SYMPLECTIC CAMELS AND UNCERTAINTY ANALYSIS

ANDREA CENSI

Abstract. Gromov's non-squeezing theorem is a relatively new results in symplectic topology that provides a non-trivial constraint on the propagation of uncertainty in a Hamiltonian system. While the application to nonlinear systems is difficult, one can apply the theorem to obtain bounds on the propagation of uncertainty in the *linearized* system, by exploiting the fact that the flow of the linearization of a Hamiltionian system is still symplectic. This report compares the results that one obtains by this linearized analysis to the actual behavior of the nonlinear system.

1. Introduction

Consider a probability distribution on the phase space of a Hamiltonian system. The evolution of the distribution through the flow of the system must respect various classical constraints. The first obvious one is that the overall probability mass must be conserved. Moreover, the fact that a Hamiltonian vector field is divergence free implies two other constraints:

(1) The volume of a finite set is invariant along the flow (Liouville's theorem).
(2) The probability density is constant along the flow ($p(\varphi_t(z)) = p(z)$).

An application of Gromov's non-squeezing theorem gives an additional non-trivial constraint, which is based on the fact that the flow of a Hamiltonian system is a symplectic map. Intuitively, a volume in phase space cannot be stretched with respect to one particular symplectic plane more than its "symplectic width" allows. In other words, it is impossible to squeeze a symplectic camel into the eye of a needle, if the needle is small enough. This is a very powerful result, which is intimately tied to the Hamiltonian nature of the system, and is a completely different result than Liouville's theorem, which only interests the overall volume and does not pose any restriction on the *shape.*

Recently, it has been shown that using this theory, practical constraints can be obtained for the linearized propagation of uncertainty in Hamiltonian systems [1, 2]. This paper summarizes these results, and compares them with a numerical nonlinear analysis. We proceed as follows. Section 2 states Gromov's non-squeezing theorem and the concept of symplectic width. Section 3 applies the theory to the analysis of the linearized flows. Section 4 presents some simulations. Conclusions are provided in Section 5.

2. Conservation laws from Gromov's non-squeezing theorem

Let $B^{2n}(r)$ denote the closed Euclidean ball in $\mathbb{R}^{2n}$ centered at 0 with radius r. Let $Z^{2n}(r)$ denote the symplectic cylinder

$$Z^{2n}(r) = B^2(r) \times \mathbb{R}^{2n-2}.$$

Note that Z^{2n} is intended to be a "symplectic" cylinder, in the sense that $B^2(r)$ is supposed to correspond to a pair of conjugate variables; e.g., q_1 and p_1. A basic version of Gromov's result can be stated as follows.

Theorem 1. Let $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a symplectic transformation that maps the ball $B^{2n}(R)$ into a subset of the cylinder $Z^{2n}(r)$. Then $r \geq R$.

More in general, one defines the linear symplectic width $w_L(\Omega)$ of a set Ω as follows:

Definition 2.

$$w_L(\Omega) = \sup_{S \in \mathrm{Sp}(n)} \{\pi R^2 | S(B^{2n}(R)) \subset \Omega\}$$

Here $\mathrm{Sp}(n)$ is the set of all *linear* symplectic maps. One can show that if $n = 1$, $\omega_L(\Omega)$ is simply the area of Ω.

This allows to generalize Theorem 1 above by saying that Ω cannot be mapped symplectically to a cylinder whose width is less than $w_L(\Omega)$. This gives directly a constraint on the joint uncertainty on a pair of conjugate variables.

If one includes the possibility of using also a nonlinear map, one defines "Gromov's width" $w_G(\Omega)$ as follows:

Definition 3.

$$w_G(\Omega) = \sup_{f \in \mathrm{Symp}(n)} \{\pi R^2 | f(B^{2n}(R)) \subset \Omega\}$$

Here $\mathrm{Symp}(n)$ is the set of all symplectic maps. One can prove that w_L and w_G agree on phase-space ellipsoids.

3. Application to uncertainty propagation in Hamiltonian systems

Hsiao and Scheeres [1] applied these ideas to the analysis of the linearized propagation of uncertainty in a Hamiltonian system. A similar technique was employed in [2] where the author tries to make explicit the ties of these bounds to the Heisenberg uncertainty principle.

