Expected Cost Optimization using Asymmetric Probability Density functions

Bertrand Pigeon*. Michel Perrier*. Bala Srinivasan*

**NSERC Environmental Design Engineering Chair in Process Integration, Department of Chemical Engineering, École Polytechnique de Montréal, C.P. 6079, Succ. Centre Ville, Montréal, Québec, Canada, H3C 3A7, (e-mail: bertand.pigeon@ polymtl.ca)*

Abstract: In the stochastic context, expected value of the cost function is optimized either by changing the mean values of the manipulated variables or by reducing their variance. An extension is to look for an optimal shape for the entire probability density function (PDF). Though the use of asymmetric PDFs is proposed in the literature, no formal proof that justifies their use has been provided. In this paper, it is shown that an asymmetric PDF is required if and only if the cost function is asymmetric and the manipulated variable is penalised. The proof uses an analytical solution of the Fokker-Planck-Kolmogorov equation derived to calculate the shape the output PDF for scalar systems. In particular, this analytical solution is adapted to a switching proportional controller. The theoretical concepts are illustrated on a simulation example, where the advantage of choosing an asymmetric PDF is shown.

Keywords: Stochastic control, Optimization, Probability Density Functions, Switching Algorithms

1. INTRODUCTION

Optimization in a stochastic context involves studying the influence of decisions variables on the expected value of the objective function. In the stochastic context, not only the mean values of the decision variables but the entire distribution plays a role in optimization. Typically, in the presence of constraints, variability is reduced first using appropriate controllers, and secondly by shifting the set point closer to the constraint. Use of minimum variance controllers for optimization purposes has been well studied in the literature (Muske, 2003).

However, shaping the entire probability density function (PDF) could be a viable option to reduce costs. The first mention of this possibility was made in Kàrnỳ (1996). Then, Wang (1998) developed a PDF shaping algorithm based on the weights of a neural B-Spline that parameterized the output PDF. This method has been improved ever since by the same authors (Wang ,2002; Wang & Zhang 2002; Wang & Wang, 2002; Guo & Wang 2005). Crespo and Sun (2002) used an analytical solution of Fokker-Plank-Kolmogorov equation in steady-state to develop a PDF shaping algorithm. On the other hand, Forbes et al. (2004) developed an algorithm based on the parametrization of the target PDF using Gram-Charlier basis functions. In all the above cited works, though the motivation is to improve an optimization objective, only the sub-problem of getting close to a target PDF is addressed. No indication is given on how to compute a target PDF that is suited for the optimization problem at hand.

It has been argued in all the above works that the advantage of PDF shaping lies in shaping it in an asymmetric manner. The necessity of an asymmetric PDF arises from the asymmetry of the objective function. This is normally due to the presence of process and operational constraints. With constraints, typically, an approach based on penalty (barrier) function is used for resolution. An additional cost is added when the constraint is violated (or in the barrier function case an additional cost is added when operated close to the constraint), which inturn causes asymmetry.

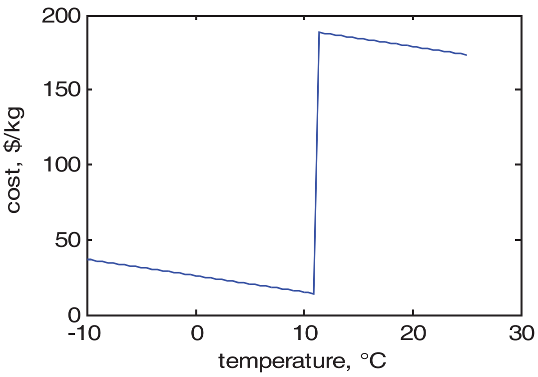

Figure 1: Example of an asymmetric objective function

Figure 1 shows an example with a penalty function where a constant penalty is added if the manipulated variable is above the constraint set at 11°C. As seen, such penalty/barrier functions cause a huge asymmetry around the optimal solution. The optimal solution without any stochastic behaviour would be on the constraint 11°C. However with process noise, a controller needs to be used to reduce the variance of the manipulated variable, and the set point must be lower than 11°C, so that only a small part of the distribution violates the constraint. The minimum variance controller tries to squeeze and shift the distribution towards the constraint. On the other hand, the PDF shaping solution tries to match the asymmetry in the objective function using an asymmetric PDF with its tail on the opposite side of the constraint.

