2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Probabilistic μ for rank-one and perturbed rank-one matrices

Máté Mánfay[1], Gary J. Balas[2], József Bokor [3] and László Gerencsér [4]

***Abstract*— The structured singular value μ has been widely studied for uncertain dynamical systems. Recently a great attention is paid to the probabilistic μ problem. Instead of computing the conservative worst-case μ we are interested in the probabilistic distribution of μ, given a probability distribution on the set of uncertainties. Traditionally this problem is solved by Monte Carlo algorithms. In this paper we propose analytic methods to compute the probabilistic μ for rank-one and perturbed rank-one matrices. We expect that these results will provide an algorithm that is not as computationally expensive as the linear cut algorithm in [1].**

I. Introduction

In this paper we consider a special case of the probabilistic μ problem, which will be called the probabilistic rank one μ problem. The rank-one problem has been studied by several authors. In [2] two cases are considered. Denoting the reference matrix by M the first case is when $M = I$, while the second case is when dyad M is a dyadic product. In both cases the elements of M are real numbers. The model studied there is a purely probabilistic one, a mixture of the worst-case and probabilistic approach is not considered. Letting Δ denote the (random) uncertainty matrix, the distribution of $\bar{\lambda}(M\Delta)$ is defined as the probabilistic μ. In this paper we restrict ourselves to highly structured uncertainties: they are supposed to be diagonal matrices with elements having uniform distribution. Arguments in favor of the uniform distribution are given in [5]. [2] gives a formula, without proof, for the probabilities $\mathbb{P}(\mu_P(M) < \gamma)$. The contribution of this paper is that we give a formula, with proof, for the distribution of the probabilistic μ in the case of rank one M for both the pure probabilistic and the mixed (probabilistic and worst-case) formulation.

Furthermore, we use these results to estimate the probability distribution of $\mu_P(M)$ for perturbed rank-one M matrices. So far we have only worked with diagonal uncertainties. The study of full block uncertainties might be based on results concerning the probability density function defined on full block uncertainties [6],[7]. In the last section we investigate how B&B and linear cut type cutting algorithms reduce the variance of Monte Carlo algorithms.

II. Pure Probabilistic μ

A. Problem Formulation

Suppose that M is a dyad, i.e. $M = uv^T$ for some $u, v \in \mathbb{R}^n$, where M^T denotes the transpose of M. The conservative definition of the structured singular value is

$$\mu_{\mathbb{D}}(M) = \left(\min_{\Delta \in \mathbb{D}} \{ \overline{\sigma}(\Delta) : \det(I - \Delta M) = 0 \} \right)^{-1},$$

where $\mathbb{D}$ is a set of block diagonal matrices, and $\overline{\sigma}(A)$ denotes the largest singular value of A. This definition corresponds to a worst case scenario.

To define a probabilistic μ consider a random block diagonal matrix Δ. More precisely, let $(Z, \mathcal{F}, \zeta)$ be a probability space, and let Δ be a random matrix of the realization of which is of the form

$$\Delta(\zeta) = \mathrm{diag}(\delta_1(\zeta), ..., \delta_n(\zeta)).$$

In this paper we suppose that δ_k-s are independent uniformly distributed random variables, such that $\delta_k \sim UNI[-a_k, a_k]$ with some $a_k > 0$. Then the probabilistic μ will be defined as

$$\mu_P(M) = \bar{\lambda}(M\Delta),$$

where $\bar{\lambda}(.)$ denotes the largest eigenvalue of a matrix. This definition is motivated by the fact that $\mu_{\mathbb{D}}(M) = \max_{\Delta} \bar{\lambda}(M\Delta)$. Our aim is to calculate the probability density function of $\mu_P(M)$.