The paper [1] contains four main ideas:

(1) The linearization of an Hamiltonian flow is still an Hamiltonian flow.
(2) The evolution of the covariance of a distribution is easily computed if the flow is linear.
(3) In the case of a Gaussian distribution, the sublevel sets of the density are ellipsoids.
(4) The symplectic width of some ellipsoids is easily computed.

Thus, they are able to use the non-squeezing theorem to prove bounds on the blocks of the evolved covariance. Let us see these steps in detail.

Linearizing the Hamiltonian system. Recall that a map $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is symplectic/canonical/a symplectomorphism, if it preserves the bilinear form Ω. This condition can be written as

$$\Omega(\mathbf{D}\varphi(z) \cdot z_1, \mathbf{D}\varphi(z) \cdot z_2) = \Omega(z_1, z_2)$$

Or, equivalently, that the matrix representation of $\mathbf{D}\varphi$ is a symplectic matrix $\boldsymbol{A}$ ($\boldsymbol{A}\mathbb{J}\boldsymbol{A}^T = \mathbb{J}$).

Now consider a reference trajectory $\overline{z}(t)$, and the linearized system around this trajectory. Letting $\xi(t) = z(t) - \overline{z}(t)$, the linearized evolution of ξ can be written as

$$\dot{\xi}(t) = A(t)\xi(t)$$

with $A_{ij}(t) = \frac{\partial \varphi_i}{\partial z_j}|_{z=\overline{z}(t)}$. This is, in general, a linear time-variant system. The solution can be written as follows:

$$\begin{aligned} \xi(t) &= \Phi(t,0)\xi(0) \\ \dot{\Phi}(t,0) &= A\Phi(t,0) \\ \Phi(0,0) &= I \end{aligned} \tag{3.1}$$

The crucial property is that $\xi \mapsto \Phi(t,0)\xi$ is still a symplectic map. This can by verified by verifying that $\Phi(0,0) = I$ is clearly symplectic, and that $f(\Phi) = \Phi\mathbb{J}\Phi^T = \mathbb{J}$ is an integral invariant of the operator flow given by (3.1).

Evolution of the covariance in linear systems. Suppose that $\xi(t)$ is a random variable (not necessarily Gaussian) with zero mean, and define $P(t)$ to be its covariance:

$$P(t) = \mathbb{E}\{\xi(t)\xi(t)^T\}$$

It is immediate to see that the relation of $P(t)$ and $P(0)$ is

$$P(t) = \Phi(t,0)P(0)\Phi(t,0)^T \tag{3.2}$$

Note that an equivalent relation does *not* hold for nonlinear systems. Depending on the system, the linearized approximation of the covariance propagation can be either optimistic or pessimistic. For flows that "destroy" information, such as those that arise in problems of tracking, the linearized evolution is typically an optimistic estimate of the real covariance. A filter that uses this approximation would become

quickly *inconsistent* (it underestimates its uncertainty). This is one of the basic problems of the Extended Kalman Filter, and whether it is a problem for the particular application must be ascertained by an ad hoc analysis.

Nevertheless, the assumption of [1] is that the covariance of the distribution evolving through the nonlinear flow can be approximated by equation (3.2) where Φ is the transition matrix of the linearized flow.

Gaussian distributions and ellipses. The non-squeezing theorems applies to (finite) *volumes* evolving in phase space. But if we consider a probability distribution, say a Gaussian, it will occupy, by definition, the whole phase space. The usual trick that one must use is to confound the probability distribution (possibly Gaussian), its covariance matrix, and the ellipsoid corresponding to a given confidence level. The trick works because the flow is assumed to be linear, and the conversion to/from the ellipse commutes with the flow. Intuitively, one can state that

$$\text{ellipse}(\varphi_t(\text{distribution})) = \varphi_t(\text{ellipse}(\text{distribution})).$$

Linear symplectic width of ellipsoids. The linear symplectic width can be computed for some classes of ellipsoids. Given a positive-definite matrix P, define the ellipsoid E_P as

$$E_P = \{\boldsymbol{x} | \boldsymbol{x}^T P^{-1} \boldsymbol{x} \leq 1\}.$$

Even if the ellipsoid is arguably one of the simplest shapes, the linear symplectic width can be computed only in special cases. It is convenient to choose coordinates such that the conjugate variables correspond to adjacent slots of $\boldsymbol{x}$, i.e. $\boldsymbol{x} = (q_1, p_1, q_2, p_2, \ldots, q_n, p_n)$. With this choice, the covariance matrix can be factorized as

