# Computing time scales from reaction coordinates by milestoning

Tony Faradjian and Ron Elber

Department of Computer Science

Cornell University

Ithaca NY 14853

**Abstract**

An algorithm is presented to compute time scales of complex processes following pre-determined milestones along a reaction coordinate. A non-Markovian hopping mechanism is assumed and constructed from underlying microscopic dynamics. General analytical analysis, a pedagogical example, and numerical solutions of the non-Markovian model are presented. No assumption is made in the theoretical derivation on the type of microscopic dynamics along the reaction coordinate. However, the detailed calculations are for Brownian dynamics in which the velocities are uncorrelated in time (but spatial memory remains).

## I Introduction.

Calculations of mechanisms and time scales of complex molecular processes have been a significant challenge in computer simulations. Typically, we consider a transition between two well-defined molecular states, A and B, and determine the rate of transition between them. The states can be two chemical species, different forms of a crystal, folded and unfolded protein conformations, and more; covering a wide range of spatial and temporal scales.

It is usually assumed that A and B are meta-stable states, meaning that a typical transition time from A to B is much longer than the relaxation times within each state to a local equilibrium. The separation of time scales into slow and rapid relaxations to equilibrium simplifies the theoretical and computational treatment, and reduces the complexity of the full problem. Even if the spatial boundary between A and B (the so-called transition state) is not known, clever methods were introduced (transition path sampling [1]) that build on the separation of time scales to compute the rate.

The procedure defined above requires a clear definition of A and B as quasi states, which reduces somewhat the general applicability of the protocol; not all interesting processes have significant time scale separation and/or maintain quasi equilibrium of the beginning and ending states. Consider for example a protein folding experiment in which an unfolded conformation is perturbed and is no longer at equilibrium. Faster folders [2], without significant free energy barrier, will diffuse to the folded state following an overall (average) attraction to the folded state without a clear boundary between the

initial and final state. More generally, diffusive motions that are not characterized by a single dominant relaxation at long times are of particular interest to this paper.

Of course, in the present manuscript we do not solve the most complex problem, and our treatment is also limited. The essential hypothesis is that a good and low-resolution reaction coordinate, $q$, can be found. By a "good" reaction coordinate we mean a one-dimensional "order parameter" along which the motion is slow compared to all other degrees of freedom. The motions along the complementary set of coordinates (say $C$) are fast in the sense that they reach statistical equilibrium a lot faster than the motion along $q$. The hypersurfaces perpendicular to the reaction coordinate have a significant separation from each other and serves as milestones or a low-resolution reaction coordinate.
We emphasize that the notion of a single reaction coordinate does not imply a single dominant time scale. The reaction coordinate need not include a single large barrier separating A and B (we consider below an example of a completely flat energy surface along the reaction coordinate). The motion along $q$ may have a continuum of relevant time scales that are still slower than typical motions in $C$.

Identifying reaction coordinates that satisfy the above requirement is far from obvious, and therefore in many studies reaction coordinates are selected in an intuitive, ad-hoc way. We have considerable experience in computing long time approximate trajectories (using the SDEL algorithms [3]) that can also serve as such coordinates. SDEL trajectories filter out rapid motions that are expected to equilibrate quickly compared to (slow) diffusive motion along the slow degree of freedom. They further provide a rough

description of the trajectories that we use for "milestoning". This is the type of $q$ we had in mind when starting this study. Hence (and in contrast to the usual notion of a reaction coordinate which is a property of the potential energy surface only), the present reaction coordinate is discrete and may depend on the underlying dynamics (e.g. the kinetic energy or the temperature of the system).

A comment about the nature of the reaction coordinate and the complementary space $C$ is in order. A complete characterization of a reaction coordinate, $q$, (e.g. the distance $r_{ij}$ between two particles $i$ and $j$) must define the complementary space to it, $C$ (figure 1). However, a single trajectory (like SDEL), which provides a sequence of points (not surfaces) in space, cannot do so. One way to address this problem is to consider only the neighborhood of the reaction coordinate such that a hyperplane perpendicular to $q$ is a sound approximation to $C$ (figure 1). We [4] have used this approach to compute a free energy profile in the past.

In the present study we use a sequence of $C$-$s$ ("milestones") along the reaction coordinate $q$ to compute time scales. We denote this sequence by $C_s \quad s = 1,...,S$. Instead of initiating (long) trajectories at the reactant state (as in straightforward Molecular Dynamics simulations), and count the trajectories that made it to the final state, we initiate (shorter) molecular dynamics trajectories at each of the hypersurfaces $C_s$. The trajectories are shorter than the usual Molecular Dynamics simulations since they are terminated when they reach for the first time one of the neighboring milestones $C_{s\pm1}$. For

brevity we will denote from now onwards the hyperplanes by their corresponding index $s$.

From the short trajectories, initiated at each of the hyperplanes and sandwiched between neighboring hypersurfaces, we compute first-passage-time distributions, $K_s^+(\tau)$ and $K_s^-(\tau)$. These are the probability densities for trajectories that at time $\tau = 0$ were initiated at hypersurface $s$, survived there for exactly $\tau$ (incubation time), and then terminated either to the right (+) or to the left (-) hypersurfaces with respect to *s*.

