

Optimization of the prefilter in Iterative Feedback Tuning for improved accuracy of the controller parameter update.*

R. Hildebrand[†], A. Lecchini^{‡,§}, G.Solari[#] and M. Gevers[#]

[†] Center for Operations Research and Econometrics (CORE)
Université Catholique de Louvain - Belgium
hildebrand@core.ucl.ac.be

[‡] Department of Engineering
University of Cambridge - UK
a1394@cam.ac.uk

[#] Centre for Systems Engineering and Applied Mechanics (CESAME)
Université Catholique de Louvain - Belgium
{solari, gevers}@auto.ucl.ac.be

Abstract—Iterative Feedback Tuning (IFT) is a data-based method for the tuning of restricted complexity controllers. At each iteration, an update for the parameters of the controller is estimated from data obtained partly from the normal operation of the closed loop system and partly from a special experiment. The choice of a prefilter for the input data to the special experiment is a degree of freedom of the method. In the present contribution, the prefilter is designed in order to enhance the robustness of the IFT update.

I. INTRODUCTION

The Iterative Feedback Tuning (IFT) is a data-based method to tune the parameters of a controller with a given structure [2][3][5][6]. The objective of IFT is to minimize a quadratic performance criterion defined on the controller parameter space. IFT consists in a stochastic gradient descent scheme. The gradient of the performance criterion is estimated from a set of data obtained partly from normal operation and partly from a special experiment on the plant. This gradient estimate is used to perform the next descent step in the iterations of the vector of controller parameters. Under suitable assumptions the algorithm converges to a local minimum of the performance criterion [2][4].

In the IFT procedure, the user is given the possibility of prefiltering the input data to the special experiment. Basically, the choice of a particular prefilter is a degree of freedom of the procedure. In the original formulation of IFT [6] this degree of freedom was not used. In practice, this corresponds to using a trivial constant prefilter. In [2][3], it has been shown that the prefilter influences the covariance of the

gradient estimate. A prefilter was derived which minimizes the weighted trace of this covariance for a given weighting matrix. The main motivation for this result was the fact that near the optimal point the asymptotic convergence rate of IFT, measured as the rate of approach of the expected cost to the minimal value, corresponds to an optimal selection of the weight. By choosing the correct weight, an optimal prefilter was derived that maximally improves the convergence rate of the algorithm. The choice of the correct weight was based on an estimate of the Hessian of the design cost function near the optimal point.

As far as we know, there are no clear results that allow one to formalize a design criterion for the convergence rate in the case where the current controller is far from the optimal one. As a matter of fact, the convergence depends on the global shape of the performance criterion. This shape is unknown and can hardly be estimated. Therefore, in this situation, the objective is mainly the improvement of the accuracy of a single IFT step. By reducing the deviation of the actual descent direction from the optimal direction at each iteration, one enhances the robustness of the whole iterative procedure.

In this contribution, we show that it is in general possible, by prefiltering, to obtain a covariance matrix that is strictly smaller than the one obtained with standard IFT step, i.e. with a trivial constant prefilter. By strictly smaller we mean that the difference between the covariance matrix obtained with standard IFT and the new one is positive definite. We propose a design criterion for the prefilter that is consistent with this goal and show how a prefilter can be computed that optimizes this design criterion. It turns out, as it was the case in [2][3], that the proposed prefilter can be estimated from data collected under normal operating conditions. Thus the computation of the prefilter does not require any special experiment on the process and does not impose any additional cost.

The paper is structured as follows. In the next section we

*Paper supported by the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture and the European Research Network on System Identification (ERNSI) funded by the European Union. The scientific responsibility rests with its authors.

[§]This work was carried out while the second author was a post doc researcher at the Centre for Systems Engineering and Applied Mechanics (CESAME) Université Catholique de Louvain - Belgium.

