49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Independent vs. Joint Estimation in Multi-Agent Iterative Learning Control

Angela Schöllig, Javier Alonso-Mora, and Raffaello D'Andrea

Abstract— **This paper studies iterative learning control (ILC) in a multi-agent framework, wherein a group of agents simultaneously and repeatedly perform the same task. The agents improve their performance by using the knowledge gained from previous executions. Assuming similarity between the agents, we investigate whether exchanging information between the agents improves an individual's learning performance. That is, does an individual agent benefit from the experience of the other agents? We consider the multi-agent iterative learning problem as a two-step process of: first, estimating the repetitive disturbance of each agent; and second, correcting for it. We present a comparison of an agent's disturbance estimate in the case of (I) independent estimation, where each agent has access only to its own measurement, and (II) joint estimation, where information of all agents is globally accessible. We analytically derive an upper bound of the performance improvement due to joint estimation. Results are obtained for two limiting cases: (i) pure process noise, and (ii) pure measurement noise. The benefits of information sharing are negligible in (i). For (ii), a performance improvement is observed when a high similarity between the agents is guaranteed.**

I. Introduction

Exploiting previous experience when repeatedly executing the same task is a natural way to improve future performance in the presence of repetitive, unmodeled disturbances. Iterative learning control (ILC), as first proposed in [1], achieves precise tracking behavior by effectively incorporating past control information (such as applied input signals and measured outputs) when calculating the feedforward control action used in the next iteration, cf. [2], [3]. One way of viewing ILC is as a two-step process of estimation and control: first identifying the unknown repetitive disturbance and later compensating for it [4]–[8]. LQG-type solutions have been proposed in [9]–[11], which estimate the tracking error and, based on this result, calculate a new input trajectory by minimizing a quadratic cost function.

While ILC has proven to be successful in a variety of industrial applications (including chemical process control, rotary systems and robotics), we have yet to identify if – and how – ILC schemes can be generalized when facing homogeneous groups of agents or assemblies of similar units (for example, robot arms in an industrial environment, or a fleet of mobile robots in a warehouse [12], [13]). In other words, how can we cope with uncertainties in a multi-agent framework? Is there a benefit of exchanging information between the agents? What kind of information sharing makes sense? Cooperative iterative learning schemes were previously proposed in [14]. Recently, ILC was applied to multi-agent systems, cf. [15], with the goal of achieving formation control. While it has been established that the joint performance of all agents is fundamental to the formation problem, this paper focuses on the potential for individual agents to improve their performance when conducting a task alongside a group of similar agents conducting the same task. Analogous questions were previously studied in the context of reinforcement learning, see [16].

The authors are with the Institute for Dynamic Systems and Control, ETH Zurich, 8092 Zurich, Switzerland. (schoellig, dandrea)@idsc.mavt.ethz.ch

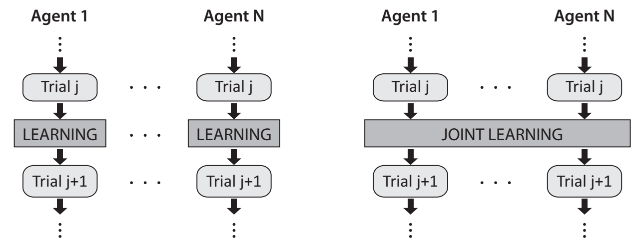

Fig. 1. Simultaneous run of all agents: independent (left) vs. joint (right) estimation and learning.

The results of our research show that the passing of information between agents has limited benefit for a large class of problems. This conclusion is based upon a comparison of independent learning with a cooperative scheme, where information of all agents is globally accessible to every agent, see Fig. I. Similarity between the agents is assured by assuming that they have the same nominal dynamics and share a common iteration-independent disturbance; however, they differ in an additional individual error component that is also constant across iterations. We introduce iteration-dependent noise terms that account for measurement and process noise, and obtain results for two limit cases: (i) pure process noise, and (ii) pure measurement noise. The benefits of information sharing are negligible in (i). For (ii), we observe a greater improvement in performance when a high similarity between the agents is guaranteed. In short: Individual agents in an ILC framework do not, in most cases, benefit significantly from information sharing when simultaneously learning the same task.

The paper is organized as follows: Section II formalizes the multi-agent iterative control problem and reduces it to a comparison of independent versus joint estimation. Section III compares both scenarios and presents the core result of the paper in terms of an upper bound on the performance improvement due to joint estimation. The work is summarized in Section IV, whereas proofs are partly presented in the Appendix.