These functions contain all the information on the underlining microscopic dynamics that we use in our model. Note that $K_s^+(\tau)$ and $K_s^-(\tau)$ do not depend on the absolute time of arrival to hypersurface $s$ (we denote the absolute time by $t$). This is an assumption that implies a stationary process and loss of memory with respect to the arrival time to $s$. In practice this construction makes it possible to sample initial conditions for the trajectories initiated at $s$ from a statistical ensemble (no direct dynamic information is required). For this assumption to hold it is necessary for the hypersurfaces to be sufficiently far from each other, so that on arrival to a new hypersurface the memory of the trajectory will be lost (figure 1).

The above condition suggests an interplay between two conflicting desires. On one hand we wish the hypersurfaces to be as far as possible (to guarantee loss of memory), and on the other hand we wish them to be as close as possible for computational efficiency (for closer hypersurfaces the trajectories initiated at hypersurface $s$ will terminate on the neighboring hypersurfaces $s \pm 1$ more quickly). Optimal distances between the hypersurfaces are chosen in the numerical examples of this paper following the above conflicting guidelines: We must guarantee loss of memory and at the same time ensure

computational efficiency. The gain in efficiency by using multiple planes compared to a straightforward simulation is two fold. First, the use of multiple planes makes it possible to use parallel (independent) computers. Parallelization is not possible in a single very long trajectory. Second, the use of the planes to initiate trajectories makes it possible to enhance the probability of rare events (e.g. being at a plane with high free energy) that otherwise will be sampled too poorly. A disadvantage of using planes from computational perspective (compared to a single trajectory) is the requirement of equilibrium at each plane. However, the other advantages more than compensate for the time required to prepare a statistical ensemble at the hyperplanes.

We therefore choose the distance between the hypersurfaces to be as short as possible while still maintaining local equilibrium in the hyperplane. Convergence is assumed when the results do not change upon increasing the hyperplane spacing. We examine the values of the rate constants as well as the degree of equilibration of the first passage time distribution. We demonstrate the impact of different choices in the numerical examples.

In the theory section below we explain how the probability densities ( $K_s^+(\tau)$ or $K_s^-(\tau)$ ) could be used to compute the time evolution of the system, a theory that is followed by numerical examples.

After this work was essentially complete we became aware of a related study by Moroni et al and van Erp et al. [5, 6]. Their procedure is conceptually similar to ours in the sense that they too divide the space into volumes along an "order parameter" (reaction coordinate). They further describe the overall dynamics by a sequence of Markov events between sequential hypersurfaces. Nevertheless, there are a number of significant

differences in the implementation of the similar idea. First, from a microscopic viewpoint we focused on the distribution of the first passage times, while their focus was on the extraction of correlation functions and local rate constants. Second and more importantly, the theory in [5] concentrates on Markovian processes while our theory evolves from a non-Markovian description. Zuckerman and Woolf demonstrated recently the importance of memory in transitional events [7]. The non-Markovian description, which is not difficult to solve in one dimension, is more general and adds flexibility to the modeling.

## II. Theory

Our goal is to compute the time evolution of the system that is given by $P_s(t)$, the probability density of finding the system at $s$ between times $t$ and $t+dt$. Below we describe two approaches to derive $P_s(t)$ from the $K_s^{\pm}(\tau)$ ($K_s^{+}(\tau)$ or $K_s^{-}(\tau)$). The incubation time, $\tau$, is defined only in the interval $[0,t]$. Starting from $s$ termination can occur only at milestones $s\pm1$. The function $P_s(t)$ is normalized to 1,

$$\sum_s P_s(t) = 1 \tag{1}$$

For convenience we define the shorthand

$$K_s(\tau) = K_s^{+}(\tau) + K_s^{-}(\tau) \tag{2}$$

for the probability of transition out of *s* after incubation $\tau$. Incubation means staying at $s$ exactly the given amount of time and no more. We also use the abbreviations

$$\begin{aligned} a^{\pm}b^{\mp} &= a^{+}b^{-} + a^{-}b^{+} \\ a^{\pm}b^{\pm} &= a^{+}b^{+} + a^{-}b^{-} \end{aligned} \tag{3}$$

We will be more explicit when there is a risk of ambiguity.

**II.1 The QK formulation**

Let $Q_s(t)$ be the probability that the system makes a transition to position *s* at time *t*. We do not consider times earlier than the initial time $t = 0$. The evolution of the system is given by:

$$\begin{aligned} P_s(t) &= \int_0^t Q_s(t') \left[ 1 - \int_0^{t-t'} K_s(\tau) d\tau \right] dt' \\ Q_s(t) &= 2\delta(t) P_s(0) + \int_0^t Q_{s\pm1}(t'') K_{s\pm1}^{\mp}(t-t'') dt'' \end{aligned} \tag{4}$$

The first line in equation (4) is the probability of arriving at *s* at time *t'* and not leaving before time *t*. The second line in equation (4) is the probability of transition to *s* expressed as a sum over the initial conditions $\left(P_s(0)\right)$ and over probabilities of previous transitions from its neighbors that follows by a transition to $s$.

