CONSTRAINT HANDLING IN REDUCED ORDER MPC APPLICATION TO PAPER MACHINES

Yaman Arkun * Apostolos Rigopoulos **

* Koç University, College of Engineering, Sariyer, Istanbul, 34450, Turkey ** Weyerhaeuser Corp., Tacoma,WA 98477-2999, USA

Abstract: This paper presents a method to handle input constraints when a large scale system is to be controlled by a model predictive control algorithm which uses a reduced order model of the process under consideration. A paper machine is used throughout to motivate and illustrate the method.

1. INTRODUCTION

In control of large scale systems controller reduction is one method to handle the computational and implementational difficulties arising in a real-time environment. In their earlier work (Arkun and Kayihan, 1998; Rigopoulos, 1999) the authors have addressed the cross-directional (CD) control of paper machines which are equipped with large number of CD actuators. In order to reject disturbances certain transformations are computed to map a given large scale input-output system to a lower dimensional subspace which captures most of the process dynamics. Next reduced order controllers are designed in this lower dimensional space and the resulting controller inputs are transformed back to the original dimension and implemented on the real plant. Arkun and Kayihan (1998) has used reduced order unconstrained IMC as controller, whereas (Rigopoulos, 1999) has used a reduced order constrained model predictive controller. The goal of this paper is to show how the original constraints are preserved during the three steps of model reduction, reduced order MPC design and final implementation.

2. RATIONALE FOR A REDUCED ORDER CONTROLLER DESIGN. A MOTIVATING EXAMPLE: PAPER MACHINE

Paper machines are equipped with large number of CD actuators (slice lips on the headbox) and scanners provide measurements of the property of interest (e.g. thickness, basis weight) at many measurement points across the paper sheet. In addition control inputs can be tightly constrained due to the physical limitations of the actuators. In this work the full order system model used for CD control is given by (see (Rigopoulos, 1999)):

$$\mathbf{y}^{N}(k) = g(q^{-1})\mathbf{G}_{s}\mathbf{u}^{N}(k) + \mathbf{d}(k)$$
(1)

where $\mathbf{y}^{N}(k) \in \mathfrak{R}^{N}$ is the measured output (sheet) to be controlled at sampling time k; $g(q^{-1})$ accounts for CD dynamics; \mathbf{G}_{s} is the steady state CD actuator gain matrix; $\mathbf{u}^{N}(k) \in \mathfrak{R}^{N}$ is the vector of CD control elements; and $\mathbf{d}(k)$ is the disturbance affecting the property of interest. Here dimension N can be very large (several hundreds).

The idea of building a reduced order representation of the full system (1) originated from the reduced order modeling of the disturbances using the method of Karhunen Loeve Expansion (KLE). KLE generates a model with only a few degrees of freedom ($\mathbf{d}^{L}(k)$) that capture the most significant disturbance patterns (Rigopoulos *et al.*, 1997)

$$\mathbf{d}(k) = \mathbf{\Phi}^{L} \mathbf{t}^{L}(k) + \mathbf{\Phi}^{N-L} \mathbf{t}^{N-L}(k)$$
(2)

where $\Phi = \{\phi_1, \dots, \phi_L, \dots, \phi_N\}$ consists of the orthonormal basis vectors. They are in fact the eigenvectors of the covariance matrix of the random process $\mathbf{d}(k)$. They can be computed since $\mathbf{d}(k)$ is available through (1) (we assume that control inputs and outputs are measured). The vector \mathbf{t} is computed by projecting \mathbf{d} onto the set of basis functions, i.e.

$$\mathbf{t}^{L}(k) = (\mathbf{\Phi}^{L})^{T} \mathbf{d}(k)$$
(3)

The subspace order L directly identifies the amount of sheet variance that is captured by using only the L most significant modes. Here the interest is in the design of a constrained feedback controller that is capable of rejecting these L significant modes. The reduced order subspace in which controller design takes place has been derived in the following way (Rigopoulos, 1999) Start with the original system (1) and substitute for disturbance its KLE (1):

$$\mathbf{y}^{N}(k) = g(q^{-1})\mathbf{G}_{s}\mathbf{u}^{N}(k) + \mathbf{\Phi}^{L}\mathbf{t}^{L}(k) + \mathbf{\Phi}^{N-L}\mathbf{t}^{N-L}(k)$$
(4)

