Cascaded Reference Governor-MPC for Motion Control of Two-Stage Manufacturing Machines

Di Cairano, S.; Goldsmith, A.; Kalabić, U.; Bortoff, S.A.

TR2018-083 July 12, 2018

Abstract

Two-stage processing machines for manufacturing are equipped with two sets of actuators for motion control with different operating ranges and bandwidths. The processing time, product quality and flexibility of the manufacturing process can be optimized by coordinating these actuators and exploiting the entire actuator operating range. We propose a motion control strategy for two-stage processing machines based on model predictive control (MPC). By exploiting timescale separation, we formulate the problem as a single-stage motion control with reference-dependent constraints. Feasibility of the MPC problem is guaranteed by using a reference governor that adjusts the feed rate. The proposed method guarantees correct processing while satisfying the actuators’ range and dynamics constraints, finite time processing of a given spatial pattern and real-time execution even with limited computational resources. Simulation and experimental results on a real processing pattern are shown for a scaled laboratory demonstration machine.

IEEE Transactions on Control Systems Technology
Cascaded Reference Governor–MPC for Motion Control of Two-Stage Manufacturing Machines

S. Di Cairano Senior Member, IEEE, A. Goldsmith Member, IEEE, U.V. Kalabić Member, IEEE, S.A. Bortoff Senior Member, IEEE,

Abstract—Two-stage processing machines for manufacturing are equipped with two sets of actuators for motion control with different operating ranges and bandwidths. The processing time, product quality and flexibility of the manufacturing process can be optimized by coordinating these actuators and exploiting the entire actuator operating range. We propose a motion control strategy for two-stage processing machines based on model predictive control (MPC). By exploiting timescale separation, we formulate the problem as a single-stage motion control with reference-dependent constraints. Feasibility of the MPC problem is guaranteed by using a reference governor that adjusts the feed rate. The proposed method guarantees correct processing while satisfying the actuators’ range and dynamics constraints, finite time processing of a given spatial pattern and real-time execution even with limited computational resources. Simulation and experimental results on a real processing pattern are shown for a scaled laboratory demonstration machine.

Index Terms—Precision Manufacturing, Motion Control, Model predictive control, Constrained Control

I. INTRODUCTION

Precision manufacturing is a relevant application area of control systems [1], [2], due to the need for high quality processing of the machined parts and for high throughput to reduce manufacturing costs. High throughput requires high speed processing, while high quality requires high precision, and a trade-off must be achieved between these conflicting requirements by hardware and software design.

Manufacturing machines are composed of a worktool capable of performing operations such as milling, cutting, welding or drilling on a piece of raw material, and of a motion control system that moves the worktool. The worktool may have a considerably large mass, also due to the size of the support structure, and it may need to move over large ranges due to processing large parts or multiple parts in a single block of raw material. Moving a large mass over a long range with high accelerations and high precision results in conflicting objectives. The motion control system aims at achieving a desirable trade-off between such objectives.

Controlling a processing machine usually involves actuating the worktool to follow a geometric path that represents the machining pattern of the part being produced. The common performance requirements are the time to process the path, the precision of the obtained machined path, which is also affected by the vibrations induced by the machine moving parts, and the energy consumed during processing. Due to the high rates of operation and relatively low computational power of the microprocessors used in factory automation, the control methods used in manufacturing have been mostly limited to PID and loop shaping filters in the servomotors. In fact, for common designs, the axes are dynamically decoupled and each axis of the single-stage processing machines can be considered single-input single-output system. Thus, the worktool path is computed off-line, based on the geometry of the parts to be machined [3], [4], and the worktool velocity, i.e., the path feed rate, is scheduled based on the desired precision. The obtained trajectories are then used in “open-loop”, i.e., without modifications during execution, relying on the feedback loops of the servomotors to reject disturbances and to compensate for uncertainty. For repeated operation, supervisory strategies such as iterative learning control have been suggested in [2]. A real-time trajectory generation may be advantageous, for instance by allowing modifications to the processed part between different runs, or even within a run, without the need to stop the processing.

In order to increase the speed of processing and the precision over a large workspace, the motion control system of high performance manufacturing machines may exploit multiple actuation stages per processing axis, where actuators with different bandwidths and operating ranges are combined to process large workpieces at high rate. Two-stage actuation has been applied in different fields ([5]–[7]). In particular, in a two-stage machine each axis is equipped with a “slow” stage with large operating range but small bandwidth and acceleration limits, and a “fast” stage with large bandwidth and acceleration limits but small operating range (see Figure 1). For each axis, the overall position of the worktool is the sum of the positions of the two stages. With this architecture the machine can rapidly process small features of the workpiece by actuating the fast stage, and still be able to process large features by superimposing the motion of the slow stage.

Trajectory generation and control for two-stage machines is more complicated than for single stage ones. Each axis of a two-stage machine is effectively modeled as a multiple input–single output system [8] subject to constraints on velocity, acceleration, and operating range. Classical methods based on frequency separation [5], [9], [10] may be significantly suboptimal and may require convoluted ad-hoc logic and iterations to handle the constraints, thus limiting the applicability for online trajectory generation. Instead, model predictive control (MPC) [11] has proven effective for controlling multivariable systems subject to constraints, and its application domain has been extended in the last few years from process and...
chemical control to systems with faster dynamics and reduced computational resources, such as in automotive, aerospace, and mechatronics [12], [13]. Recently some MPC applications in precision manufacturing have been developed. In particular, [14]–[16] (see also references therein) propose path following MPC algorithms, and their applications to contouring control, that operate in the spatial domain, and hence require (explicitly or implicitly) the linearization of the spatial nonlinear dynamics. These approaches have been developed for single-stage machines where the algorithms operate on a linearized (i.e., approximate) model along a path known in advance, and the machining error is a performance measure, and hence not subjects to hard bounds.

However, application of path following MPC to two stage processing machines is not straightforward. The reasons include the timescale separation of the stages (up to 100x), the high control rates of the fast stage, (up to 200kHz), the accuracy being not a performance metric, but rather a constraint to be enforced on all points of the path, on the order of hundreds of thousands, and the lack of an a-priori defined path for linearization, as the system trajectory is not uniquely defined due to the overactuation. In addition, for real-order of hundreds of thousands, and the lack of an a-priori constraint satisfaction, in a way that maintains the spatial domain, and hence require (explicitly or implicitly) the linearization of the spatial nonlinear dynamics. These approaches have been developed for single-stage machines where the algorithms operate on a linearized (i.e., approximate) model along a path known in advance, and the machining error is a performance measure, and hence not subjects to hard bounds.

In this paper we propose a strategy for controlling two-stage processing machines equipped with a small-range fast actuator (fast stage) and a large-range slow actuator (slow stage) per axis. An example of such machine is described in Section II. Rather than controlling both actuators in the same loop, in Section II we exploit the timescale separation between the stages to re-formulate the machine control problem into controlling the slow stage by a tracking MPC with constraints that depend on the reference trajectory, which is generated for an “ideal” single stage, that has the range of the slow stage and the bandwidth of the fast stage.

Due to the dependency of the constraints on the reference, the reference trajectory may need to be modified for ensuring feasibility of MPC, which is a recently studied problem, see, e.g., [18], [19]. Unfortunately, such methods cannot be directly applied to the problem considered here because the modification to the setpoint will in general cause a modification to the spatial path, which results in an incorrectly machined part. As described in Section III, we modify online the upcoming segment of the ideal trajectory to be feasible and to ensure future constraint satisfaction, in a way that maintains the spatial shape of the pattern to be processed by using a particular reference governor [20]. The modified trajectory is provided to an appropriately designed MPC, which is recursively feasible and achieves finite-time processing of the machining path.

As discussed in Section IV, a benefit of the proposed approach is that it only requires the evaluation of linear inequalities for the reference governor, and the solution of a convex quadratic program for MPC, which can be achieved by low complexity solvers. Furthermore, the control architecture allows for real-time operation where a recursively feasible solution is found within a pre-assigned computing time.

Fig. 1. Schematic of the front view of a two-stage machine.

In Section V we evaluate the algorithm on a laboratory experiment built with standard off-the-shelf components with scaled slow stage with respect to existing processing machines. We draw our conclusions in Section VI hinting at the effectiveness of MPC in manufacturing applications.

