Early Termination of Convex QP Solvers in Mixed-Integer Model Predictive Control for Real-Time Decision Making

Integer Programming and Combinatorial Optimization 2021

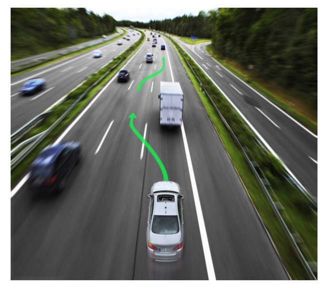
Jiaming Liang (Georgia Tech), Stefano Di Cairano, Rien Quirynen

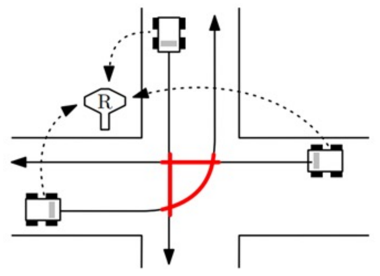
Mixed-Integer MPC for Real-Time Decision Making

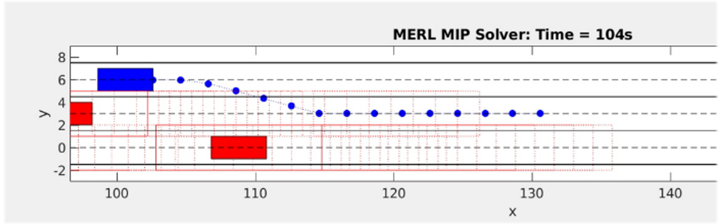
MI-MPC provides a general-purpose modeling framework for real-time decision making.

We are particularly interested in an MI-MPC formulation of a high-level motion planning task for an autonomous vehicle, including discrete decisions resulting from lane changes, static and dynamic obstacles

The MI-MPC framework solves an MIQP problem at every sampling time instant.

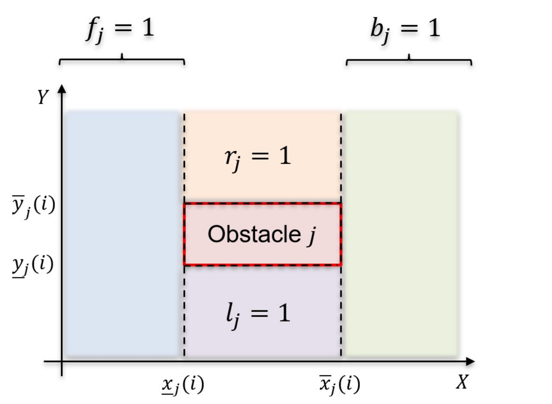






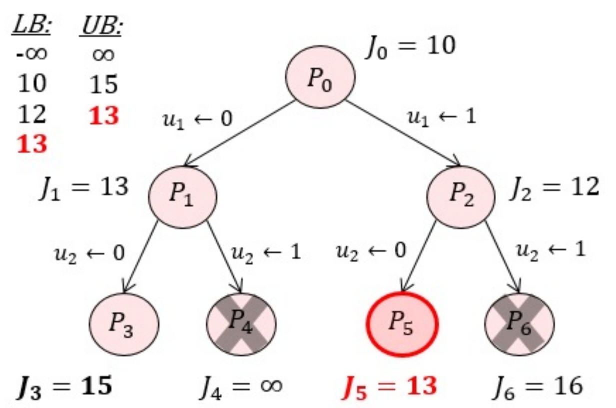
- Switches in system dynamics, e.g., contacts
- Discrete decisions, e.g., pass or stay in lane
- Quantized decisions, e.g., on/off actuation
- Disjoint constraint sets, e.g., obstacle avoidance

For example, using big-M formulation and 4 binary variables



Branch-and-Bound Algorithm for MIQP

- Convex QP relaxations solved to obtain lower bounds (LB)
- Each integer-feasible solution forms an upper bound (UB) for the MIQP solution
- A node can be pruned due to LB > UB (P_6) or infeasibility (P_4)



Early termination of QP solvers in B&B: aim to prune node without need to solve convex QP

- If dual feasible objective $\psi(\cdot) > UB$, then primal optimal objective $\phi^\star > UB$:

$$\psi(\mu,\lambda) \le \psi^\star \le \phi^\star \le \phi(\boldsymbol{x},\boldsymbol{y})$$

- Terminate the QP solver before convergence.
- Also effective in detecting primal infeasibility.