Perform an orthogonal projection from \Re^N to \Re^L by multiplying both sides by $(\Phi^L)^T$:

$$\mathbf{y}^{L}(k) = g(q^{-1})(\mathbf{\Phi}^{L})^{T}\mathbf{G}_{s}\mathbf{u}^{N}(k) + \mathbf{t}^{L}(k) \qquad (5)$$

Defining

$$\mathbf{u}^{L}(k) = (\mathbf{\Phi}^{L})^{T} \mathbf{G}_{s} \mathbf{u}^{N}(k)$$
(6)

the reduced order model for controller design becomes

$$\mathbf{y}^{L}(k) = g(q^{-1})\mathbf{u}^{L}(k) + \mathbf{t}^{L}(k)$$
(7)

Once the optimal solution $\mathbf{u}^{L^*}(k)$ is computed for (7) it needs to be projected to the full order system (1) so that it can be implemented on the real plant i.e

$$\mathbf{u}^{N}(k) = \tilde{\mathbf{A}}\mathbf{u}^{L}(k) \tag{8}$$

For an uncontrained minimum variance type controller Rigopoulos (1999) has shown that the following transformation is optimal:

$$\tilde{\mathbf{A}} = \mathbf{G}_{s}^{+} \boldsymbol{\Phi}^{L} \tag{9}$$

where \mathbf{G}_{s}^{+} is a generalized inverse of \mathbf{G}_{s} . Final feedback configuration is schematically shown in Figure 1.

In many applications disturbances may not be classified as stationary. In this case KLE and above transformations can still be applied using the most recent disturbance data; thus, they become time-dependent and we use subscript *k* to denote the time dependence of retained basis functions Φ_k^L .



Fig. 1. Feedback configuration.

The transformation of the full order system to a system that has, potentially, much fewer variables (transformed actuators) is done without explicitly considering the original actuator constraints. This is because, there is no way of a-*priori* knowing which of the original constraints would be active at the optimum, at every iteration. Finding the active set would amount to solving the full-order system. When the actuator constraints are mapped onto the reduced order space, the (transformed) feasible region may be empty. The proposed reduced order design is able to recover from this by splitting the problem in two steps:

- Step 1: The original N dimesional input/output system is transformed into an M_k -dimensional system with $L \leq M_k$ disturbance modes. The subscript 'k', which denotes sampling time, is included to explicitly show that the input/output dimensions of the transformed system are timevarying. The same transformation matrix that was used in the unconstrained case is also used here to map the actuator values of the reduced order system to the original full order system. Construct the mapping of constraints from the full to the reduced order system. Check for feasibility of that set, possibly by doing a phase-I simplex. If the set is feasible, go to step 2. Otherwise, increase M_k by one, and redo this step. Since the original full order system is considered to be always feasible, there will be an $M_k \leq N$ value for which the reduced order system will also be feasible.
- **Step 2:** Construct all other quantities necessary to form the objective function of the QP for the MPC, and solve the QP, *using the feasible solution of step 1 as the initial value.*

Since this method reduced the dimensions of both input and output spaces, the resulting system is (potentially much) smaller in size than the original one, hence the reduced memory requirements. It is also faster, because although the number of constraints stays the same in the reduced order system, the number of decision variables (transformed actuators) has (considerably) decreased, thus, it takes less time to compute the active set, and thus the optimum solution.

2.1 Actuator Constraints and their Impact on the Reduced Order Controller Design

There are three types of actuator constraints that are usually encountered in the production of paper and other sheet forming processes:

• Lower and upper bound constraints

$$\mathbf{u}_{\min} \le \mathbf{u}^N(k) \le \mathbf{u}_{\max} \tag{10}$$

where usually, because of the problem geometry and that the actuators are expressed in deviation form, $\mathbf{u}_{\text{max}} = -\mathbf{u}_{\text{min}} \ge \mathbf{0}$, and $\mathbf{u}_{\text{max}} = \mathbf{I}_N u_{\text{max}}$.

Adjacent actuator constraints

$$\mathbf{m}_{\min} \le D \mathbf{u}^{N}(k) \le \mathbf{m}_{\max} \tag{11}$$

where $D \in \Re^{N \times N}$, and for the same reasons as above, $\mathbf{m}_{\text{max}} = -\mathbf{m}_{\text{min}} \ge \mathbf{0}$, with $\mathbf{m}_{\text{max}} = \mathbf{I}_N m_{\text{max}}$. In paper machines this constraint effectively penalizes the bending stress of the slice lip.