Notation: $\mathbb{R}$, $\mathbb{R}_{+}$, $\mathbb{R}_{+}^{*}$ and $\mathbb{Z}$, $\mathbb{Z}_{+}$, $\mathbb{Z}_{+}^{*}$ are the sets of real, nonnegative real, positive real, and integer, nonnegative integer, positive integer numbers, and we use notations like $\mathbb{Z}_{[a, b]} = \{z \in \mathbb{Z} : a \leq z < b\}$ to denote intervals. By $[a]_{i}$ we denote the $i$-th component of $a$, for $a \in \mathbb{R}^{n}$, $b \in \mathbb{R}^{m}$, $(a, b) = [a' \ b']' \in \mathbb{R}^{n+m}$ is the stacked vector, and $I$, $1$, and $0$ are the identity, the all-1, and the zero matrices of appropriate size, respectively. Relational operators between vectors are intended componentwise, while for matrices denote (semi)definiteness. Given a set $\mathcal{A}$ and $(a, b) \in \mathcal{A}$ we denote by $\mathcal{A}(b)$ the section of $\mathcal{A}$ in the coordinates of $b$ at the values of $b$. We denote the Minkowski set sum by $\oplus$, $B(\rho)$ denotes the norm-ball of radius $\rho \in \mathbb{R}_{+}$ centered at the origin. For a discrete-time signal $x \in \mathbb{R}^{n}$ with sampling period $T$, $x_{t}$ is the value at sampling instant $t$, i.e., at time $tT$, and $x_{k|t}$ denotes the predicted value of $x$ at sample $t + k$, i.e., $x_{t+k}$, based on data at sample $t$, where $x_{0|t} = x_{t}$. We denote the convolution operator by $\ast$, and, with a little abuse of notation, $y(t) = G(t) \ast u(t) = \int_{0}^{t} G(t - \tau)u(\tau)d\tau$.

II. ARCHITECTURE, MODEL AND CONTROL PROBLEM

We consider a two-stage dual-axis (i.e., 2D) processing machine as in the schematics in Figure 1. The objective of the machine is to process blocks of raw materials at a high rate and with high precision to manufacture parts that have small and large features. The two-stage machine aims at resolving some of the conflicting requirements in precision manufacturing. Due to small features, e.g., less than a millimeter, and the requirement to process at high rate, the worktool may need to sustain large accelerations, on the order of several $g$. However, due to large features, e.g., more than a meter, and possibly multiple parts being produced from a single block of raw material, the worktool must have a large operating range and as a consequence a large mass, up to several hundred kilograms. Large accelerations and large operating range (i.e., large mass) are obviously in conflict. Thus, two-stage machines combine slow stages and fast stages, see Figure 1, that are implemented by different actuators, such as motors, piezoelectric actuators, electromagnetic actuators, in closed-loop with their servocontrollers. The fast stages have smaller operating range, and as a consequence are smaller and less massive, and thus

---

1Preliminary studies related to this research appeared in [17].
can achieve larger accelerations. On the other hand the slow stages have larger operating ranges, and as a consequence, are more massive, and can achieve smaller accelerations. The rationale for such an architecture is that the small features of the machined part can be processed by high-acceleration movements of the fast stage, while large features can still be processed by superimposing the large range movements of the slow stage to the fast stage movements.

The objective of trajectory generation and control for a two-stage machine is to coordinate the motion of the slow and fast stages for each processing axis, to achieve the desired processing pattern as fast as possible within the given precision specifications. Additional secondary objectives are to reduce the acceleration for the massive slow stage, thus reducing power consumption and the vibrations induced by the forces to accelerate the large mass.

### A. Motion models of two-stage axes

A schematic of the control architecture of the machine considered here is shown in Figure 2, where the control algorithm generates commands for the slow stage and the fast stage of two processing axes (x, y). The commands are then provided to the corresponding stage actuators, which follow them using conventional feedback loops. The combined effects of the slow stage result in the worktool achieving the desired processing pattern.

The motion model of the stage in closed-loop with the servocontroller is

\[ y^i_j(t) = G^i_j(t) * u^i_j(t), \quad j \in \{s, f\}, \quad i \in \{x, y\}, \]  

where \( y \) is the position, \( u \) is the position command, \( G \) is the transfer function with unitary dc-gain, \( j \in \{s, f\} \) is the index of the stage (slow/fast) and \( i \in \{x, y\} \) is the index of the processing axis (x/y).

For the considered architecture, the position of the worktool is the sum of the positions of the two stages in each axis,

\[ y^i(t) = y^i_s(t) + y^i_f(t) = G^i_s(t) * u^i_s(t) + G^i_f(t) * u^i_f(t), \quad i \in \{x, y\}. \]

The stages are subject to constraints due to physical limitations and operating requirements. The stage operating ranges are constrained due to the actuator mechanical limitations,

\[ \overline{y}_j^i \leq y^i_j \leq \overline{y}_j^i, \]  

where \( \overline{y}_j^i, j \in \{s, f\}, i \in \{x, y\} \) determine (symmetric) upper and lower bounds. Velocities and accelerations are also constrained

\[ \overline{y}_j^i \leq \dot{y}^i_j \leq \overline{y}_j^i \]  

\[ \overline{y}_j^i \leq \ddot{y}^i_j \leq \overline{y}_j^i, \]  

where \( \overline{y}_j^i, \dot{y}^i_j, \ddot{y}^i_j, j \in \{s, f\}, i \in \{x, y\} \).

The differences between the slow and fast stages reside in the bandwidth of the transfer functions (1), and in the constraints (3), (5). The fast stages have much larger bandwidths, due to the reduced actuation range,

\[ BW_3(G^i_f) \gg BW_3(G^i_s), \quad i \in \{x, y\}, \]

where \( BW_3(\cdot) \) denotes the 3-db bandwidth. The slow and fast stage frequency responses for the machine considered here are showed in Figure 3, where \( BW_3(G^i_f) = 50BW_3(G^i_s) \), \( i \in \{x, y\} \). The fast stage range is much smaller than that of the slow stage, but it allows for much larger accelerations,

\[ \overline{y}_j^i \ll \overline{y}_j^s, \quad \overline{y}_j^i \gg \overline{y}_j^s, i \in \{x, y\}. \]

Due to the different bandwidths, the sampling periods for the discrete-time servocontrollers of the fast and slow stage, and hence for the models in (1), are different, i.e., \( T_f \ll T_s \). Specifically here, \( T_s = M \cdot T_f \), where \( M \in \mathbb{Z}_+ \) and \( M \gg 1 \).

### B. Tracking control problem definition

The objective for the control system in Figure 2 is to make the worktool follow a geometric path representing the pattern to be processed within a desired spatial accuracy \( \varepsilon \in \mathbb{R}_+ \), e.g., on the order of microns, while satisfying all the constraints, and can be formulated as follows.

**Problem 1:** Given (2) subject to (3), (4), (5), and the regular geometric curve \( p(\sigma) = [p^x(\sigma) \ p^y(\sigma)]^T \), where \( \sigma \in \mathbb{R}_{[0,1]} \) is the (normalized) path parameter [15] and \( \varepsilon \in \mathbb{R}_+ \) is the
accuracy tolerance, control the commands \( u^j_i, i \in \{x, y\}, j \in \{s, f\} \) such that
\[
\| (p^x(\sigma), y^x(\sigma)) - (p^y(\sigma), y^y(\sigma)) \|_\infty \leq \varepsilon, \forall \sigma \in \mathbb{R}_{[0,1]}.
\] (6)

A holistic approach for solving Problem 1 is to formulate a path following MPC, (see, e.g., [14], [16], [21] and references therein). However, in the spatial domain the dynamics and the constraints are nonlinear, and hence the controller has to solve a non-convex optimization problem at every iteration. Thus, the command calculation in the path following approaches in [2], [16], [21], [22] is based on the repeated linearization or approximation of the nonlinear problem along the path, which is reasonable when the tracking error is a performance variable, i.e., there is no bound that the error must satisfy. On the other hand, due to the linearization and the usage in the cost function, a bound on the tracking error is difficult to guaranteed. In fact, [14], [16], [21] do not impose tracking error constraints which is however the case here due to (6). Besides this, there are others complicating factors in two-stage machines. The timescale separation of the stages imposes a prediction horizon related to slow stage bandwidth, and a sampling period based on the fast stage bandwidth, thus resulting in a large optimization problem with poor numerical conditioning. Furthermore, since the machine has multiple stages, the stage trajectory is not uniquely determined by the processing pattern, and hence it is not possible to determine a “good” nominal trajectory for linearization. Thus, a significant linearization error may occur, causing the violation of (6). For instance the path following MPC in [14] assumes the prediction model to be square and accordingly reduces the degrees of freedom in the application, while multi-stage machines have more inputs than outputs.