B. Analysis

Theorem II.1. *If M is a dyad and the uncertainties form a diagonal matrix with $\Delta = \mathrm{diag}(\delta_1, ..., \delta_n)$, where δ_k-s are independent and $\delta_k \sim \mathrm{UNI}[-a_k, a_k]$, then the probabilistic density function of $\mu_P(M)$ is*

$$f_{\mu_P}(t) = \left(\prod_{k=1}^{n} 2m_k \right)^{-1} \sum_{s \in S} \left[(-1)^{\sum_{k=1}^{n} s_k} \times \mathrm{sgn}\left(t + \sum_{k=1}^{n} (-1)^{s_k} m_k \right) \frac{(t + \sum_{k=1}^{n} (-1)^{s_k} m_k)^{n-1}}{(n-1)!} \right], \tag{1}$$

where $S = \{(s_1, ..., s_n) : s_k \in \{0, 1\}\}$ and $m_k = |u_k v_k a_k|$.

Proof:

First note that $\mu(M) = \bar{\lambda}(\Delta(\zeta)M)$, since Δ is diagonal. Furthermore $\Delta(\zeta)M$ is a dyad, so it has only one eigenvalue. Hence,

$$\bar{\lambda}(\Delta(\zeta)M) = \sum_{k=1}^{n} u_k v_k \delta_k(\zeta).$$

Here $u_k v_k \delta_k \sim \mathrm{UNI}[-|u_k v_k a_k|, |u_k v_k a_k|]$, for simplicity let $m_k = |u_k v_k a_k|$. Let φ_X and f_X denote the characteristic function and the probability density function of a random

[1]MTA SZTAKI, Budapest, Hungary; Central European University, Budapest, Hungary
[2]University of Minnesota, Minneapolis, USA
[3]MTA SZTAKI, Budapest, Hungary; MTA-BME Control Engineering Research Group
[4]MTA SZTAKI, Budapest, Hungary

variable X, respectively. Then

$$\varphi_{u_k v_k \delta_k}(\omega) = \int_{-\infty}^{\infty} \frac{1}{2m_k} e^{i\omega t} \mathbb{I}\left[t \in [-m_k, m_k]\right] dt = \quad (2)$$

$$\frac{1}{2m_k} \int_{-m_k}^{m_k} e^{i\omega t} dt = \frac{1}{2m_k} \frac{e^{i\omega m_k} - e^{-i\omega m_k}}{i\omega}. \quad (3)$$

Calculate the characteristic function of μ_P

$$\varphi_{\mu_P}(\omega) = \mathbb{E}\left(\exp(i\omega\mu)\right) = \mathbb{E}\left(\exp\left(\sum_{k=1}^{n} i\omega m_k \delta_k\right)\right) =$$

$$= \prod_{k=1}^{n} \mathbb{E}\left(\exp\left(i\omega m_k \delta_k\right)\right) = \prod_{k=1}^{n} \frac{1}{2m_k} \frac{e^{i\omega m_k} - e^{-i\omega m_k}}{i\omega}$$

$$= \prod_{k=1}^{n} (2m_k)^{-1} \sum \frac{\exp(i\omega \sum_{k=1}^{n} \pm m_k)}{(i\omega)^n} \quad (4)$$

A typical term in (4) is of the form

$$\frac{\exp(i\omega \sum_{k=1}^{n} \pm m_k)}{(i\omega)^n}.$$

Since φ_{μ_P} is the Fourier transform of f_{μ_P} to get the probability density function we invert the transformation. This yields

$$\mathscr{F}^{-1}\left(e^{i\omega c}(i\omega)^{-n}\right) = \operatorname{sgn}(c+t) \frac{(c+t)^{n-1}}{(n-1)!}.$$

Using this result we obtain the probability density function of μ_P:

$$f_{\mu_P}(t) = \left(\prod_{k=1}^{n} 2m_k\right)^{-1} \sum_{s \in S} \left[(-1)^{\sum_{k=1}^{n} s_k} \times \right.$$

$$\left. \operatorname{sgn}\left(t + \sum_{k=1}^{n} (-1)^{s_k} m_k\right) \frac{(t + \sum_{k=1}^{n} (-1)^{s_k} m_k)^{n-1}}{(n-1)!}\right], \quad (5)$$

where $S = \{(s_1, ..., s_n) : s_k \in \{0, 1\}\}$. □

Remark 1: If we assume that δ_k-s are independent and have normal distribution, then the distribution of μ_P is normal too.