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ & \ddots & \\ & & P_{nn} \end{bmatrix},$$

where the single block P_{ij} is the covariance between (q_i, p_i) and (q_j, p_j):

$$P_{ij} \triangleq \begin{bmatrix} \text{cov}(q_i, q_j) & \text{cov}(q_i, p_j) \\ \text{cov}(p_i, q_j) & \text{cov}(p_i, p_j) \end{bmatrix}$$

Proposition. *Suppose that P is a diagonal matrix. Then the symplectic width is*

$$\omega_L(E_P) = \min_i \{\pi \sigma_{q_i} \sigma_{p_i}\}$$

The proof of this is obvious by Definition 2, noting that $\pi \sigma_{q_i} \sigma_{p_i}$ is the area of the ellipsoid projected onto the q_i, p_i plane. More generally, suppose that P can be diagonalized as $P = SDS^T$, where D is diagonal and S is symplectic. By the definition of symplectic width, $w_L(E_P) = w_L(E_D)$, and $w_L(E_D)$ can be computed according to the previous result.

In all other cases, the symplectic width can be computed by finding a symplectic transformation that transforms an ellipse in $\mathbb{R}^{2n}$ to a standard ellipse in $\mathbb{C}^n$. This transformation, however, can be found only numerically as the solution of a nonlinear problem. Details can be found in [1].

Putting it all together. The main theorem of [1] is stated as follows:

Proposition 4. *Consider a dynamical system which is linear and Hamiltonian, with state transition matrix $\Phi(t,0)$. Let $P(t)$ be the covariance matrix of a zero-mean probability distribution. Then the following bound holds:*

$$\det(P_{ii}(t)) \geq \left(\frac{w_L(E_{P(0)})}{\pi}\right)^2.$$

Proof. The covariance evolves as in (3.2). Due to the equivalence of ellipsoids and covariances, and the fact that $\Phi(t,0)$ is symplectic, we can say that $w_L(E_{P(t)}) = w_L(E_{P(0)})$. Moreover, the area A_i of the projection of $E_{P(t)}$ on each symplectic plane must be at least larger than $w_L(E_{P(t)})$: $A_i = \pi\sqrt{\det(P_{ii}(t))} \geq w_L(E_{P(t)}) = w_L(E_{P(0)})$. □

4. Numerical results

In this section, we present some numerical results, by reproducing the simulations in [1] and comparing the linearized analysis with a nonlinear analysis using Monte Carlo simulations.

Model. The simplest system to which we can apply the non-squeezing theorem usefully must have dimension 4, otherwise there would be only one symplectic plane, and the non-squeezing theorem would degenerate, in some cases, to conservation of volume. Consider the case of the two-body problem, a classical Hamiltonian system. Let $r \in \mathbb{R}^2$ be the position of the orbiter in the elliptic plane. Then the dynamical system can be written as

$$\ddot{r} = -\frac{\mu}{||r||^3} r,$$

with μ being the normalized mass. One linearizes the system along a reference trajectory $\overline{r}(t)$. Let $\delta r(t) = r(t) - \overline{r}(t)$. The linearized evolution of $\delta r(t)$ can be written as

$$\delta\ddot{r} = -\frac{\mu}{||\overline{r}||^3}\left(I - 3\frac{\overline{r}\overline{r}^T}{||\overline{r}||^2}\right)\delta r \triangleq \mathbb{V}(t)\delta r,$$

Letting $z(t) = [\delta r, \delta\dot{r}] = (x, y, \dot{x}, \dot{y})$, the linearized system is:

$$\dot{z}(t) = \begin{bmatrix} 0 & I \\ \mathbb{V}(t) & 0 \end{bmatrix} z(t) \triangleq A(t)z(t).$$

One verifies that $\mathbb{V}(t)$ is symmetric, and that $A(t)$ is symplectic.

Methods. In the simulations, we consider an initial Gaussian distribution with mean $\overline{r}(0)$ and covariance $P(0)$. We compute the covariance in two ways:

- According to [1], we compute the linearized covariance $\tilde{P}(t)$ by first integrating the reference trajectory $\overline{r}(t)$, and then computing $\Phi(t, 0)$ by integrating the flow (3.1).
- We compute the covariance $P(t)$ numerically by sampling from the distribution $\mathcal{N}(\overline{r}(0), P(0))$, integrating the trajectory of each sample, and computing the covariance.