Instead, we formulate a modified problem whose solution also solves Problem 1 as described next. First, we consider the “ideal” single stage machine that has accuracy \( \varepsilon \), the range of the slow stage and the dynamics, velocity, and acceleration limits of the fast stage. Being this a single stage machine, we determine a path feed rate by standard methods resulting in the trajectory as the sequence of points
\[
\{q(hT_f)\}_h = \{(q^x(hT_f), q^y(hT_f))\}_h, \ h \in \mathbb{Z}_{[0,\bar{h}]},
\] (7)
where \( h \) is the point index and \( \bar{h} \in \mathbb{R}_+ \) is the total number of points. The trajectory in (7) is such that \( y^f(t) = G^f(t) \ast q^f(t) \), \( i \in \{x, y\} \) is feasible for the ideal machine that has all the strengths of the two stages and none of the weaknesses, i.e., \( \{q(hT_f)\}_h \) satisfies (5) for \( j = f \), (3) for \( j = s \), and (6).

In addition to those already enforced by (7), the machine needs only to satisfy the actual constraint (3) for \( j = f \) by a motion of the slow stage that satisfies (3), (4), (5), for \( j = s \). To this end we control (1), (3), (4), (5), for \( j = s \) such that
\[
-\bar{\gamma}_f \leq y^s_i(t) - q^f_i(t) \leq \bar{\gamma}_f, \ i \in \{x, y\}.
\] (8)

In an MPC context, this amounts to solving in receding horizon with sampling period \( T_s \)
\[
\begin{align*}
\min_{U^s_k} & \quad F(x^s_N|t, q^s_N|t) + \sum_{k=0}^{N-1} L(x^s_{k+1|t}, q^s_{k+1|t}, u^s_{k+1|t}) \quad (9a) \\
\text{s.t.} & \quad x^s_{k+1|t} = f^s_k(x^s_k|t, u^s_k|t) \quad (3), (4), (5), \ j = s \quad (9b) \\
& -\gamma_f \leq q^s_i|t - q^f_i|t \leq \gamma_f \quad (9c) \\
& H(x^s_k|t, q^s_k|t, u^s_k|t) \leq 0 \quad (9d)
\end{align*}
\]
for \( i \in \{x, y\} \), where \( x^s_i \) is the state that includes at least state position, velocity and acceleration, \( f_s \) is the state update equation representing (1) for \( j = s \), \( N \in \mathbb{Z}_{0,s} \) is the prediction horizon, \( U^s_{st} = [u^s_0|t \ldots u^s_{N-1}|t] \), \( F, L \) are the terminal and stage cost, respectively, \( H \) models additional constraints, and perfect preview of the reference \( q \) for \( N \) steps is available.

A complicating feature of (9) is that constraint (9e) depends on the reference trajectory. Thus, feasibility of (9) depends on an exogenous variable and cannot be guaranteed for an unmodifiable reference. Even if (9) is feasible at time \( t \), there is no guarantee of recursive feasibility, i.e., that (9) is feasible for \( \tau > t \). However, since the timing along the path is not fixed, there is some freedom in modifying the reference as long as such modifications do not change the spatial path. For instance, changing the feed rate, the rate at which the points of the path are processed, does not change the spatial path. Thus, we formulate the following alternative to Problem 1.

**Problem 2:** Given \( \{q(hT_f)\}_h, \ h \in \mathbb{Z}_{[0,\bar{h}]} \) that satisfies (4), (5) for \( j = f \), (3) for \( j = s \), and such that \( y^f(t) = G^f(t) \ast q^f(t) \) satisfies (6), design \( F, L, H \) in (9) and construct \( \{r(T_s)|t = \{(r^x(T_s), r^y(T_s))|t\} \) such that \( y^f(t) = G^f(t) \ast r^f(t) \) satisfies (6), and (9) is strictly convex and recursively feasible when \( q \) is substituted by \( r \). Also, when \( \{q(hT_f)\}_h \) is finite, the processing time is finite.

Problem 2 involves the simultaneous modification of the reference and generation of the command to track such a modified reference, and has attracted interest in recent years, see, e.g., [18], [19]. Unfortunately, the previously proposed methods modify the reference within the MPC problem in ways that do not guarantee that the geometry of the path is maintained, and hence (6) may be violated. Instead, we solve Problem 1 by cascading a tracking MPC to a “spatial” reference governor (SPRG), which does not change the spatial shape of the reference and hence does not affect the satisfaction of (6). In practice, the effect of the SPRG is to change the feed rate to ensure that the constraints can be satisfied. It is worth noting that the path tracking error of (1) is inversely proportional to the feed rate. Hence, if \( \{q(hT_f)\}_h \) satisfies (6) and the feed rate is only reduced to generate \( \{r(T_s)|t \), then also the latter satisfies (6).

### III. SPRG-MPC FOR TWO-STAGE PROCESSING MACHINES

We propose a control system design for solving Problem 2, and consequently Problem 1, with the block diagram shown in Figure 4. Using conventional methods, we generate offline a reference trajectory for an ideal single stage machine with dynamics, velocity and acceleration constraints of the fast stage, and the range of the slow stage, thus obtaining an
idealized, usually infeasible, trajectory. The spatial SPRG developed in Section III-A modifies the reference to obtain a feasible reference that does not change the spatial processing pattern. The feasible reference is provided to a MPC algorithm developed in Section III-B that optimizes the trajectory with guaranteed recursive feasibility. The MPC feeds the command to the slow stage actuator while the fast stage actuator commands are obtained as the difference between the slow stage trajectory and the feasible reference, with the implementation discussed in Section IV-C. The predicted terminal state is provided to the SPRG for continuing operation.

A. Spatial Reference Governor

In order to obtain a reference trajectory that guarantees the feasibility of the MPC problem (9) and ensures satisfaction of (6) we develop a reference governor that operates on the spatial points \( \{q(h)\}_h \), represented by the SPRG block in Figure 4. First, we recall few useful notions.

Definition 1: Given \( x(t+1) = f(x(t)), x \in \mathbb{R}^n \), with \( z = h(x(t)), z \in \mathbb{R}^n \), such that \( z \in \mathcal{Z} \subset \mathbb{R}^n \), a constraint admissible set \( \mathcal{S}_\infty \subset \mathbb{R}^n \) is a set such that

\[
x(t) \in \mathcal{S}_\infty \Rightarrow h(x(t)) \in \mathcal{Z}, \forall \tau \geq t.
\]  

(10)

Given any constraint admissible set \( \mathcal{S}_\infty \), \( x \in \mathcal{S}_\infty \) implies that \( f(x) \in \mathcal{S}_\infty \), i.e., \( \mathcal{S}_\infty \) is positive invariant (PI) for \( x(t+1) = f(x(t)) \). The maximal constraint admissible set \( \mathcal{O}_\infty \) is the union of all constraint admissible sets.

Let the discrete-time state-space formulation of the slow stage motion, i.e., (1) for \( j = s \), sampled with period \( T_s \) be

\[
x^i(t+1) = A^i x^i(t) + B^i u^i_s(t), \quad i \in \{x,y\}
\]  

(11)

where \( y^i_s, \hat{y}^i_s, \bar{y}^i_s \) are components of the state vector \( x^i \), and define \( z^i = C^i x^i + C^i r^i = [y^i_s, \hat{y}^i_s, \bar{y}^i_s, -r^i]^T \). Thus, (11) with \( u^i_s(t) = r^i(t) \) and constant reference dynamics, and subject to (3), (4), (5), for \( j = s, \) and (8), results in

\[
\begin{align*}
x^i(t+1) &= A^i x^i(t) + B^i r^i(t) \\
r^i(t+1) &= r^i(t) \\
z^i(t) &= C^i x^i(t) + C^i r^i(t) \\
H^i z^i(t) &\leq K^i.
\end{align*}
\]  

(12a)  

(12b)  

(12c)  

(12d)

Given \( r^i \), let \( x^i_r(r^i) \) denote the corresponding steady state equilibrium. We use the following result (see, e.g., [20]).

