. 6). Naproti tomu ani nelineární rastr rychlosti, ani zrychlení regulační odchylky nepřinesly další zlepšení kvality regulace.

Na obr. 7 je zobrazen průběh regulace motoru, který je zatěžován sinusovým zátěžným momentem v rozmezí $0-5Nm$ a proměnnou frekvencí. Při frekvenci $0Hz$ byl zátěžný moment roven po celou dobu simulace $2.5Nm$. K rozběhu motoru došlo za necelých $0.15s$ a otáčky byly regulovány s nepatrnou regulační odchylkou.



Obr. 7.: Rozběh AM se zátěžným momentem $0-5Nm$

5. Závěr

Metoda využívá schopnosti učení postupně zlepšovat chování asynchronního elektromotoru. Z výsledků experimentů vyplynulo, že 3D stav prostředí je výhodnější než 1D i 2D stav. Prostorové prostředí reprezentované modelem asynchronního elektromotoru je dobře říditelné nelineární tabulkou o velikosti $(12,08,04)$ s koeficientem nelinearity $r = 0.01$. Výsledky provedených simulací rozběhu a řízení (viz. obr. 7) jsou velmi slibné.

Poděkování:

Tato práce je podporována projekty MSM 262100024 a CEZ: J22/98:261100009.

Použitá literatura:

- [1] Březina T., Krejsa J.: The control of active magnetic bearing using two-phase Q-learning, Engineering Mechanics 2003, Svatka 2003.
- [2] Marada T., Singule V.: Využití Q-učení pro řízení asynchronního elektromotoru, Vybrané problémy elektrických strojů a pohonů, Lomnice u Tišnova 2003.
- [3] Sutton, Richard S., Barto, Andrew G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, 1998.