Noncooperative Model Predictive Control for Affine-Quadratic Games

Marleen Stieler[1,*], **Michael H. Baumann**[1], and **Lars Grüne**[1]

[1] Chair of Applied Mathematics, Department of Mathematics, University of Bayreuth, 95440 Bayreuth

Nash strategies are a natural solution concept in noncooperative game theory because of their 'stable' nature: If the other players stick to the Nash strategy it is never beneficial for one player to unilaterally change his or her strategy. In this sense, Nash strategies are the only reliable strategies.

The idea to perform and analyze Model Predictive Control (MPC) based on Nash strategies instead of optimal control sequences is appealing because it allows for a systematic handling of noncooperative games, which are played in a receding horizon manner. In this paper we extend existence and uniqueness results on Nash equilibria for affine-quadratic games. For this class of games we moreover state sufficient conditions that guarantee trajectory convergence of the MPC closed loop.

1 Setting and Preliminary Result

In this paper we are considering dynamic s-player games, $s \in \mathbb{N}$, in discrete time with affine dynamics

$$x(k+1, x_0) = f(x(k, x_0), u(k)) = Ax(k, x_0) + \sum_{i=1}^{s} B_i u_i(k) + c, \text{ or briefly } x^+ = Ax + \sum_{i=1}^{s} B_i u_i + c, \tag{1}$$

in which $A \in \mathbb{R}^{n\times n}$ and $B_i \in \mathbb{R}^{n\times m_i}$. The state $x \in \mathbb{R}^n$ and controls $u_i \in \mathbb{R}^{m_i}$ are unconstrained. Each player $i \in \{1, \dots, s\}$ can influence the common system dynamics (1) through his or her input u_i and acts according to the stage costs

$$\ell_i(x, u_i) = \frac{1}{2}\left[(x - x_i^*)^T Q_i (x - x_i^*) + u_i^T R_i u_i\right] \tag{2}$$

for positive semidefinite Q_i ($Q_i \succeq 0$), positive definite R_i ($R_i \succ 0$), and a desired state $x_i^* \in \mathbb{R}^n$. The definiteness assumptions can be weakened. If this affine-quadratic game is played over $N \in \mathbb{N}$ time stages, each player aims to minimize the cost functional $J_i^N : \mathbb{R}^n \times (\mathbb{R}^{m_1})^N \times \dots \times (\mathbb{R}^{m_s})^N \to \mathbb{R}$ defined by $J_i^N(x, \mathbf{u}) = \sum_{k=0}^{N-1} \ell_i(x^{\mathbf{u}}(k, x), u_i(k))$ along the solution $x^{\mathbf{u}}(\cdot, x)$ to (1) for initial value x. Note that the cost functional J_i^N of player i does not only depend on $\mathbf{u}_i \in (\mathbb{R}^{m_i})^N$ but also on the control sequences of the other players that enter indirectly through the system dynamics.

Definition 1.1 (Nash Equilibrium) A control sequence $\mathbf{u}^{e,N} \in (\mathbb{R}^{m_1})^N \times \dots \times (\mathbb{R}^{m_s})^N$ is said to be a *Nash equilibrium (NE)* of length N for initial value $x \in \mathbb{R}^n$ if for all $i \in \{1, \dots, s\}$ and all $\mathbf{u}_i \in (\mathbb{R}^{m_i})^N$ it holds
$J_i^N(x, \mathbf{u}_1^{e,N}, \dots, \mathbf{u}_i^{e,N}, \dots, \mathbf{u}_s^{e,N}) \le J_i^N(x, \mathbf{u}_1^{e,N}, \dots, \mathbf{u}_i, \dots, \mathbf{u}_s^{e,N})$.

A NE is a control strategy with the property that it is never beneficial for one player to unilaterally deviate from this strategy. This does *not* imply that there is no strategy which improves all players' objective function simultaneously. Even though many papers such as [1–4] focus on affine- or linear-quadratic dynamic games, it seems that the case of 'true' conflict is typically not dealt with in the literature. This means that all x_i^* in (2) are assumed to be identical in these references. In order to calculate NEs for our setting we perform the coordinate transformations $y_i := x - x_i^*$ for all $i \in \{1, \dots, s\}$. This way we obtain the augmented system

$$y^+ = \begin{pmatrix} y_1^+ \\ \vdots \\ y_s^+ \end{pmatrix} = \begin{pmatrix} Ay_1 + \sum_{i=1}^{s} B_i u_i + c + (A - Id)x_1^* \\ \vdots \\ Ay_s + \sum_{i=1}^{s} B_i u_i + c + (A - Id)x_s^* \end{pmatrix} =: \bar{A}y + \sum_{i=1}^{s} \bar{B}_i u_i + \bar{c}$$

and stage costs $\ell_i(x, u_i) = \frac{1}{2}\left[y_i^T Q_i y_i + u_i^T R_i u_i\right] =: \frac{1}{2}\left[y^T \bar{Q}_i y + u_i^T R_i u_i\right] =: \bar{\ell}_i(y, u_i)$. It can easily be seen that any NE to the transformed game is a NE to the original game. By means of the transformation we can proceed similarly to e.g. [1, Thm. 6.2] to calculate NEs for our game:

Theorem 1.2 (NEs for Affine-Quadratic Games with 'True' Conflict) *Consider the s-player game defined by* (1) *and* (2) *with horizon $N \in \mathbb{N}$. Consider the backward matrix iterations*

$$\Lambda^k = Id + \sum_{i=1}^{s} \bar{B}_i R_i^{-1} \bar{B}_i^T M_i^{k+1}, \qquad M_i^k = \bar{Q}_i + \bar{A}^T M_i^{k+1} (\Lambda^k)^{-1} \bar{A},\ M_i^N = 0 \tag{3}$$

* Corresponding author: E-mail marleen.stieler@uni-bayreuth.de, phone +49 (0)921 55 3285, fax +49 (0)921 55 5361

for $k = N-1, \dots, 0$. If all Λ^k are invertible, then for each initial value $x \in \mathbb{R}^n$ there exists a unique NE $\mathbf{u}^{e,N}$ to the transformed (and thus also to the original) game.