Result 1: For \( i \in \{x,y\} \), consider (12) where \( C^i = [C^i_x, C^i_y], A^i = [A^i_x, A^i_y], (C^i_x, A^i_x) \) is observable and \( \mathcal{Z}^i = \{z : H^i z^i \leq K^i\} \) is a polytope, i.e., closed and bounded. Let \( \mathcal{R}^i \) be the set of commands that are steady state admissible, that is, for all \( r^i \in \mathcal{R}^i \) the corresponding equilibria \( x^i_r(r^i) \) satisfy \( C^i x^i_r(r^i) + C^i r^i \in \text{int}(\mathcal{Z}^i) \). Then, with arbitrary precision, the maximum constraint admissible set for \( r^i \in \mathcal{R}^i \) is the polytope

\[
\mathcal{O}_\infty^i = \{ (x^i, r^i) : H^i x^i + H^i r^i \leq K^i \}.
\]  

(13)

defined by a finite number of constraints.

\( \mathcal{O}_\infty^i \) can be computed by solving offline a sequence of linear programs [20]. We denote the section of \( \mathcal{O}_\infty^i \) at reference value \( r^i \) by \( \mathcal{O}_\infty^i (r^i) \), which is the set of states that are in the maximum constraint admissible set when the reference is \( r^i \), i.e., \( x^i \in \mathcal{O}_\infty^i (r^i) \) iff \( (x^i, r^i) \in \mathcal{O}_\infty^i \).

Given \( x^i, i \in \{x,y\} \), the SPRG uses Result 1 to choose \( r^i \) in the sequence of points to be processed, \( \{q(h_T)\}_h \). Let

\[
\kappa(x, \mu, \{q(h)\}_h) = \max_{\eta \in [0, \mu]} \eta \quad \text{s.t.} \quad (x^i, q^i(\mu + \eta)) \in \mathcal{O}_\infty^i \\
\min q^i(\mu + \eta) \geq [C^i_x x^i]_1 - \bar{y}_f^i \\
\max q^i(\mu + \eta) \leq [C^i_x x^i]_1 + \bar{y}_f^i
\]  

(14a)  

(14b)  

(14c)  

(14d)

and let \( \mu(t) \in \mathbb{Z}_{0+} \) be the index of the last processed point within the \( t^\text{th} \) sampling interval, i.e., \( r(t) = q(\mu(t)) \), then

\[
\mu(t) = \kappa(x(t), \mu(t-1), \{q(h)\}_h) \\
r^i(t) = q^i(\mu(t)), \quad i \in \{x,y\}
\]  

(15a)  

(15b)

The SPRG (15) constructs the reference by solving (14), searching how many points can be processed until the next sampling period without violating the constraints, and ensuring that the selected reference can be maintained as a target in the future without violating the constraints. Since constraint satisfaction is ensured only at the sampling instants, (14c), (14d) are included to enforce that all the points \( \{q(\mu+j)\}_{j=0}^{\mu} \) to be processed by the fast stage during the sampling interval are within the fast stage range. Instead of enforcing \( ||y^i_s - q(\mu+j)|| \leq y_f^i \) for all \( j \in \mathbb{Z}_{[0,\mu]} \), (14c), (14d) are enforced on the worst case values to reduce the computations. Let \( q^i(t) = \min_{\mu \in [\mu(t-1), \mu(t)]} q^i(\mu), \quad \overline{\nu}(t) = \max_{\mu \in [\mu(t-1), \mu(t)]} \overline{q}(\mu) \) be the worst case intersampling bounds that are fixed once the reference has been selected, since they depend only on the points that are processed in the current sampling interval. While (14) seems a difficult problem, it is solved by checking the inequalities while scanning the points in \( \{q(h)\}_h \), and, in doing that, the computation of \( \overline{\nu}(t) \) and \( \overline{q}(\mu) \) is straightforward.

By using (15), the feasible reference is such that \( r(t) \in \{q(h)\}_h \) for all \( t \in \mathbb{Z}_{0+} \), and hence the spatial pattern is not deformed, but its traversal is slowed down due by reducing the feed rate. The maximum of \( M \) points in (14) is imposed due to the maximum number of points that can be processed by the
fast stage in a single sampling period of the slow stage. This is needed for satisfying (6), since starting from a trajectory \( \{q(h)\}_h \) that enforces (6), the trajectory \( \{r(t)\}_t \) should not be faster than \( \{q(hT_f)\}_h \) to prevent the fast stage tracking error from increasing. Hence, a maximum of \( M \) points of \( \{q(h)\}_h \) can be processed during each slow stage sampling period, which is the ratio between \( T_s \) and \( T_f \).

**Theorem 1:** Let \( \{q(hT_f)\}_{h=0}^{\infty} \) be a finite-time trajectory such that for all \( h \in \mathbb{Z}_{[0,\bar{h}]} \), \( (x'(q'(h)), q'(h+1)) \in \text{int}(\mathcal{O}_\infty) \), and let \( x(0) \) be such that \( (x'(0), q'(0)) \in \mathcal{O}_\infty \), \( i \in \{x,y\} \). For (12) in closed-loop with (15), (14) is recursively feasible, \( (12d) \) is satisfied and there exists \( i \in \mathbb{Z}_+ \) such that \( r(\bar{T}_s) = q(hT_f) \). □

For this paper, all the proofs are in the Appendix.

**Remark 1:** The reference governor (14), (15), preserves the geometry, i.e., (6) by: (i) selecting only points on the trajectory \( \{q(h)\}_h \) rather than selecting points close to \( \{q(h)\}_h \) but not necessarily on it, and (ii) computing the reference for both axis at the same time, even if the dynamics of the axes are decoupled, allowing the geometry of the path to couple them. The operation of the reference governor amounts to a (nonlinear) transformation on the curve parameter, i.e., in case of \( \{q(h)\}_h \), the time, which does not change the curve. Thus, the reference governor “stretches” the time, to reduce the processing speed when needed to guarantee constraint satisfaction.

**B. Tracking MPC for two-stage machines**

Next, we construct the MPC block in Figure 4 by appropriately specifying the optimization problem (9) to operate with the SPRG. For the ease of notation, in what follows we omit the superscript \( i \in \{x,y\} \), when clear from context, and the subscript \( s \), since all the variables refer to the slow stage.

Given the reference trajectory \( \bar{R}_t = [r_{0|t} \ldots r_{N|t}] \), and the corresponding intersampling worst case bounds \( W_{\min}^{\max} = [\bar{Q}_{0|t} \ldots \bar{Q}_{N|t}] \), the MPC finite horizon optimal control problem is

\[
\begin{align*}
\min_{U_t} & \quad F(x_{N|t}, r_{N|t}) + \sum_{k=0}^{N-1} L(x_{k|t}, r_{k|t}, u_{k|t}) \\
\text{s.t.} & \quad x_{k+1|t} = Ax_{k|t} + Bu_{k|t} \\
& \quad Hc_{k}x_{k|t} + Hc_{q}r_{k|t} \leq K \\
& \quad \{c_{x}x_{k|t}\} \geq \bar{q}_{k|t} - \gamma_f \\
& \quad \{c_{x}x_{k|t}\} \leq q_{k|t} + \gamma_f \\
& \quad x_{N|t} \in \mathcal{O}_\infty(r_{N|t}) \\
& \quad \bar{x}_{0|t} = \bar{x}(t).
\end{align*}
\]

(16a)

(16b)

(16c)

(16d)

(16e)

(16f)

(16g)

(16h)

where \( F(x,r) \geq 0 \) for all \( x, r, \) and \( F(x,r) = 0 \) iff \( x = (x_1(r), r) \), i.e., at the equilibrium for \( r \) where \( y = r \), \( L(x,r,u) \geq 0 \) for all \( x, r, u, \) and \( L(x,r,u) = 0 \) iff \( \bar{x} = x_1(r) \) and \( u = r \), since the dc-gain is 1, \( U_t = [u_{k|t} \ldots u_{N-1|t}] \) is the command sequence to be optimized at time \( t \) and we denote the optimal command sequence at time \( t \) by \( U^*_t \).

Next we prove some of the control strategy shown in Figure 4 where: (i) the SPRG implemented by (15) computes a new reference step, together with its worst case intersampling bounds, from the previous prediction of the MPC terminal state, (ii) the MPC shifts the previous reference by one time step and appends the new reference step provided by SPRG, (iii) solves (16), and (iv) provides the new terminal state to the SPRG.