978-1-4244-7744-9/10/$26.00 ©2010 IEEE

II. Problem Statement

A. Motivation

We begin by considering a group of N agents that simultaneously and repeatedly perform the same task. A common way of describing an agent's dynamics during a single run is the so-called lifted system representation [17]–[19]. For each agent $i \in \mathcal{I} = \{1, 2, \ldots, N\}$, the input-state relationship is modeled by a static matrix equation,

$$x^i = F^i u^i + d^i \,, \tag{1}$$

which maps a given discrete-time input signal $u^i = \left[\, u^i(0),\, u^i(1),\, \ldots,\, u^i(T) \,\right]^{\mathrm{T}} \in \mathbb{R}^{(T+1)n_u}$ to the corresponding lifted states $x^i \in \mathbb{R}^{(T+1)n_x}$. In this context, $(T+1)$ samples represent a single iteration and n_u and n_x denote the dimension of the input and state, respectively. The vectors x^i and u^i are defined as the deviation from the desired task trajectory and the corresponding nominal input, see for example [8]. The vector d^i represents an exogenous disturbance *constant across iterations*, which captures model errors along the trajectory as well as repeating disturbances and nonzero initial conditions [3], [20], [21]. We include a *trial-uncorrelated* noise signal ξ^i in model (1) to account for process noise, which varies from trial to trial. Introducing the iteration index $j \in \{1, 2, \ldots\}$, the state in the jth trial is given by

$$x^i_j = F^i u^i_j + d^i + \xi^i_j \,. \tag{2}$$

where ξ^i_j is assumed to be zero-mean Gaussian white noise. The vector d^i is viewed as an agent-dependent, normally-distributed random signal.

The agents' output y^i_j is corrupted by measurement noise and similarly represented in the lifted domain,

$$y^i_j = G^i x^i_j + \mu^i_j \,, \tag{3}$$

where μ^i_j is zero-mean Gaussian white noise.

Note that (2) and (3) might be the result of linearizing the agent dynamics about a desired task trajectory. Refer to [8], [22] for a more detailed derivation.

In the above context, the goal of the iterative learning algorithm is to make the state x^i_j (that is, the deviation from the desired task trajectory) small or, more precisely, to reduce x^i_j with an increasing number of iterations j. The performance of each individual agent is gradually improved by taking into account all information on previous iterations when estimating the disturbance vector d^i. As the accuracy of the disturbance estimate increases, a more appropriate open-loop input is determined, thereby compensating for the deficiencies in the modeled dynamics represented by the matrix F^i. From x^i_j conclusions can be drawn as to the performance of execution j.

We now consider a homogeneous fleet of agents with the same nominal dynamics:

$$\begin{aligned} F^i &= F \\ G^i &= I \qquad \forall\, i \in \mathcal{I} \,, \end{aligned} \tag{4}$$

where I denotes the identity matrix. That is, the state is assumed to be measured directly. Differences between the agents are captured in the disturbance vector d^i, which is composed of a common part d^0 identical for all agents, and an individual part $d^{i,ind}$,

$$d^i = d^0 + d^{i,ind} \quad \forall\, i \in \mathcal{I} \,. \tag{5}$$

In this context, the question arises: Does an individual agent benefit from sharing information with its companions? To what degree can the disturbance estimate d^i be improved by taking into account the measurements of the other agents?

B. Simplified Model

Our main objective and central problem is to identify the disturbance d^i for each agent i in the presence of both process and measurement noise. Based on the disturbance estimate, a correcting input u^i_j can be found that best compensates for the repetitive disturbance using a problem-specific optimization criterion, see for example [8]. Importantly, the correcting input u^i_j applied in each iteration is known. Focusing on the estimation problem, we consider a condensed form of the above multi-agent system representation (2)-(3),

$$x^i_j = d^i + \xi^i_j \tag{6}$$

$$y^i_j = x^i_j + \mu^i_j \,, \tag{7}$$

which features the key noise and disturbance characteristics, but omits the known part Fu^i_j, without loss of generality. Equations (6) and (7) are summarized by

$$y^i_j = d^i + \upsilon^i_j \,, \tag{8}$$

where $\upsilon^i_j = \xi^i_j + \mu^i_j$ captures both process and measurement noise.