The tolerance of the integrator (Runge-Kutta 4-5) is chosen low enough such that numerical errors are not an issue.

Results. By choosing the initial point $\overline{r}(0) = (1, 0, 0, 1)$, and normalized mass $\mu = 1$, one obtains a circular trajectory. The initial covariance matrix is chosen as $P(0) = \mathrm{diag}(0.01^2, 0.01^2, 0.01^2, 0.01^2)$, which is the value used in [1]. The results for this case are shown in Fig. 5.1 for a time interval of 2π, and in Fig. 5.3 for a much larger interval. Fig. 5.2 and Fig. 5.4 show the results for an elliptic reference trajectory, starting at at $\overline{r}(0) = (1, 0, 0, 0.5)$.

For each Figure, subfigure (a) shows the initial sample distribution, and subfigure (b) the final distribution. Subfigure (c) shows the curves $\det(\tilde{P}_{x,\dot{x}}(t))$ and $\det(P_{x,\dot{x}}(t))$, corresponding to the linearized and Monte-Carlo-computed covariances. Analogously, subfigure (d) shows the same for $y, \dot{y}$.

We note the following:

- The simulations appear to be correct because the results of [1] are very similar (nevertheless, they are only *qualitatively* equivalent, as the precise value of $\overline{r}(0)$ was not specified in the paper and had to be guessed).

- The simulations confirm the result of Proposition 4: $\det(\tilde{P}_{x,\dot{x}}(t))$ does not decrease below the initial value. We note, however, that the bound appears to be very loose — note (c) and (d) use a logarithmic scale — because it grows by orders of magnitudes after only one revolution.

However, comparing the results of the linearized and nonlinear analysis, we must conclude that the linearized analysis is not adequate to explain the nonlinear behavior. For the value of $P(0)$ used in [1], the nonlinear flow is different enough to give completely different results (e.g., Fig. 5.1c).

One could say that the linearized analysis would be valid if the initial covariance would be small enough. This is true for small time spans, and in fact we see from the plots in Fig. 5.5 that, by scaling the covariance of 2 orders of magnitudes, $P(0) = \mathrm{diag}(0.001^2, 0.001^2, 0.001^2, 0.001^2)$, the evolutions of $\det(\tilde{P}_{x,\dot{x}}(t))$ and $\det(P_{x,\dot{x}}(t))$ appear to be more similar. Even though, theoretically, I should have obtained approximately the *same* results for even smaller initial covariances, this did not happen, and I blame it on numerical errors.

However, the linearized analysis would be misleading for long enough time spans no matter how small the initial covariance. This can be seen in Fig. 5.3 and 5.4. In the nonlinear system, if the initial uncertainty is non-zero in all directions $x, \dot{x}, y, \dot{y}$, then the sample trajectories will have a continuum of energies and orbiting periods. We assume that the initial uncertainty is small enough such that no sample can escape the gravitational pull. In particular, because the periods will be incommensurable, the uncertainty will be spread out over an annulus in x and y. For the velocities, these will stay in a certain (bounded) set, compatible with the uncertainty in the energy. Therefore, after enough time has passed, the initial distribution will be spread out along the reference orbit. The actual shape of the distribution will be complicated — recall that the flow is a diffeomorphism, therefore the connectivity of the uncertainty volume will be preserved.

In conclusion, no matter how small the uncertainty, the actual covariance under the nonlinear flow will tend to a certain limit. This can be clearly seen in the plots (b) of Fig. 5.3 and 5.4. On the other hand, as can be seen in the plots (c) and (d) of those figures, $\det(\tilde{P}_{x,\dot{x}}(t))$ and $\det(\tilde{P}_{y,\dot{y}}(t))$ will grow without bound. Remember, though, that $\det \tilde{P}(t)$ remains constant, as the volume of the ellipsoid is conserved. What happens is that the ellipsoid is stretched excessively in a certain direction, by increasing the correlation between non-conjugate variables. Therefore, the linearized analysis leads to qualitatively incorrect conclusions.