**Theorem 2:** Consider the MPC that at any time \( t \in \mathbb{Z}_{0+} \) solves (16), where \( r_{k|t} = r_{k+1|t-1}, k \in \mathbb{Z}_{[0,N-1]} \), \( r_{N|t} = q(\mu_{N|t}) \), and \( \mu_{N|t} = \kappa(x_{N|t-1}, \mu_{N-1|t}, \{q(h)\}_h) \). Let (16) be feasible at time \( t \in \mathbb{Z}_{0+} \), then (16) is feasible for all \( t \geq 0 \).

**Theorem 3:** Let \( \{q(hT_f)\}_{h=0}^{\infty} \) be a finite-time trajectory such that for all \( h \in \mathbb{Z}_{[0,\bar{h}]} \) the equilibrium satisfies \( (x_r(q(h)), q(h+1)) \in \text{int}(\mathcal{O}_\infty) \), and let \( (x(0), q(0)) \in \mathcal{O}_\infty \). Consider the MPC that at every iteration solves (16) where

\[
\begin{align*}
 r_{k|t} & = r_{k+1|t-1}, k \in \mathbb{Z}_{[0,N-1]} \quad r_{N|t} = q(\mu_{N|t}), \\
 \mu_{N|t} & = \kappa(x_{N|t-1}, \mu_{N-1|t}, \{q(h)\}_h).
\end{align*}
\]

Assume that

\[
\forall(x,r) \in \mathcal{O}_\infty, \exists \mu \in \mathbb{R} : \quad F(Ax + Bu, r) + L(x, r, u) - F(x, r) \leq 0
\]

(17)

Then, there exists a finite \( t \in \mathbb{Z}_{0+} \) when \( r_{N|t} = q(h) \) and \( (x_{N|t}, r_{N|t}) \in \mathcal{O}_\infty \). □

The assumption in (17) is related to the existence of a control Lyapunov function [11] for the tracking problem. Since the SPRG keeps the reference constant until it can move to the next point and the prediction model is asymptotically stable, for linear-quadratic MPC a simple choice for \( F \) is based on solving the Lyapunov equation with respect to the origin as steady state equilibrium, and use such solution to construct \( F \) as function of \( x \) and \( r \) by translation of steady state state and input. Feasibility of this choice is ensured by using \( \mathcal{O}_\infty \) as terminal set in (16).

**IV. CONTROL SYSTEM IMPLEMENTATION**

In this section we describe the implementation of the numerical algorithms for the control system designed in Section III, in particular the computation of the SPRG and the numerical solver for the MPC optimal control problem (16). Due to the real-time nature of the control system, the numerical algorithms must terminate in a given fixed time, which requires additional considerations. Finally, we briefly discuss the command generation for the fast stage.

**A. Implementation of SPRG and MPC**

The computation of the feasible reference by SPRG requires solving (14). Since \( r^i \) is a scalar the easiest way to solve (14) is by direct evaluation. Thus, solving (14) amounts to the following steps: (i) substitute \( x^i \) in the constraints, eliminate redundant constraints and set \( \eta = M \); (ii) compute \( \eta, q \) from \( \mu \) and \( \eta \); (iii) if \( q(\mu + \eta) \) satisfies the constraints, both the maximum constraint admissible set and the worst case intersampling bounds, then \( r = q(\mu + \eta) \), and terminate; otherwise \( \eta = \eta + 1 \) and go to (ii).

It can be noted that \( x^i \) is given, and hence constraint (14b) can be transformed into \( H^r \leq K_r - H_x \leq x^i \), which, after minor manipulations, results in all but two constraints to
be redundant, since $r^i$ is scalar. Thus, a total of 4 constraints have to be verified: two for the maximum constraint admissible set and two for the worst case intersampling bounds. Finally, instead of the decremental search, different updates for $\eta$ can be applied, such as bi-section. Decremental search is particularly favorable for the situations when it is expected $\eta$ to be very close to its maximum value, i.e., the feed rate is only minimally reduced.

As for the MPC implementation, they key element is the solver for (16). When $F$, $L$ are convex quadratic functions,

$$
L(x,v) = \begin{bmatrix} x' & r' \end{bmatrix} Q \begin{bmatrix} x \\ r \end{bmatrix} + \begin{bmatrix} u' & r' \end{bmatrix} R \begin{bmatrix} u \\ r \end{bmatrix}
$$

$$
F(x,v) = \begin{bmatrix} x' & r' \end{bmatrix} P \begin{bmatrix} x \\ r \end{bmatrix},
$$

where $\varrho_p$, $\varrho_v$, $\varrho_u$, $\varrho_a$ are positive weights on slow stage position, velocity, acceleration and command, and $P \geq 0$ is designed to satisfy the assumptions in Theorem 3, (16) results in a convex parametric quadratic program

$$
\min_{\zeta} \frac{1}{2} \zeta' Q\zeta + \theta' C_p\zeta + \frac{1}{2} \theta' \Omega_p \theta
$$

s.t. $G_p \zeta \leq S_p \theta + W_p.$

where at time $t$, $\zeta = U_t$ and $\theta \in \mathbb{R}^{n_s}$ is the parameter vector, $\theta = [x' R_t' W_t^{\max} W_t^{\min}, r'].$

For solving (19) one can iterate [23]

$$
\xi(t+1) = \left( \frac{Q_d + \phi}{F_d^T} \right)^{-1} \left( \frac{Q_d + \phi \xi(t) + F_d^T \xi(t)}{t} \right)
$$

where $F_d = S_p \theta + W_d$, and $Q_d = G_p Q_p^{-1} G_p$, $S_d = (G_p Q_p^{-1} C_p + S_p)$, $W_d = W_p$, $\Omega_d = C_p Q_p^{-1} C_p - \Omega_p$ are the matrices of the dual problem of (19), and $A^+$ and $A^-$ are the positive and negative parts of matrix $A = A^+ - A^-$. The solution of (19) is constructed from the fixed point $\xi^*$ of (20)

$$
\zeta(t)^* = \Psi_{d2p} (\theta, \xi^*) = \Gamma_d \theta + \Xi_d \xi^*,
$$

where $\Xi_d = -Q_p^{-1} C_p$, $\Gamma_d = -Q_p^{-1} C_p$.

### B. Real-time Implementation

Both SPRG and MPC can be implemented and verified even in embedded platforms with minimal capabilities, because SPRG only requires verifying satisfaction of linear inequalities, and MPC only requires computing the simple iteration (20).

For real-time operation, it is necessary that the number of worst case operations computed by the control algorithm is deterministic, and fits in the available computing time. The number of operations of the SPRG is simple to bound since given that $\eta \leq M$, at most $M$ iterations will be performed, and in each iteration a known number of inequalities is evaluated.

For MPC, computing the number of operations in each control cycle is in general difficult. Some notable exceptions are explicit MPC, see, e.g., [12], and some recently developed low complexity solvers that allows for complexity certification, see, e.g., [24], [25]. These methods estimate an upper bound to the number of iterations (and operations) to compute the MPC solution. Then, a processor must be selected to fit the computational requirements. This often results in an over-dimensioned processor, even by one or more orders of magnitude, due to the estimate being an upper bound of the worst case. Here, we use a different approach that aims at not imposing constraints on the processor, but rather uses all the available computing time, and if that is not sufficient, it exploits a readily available backup solution. Such an approach is beneficial as it avoids over-dimensioning the processor for an unlikely worst case, and hence avoids unnecessary costs.

Due to the simplicity of the solver in Section III, we can compute the number of operations per iteration, so that, given any processor, the maximum number $\bar{\ell} \in \mathbb{Z}_+$ of iterations (20) per computing cycle can be determined. Due to hard real-time requirements, $\bar{\ell}$ in $\mathbb{Z}_+$ may not be enough to always solve to optimality. However, in case the solver does not return a solution within $\bar{\ell}$ iterations, the SPRG solution can be used to construct a feasible solution. Let $\xi(\ell)$ be the candidate solution from (20), $\bar{\xi}$ be the corresponding solution of (19) from (21), and $U_\ell$ be the corresponding optimal control sequence. If $U_\ell$ is feasible (possibly optimal), it will be used, because the constraints are satisfied and the terminal constraint guarantees recursive feasibility for any feasible solution. If $U_\ell$ is not feasible, we exploit the previous feasible solution and the current reference.