Moreover, assuming independence of the single entries in the vectors d^i and υ^i_j and identical noise characteristics, the problem reduces to the scalar case,

$$y^i_j = d^0 + d^{i,ind} + \upsilon^i_j \,, \tag{9}$$

where all variables are scalar-valued. The probability distributions are given by

$$\begin{aligned} d^0 \quad &\sim\ \mathcal{N}(0,\, \alpha) \\ d^{i,ind} &\sim\ \mathcal{N}(0,\, \beta) \\ \upsilon^i_j \quad &\sim\ \mathcal{N}(0,\, 1) \,, \qquad \alpha, \beta \geq 0 \,, \end{aligned} \tag{10}$$

and υ^i_j, d^0, and $d^{i,ind}$ are assumed to be mutually independent for all $i \in \mathcal{I}$ and $j \in \{1,\, 2,\, \ldots\}$. The notation $\mathcal{N}(0,\, \alpha)$ represents a normal distribution with mean 0 and variance α. Note that in (10), the variance of the individual disturbance $d^{i,ind}$ is assumed to be identical for all agents $i \in \mathcal{I}$. Without loss of generality, the variances are normalized such that υ^i_j is 1.

C. Independent vs. Joint Estimation

As the number of trials and measurements increases, more information about the system is collected, allowing an increasingly accurate estimate of the agents' constant noise terms d^i, $i \in \mathcal{I}$. Two limiting approaches might be taken when solving the estimation problem: (I) independent estimation, and (II) joint estimation.

In the case of independent estimation (I), each agent i individually estimates its disturbance d^i taking only its own measurements $y_j^i,\ j \in \{1, 2, \dots\}$ into account.

In the joint case (II), the acquired measurement data is fully exchanged between all agents. Based on this global knowledge, we can design a joint estimation scheme that exploits the measurements of all agents and provides estimates d^i for every agent $i \in \mathcal{I}$. A vector D, reflecting the estimation objective in this case, is defined as: $D = [\, d^0,\, d^1,\, \dots,\, d^N \,]^{\mathrm{T}} \in \mathbb{R}^{(N+1)}$. The measurements of all agents in the jth trial are combined in $Y_j = [\, y_j^1,\, y_j^2,\, \dots,\, y_j^N \,]^{\mathrm{T}}$, and analogously, the noise vector $V_j = [\, \upsilon_j^1,\, \upsilon_j^2,\, \dots,\, \upsilon_j^N \,]^{\mathrm{T}}$ is introduced. Based on this representation, the joint estimation problem can be formulated as a Kalman filter problem, cf. [23], [24]:

$$\begin{aligned} D_j &= D_{j-1} \qquad\qquad \forall\, j \geq 1 \\ Y_j &= H\, D_j + V_j\,, \end{aligned} \tag{11}$$

where $H = [\, \mathbf{0},\, I \,]$ is a matrix with zeros in the first column concatenated with an identity matrix of appropriate dimensions. The Kalman filter returns an unbiased state estimate $\widehat{D}_j$ for $j \geq 1$ that minimizes the error covariance matrix

$$P_j = E\left[(D_j - \widehat{D}_j)(D_j - \widehat{D}_j)^{\mathrm{T}} \right], \tag{12}$$

of trial j, taking measurements $Y_m,\ 1 \leq m \leq j$, into account. $E\,[\cdot]$ denotes the expected value. The initial values are obtained from (10); in particular,

$$\widehat{D}_0 = [\, 0,\, 0,\, \dots,\, 0 \,]^T \tag{13}$$

and the initial covariance matrix $P_0 = [\, p_0^{(k,l)} \,],\ k, l \in \mathcal{K} = \{0,\, 1,\, \dots,\, N\}$, is

$$P_0 = E\left[\, D_0 D_0^{\mathrm{T}} \,\right] \tag{14}$$

with

$$p_0^{(k,l)} = E\left[\, d^k d^l \,\right] = E\left[\, \left(d^0 + d^{k,ind}\right)\left(d^0 + d^{l,ind}\right) \,\right],$$

where $d^{0,ind} = 0$. Recalling (10) and the mutual independence of d^0 and $d^{i,ind}$ for all $i \in \mathcal{I}$, the initial covariance is given by

$$p_0^{(k,l)} = \begin{cases} \alpha + \beta & \text{for} \quad k = l \geq 1 \\ \quad \alpha & \text{otherwise}\,. \end{cases} \tag{15}$$

Note that the above derivations do not place further assumptions or restrictions on how information is shared between agents: the information y_j^i of each agent is available to every other agent. In other words, we are investigating the ideal case of centralized, joint estimation within an optimal filtering context.

Equally important, the independent estimation problem (I) is just a special case of the cooperative framework (II) with $N = 1$.