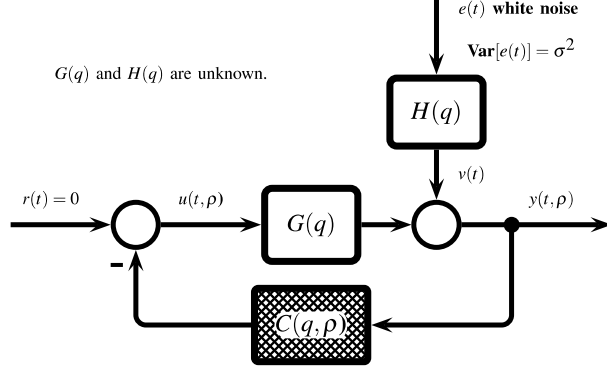


Fig. 1. The control system under normal operating conditions.

recall some results on the statistical properties of the gradient estimate in IFT. This enables us in Section 3 to establish a design criterion that has to be minimized with respect to the prefilter in order to strictly reduce the covariance of the gradient estimate. We also show how to compute a prefilter that is optimal with respect to this design criterion. In Section 4, we demonstrate, by a simulation example, the gain in accuracy between the use of the optimal prefilter and the use of the trivial constant prefilter. Finally, we draw some conclusions in the last section.

II. THE IFT PARAMETER UPDATE

We assume that the plant to be controlled is a SISO linear time-invariant system; its transfer function is denoted by $G(q)$. The output of the plant is affected by an additive stochastic disturbance $v(t) = H(q)e(t)$ where $H(q)$ is a monic, stable and inversely stable transfer function and $e(t)$ is zero mean white noise with variance σ^2 . The transfer functions $G(q)$ and $H(q)$ are unknown. We consider the closed loop system depicted in Figure 1, where $C(q, \rho)$ belongs to a parameterized set of controllers with parameter $\rho \in \mathbf{R}^n$. The transfer function from $v(t)$ to $y(t, \rho)$ is denoted by $S(q, \rho)$. The reference signal $r(t)$ is set at zero under normal operating conditions. The goal is to find a minimizer for the cost function

$$J(\rho) = \frac{1}{2} \mathbf{E} [y(t, \rho)^2 + \lambda u(t, \rho)^2], \quad (1)$$

where $\lambda \geq 0$ is a penalty on the control effort chosen by the user.

The IFT method is an iterative procedure that gives a solution to this problem. It is based on the construction of an unbiased estimate of the gradient of $J(\rho)$ from data collected on the plant. The cost function $J(\rho)$ is minimized with an iterative stochastic gradient descent scheme of Robbins-Monro type [1]. Under some suitable assumptions [2][6], the sequence of controllers converges to a local minimum of $J(\rho)$.

Since every iteration proceeds in a similar fashion, in the sequel it suffices to consider the stage where $i = 0$. Therefore

we will denote the current controller parameter by ρ_0 and the updated parameter by ρ_1 . The IFT parameter update rule is given by

$$\rho_1 = \rho_0 - \gamma R^{-1} \text{est}_N \left[\frac{\partial J}{\partial \rho}(\rho_0) \right] \quad (2)$$

where γ is a positive step size and R is a positive definite matrix. The reader is referred to [6] for the algorithm to construct the gradient estimate. Here it suffices to recall that, in order to construct $\text{est}_N \left[\frac{\partial J}{\partial \rho}(\rho_0) \right]$, first a batch $\{u^1(t, \rho_0), y^1(t, \rho_0)\}_{t=1, \dots, N}$ of N data is collected with the controller $C(q, \rho_0)$ in the loop under normal operating conditions. Then, this batch of data is used to construct the reference signal $r(t) = -K(q)y^1(t, \rho_0)$ which is applied to the reference input of the system, see Figure 1; this is a “special experiment”, which deviates from normal operating conditions. The choice of the prefilter $K(q)$ is a degree of freedom of the algorithm and is basically left to the user. In the original formulation of IFT the prefilter was not used, which corresponds to setting $K(q) = \text{const}$.

The prefilter $K(q)$ can influence the statistical properties of the IFT update as has been shown in [6]. More specifically, we have the following proposition.

Proposition 2.1: Let $\bar{P} = \lim_{N \rightarrow \infty} N \text{Cov}[\rho_1]$. Then \bar{P} can be decomposed as $\bar{P} = \bar{E} + \bar{S}$ where \bar{E} is given by

$$\begin{aligned} \bar{E} = & \gamma^2 R^{-1} \left[\frac{\sigma^4}{2\pi} \int_{-\pi}^{\pi} \frac{|S(e^{j\omega}, \rho_0)H(e^{j\omega})|^4}{|K(e^{j\omega})|^2} [1 + \right. \\ & \left. + \lambda |C(e^{j\omega}, \rho_0)|^2]^2 \frac{\partial C}{\partial \rho}(e^{j\omega}, \rho_0) \frac{\partial C^*}{\partial \rho}(e^{j\omega}, \rho_0) d\omega \right] R^{-T} \end{aligned}$$

and \bar{S} is a constant matrix which does not depend on the choice of $K(q)$.