**Corollary 1:** Let $U_{t-1}$ be a feasible solution for $t \in \mathbb{Z}_+$. The solution $U_t$ where $\bar{u}_{k[t]} = u_{k+1[t-1]}$ and $u_{N[t]} = r_{N[t]}$ is feasible for (16).

According to Corollary 1, (15) provides a backup strategy for when a feasible solution is not found within the available iterations. By Corollary 1, one can dimension the processor for a certain probability to terminate within $\bar{\ell}$ iterations, and then use (15) as backup strategy.

### C. Fast Stage Command Generation

The fast stage receives as targets the points that have to be processed during a slow stage interval, i.e., for $t \in [(k-1)T_s, kT_s]$, these are $\{q_{h^i \mu(k)}\}$. The curve visiting these points is sampled to obtain a time reference $r_f^i(\tau)$, $i \in \{x,y\}$ with sampling frequency $T_f$. The fast stage command is generated while accounting for the motion of the slow stage as

$$
u_f^i(\tau) = r_f^i(\tau) - \bar{G}_i^* * \hat{u}_s^i(\tau),$$

where $\bar{G}_i^*$ is a discrete-time representation of $G_i^*$ with sampling period $T_f$, which is realized in the upsampling block shown in Figure 4, and $\hat{u}_s^i(\tau)$ is an up-sampling to period $T_f$ of $u_s^i(t)$, so that $\hat{u}_s^i(\tau) = u_s^i(t)$ for all $\tau \in \mathbb{Z}_{\lceil N[M,(M+1)t-1]\rceil}$. In (22) the fast-stage dynamics are infinitely fast or at least significantly faster than the slow stage ones. Alternatively, the motion of the slow stage can be compensated for by including in the upsampling block in Figure 4 an appropriate matching filter

$$u_f^i(\tau) = r_f^i(\tau) - W_f^i(\tau) * u_s^i(\tau), \quad W_f^i(\tau) = \frac{\bar{G}_i^*(\tau)}{G_f^i(\tau)},$$

(23)
Then, the overall worktool motion results in
\[ y^f(\tau) = G^f_j(\tau) * r^f_j(\tau) + \frac{G^f_j(\tau)}{s} W^f_j(\tau) * u^*_{s}(\tau) - G^f_j(\tau) * u^*_{s}(\tau), \]
and hence \( y^f(\tau) = G^f_j(\tau) * r^f_j(\tau) \) which shows that the motion of the worktool is exactly the same as the one of the ideal machine having the dynamics of the fast stage. While in general matching filters such as (23) are not recommended, due to the fact that models are imprecise, in precision manufacturing the models of the stage motion are of very high precision. Also, while the matching filter does not account for the plant constraints, here the velocity, acceleration, and range constraints are enforced by the ideal reference generation and the SPRG-MPC of Section III. While discussed here for completeness, if the stages are well separated in frequency, the impact of the matching filter (23) is negligible.

V. SIMULATION AND EXPERIMENTAL RESULTS

For evaluating the proposed control design, we have built a laboratory prototype two-stage dual-axis machine using Mitsubishi Electric components for manufacturing applications. The workspace is approximately one quarter of that of a real precision manufacturing machine, and in the prototype we have retained the ratios of relevant quantities such as motor power, stage inertia, workspace area, etc. The slow stage actuation is implemented by 3 servomotors, with 2kW power each, driven by servomotor amplifiers. In the frequency range of interest, the closed-loop slow stage response model from position command to position is the 3rd order function whose Bode plot is shown in Figure 3. The range of the slow stage is \( y^s = 0.4\text{m}, i \in \{x, y\} \). The maximal velocity and the maximal acceleration for the slow stage are \( \dot{y}^s = 0.7\text{m/s} \) and \( \ddot{y}^s = 1.5\text{g}, i \in \{x, y\} \), respectively. The fast stage is implemented by a low inertia high speed servomotor that in the frequency range of interest has the 3rd order closed-loop response model shown in Figure 3. The range of the fast stage is \( y^f = 35\text{mm}, i \in \{x, y\} \). A high-end laser diode is used to mark the position of the worktool actuated by the two-stage machine. The microcontroller is based on the TMS320C6678 digital signal processor, for which we reserved only 20% of the computing time for each control loop.

For testing, we have designed using a CAD program (SolidWorks) a set of parts that needs to be machined on a single batch of material. The parts are composed of common shapes machined in manufacturing applications, and include both small and large features. Thus, processing the parts requires both, short range high acceleration movements, and long range smooth movements. The resulting worktool path is shown in Figure 6. The CAD design is then processed by a CAM program realizing its conversion into instructions for a standard numerical motion controller, that is appropriately configured and upsampled to \( T^f \), to generate trajectories for the ideal machine with range of the slow stage and dynamics of the fast stage. This results in the ideal reference trajectory \( \{q(\bar{h})\}_{h=1}^{h} \), with \( h = 232100 \) being the total number of processing points, and an ideal processing time of 46.42s. The accuracy bound for all the tests reported next is \( \varepsilon = 50\text{µm} \), i.e., \( |y^f(t) + y^s(t) - r^i(t)| \leq 50\text{µm} \) for all \( t \in \mathbb{R}_0^+ \). The reference generation enforces such a precision using classical approaches for the ideal single stage machine, and the method proposed here only reduces the feed rate, which is inverse proportional to the tracking error, hence further increasing the precision. It shall be noted that for this application the accuracy is a hard bound, as opposed to being the performance objective as in path following MPC [14], [16], and all tests reported next satisfy such bound.

The controller is designed with the quadratic cost function (18) in (16), \( M = 150, T^s = 30\text{ms} \), and a prediction horizon \( N = 20 \) for the base design. Based on the target processor and on the computational load of iteration (20), we have imposed a limit of \( \bar{\ell} = 500 \) iterations (20) per control cycle per axis, which can easily fit in the available computing time during the control cycle. In practice, such a limit is very conservative for the considered processor, but it has been chosen to demonstrate how Corollary 1 ensures real-time feasibility in the presence of limited computation.

![Fig. 5. Experimental two-stage machine used for experimental validation with key components highlighted. Slow stage servomotors, orange. Slow stage range, red. Slow stage moving base, blue. Fast stage base, actuator and worktool, green.](image1)

![Fig. 6. Worktool path resulting from CAD design of the parts to be machined.](image2)
A. Simulations

First we have performed an extensive simulation study, testing several controller calibrations in closed-loop with a high fidelity model of the two-stage machine, composed of the mechanical models and the servo models that result in the closed-loop axis stage frequency response shown in Figure 3. The results for our best calibration, i.e., the calibration that achieves a desirable trade off between fast processing of the machined part, and low acceleration of the slow stage, and hence limited vibrations, are reported in Figures 7–11.

Fig. 7. Processed path (red) covering the desired path, slow stage motion (black), and points where the reference governor reduces the feed rate (green).

Figure 7 shows the processed path, the slow stage motion, and the points where the reference governor reduces the number of points processed per sample. Most reductions occur on long straight segments when the reference governor needs to guarantee the capability of braking should the reference trajectory change direction suddenly, e.g., due to a corner. The slow stage position is shown in Figure 8 for x and y coordinates together with the constraints related to the maximal allowed distance from the reference.

The smoothness of the slow stage motion stage can be further noted in Figures 9, reporting the velocity and acceleration profiles. For this particular calibration the controller keeps the motion of the slow stage far from the velocity and acceleration constraints. Instead, the fast stage range constraints, i.e., the slow stage tracking error constraints, are often active, as seen in Figure 8, and more easily in Figure 10. Finally, Figure 11 shows for each step the ratio of processed points with respect to the maximum $M = 150$, and the number of iterations executed by the QP solver, where a value above 500 is shown.

Fig. 8. Position of the slow stage x, y axes (blue) and constraints (black) due to the reference and fast stage range.

Fig. 9. Velocity and acceleration of the slow stage x axis (blue), y axis (black) and constraints (red).

Fig. 10. Position of the fast stage for x and y axes and range limits (red).
to indicate that the algorithm did not find a feasible solution within the allowed 500 iterations and the feasible solution was generated by Corollary 1. In these cases, the constraints are satisfied but the trajectory tends to be less smooth, see for instance immediately after \( t = 30 \text{s} \) where the iteration limit is exceeded multiple times first on the \( y \) and then on the \( x \) axis, which results in larger accelerations and velocities.