A solution of equation (4) solves the stated problem. That is, the equation determines $P_s(t)$ from $P_s(0)$ and $K_s^{\pm}(\tau)$. In principle, we may stop here the theoretical development and continue to the examples. Nevertheless, we believe that the novel mathematical structure and the formulation of the non-Markovian model in equation (4) warrant further elaborations. As with other new formulations it is desirable to develop more than one theoretical viewpoint. We therefore derive below an equivalent pair of integral equations that facilitate further understanding of the underlying model. In fact, in a follow-up study (D. Shalloway and A. Faradjian, to be published), an alternative more efficient numerical procedure is proposed which is based on the PJ formulation below.

### II.2 The PJ formulation

Let $\overline{P}_s(t,\tau)$ be the probability that the system is at position *s* at time *t* with incubation $\tau$ :

$$\overline{P}_s(t,\tau)=\begin{cases} Q_s(t-\tau)\left[1-\int_0^{\tau}K_s(\eta)d\eta\right] & \tau\in[0,t] \\ 0 & \tau\notin[0,t] \end{cases} \tag{5}$$

The entity of interest $P_s(t)$ is given by

$$P_s(t)=\int_0^t \overline{P}_s(t,\tau)\,d\tau \tag{6}$$

Let $J_s^{\pm}(\tau)$ be the probability that the system makes a transition from *s* to $s\pm1$ after incubation $\tau$ , *given* incubation $\tau$ . It is directly related to the input distributions - $K_s^{\pm}(\tau)$

$$J_s^{\pm}(\tau)=\frac{K_s^{\pm}(\tau)}{1-\int_0^{\tau}K_s(\eta)d\eta} \tag{7}$$

Again, for convenience we will use $J_s=J_s^{+}+J_s^{-}$ . The following equation and equation (6) are equivalent to equations (4).

$$\overline{P}_s(t,\tau)=2\delta(\tau)P_s(0)+\int_0^{(t-\tau)}\overline{P}_{s\pm1}(t-\tau,\eta)J_{s\pm1}^{\mp}(\eta)d\eta-\int_0^{\tau}\overline{P}_s(t-\tau+\eta,\eta)J_s(\eta)d\eta \tag{8}$$

The equivalence is shown in **Appendix I**.

We refer to equation (4) as the *QK picture* and to equation (8) as the *PJ picture*. The PJ picture was derived (reversibly) from the QK approach and the two approaches are therefore equivalent. It is also possible to derive the PJ picture directly, on an intuitive ground using probability balance. Consider the probability density of the system to be in

position $s$, at time $t$, and incubation time $\tau$ -- $\overline{P}_s(t,\tau)$. This density, at times different from zero, is growing or diminishing according to fluxes in and out the site $s$. On one hand we need to consider the transition into $s$ from sites $s \pm 1$ precisely at time $t-\tau$. This transition occurs from all range of incubation times $\eta$ in sites $s \pm 1$, and terminates at $s$ exactly after incubation time $\tau$. The second integral on the right hand side equation is the loss to the nearby states of a system initiated at $s$ in time $t-\tau$.

As a side comment it is interesting to note that $\overline{P}_s(t,t)$ is a non-increasing function of time. The probability of remaining at position s from time 0 until time *t* is

$$\overline{P}_s(t,t) = 2\delta(0)P_s(0) - \int_0^t J_s(\eta)\overline{P}_s(\eta,\eta)d\eta \tag{9}$$

$\overline{P}_s(t,t)$ is the probability density that the system will stay at $s$ for a total time $t$ and incubation time $t$. Since both $\overline{P}_s(t,t)$ and $J_s(\eta)$ are non-negative, $\overline{P}_s(t,t)$ is a non-increasing function of $t > 0$.

**II.3 Integro-differential equation**

The PJ picture can be simplified somewhat by deriving an integro-differential equation equivalent to equation (8). Let $\Delta$ be an infinitesimal time interval. Then the relation

$$\overline{P}_s(t,\tau) = \overline{P}_s(t-\Delta,\tau-\Delta)[1-J_s(\tau)\Delta] \quad \tau \neq 0 \tag{10}$$

follows intuitively from the meaning of $P_s(t,\tau)$. Adding $0 = P_s(t,\tau-\Delta) - P_s(t,\tau-\Delta)$ to the left hand side of this equation and dividing by $\Delta$, we obtain

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} + J_s(\tau)\right]\overline{P}_s(t,\tau) = 0 \quad \tau \neq 0 \tag{11}$$

For $\tau = 0$ we have from equation (8)

$$\overline{P}_s(t,0) = 2\delta(t)\overline{P}_s(0) + \int_0^t \overline{P}_{s\pm1}(t,\eta)J^{\mp}_{s\pm1}(\eta)d\eta \tag{12}$$

Combining equations (11) and (12) provides the final formula of this section

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} + J_s(\tau)\right]\overline{P}_s(t,\tau) = \delta(\tau)\left[2\delta(t)\overline{P}_s(0) + \int_0^t \overline{P}_{s\pm1}(t,\eta)J^{\mp}_{s\pm1}(\eta)d\eta\right] \tag{13}$$

Equation (12) resembles a diffusion equation (without a memory) with the additional uni-directional diffusion guided by the operator $\frac{\partial}{\partial \tau}$. This observation motivated us to look for an alternative formulation that casts the problem as Markovian process in a different space.