Rate constraints

$$\Delta \mathbf{u}^{N}(k) | \le \Delta \mathbf{u}_{\max} \tag{12}$$

where, $\Delta \mathbf{u}^{N}(k) = \mathbf{u}^{N}(k) - \mathbf{u}^{N}(k-1)$. Again, it is common to have $\Delta \mathbf{u}_{max} = \mathbf{I}_{N}\Delta u_{max}$. This constraint is imposed in order to avoid drastic changes in the magnitude of each actuator *within two consecutive time periods*, which can lead to excessive wear and tear of the actuator hardware.

It is assumed that the above set of inequalities is *always consistent*, i.e. for the operating conditions for which the system was designed, there always exists a feasible solution vector $\mathbf{u}^{N}(k)$. The situation is different in the case of the reduced order constrained controller design. In particular, the above set of inequalities in the transformed domain becomes

$$\mathbf{u}_{\min} \leq \mathbf{\hat{A}}_k \mathbf{u}^L(k) \leq \mathbf{u}_{\max} \tag{13}$$

$$\mathbf{m}_{\min} \leq D\tilde{\mathbf{A}}_k \mathbf{u}^L(k) \leq \mathbf{m}_{\max} \tag{14}$$

$$-\Delta \mathbf{u}_{\max} + \tilde{\mathbf{A}}_{k-1} \mathbf{u}^{L}(k-1) \leq \tilde{\mathbf{A}}_{k} \mathbf{u}^{L}(k)$$

$$\leq \Delta \mathbf{u}_{\max} + \mathbf{A}_{k-1} \mathbf{u}^{L} (k-1) \tag{15}$$

Thus the *number of constraints has stayed the same*, but the *number of decision variables was reduced from N to L*. This point is very important because, assuming that the objective function is quadratic, the computation time for solving a QP problem will depend heavily on the number of variables, because the latter sets the upper bound on the number of constraints that may be active at the optimum. Finding the active set is one of the most time consuming operations, especially in the presence of tight constraints. This is why the reduced order controller design with a low ratio $\alpha = \frac{L}{N}$ becomes so appealing.



Fig. 2. Contour plot of the objective function arising from a controller with no dynamics $(g(q^{-1}) = 1)$, N = 2, and L = 1.

On the other hand, because transformation matrix $\tilde{\mathbf{A}}$ was constructed based on the unconstrained case, *some of the above inequalities may become infeasible.*

Even in the case where no rate constraints are imposed, the closed loop performance obtained from application of the reduced order controller may be unacceptable. This arises again from the fact that the range of $\tilde{\mathbf{A}}_k$ forces $\mathbf{u}^L(k)$ to lie in that region of \Re^N where significant disturbances also lie. But in the presence of constraints this locus may be far from the full order constrained minimum. Fig. 2 shows the problem for a simple example with no dynamics (g = 1), N = 2, L = 1 and a quadratic objective function. The feasible region is the rectangle $S = \{(x, y), -1 \le x \le 1, -1$ $y \leq 1$. The unconstrained minimum denoted by $\mathbf{u}_{unc}^{N\star}$ is located at $[3 - 1]^T$, and it is the same for the full and for the reduced controller designs. In the presence of constraints, however, the optimal full order minimum $\mathbf{u}^{N\star} = [1 - 1]^T$ with an objective function minimum of -9. On the other hand, the locus of $\tilde{\mathbf{A}}\mathbf{u}^L$ is the straight line that passes through points $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ and $\mathbf{u}_{unc}^{N\star}$. The constrained optimum arising from the reduced order controller is at $\mathbf{u}_{\star}^{N} = \begin{bmatrix} 1 & -\frac{1}{3} \end{bmatrix}^{T}$ and is equal to -7.22, which corresponds to a 20% drop in performance over the full order case that may be deemed unacceptable.