Proof: See [2]. \square

This result shows how the covariance of the parameter update depends on the prefilter $K(q)$. It is the sum of a constant term and a term that is a frequency weighted integral of the inverse of $K(q)$. In the next section we will show how to choose $K(q)$ in order to make this covariance smaller than what is obtained with a constant prefilter.

III. DESIGN OF AN OPTIMAL PREFILTER

By Proposition 2.1 we can influence the covariance term \bar{E} by choice of the prefilter $K(q)$. Our goal is to make \bar{E} as small as possible by choosing $K(q)$ appropriately. However, here we deal with a matrix-valued object, minimization of which has no well-defined meaning. Nevertheless, there exists a partial ordering on the space of symmetric matrices. Namely, if the difference $\bar{P}_2 - \bar{P}_1$ of two symmetric matrices \bar{P}_1, \bar{P}_2 is positive definite, we can say that \bar{P}_1 is strictly smaller than \bar{P}_2 . Clearly, if we have a choice between two different prefilters $K(q)$ yielding two different covariance matrices, which are comparable in this sense, then it is preferable to use that prefilter which leads to the smaller covariance matrix.

Specifically, if we can find a prefilter which leads to a covariance matrix \bar{P} which is strictly smaller than the covariance matrix obtained by using no prefilter at all, then it is preferable to use this prefilter. In this section we point out a subset of prefilters which lead to such covariance matrices, and we propose an algorithm to construct the prefilter in that set which leads to the smallest covariance matrix.

We shall proceed as follows. First we clarify the structure of the set of all covariance matrices \bar{P} which can be obtained at all by using all possible prefilters $K(q)$. This will give us clues for the construction of a prefilter which leads to a covariance matrix that is smaller than the one obtained with some given prefilter. Specifically, we will construct a prefilter that yields a covariance matrix which is smaller than the one obtained by using no prefilter at all (i.e. by using the trivial constant prefilter).

A. The set of achievable covariance matrices

By Proposition 2.1 we can write

$$\bar{E} = \int_0^\pi \frac{1}{|K(e^{j\omega})|^2} M(\omega) d\omega, \quad (3)$$

i.e. \bar{E} is a weighted integral over a frequency-dependent real positive semidefinite matrix $M(\omega)$, which does not depend on the prefilter $K(q)$. The prefilter is assigned the role of a positive weighting function. Thus, by choice of a suitable prefilter, we can assign to the expression \bar{E} a value that is arbitrarily close to any given matrix in the convex conic hull of the matrix-valued curve $M(\omega)$. By Proposition 2.1, the closure of the set of all asymptotic covariance matrices \bar{P} that can be achieved by choosing a prefilter is an affine closed convex cone with an offset \bar{S} . Let us denote this cone by \mathcal{C} .

Since \bar{E} is inversely proportional to the squared magnitude of the prefilter $K(q)$, we could make it as small as we wish by choosing a prefilter with a sufficiently large magnitude. However, this would be at the cost of a higher input energy for the reference signal $r(t) = -K(q)y^1(t, \rho_0)$ of the special experiment and this would represent a bigger perturbation of the process, which is not desirable. Thus we have to restrict the set of allowed prefilters $K(q)$ by imposing some bound α on this input energy $\mathbf{E}[r(t)^2]$. This bound represents the level of acceptable perturbation of the normal operating conditions during the special experiment. Since $r(t) = -K(q)y^1(t, \rho_0)$, we impose the following restriction on the magnitude of $K(q)$: $\frac{\sigma^2}{2\pi} \int_{-\pi}^\pi |K(e^{j\omega})|^2 |S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2 d\omega \leq \alpha$. This can be written as

$$\int_0^\pi |K(e^{j\omega})|^2 w(\omega) d\omega \leq 1, \quad (4)$$

where $w(\omega) = \frac{\sigma^2}{\alpha\pi} |S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2$ is a positive frequency-dependent scalar function. The set of covariance matrices \bar{P} that can be obtained by prefiltering under restriction (4) is naturally smaller than the entire cone \mathcal{C} . We shall now investigate this set.