5. Conclusions

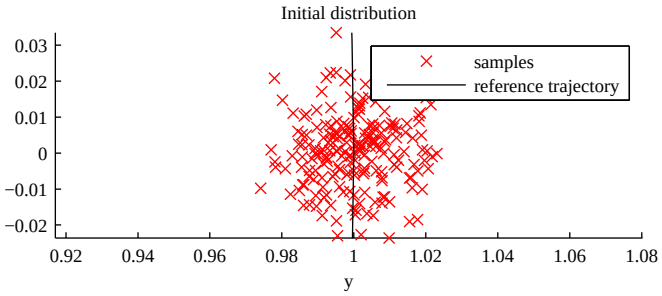
The propagation of uncertainty in Hamiltonian systems must respect several classical constraints. Gromov's non-squeezing theorem provides an absolutely non-trivial constraint on the propagation of uncertainty. However, the problem of applying the theorem to nonlinear systems is that the symplectic width cannot be readily computed for most shapes — even the computation of the linear symplectic width for a generic ellipsoid requires the solution of a nonlinear problem.

The low-hanging fruit used by [1] is that the linearized flow of a Hamiltonian system is still symplectic. Therefore, the theory can be applied to the local analysis of uncertainty propagation. However, the resulting analysis has very local validity and cannot be easily extended to the nonlinear case. The reason is that, while Hamiltonian systems could be considered "nice", in the sense that they preserve phase volumes, and the probability along the flow, the evolution of the shape of a volume can be quite complicated. For example, in the two body problem, two samples arbitrarily close may have incommensurable periods, and therefore end up at opposite places in the space after enough time. Thus it seems likely that several advances to the current state of the art would be necessary before obtaining practical results for the nonlinear case.

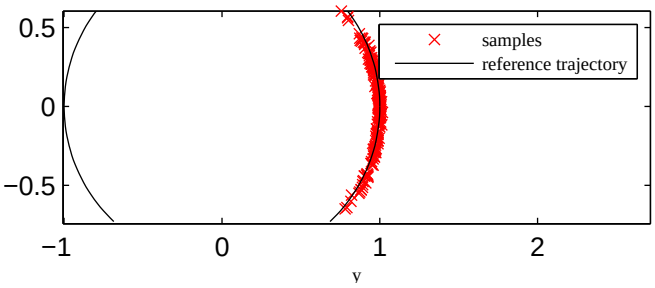
References

[1] F.-Y. Hsiao and D. Scheeres, "Fundamental constraints on uncertainty evolution in hamiltonian systems," *IEEE Transactions on Automatic Control*, vol. 52, pp. 686–691, April 2007.

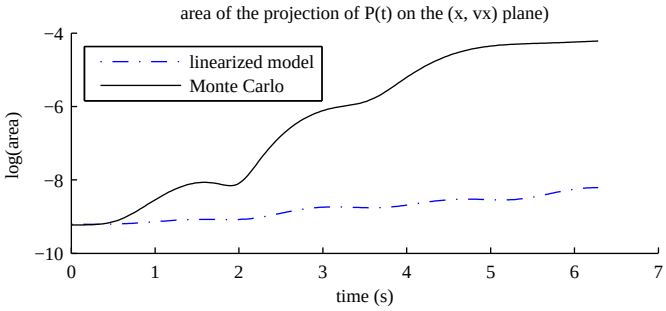
[2] M. de Gosson, "The symplectic camel and the uncertainty principle: The tip of an iceberg?," *Foundations of Physics*, vol. 39, pp. 194–214, Feb. 2009.



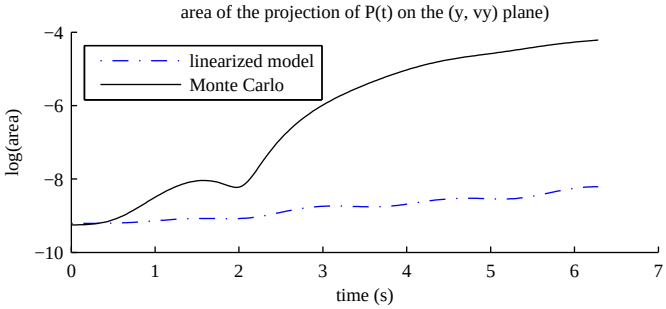
(a) Initial distribution of samples, superimposed on reference trajectory