In both cases, (I) and (II), the variance of an individual's disturbance estimate at iteration j is given by

$$E\left[(d^i - \widehat{d}_j^{\,i})^2 \right] = p_j^{(i,i)} = p_j^{(1,1)}, \quad \forall\, i \in \mathcal{I}, \tag{16}$$

where $\widehat{D}_j = [\, \widehat{d}_j^{\,i} \,],\ i \in \mathcal{I}$, and $P_j = [\, p_j^{(k,l)} \,],\ k, l \in \mathcal{K}$. The variance is identical for all agents, since for each agent the same assumptions on the dynamics (9) and the initial noise characteristics (10) are made. The variance of an individual's disturbance (16) indicates the quality of the disturbance estimate. In the general case, (2)-(3), this value influences the effectiveness of the disturbance compensation, since the input update rule of the ILC algorithm is based on the current estimate $\widehat{d}_j^{\,i}$; for example by a relation as follows, see [8]:

$$u_{j+1}^i = \arg\min_u \ ||\, F^i u + \widehat{d}_j^{\,i} \,||\,. \tag{17}$$

Below, we distinguish between the individual disturbance variance $p_j^{(1,1)}$ in case of joint and independent estimation, where the latter is given when evaluating $p_j^{(1,1)}$ for $N = 1$, i.e.

$$p_j^{(1,1)} \Big|_{N=1}\,. \tag{18}$$

Thus, the initial question can be reformulated: To what degree does joint estimation benefit the individual learning of an agent?

III. Result

We compared the learning performance based on (I) independent and (II) joint estimation, via the variance of the state x_j^i given all past measurements. This value indicates the accuracy of the tracking behavior in each iteration j. We investigated the benefits of information sharing and used, as our basis for the investigation, two limiting cases of (8): (i) encountering *pure process noise*, and (ii) dealing with *measurement noise only*. From these benchmark examples, we were able to deduce properties for the general mixed-noise case and draw conclusions about the advantages of passing information in an ILC framework.

In order to compare the independent estimation result (I) with the joint estimation result (II), we derived an analytical expression for $p_j^{(1,1)}$.

Proposition 1. The error variance of an agent's disturbance $p_j^{(1,1)}$ can be expressed in terms of the initial variances α and β, the number of agents N, and the iteration j,

$$p_j^{(1,1)} = \frac{\alpha + \beta + j\beta^2 + jN\alpha\beta}{(1 + j\beta)(1 + j\beta + jN\alpha)}\,. \tag{19}$$

The result is obtained by solving the Kalman filter equations for (11) with initial conditions (13) and (15). ∎

A detailed proof is found in the Appendix.

Next, we use the relation (19) to derive an upper bound on the performance improvement due to joint estimation. Two limiting cases are considered: (i) pure process noise and (ii) pure measurement noise.

A. Pure Process Noise

Perfect measurements are assumed, i.e. $\mu_j^i = 0$ in (7) and υ_j^i is interpreted as pure process noise, $\upsilon_j^i = \xi_j^i$. The performance of independent (I) vs. joint (II) estimation is analyzed through the variance of the state estimate. As mentioned in Section II-A, the goal of ILC is to reduce the value x_j^i. This is most easily achieved if the variance in the estimate of x_j^i is small. That is, the variance of the state estimate can be used as a measure of learning performance. Given (6) and (10), the best estimate of the state $\widehat{x}_j^i$ at iteration j is equal to the current disturbance estimate $\widehat{d}_j^i$,

$$\widehat{x}_j^i = \widehat{d}_j^i, \tag{20}$$

since the noise ξ_j^i has zero mean. Recalling the noise characteristics (10) and the previous assumption of mutual independence between d^i and υ_j^i, we obtain the variance of state estimate from the sum of the variance of the estimate $\widehat{d}_j^i$ and the variance of $\xi_j^i = \upsilon_j^i$. That is, with (16),

$$\begin{aligned} E\Big[(x_j^i - \widehat{x}_j^i)^2\Big] &= E\Big[(d^i + \xi_j^i - \widehat{d}_j^i)^2\Big] \\ &= p_j^{(1,1)} + 1\,. \end{aligned} \tag{21}$$

We introduce the performance index (for the pure process noise case) as the ratio of the state variance in the independent case vs. the joint case,

$$R^{\text{proc}} = \frac{p_j^{(1,1)}\big|_{N=1} + 1}{p_j^{(1,1)} \qquad + 1}\,, \tag{22}$$

using the notation of (18).