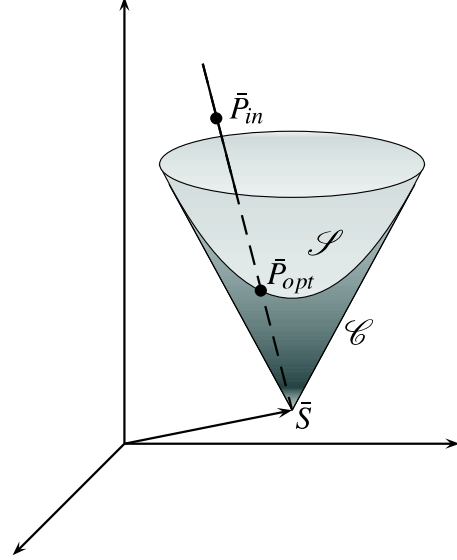


Fig. 2. Representation of the optimization problem.

Let $\bar{P}_{in} = \bar{S} + \bar{E}_{in} \in \mathcal{C} \setminus \{\bar{S}\}$ be an arbitrary covariance matrix which can be achieved by prefiltering with some prefilter. Then the ray $\{\bar{S} + \kappa \bar{E}_{in} | \kappa \geq 0\}$ is contained in the cone \mathcal{C} . It is easily seen that there exists a unique matrix $\bar{P}_{opt} = \bar{S} + \kappa_{opt} \bar{E}_{in}$ on this ray such that any matrix $\bar{S} + \kappa \bar{E}_{in}$ with $\kappa \geq \kappa_{opt}$ can be approximated arbitrarily well by choosing prefilters satisfying constraint (4), but matrices $\bar{S} + \kappa \bar{E}_{opt}$ with $\kappa < \kappa_{opt}$ cannot. The union of these matrices \bar{P}_{opt} for all rays in \mathcal{C} forms a section \mathcal{S} of the cone \mathcal{C} (see Fig. 2).

Proposition 3.1: The section \mathcal{S} is a convex hypersurface.

Proof. It is sufficient to show that the closure of the set of all covariance matrices that can be achieved by prefiltering under restriction (4) is convex. Let $K_1(q)$, $K_2(q)$ be two prefilters satisfying (4) and yielding covariance matrices $\bar{P}_1, \bar{P}_2 \in \mathcal{C}$, respectively. Let $K_\tau(q)$ be such that $\frac{1}{|K_\tau(q)|^2} = \tau \frac{1}{|K_1(q)|^2} + (1-\tau) \frac{1}{|K_2(q)|^2}$, $\tau \in [0, 1]$. Then by convexity of the function $f(x) = \frac{1}{x}$ every filter $K_\tau(q)$ satisfies restriction (4). Thus the matrices on the line segment between the covariance matrices \bar{P}_1, \bar{P}_2 can be approximated arbitrarily well, along with the endpoints of this segment. \square

Prefilters $K(q)$ satisfying (4) and leading to covariance matrices on the surface \mathcal{S} are optimal in the following sense. Suppose we are given some prefilter $K_{in}(q)$ satisfying (4) and leading to a covariance matrix $\bar{P}_{in} = \bar{S} + \bar{E}_{in}$ in the convex hull of \mathcal{S} . Then the covariance matrix $\bar{P}_{opt} = \bar{S} + \kappa_{opt} \bar{E}_{in} \in \mathcal{S}$ is the *smallest* covariance matrix on the ray $\{\bar{S} + \kappa \bar{E}_{in} | \kappa \geq 0\}$ which can be approximated by using prefilters satisfying (4). In particular, we have $\kappa_{opt} \leq 1$ and \bar{P}_{opt} is smaller than \bar{P}_{in} . In this sense a prefilter $K_{opt}(q)$ which satisfies (4) and produces the covariance matrix \bar{P}_{opt} is optimal along the given ray. In the next subsection we shall construct this optimal prefilter.