Remark 2: Similar formula for the density function can be obtained if $\delta_k \sim \mathrm{UNI}[-a_k, b_k]$, and $a_k \neq b_k$.

Remark 3: Note that for rank-one reference matrices our defintion of probabilistic μ coincide with the definition seen in [2], i.e. with the distribution of $\bar{\lambda}(M\Delta)$.

C. Numerical Results

In Fig 1. and Fig 2. one can see the probability density function (pdf) of μ_P calculated with the previous formula and with Monte Carlo methods. The pure probabilistic case was considered that was formulated in the previous section. The actual pdf was calculated using (1) with grid size 0.01, the Monte Carlo simulation generated 10000 random diagonal matrices with elements having uniform distribution on the interval $[-1, 1]$. In Fig 3. $u = (0.1, 0.5, 0.7, 0.2, 1, 0.9)^T$ and $v = (0.4, 0.9, 0.2, 0.3, 0.1, 0.9)^T$. In Fig 4. $u = (0.1, 0.5, 0.7, 0.2)^T$ and $v = (0.4, 0.9, 0.2, 0.3)^T$ were used.

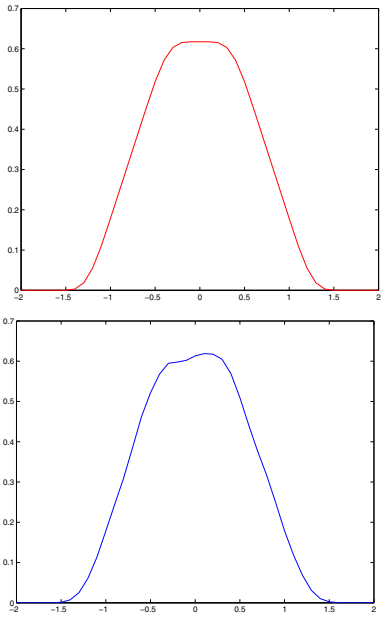

Fig. 1. Actual pdf (above) vs MC simulation (below)

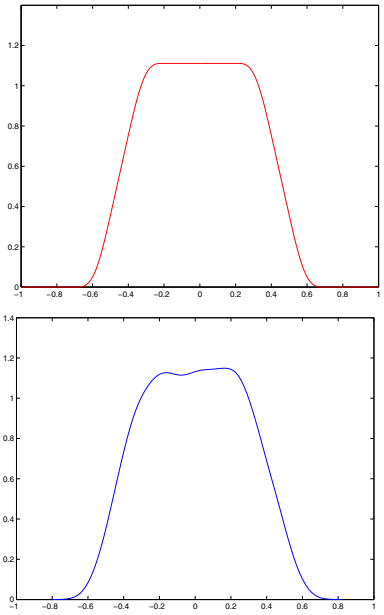

Fig. 2. Actual pdf (above) vs MC simulation (below)

III. Mixed μ

In this section we propose a formulation of the mixed worst case and probabilistic μ problem. Suppose that the uncertainty structure is of the form

$$\mathbf{\Delta} = \begin{pmatrix} \mathbf{\Delta_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta_2} \end{pmatrix},$$

where $\mathbf{\Delta}_1$ is a worst case formulation, and a probability distribution is given on $\mathbf{\Delta}_2$. For a given M and ζ define

$$\mu_\zeta(M) = \Big(\min_{\Delta_1 \in \mathbf{\Delta}_1} \Big\{ \overline{\sigma}(\Delta) : \det(I - \Delta M) = 0, \text{ with: } \Delta = \operatorname{diag}(\Delta_1, \Delta_2(\zeta)) \Big\} \Big)^{-1}. \tag{6}$$

Then the distribution of the mixed μ is defined as the distribution of $\mu_{\mathbf{\Delta_1}}(M)$ defined by, $\mu_{\mathbf{\Delta_1}}(M)(\zeta) = \mu_\zeta(M)$. Given this definition one can consider problems such as calculating contour lines of

$$\mathbb{P}\left(\mu_{\mathbf{\Delta_1}}(M) < \gamma\right).$$

A. Rank-one case:

Suppose that M is a dyad, i.e. $M = uv^T$ for some $u, v \in \mathbb{R}^n$, where M^T denotes the conjugate transpose of M.

