International Journal of Control, Automation, and Systems (2011) 9(3):415-424 http://www.springer.com/12555
DOI 10.1007/s12555-011-0300-6

Invited Paper

Model Predictive Control: Review of the Three Decades of Development

Jay H. Lee

Abstract: Three decades have passed since milestone publications by several industrialists spawned a flurry of research and industrial / commercial activities on model predictive control (MPC). This article reviews major developments and achievements during the three decades and attempts to put a perspective on them. The first decade is characterized by the fast-growing industrial adoption of the technology, primarily in the refining and petrochemical sectors, which sparked much interest and also confusion among the academicians. The second decade saw a number of significant advances in understanding the MPC from a control theoretician's viewpoint, which included state-space interpretations / formulations and stability proofs. These theoretical triumphs contributed to the makings of the second generation of commercial software, which was significantly enhanced in generality and rigor. The third decade's main focus has been on the development of "fast MPC," a term chosen to collectively describe the various efforts to bring orders-of-magnitude improvement in the efficiency of the on-line computation so that the technology can be applied to systems requiring very fast sampling rates. Throughout the three decades of the development, theory and practice supported each other quite effectively, a primary reason for the fast and steady rise of the technology.

Keywords: Constrained control, model predictive control, on-line optimization, review, state-space control.

1. INTRODUCTION

Three decades have passed since the seminal paper on Dynamic Matrix Control (DMC) appeared [1]. The publication, along with earlier papers reporting similar ideas [2], generated an unprecedented level of excitement within the process control community and ushered in the era of Model Predictive Control. Three decades after, it is regarded by many as one of the most important developments in process control.

MPC has had a tremendous impact on industrial control practice. Nowadays it is found in the control rooms of almost every refinery and petrochemical plant. The first generation of MPC algorithms were aimed to solve multivariable constrained control problems typical for the oil and chemical industries. In the pre-MPC era, in absence of systematic methods to handle hard constraints, engineers had to resort to ad hoc methods, e.g., single loop controllers augmented by various selectors and other peripherals. MPC provided a first systematic solution to these problems and were immediately welcomed by industry. The industrial interest spawned a number of specialized vendors, DMCC, Setpoint, and others, which were later consolidated and merged into bigger vendors of Aspen Tech, Honeywell, and Invensys. Qin and Badgwell [3] report in their survey a total of nearly five thousand installations by the top five MPC vendors. Given that the survey was conducted more than 10 years ago and the reported figures did not include "in-house" applications and those by numerous other smaller vendors, the real figures today may very well be orders-of-magnitude higher.

Equally important is the impact MPC made on control research. Beginning with the attempt to understand the DMC, the research community has made several significant strides over the three decades. Though the idea of receding horizon control dates back to the 60s [4], constrained optimization became a part of the algorithm only during the 80s when the cheap microprocessors became readily available. This added feature, while extremely useful for industrial process control problems, made the control law implicit and nonlinear, and MPC for a while appeared to defy mathematical analysis. Adding to the difficulty was the finite impulse response or step response model, which most vendors adopted despite their awkwardness. Through the efforts of many researchers, however, MPC now rests on a firm theoretical foundation, with rigorous closed-loop stability conditions and performance guarantees like other classical optimal control techniques, e.g., the Linear Quadratic Regulator (LQR). Attesting to the MPC's impact on the control research is the recent announcement of the first recipient of High Impact Paper Award, a major new IFAC award to acknowledge the impact of a paper in any official

Manuscript received April 29, 2011.
Jay H. Lee is with the Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Korea (e-mail: jayhlee@kaist.ac.kr).

© ICROS, KIEE and Springer 2011

IFAC journals on the broad areas of Automatic Control theory and application: It was given to a review paper on stability and optimality of MPC, coauthored by Mayne, Rawlings, Rao, and Scokaert [5], which currently shows the ISI citation number of over 2000, a figure seldom heard of for a control theory paper.

The aim of this paper is to survey major developments in MPC during the past three decades and put a perspective on them. The readers should be able to see how the MPC first started, why it was able to penetrate into industrial control rooms so fast, what theoretical problems were formulated and solved along the years, and where the recent thrusts have been. It is intended serve as a starting place for more in-depth investigation by the interested readers. One may ask why another review is needed at this time when a large number of review papers and books have already been written on the subject [3,5,6-11]. However, all these reviews are at least a decade old and miss some of the significant advances that occurred in recent years. Such is the effort to improve the efficiency of the on-line optimization part of the MPC, by orders of magnitude, so that it can be applied to application problems that require very fast sampling, e.g., those involving mechanical and mechatronic systems. In addition, it is always enjoyable to look back at a success story like the one of MPC, which exemplifies how theoreticians and practitioners can support each other to spur progress.

The literature of MPC has grown too vast over the years to allow for a comprehensive review. Hence, the references provided here are intended to be only representative. Readers will find references that will serve as helpful starting points for further exploration of the literature. The next three sections will describe the three decades of MPC, one at a time, emphasizing the main driving factors and the major achievements. The paper ends with a summary and some perhaps subjective and speculative view of the author.

2. FIRST DECADE: INDUSTRIAL SUCCESS STORIES

2.1. Pre-DMC years

Essential features of MPC can be seen in some of the earliest installations of computer based supervisory control, which date back to 1950s. They include projects by various oil and petrochemical industries, including Texaco (with Dr. Wooldridge), Monsanto, B.F. Goodrich, Riverside Cement, Union Carbide, etc. [12]. Notable among them is a project between Standard Oil Co. of California and IBM, in which data from a fluid catalytic cracking unit at El Segundo were sent via teletype to an IBM 7090 mainframe computer, located in San Francisco. The optimal process settings were computed and sent back to El Segundo every 15-20 minutes, which were then implemented manually by the operators. Following the success of the approach, on-site process computer was installed to remove the need for the telecommunication and automate the adjustments.