B. The optimal prefilter

The problem of finding the optimal prefilter $K_{opt}(q)$ corresponding to a given initial covariance matrix \bar{P}_m can be cast as the following optimization problem (compare (3), (4)). Given an initial prefilter $K_{in}(q)$ which realizes a covariance matrix $\bar{P}_m = \bar{S} + \bar{E}_{in}$, minimize κ , by choice of $K(e^{j\omega})$, under the constraints

$$\int_0^\pi |K(e^{j\omega})|^2 w(\omega) d\omega \leq 1, \quad \int_0^\pi \frac{1}{|K(e^{j\omega})|^2} M(\omega) d\omega = \kappa \bar{E}_{in}.$$

A natural choice of the initial prefilter $K_{in}(q)$ is the trivial constant prefilter with maximum gain satisfying the energy constraint.

We shall now solve this problem. By associating a scalar Lagrange multiplier $\lambda_0 \geq 0$ and a matrix-valued Lagrange multiplier Λ with the two constraints, we obtain the Lagrange function

$$L = \kappa + \lambda_0 \left(\int_0^\pi |K(e^{j\omega})|^2 w(\omega) d\omega - 1 \right) + \left\langle \Lambda, \int_0^\pi \frac{1}{|K(e^{j\omega})|^2} M(\omega) d\omega - \kappa \bar{E}_{in} \right\rangle.$$

By setting the gradient of L with respect to the design variables κ and $K(\omega)$ to zero and inserting the resulting equations into the constraints, we get

$$|K_{opt}(\omega)|^2 = \frac{\sqrt{\langle \Lambda, M(\omega) \rangle}}{w(\omega) \int_0^\pi \sqrt{\langle \Lambda, M(\omega) \rangle} w(\omega) d\omega}, \quad (5)$$

$$\begin{aligned} & \frac{\partial}{\partial \Lambda} \left(-\ln \int_0^\pi \sqrt{\langle \Lambda, M(\omega) \rangle} w(\omega) d\omega + \frac{1}{2} \langle \Lambda, \bar{E}_{in} \rangle \right) = \\ & = \frac{\partial f(\Lambda)}{\partial \Lambda} = 0. \end{aligned}$$

The function $f(\Lambda)$ is convex with respect to Λ . In order to determine Λ we thus have to solve the following convex optimization problem.

$$\text{minimize } f(\Lambda) \quad \text{s.t. } \langle \Lambda, M(\omega) \rangle \geq 0 \quad \forall \omega. \quad (6)$$

The matrix-valued function $M(\omega)$ is rational in $\cos(\omega)$. Therefore the constraint $\langle \Lambda, M(\omega) \rangle > 0 \quad \forall \omega$ is semidefinite representable and can be formulated as an LMI (Linear Matrix Inequality). Problem (6) is thus a standard convex optimization problem for which efficient numerical solution algorithms are available. Note that the gradient of f tends to infinity as Λ approaches the boundary of the feasible set. Hence at the optimum, if it exists, the inequality in (6) is strict.

Before discussing how to minimize $f(\Lambda)$ in practice, let us turn to the question of the existence of a solution. The matrix \bar{P}_m is not contained in the convex cone \mathcal{C} if and only if there exists a separating linear functional, i.e. a feasible Λ such that $\langle \Lambda, \bar{E}_{in} \rangle < 0$. If this is the case, then we will necessarily encounter such a Λ in the course of the minimization procedure, because the logarithm grows slower

than a linear function. If \bar{P}_m lies in the interior of \mathcal{C} , then there exists a minimizer Λ^* of $f(\Lambda)$.

Remark: A feasibility check can be more easily performed by solving the auxiliary problem of minimizing the linear functional $\langle \Lambda, \bar{E}_{in} \rangle$ over all Λ such that $\langle \Lambda, M(\omega) \rangle \geq 0$ for all $\omega \in [0, \pi]$. If $M(\omega)$ is rational in $\cos(\omega)$, then this is a standard semidefinite program. Thus we obtain a simple tool to check whether a given covariance matrix can be achieved with some prefilter or not.