**II.4 Pseudo-Markovian view**

The system state has hitherto been defined as the position *s*. We now redefine it to be a tuple $(s,\tau)$ consisting of both position *s* and incubation $\tau$. Such a state is called an *incubative state*. To mark the promotion of $\tau$, we modify the notation of probabilities as follows:

$$\overline{P}_s(t,\tau) \to P(s,\tau;t) \tag{14}$$

Since *s* and $\tau$ are all that is needed to specify a transition probability, this redefinition of states recasts this CTRW (Continuous Time Random Walk) as a Markov process. We call this conceptual reformulation the *pseudo-Markovian view*. It is advantageous because it brings the theory of Markov processes to bear on the problem. As an example we derive below a differential Chapman Kolmogorov equation.

Every stochastic process can be defined in terms of conditional transition probabilities. Let $\Delta t$ be an infinitesimal time interval. A suggestive CTRW formula is defined by the following conditional probabilities:

$$P(x,\tau_x;t+\Delta t \mid z,\tau_z;t) = \delta_{x,z}\delta(\tau_x-\tau_z-\Delta t)\left[1-J_z(\tau_z)\Delta t\right]+\delta_{x,z\pm 1}\delta(\tau_x)J_z^{\pm}(\tau_z)\Delta t \qquad (15)$$

where $\pm$ indicates summation over both plus and minus. We now prove that equation (13) is equivalent to the PJ picture. We do so by deriving the differential Chapman-Kolmogorov equation for equation (15) and showing that it is identical to equation (13). The derivation follows [8].

We begin by defining the transition probability per unit time:

$$W(x,\tau_x \mid z,\tau_z;t) = \lim_{\Delta t\to o}\frac{1}{\Delta t}P(x,\tau_x;t+\Delta t \mid z,\tau_z;t) \qquad (16)$$

where $(x,\tau_x)$ and $(z,\tau_z)$ refer to *different* states. Using equation (15), we get

$$W(x,\tau_x \mid z,\tau_z;t) = \delta(\tau_x)[\delta_{x,z+1}J_z^{+}(\tau_z)+\delta_{x,z-1}J_z^{-}(\tau_z)] \qquad (17)$$

(We have expanded the sum over $\pm$.) Now let the distance between states $(x,\tau_x)$ and $(z,\tau_z)$ be bounded by $\varepsilon$. Then the functions $A_i$ defined by

$$\begin{aligned}
&\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int dx d\tau_x (x-z)P(x,\tau_x;t+\Delta t \mid z,\tau_z;t) = A_1(z,\tau_z;t)+O(\varepsilon)\\
&\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int dx d\tau_x (\tau_x-\tau_z)P(x,\tau_x;t+\Delta t \mid z,\tau_z;t) = A_2(z,\tau_z;t)+O(\varepsilon)
\end{aligned} \qquad (18)$$

evaluate to

$$\begin{aligned}
A_1(z,\tau_z;t) &= 0\\
A_2(z,\tau_z;t) &= 1
\end{aligned} \qquad (19)$$

The functions $B_{ij}$, similarly defined but second-order, are all zero. Using $A_i$, $B_{ij}$, and $W$, we have the following differential Chapman-Kolmogorov equation:

$$\left[\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau_z}+J_z(\tau_z)\right]P(z,\tau_z;t\,|\,y,\tau_y;t')=\delta(\tau_z)\int_0^t P(z\pm 1,\eta;t\,|\,y,\tau_y;t')J_{z\pm 1}^{\mp}(\eta)d\eta \quad (20)$$

Apart from the fact that equation (20) involves probabilities that are conditional (and hence does not involve the initial system distribution $P_s(0)$), it is identical to equation (13).

## III. Examples

The first example that we considered is primarily pedagogical, showing the relationship of our formulation to the more familiar Master equation. The next three examples are somewhat more complex and are solved numerically, either by exact binning of (long time) trajectories, or using the present QK formulation with different number of slices along the reaction coordinates. Two of the numerical examples are for one-dimensional problems and the third example is for diffusion in two dimensions.

Note that in the present manuscript we did not consider the difficulties in the choice and use of hypersurfaces for milestoning and choose examples in which hyperplanes are appropriate. The use of hyperplanes (an approximation to general hypersurfaces) will be examined in detail in future works.

### III.1 Simulating Markov processes

As a pedagogical example, we specialize our formalism to a pure Markov process, modeled by a Master equation with nearest neighbor transitions.

$$\frac{dP_s}{dt}=-\left(k_s^+ + k_s^-\right)P_s + k_{s+1}^- P_{s+1} + k_{s-1}^+ P_{s-1} \equiv -k_s P_s + k_{s\pm 1}^{\mp} P_{s\pm 1} \quad (21)$$

As before $s$ is used to denote a state, and $P_s$ is the probability of finding the system at state $s$. The parameters $k_s^{\pm}$ are local rate constants in the backward and forward direction from state $s$. To show the reduction it is useful to keep in mind both formulations (QK and PJ) and to switch between them to make the derivation simpler. We consider equations (5) and (8). We start however with the observation that for a Markov process, the first passage-time distribution function takes the form

$$K_s^{\pm}(\tau) = k_s^{\pm} \exp\left[-\left(k_s^{+} + k_s^{-}\right)\tau\right] = k_s^{\pm} \exp\left[-k_s \tau\right] \tag{22}$$