QP Formulations and Infeasible IPM Solver

Primal QP Formulation

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \phi(\boldsymbol{x},\boldsymbol{y}) := \frac{1}{2}\boldsymbol{x}^\top Q \boldsymbol{x} + h_{\boldsymbol{x}}^\top \boldsymbol{x} + h_{\boldsymbol{y}}^\top \boldsymbol{y} \\ \text{s.t.} \quad & G_{\boldsymbol{x}}\boldsymbol{x} + G_{\boldsymbol{y}}\boldsymbol{y} \le g, \\ & F_{\boldsymbol{x}}\boldsymbol{x} + F_{\boldsymbol{y}}\boldsymbol{y} = f, \end{aligned}$$

Dual QP Formulation

$$\hat{h}(\mu,\lambda) := h_{\boldsymbol{x}} + G_{\boldsymbol{x}}^\top \mu + F_{\boldsymbol{x}}^\top \lambda,$$

$$\begin{aligned} \max_{\mu,\lambda} \quad & \psi(\mu,\lambda) := -\frac{1}{2}\|\hat{h}(\mu,\lambda)\|^2_{Q^{-1}} - \begin{bmatrix} g \\ f \end{bmatrix}^\top \begin{bmatrix} \mu \\ \lambda \end{bmatrix} \\ \text{s.t.} \quad & G_{\boldsymbol{y}}^\top \mu + F_{\boldsymbol{y}}^\top \lambda = -h_{\boldsymbol{y}}, \\ & \mu \ge 0, \end{aligned}$$

- *Infeasible IPM:* Newton-type iteration

$$\begin{bmatrix} H & F^\top & G^\top \\ F & 0 & 0 \\ G & 0 & -W^k \end{bmatrix} \begin{bmatrix} \Delta z^k \\ \Delta\lambda^k \\ \Delta\mu^k \end{bmatrix} = -\begin{bmatrix} r_z^k \\ r_F^k \\ \tilde{r}_G^k \end{bmatrix}$$

- *Problem:* infeasible IPM iterations generally do not satisfy dual feasibility until convergence
- *Proposed solution:* computationally efficient projection to obtain dual feasible solution guess for early termination of infeasible IPM

Infeasible IPM: Projection to Dual Feasibility

Equality-constrained optimization for minimum-norm projection on constraint

$$\min_{\Delta\lambda,\Delta\mu} \quad \frac{1}{2}\|\Delta\lambda\|^2 + \frac{1}{2}\|\Delta\mu\|^2, \quad \text{s.t.} \quad F_{\boldsymbol{y}}^\top \Delta\lambda + G_{\boldsymbol{y}}^\top \Delta\mu = -r_{\boldsymbol{y}}^k$$

But projection does not guarantee nonnegativity of Lagrange multipliers, i.e., $\mu \geq 0$

Proposed approach: modified optimization problem for projection on constraint

Solve equality-constrained optimization problem to compute projection step

$$\begin{aligned} \min_{\Delta\boldsymbol{x},\Delta\lambda,\Delta\mu} \quad & \frac{1}{2}\|\Delta\boldsymbol{x}\|_Q^2 + \frac{1}{2}\|\Delta\lambda\|^2_{\epsilon_{\text{dual}}} + \boxed{\frac{1}{2}\|\Delta\mu\|^2_{W^k}} \\ \text{s.t.} \quad & \begin{bmatrix} Q \\ 0 \end{bmatrix} \Delta\boldsymbol{x} + F^\top \Delta\lambda + G^\top \Delta\mu = -\begin{bmatrix} 0 \\ r_{\boldsymbol{y}}^k \end{bmatrix}, \end{aligned}$$

$$W^k = \text{diag}\left(w^k\right) \text{ and } w_i^k = \frac{s_i^k}{\mu_i^k} > 0$$

$$\Downarrow$$

$$\|\Delta\mu\|^2_{W^k} = \sum_i \left(\frac{s_i^k}{\mu_i^k}\Delta\mu_i^2\right)$$