Despite the potential benefits seen in these projects, computer based control did not see a widespread use within the process industries, probably due to the prohibitively high level of effort and expense associated with it. In the 60s and early 70s, the idea of MPC continued to show up sporadically in the literature. Propoi [13] proposed to use linear programming to control linear systems with hard constraints. Lee and Markus [4] made the following prescient statement succinctly summarizing the essence of MPC, much ahead of its time: "One technique for obtaining a feedback controller synthesis from knowledge of open-loop controllers is to measure the current control process state and then compute very rapidly for the open-loop control function. The first portion of this function is then used during a short time interval, after which a new measurement of the process state is made and a new open-loop control function is computed for this measurement. The procedure is then repeated."

2.2. Early industrial algorithms

It wasn't until the mid-70s when cheaper, more reliable and more powerful microprocessors and distributed control systems arrived, that computer based process control took on a real life. Not coincidentally, several practitioners around this time started reporting phenomenal success they had with the use of computer-model-based control in controlling various processing units in refineries. It is difficult to attribute the birth of MPC to a single person as several seminal publications appeared, more or less independently, around a same time. Richalet and coworkers [2] introduced a technique called Model Heuristic Predictive Control and reported successful applications to a dozen large-scale industrial processes including a fluid catalytic cracking column, a steam generator, and a PVC plant. Around the same time, Charlie Cutler was generating a lot of interest after reporting the use of a multivariable model-based control algorithm called Dynamic Matrix Control (DMC) in the 1979 AIChE Annual Meeting and the 1980 Joint Automatic Control Conference [1]. MHPC and DMC differed in that the former employed a finite impulse response (FIR) model and a 'reference trajectory' (i.e., a path that specifies how each CV should return from the current position to the desired setpoint) along with 'coincidence points' (i.e., the points in the prediction window at which the output should be on the specified reference trajectory), whereas the latter employed a truncated step response (TSR) model and least squares minimization of errors with respect to a constant setpoint. Hence, the main tuning knobs for MHPC were the time constants of the reference trajectories, while for DMC they were the weight parameters in the least squares minimization.

Fueled by these publications and the significantly lowered barrier to implementation, the use of model-based computer control spread rapidly across refineries and petrochemical plants in the Western world. Techniques used in different companies had their own idiosyncratic features and were referred by various acronyms like DMC, MAC, ID-COM, etc. These later became trade names for the commercial software products and continue to be used today, but the generally agreed collective term for these techniques is Model Predictive Control (MPC).

Most of these algorithms were heuristic in nature. They employed time-domain response based models, e.g., FIR and TSR, were completely deterministic without any explicit disturbance model, and lacked stability guarantees and systematic tuning guidelines. It was wrongly claimed by the practitioners that the use of the time-domain response models was what imparted the MPC with the robustness needed for the practical control problems. Initially, process control researchers struggled to understand the essence of these "strange" algorithms, which looked far from being conventional. Despite vigorous attempts, the implicit nonlinear control law appeared to defy a traditional mathematical analysis. However, these struggles led to the development of Internal Model Control, which failed to shed light on the constrained control but led to many insights on robust control [14,15].

2.3. Generalized predictive control

Independent from the development in the process industries, the adaptive control community saw a rise of its own version of MPC called Generalized Predictive Control (GPC). The motivations behind the developments of GPC and DMC were very different: Whereas DMC was conceived to handle multivariable constrained control problems typical of the oil and petrochemical industries, GPC was intended to offer a new alternative to the self-tuning regulator, mainly to overcome its robustness problem. The first complete exposition of GPC appeared as a two part paper in Automatica [16,17]. GPC naturally employed a transfer function model, like much of the work in adaptive control, and stochastic aspects played a key role from the very beginning. Because of this, it was awkward to apply to multi-variable control problems and lacked the inclusion of constraints, an important feature needed for most process control problems. Though regarded to be theoretically more sound than DMC and its contemporaries, GPC went largely unnoticed by the practitioners.

3. SECOND DECADE: FOUNDING OF MPC THEORY

The second decade of the MPC saga can be summarized by two phrases: the continued penetration of the commercial MPC algorithms into a wide variety of industries and the development of a sound theoretical foundation for it.

3.1. Exemplary algorithm

During the late 80's, researchers had begun to realize that, counter to the industrialists' claims, the FIR or TSR models were not what imparted MPC with the observed robustness, and hence needed not be essential features of MPC. In fact, employing a state-space model may be beneficial in several ways [18,19]. Hence, the research literature on MPC began to adopt the following formulation as the standard.

Consider the following state-space system:

$$x(i+1) = Ax(i) + Bu(i). \tag{1}$$

At each sample time, the control input is determined by solving the following optimization with the initial state x_0 set equal to the measured (or estimated) state value:

$$\min_{\hat{u}(0),\cdots,\hat{u}(p-1)} \left\{ V_p \triangleq \sum_{i=0}^{p-1} \left[\hat{x}^T(i)Q\hat{x}(i) + \hat{u}^T(i)R\hat{u}(i) \right] + \hat{x}^T(p)Q_t\hat{x}(p) \right\} \tag{2}$$

with constraints

$$\hat{x}(i+1) = A\hat{x}(i) + B\hat{u}(i), \quad \hat{x}(0) = x_0 \tag{3}$$

and

$$\hat{u}(i) \in \mathbb{U}, i = 0,\cdots,p-1 \tag{4}$$

$$\hat{x}(i) \in \mathbb{X}, i = 1,\cdots,p-1 \tag{5}$$

$$\hat{x}(p) \in \mathbb{Z} \tag{6}$$

$$i = 0,\cdots,p-1.$$

In the above, the notations $\hat{x}$ and $\hat{u}$ are used to distinguish the predicted state and computed input from the actual state x and implemented input u. It is often assumed, for simplicity of the argument, that $Q > 0$, $Q_t > 0$ and $R > 0$. In the simplest case, feasible sets, $\mathbb{X}$ and $\mathbb{U}$, are defined by their upper and lower bounds on x and u. The flexibility of assigning a tighter feasible region for the terminal state (i.e., $\mathbb{Z} \subset \mathbb{X}$) is allowed. In general, these sets are assumed to be convex, compact and to include the origin as an interior point so that the origin is a feasible stationary point for the system. The feasible sets are also assumed to be time invariant. The resulting problem is a constrained least squares problem, which is to be solved using a numerical optimization technique.