(b) Final distribution of samples

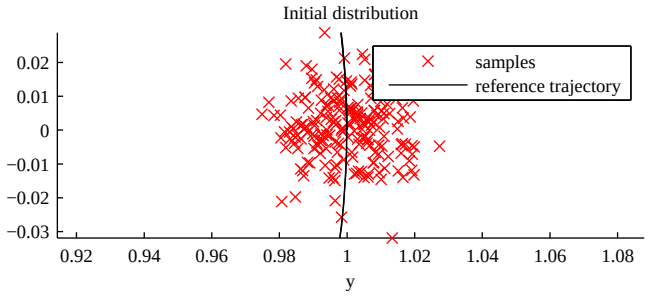


(c) Evolution of uncertainty on the $x, \dot{x}$ symplectic plane.

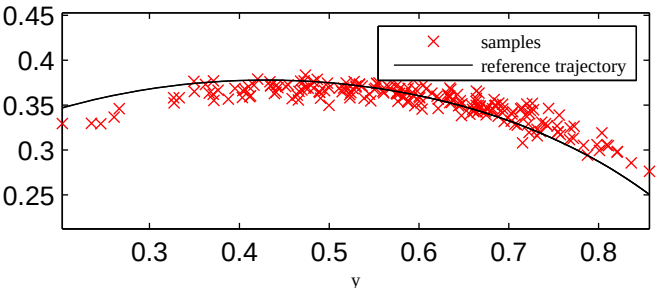


(d) Evolution of uncertainty on the $y, \dot{y}$ symplectic plane.

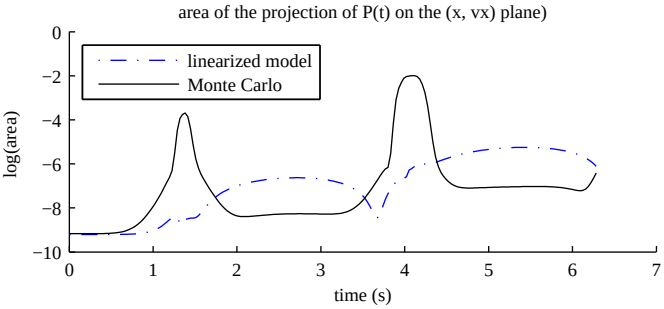
Figure 5.1: Case of a circular trajectory, short time span.



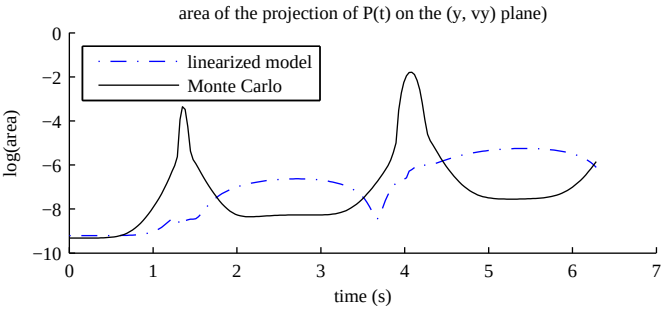
(a) Final distribution of samples



(b) Final distribution of samples.

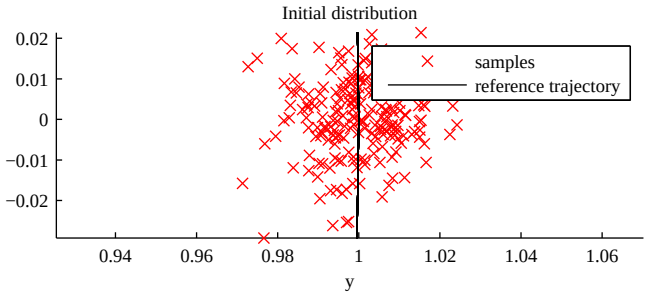


(c) Evolution of uncertainty on the $x, \dot{x}$ symplectic plane.

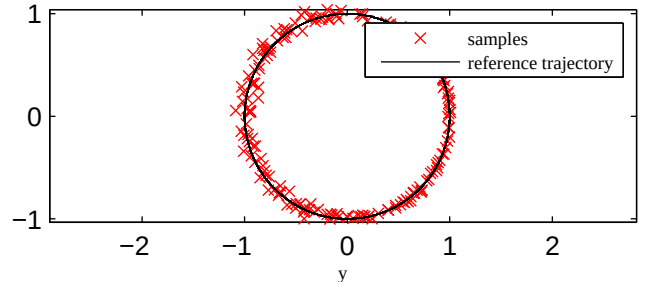


(d) Evolution of uncertainty on the $y, \dot{y}$ symplectic plane.