The following theorem can be stated:

Theorem 1. The bounds on the performance improvement due to joint estimation (vs. independent estimation) are given by

$$1 \,\le\, R^{\text{proc}} \,\le\, \frac{1+j}{j} \qquad \forall\, \alpha,\, \beta,\, N,\, j\,, \tag{23}$$

where the best performance improvement occurs when $N \to \infty$, $\alpha \to \infty$, and $\beta = 0$. In this case, $R^{\text{proc}} = (1+j)/j$. ■

Interpretation of the result:

– The performance improvement due to joint estimation has an upper bound which is valid for all possible combinations of α, β, N, and j.
– Joint estimation is most beneficial if the agents' common disturbance component dominates and the individual noise component is negligible compared to the process noise; this corresponds to a large common noise variance α and a small individual component $\beta \ll 1$.
– The largest possible improvement in performance is a factor of 2, which is obtained only in the first iteration. With more iterations, the performance index rapidly decays to 1. That is, the more often the agents perform a task, the less beneficial the exchange of information.
– Intuitively, the result shows that if the agents are different, the measurements of the other agents do not provide significant information for an individual's performance improvement. If the agents are almost identical, 'averaging' the measurements of the agents via a joint estimation still has no 'visible' effect, since the process noise directly corrupts the value of interest, x_j^i, see (6).

Moreover, independent estimation and learning (I) is robust to uncertainties in the initial noise assumptions (10). Note that the variance of an individual's disturbance in the independent case depends solely on the sum $(\alpha + \beta)$, cf. (19) with $N = 1$. In other words, the assumption on how the disturbance d^i is decomposed in d^0 and $d^{i,ind}$, does not enter the result. It does, however, affect the joint estimation.

To conclude, there is little benefit of sharing information in the case of pure process noise.

Proof: Based on the closed-form representation in (19), Theorem 1 is proven by introducing R^{proc} as a function of j, α, β, and N,

$$R_j^{\text{proc}}(\alpha,\beta,N) = \frac{p_j^{(1,1)}(\alpha,\beta,1) + 1}{p_j^{(1,1)}(\alpha,\beta,N) + 1}\,. \tag{24}$$

Recalling the properties

$$\alpha, \beta \ge 0 \quad \text{and} \quad j, N \in \{1,\, 2,\, \dots\}\,, \tag{25}$$

we note that $p_j^{(1,1)}(\alpha,\beta,N) \ge 0$ for all possible arguments. By taking partial derivatives of $R_j^{\text{proc}}(\alpha,\beta,N)$, it can be shown that

$$\frac{\partial R_j^{\text{proc}}(\alpha,\beta,N)}{\partial N} \ge 0 \tag{26}$$

and $R_j^{\text{proc}}(\alpha,\beta,N)$ is bounded by

$$R_j^{\text{proc}}(\alpha,\beta,\infty) := \lim_{N\to\infty} R_j^{\text{proc}}(\alpha,\beta,N) \tag{27}$$

with

$$R_j^{\text{proc}}(\alpha,\beta,\infty) = \frac{1 + \frac{\alpha+\beta}{1+j(\alpha+\beta)}}{1 + \frac{\beta}{1+j\beta}}\,.$$

Secondly, it is shown that

$$\frac{\partial R_j^{\text{proc}}(\alpha,\beta,\infty)}{\partial \alpha} \ge 0 \tag{28}$$

with

$$R_j^{\text{proc}}(\infty,\beta,\infty) := \lim_{\alpha\to\infty} R_j^{\text{proc}}(\alpha,\beta,\infty) = \frac{1 + \frac{1}{j}}{1 + \frac{\beta}{1+j\beta}}\,,$$

that is $R_j^{\text{proc}}(\alpha,\beta,N) \le R_j^{\text{proc}}(\infty,\beta,\infty)$. Finally, with

$$\frac{\partial R_j^{\text{proc}}(\infty,\beta,\infty)}{\partial \beta} \ge 0\,, \tag{29}$$

and

$$R_j^{\text{proc}}(\infty,0,\infty) = 1 + \frac{1}{j}\,,$$

statement (23) is proven,

$$R_j^{\text{proc}}(\alpha,\beta,N) \le R_j^{\text{proc}}(\infty,0,\infty)$$

for all α, β, N, j. The lower bound is obtained for $N = 1$, cf. (26). Matlab and Mathematica files for reproducing the results below are available at `www.idsc.ethz.ch/Downloads/multiagentILC`.