For the convenience of presentation, we will refer to the above finite horizon optimal control problem as $J_p(x_0)$; the same notation will also be used to denote the optimal cost.

The strategy of solving $J_p(x_0)$ at each sample time with x_0 set equal to $x(k)$ (the state feedback signal), and implementing the resulting solution (i.e., setting $u(k)= \hat{u}(0)$) defines a *state feedback* control law, which can be expressed as

$$u(k) = \mathcal{U}^*(x(k)), \tag{7}$$

where $\mathcal{U}^*(\cdot)$ represents the solution $u(0)$ to the convex program $J_p(x_0)$ with $x_0= x(k)$ where $x(k)$ is the state feedback signal. Stability and optimality results in the literature typically apply to the above control law. When viewed in this manner, we suddenly realize that MPC is not that different from the optimal control problems studied in the 50s and 60s. However, in the classical approaches, the aim was to derive an explicit form of an optimal feedback law *off-line*. This required a solution to the Hamilton-Jacobi-Bellman (HJB) equation, which is generally not solvable except for a few special cases, e.g., the celebrated Linear Quadratic Regulator problem. Because of this, most interesting optimal control problems had remained unsolved, practically speaking. The MPC approach bypasses the need to solve the HJB equation by performing an open-loop optimal control calculation *on-line* based on the state feedback.

In most practical problems, measurements of state variables are not directly available. To extend the above to the output feedback case, one typically employs a state estimator and uses the estimated state, denoted hereafter by $x(k|k)$, instead of the actual state $x(k)$.

3.2. Stability results

It is relatively easy to prove that the "infinite-horizon" control law resulting from setting p=m=∞ is closed-loop stable, as long as the starting state is in the stabilizable set. To prove this, one can simply use the fact that the optimal cost function $J_\infty(x(k))$ qualifies as a Lyapunov function. Such a control law, though desirable in theory, cannot be implemented in practice as it would require solving an infinite dimensional optimization problem at each sample time. Then, the main question becomes how to approximate this infinite horizon feedback law by solving a finite horizon problem of (2)-(6), *without losing the stability property and also significantly sacrificing the performance*. In addition to the stability, one must also ensure that the constrained optimization remains feasible throughout.

For stability and performance of the feedback law defined by (2)-(6), the choices of the terminal weight Q_t and the terminal constraint (6) turn out to be critical. Almost all stability results in the MPC literature employ conditions on one or both. Perhaps the most straightforward way to approximate the infinite horizon control problem with a finite horizon problem is to add an equality constraint on the terminal state $x(p)$=0. Some of the earliest stability results used such a constraint [20,21]. Though easiest to handle from the viewpoint of mathematical analysis, a significant drawback of this approach is that the terminal equality constraint can be quite severe; it is hard to satisfy and artificially imposing such a strict constraint can lead to substantial performance loss. For example, for such an approach to work, the underlying system needs to be reachable, instead of just being stabilizable.

Another interesting approach to achieve stability was suggested by Rawlings and Muske [22], who showed that choosing the terminal weight Q_t as the solution to a Lyapunov equation and the terminal set $\mathbb{Z}$ as an *output admissible set* for the open-loop system $x(i+1)$=$Ax(i)$ make the finite horizon control law equivalent to the infinite horizon control law (but with the control moves constrained to be zero beyond the horizon p, (i.e., $\hat{u}(p)=\hat{u}(p+1)=\cdots=0$). The concept of output admissible set was first introduced by Gilbert and Tan [23] and it is imposed here to ensure that the state constraints are satisfied beyond the finite prediction horizon. Such a set $\mathbb{Z}\subset\mathbb{X}$ must be positively-invariant (i.e., if $x\in\mathbb{Z}\rightarrow Ax\in\mathbb{Z}$). To make the constraint least restrictive, it is desired to seek a maximal such set, i.e., the largest possible set with the stated properties. However, the maximal set is neither simple-shaped nor easy to find; hence, a tight subset of manageable complexity is to be employed instead. With these choices of Q_t and $\mathbb{Z}$, they go to prove the closed-loop stability of such a control law, by showing that the cost function $J_p(x(k))$ is a Lyapunov function. Note that, since zero control action is assumed beyond the prediction horizon, for an open-loop unstable system, all the unstable modes must be "zeroed in" at the end of the horizon. This requirement translates into equality constraints.

A more general approach is to assume that some locally stabilizing control law (usually a linear feedback of the form $u(i)$=$Fx(i)$) is enforced beyond the prediction horizon. In this case, the terminal set $\mathbb{Z}$ must be chosen as an *output admissible set* for the closed-loop system $x(k+1)$=$(A+BF)x(k)$. Such a set $\mathbb{Z}\subset\mathbb{X}$ must be positively-invariant for the system $x(i+1)$=$(A+BF)x(i)$. (i.e., if $x\in\mathbb{Z}\rightarrow(A+BF)x\in\mathbb{Z}$), and $Fx(k)\in\mathbb{U},\forall x\in\mathbb{Z}$. Again, to make the constraint least restrictive, it is desired to find a maximal such set, but a more practical approach would be to seek a balance between optimality and complexity. In addition, the terminal weight should be chosen so that $x^T(p)Q_t x(p)$ represents the infinite horizon cost under the assumed feedback law of $u(i)$=$Fx(i)$ with the starting state $x(p)$. For the linear system, this involves solving a Lyapunov equation for the system $x(i+1)$=$(A+BF)x(i)$. Keerthi and Gilbert [24] are the ones who first proposed the approach.