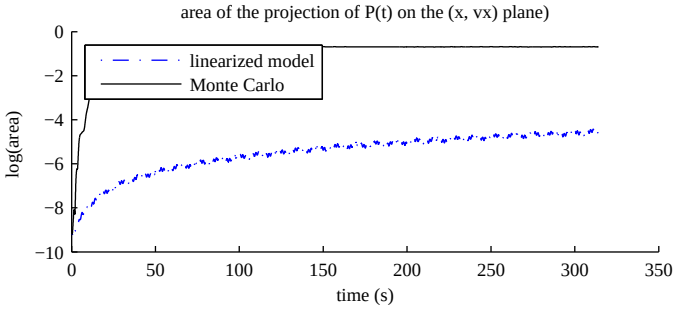
Figure 5.2: Case of an elliptic trajectory, short time span.



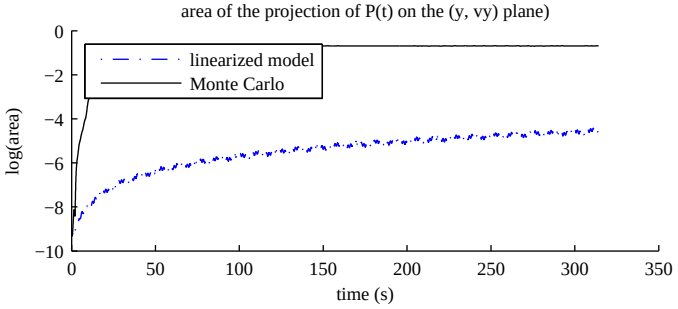
(a) Initial distribution of samples, superimposed on reference trajectory



(b) Final distribution of samples

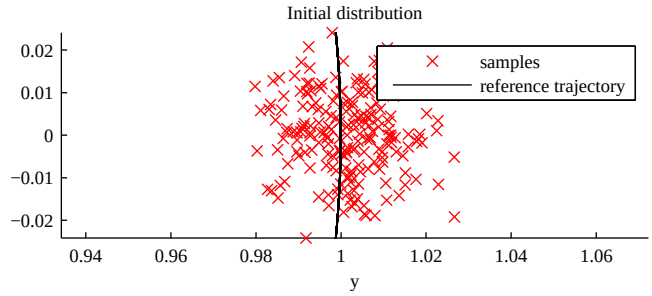


(c) Evolution of uncertainty on the $x, \dot{x}$ symplectic plane.

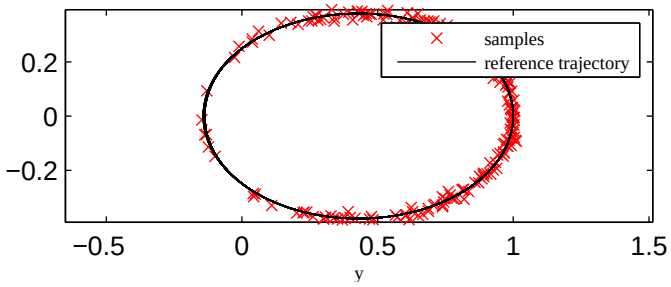


(d) Evolution of uncertainty on the $y, \dot{y}$ symplectic plane.

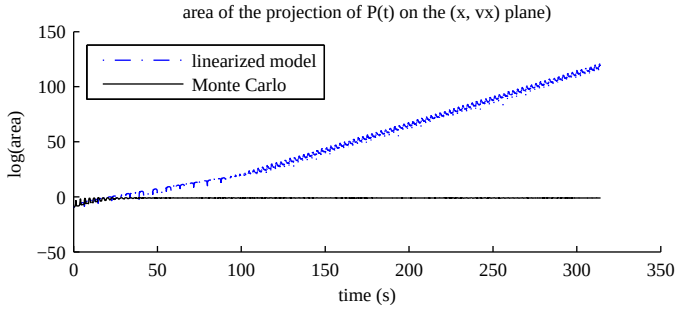
Figure 5.3: Case of a circular trajectory, long time span.