A modification to the reduced order design will be described now that addresses any feasibility issues and allows for improvement of the closed loop performance. The key idea is to *allow the actuator subspace order to be different from the disturbance subspace order.* Thus the reduced order system will have *L* disturbances, and $M_k \ge L$ inputs and outputs. Following an approach very similar to the one used for the derivation of Eq. (7) one gets:

$$\mathbf{y}^{M_k}(k) = g(q^{-1})\mathbf{u}^{M_k}(k) + \mathbf{R}_k^T \mathbf{t}^L(k)$$
(16)

where $\mathbf{R}_k = [\mathbf{I}_L \ \mathbf{0}_{L \times (M_k - L)}].$

With regard to the above system the following observations are in order

 Obtaining u^N_{*}(k) from u<sup>M_k*(k) presents the same problems as before. Again, a linear relation is postulated
</sup>

$$\mathbf{u}^{N}(k) = \tilde{\mathbf{A}}_{k} \mathbf{u}^{M_{k}}(k) \tag{17}$$

where $\tilde{\mathbf{A}}_k \in \Re^{N \times M_k}$. Eq. 16 can be written more explicitly as

$$\begin{bmatrix} \mathbf{y}^{L}(k) \\ \mathbf{y}^{M_{k}-L}(k) \end{bmatrix} = g(q^{-1}) \begin{bmatrix} \mathbf{u}^{L}(k) \\ \mathbf{u}^{M_{k}-L}(k) \end{bmatrix} + \begin{bmatrix} \mathbf{I}_{L} \\ \mathbf{0}_{(M_{k}-L)\times L} \end{bmatrix} \mathbf{t}^{L}(k) \quad (18)$$

where the top block is completely separate from the bottom and identical to the *L*-dimensional system. Thus, one can apply the results of that design directly to get A^L_k = G⁺_k Φ^L_k, where A^L_k = [A^L_k A^{M_k-L}].
A^{M_k-L} is obtained by considering the con-

Â^{M_k-L}_k is obtained by considering the constrained minimization problem. In particular, the inclusion of additional inputs u^{M_k-L}(k) is done so as to increase the number of elements of set S_u ⊂ ℜ^N where u^N_k(k) belongs. This increase is maximized by appropriately selecting Ã^{M_k-L}_k to maximize the number of elements that belong to the range of Ã_k. This is accomplished by designing Ã^{M_k-L}_k to be orthogonal to Ã^L_k = G⁺_sΦ^L_k, which is guaranteed by setting Ã^{M_k-L}_k = G⁺_sΦ^{M_k-L}_k, since Φ^{M_k-L}_k is orthogonal to Φ^L_k. Finally,

$$\tilde{\mathbf{A}}_k = \mathbf{G}_s^+ \mathbf{\Phi}_k^{M_k} \tag{19}$$

2.2 Implementation of Constrained Control through State-Space Model Predictive Control

In the present work state space MPC as detailed in (Ricker, 1992) is used. Here we will present only the important features which are unique to our problem setting. The control algorithm is based on the reduced order system given by Eq. 16:

$$\mathbf{y}^{M_k}(k) = g(q^{-1})\mathbf{u}^{M_k}(k) + \mathbf{R}_k^T \mathbf{t}^L(k)$$
$$= \mathbf{y}_u^{M_k}(k) + \mathbf{y}_d^{M_k}(k)$$
(20)

State-space transformation of $\mathbf{y}_d^{M_k}(k)$ **:** Assuming that the disturbance subspace *L* has been selected, $\mathbf{y}_d^{M_k}(k)$ is converted to state-space (Rigopoulos, 1999)

$$\mathbf{x}_d(k+1) = \mathbf{A}_d(k)\mathbf{x}_d(k) + \mathbf{\Gamma}_d(k+1)\mathbf{e}(k+1)$$
(21)

$$\mathbf{y}_d^{M_k}(k) = \mathbf{R}_k^T \tilde{\mathbf{C}}_d(k) \mathbf{x}_d(k)$$
(22)

where $\mathbf{A}_d(k)$, and the size of $\Gamma_d(k)$ depend on the AutoRegressive (AR) modeling of $\mathbf{t}^L(k)$; and $\tilde{\mathbf{C}}_d(k) = (\mathbf{\Phi}_k^L)^T \mathbf{C}_d(k)$. The order of the AR model selected remains constant throughout the simulation. Thus, although the size of $\mathbf{y}_d^{M_k}(k)$ is determined by the size of

 \mathbf{R}_k which depends on M_k , the sizes of the vectors and matrices of state equation (21) remain unaffected. As a consequence, varying M_k presents no problem to the state-space modeling of $\mathbf{y}_d^{M_k}(k)$.