Condition 1 Let

$$\mathbf{\Delta} = \operatorname{diag}(\delta_1, ..., \delta_k, \delta_{k+1}, ..., \delta_n)$$

where δ_l-s are independent random variables for $k+1 \leq l \leq n$, such that $\delta_l \sim \text{UNI}[-a_l, a_l]$ with some $a_l > 0$, and $\mathbf{\Delta_1} = \operatorname{diag}(\delta_1, ..., \delta_k)$ is a worst case formulation with $\delta_i \in [-a_i, a_i]$.

We are to calculate the distribution of $\mu_{\mathbf{\Delta_1}}(M)$. Since M is a dyad

$$\mu(M) = \overline{\lambda}(\Delta M).$$

It follows that

$$\mu_\zeta(M) = \sup_{\{\delta_l, l \leq k\}} \left\{ \sum_{l=1}^{k} u_l v_l \delta_l + \sum_{l=k+1}^{n} u_l v_l \delta_l(\zeta) \right\}.$$

Hence,

$$\mu_\zeta(M) = \sum_{l=1}^{k} |u_l v_l| a_l + \sum_{l=k+1}^{n} u_l v_l \delta_l(\zeta).$$

Using our previous result concerning the pure probabilistic case we obtain the probabilistic density function of $\mu_{\mathbf{\Delta_1}}$:

Proposition III.1. *Under Condition 1 the probability density function of the mixed rank-one μ is given by*

$$f_{\mu_{\mathbf{\Delta_1}}}(t) = \left(\prod_{l=k+1}^{n} 2m_l \right)^{-1} \sum_{s \in S} \left[(-1)^{\sum_{k=1}^{n} s_k} \times \operatorname{sgn}\left(t - \sum_{l=1}^{k} m_k + \sum_{l=k+1}^{n} (-1)^{s_l} m_l \right) \times \frac{\left(t - \sum_{l=1}^{k} m_k + \sum_{l=k+1}^{n} (-1)^{s_l} m_l\right)^{n-k-1}}{(n-k-1)!} \right], \tag{7}$$

where $S = \{(s_{k+1}, ..., s_n) : s_k \in \{0, 1\}\}$ *and* $m_k = |u_k v_k a_k|$.

B. Numerical Results

In Fig 3. and Fig 4. the probability density function (pdf) of μ_P is plotted using the previous formula and Monte Carlo methods. Here, we consider the mixed probabilistic-worst case formulation that was defined in the previous section. The actual pdf was calculated using (7) with grid size 0.01, the Monte Carlo simulation generated 30000 random diagonal matrices with $n - l$ elements having uniform distribution on interval $[-1, 1]$. In Fig 1. $u = (0.1, 0.5, 0.7, 0.2, -0.4, -0.7)^T$, $v = (0.4, 0.9, 0.2, -0.3, 1, 0.25)^T$ and $l = 2$. In Fig 2. $u = (1, 0.2, -0.7, 0.45, -0.6, -0.1, .55, -0.2)^T$, $v = (-1, -0.35, 0.37, 0.4, 0.15, -0.3, 0.95, -0.3)^T$ and $l = 3$. Fig 5. was generated using the same u and v as for Fig 1., and one can see the difference between the pdf of the pure probabilistic μ and the pdf of the mixed μ with $l = 3$.