Consider equation (4), with the above expression for the first-passage-time distribution we can do the integral in the first line.

$$P_s(t) = \int_0^t Q_s(t')\left[1 - \int_0^{t-t'} k_s \exp\left[-k_s \eta\right] d\eta\right] dt' = \int_0^t Q_s(t') \exp\left[-k_s\left(t - t'\right)\right] dt' \tag{23}$$

On the other hand differentiating the first line of equation (4) with respect to time we have

$$\frac{dP_s(t)}{dt} = Q_s(t) - \int_0^t Q_s(t')\, K_s(t - t')\, dt' \qquad t \neq 0 \tag{24}$$

Substituting the second line of equation (4) (the expression of $Q_s(t)$) in the above equation we obtain for $t \neq 0$

$$\frac{dP_s(t)}{dt} = \int_0^t \left[Q_{s\pm1}(t')\, K_{s\pm1}^{\mp}(t - t') - Q_s(t')\, K_s(t - t')\right] dt' \qquad t \neq 0 \tag{25}$$

Note the similarity of equation (25) to a generalized master equation with memory [9, 10]. However, the right hand side of equation (25) includes $Q$ and not $P$ in contrast to the generalized Master equation. Nevertheless, the two formulations are equivalent as was pointed out to us by Attila Szabo. The QK formulation is the simplest to apply in the

simulation since $K_s^{\pm}(\tau)$ is obtained directly from trajectories. In **Appendix II**, we show the equivalence between equation (25) and the generalized Master equation.

Returning to the Markov limit we substitute the explicit expression of $K$ (equation (22)) into equation (25) and obtain

$$\frac{dP_s(t)}{dt} = \int_0^t \left[ Q_{s\pm1}(t') k_{s\pm1}^{\mp} e^{-k_{s\pm1}(t-t')} - Q_s(t') k_s e^{-k_s(t-t')} \right] dt' \tag{26}$$

The terms in the integral of the above equation could be identified using equation (23).

$$\frac{dP_s(t)}{dt} = k_{s\pm1}^{\mp} P_{s\pm1}(t) - k_s P_s(t) \tag{27}$$

Hence, we finally recover the Master equation.

### III.2 Numerical simulations

We tested numerically the formalism of the above memory equations on two models: one-dimensional free diffusion, and two-dimensional diffusion through an entropic barrier. In all the calculations below the QK integral equation were used. The calculation of $J_s^{\pm}(\tau) \equiv K_s^{\pm}(\tau) \Big/ \left[ 1 - \int_0^{\tau} K_s(\eta) d\eta \right]$ (equation (7)) in the PJ formalism was proven difficult numerically due to significant noise in the estimated $J_s^{\pm}(\tau)$ for large $\tau$ for which we have limited statistics. This noise affects less the QK formulation, since $K_s^{\pm}(\tau)$ decays to 0 at large $\tau$. Nevertheless, there are alternative routes to solve the PJ equations that are described in a follow up study (Shalloway and Faradjian, to be published).

In both numerical examples, computer simulations were used to calculate exact long-time trajectories as well as first-passage-time trajectories that begin at reaction coordinate position *s* and terminate at neighboring planes $s \pm 1$ (for all *s*). All ensembles consisted of 5000 trajectories, and were generated using a Brownian integrator [11] with friction coefficient $\gamma = 0.1$ and diffusion constant $D = T/\gamma$ ($m = k_B = 1$). Langevin timesteps ranged from $10^{-7}$ to $10^{-5}$. To compute the first passage time distributions $K_s^{\pm}(\tau)$, histograms of the first passage times were repeatedly reconstructed using successively smaller bin widths in $\tau$, until they converged, i.e. when further reduction in the bin size (and an increase in the number of trajectories) did not change the simulation results.

It is important to note that even though Brownian dynamics has no memory in velocity space, memory in coordinate space (relaxation in the plane perpendicular to the reaction coordinate) is not guaranteed and requires careful evaluation.

We compared exact (i.e., long-time) trajectories up to a given time $t_f$ and the resulting evolution of the system distribution to that computed by the QK equations up to the same time. We consider the rate constant for flow of probability across a division of the configuration space into two equal and symmetric halves. That is, the space was divided into two symmetric "states," *A* and *B*, and the rate of reaction $k(T)$ from *A* to *B* was computed assuming two-state kinetics:

$$kt = -\ln(1 - P_B(t)/P_B^{eq}) \tag{28}$$

where $P_B(t)$ is the total probability of state *B* at time *t*, and $P_B^{eq}(t)$ is its equilibrium probability. Since the initial population is all at the $A$ state we also have $P_B(0)=0$. The rate constant *k* was extracted from a linear fit.

### III.2.1 One dimensional free diffusion

In this model, particles (trajectories) are confined to the unit interval $[0,1]$ using a "steep wall" potential

$$U(x)=\begin{cases} 0,\ x\in[0,1] \\ W(x-1)^2,\ x>1 \\ Wx^2,\ x<0 \end{cases} \tag{29}$$

where $W=10^4$. The reaction coordinate is the unit interval and is divided by points $x_s=(s-1)/(M-1),\ s=1,2,\ldots M$. Three different sets of planes were used: $M=4,\ 8,\ 16$. Simulations were run for four temperatures: $T=0.025,\ 0.05,\ 0.1,\ 0.2$. Here we chose $t_f=1$. The initial condition for the exact trajectories was $x=0$; to integrate the QK equations the initial condition was a delta function at the same position. Note that in the one-dimensional case there is no memory in the coordinates perpendicular to the reaction coordinate (there are no such coordinates!), and therefore deviation from the exact results are due only to the discretization of configuration space and our numerical procedure.