Indirectly enforces positivity constraints $\mu \geq 0$

Active inequality constraints:

$s_i \to 0$ and $\mu_i > 0 \Rightarrow w_i = \frac{s_i}{\mu_i} \to 0$ (small penalty)

Inactive inequality constraints:

$\mu_i \to 0$ and $s_i > 0 \Rightarrow w_i = \frac{s_i}{\mu_i} \to \infty$ (large penalty)

$\equiv$

is equivalent to solving KKT system to compute Newton-type search direction

$$\begin{bmatrix} H & F^\top & G^\top \\ F & -\epsilon_{\text{dual}} I & 0 \\ G & 0 & -W^k \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{z} \\ \Delta\lambda \\ \Delta\mu \end{bmatrix} = -\begin{bmatrix} \begin{bmatrix} 0 \\ r_{\boldsymbol{y}}^k \end{bmatrix} \\ 0 \\ 0 \end{bmatrix},$$

Projection corresponds to IPM iteration with different right-hand side: obtain dual feasibility while maintaining optimality conditions

$\equiv$

is equivalent to solving reduced KKT system to compute search direction Δz

$$\boxed{H + \frac{1}{\epsilon_{\text{dual}}} F^\top F + G^\top W^{k^{-1}} G} \Delta z = -\begin{bmatrix} 0 \\ r_y^k \end{bmatrix},$$

which is used to compute projection step for Lagrange multipliers $(\Delta\mu^k, \Delta\lambda^k)$:

$$\Delta\lambda = \frac{1}{\epsilon_{\text{dual}}} F\Delta z, \quad \Delta\mu = W^{k^{-1}} G\Delta z$$

We can reuse the KKT matrix factorization between IPM iterations and projection steps for computational efficiency

Early Termination of IPM: Infeasibility detection

- Certificate of primal infeasibility (i.e., unboundedness of dual) the following set of equations is strictly infeasible

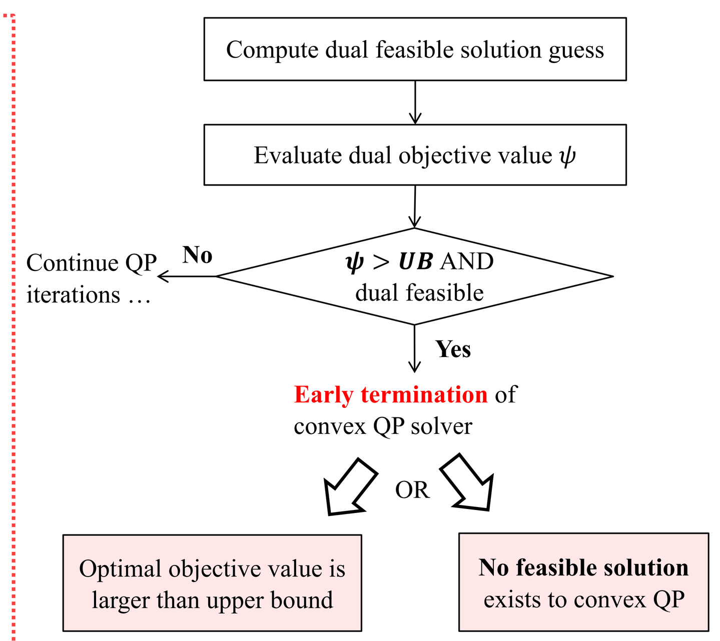
$$Gz < g, \quad Fz = f,$$

if and only if there exists a pair $(\tilde{\mu}, \tilde{\lambda})$ such that (Farkas' lemma)

$$G^\top \tilde{\mu} + F^\top \tilde{\lambda} = 0, \quad g^\top \tilde{\mu} + f^\top \tilde{\lambda} < 0, \quad \tilde{\mu} > 0$$