Though intuitive arguments were given for using asymmetric PDFs, no formal results are available to distinguish the cases where an asymmetric PDF would be more beneficial than the symmetric one. So, the main question asked in this paper is, "which class of problems requires an asymmetric PDF?" It is shown that not only the asymmetry of the objective function but also an input weighting is needed to necessitate an asymmetric PDF. The importance of input weighting is one of the core contributions of this paper. In the minimum variance controller, by reducing the variability of the output variable, the variability of the manipulated variables would increase, straining the process equipment. Contrarily, with an asymmetric PDF, the set point can be shifted toward the constraint and with less impact on the manipulated variables.

This paper first presents an analytical solution of the Fokker-Planck-Kolmogorov (FPK) equation for general scalar systems. This analytical solution is then applied to the switching controller case, using which, the optimality or non-optimality of symmetric solution is ascertained. The last section is devoted to a simulation example where the improvement in cost using an asymmetric controller is shown.

2. PROBLEM FORMULATION

2.1Optimization problem formulation

Consider the dynamic system given by equation (1), where u is the scalar manipulated variable, x the scalar state variable, and w the zero-mean Gaussian process noise input with standard deviation η .

$$\dot{x} = f(x) + g(x)u + w \tag{1}$$

The functions $f(x)$ and $g(x)$ represent the unforced and the forced parts of the system dynamics. Consider the optimization of the above system at steady state:

$$\begin{gathered} \min_u \Phi(x,u) \\ C(x,u) \leqslant 0 \\ f(x) + g(x)u = 0 \end{gathered} \quad , \tag{2}$$

where Φ is the function to be optimized, C the constraints. Note that the optimization considers the system equations without noise at steady state as equality constraints.

In the context of this paper, a penalty function is introduced to handle the constraints as show below:

$$\begin{gathered} \min_u [\bar{\Phi}(x,u) + D(C(x,u))] = \phi(x,u) \\ f(x) + g(x)u = 0 \end{gathered} \quad , \tag{3}$$

where $D(.)$ is any appropriate penalty function and $\phi(.)$ the augmented cost. As discussed earlier, $D(.)$ is asymmetric which would lead to an asymmetry in the cost function.

In the context of this paper, x is considered stochastic due to the presence of the noise term w. So, the expectation of the cost function needs to be calculated for optimization purposes. The cost function that is minimized is given by:

$$J = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi(x,u)\, p(x,u)\, dx du \quad , \tag{4}$$

where $p(x,u)$ is the joint probability density function.

2.3 Controllers for PDF shaping

In this section, the nonlinear controller used for PDF shaping is presented. Nonlinearity is crucial since, if the process and the controller were linear, and the input is Gaussian, the output PDF would just be Gaussian.

In order to have a full control on the nonlinearity, all the system nonlinearities are eliminated by feedback linearization. In addition, the controller $h(x)$ is used to bring the state to its desired set point. Then, the controller would introduce the nonlinearities required to shape the PDFs. For the system under consideration, the linearizing feedback is given by:

$$u = \frac{-f(x) + h(x)}{g(x)} \quad . \tag{5}$$

Here, a switching controller of the following form will be studied. The nonlinearity arises from the gain schedule and results in an asymmetrical PDF.

$$h(x) = \begin{Bmatrix} k_1(x_{sp} - x) & if & x < x_{sp} \\ k_2(x_{sp} - x) & if & x > x_{sp} \end{Bmatrix} \quad . \tag{6}$$

To simplify the development, no measurement noise is considered, while a zero-mean Gaussian measurement noise, z, with standard deviation λ will be added to the set point. Thus, the the system reads

$$\begin{gathered} \dot{x} = h(x) + w + k_{cont}(x) z \\ k_{cont}(x) = \begin{Bmatrix} k_1 & if & x < x_{sp} \\ k_2 & if & x > x_{sp} \end{Bmatrix} \end{gathered} \tag{7}$$

3.ANALYTICAL SOLUTION OF THE SCALAR FPK EQUATION FOR SWITCHING CONTROLLER

In this section, the analytical solution of the FPK equation will be developed for the general $h(x)$ and later exploited to suit the switching controller.