One potential choice of the local controller is the unconstrained Linear Quadratic Regulator, which is optimal for the unconstrained infinite horizon problem. LQR can be easily derived by solving a Riccati equation. The resulting MPC control law, with the corresponding terminal weight and the terminal constraint, is indeed an optimal feedback controller for the infinite horizon problem if the terminal constraint remains inactive [25-27]. If the terminal constraint does become active, it loses the optimality property as it implies that the computed moves may be different from the optimal ones without the constraint. Hence, for the control law to be optimal for the constrained LQ problem, a sufficiently large window of size p must be used so that the terminal constraint is satisfied *automatically*. Such a p may depend on x and hence may have to be calculated on-line, however.

Besides the approaches using terminal penalty / constraints, other approaches have been proposed. One notable alternative approach employs a contraction constraint, that requires the size of the state to shrink over the prediction horizon [28]. More generally, it chooses a positive-definite function of the state and requires this function to decrease over time in the optimization. To ensure feasibility, the size of the prediction horizon is not set a priori but is treated as an optimization variable. The whole computed input sequence can be implemented in open loop until the end of the horizon, as originally suggested, or the optimization can be repeated after some time as suggested in [29]

Closed-loop stability results for the above were proved for the state feedback case. In practice, most systems do not allow direct measurements of state variables, so a state observer of some sort must be employed. Since the MPC controller even for a linear system is nonlinear due to the inclusion of constraints in the optimization, the usual separation result between an observer and a state feedback regulator does not hold. However, it turns out that the nonlinearity of the controller is not too severe (i.e., it is Lipschitz continuous) and a combination of a stable state-feedback MPC control law and a stable observer guarantees stability [30].

In retrospect, many related results existed before the 90s that could have been used to establish the stability of MPC. These results, however, were somewhat fragmented and missed one or more key components of MPC: They were developed under different assumptions in the context of different system types - unconstrained and constrained, linear and nonlinear, time-invariant and time-varying. It is during the 90s that these results were rediscovered and brought together in the context of a well-defined MPC algorithm. For those who wish to delve further into the issues of feasibility, stability, and optimality of MPC, an excellent review exists [5].

3.3. Robust MPC

MPC, being a feedback control method, has some inherent robustness, which was analyzed by several researchers [31-34]. However, when a quantitative description of the model uncertainty is available, it may be beneficial to consider all possible future trajectories under the given uncertainty description in the optimal control calculation. MPC based on a model with an uncertainty description, e.g., disturbance and parameter bounds, was actively studied during this period. Earliest of such works formulated robust MPC as a min-max problem, which attempts to find an input trajectory that minimizes the worst-case error over the output trajectories possible for the given model set [35,36]. However, several examples pointed out that the receding horizon control law resulting from such a formulation was not robust at all. It was immediately realized that the lack of robustness was due to the limitation of the open-loop optimal control formulation: It failed to account for the fact that the control calculation would be repeated in a receding horizon fashion, with feedback update. Lee and Yu [37] presented an argument indicating the deficiency of the open-loop formulation and presented an alternative formulation based on dynamic programming. With some modification, they were able to formulate a MPC algorithm that solves a convex program at each time and guarantees robust stability.

It was argued that a more logical approach was to perform the minimization over feedback control laws rather than control inputs [38]. However, such an approach cannot be implemented directly since the possible control laws do not yield a finite-dimensional parametrization. Kothare, Balakrishnan and Morari [39] presented an interesting formulation where the minimization at each sample time searched over all linear state feedback laws to minimize the worst case error. The problem was formulated as a Linear Matrix Inequality (LMI), which is convex and can be solved through semi-definite programming. More general extensions of the approach appeared subsequently [40] but all these methods, though yielding useful insights, are short of being practically implementable due to the need to solve the min-max problem on-line.

3.4. Nonlinear MPC

A logical extension of the above is the use of a nonlinear model, given the ubiquitousness of nonlinear control problems and the lack of a universally accepted nonlinear control solution. Extending the MPC formulation for constrained linear systems to nonlinear systems is conceptually straightforward but met with practical difficulties. Most of the stability results for the constrained linear systems apply to nonlinear systems without modification. In fact, many of the earlier stability results for constrained optimal control [24,41] were developed in the context of a general nonlinear system. However, the implementation is greatly complicated by the computational complexity in finding a globally optimal solution to a nonconvex optimization problem. Despite the much progress made in the area of nonlinear programming, computational complexity remains as a major obstacle for designing a practically implementable nonlinear MPC algorithm with guaranteed stability. Naturally the researchers focused on finding a formulation that does not require a globally optimal solution to be found, just a feasible solution. This idea was first explored by Mayne and Michalska [41]. In their approach once a feasible solution is found, the subsequent calculation preserves the feasibility and tries to merely improve the cost. Chen and Allgower [42] presented an approach called quasi-infinite-horizon MPC, where a quadratic terminal penalty corresponding to the infinite horizon cost of the linearized system is imposed. Because a terminal constraint is used to force the state to lie within a prescribed terminal region, within which the system is stabilized by the linear feedback, feasibility alone implies the asymptotic stability.