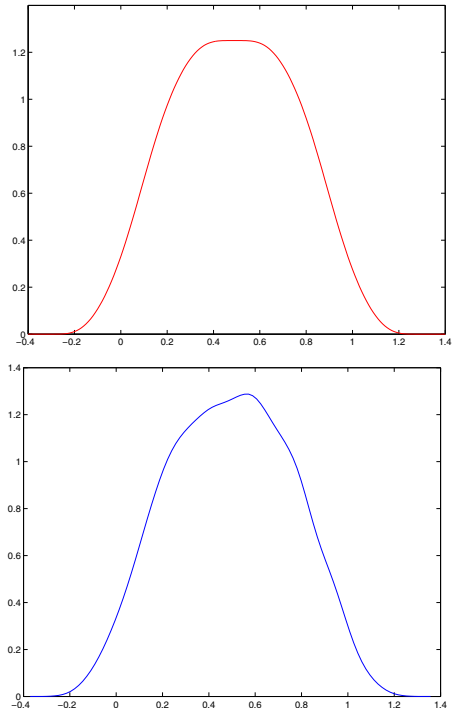

Fig. 3. Actual pdf (above) vs MC simulation (below)

IV. PERTURBED RANK-ONE PROBLEM

Let us suppose again that M is a rank-one matrix, that is $M = uv^T$ for some $u, v \in \mathbb{R}^n$. Perturb M with the diagonal matrix $\delta M = \operatorname{diag}(\delta m_1, \ldots, \delta m_n)$. Our aim is to estimate $\mathbb{P}(\mu_P(M + \delta M) < \gamma)$ using results seen in the previous sections. Now let the distribution of μ_P defined as distribution of $\overline{\lambda}(\Delta M)$. We will use the following theorem.

Theorem IV.1. *[8] Let λ_0 be a simple eigenvalue of $Z_0 \in \mathbb{C}^{n \times n}$, and u_0 be an associated eigenvector. Then a function λ and a vector function u are defined for all Z in some neighborhood $N(Z_0)$ of Z_0 such that*

$$\lambda(Z_0) = \lambda_0, u(Z_0) = u_0,$$

and

$$Zu = \lambda u, u_0^T u = 1, Z \in N(Z_0),$$

Moreover, λ and u are ∞ times differentiable in $N(Z_0)$, and

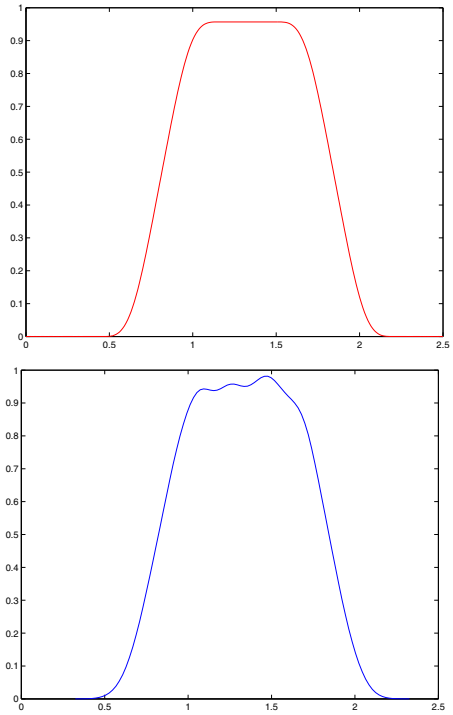

Fig. 4. Actual pdf (above) vs MC simulation (below)

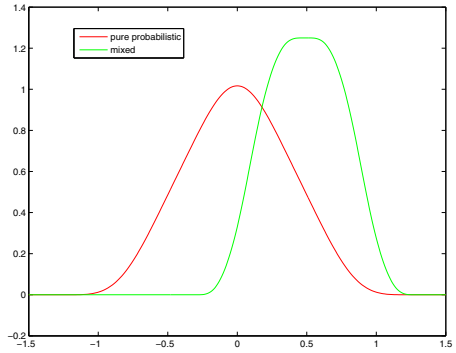

Fig. 5. PDF for the mixed and pure case

the differential at Z_0 is

$$d\lambda = v_0^T (dZ) u_0 / v_0^T u_0, \tag{8}$$

where v_0 is the eigenvector associated with the eigenvalue $\overline{\lambda}_0$ of Z_0^T.