B. Pure Measurement Noise

We studied the system properties under the assumption of zero process noise, i.e. $\xi_j^i = 0$ in (6), and interpreted υ_j^i as pure measurement noise, $v_j^i = \mu_j^i$. Here, the noise term in (21) disappears and the ratio of the state variances is given by

$$R^{\text{meas}} = \frac{p_j^{(1,1)}\big|_{N=1}}{p_j^{(1,1)}}\,. \tag{30}$$

The following theorem can be stated:

Theorem 2. The bounds on the performance improvement due to joint estimation (vs. independent estimation) are given by

$$1 \,\le\, R^{\text{meas}} \,\le\, N \qquad \forall\, \alpha,\, \beta,\, N,\, j\,, \tag{31}$$

where the best performance improvement occurs when $\alpha \to \infty$ and $\beta = 0$, for all N, j. In this case, $R^{\text{meas}} = N$. ∎

Interpretation of the result:

- Again, an upper bound of the performance index is found which is valid for all possible combinations of α, β, N, and j. However, the upper bound does not depend on the number of iterations.
- Joint estimation is most beneficial if the agents' common disturbance component dominates and the individual noise component is negligible compared to the measurement noise; this corresponds to a large common noise variance α and a negligible individual component $\beta \ll 1$. The largest possible improvement in performance is a factor of N.
- Intuitively, the result shows that if the agents are very similar ($\beta \ll 1$), joint estimation has a 'visible' effect. The measurement noise is 'averaged out'. It does not corrupt the performance result, x_j^i, directly, see (7). A significant improvement in the individual's performance can be achieved.

Joint estimation is beneficial when considering a large group of almost identical agents, where the individual disturbance is small compared to the measurement noise.

Note that the performance index of the *mixed noise case* falls between R^{proc} and R^{meas}.

Proof: The proof of Theorem 2 proceeds similarly as the proof in Section III-A. With (19), the performance index R^{meas} is given as a function of j, α, β, and N,

$$R_j^{\text{meas}}(\alpha,\beta,N) = \frac{p_j^{(1,1)}(\alpha,\beta,1)}{p_j^{(1,1)}(\alpha,\beta,N)}\,. \tag{32}$$

Partial derivatives are directly computed, where

$$\frac{\partial R^{\text{meas}}}{\partial \beta} \le 0\,, \quad \frac{\partial R^{\text{meas}}}{\partial \alpha} \ge 0\,, \quad \frac{\partial R^{\text{meas}}}{\partial N} \ge 0\,, \tag{33}$$

with (25). In addition, the limiting property for $\beta = 0$ is

$$R_j^{\text{meas}}(\alpha,0,N) = \frac{1+\alpha j N}{1+\alpha j}$$

and

$$\lim_{\alpha\to\infty} R_j^{\text{meas}}(\alpha,0,N) = N \qquad \forall\, j,\, N\,.$$

The lower bound is obtained for $N = 1$, cf. (33); see also `www.idsc.ethz.ch/Downloads/multiagentILC`.

IV. Conclusion

In this paper we considered a group of agents which share the same dynamics and a common iteration-independent disturbance, but differ in an additional individual error component. In the context of having these agents learn to perform an identical task, we asked: How beneficial is it to exchange experience in order to improve an individual agent's learning performance? We considered two cases: (I) independent learning without information exchange and (II) learning based on full information exchange between agents. In the proposed framework, the question can be reduced to the comparison of the disturbance estimate in case of independent estimation (I) and when solving a global estimation problem for (II). An upper bound for the performance improvement due to information exchange is derived analytically and reflects the limited benefit of sharing information in the given setup. In the best case – where the noise is due to measurement noise only, the agent's common disturbance dominates, and the individual disturbance component is small compared to the noise – joint estimation improves the performance by a factor equal to the number of agents. That is, instead of one agent performing a task N times, N agents performing the task once results in the same accuracy for the disturbance estimate. For the general case and, in particular, in the presence of process noise or a large individual disturbance component, the benefits are shown to be limited.

Appendix

We derive an explicit representation of the variance $p_j^{(1,1)}$ that depends only on α, β, j, and N as presented in Proposition 1. Matlab and Mathematica files for reproducing the results below are available at `www.idsc.ethz.ch/Downloads/multiagentILC`.

Proof: A closed form of the covariance matrix P_j is derived, cf. (15) and (16). Since noise is assumed to have the same characteristics for each agent, by symmetry,

$$p_j^{(k,l)} = \begin{cases} p_j^{(0,0)} & \text{if} \quad k=l=0 \\ p_j^{(0,1)} & \text{if} \quad kl=0 \;\text{ and }\; k\neq l \\ p_j^{(1,1)} & \text{if} \quad k=l\neq 0 \\ p_j^{(1,2)} & \text{otherwise}\,. \end{cases} \tag{34}$$