3.1 General case

Consider the system (8) where the two random variables are Brownian processes :

$$dx = h(x)dt + \eta\ d\beta_w + k_{cont}(x)\lambda\ d\beta_z, \quad t \geq t_0 \quad , \qquad (8)$$

where η and λ are the standard deviations of the process and measurement noise respectively, $d\beta_w$ and $d\beta_z$ are unit variance Brownian processes. These two noises can be clubbed together into a general equation as follows:

$$dx = h(x)dt + \rho(x)\ d\beta, \quad t \geq t_0 \qquad (9)$$

where $\rho(x)$ represents the agglomerated standard deviation.

The evolution of its probability density function of x is given by the Fokker-Planck-Kolmogorov equation (Jazwinsky (1968)):

$$\frac{\partial p(x,t)}{\partial t} = \frac{-\partial[p(x,t)h(x)]}{\partial x} + \frac{1}{2}\frac{\partial^2[p(x,t)\rho(x)^2]}{\partial x^2} \quad . (10)$$

with boundary conditions $lim_{x\to\infty} p = lim_{x\to-\infty} p = 0$ $lim_{x\to\infty}\frac{\partial p}{\partial x} = lim_{x\to-\infty}\frac{\partial p}{\partial x} = 0$, $\int_{-\infty}^{\infty} p(x)dx = 1$.

At steady state, this equation reads

$$\frac{d\,p\,h}{dx} = \frac{1}{2}\frac{d^2\rho^2 p}{dx^2} \quad . \qquad (11)$$

By integrating both sides of the equation:

$$2\,ph = \frac{d\rho^2 p}{dx} + c \quad . \qquad (12)$$

Using the boundary conditions it can be seen that, $c = 0$. Rearranging the terms gives,

$$\frac{dp}{p} = \left(\frac{2h}{\rho^2} - \frac{2}{\rho}\frac{d\rho}{dx}\right)dx \quad . \qquad (13)$$

The solution of the above equation is given by:

$$p(x) = p_0 e^{\int_{-\infty}^{\infty}\left(\frac{2h}{\rho^2} - \frac{2}{\rho}\frac{d\rho}{dx}\right)dx} \quad , \qquad (14)$$

where p_0 is the normalizing constant to render the integral of the probability to 1.

3.2 Switching controller case

The analytical solution developed in Section 3.1 is applied to a case of the switching controller. Let p_{sp} be the value of the probability density function at $x = x_{sp}$. From (6) and (7) it can be seen that to the left of the set point

$$h(x) = k_1(x - x_{sp}),\ \rho^2(x) = \eta^2 + k_1^2\lambda^2 \quad , \qquad (15)$$

and to the right

$$h(x) = k_2(x - x_{sp}),\ \rho^2(x) = \eta^2 + k_2^2\lambda^{2.} \quad . \qquad (16)$$

Thus, it can be seen that

$$p(x) = \begin{cases} p_{sp}\,e^{\frac{-2k_1(x-x_{sp})^2}{\eta^2+k_1^2\lambda^2}} & \text{for } x < x_{sp} \\ p_{sp}\,e^{\frac{-2k_2(x-x_{sp})^2}{\eta^2+k_2^2\lambda^2}} & \text{for } x \geq x_{sp} \end{cases} \quad . \qquad (17)$$

This can be interpreted as Gaussian function where the two branches are not symmetric. The variance on one side is different from that of the other. The variances on either side can be computed as follows:

$$\sigma_1 = \frac{\sqrt{\eta^2 + k_1^2\lambda^2}}{2\sqrt{k_1}} \quad \text{and} \quad \sigma_2 = \frac{\sqrt{\eta^2 + k_1^2\lambda^2}}{2\sqrt{k_2}} \quad . \qquad (18)$$

Also, the normalisation constant can be computed analytically as follows:

$$p_{sp} = \frac{2}{\sqrt{2\pi}(\sigma_1 + \sigma_2)} \quad . \qquad (19)$$

From the expression of $h(x)$ it can also be shown that

$$p(h) = \begin{cases} p_{h0}\,e^{\frac{-2h^2}{k_1(\eta+k_1\lambda)^2}} & \text{for } h < 0 \\ p_{h0}\,e^{\frac{-2h^2}{k_2(\eta+k_2\lambda)^2}} & \text{for } h \geq 0 \end{cases} \qquad (20)$$

with

$$\sigma_{1h} = \frac{\sqrt{\eta^2 + k_1^2\lambda^2}\sqrt{k_1}}{2} \quad , \quad \sigma_{2h} = \frac{\sqrt{\eta^2 + k_2^2\lambda^2}\sqrt{k_2}}{2} \qquad (21)$$

$$p_{h0} = \frac{2}{\sqrt{2\pi}(\sigma_{1h} + \sigma_{2h})} \quad . \qquad (22)$$

4.NON-OPTIMALITY OF THE SYMMETRIC SOLUTION

In this section, it is be shown that a symmetric PDF is sufficient even for an asymmetric objective function, when there is no input weighting. Also, when the objective function is symmetric, with or without input weighting a symmetric PDF is indeed optimal. However, when there is asymmetry and input weighting, then it is shown that a symmetric solution is not optimal.