In our case this theorem implies that

$$d\overline{\lambda}(M) = \frac{v^T (dM) u}{v^T u}, \tag{9}$$

because u is the eigenvector associated with $v^T u$ of M, and v is the eigenvector associated with $u^T v$ of M^T. Write

$$\begin{aligned} \overline{\lambda}\left(\Delta(\zeta)(M+\delta M)\right) = \overline{\lambda}\left(\Delta(\zeta) M + \Delta \delta M)\right) = \\ \overline{\lambda}\left(X(\zeta) + \delta X(\zeta))\right), \end{aligned} \tag{10}$$

with $X(\zeta) = \Delta(\zeta) M$ and $\delta X(\zeta) = \Delta(\zeta) \delta M$. Then if δM is sufficiently small, then the following first order approximation can be used

$$\overline{\lambda}(\Delta(\zeta)(M+\delta M)) \approx v^T \Delta(\zeta) u + \frac{v^T (\Delta(\zeta) \delta M) u}{v^T u}, \tag{11}$$

the right hand side reads as

$$v^T \Delta(\zeta)(I + \delta M') u, \tag{12}$$

with $\delta M' = \delta M / (v^T u)$. Note that on the one hand $\Delta(\zeta)(I + \delta M')$ is a diagonal matrix, on the other hand the non-zero entries of the random matrix $\Delta(I + \delta M')$ have uniform distribution. More precisely

$$\begin{aligned} (\Delta(I+\delta M'))_{k,k} \sim \\ \mathrm{UNI}[-a_k(1 + \delta m_k/(v^T u)), a_k(1 + \delta m_k/(v^T u))] \end{aligned} \tag{13}$$

holds. It follows that the value of $\mathbb{P}\left(\mu_P(M+\delta M) < \gamma\right)$ where the uncertainty is a diagonal matrix with its k^{th} element in the diagonal having uniform distribution on $[-a_k, a_k]$ can be approximated with $\mathbb{P}\left(\mu_P(M+\delta M) < \gamma\right)$ where the uncertainty is a diagonal matrix with its k^{th} element in the diagonal having uniform distribution on $[-a_k(1+\delta m_k/(v^T u)), a_k(1+\delta m_k/(v^T u))]$. Since this later problem was solved in the previous section, using II.1 we obtain

$$\begin{aligned} \mathbb{P}\left(\mu_P(M+\delta M) < \gamma\right) \approx \\ F_{\mu_P}(\gamma) - F_{\mu_P}(\gamma_0), \end{aligned} \tag{14}$$

where $\gamma_0 = -\sum_{k=1}^n \left| a_k (1 + \delta m_k / (v^T u)) \right|$ and $F_{\mu_P}(t)$ is the following cdf:

$$\begin{aligned} F_{\mu_P}(t) = \left(\prod_{k=1}^n 2\overline{m}_k \right)^{-1} \sum_{s \in S} \left[(-1)^{\sum_{k=1}^n s_k} \times \right. \\ \left. \operatorname{sgn}\left(\gamma + \sum_{k=1}^n (-1)^{s_k} \overline{m}_k \right) \frac{(\gamma + \sum_{k=1}^n (-1)^{s_k} \overline{m}_k)^n}{n!} \right] \end{aligned} \tag{15}$$

with $S = \{(s_1, ..., s_n) : s_k \in \{0, 1\}\}$ and $\overline{m}_k = \left| a_k v_k u_k (1 + \delta m_k / (v^T u)) \right|$.

V. A NOTE ON MONTE CARLO ALGORITHMS

Several paper deals with the problem of solving probabilistic μ problems using Monte Carlo (MC) algorithms. In [9] axially align cuts have been performed to divide a μ problem into two μ problems. In [3] authors propose to use linear cuts to compute upper bounds. This linear cut algorithm gives an exact solution for rank-one matrices. Here we will study how these type of algorithms reduce the variance of the Monte Carlo simulation. Let us formulate the following problem. Suppose we are given a hypercube $C \subset \mathbb{R}^d$ with volume 1. Let $\mu : C \to \mathbb{R}$, and define

$$D = \{x \in C : \mu(x) < \gamma\}. \tag{16}$$

Our aim is to calculate $\mathrm{Vol}(D)$, i.e. the volume of D. The interesting feature of this problem lies in the fact that if we choose a $S \subset C$ we are able to decide whether $\mu(x) < \gamma$ holds for all $x \in S$. Regions with points satisfying $\mu(x) < \gamma$ will be called *good* regions. Let $p = \mathrm{Vol}(D)$ and $q = \mathrm{Vol}(C - D)$, clearly $p + q = 1$ holds. We are given to

alternatives.