We obtain the previous values by solving the filter equations, cf. [23], [24],

$$\begin{aligned} Q_j &= H P_{j-1} H^T + I \\ K_j &= P_{j-1} H^T Q_j^{-1} \\ P_j &= (I - K_j H) P_{j-1}, \end{aligned} \tag{35}$$

where $Q_j = [\,q_j^{(k,l)}\,],\; k,l \in \mathcal{I}$ and $K_j = [\,k_j^{(k,l)}\,],\; k \in \mathcal{K}$, $l \in \mathcal{I}$. With (34) and (35), the matrix Q_j and its inverse $Q_j^{-1} = [m_j^{(k,l)}]$ are directly computed,

$$q_j^{(k,l)} = \begin{cases} 1 + p_{j-1}^{(1,1)} & \text{if} \quad k=l \\ p_{j-1}^{(1,2)} & \text{otherwise} \end{cases} \tag{36}$$

$$m_j^{(k,l)} = \begin{cases} m_j^{(1,1)} & \text{if} \quad k = l \\ m_j^{(1,2)} & \text{otherwise}\,, \end{cases} \tag{37}$$

where

$$\begin{aligned} m_j^{(1,1)} &= \frac{1 + p_{j-1}^{(1,1)} + \left(N-2\right) p_{j-1}^{(1,2)}}{n_1 n_2} \\ m_j^{(1,2)} &= \frac{-p_{j-1}^{(1,2)}}{n_1 n_2} \end{aligned} \tag{38}$$

with

$$\begin{aligned} n_1 &= \left(1 + p_{j-1}^{(1,1)} - p_{j-1}^{(1,2)}\right) \\ n_2 &= \left(1 + p_{j-1}^{(1,1)} + \left(N-1\right) p_{j-1}^{(1,2)}\right)\,. \end{aligned} \tag{39}$$

With this, the filtering matrix K_j is given by

$$k_j^{(k,l)} = \begin{cases} k_j^{(0,1)} & \text{if} \quad k = 0 \\ k_j^{(1,1)} & \text{if} \quad k = l \\ k_j^{(1,2)} & \text{otherwise}\,. \end{cases} \tag{40}$$

where,

$$\begin{aligned} k_j^{(0,1)} =&\ p_{j-1}^{(0,1)} \left(m_j^{(1,1)} + \left(N-1\right) m_j^{(1,2)}\right) \\ k_j^{(1,1)} =&\ p_{j-1}^{(1,1)} m_j^{(1,1)} + \left(N-1\right) p_{j-1}^{(1,2)} m_j^{(1,2)} \\ k_j^{(1,2)} =&\ p_{j-1}^{(1,1)} m_j^{(1,2)} + p_{j-1}^{(1,2)} m_j^{(1,1)} \\ &+ \left(N-2\right) p_{j-1}^{(1,2)} m_j^{(1,2)} \end{aligned} \tag{41}$$

From (35), the following values for P_j are found,

$$p_j^{(k,l)} = \begin{cases} p_j^{(0,0)} & \text{if} \quad k = l = 0 \\ p_j^{(0,1)} & \text{if} \quad kl = 0 \text{ and} \quad l \neq k \\ p_j^{(1,1)} & \text{if} \quad k = l \neq 0 \\ p_j^{(1,2)} & \text{otherwise} \end{cases} \tag{42}$$

with

$$\begin{aligned} p_j^{(0,0)} &= p_{j-1}^{(0,0)} - N p_{j-1}^{(0,1)} k_j^{(0,1)} \\ p_j^{(0,1)} &= p_{j-1}^{(0,1)} + p_{j-1}^{(1,1)} k_j^{(0,1)} - \left(N-1\right) p_{j-1}^{(1,2)} k_j^{(0,1)} \\ p_j^{(1,1)} &= \left(1 - k_j^{(1,1)}\right) p_{j-1}^{(1,1)} - \left(N-1\right) p_{j-1}^{(1,2)} k_j^{(1,2)} \\ p_j^{(1,2)} &= \left(1 - k_j^{(1,1)}\right) p_{j-1}^{(1,2)} - k_j^{(1,2)} p_{j-1}^{(1,1)} \\ &\quad - \left(N-2\right) p_{j-1}^{(1,2)} k_j^{(1,2)} \end{aligned}$$

We prove the desired symmetry and obtain the following values for (34) by induction, using (42) with initial condition (15):

$$\begin{aligned} p_j^{(0,0)} &= \frac{(1+j\beta)\alpha}{1+j\beta+jN\alpha} \\ p_j^{(0,1)} &= \frac{\alpha}{1+j\beta+jN\alpha} \\ p_j^{(1,1)} &= \frac{\alpha+\beta+j\beta^2+jN\alpha\beta}{(1+j\beta)(1+j\beta+jN\alpha)} \\ p_j^{(1,2)} &= \frac{\alpha}{(1+j\beta)(1+j\beta+jN\alpha)}\,. \end{aligned} \tag{43}$$

The only value of interest is $p_j^{(1,1)}$.