Consider equation (4). Since, u is a function of x, the objective function $\phi(x,u)$ is just a function of x. In particular, consider a special case where the squared deviation of the control action $h(x)$ is included in the cost function. The remaining part of the objective function is termed $l(x)$. So,

$$\phi(x,u) = l(x) + \gamma h^2(x) \qquad (23)$$

Due to the imposed control structure, the degree of freedom for the optimization problem is no longer u, but the parameters x_{sp}, k_1 and k_2. So, the optimization problem reads,

$$\min_{x_{sp},k_1,k_2} J=\int_{-\infty}^{+\infty} l(x)\,p(x)\,dx+\int_{-\infty}^{+\infty} \gamma h^2\,p(h)\,dh \quad . \tag{24}$$

The proof of non-optimality proceeds by deriving the necessary conditions of optimality of the above optimization problem by considering that k_1 and k_2 are varied independently. Then an additional condition of symmetry, i.e. $k_1 = k_2$ is imposed. This gives four conditions (3 necessary conditions and one condition of symmetry) for three variables. If these four conditions are consistent then the symmetric solution is indeed optimal. On the other hand, if it leads to an inconsistency or contradiction then it shows that the symmetric solution is not optimal in the case considered.

Theorem 1: The symmetric switching controller is locally optimal if and only if (i) l(x) is symmetric around the optimum, i.e., the third derivative evaluated at the optimum is zero, or (ii) the input weighting γ is zero.

Proof: Without loss of generality let $x = 0$, $l(x) = 0$, $J = 0$ be the optimum in the absence of noise. Consider the third order Taylor series expansion of $l(x)$ around $x = 0$. The first two terms are zero since $l(0) = 0$ and the first derivative is zero due to optimality. Thus the expansion is given by

$$l(x)=\alpha x^2+\delta x^3 \quad , \tag{25}$$

where α and δ are the second and third derivatives, respectively, at the origin. The expected cost (5) is then given by,

$$J=\alpha\int_{-\infty}^{+\infty} x^2\,p(x)\,dx+\delta\int_{-\infty}^{+\infty} x^3\,p(x)\,dx+\gamma\int_{-\infty}^{+\infty} h^2\,p(h)\,dh \quad . \tag{26}$$

Analytical expressions for all the three terms can be obtained.

$$\int_{-\infty}^{+\infty} x^2\,p\,dx=x_{sp}^2-\frac{4\,x_{sp}(\sigma_1-\sigma_2)}{\sqrt{2\pi}}+(\sigma_1^2-\sigma_1\sigma_2+\sigma_2^2) \tag{27}$$

$$\int_{-\infty}^{+\infty} x^3\,p\,dx=x_{sp}^3-\frac{6\,x_{sp}^2}{\sqrt{2\pi}}(\sigma_2-\sigma_1)+3\,x_{sp}(\sigma_2^2-\sigma_1\sigma_2+\sigma_1^2) +\frac{4}{\sqrt{2\pi}}(\sigma_2-\sigma_1)(\sigma_1^2+\sigma_2^2) \tag{28}$$

$$\int_{-\infty}^{+\infty} h^2\,p(h)\,dh=(\sigma_{h1}^2-\sigma_{h1}\sigma_{h2}+\sigma_{h2}^2)-\frac{2}{\pi}(\sigma_{h1}-\sigma_{h2})^2 \tag{29}$$