The proof is similar to the proof of [1, Thm. 6.2] and omitted here. We note that there exist explicit formulas for the NE as well as for the corresponding trajectory.

2 Main Result and Example

We propose the following noncooperative MPC algorithm. For more details on MPC we refer to [5].

Algorithm 2.1 (Noncooperative MPC) *At each time instant $n \in \mathbb{N}_0$ and for fixed horizon $N \in \mathbb{N}$:*

1. *Set $x := x(n)$ and calculate a Nash equilibrium $\mathbf{u}^{e,N}$ of length N for initial value x for the s-player game.*
2. *For each player $i \in \{1, \dots, s\}$ define the MPC-feedback $\mu_i^N(x) := u_i^{e,N}(0)$ and apply it to the system, i.e. $x(n+1) = f(x, \mu^N(x))$.*

Theorem 2.2 (Convergence of MPC trajectories) *Consider the affine-quadratic s-player game given by the dynamics* (1) *and stage costs* (2) *and a horizon $N \in \mathbb{N}$. Assume that all Λ^k, $k = 0, \dots, N-1$, in* (3) *are invertible and $\|\bar{A}(\Lambda^0)^{-1}\| \leq 1$. Assume moreover that all eigenvalues λ of $(\Lambda^0)^{-1}\bar{A}$ fulfill either $|\lambda| < 1$, or $\lambda = 1$ and λ is semisimple/nondefective*[1]. *Then for each $x_0 \in \mathbb{R}^n$ the MPC closed-loop trajectory of Algorithm 2.1 converges.*

Sketch of the Proof. The matrix iterations in (3) only depend on the data of the game but not on the current time or state. Thus, they are identical in each iteration of Algorithm 2.1. This is why an explicit formula for the MPC closed-loop trajectory can easily be given. The condition $\|\bar{A}(\Lambda^0)^{-1}\| \leq 1$ and the conditions on the eigenvalues of $(\Lambda^0)^{-1}\bar{A}$ are sufficient conditions for the closed-loop trajectory to converge. The limit can explicitely be calculated. □

Example 2.3 We consider a simple model of the room temperature $x \in \mathbb{R}$, controlled by two persons. The dynamics are given by (1) with $A = 0.8$, $B_1 = B_2 = 1$ and $c = 0$. For the stage costs (2) we use the values $Q_i = 1$, $R_1 = 2$, $R_2 = 1$, $x_1^* = 23$, $x_2^* = 17$, i.e. both persons have different desired temperatures. We execute Algorithm 2.1 with $N = 5$. For these parameters the assumptions of Theorem 2.2 are satisfied. Figure 1 illustrates the convergence of the MPC closed-loop trajectories. We note that we also observe convergence if the conditions in Theorem 2.2 are not satisfied, which is our motivation to investigate less restrictive conditions in future research.

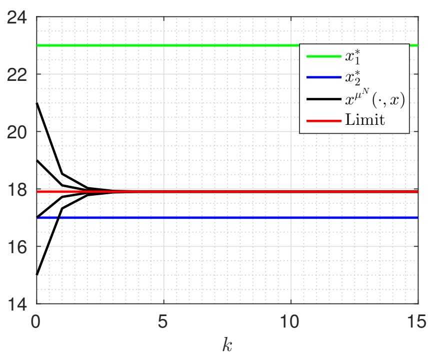

Fig. 1 Closed-loop trajectories of Algorithm 2.1 for $N = 5$ and different initial values (black) and the theoretically calculated limit (red).

All the statements in this paper and the corresponding proofs can be found in [6].

Acknowledgment

Lars Grüne and Marleen Stieler are supported by DFG Grant Gr 1569/13-1. Michael H. Baumann is supported by Hanns-Seidel-Stiftung e.V. (HSS), funded by Bundesministerium für Bildung und Forschung (BMBF).

References

[1] T. Başar and G. J. Olsder, Dynamic Noncooperative Game Theory, 2nd edition (SIAM, 1999).
[2] J. C. Engwerda, Journal of Economic Dynamics and Control **22**, 729–762 (1998).
[3] G. Jank and H. Abou-Kandil, IEEE Trans. Autom. Control **48**(2), 267–271 (2003).
[4] D. Kremer, Non-symmetric Riccati theory and noncooperative games (Wissenschaftsverlag Mainz in Aachen, 2003).
[5] L. Grüne and J. Pannek, Nonlinear Model Predictive Control: Theory and Algorithms, 2nd edition (Springer, 2017).
[6] M. Stieler, Performance Estimates for Scalar and Multiobjective Model Predictive Control Schemes, Submitted as PhD Dissertation, 2018, Currently under review.

[1] This means that the eigenvalue is a root of multiplicity one in the minimal polynomial.