To consider rate calculations we divide the space into two. Coordinates smaller than $1/2$ are considered the $A$ state (reactant), and coordinate larger or equal to $1/2$ are the state $B$ (product). Comparisons of rate calculations using the formula $kt=-\ln(1-P_B(t)/P_B^{eq})$ as a function of time $t$ are presented; where $k$ is the rate

constant, $P_B(t)$ the reactant probability at time $t$, and $P_B^{eq}$ is the equilibrium distribution of state $B$. Linear fits to these curves yield the rates shown in table 1.

** PLACE TABLE 1 HERE **

In figure 2 we show a few rate calculations in greater details. The time ranges shown in these latter curves are those for which $P_B(t)$ is still sufficiently different from $P_B^{eq}$ to avoid numerical problems in the formula $kt = -\ln(1 - P_B(t)/P_B^{eq})$.

**III.2.2 Comparison to a Markov model**

It is of interest to compare the results of the QK model to the simpler stochastic approach of a Markov process. We have shown in section III.1 that the reduction is straightforward. If we choose $K_s^{\pm}(\tau)$ to be the simple exponential function - $K_s^{\pm}(\tau) = k_s^{\pm} \exp[-k_s \tau]$, then $J_s^{\pm}(\tau) \equiv k_s^{\pm}$ is a constant and the process is Markovian. As an example we fit exponential expressions to $K_s^{\pm}(\tau)$ (we use $k_s = k_s^+ + k_s^-$ as an independent parameter, and assume $k_s^+ = k_s^-$ -- appropriate for free diffusion). The numerical data is from the simulations of first-passage-time distributions of the previous section (free diffusion). While one can argue that a different fit may have been used, this is a reasonable choice given the overall success of the QK formulation (note that the QK formulation does not require a fit since the numerical $K_s^{\pm}(\tau)$ is used directly).

The resulting Markovian model was used to solve for the rate in the free diffusion case. The results were considerably worse than those of the QK model. For temperature $T = 0.1$ we estimate by straightforward Brownian trajectories an exact rate constant of

9.9. With the milestoning approach we estimate the rate constants of 9.5, 9.5 and 9.2 for 4, 8, and 16 milestones respectively. The Markov model (derived as described above) provides rate constants of 6.4, 8.95 and 9.4 estimated using 4, 8, and 16 planes (Markovian milestones). The statistical errors for the Markovian calculations are roughly $\pm 0.2$. When the separation between the milestones is large, the Markov model is less accurate since it does not account correctly for the early delay of the first-passage-time (see figure 3 for numerically computed memory functions for the two dimensional case). The Markov process becomes more accurate as the milestones are getting closer. However, using a large number of milestones can be problematic since the local equilibrium assumption at a milestone may be invalid. The non-Markovian model is therefore more flexible and with a wider range of applicability.

**III.2.3 One dimensional harmonic oscillator**

The one-dimensional harmonic oscillator differs from the previous example by having a non-zero potential term making it possible for us to consider non-uniform (in space) equilibrium properties. Brownian dynamics is considered on a potential $U(x) = (1/2)x^2$, a friction constant of $\gamma = 0.1$, and temperature $T = 0.1$. In Figure 4 we demonstrate that the equilibrium state of the QK model agrees well with what we know for harmonic oscillator in the canonical ensemble and with the binning of a single long Brownian trajectory. The initial condition was a delta function at $x = -1$. It is not obvious that the non-Markov model with the numerically derived $K_s^{\pm}(\tau)$ approaches the correct statistical equilibrium, and the present numerical example (together with other tests that we worked out) is reassuring.

### III.2.4 Two dimensional entropic barrier

Our second test case involves the potential

$$U(x,y) = x^6 + y^6 + e^{-(x/\sigma_x)^2}(1 - e^{-(y/\sigma_y)^2}) \tag{30}$$

where $\sigma_x = \sigma_y = 0.1$. As shown in the contour plot in Figure 5, this potential energy divides the two-dimensional configuration space into two equal and symmetric wells, separated by a barrier with a small opening at the origin. The reaction coordinate is the *x*-axis and is divided by planes (lines) parallel to the *y*-axis, passing through points $x_s = -1 + 2(s-1)/(M-1),\ s = 1, 2, \ldots M$. We used three sets of planes: $M = 4,\ 8, 16$; and we used five temperatures: $T = 0.0125,\ 0.025,\ 0.05,\ 0.1,\ 0.2$. Here we chose $t_f = 10$. The initial condition for the exact (long-time) computation was a canonical ensemble prepared in the plane $x = -1$. To integrate the memory equations the initial condition was a delta function at the same position. Similarly to the calculations in one dimension (table I and figure 2) we divide the space into two: a reactant (negative $x$) and a product (positive $x$).