In order to solve (6) and compute this prefilter in practice, one needs an estimate of the unknown spectral density $|S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2$ of the signal $y(t, \rho_0)$ which is the output of the plant under normal operating conditions, i.e. with zero reference signal. This quantity is in fact the only unknown part in $M(\omega)$ and $w(\omega)$. Such estimate can be obtained with standard techniques in the time or in the frequency domain [7], [8]. Note that since the data needed to estimate $|S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2$ do not stem from a special experiment they are available in large amounts. In fact, periods of normal operating conditions can be interlaced with the IFT special experiments. By assuming these periods to be much longer than the length of the special experiment from which the gradient is estimated, the contribution of the variability in the estimate of $|S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2$ to the variability of the gradient estimate can be considered as being negligible. Having an estimate of $|S(e^{j\omega}, \rho_0)H(e^{j\omega})|^2$ one can solve (6). Having determined Λ , one can compute the magnitude of the prefilter according to (5). Then, there exist standard tools to approximate a given magnitude function by a minimum phase filter. Before closing this section, let us remind the reader that the optimal prefilter is guaranteed to yield a smaller covariance matrix of the updated parameter vector than the one which would be obtained with the standard IFT procedure; it is thus optimal in the sense described in the previous subsection.

IV. SIMULATION EXAMPLE

Consider the system described by $G(q) = \frac{q^{-1} - 0.5q^{-2}}{1 - 0.3q^{-1} - 0.28q^{-2}}$ and $H(q) = \frac{1}{1 + 0.9q^{-1}}$ with $\sigma^2 = 1$. Let the class of controllers be $C(q, \rho) = \frac{\rho^1}{1 + \rho^2 q^{-1}}$ and set $\lambda = 0.6$ in (1). Let the current stabilizing controller be given by: $\rho_0 = [-1.25 \quad 0.83]$. We assume that the constraint on the reference signal $r(t)$ of the special experiment is that this signal must have the same energy as the output of the plant in normal operating conditions. In the following, we quantify the accuracy improvement on the first parameter vector update ρ_1 , when the optimal prefilter $K_{opt}(q)$, given by (5), is used instead of the trivial constant prefilter $K_{in}(q)$, both of which having the maximum gain satisfying the energy constraint on $\mathbf{E}[r(t)^2]$.

We consider the IFT update (2) with step size $\gamma = 1$ and $R = I$. By using Proposition 2.1 and results from [2] we can find the numerical value of $\bar{P} = \lim_{N \rightarrow \infty} N \text{Cov}[\rho_1]$.

For the constant prefilter $K_{in}(q)$ we obtain

$$\bar{P}_{in} = 10^4 \times \begin{bmatrix} 0.0524 & -0.3019 \\ -0.3019 & 2.2765 \end{bmatrix} = \bar{S} + \bar{E}_{in}$$

in which the $K(q)$ -dependent component is given by

$$\bar{E}_{in} = 10^4 \times \begin{bmatrix} 0.0472 & -0.3053 \\ -0.3053 & 2.2735 \end{bmatrix}.$$

For the optimal prefilter $K_{opt}(q)$ we have

$$\bar{P}_{opt} = 10^3 \times \begin{bmatrix} 0.1245 & -0.4309 \\ -0.4309 & 3.4975 \end{bmatrix} = \bar{S} + \bar{E}_{opt}$$

where $\bar{E}_{opt} = \kappa_{opt} \bar{E}_{in}$ with $\kappa_{opt} = 0.15$. The difference in the total covariance for the two cases is then

$$\bar{\Delta} = \bar{P}_{in} - \bar{P}_{opt} = 10^4 \times \begin{bmatrix} 0.0399 & -0.2588 \\ -0.2588 & 1.9267 \end{bmatrix}.$$

The improvement in the use of the optimal prefilter is shown by the fact that $\bar{\Delta}$ is positive definite. Thus, whatever scalar measure of the covariance matrix one might use to evaluate the spread of ρ_1 , the use of $K_{opt}(q)$ leads to a strict improvement.

The above theoretical values can be verified by a Monte-Carlo simulation. The parameter vector ρ_1 has been extracted 2048 times, by performing 2048 times the update step (2) with a different noise realization each one of length $N = 1024$. The 2048 parameter vectors obtained in this way are shown in Figure 3 for the case of the constant prefilter. The corresponding sampled estimate of $NCov[\rho_1]$ is given by

$$\hat{P}_{in} = 10^4 \times \begin{bmatrix} 0.0523 & -0.3003 \\ -0.3003 & 2.2338 \end{bmatrix}.$$

The parameter vectors obtained for the case of the optimal prefilter are shown in Figure 4. In this case, the corresponding sampled estimate of $NCov[\rho_1]$ is given by

$$\hat{P}_{opt} = 10^3 \times \begin{bmatrix} 0.1270 & -0.4344 \\ -0.4344 & 3.5599 \end{bmatrix}.$$

Observe that the results of the Monte-Carlo simulation are really close to the theoretical results.