3.5. Industrial applications

During this period, MPC continued to find its way into industrial control rooms. Many first-generation MPC vendors, which were small startup companies, got merged into bigger vendors like Aspen Tech and Honeywell. In addition, some vendors, notably Shell Global Solutions, parted from the traditional time-response models and introduced software adopting more standard state-space models and all the associated tools (e.g., Kalman filter). This new generation of industrial MPC algorithms looked much closer to the classical optimal control algorithms like LQG but with the added benefit of constraint handling.

At CPC 5, Qin and Badgwell reported a survey involving five major MPC vendors (DMC, Setpoint, Adersa, Honeywell Profimatics, and Treiber Controls) [43]. The result shows 2233 applications, mostly concentrated in the refining and petrochemical industries, but a significant number of applications were also reported for other industrial sectors like chemicals, pulp and paper, food, and mining. Approximately five years later, in another similar survey conducted by the same people involving essentially the same set of vendors [3], the total number of applications doubled, indicating a growth rate of about 15%. Their survey also showed a more widespread use of nonlinear MPC, total number of applications being around 100. Unlike the linear MPC, the majority of the nonlinear MPC applications were found in chemicals, air/gas, and polymer industries.

3.6. Summary

This decade can be summarized as a period of enlightenment. Formulations studied in the literature became more standardized and mathematical properties like stability were successfully established. With these came insights and a perspective that placed MPC among some of the best known optimal control techniques. Meanwhile, in industries, applications continued to flourish and small startup vendors got bought out by household names. Though the researchers agreed on a standard form of the algorithm for the purpose of theoretical investigation, an important attraction of MPC is the flexibility that the use of on-line optimization affords in specifying the performance measure and constraints. Hence software that different vendors marketed continued to have many idiosyncratic features, which made analysis and comparison of commercial software more difficult.

4. THIRD DECADE: DIVERSIFICATION THROUGH FAST MPC

4.1. MPC for hybrid systems and systems with logical constraints

The second decade ended with the emergence of "Hybrid MPC," which is aimed at systems with logical rules and continuous dynamics. Bemporad and Morari [44] proposed a standard mathematical description for such systems and coined the term "mixed logical dynamical (MLD)" systems. Due to the presence of logical rules, both the system description and constraints contained binary variables and a standard formulation of MPC based on them yielded a mixed integer linear programming (MILP) or a mixed integer quadratic programming (MIQP) problem to be solved on-line. Though MILP or MIQP are well-studied optimization problems due to their prevalence in scheduling applications, and reliable general commercial solvers (e.g., CPLEX by ILOG) exist, it is a whole different matter to apply them on-line. This is especially true since many hybrid systems of interest were mechanical systems requiring much faster sampling than process applications did. This shortcoming, however, had a beneficial effect of motivating the development of *Explicit MPC*, which was aimed at replacing the on-line optimization with table lookup.

4.2. Explicit MPC

Bemporad *et al.* [45] presented a technique to determine the linear quadratic regulator for constrained systems through off-line multi-parametric linear programming (mp-LP) and multi-parametric quadratic programming (mp-QP). The control law was shown to be piecewise linear and continuous, and could be implemented as a lookup table, i.e., different linear state feedback laws apply to different polyhedral regions. Hence the on-line control computation is reduced to determining the region associated with the current state and then applying the stored control law associated with that region. The multi-parametric programming approach was later used to derive an explicit form of MPC control laws for hybrid systems and uncertain systems [46].

One problem with the approach was that, as the horizon size, the number of states, and the number of constraints grew, the number of polyhedral regions grew quickly, making the lookup table approach difficult to implement in practice. Therefore various explicit MPC design and search algorithms with some minor sacrifice of optimality were proposed. Johansen and Grancharova [47] presented a method to determine an approximate explicit piecewise linear state feedback by imposing an orthogonal search tree structure on the partition. They showed that this led to a real-time computational complexity that is logarithmic in the number of regions in the partition while maintaining asymptotic stability and constraint fulfillment. Tondel *et al*. [48,49] extended this result to develop a new exploration strategy that avoids unnecessary partitioning by analyzing the geometry of the polyhedral partition and its relation to active constraints at the optimum of the quadratic program. Bemporad and Filippi [50] presented a method to solve the mp-QP approximately, allowing to tradeoff between optimality and a smaller number of partitions. They provide analytical error bounds and also showed that the key properties like closed-loop stability and constraint fulfillment were preserved. Johansen, Petersen and Sluphaug [51] developed a similar method to obtain explicit suboptimal solutions, with lower numbers of regions. More recently, Canale, Fagiano, and Milanese [52] proposed to use set membership methodologies to approximate MPC laws, with an explicit bound for approximation errors and preserving key properties like stability and feasibility. Finally, Jones and Morari [53] introduced a method to compute approximate explicit control laws that trade-off complexity against approximation error based on a polyhedral approximation to the optimal cost function and barycentric interpolation. The controller was shown to be feasible and stabilizing.

4.3. Fast optimization

Aside from the effort to calculate an explicit MPC control law off-line, a number of researchers aimed to achieve a speedup in the on-line optimization through the development of customized optimization algorithms.

Customized algorithms exploiting the particular structure of the MPC problem were designed and they were shown to be orders of magnitude faster than a generic optimization solver. Rao, Wright, and Rawlings [54] proposed an interior point method, of which the cost grows linearly in the horizon length, compared with cubic growth for a naive approach. Later, the same group proposed, mainly for large-scale problems, to implement partial enumeration (PE) that combines a table storage and on-line optimization [55]. Ferrau, Bock and Diehl [56] exploited solution information of the previous QP under the assumption that the active set does not change much from one QP to the next. They applied the method to a challenging test problem composed of 240 variables and 1191 equality constraints, and showed that the CPU time was well below 100ms per QP, one order of magnitude faster than a standard QP solver. More recently, Wang and Boyd [57] showed that computation of the control actions for a problem with 12 states, 3 controls, and horizon of 30 time steps (resulting in a QP of 450 variables and 1284 constraints) could be solved in around 5 milliseconds. Other examples were presented showing impressive numbers for the speedup. The speedup was achieved mainly by using an approximate primal barrier interior point method adored with several customized features like fast Newton step computation, warm start and a fixed barrier parameter.