- Instead, our proposed early termination technique can be used for infeasibility detection and requires limited computational cost (projection based on reuse of KKT matrix factorization).
- Intuition behind using early termination for infeasibility detection:

Proposition 4.3: If the sequence of IPM iterates $\{(z^k,\mu^k,\lambda^k,s^k)\}$ satisfy $\mu^{k^\top} s^k \le \mu^{0^\top} s^0$ and $\|\mu^k\| \to \infty$, then the dual objective $\psi(\mu^k,\lambda^k) \to \infty$.



Algorithm 1 Early termination for IPM in B&B method.

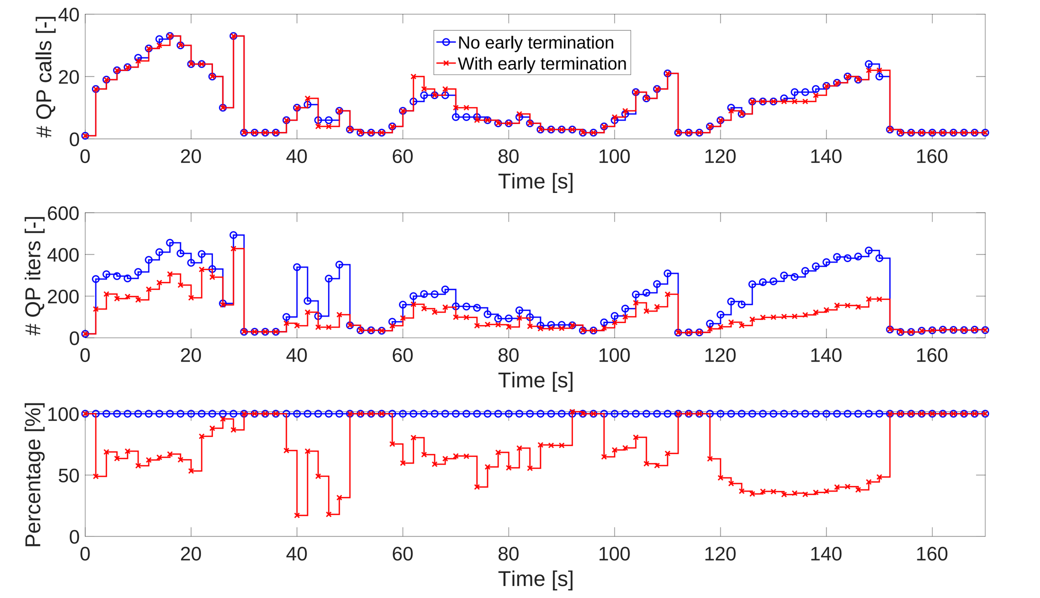
Input: Warm start $\{(z^0,\mu^0,\lambda^0,s^0)\}$, tol, and UB.
while $\max\{\tau^k, \|r^k\|\} > tol$ **do**
    if $\psi(\mu^k,\lambda^k) >$ UB & dual_feasible **then**
        break while loop. ▷ Early termination
    else if $\psi(\mu^k,\lambda^k) >$ UB **then**
        Compute projection step $(\Delta\mu, \Delta\lambda)$ in (13).
        $\mu \leftarrow \mu^k + \Delta\mu$, $\lambda \leftarrow \lambda^k + \Delta\lambda$, and
        $r_{\boldsymbol{y}} \leftarrow F_{\boldsymbol{y}}^\top \lambda + G_{\boldsymbol{y}}^\top \mu + h_{\boldsymbol{y}}$.
        if $\mu > 0$ & $\|r_{\boldsymbol{y}}\| < tol$ **then**
            $\mu^k \leftarrow \mu$, $\lambda^k \leftarrow \lambda$, $r_{\boldsymbol{y}}^k \leftarrow r_{\boldsymbol{y}}$, and
            dual_feasible ← 1.
            if $\psi(\mu^k,\lambda^k) >$ UB **then**
                break while loop. ▷ Early termination
            end if
        end if
    end if
    Perform an IPM iteration (8), e.g., see [18].
end while