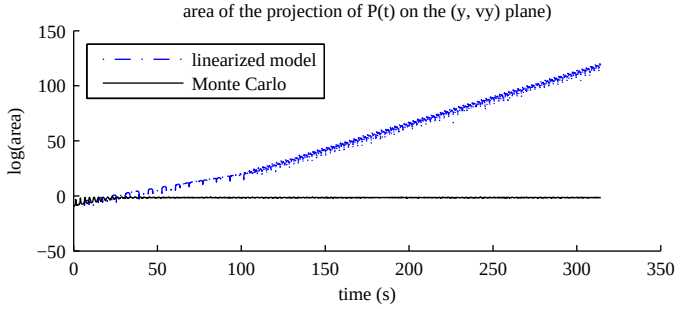
(a) Initial distribution of samples, superimposed on reference trajectory



(b) Final distribution of samples

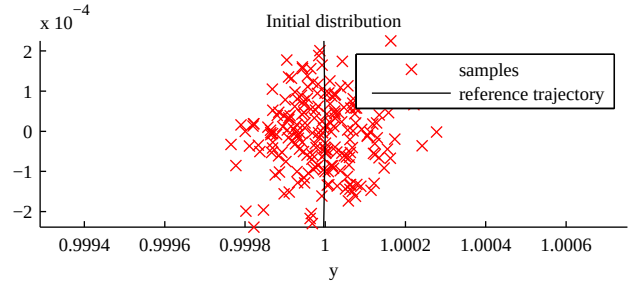


(c) Evolution of uncertainty on the $x, \dot{x}$ symplectic plane.

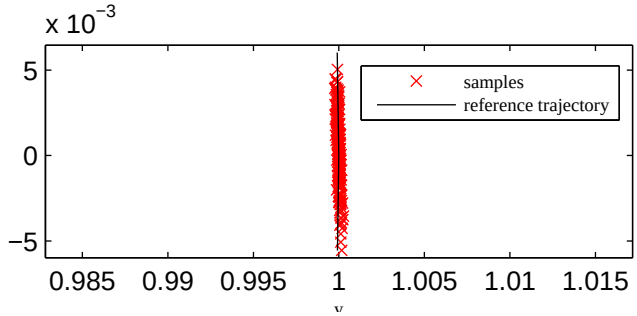


(d) Evolution of uncertainty on the $y, \dot{y}$ symplectic plane.

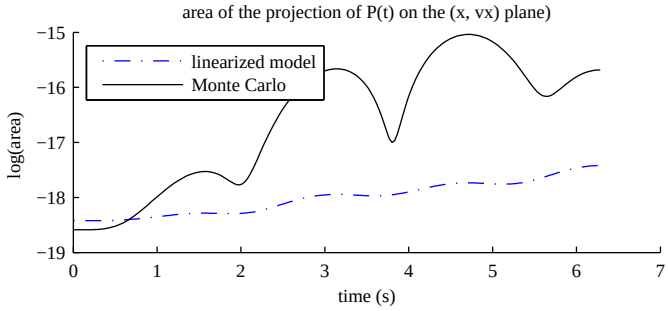
Figure 5.4: Case of elliptic trajectory, long time span.



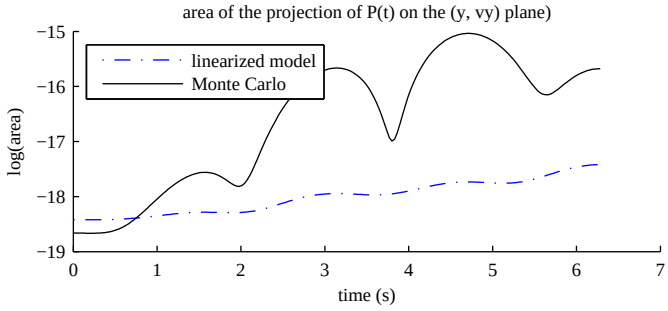
(a) Initial distribution of samples, superimposed on reference trajectory



(b) Final distribution of samples



(c) Evolution of uncertainty on the $x, \dot{x}$ symplectic plane.



(d) Evolution of uncertainty on the $y, \dot{y}$ symplectic plane.

Figure 5.5: Case of a circular trajectory, with very small initial uncertainty.