Compared to the explicit method, this line of approach has an important added advantage that weight parameters, horizon size, and model parameters can be changed on the fly. For the explicit method, an entirely new lookup table would have to constructed by solving a new multiparametric programming problem.

4.4. Application

A major shift seen from the past decades is in the type of systems where MPC is applied. A significant number of applications involving mechanical and electronic systems are now being reported in the literature. This is made possible by the aforementioned developments, which enable the implementation of MPC at sample rates that are orders-of-magnitude faster than in traditional process applications. The reported applications include vehicle traction control [58], suspension [59], direct injection stratified charge engines [60], ducted fan in a thrust-vectored flight control experiment [61], automotive powertrains [62,63], magnetically actuated mass spring damper system [64], power converters [65,66], multicore thermal management [67], and so on.

5. CONCLUSIONS

Since DMC and IDCOM were introduced to the control community more than three decades ago, MPC has been in the center stage of control research and application. Over the three decades, we witnessed MPC being transformed from industrial heuristics to one of the most influential control techniques, both in terms of research and practice. MPC is now a part of every refinery and chemical plant operation. In addition, its scope of application continues to broaden, to include mechanical and mechatronic systems that require sampling rates that are orders of magnitude faster. On the theoretical front, after a decade of much confusion, the research community successfully worked out theories to allow design of MPC laws with guaranteed stability and optimality. Many of the contributing pieces were not new and they were rediscovered and tailored to form a unified theory. Now MPC is accepted as the most effective and complete technique for constrained control problems. Throughout, theory and practice supported each other exceptionally well to spur the progress.

What remains now after the three decades of so much progress? One can only speculate where a next major breakthrough will come in this area. It may be in the robust MPC area, which so far yielded some useful theories but lacks a practically implementable method. MPC as we know thus far performs *open-loop* optimization on-line repeatedly based on trajectories predicted with on-line measurements. Since this formulation lacks the proper accounting of future feedbacks, it appears inherently limited to handle uncertain systems in an optimal manner. Dynamic programming provides a rigorous solution to such optimal control problem but it has its own limitation arising from the "curse of dimensionality." Approximate dynamic programming developed by the artificial intelligence community holds some promise in this regard [68]. Successfully solving this problem will open up a whole new set of applications including those involving scheduling, investment, and other economic decisions. This problem still remains as an exciting challenge.

REFERENCES

[1] C. R. Cutler and B. L. Ramaker, "Dynamic matrix control - a computer control algorithm," *Proc. Joint Automatic Control Conf.*, San Francisco, CA, 1980.

[2] J. Richalet, A. Rault, J. D. Testud, and J. Papon, "Model predictive heuristic control: applications to industrial processes," *Automatica*, vol. 14, no. 5, pp. 413-428, 1978.

[3] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, pp. 733-764, 2003.

[4] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, Wiley, New York, NY, 1967.

[5] D. Q. Mayne, J. Rawlings, C. Rao, and P. Scokaert, "Constrained model predictive control: stability and optimality," *Automatica*, vol. 36, pp. 789-814, 2000.

[6] C. E. Garcia, D. M. Prett, and M. Morari, "Model predictive control: theory and practice - a survey," *Automatica*, vol. 25, no. 3, pp. 335-348, 1989.

[7] J. H. Lee and B. Cooley, "Recent advances in model predictive control and related areas," *Proc. of the 5th International Conference on Chemical Process Control*, aIChE Symposium Series, vol. 93, pp. 201-216, 1997.

[8] D. Q. Mayne, "Nonlinear model predictive control: an assessment," *Proc. of the 5th International Con-*