The optimality condition requires that the derivatives of J with respect to x_{sp}, k_1 and k_2 be zero, the expressions for which can be readily obtained. To analyse the symmetric solution, consider $k_1 = k_2 = k$. Substituting this in the derivatives leads to

$$\frac{\partial J}{\partial x_{sp}}=2\alpha x_{sp}+3\delta x_{sp}^2+\frac{3\delta}{4k}(\eta^2+k^2\lambda^2)=0 \quad , \tag{30}$$

$$\frac{\partial J}{\partial k_1}-\frac{\partial J}{\partial k_2}=\frac{(\eta^2-k^2\lambda^2)}{\sqrt{2\pi k^3}}\left(2\alpha x_{sp}+3\delta x_{sp}^2+\frac{\delta}{k}(\eta^2+k^2\lambda^2)\right)=0 \quad , \tag{31}$$

and

$$\frac{\partial J}{\partial k_1}+\frac{\partial J}{\partial k_2}=\frac{\eta^2-k^2\lambda^2}{4k^2}(\alpha+3\delta x_{sp})+\frac{\gamma}{4}(\eta^2+3k^2\lambda^2)=0 \quad . \tag{32}$$

It can be seen that there are 3 equations for 2 unknowns, k and x_{sp}. Replacing the terms with x_{sp} in (31) using (30), it can be seen that

$$\frac{\partial J}{\partial k_1}-\frac{\partial J}{\partial k_2}=\frac{(\eta^2-k^2\lambda^2)}{\sqrt{2\pi k^3}}\frac{\delta}{4k}(\eta^2+k^2\lambda^2)=0 \quad . \tag{33}$$

Only if part: $\delta\neq 0, \gamma\neq 0 \Rightarrow non-optimality$

When $\delta\neq 0$, the only solutions of (33) are $k=\pm\eta/\lambda$ But, plugging these values of k in the sum of derivatives lead to $\gamma=0$. So, if $\gamma\neq 0$ the symmetric controller is not optimal .

If part: $\gamma=0 \Rightarrow optimality$

When $\gamma=0$ note that $k=\pm\eta/\lambda$ satisfies all the three necessary conditions of optimality .

If part: $\delta=0 \Rightarrow optimality$

Since $\delta=0$, (31) gives $x_{sp} = 0$. (33) is not useful in determining k. However from (32), it can be seen that the following 4th order equation can be used to compute k.

$$3\gamma\lambda k^4+(\gamma\eta+\lambda\alpha)k^2-\alpha\eta=0 \tag{34}$$

■

5. EXAMPLE

In this section, an asymmetric example with input weighting is presented. The optimal switching controller is computed using the output PDF obtained through the analytical solution. It will be shown that such a controller indeed leads to an asymmetric PDF.

A cost function analogous to the one in Figure 1 is considered here.

$$\begin{gathered}\phi(x)=26-10x+D(c(x))+10h^2(x)\\ D(c(x))=0 \;\; if\; x\leq 11\\ D(c(x))=10^5 \;\; if\; x>11\end{gathered} \tag{35}$$

The system dynamics is given by

$$\dot{x}=-0.4x+0.2u+w \quad , \tag{38}$$

where the process noise w has a mean of 0 and a standard deviation $\eta=1$. A measurement noise of standard deviation $\lambda=0.01$ was considered. Though it is unrealistic to consider a ratio of 100 between the standard deviations of process and measurement noises, it is required in this case to prove the principle. The asymmetric PDF gives better results only in a narrow range of parameter values and so is such a choice made.

5.1 Controller design

The controller (6) is used here. It has 3 parameters; gains k_1, k_2 and the set point x_{sp}. These parameters are found via non linear programming where the equation (24) is minimized. Equation (24) for the given example can be written as follows:

$$J = 26 - 10 \int_{-\infty}^{+\infty} x\, p(x)\, dx + 10^5 \int_{11}^{+\infty} p(x)\, dx + 10 \int_{-\infty}^{+\infty} v^2\, p(v)\, dv \tag{39}$$

Also, in this case an analytical expression for all the three terms can be derived using *p(x)* given in (20). The analytical expression of the last term is already provided in (31). The expressions for the other terms are given as follows:

$$\int_{c}^{+\infty} p(x)\, dx = \frac{\sigma_2}{\sigma_1 + \sigma_2} \left(1 + \mathrm{erf}\left(\frac{x_{sp} - c}{\sqrt{2}\,\sigma_2} \right) \right) \tag{48}$$

$$\int_{-\infty}^{+\infty} x\, p(x)\, dx = x_{sp} + \sqrt{\frac{2}{\pi}} (\sigma_2 - \sigma_1) \quad . \tag{49}$$

5.2 Results

The optimal parameters for a switching controller and a constant gain control have been found numerically. For calculating the optimal single-gain controller, the same calculations are used with $k_1 = k_2$. The optimal gains and the value of the cost function are presented in Table 1. It can be seen that with the switching controller, the cost is reduced by around 6.7%. It is because by having 2 gains, the controller can be aggressive on one side, the side of the constraint, while having a low gain and thereby low input variance on the other side.

