# Linear Gaussian Quadratic Regulation under Poisson Distributed Intermittent State Observations

by

Adès, M.*, Caines, P.E.*, and Malhamé, R.P.**
* Electrical Engineering Department, Mc Gill University
Montréal, Québec, Canada
Electrical and Computer Engineering Department,
École Polytechnique de Montréal,
Montréal, Québec, Canada

## I. Abstract

The optimal control problem of linear stochastic continuous-time systems with a finite budget of N exact full state observations interspersed at Poisson distributed instants of time is considered. The length of the control horizon is random. The horizon lies at the instant where the last observation is gathered.

It is shown that this problem has intimate connections with that of linear quadratic Gaussian regulation with an exponentially discounted cost. Also, the optimal control is shown to be made up of a sequence of piecewise open loop controls corresponding to linear feedback of the state predictors based on the most recent information. The feedback gains are piecewise-constant and are obtained from a sequence of algebraic Riccati equations. They are computed off-line.

## II-Introduction

The solution to the problem of optimally controlling a discrete or continuous-time stochastic Gaussian (finite-state) dynamic system under full or partial state observation is by now very well known (see[1] or [2] for example). In the existing literature, state observations whether noisy or exact, partial or total, occur either continuously or at regularly spaced intervals of time.

In this paper, we consider instead the case where observations, while corresponding to exact full-state observations, occur at random intervals of time which are Poisson distributed with a given fixed mean interarrival time. The controller is allowed a finite budget of N observations. The length of the control horizon is random as it ends with the gathering of the $N^{th}$ observation.

Interest in this problem stems from at least two practical control situations; both situations involve reliance on a highly trained technician with expensive hourly rates for manual measurements of the controlled process, and constraints on the total measurements budget over the finite control horizon. However, in the first situation, while regularly spaced measurements are planned for initially, it is the technician who has random availability, thus introducing Poisson distributed measurement times of the process. In the second situation, while more control can be exerted over the technician's availability, it is desired to investigate the appropriateness of randomizing observation times of the process, while keeping the total mesurements budget fixed. Indeed, one can legitimately ask whether the choice of random independant interarrival observation times with a given fixed mean and maximum variance may do better from the point of view of overall control performance than regularly spaced fixed observation times for a fixed total number of observations. In particular, one such interarrival observation time distribution is exponential, thus making the resulting observation times process Poisson. We note that, in principle, the comparison of the performance of this latter scheme with the original fixed observation time scheme may be carried out using the results of this paper.

The remainder rest of the paper is organized as follows. In section 3, we formulate the linear Gaussian quadratic regulator problem with a finite budget of Poisson distributed (full state) observation times. In section 4, and in order to limit complexity ,we present results only for the scalar case. A dynamic programming framework is developed whereby it is shown that at stage i (i.e. following the $i^{th}$ observation), the structure of the optimal control is obtained from the solution of an infinite horizon linear Gaussian exponentially weighted quadratic regulator problem, the parameters of which are obtained as the result of a backward recursion from the parameters at stage i+1. The solution corresponds to a sequence of piecewise constant gain feedback controls linear, not in the system state per se, but in the predictor of that state based on the most recent observation and the calculated control structure. Note that all feedback gains can be computed off line, thus making the overall controller a gain scheduling type of controller. In section 5, conclusions and suggestions for future work are summarized.

## III-Problem statement

We consider a linear time-invariant stochastic system evolving according to the following Itô stochastic state equation:

$$dx(t) = A\, dt + B\, u(t)\, dt + G\, dw(t) \qquad (1)$$

where x,u are respectively nx1 and mx1 vectors, and w is a normalized zero mean standard vector brownian motion, while A, B, and G are matrices of the appropriate dimensions.

We associate with (1) a full-state, exact, observation equation:

$$z(t_i) = x(t_i) \ , \ i=0,1,2 \ldots\ldots ,N \qquad (2)$$

where the $t_i$'s are Poisson distributed with mean interarrival time $1/\lambda$. Note that we set $t_0 = 0$.