ference on Chemical Process Control, aIChE Symposium Series, vol. 93, pp. 201-216, 1997.
[9] M. Morari and J. H. Lee, "Model predictive control: past, present, and future," *Computers and Chemical Engineering*, vol. 23, pp. 667-682, 1999.
[10] R. R. Bitmead, M. Gevers, and V. Wertz, *Adaptive Optimal Control*, Prentice Hall, Englewood Cliffs, NJ, 1990.
[11] E. Camacho and C. Bordons, *Model Predictive Control*, Springer, Berlin, 1998.
[12] T. M. Stout and T. J. Williams, "Pioneering work in the field of computer process control," *IEEE Annals of the History of Computing*, vol. 17, no. 1, pp. 6-18, 1995.
[13] A. I. Propoi, "Use of linear programming methods for synthesizing sampled-data automatic systems," *Automn. Remote Control*, vol. 24, no. 7, pp. 837-844, 1963.
[14] C. E. Garcia and M. Morari, "Internal model control 1. a unifying review and some new results," *Ind. Eng. Chem. Process Des. and Dev.*, vol. 21, pp. 308-322, 1982.
[15] C. E. Garcia and M. Morari, "Internal model control 3. multivariable control law computation and tuning guidelines," *Ind. Eng. Chem. Process Des. and Dev*., vol. 24, pp. 484-494, 1985.
[16] D. W. Clarke, C. Mohtadi, and P. S. Tuffs, "Generalized predictive control-I. the basic algorithm," *Automatica*, vol. 23, pp. 137-148, 1987.
[17] D. W. Clarke, C. Mohtadi, and P. S. Tuffs, "Generalized predictive control-II. Extensions and interpretations," *Automatica*, vol. 23, pp. 149-160, 1987.
[18] J. H. Lee, M. Morari, and C. E. Garcia, "State space interpretation of model predictive control," *Automatica*, vol. 30, pp. 707-717, 1994.
[19] K. R. Muske and J. B. Rawlings, "Model predictive control with linear models," *AIChE J.*, vol. 39, no. 2, pp. 262-287, 1993.
[20] D. L. Kleinman, "Stabilizing a discrete, constant, linear system with application to iterative methods for solving the Ricatti equation," *IEEE Trans. on Automatic Control*, vol. AC-19, pp. 252-254, 1974.
[21] W. Kwon and A. Pearson, "A modified quadratic cost problem and feedback stabilization of a linear system," *IEEE Trans. on Automatic Control*, vol. 22, no. 5, pp. 838-842, 1977.
[22] J. B. Rawlings and K. R. Muske, "The stability of constrained receding horizon control," *IEEE Trans. on Automatic Control*, vol. 38, pp. 1512-1516, 1993.
[23] E. G. Gilbert and K. T. Tan, "Linear systems with state and control constraints: the theory and application of maximal output admissable sets," *IEEE Trans. on Automatic Control*, vol. 36, no. 9, pp. 1008-1020, 1991.
[24] S. Keerthi and E. Gilbert, "Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: stability and moving-horizon approximations," *Journal of Optimization Theory and Applications*, vol. 57, pp. 265-293, 1988.
[25] D. Chmielewski and V. Manousiouthakis, "On constrained infinite-time linear quadratic optimal control," *Systems and Control Letters*, vol. 29, pp. 121-129, 1996.
[26] M. Sznaier and M. J. Damborg, "Heuristically enhanced feedback control of constrained discrete-time linear systems," *Automatica*, vol. 26, pp. 521-532, 1990.
[27] P. O. M. Scokaert and J. B. Rawlings, "Constrained linear quadratic regulation," *IEEE Trans. on Automatic Control*, vol. 43, no. 8, pp. 1163-1169, 1998.
[28] E. Polak and T. H. Yang, "Moving horizon control of linear systems with input saturation and plant uncertainty, parts 1 and 2," *International Journal of Control*, vol. 58, no. 3, pp. 613-663, 1993.
[29] S. L. de Oliveira Kothare and M. Morari, "Contractive model predictive control for constrained nonlinear systems," *IEEE Trans. on Automatic Control*, vol. AC-45, pp. 1053-1071, 2000.
[30] P. O. M. Scokaert, J. B. Rawlings, and E. S. Meadows, "Discrete-time stability with perturbations: applications to model predictive control," *Automatica*, vol. 33, no. 3, pp. 463-470, 1997.
[31] E. Zafiriou, "On the closed-loop stability of constrained QDMC," *Proc. of American Control Conf.*, Boston, MA, pp. 2367-2372, 1991.
[32] E. Zafiriou, "On the effect of tuning parameters and constraints on the robustness of model predictive controllers," Y. Arkun, W. Ray (Eds.), *Proc. of 4th International Conference on Chemical Process Control - CPC IV*, South Padre Island, TX, pp. 363-393, 1991.
[33] J. H. Lee and Z. H. Yu, "Tuning of model predictive control for robust performance," *Comp. and Chem. Engr*., vol. 18, pp. 15-37, 1994.
[34] J. A. Primbs and V. Netistic, "A framework for robustness analysis of constrained finite receding horizon control," *Proc. of American Control Conf.*, pp. 2718-2722, 1998.
[35] P. J. Campo and M. Morari, "Robust model predictive control," *Proc. of the American Control Conference*, pp. 1021-1026, 1987.
[36] J. C. Allwright and G. C. Papavasiliou, "On linear programming and robust model-predictive control using impulse-responses," *Systems and Control Letters*, vol. 18, pp. 159-164, 1992.
[37] J. H. Lee and Z. Yu, "Worst-case formulation of model predictive control for systems with bounded parameters," *Automatica*, vol. 33, no. 5, pp. 763-781, 1997.
[38] P. O. M. Scokaert and D. Q. Mayne, "Min-max feedback model predictive control for constrained linear systems," *IEEE Trans. on Automatic Control*, vol. 48, no. 8, pp. 648-654, 1998.
[39] M. V. Kothare, V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, vol. 32, no. 10, pp. 1361-1379, 1996.
[40] Z. Wan and M. V. Kothare, "Efficient robust constrained model predictive control with a time varying terminal constraint set," *Systems and Control*