Table 1: Results of the example

Controller type	Switching controller	Single gain controller
Set point	10.6	10.53
k_1	0.41	2.94
k_2	4.99	2.94
Cost	5.39	5.78

Figure 2 shows the output PDF for the both controllers. It can be seen that the single proportional controller leads to a symmetric Gaussian PDF, while with the switching controller results in an asymmetric PDF. It is equally interesting to see in Figure 3 that the asymmetry in the input PDF is reversed.
It can be explained by the fact that closer to the constraint, the input works hard and has a larger variance, while far from the constraint, the input does not work in order to reduce the cost by decreasing its variance.

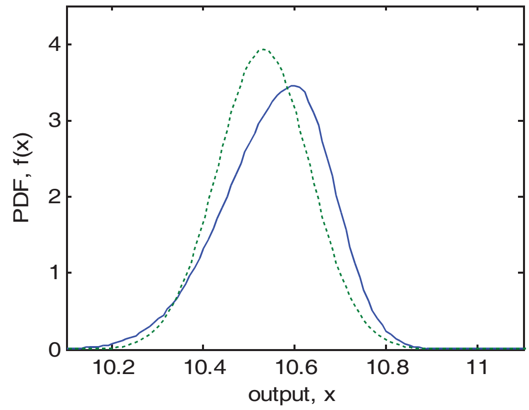

Figure 2: Output PDF with a switching controller (solid line) and with a non-switching controller (dotted line)

Several tests were performed with varying penalties, with varying input weights, and varying measurement noise levels. Figure 4 shows the effect changing the penalty. It can be seen that increasing the weighting for the penalty increases the difference between the cost functions of the symmetric and asymmetric PDF. This tendency can be attributed to the fact that increasing the penalty increases the asymmetry of the cost function. Note that the x axis is logarithmic, i.e, a small increase in the difference calls for a order of magnitude change in the weighting.

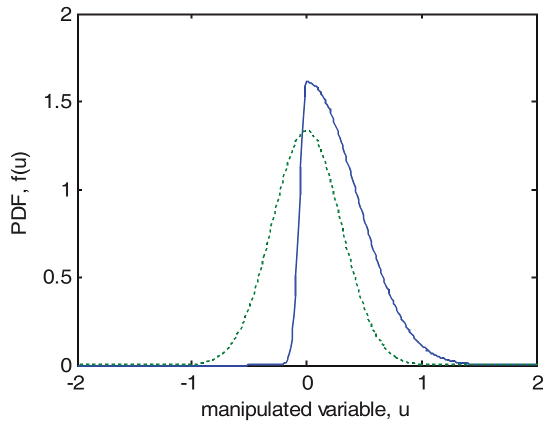

Figure 3: Manipulated variable PDF with a switching controller (solid line) and with a non-switching controller (dotted line)

Figure 5 shows the effect of changing the input weight. An interesting effect can be observed here. The difference first increases, while it decreases after reaching a maximum. Intuitively, when the input weight is zero, the symmetric solution is indeed optimal and there can be no gain by using an asymmetric controller. On the other hand, since the input weighting is symmetric, for large input weightings the asymmetry of the cost function becomes negligible and so a symmetric controller is again optimal.

Figure 6 shows the influence of measurement noise on the difference. The larger the measurement noise, lesser is the gain that can be obtained by using an asymmetric PDF. This is due to the fact that with increasing measurement noise the minimum variance controller as such has a fairly low gain and not much manoeuvrability is left.