Simulation Results: Real-Time Vehicle Decision Making

QP Calls and Total Iterations

blue: No early termination
red: Early termination

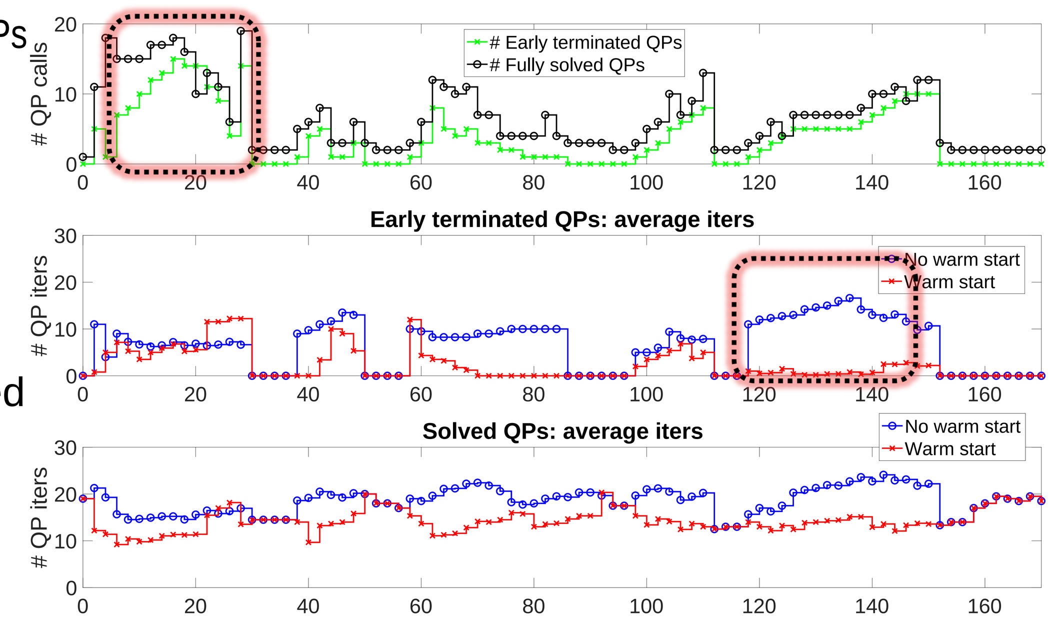
- ➤ Early termination reduces total number of QP iterations by 42%.



Early Terminated versus Fully Solved QPs

green: Early terminated QPs
black: Fully solved QPs
blue: No warm start
red: Warm start

- ➤ Early termination happens often, 36% of QPs are early terminated
- ➤ Early termination benefits from warm starting IPM and takes less QP iterations.



Infeasibility Detection

Number of IPM Iterations for Certificate Versus Early Termination With and Without Warm Starting

	QP # 1	QP # 2	QP # 3
Certificate of primal infeasibility	40	45	38
Early termination: cold started	10	12	10
Early termination: warm started	0	0	11

- ➤ Early termination requires considerably less IPM iterations than the computation of a certificate of infeasibility.
- ➤ Warm starting can reduce the number of IPM iterations further and it can lead to immediate termination, i.e., termination at 0 iterations.

Conclusions

An efficient early termination strategy based on a projection step tailored to IPMs, in order to reduce the computational cost within B&B method in solving MI-MPC.

Early termination of QP solvers in MI-MPC works well in

- terminating QPs whose objective value > UB;
- detecting infeasible QPs.

Early termination is performed by using Newton-type IPM iterations

- reuses KKT matrix factorizations for computational efficiency;
- intuitively guarantees the inequality constraint, i.e., positivity in μ;
- projection also makes progress towards convergence.