V. CONCLUSIONS

In this contribution we have exploited the possibility to prefilter the reference input in the special experiment of the IFT procedure in order to enhance the robustness of the algorithm. This was performed by altering the covariance of the gradient estimate, which can be directly influenced by the choice of a prefilter in this experiment. Based on the relationship between the prefilter and the covariance of the parameter update, we have formulated a design criterion for the prefilter. The problem of optimizing the prefilter with respect to this criterion amounts to a convex optimization problem with LMI constraints.

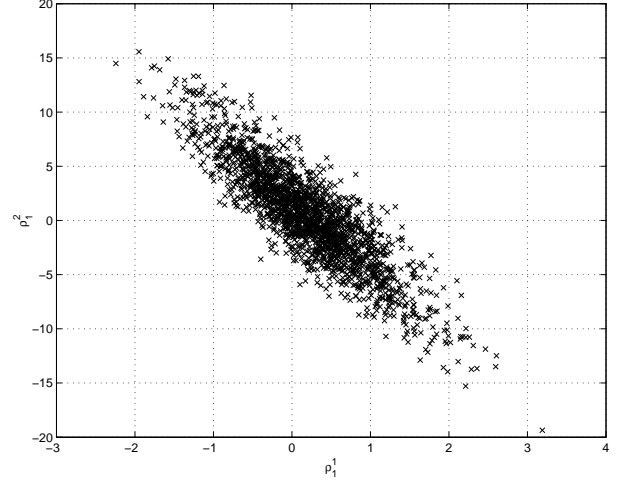


Fig. 3. The 2048 parameters ρ_1 obtained using the constant prefilter $K_{in}(q)$.

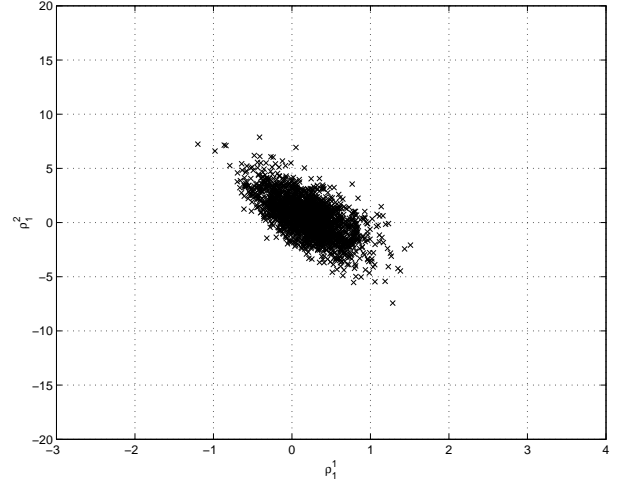


Fig. 4. The 2048 parameters ρ_1 obtained using the optimal prefilter $K_{opt}(q)$.

VI. REFERENCES

- [1] Julius R. Blum. Multidimensional stochastic approximation methods. *Annals of Mathematical Statistics*, 25:737–744, 1954.
- [2] R. Hildebrand, A. Lecchini, G. Solari, and M. Gevers. Convergence analysis and optimal prefiltering in Iterative Feedback Tuning. *Submitted*, August 2002.
- [3] R. Hildebrand, A. Lecchini, G. Solari, and M. Gevers. Optimal prefiltering in Iterative Feedback Tuning. *CD-ROM 13th IFAC Symposium on System Identification*, pages 513–518, 2003.
- [4] H. Hjalmarsson. Performance analysis of iterative feedback tuning. Technical report, Dept of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden, 1998.
- [5] H. Hjalmarsson. Iterative Feedback Tuning - an

overview. *International Journal of Adaptive Control and Signal Processing*, 16(5):373–395, Special Issue: New Ideas in Recursive Estimation and stochastic Adaptive Control 2002.

- [6] H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative Feedback Tuning: theory and applications. *IEEE Control Systems*, 18(4):26–41, August 1998.
- [7] L. Ljung. *System Identification: Theory for the user*. Prentice Hall, 1999.
- [8] R. Pintelon and J. Schoukens. *System Identification: a frequency domain approach*. IEEE Press, 2001.