<table>
<thead>
<tr>
<th>Method</th>
<th>Time [s]</th>
<th>Maximum accel. [m/s(^2)]</th>
<th>Mean accel. [m/s(^2)]</th>
<th>Mean vel. [cm/s]</th>
<th>Mean fast stage motion [mm]</th>
<th>Slow stage work per mass [W/kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFS</td>
<td>55.98</td>
<td>(6.18, 5.113)</td>
<td>(0.66, 0.54)</td>
<td>(7.71, 7.49)</td>
<td>(11.22, 10.74)</td>
<td>(3.02, 2.55)</td>
</tr>
<tr>
<td>MPC-20-B</td>
<td>50.38</td>
<td>(6.18, 10.73)</td>
<td>(0.08, 0.11)</td>
<td>(3.23, 4.13)</td>
<td>(15.35, 16.50)</td>
<td>(0.34, 0.82)</td>
</tr>
<tr>
<td>MPC-20-H</td>
<td>48.25</td>
<td>(7.10, 12.73)</td>
<td>(0.84, 0.70)</td>
<td>(8.59, 8.35)</td>
<td>(5.15, 4.24)</td>
<td>(4.35, 4.00)</td>
</tr>
<tr>
<td>MPC-20-L</td>
<td>54.81</td>
<td>(5.87, 10.11)</td>
<td>(0.07, 0.08)</td>
<td>(2.83, 3.66)</td>
<td>(17.28, 19.98)</td>
<td>(0.26, 0.43)</td>
</tr>
<tr>
<td>MPC-03-B</td>
<td>51.35</td>
<td>(14.66, 15.13)</td>
<td>(0.49, 0.59)</td>
<td>(3.02, 3.90)</td>
<td>(20.92, 24.89)</td>
<td>(5.32, 5.77)</td>
</tr>
<tr>
<td>MPC-03-H</td>
<td>48.58</td>
<td>(13.08, 12.94)</td>
<td>(1.21, 1.18)</td>
<td>(7.86, 7.70)</td>
<td>(9.88, 9.55)</td>
<td>(9.80, 9.55)</td>
</tr>
<tr>
<td>MPC-03-L</td>
<td>55.71</td>
<td>(12.59, 12.57)</td>
<td>(0.45, 0.54)</td>
<td>(2.68, 3.50)</td>
<td>(22.21, 25.53)</td>
<td>(4.93, 5.26)</td>
</tr>
<tr>
<td>SPRG</td>
<td>50.75</td>
<td>(15.69, 15.69)</td>
<td>(1.99, 1.66)</td>
<td>(12.53, 11.02)</td>
<td>(3.51, 3.07)</td>
<td>(10.57, 14.26)</td>
</tr>
</tbody>
</table>

Fig. 11. Top: Feed rate ratio with respect to maximum (\( \eta/M \)). Bottom: MPC iterations (\( x \) axis blue, \( y \) axis black), where 600 indicates that a feasible solution is not found within \( \ell = 500 \) iterations and Corollary 1 is used.

In addition to the results in Figures 7–11, representing the most desirable trade off, we have evaluated the proposed algorithm for different controller calibrations to analyze how the performance is affected by changes in the calibration of the cost function weights (\( \text{specifically, } Q \) in (18)) and the prediction horizon (\( N \) in (16)), and to ensure that the controller operates appropriately under such conditions, that may cause different set of constraints to be active. Table I reports the summary of the results, when using 3 calibrations for the MPC cost function: Best (B, whose results are shown in Figures 7-11) representing a trade off between aggressive tracking and smooth motion, and High (H) and Low (L) aggressiveness, favoring aggressive reference tracking by increasing the tracking weight, and low acceleration by increasing the acceleration weight, respectively, also with a shorter horizon \( N = 3 \). For completeness we have also included the results obtained by running the SPRG alone, i.e., \( u_i(t) = r_i(t), i \in \{x, y\} \), which is feasible by Theorem 1, and the results obtained by applying

![Graphs showing feed rate ratio and MPC iterations](image1)

![Graphs showing position results for MPC-20-H](image2)
a linear frequency separation (LFS) method, along the lines of conventional methods used in production machines. In the LFS approach, the command for the slow stage is obtained by low pass filtering the reference trajectory (see, e.g., [9]), and the command of the fast stage is obtained as the difference between the slow stage position and the reference trajectory. The filter is tuned offline to avoid violation of the acceleration and velocity constraints and the amount of processed points at each step is selected as the largest value, not larger than $M$, that can be reached by the fast stage range.

![Fig. 13. Velocity and acceleration of the slow stage x axis (blue), y axis (black) and constraints (red). Top: MPC-20-H. Bottom: MPC-3-H.](image)

The results in Table I compare the algorithms in terms of machining completion time, mean fast stage motion during the slow stage period, and acceleration, velocity, mechanical work per unit mass, i.e., the integral of the product of velocity and acceleration, of the slow stage. The base MPC shown in Figures 7–11 improves the processing time by more than 10% with respect to the frequency separation method, and still results in lower mean acceleration, velocity, and energy of the slow stage. When compared to using the SPRG alone, the base MPC provides a slightly better processing time, but above all a much smoother motion of the slow stage.

![Fig. 14. Processed path (red) covering the desired path, slow stage motion (black), and points where the reference governor reduces the feed rate (green). Top: MPC-3-L. Bottom: MPC-3-H.](image)

When the cost function is calibrated for a more aggressive tracking (H), the MPC processing time decreases but the mean acceleration, velocity and energy increase, while obviously the opposite occurs in the case of less aggressive tracking (L). Figure 12 shows how in the case of aggressive tracking (H) the slow stage stays away from the tracking error bounds, which results in shorter motion of the fast stage. On the other hand, Figure 13 shows that the velocity and acceleration of the slow stage become larger. Figure 13 also shows that in the case with short horizon ($N = 3$), the MPC uses less predictive information and hence has less capability of smoothing the slow stage motion, resulting in even larger acceleration and velocity. Since the work per unit mass is the product of velocity and acceleration, the previous consideration results in an increase in energy consumption for the designs with more aggressive tuning and shorter horizon. The quantitative values for all these effects are shown in Table I, in terms of mean and maximal acceleration, mean velocity and mean motion of the fast stage. In addition, Table I shows that when the MPC horizon becomes shorter, the processing time is in general only minimally affected, but the smoothness of the motion is significantly affected. Still, all the constraints are satisfied. The path produced by the MPC with short horizon...
in the case of low (L) and high (H) aggressiveness are shown in Figure 14, where the first tends to make the slow stage move the least amount, until a sharp motion is required not to violate constraints, resulting in several sharp turns, while the second tends to remain close to the pattern shape, due to the high tracking weight. This also results in much more frequent feed rate reductions by SPRG in the first case than in the second. However, in all the cases MPC enforces all the constraints and hence processes the pattern correctly. This is also the case for shorter prediction horizon, even though for $N < 3$ the difference between MPC and SPRG is small.

B. Experimental validation

We have tested the algorithm in our experimental laboratory two-stage machine. The results are consistent with those obtained in simulation due to the high precision of the components and to the high fidelity of the simulation model. Note that in all the experiments we have added an initial 2s stop period to make the results easier to see, which can be noted for instance in the initial spikes in velocity and accelerations that are not easily visible in the simulation plots in Section V-A, due to being at the left border of the figures. The results obtained for the best calibration with $N = 20$ are reported in Figures 15–18. These results are consistent with those obtained in simulation in terms of both, completion time and constraints satisfaction. This is expected due to both, the precision of the model and inner-loop servo controllers, and the feedback action of MPC that rejects model errors and disturbances. The only differences can be noted between Figure 9 and Figure 17, where the acceleration in Figure 17 shows slightly larger spikes than reported in Figure 9. This happens because, despite the velocity and acceleration are available to the control loop, the acquisition board only allows logging of the position signal. Hence, velocities and accelerations shown in the figures are obtained by low pass filtered numerical differentiation, similar to [14]. This amplifies the noise, which despite being small because the encoders are high precision, is still present.

We have also tested the other MPC calibrations reported in Table I, and all behaved as expected. In Figures 19–22 we report the results for the MPC with aggressive (H) tuning and short ($N = 3$) horizon, which is probably the most critical to run due to the high accelerations, and short horizon that allows fewer degrees of freedom to compensate for errors. Figures 19–22 show that the MPC works well also in this case and the results are consistent with the simulations, both in terms of processing time and constraint satisfaction.