In figure 6 we examine the basic assumption that the first-passage-time is significantly longer than the relaxation time to equilibrium within a milestone. We are showing a favorable case using four milestones, the calculations with 16 planes fails to show separation of time scales. These results are consistent with the observations for rate calculations in which 4 planes yield accurate rate constants and 16 planes do not.

Comparison of the rate calculations for different temperatures and different number of milestones is given in table II. Detailed time courses for two temperatures (0.0125 and 0.2) are shown in figure 7. The $M = 8$ and $M = 16$ predictions do not reliably follow the

exact curve across all temperatures, as does the most widely-spaced plane division, $M = 4$. Together with the results presented in figure 6 we interpret the above observations as verifying the fundamental assumption of separation of timescales: for $M = 8$ and $M = 16$, the planes are close enough to violate the condition that the time scale for relaxation for coordinates perpendicular to the reaction coordinate is much faster than the relaxation time along the reaction coordinate (i.e. $\tau_{\perp} \ll \tau_{\parallel}$). (see table II).

** PLACE TABLE 2 HERE **

Finally, a log-log plot (figure 8) of the rate constant versus the temperature suggests that all the discretizations we considered are doing (perhaps surprisingly) reasonably well. While $M = 4$ is the superior division, $M = 8$ is a close second. Note however, that for an order of magnitude estimate of the rate the cheapest version (16 planes) may do.

## IV. Final remarks

We introduced a new algorithm to compute time scales of long-time processes if a reaction coordinate (or an order parameter) is known. The reaction coordinate is used to guide the trajectory calculations by "milestones" (hyperplanes) of the process. A specific implementation of a Markovian milestoning was found inaccurate compared to the non-Markov QK calculations. The essential assumption of our algorithm is the loss of spatial (and velocity) memory upon transitions between the "milestones". This assumption is about separation of time scales: slow time scales that are associated with motion along the reaction coordinate, and fast relaxation times to statistical equilibrium along other coordinates.

It is this assumption that will have to be examined carefully in future studies of much larger systems. SDEL trajectories [3], which are approximate classical mechanical paths with poor resolution of fast degrees of freedom, are potential candidates for such "reaction coordinates". Another potential difficulty in using milestones and hyperplanes to compute rates is the approximate nature of the hyperplanes to the true hypersurfaces orthogonal to the reaction coordinate. This difficulty was not addressed in the present manuscript and will be studied in future work.

**Acknowledgements**

This research was supported by NIH grant GM59796. We are grateful to Attila Szabo for the discussion, suggestions, and for pointing out the equivalence of our equation (4) and the generalized Master equation. Useful comments on an earlier version of the manuscript from David Shalloway, Dan Zuckerman, and Art Voter are acknowledged.

**Appendix I**

We derive the QK integral equations from the PJ integral equations. Using the definition of $\overline{P}_s(t,\tau)$

$$\overline{P}_s\left(t,\tau\right)=Q_s\left(t-\tau\right)\left[1-\int_0^{\tau}K_s\left(\eta\right)d\eta\right] \tag{A.I.1}$$

We integrate over $\tau$ from 0 to $t$ and change variables from $\tau$ to $t-t'$ we have

$$P_s\left(t\right)=\int_0^{t}Q\left(t'\right)\left[1-\int_0^{t-t'}K_s\left(\eta\right)d\eta\right]dt' \tag{A.I.2}$$

which is identical to the first line in the QK formulation (equation (4)). We also need to show that the integral equation of the PJ formulation (equation (8))

$$\overline{P}_s(t,\tau)=2\delta\left(\tau\right)P_s\left(0\right)+\int_0^{(t-\tau)}\overline{P}_{s\pm1}(t-\tau,\eta)J_{s\pm1}^{\mp}(\eta)d\eta-\int_0^{\tau}\overline{P}_s(t-\tau+\eta,\eta)J_s(\eta)d\eta \tag{A.I.3}$$

implies the second line in the QK formulation (equation (4)).

$$Q_s(t)=2\delta\left(t\right)P_s\left(0\right)+\int_0^{t}Q_{s\pm1}(t'')K_{s\pm1}^{\mp}(t-t'')dt'' \tag{A.I.4}$$

Substituting the definitions of $\overline{P}_s(t,\tau)$ (equation A.I.1) and $J_s^{\pm}(\tau)$ (equation (7)) into equation (A.I.3) we obtain

$$Q_s(t-\tau)\left[1-\int_0^{\tau}K_s(\eta)\,d\eta\right]=\int_0^{t-\tau}Q_{s\pm1}(t-\tau-\eta)\,K_{s\pm1}^{\mp}(\eta)\,d\eta-\int_0^{\tau}Q_s(t-\tau)\,K_s(\eta)\,d\eta \tag{A.I.5}$$

Factors of $1-\int_0^{\eta}K_{s\pm1}(\eta')\,d\eta'$ from the definitions of $\overline{P}_s(t,\tau)$ and $J_s^{\pm}(\tau)$ have cancelled each other out in both integrals on the right-hand side. The factor $Q_s(t-\tau)$ in the second

right-hand integral is a constant with respect to the integration. It cancels the same term on the left-hand side. What remains is

$$Q_s(t-\tau) = \int_0^{t-\tau} Q_{s\pm1}(t-\tau-\eta)\, K_{s\pm1}^{\mp}(\eta)\, d\eta \tag{A.I.6}$$

Now we make the change of variables $t' \equiv t-\tau$ to get

$$Q_s(t') = \int_0^{t'} Q_{s\pm1}(t'-\eta)\, K_{s\pm1}^{\mp}(\eta)\, d\eta \tag{A.I.7}$$

Another change of variables, $t'' \equiv t'-\eta$, yields an equation that is identical to the second QK integral equation. QED.