Under this randomized observation structure, the objective is to construct a control law which optimizes the following quadratic expected cost functional:

$$J(x(t_0)) = E_{W,t_N} \int_{t_0}^{t_N} [\, ( x'Q\, x + u'R\, u )\, dt \mid x(t_0)] \ , \qquad (3)$$

where Q, R are symmetric and respectively positive definite, and positive semi-definite matrices of the appropriate sizes. Note that for notational simplicity, we shall consider that knowing $x(t_i)$ corresponds in fact to knowledge of both $x(t_i)$ and $t_i$, that is to say, the $i^{th}$ observation is $(t_i, x(t_i))$.

In the next section, we develop a dynamic programming solution to this optimal control problem in the scalar case.

## IV-Optimal control synthesis

Consider the following scalar version of (1):

$$dx(t) = a\, dt + b\, u(t)\, dt + g\, dw(t) \ , \qquad (4)$$

with the randomized observation equation:

$z(t_i) = x(t_i)$, where the $t_i$'s, $i = 0, 1, \ldots, N$ are as in (2) above.

It is desired to construct the control law, within the class $\mathcal{B}$ of admissible control laws (defined to be that of control functions measurable with respect to the σ-field of observations, and yielding a finite performance index), which minimizes the following cost functional:

$$J(x(t_0)) = \min_{u \in \mathcal{B}} E_{W,t_N} [ \int_{t_0}^{t_N} (x^2(t) + r\, u^2(t))dt \mid x(t_0)] \ ,$$
$$r > 0 \, . \qquad (5)$$

We consider a dynamic programming formulation of the optimization problem, whereby the $i^{th}$ stage starts with the occurence of the $i^{th}$ measurement, and the optimal cost-to-go for the $i^{th}$ stage is given by:

$$V(x(t_i), i) \triangleq V(x(t_i), t_i ,i)$$

$$= \min_{u \in \mathcal{B}} E_{W,t_N} [ \int_{t_i}^{t_N} (x^2(t) + r\, u^2(t))\, dt \mid x(t_i) ] \qquad (6)$$

Notice that the optimal cost-to-go associated with the $N^{th}$ stage is given by:

$$V(x(t_N), N) = 0 \qquad (7)$$

At the $(N\text{-}1)^{th}$ stage, the cost is given by:

$$V(x(t_{N-1}), N\text{-}1) = \min_{u \in \mathcal{B}} E_{W,t_N} [ \int_{t_{N-1}}^{t_N} (x^2(t) + r\, u^2(t))\, dt$$
$$+ V(x(t_N), N)] \mid x(t_{N-1}) ]$$

$$= \min_{u \in \mathcal{B}} E_{W,t_N} [ \int_{t_{N-1}}^{t_N} (x^2(t) + r\, u^2(t))\, dt \mid x(t_{N-1}) ] \qquad (8)$$

Now, upon conditioning on $t_N$ in (8), and using the independence of the w(t) process and the sequence of $t_i$'s, we can write:

$$V(x(t_{N-1}), N\text{-}1) = \min_{u \in \mathcal{B}} E_W \int_{f_{N-1}}^{\infty} \lambda e^{-\lambda(t_N - t_{N-1})} [ \int_{t_{N-1}}^{t_N} (x^2(t)$$
$$+ r\, u^2(t))\, dt]\, dt_N \mid x(t_{N-1})] \qquad (9)$$

Using Fubini's theorem in (9), one can interchange the order of the double integration to obtain:

$$V(x(t_{N-1}), N-1) = \min_{u \in \mathcal{B}} E_w\left[ \int_{t_{N-1}}^{\infty} [(x^2(t) + r\,u^2(t)] \int_t^{\infty} \lambda e^{-\lambda(t_N - t_{N-1})} dt_N \, dt \,|x(t_{N-1})\right]$$

$$= \min_{u \in \mathcal{B}} E_w\left[ \int_{t_{N-1}}^{\infty} e^{-\lambda(t_N - t_{N-1})} (x^2(t) + r\,u^2(t))\, dt \mid x(t_{N-1})\right] , \quad (10)$$

which is effectively an infinite horizon dicounted linear quadratic regulator problem, with perfect knowledge of the initial state, but no further observations. The resulting optimal control law is well known [4], and by a separation theorem [2], corresponds to a linear state feedback $u = -k_{N-1}\,\hat{x}(t|t_{N-1})$, where $\hat{x}(t|t_{N-1})$ is the predictor of x(t) ,given the initial state $x(t_{N-1})$, with the following dynamics:

$$d\hat{x}(t \mid t_{N-1}) = (a - k_{N-1})d\hat{x}(t \mid t_{N-1}). \quad (11)$$

The gain $k_{N-1}$ is obtained from the following algebraic Riccati equation:

$$r^{-1}b^2k_{N-1}^2 - (2a + \lambda)k_{N-1} - 1 = 0 \quad (12)$$

Thus, the resulting closed-loop system evolves according to:

$$Dx(t) = (a\,x(t) - k_{N-1}\,\hat{x}(t|t_{N-1}))dt + g\,dw(t)$$

$$d\hat{x}(t \mid t_{N-1}) = (a - k_{N-1})\,\hat{x}(t|t_{N-1})dt , \quad (13)$$

with the following optimal cost functional:

$$V(x(t_{N-1}), N-1) = E_w\left[ \int_{t_{N-1}}^{\infty} e^{-\lambda(t - t_{N-1})} [x(t) u(t)] \begin{bmatrix} 1 & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \,\middle|\, x(t_{N-1}) \right]$$

$$= E_w \int_{t_{N-1}}^{\infty} \left[ \text{trace}\, e^{-\lambda(t - t_{N-1})} \begin{bmatrix} 1 & 0 \\ 0 & k_{N-1}^2 t^2 \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t \mid t_{N-1}) \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t \mid t_{N-1}) \end{bmatrix}' \,\middle|\, x(t_{N-1}) \right] dt \quad (14)$$

In order to compute the above integrals, we rely on the following identity:

$$E[(X\,X')] = P(t) + E(X)\,E(X)' \quad (15)$$

where X is the state vector in (13), $P(t) = E[(X-E(X))(X-E(X))']$, as well as the following covariance propagation equation [3]:

$$\dot{P}(t) = A\,P(t) + P(t)\,A' + G\,G' , \quad P(0) = 0 \quad (16)$$

while A and G are the A matrix in (13), and the [g 1]' vector respectively.

Notice that the cost in (14) can now be expressed as:

$$V\,x(t_{N-1}, N-1) = \int_{t_{N-1}}^{\infty} e^{-\lambda(t - t_{N-1})} \begin{bmatrix} P_{11} + (k_{N-1}t)^2 P_{22} \\ + \hat{x}^2(t|t_{N-1})(1 + 1_{N-1}r)^2 \end{bmatrix} dt \quad (17)$$

Furthermore, Laplace transformation of (16) yields:

$$sP(s) - P(0) = A\,P(s) + P(s)\,A' + GG' \quad (18)$$

and by setting $s = \lambda$ , one can compute the integrals in (17).

(15) - (18) yield after appropriate calculations:

$$V(x(t_{N-1}), N-1) = (\lambda - 2(a - k_{N-1}))^{-1}[1 + (k_{N-1}^2 r)]x(t_{N-1}) + \frac{g^2}{\lambda - 2a}$$

$$= \alpha_{N-1} x^2(t_{N-1}) + \beta_{N-1} \quad (19)$$

At this stage, we postulate that the cost structure will remain quadratic in the initial state, and establish that result through a backwards recurrence.

Thus, let:

$$V(x(t_{i+1}), i+1) = \alpha_{i+1}\;x(t_{i+1}) + \beta_{i+1} , \quad (20)$$

for some $0 \le i < N$, and given constants $\alpha_{i+1}$, $\beta_{i+1}$.

We show that:

$$V(x(t_i), i) = \alpha_i\,x^2(t_i) + \beta_i , \quad (21)$$

For some constants $\alpha_i$, $\beta_i$.