Overall, the advantage of SPRG-MPC over LFS appears to be due to the capability to exploit the entire range of the operating constraints of the machine. SPRG-MPC actively enforces constraints during processing by slowing down the axis motion only when constraints would otherwise be violated. On the other hand LFS is tuned to avoid constraints, which results in slowed down motion in any condition, due to LFS being based on linear filtering as opposed to the nonlinearities
induced by constraint enforcement in SPRG-MPC.

VI. CONCLUSIONS

We have proposed a method for trajectory generation and control for two-stage precision manufacturing machines where stage actuators with different dynamics and constraints must be coordinated to achieve fast and precise processing. We have developed an architecture based on a cascade of a spatial reference governor, that modulates the reference to achieve feasibility without deforming the physical shape of the pattern being processed, and MPC that seeks the optimal trade-off between processing time and motion smoothness. For the proposed method, we have proved recursive feasibility and finite time processing termination, and we have shown how to achieve real-time processing in any reasonably sized microprocessor. The approach was evaluated in simulation and on a laboratory experiment where the slow stage has been built with scaled dimensions with respect to real processing machines.

Besides developing a solution for a specific relevant application, this paper shows that MPC can be effective for applications in precision manufacturing. For many years such a domain was not a primary target for MPC due to relatively simple processing machine architectures, the limited computing capabilities and the slow optimization algorithms. However, the constant need for higher performance and quality, the increases in computing power, and the recent algorithmic and theoretical developments in MPC have induced transformations in the field that may position MPC to be very effective in manufacturing applications.

APPENDIX

Proof of Theorem 1: Constraint satisfaction and recursive feasibility are obtained by induction based on the fact that at
every \( \tau \in \mathbb{Z}_{0+} \) (15) selects \( r \) based on (14), which exploits the (maximal) constraint admissible set. Due to Definition 1 if \( (x^i(t), r^i(t)) \in \mathcal{O}_\infty \), \( i \in \{x, y\} \), the constraints in (12) are satisfied at time \( t \). For \( t = 0 \), \( (x^i(0), r^i(0)) \in \mathcal{O}_\infty \) for \( r^i(0) = q^i(0) \). For some \( t \in \mathbb{Z}_{0+} \), let (14) have a feasible solution. By (15), this results in \( r(t) \) such that \( (x^i(t), r^i(t)) \in \mathcal{O}_\infty \). Since \( \mathcal{O}_\infty \) is positive invariant for (12), i.e.,
\[
 r(\tau) = r(t), \quad \forall \tau \geq t \implies (x^i(\tau), r^i(\tau)) \in \mathcal{O}_\infty, \quad \forall \tau \geq t,
\]
(14) is feasible for all \( \tau \geq t \), at least by setting \( r(\tau) = r(t) \), since \( \min q^i = \max q^i = r^i \) and hence that (14c), (14d) are automatically enforced by enforcing (14b). Thus,
\[
(x^i(\tau), r^i(\tau)) \in \mathcal{O}_\infty, \forall \tau \in \mathbb{Z}_{0+}.
\]
For finite termination, since \( (x^i_0(q(h)), q^i(h + 1)) \in \text{int}(\mathcal{O}_\infty) \)
\[
\exists \bar{\rho} > 0 : (x^i_0(q(h)) \oplus B(\bar{\rho}), q^i(h + 1)) \subseteq \text{int}(\mathcal{O}_\infty).
\]
Let \( (x^i(t), r^i(t)) \in \mathcal{O}_\infty \), and \( r^i(t) = q^i(h) \), and assume that for all \( \tau \in \mathbb{Z}_{0+}, (x^i(t + \tau), q^i(h + 1)) \notin \mathcal{O}_\infty \), thus, for some \( \rho > 0 \),
\[
x^i(t + \tau) \notin x^i_0(q(h)) \oplus B(\rho), \forall \tau \in \mathbb{Z}_{0+}.
\]
However, since \( A^i \) is asymptotically stable, for any \( \rho \geq 0 \) there exists \( \tau \in \mathbb{Z}_{0+} \) such that \( \|x^i(t + \tau) - x^i_0(q(h))\| \leq \rho \), which contradicts the previous statement. Then, necessarily
\[
\exists \tau \in \mathbb{Z}_{0+} : x^i(t + \tau) \in \{x^i_0(q(h))\} \oplus B(\rho),
\]
and hence \( (x^i(t + \tau), q^i(h + 1)) \in \mathcal{O}_\infty \) and (14c), (14d) are satisfied because \( (x^i(t + \tau), q^i(h)) \in \mathcal{O}_\infty \) and \( (x^i(t + \tau), q^i(h + 1)) \in \mathcal{O}_\infty \). Thus, \( \tau \) is a finite time for moving to a new point. By repeating such reasoning for every point in \( \{q(h)\}_{h=0} \), and since \( h \) is finite, the time to complete the processing of all points is finite.

**Proof of Theorem 2:** Let (16) be feasible at \( t \) and \( U^*_t = [u^*_0 \ldots u^*_N-1|] \) be the optimal input sequence. Since \( x_N[t] \in \mathcal{O}_\infty(r_N[t]) \), by Theorem 1 (14) is feasible at \( t + 1 \) and \( r_{N[t + 1]} = q(\mu_{N[t + 1]}, \mu_{N[t + 1]} = \kappa(x_N[t] \mu_{N[t]}, \{q(h)\}) \).

Then, \( \tilde{U}_{t+1} = [\tilde{u}_{0[t+1]} \ldots \tilde{u}_{N-1[t+1]}] \) such that \( \tilde{u}_{k[t+1]} = u^*_k[t+1], \quad k \in \mathbb{Z}_{[0,N-2]} \), \( \tilde{u}_{N-1[t+1]} = r_{N[t+1]} \) is a feasible input sequence, since the corresponding state trajectory \( [\tilde{x}_{0[t+1]} \ldots \tilde{x}_{N[t+1]}] \) is such that \( \tilde{x}_{k[t+1]} = x^*_k[t+1], \quad k \in \mathbb{Z}_{[0,N-1]} \), \( x^*_k[t+1] \in \mathcal{O}_\infty(r_{N[t+1]}) \), then \( \tilde{x}_{N[t+1]} \in \mathcal{O}_\infty(r_{N[t+1]}) \), hence (16g) is satisfied, and satisfaction of (14c), (14d) implies satisfaction of (16e), (16f). Thus, \( \tilde{U}_{t+1} \) is feasible for (16), and the reasoning can be repeated at any future time instant.

**Proof of Theorem 3:** Assume that the statement is not true. Then for some \( \tau \in \mathbb{Z}_{0+}, r_{N[t]} = r_{N[t]}, \) for all \( t > \tau \). Because of the assumptions, since \( r(t) \) is constant for any \( t > \tau \),
\[
\forall \rho > 0, \exists \rho \in \mathbb{R}, t_0 > \tau : \|x(t_0) - x_c(r(t_0))\| < \rho,
\]
and by the reasoning in Theorem 1, \( x(t_0) \in \mathcal{O}_\infty(q(h + 1)) \) which contradicts \( r_{N[t]} = r_{N[t]} \) for all \( t > \tau \). Considering also that (14c), (14d) are satisfied because \( x(t_0) \in \mathcal{O}_\infty(q(h)) \) and \( x(t_0) \in \mathcal{O}_\infty(q(h + 1)) \), for every point \( q(h) \), after a finite time the next point \( q(h + 1) \) is admissible for processing, and, by repeating the same reasoning, in finite time \( q(h) \) is admissible, thus reaching completion of the processing path.

**Proof of Corollary 1:** The proof is similar to that of Theorem 3. Since \( \tilde{U}_{t-1} \) was feasible at time \( t - 1 \) the first \( N - 1 \) steps of \( \tilde{U}_t \) are such that the constraints in (16) are satisfied. Since \( r_{N[t]} \) is chosen such that \( (x_{N[t]} \in \mathcal{O}_\infty, \) \( Ax_{N-1[t]} + Br_{N[t]} \in \mathcal{O}_\infty(r_{N[t]}), \)
by the positive invariance of \( \mathcal{O}_\infty \). Thus, \( \tilde{U}_t \) also satisfies (16g).

**Acknowledgements**

The authors thank Dr. V. Shilpiekandula and Dr. S. Haghighat for some of the early developments in this research, Dr. M. Brand for seminal work in developing the QP solver, and Dr. W. Vratis and J. Barnwell for their work in building the laboratory experimental setup.

**References**