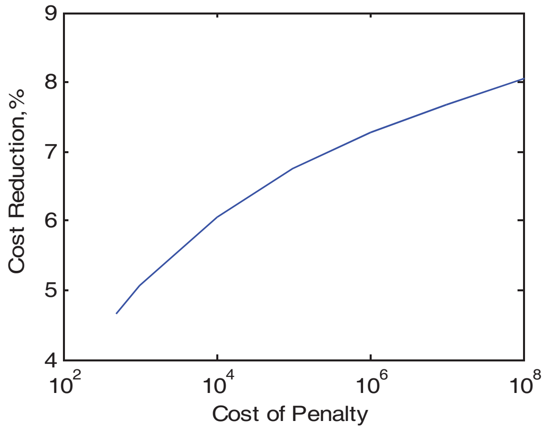

Figure 4: Effect of the weighting of the penalty on the cost reduction due to switching controller

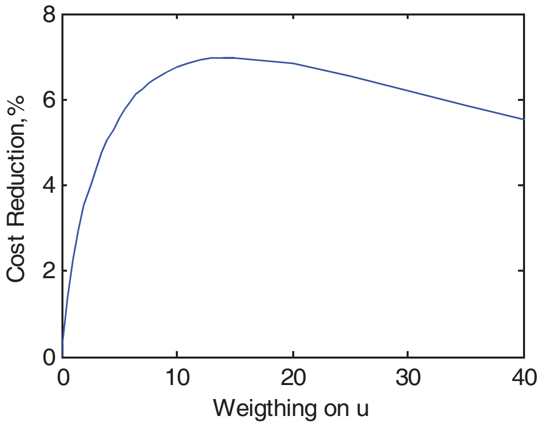

Figure 5 Effect of the input weighting on the cost reduction due to switching controller

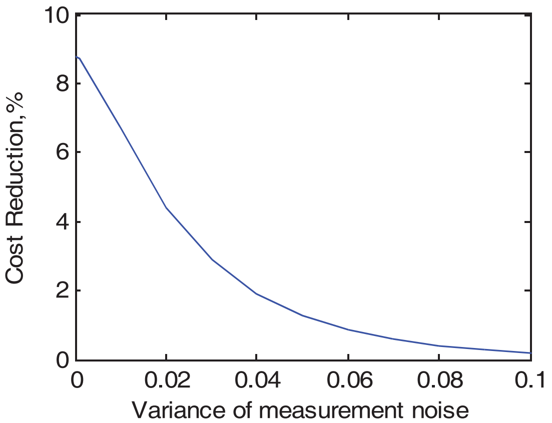

Figure 6: Effect of the measurement noise on the cost reduction due to switching controller

6. CONCLUSION

This paper showed the non-optimality of a symmetric PDF when the cost function was asymmetric and the manipulated variable was constrained. The result is derived using the analytical solution of the FPK equation for a scalar system and a switching controller. Finally, a numerical example was shown where the asymmetric PDF gave a better result than the symmetric one.

The importance of this result lies in the fact that it clearly demarks the cases where an asymmetric PDF is required. Also, a simple switching controller structure for PDF shaping is proposed that can be easily implemented in an industrial context. Finally, the analytical solution of the FPK equation is not only limited to PDF shaping, but could have more impact in the general context of stochastic optimization.

ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Environmental Design Engineering Chair at École Polytechnique and Le Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT).

REFERENCES

Crespo, L.G., Sun, J.Q. (2002). Nonlinear Stochastic Control via Stationary Probability Density Functions. Proceedings of the American Control Conference. Volume 3. 2029-2034

Forbes, M.G. , Guay M., Forbes, J.F. (2004). Control Design for First-order Processes: Shaping the Probability Density of the Process State. *Journal of Process Control.* 14(4). 399-410

Guo, L., Wang, H. (2005). Generalized discrete-time PI control of output PDF using square root B-spline expansion. *Automatica.* 41(1). 159-162

Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory.* New-York: Dover Publications.376 p.

Kàrnỳ, M. (1996). Towards Fully Probabilistic Control Design. *Automatica.* 32(12). 1719-1722

Muske, K.(2003). Estimating the Economic Benefit from Improved Process Control. *Industrial Engineering Chemical Research.* 42(20). 4535-4544

Wang, H. (1998). Robust Control of the Output Probability Density Functions for Multivariable Stochastic Systems. *Proceedings of the 37st Conference on decision on control.* volume 2. 1305-1310

Wang, H. (2002). Minimum Entropy Control of Non-Gaussian Dynamic Stochastic Systems. *IEEE Transactions on Automatic Control.* 47(2). 398-403

Wang, H., Zhang, J. H. (2002). Control of the Output Stochastic Distributions via Lyapunov Function Analysis. *Proceedings of the 2002 IEEE International Conference on Control Applications.* 927-931.