Indeed,

$$V(x(t_i), i) = \min_{u \in \mathcal{B}} E_{W,t_{i+1}} [ \int_{t_i}^{t_{i+1}} (x^2(t) + ru^2(t))dt + V(x(t_{i+1}), i+1) \mid x(t_i)]$$

$$= \min_{u \in \mathcal{B}} E_W [ \int_{t_i}^{\infty} \lambda\, e^{-\lambda(t_{i+1}-t_i)} \int_{t_i}^{t_{i+1}} (x^2(t) + ru^2(t))dt_{i+1} + \alpha_{i+1}\, x^2(t_{i+1}) + \beta_{i+1} \mid x(t_i)]$$

$$= \min_{u \in \mathcal{B}} [ E_W [ \int_{t_i}^{\infty} [x^2(t) + ru^2(t)] [ \int_{t}^{\infty} \lambda\, e^{-\lambda(t_{i+1}-t_i)}\, dt_{i+1}] dt \mid x(t_i)] + E_W[\alpha_{i+1} \int_{t_i}^{\infty} \lambda\, e^{-\lambda(t-t_i)}\, x^2(t)dt + \beta_{i+1} \mid x(t_i)]]$$

$$= \min_{u \in \mathcal{B}} E_W [ \int_{t_i}^{\infty} \lambda\, e^{-\lambda(t-t_i)} [x^2(t)[1 + \alpha_{i+1}\lambda] + ru^2(t)]dt + \beta_{i+1} \mid x(t_i)] \qquad , \qquad (22)$$

where, once again, one recognizes an infinite horizon discounted linear quadratic Gaussian optimal control problem with perfect knowledge of the initial state and no further observation .

The optimal control is again a linear feedback on the predictor $\hat{x}(t|t_i)$ of the state under the control structure . Thus,

$$u(t) = -k_i\, \hat{x}(t \mid t_i) \qquad , \qquad (23)$$

with :

$$d\hat{x}(t \mid t_i) = (a - k_i)\, \hat{x}(t \mid t_i)dt \qquad , \qquad (24)$$

and $k_i$ is obtained from the following Riccati equation:

$$0 = r^{-1}\, b^2\, k_i^2 - (2a + \lambda)k_i - (1 + \alpha_{i+1}\lambda)$$

The resulting cost functional is then given by:

$$V(x(t_i), i) = E_W [ \int_{t_i}^{\infty} \lambda e^{-\lambda(t-t_i)} [x^2(t)\, [1 + \alpha_{i+1}\lambda] + ru^2(t)]dt + \beta_{i+1} \mid x(t_i)]$$

$$= E_W[ \int_{t_i}^{\infty} \lambda e^{-\lambda(t-t_i)} [ [ (1 + \alpha_{i+1}\lambda) + r\,k_i^2 ]\; x^2(t) ]dt \mid x(t_i) ] + \beta_{i+1}$$

$$= [ -2(a-k_i) ] [ 1 + \alpha_{i+1}\lambda + r\,k_i^2 ]\, x^2(t_i) + \beta_{i+1} + \frac{g^2}{\lambda - 2a} \qquad (25)$$

Thus:

$$\alpha_i = [ -2( a-k_i )^{-1} ] [ 1 + \alpha_{i+1}\lambda + r\,k_i^2 ]$$

$$\beta_i = \beta_{i+1} + \frac{g^2}{\lambda - 2a} \quad \text{for} \quad i = 0, ...., N-1$$

with:

$$\alpha_N = 0$$
$$\beta_N = 0 \quad ,$$

furthermore,

$$k_i^2 - \frac{(2a+\lambda)}{b^2} rk_i - \frac{(1+\alpha_{i+1}\lambda)}{b^2} r = 0 \qquad (26)$$

and

$$v(x(t_i), i) = \alpha_i\, x^2(t_i) + \beta_i \quad . \qquad (27)$$

This establishes our main result.

## V-Conclusion

We have developed results on linear quadratic regulation with intermittent Poisson distributed observations with a finite number of observations N. The behavior of the optimal controller as N goes to infinity is an issue remaining to be explored. Also, the results lead to further questions, among which the following one : given a finite budget of N observations which should on average span a fixed length of horizon T, with some given maximal relative variance, is there an optimal (fixed) distribution of observation times which will minimize the resulting cost functional (and would therefore be most advisable)?

## VI-References


[1] P.E. Caines, Linear Stochastic Systems . New York : Wiley and Sons, 1988.

[2] A. Bagchi, Optimal Control of Stochastic Systems . New York : Prentice-Hall, 1993.

[3] A. Gelb, Applied Optimal Estimation. Cambridge, Massachussets : M.I.T. Press, 1974.

[4] D.P. Heyman, M.J. Sobel, Handbook in Operations Research and Management Science, North-Holland, 1990.