Letters, vol. 48, pp. 375-383, 2003.
[41] D. Mayne and H. Michalska, "Receding horizon control of nonlinear systems," *IEEE Trans. on Automatic Control*, vol. AC-35, pp. 814-824, 1990.
[42] H. Chen and F. Allgower, "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability," *Automatica*, vol. 14, pp. 1205-1217, 1998.
[43] S. J. Qin and T. A. Badgwell, "An overview of industrial model predictive control technology," *Proc. of Fifth Int. Conf. on Chemical Process Control*, Tahoe City, CA, aIChE Symposium Series, vol. 93, pp. 232-256, 1997.
[44] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407-427, 1999.
[45] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3-20, 2002.
[46] A. Bemporad, F. Borrelli, and M. Morari, "Piecewise linear optimal controllers for hybrid systems," *Proc. of American Control Conference*, Chicago, IL, pp. 1190-1194, 2000.
[47] T. A. Johansen and A. Grancharova, "Approximate explicit constrained linear model predictive control via orthogonal search tree," *IEEE Trans. on Automatic Control*, vol. 58, no. 5, pp. 810-815, 2003.
[48] P. Tondel, T. A. Johansen, and A. Bemporad, "An algorithm for multiparameteric quadratic programming and explicit MPC solution," *Automatica*, vol. 39, no. 5, pp. 489-497, 2003.
[49] P. Tondel, T. A. Johansen, and A. Bemporad, "Evaluation of piecewise affine control via binary search tree," *Automatica*, vol. 39, no. 5, pp. 945-950, 2003.
[50] A. Bemporad and C. Filippi, "Suboptimal explicit receding horizon control via approximate multiparametric quadratic programming," *Journal of Optimization Theory and Applications*, vol. 117, no. 1, pp. 9-38, 2003.
[51] T. A. Johansen, I. Petersen, and O. Slupphuag, "Explicit sub-optimal linear quadratic regulation with state and input constraints," *Automatica*, vol. 38, pp. 1099-1111, 2002.
[52] M. Canale, L. Fagiano, and M. Milanese, "Set membership approximation theory for fast implementation of model predictive control laws," *Automatica*, vol. 45, pp. 45-54, 2009.
[53] C. Jones and M. Morari, "Polytopic approximation of explicit model predictive controllers," *IEEE Trans. on Automatic Control*, vol. 55, no. 11, pp. 2542-2553, 2010.
[54] C. V. Rao, S. J. Wright, and J. B. Rawlings, "Application of interior-point methods to model predictive control," *Journal of Optimization Theory and Applications*, vol. 99, pp. 723-757, 1998.
[55] G. Pannocchia, J. B. Rawlings, and S. J. Wright, "Fast, large-scale model predictive control by partial enumeration," *Automatica*, vol. 43, pp. 852-860, 2006.
[56] H. J. Ferrau, H. G. Bock, and M. Diehl, "An on-line active set strategy for to overcome the limitations of explicit MPC," *International Journal of Robust and Nonlinear Control*, vol. 18, pp. 816-830, 2008.
[57] Y. Wang and S. Boyd, "Fast model predictive control using on-line optimization," *IEEE Trans. on Control System Technology*, vol. 18, pp. 267-278, 2010.
[58] F. Borrelli, A. Bemporad, M. Fodor, and D. Hrovat, "An MPC/hybrid system approach to traction control," *IEEE Trans. on Control System Technology*, vol. 14, pp. 541-552, 2006.
[59] N. Giorgetti, A. Bemporad, H. E. Tseng, and D. Hrovat, "Hybrid model predictive control application towards optimal semiactive suspension," *Proc. of IEEE Int. Symp. on Industrial Electronics*, Dubrovnik, Croatia, pp. 391-398, 2005.
[60] N. Giorgetti, G. Ripaccioli, A. Bemporad, I. V. Kolmanovsky, and D. Hrovat, "Hybrid model predictive control of direct injection stratified charge engines," *IEEE/ASME Trans. on Mechatronics*, vol. 11, no. 5, pp. 499-506, 2006.
[61] W. B. Dunbar, M. B. Milam, R. Franz, and R. M. Murray, "Model predictive control of a thrust-vectored flight control experiment," *Proc. of the 15th IFAC World Congress*, Barcelona, Spain, 2002.
[62] A. Lagerberg and B. Egardt, "Model predictive control of automotive powertrains with backlash," *Proc. of the 16th IFAC World Congress*, Prague, Czech Republic, 2005.
[63] B. Saerens, M. Diehl, J. Swevers, and E. Van den Bulck, "Model predictive control of automotive powertrains - first experimental results," *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, p. ThC18.2, 2008.
[64] S. Di Cairano, A. Bemporad, I. V. Kolmanovsky, and D. Hrovat, "Model predictive control of magnetically actuated mass spring dampers for automotive applications," *International Journal of Control*, vol. 80, no. 11, pp. 1701-1716, 2007.
[65] S. Richter, S. Mariethoz, and M. Morari, "High-speed on-line MPC based on a fast gradient method applied to power converter control," *Proc. of American Control Conference*, Baltimore, MD, pp. 4737-4743, 2010.
[66] S. Kouro, P. Cortes, R. Vargas, U. Ammann, and J. Rodriguez, "Model predictive control - a simple and powerful method to control power converters," *IEEE Trans. on Industrial Electronics*, vol. 56, no. 6, pp. 1826-1838, 2009.
[67] F. Zanini, C. N. Jones, D. Atienza, and G. De Micheli, "Multicare thermal management using approximate explicit model predictive control," *Proc. of IEEE ISCAS*, Paris, France, pp. 3321-3324, 2010.
[68] W. C. Wong and J. H. Lee, "Post-decision-state-based approximate dynamic programming for robust predictive control of constrained stochastic processes," *Industrial Engineering & Chemistry Research*, vol. 50, pp. 1389-1399, 2011.

Jay H. Lee obtained his B.S. degree in Chemical Engineering from the University of Washington, Seattle, in 1986, and his Ph.D. degree in Chemical Engineering from California Institute of Technology, Pasadena, in 1991. From 1991 to 1998, he was with the Department of Chemical Engineering at Auburn University, AL, as an Assistant Professor and an Associate Professor. From 1998-2000, he was with School of Chemical Engineering at Purdue University, West Lafayette and then with the School of Chemical Engineering at Georgia Institute of Technology, Atlanta. Starting this fall, he is the Head of the Chemical and Biomolecular Engineering Department at KAIST, Korea. He has held visiting appointments at E. I. Du Pont de Numours, Wilmington, in 1994 and at Seoul National University, Seoul, Korea, in 1997. He was a recipient of the National Science Foundation's Young Investigator Award in 1993 and was elected as an IEEE Fellow in 2010. He published over 120 manuscripts in SCI journals with more than 2500 ISI citations. His research interests are in the areas of system identification, state estimation, robust control, model predictive control and approximate dynamic programming.