The first alternative is that we do not apply any cuts and implement a Monte Carlo algorithm with N steps. At step k we choose a random point x_k uniformly from C and check if $\mu(x_k) < \gamma$ holds. Define

$$M_N = \frac{1}{N}\sum_{k=1}^{N} \mathbb{I}[\mu(x_k) < \gamma], \tag{17}$$

then the variance of this MC simulation is

$$\mathrm{Var}(M_N) = \frac{1}{N} pq. \tag{18}$$

The second alternative is that at first choose a region $S \subset C$ and check whether it is *good* or not. If we have chosen a good region then we implement a MC algorithm on $C - S$. This way the variance of the MC simulation is given by

$$\frac{1}{N}\Big(p - \mathrm{Vol}(S)\Big)q. \tag{19}$$

Clearly, provided we chose k *good* disjoint regions with volume $V_1, \ldots, V_k$, then we can perform a MC algorithm with variance

$$\frac{1}{N}\Big(p - \sum_{i=1}^{k} V_i\Big)q. \tag{20}$$

This result suggest us that the bigger *good* region we choose the more decrease in variance we achieve. Furthermore, it is not effective to continue the region selection process if only small *good* regions can be chosen.

To compare these two alternatives suppose that the selection of a region is computationally as expensive as the completion of m MC steps. Thus selecting k *good* regions then performing N MC steps is computationally as expensive as performing $N + mk$ MC steps. The corresponding variances are

$$\mathrm{Var}_1 = \frac{1}{N + mk} pq \tag{21}$$

$$\mathrm{Var}_2 = \frac{1}{N}\Big(p - \sum_{i=1}^{k} V_i\Big)q \tag{22}$$

Clearly we choose the alternative with less variance. Obviously these calculations are problem specific, but our results shown in this section may facilitate the selection of the optimal MC method.

VI. Conclusion

We have studied the probabilistic μ problem for rank-one matrices. We obtained analytic results for the probability density function of μ for both the pure probabilistic and the mixed worst-case, probabilistic formulation. We have seen that by using these results we may approximate the distribution of μ for diagonally perturbed rank-one matrices as well. The study of full block perturbations is the subject of our future investigations.

References

[1] X. Zhu Probabilistic μ upper bound using linear cuts *In Proceedings of the 14th IFAC World Congress*

[2] X. Zhu,Y. Huang, and J. Doyle Soft vs. hard bounds in probabilistic robustness analysis *In Proc. IEEE Conf. Decision Control, Kobe, Japan, 1996*

[3] X.Zhu and S.H. Khatri and P.A. Parrilo μ with Linear Cuts: Upper Bound Computation *In Proceedings of the American Control Conference* 1999

[4] X. Chen, K. Zhou, J.L. Aravena On the binomial confidence interval and probabilistic robust control *Automatica* Volume 40, Issue 10, October 2004, Pages 17871789

[5] B. R. Barmish and C. M. Lagoa The uniform distribution: a rigorous justification for its use in robustness analysis *Mathematics of Control, Signals and Systems* 10:203222, 1997

[6] G. Calafiore, F. Dabbene, and R. Tempo Randomized algorithms for probabilistic robustness with real and complex structured uncertainty *IEEE Transactions on Automatic Control* AC- 45:22182235, 2000

[7] B. T. Polyak and P. S. Scherbakov Random spherical uncertainty in estimation and robustness *IEEE Transactions on Automatic Control* AC-45:21452150, 2000.

[8] Jan. R. Magnus On differentiating eigenvalues and eigenvectors *Econometric Theory* 1:179-191, 1985.

[9] S. Khatri and P. A. Parrilo Guaranteed bounds for probabilistic μ *Proceedings of the 37th IEEE Conference on Decision and Control*, 1998