References

[1] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *Journal of Robotic Systems*, vol. 1, no. 2, pp. 123–140, 1984.

[2] Z. Bien and J. X. Xu, *Iterative learning control: analysis, design, integration and applications*. Kluwer Academic Publishers Norwell, MA, USA, 1998.

[3] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 96–114, 2006.

[4] J. H. Lee, K. S. Lee, and W. C. Kim, "Model-based iterative learning control with a quadratic criterion for time-varying linear systems," *Automatica*, vol. 36, no. 5, pp. 641–657, 2000.

[5] M. Phan and R. Longman, "Higher-order iterative learning control by pole placement and noise filtering," in *Proceedings of the IFAC World Congress*, 2002, pp. 1899–1904.

[6] M. Norrlöf, "An adaptive iterative learning control algorithm with experiments on an industrial robot," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 2, pp. 245–251, 2002.

[7] ——, "Disturbance rejection using an ILC algorithm with iteration varying filters," *Asian Journal of Control*, vol. 6, pp. 432–438, 2004.

[8] A. Schöllig and R. D'Andrea, "Optimization-based iterative learning control for trajectory tracking," in *Proceedings of the European Control Conference (ECC)*, 2009, pp. 1505–1510.

[9] K. S. Lee, J. Lee, I. Chin, J. Choi, and J. H. Lee, "Control of wafer temperature uniformity in rapid thermal processing using an optimal iterative learning control technique," *Industrial and Engineering Chemistry Research*, vol. 40, no. 7, pp. 1661–1672, 2001.

[10] I. Chin, S. J. Qin, K. S. Lee, and M. Cho, "A two-stage iterative learning control technique combined with real-time feedback for independent disturbance rejection," *Automatica*, vol. 40, no. 11, pp. 1913–1922, 2004.

[11] J. K. Rice and M. Verhaegen, "A structured matrix approach to efficient calculation of LQG repetitive learning controllers in the lifted setting," *International Journal of Control*, vol. 83, no. 6, pp. 1265–1276, 2010.

[12] E. Guizzo, "Three engineers, hundreds of robots, one warehouse," *IEEE Spectrum*, vol. 45, no. 7, pp. 22–29, 2008.

[13] P. Wurman, R. D'Andrea, and M. Mountz, "Coordinating hundreds of cooperative, autonomous vehicles in warehouses," *AI Magazine*, vol. 29, no. 1, pp. 9–19, 2008.

[14] H.-S. Ahn, K. L. Moore, and Y. Chen, *Iterative Learning Control: Robustness and Monotonic Convergence for Interval Systems (Communications and Control Engineering)*, 1st ed. Springer, 2007.

[15] H.-S. Ahn and Y. Chen, "Iterative learning control for multi-agent formation," in *Proceedings of the International Joint Conference on Control, Automation, and Systems of the Society of Instrument and Control Engineers and the Institute of Control, Robotics and Systems (ICCAS-SICE)*, 2009, pp. 3111–3116.

[16] M. Tan, "Multi-agent reinforcement learning: Independent vs. cooperative agents," in *Proceedings of the 10th International Conference on Machine Learning*, 1993, pp. 330–337.

[17] M. Q. Phan and R. W. Longman, "A mathematical theory of learning control for linear discrete multivariable systems," in *Proceedings of the AIAA/AAS Astrodynamics Conference*, 1988, pp. 740–746.

[18] R. Tousain, E. van der Meche, and O. Bosgra, "Design strategy for iterative learning control based on optimal control," in *Proceedings of the 40th IEEE Conference on Decision and Control*, vol. 5, 2001, pp. 4463–4468.

[19] B. Bamieh, J. B. Pearson, B. A. Francis, and A. Tannenbaum, "A lifting technique for linear periodic systems with applications to sampled-data control," *Systems & Control Letters*, vol. 17, no. 2, pp. 79–88, 1991.

[20] J. Hätönen, D. Owens, and K. Feng, "Basis functions and parameter optimisation in high-order iterative learning control," *Automatica*, vol. 42, no. 2, pp. 287–294, 2006.

[21] K. Barton and A. Alleyne, "A cross-coupled iterative learning control design for precision motion control," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 6, pp. 1218–1231, 2008.

[22] M. Butcher, A. Karimi, and R. Longchamp, "Iterative learning control based on stochastic approximation," in *Proceedings of the 17th IFAC World Congress*, 2008, pp. 1478–1483.

[23] C. K. Chui and G. Chen, *Kalman Filtering: with Real-Time Applications (Springer Series in Information Sciences)*. Springer, 1998.

[24] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2007.