State-space transformation of $\mathbf{y}_{u}^{M_{k}}(k)$ **:** Because $\mathbf{y}_{u}^{M_{k}}(k) = g(q^{-1})\mathbf{u}^{M_{k}}(k)$ is a *decoupled* system, its state-space equivalent description will be in terms of *block-diagonal* matrices:

$$\mathbf{x}_{u}^{(nM_{k+1})}(k+1) = \mathbf{A}_{u}^{(\bar{n}M_{k})}\mathbf{x}_{u}^{(nM_{k})}(k) + \mathbf{B}_{u}^{M_{k}}\mathbf{u}^{M_{k}}(k)$$
(23)
$$\mathbf{y}_{u}^{M_{k}}(k) = \mathbf{C}_{u}^{M_{k}}\mathbf{x}_{u}^{(nM_{k})}(k)$$
(24)

where, for example $\mathbf{A}_{u}^{(\bar{n}M_{k})} = \text{diag}[\mathbf{A}^{\bar{n}}\cdots\mathbf{A}^{\bar{n}}]$; each $\mathbf{A}^{\bar{n}} \in \mathfrak{R}^{n \times n}$, with rank $[\mathbf{A}^{\bar{n}}] = \bar{n}$ contains the necessary states for the modeling of each $y_{u,i}^{M_{k}}(k) = g(q^{-1})u_{i}^{M_{k}}(k)$.

Now, suppose that at iteration k + 1 the feasibility of constraints imposes $M_{k+1} = M_k + 1$ leading to a unit increase of inputs $\mathbf{u}^{M_{k+1}} = \begin{bmatrix} \mathbf{u}^{M_k} \\ u^{new} \end{bmatrix}$ at time k + 1. However, because of the block diagonal nature of all the state-space matrices involved, the states appropriated to the modeling of the first M_k elements of vector $\mathbf{u}^{M_{k+1}}(k+1)$ will not be influenced by the new states that must be introduced for the additional input. This observation is important because it shows that when the state order increases from nM_k to $nM_k + n$ due to the introduction of additional input u^{new} , only the new states $nM_k + 1$ to nM_k need to be initialized (to zero), before the state equation can be used to compute $\mathbf{x}_u^{(nM_{k+2})}(k+2)$. The first nM_k states evolve normally.

2.3 Computational Issues and Efficiency

Consider a paper machine with N = 100 CD actuators where the full order constrained controller is to be applied with p (prediction horizon) $= m_h$ (move horizon) = 4 and all three types of constraints are present. Then, an optimization problem with 400 variables and 2,400 constraints would have to be solved *at every sampling time*. Even when a reduced order controller were to be implemented with $M_k = 30$, the task would not be trivial. Therefore, special attention needs to be put on the selection of the most suitable optimization algorithm.

In selecting the most appropriate solver one must, consider the problem that arises from a potential infeasibility due to the dimensionality reduction. The solver should be able to identify this problem as quickly as possible and compensate for it by gradually increasing the actuator subspace order M_k in order to obtain a feasible region. In this "internal" loop one only needs to update matrix \hat{A} and increase the size of $\Delta U(k)$ before re-checking for feasibility. Updating all other quantities including the objective function should be done only once, *after* feasibility has been ensured.

Of course, should the user decide to reduce M_k in an attempt to speed up the computations, the same type of feasibility test must be made prior to accepting a lower M_k value. Evidently, only a primal active set method works along those lines, because of its inherent need for the *a-priori* calculation of a feasible point. According to this strategy, a phase I simplex is performed repetitively increasing M_k by one until an initial feasible point (thus a feasible region) is obtained. Then, after all appropriate matrices are updated, a search for the optimum is initiated. This procedure may become even more efficient by using an interior point (IP) method to obtain the solution to the phase I simplex problem, instead of using a standard Dantzig-type LP solver.

Another point that can have significant impact on the overall performance of the QP solver is "hot" starts. In the presence of relatively tight input constraints, the optimal solution vector does not vary significantly from one sampling time to the next. Thus, in addition to ensuring that an initial feasible point is available, one can further benefit from providing an initial point that is relatively close to the optimum.

3. EXAMPLES

A paper machine with N = 200 CD actuators and measurements positions is considered, with $g(q^{-1}) = \frac{q^{-1}}{1-0.2q^{-1}}$, and

$$\mathbf{G}_s = \text{Toeplitz}[2.0\ 0.8\ -1.0\ -0.8\ -0.6\ -0.4\ 0\ \cdots\ 0]$$

The full disturbance $\mathbf{d}(k)$ consists of 200 CD and 300 MD positions. The last 100 full MD s are illustrated in Fig. 3. For the modeling equations used to create this the reader is referred to (Rigopoulos *et al.*, 1997). Using the first 200 full data, KLE indicated that only L = 3 modes were necessary to capture the significant disturbance patterns. Also, an AR(2) model was sufficient for the modeling of temporal vectors $\mathbf{t}^{L}(k)$.



Table 1. Description of the sets of constraints considered in the example.

Constraints Set	$u_{\rm max} = -u_{\rm min}$	$m_{\rm max} = -m_{\rm min}$	$\Delta u_{\rm max}$
set 1 — tight	0.15	0.1	0.005
set 2 — moderate	0.15	0.1	∞
set 3 — light	0.25	~	~

The parameters of the AR model and basis functions Φ_k^L were also updated at every sampling time. Table 1 gives the values of the various types of constraints used.

For constraints sets 2 and 3 M_k remained constant throughout the simulation, as no feasibility issues were encountered. Feasibility problems were encountered, however, when rate constraints were imposed (set 1), and M_k had to be increased in order to be able to compute a solution vector $\mathbf{u}^{M_k}(k)$. Also, for each simulation the time allocated to solving the QP was recorded, as well as the closed loop.

Fig. 4(a) illustrates the standard deviation of the CD profile as a function of the actuator subspace order $M_k = 10$, 30 and 150 for constraint set 3. The open loop standard deviation is also included for comparison. Clearly, the controller was able to reduce the CD variability significantly even when only 10 transformed manipulated variables $\mathbf{u}^{M_k}(k)$ were used. What is more important, however, is the very little improvement achieved by the five-fold increase to the subspace order (from 30 to 150) that shows the efficiency of the reduced order controller design.

Fig. 5(a) and (b) show the number of active constraints at the optimum when the simulation was started with initial subspace order $M_1 = 10$ and $M_1 = 30$, respectively. An increase in the number of active constraints indicates an increase in M_k because an infeasible constraint region was encountered. Notice that in almost all sampling times the number of active constraints is equal to the number of degrees of freedom, illustrating how tight the constraints are.

Finally the trade-off between closed loop performance and computation time is illustrated in Fig. 6 by comparing the drop in the sum-of-square (SSE) errors vs. the increase in computation time. For example, in the case of tight constraints, increasing M_k from 50 to 100 would result to a less than 1% performance improvement, while requiring 340% more computation time! This clearly shows the significant computational advantages associated with the implementation of the reduced order controller design.

4. CONCLUSIONS

We have presented a method to handle input constraints when an MPC algorithm is derived based on a reduced order approximation of the process but implemented on the real process. A paper machine example is used to demonstrate the method for CD control.





Fig. 4. Standard deviation as a function of initial subspace order for (a) constraints set 3 and (b) constraints set 1.

REFERENCES

- Arkun, Yaman and Ferhan Kayihan (1998). A novel approach to full CD profile control of sheetforming processes using adaptive PCA and reduced-order IMC design. *Computers Chem. Engng* 22(7–8), 945–962.
- Ricker, N. Lawrence (1992). Model-predictive control: State of the art. In: *Proceedings of the CPC-IV Conference*. Padre Island, Texas. pp. 271–296.
- Rigopoulos, A. (1999). Application of principal component analysis in the identification and control of sheet-forming processes. *PhD Thesis, Georgia Institute of Technology, Atlanta, GA*.
- Rigopoulos, Apostolos, Yaman Arkun and Ferhan Kayihan (1997). Identification of full profile disturbance models for sheet forming processes. *AIChE J.* **43**(3), 727–739.



Fig. 5. Number of active constraints for simulations with initial actuator subspace order (a) $M_1 = 10$ and (b) $M_1 = 30$, using constraints set 1.



Fig. 6. Trade-off between performance as expressed by the sum-of-square improvement with increasing initial M_k shown in (a), and computation time increase as a function of initial M_k shown in (b).