**Appendix II.**

Here we show the equivalence of equation (4) and the generalized Master equation. Attila Szabo pointed this equivalence to us.

The Laplace transform of the first of equations (4) is

$$u\tilde{P}_s(u) = \left(1-\tilde{K}_s(u)\right)\tilde{Q}_s(u) \tag{A.II.1}$$

where $u$ denotes the Laplace variable and $\tilde{A}$ denotes the Laplace transform of $A$.

On the other hand the Laplace transform of equation (25) is

$$u\tilde{P}_s(u) - P_s(0) = \tilde{Q}_{s\pm1}(u)\,\tilde{K}_{s\pm1}^{\mp}(u) - \tilde{Q}_s(u)\,\tilde{K}_s(u) \tag{A.II.2}$$

Equation (A.II.1) is used to eliminate $\tilde{Q}_s(u)$ from equation (A.II.2)

$\left(\tilde{Q}_s(u) = uP_s(u)\Big/\left(1-\tilde{K}_s(u)\right)\right)$.

$$u\tilde{P}_s(u) - P_s(0) = uP_{s\pm1}(u)\Big/\left(1-\tilde{K}_{s\pm1}(u)\right)\tilde{K}_{s\pm1}^{\mp}(u) - uP_s(u)\Big/\left(1-\tilde{K}_s(u)\right)\tilde{K}_s(u) \tag{A.II.3}$$

We now define the Laplace transform of the function $R_s^{\pm}(\tau)$

$$\tilde{R}_s^{\pm} = \frac{u\tilde{K}_s^{\pm}(u)}{\left(1-\tilde{K}_s(u)\right)} \qquad \text{(A.II.4)}$$

Substituting equation (A.II.4) in (A.II.3) and using the inverse Laplace transform we obtain

$$\frac{dP_s(t)}{dt} = \int_0^t \left[-R_s(\tau)P_s(t-\tau) + R_{s\pm1}^{\mp}(\tau)P_{s\pm1}(t-\tau)\right]d\tau \qquad \text{(A.II.5)}$$

which is the definition of the generalized Master equation. QED.

We note that it is not obvious to us how to compute $R_s^{\pm}(\tau)$ directly from Molecular Dynamics simulations, while the calculations of $K_s^{\pm}(\tau)$ are straightforward.

**Tables Legends**

**Table 1.** Comparison of exact rate constants computed as a function of temperature using different simulation protocols (the QK formulation with 4, 8, and 16 slices, and exact trajectories). In the QK formulation we assume loss of spatial memory between the slices. This loss of memory is self-evident in one-dimensional Brownian dynamics. Therefore the simulation below only tests the numerical procedure, and the space discretization.

**Table 2**. Calculations of rate constants for a two-state transition in a two dimensional system. We calculate the rate constant (as a function of temperature) using the QK formulation with 4, 8, and 16 planes and an exact ensemble of Brownian trajectories. Note that as we add more planes the assumption of memory loss get worse. The errors (as expected) are reduced for higher temperatures and are significantly less than a factor of two.

Table 1: Rate constants as a function of temperature: 1D system

| **T** | *k (M=4)* | *k (M=8)* | *k (M=16)* | *k (exact)* |
|---|---|---|---|---|
| 0.025 | 2.417 (0.005) | 2.467 (0.001) | 2.3329 (0.0003) | 2.367 (0.027) |
| 0.050 | 4.92 (0.04) | 4.756 (0.004) | 4.800 (0.001) | 4.86 (0.15) |
| 0.100 | 9.47 (0.06) | 9.460 (0.006) | 9.241 (0.002) | 9.89 (0.21) |
| 0.200 | 19.15 (0.05) | 18.526 (0.006) | 19.299 (0.003) | 19.15 (0.58) |

Table 2. Rate constants as a function of temperature : 2D system

| **T** | *k (M=4)* | *k (M=8)* | *k (M=16)* | *k (exact)* |
|---|---|---|---|---|
| 0.0125 | 0.0842 (0.0006) | 0.08737(0.00002) | 0.12189(0.00001) | 0.0894 (0.0009) |
| 0.0250 | 0.166 (0.001) | 0.2199 (0.0003) | 0.18432(0.00005) | 0.164 (0.003) |
| 0.0500 | 0.313 (0.001) | 0.4921 (0.0002) | 0.32004(0.00004) | 0.305 (0.004) |
| 0.1000 | 0.695 (0.002) | 0.9374 (0.0002) | 0.74278(0.00005) | 0.635 (0.011) |
| 0.2000 | 1.740 (0.004) | 1.9662 (0.0005) | 2.1875 (0.0003) | 1.78 (0.04) |

**Figure Legends**

Figure 1.

A schematic drawing of a reaction coordinate on a two dimensional energy surface and sequential planes perpendicular to it (representing the states $s$). Also shown is a trajectory (dashed line) that starts at the second slice from the right and terminates at the third slice from the right. Trajectories initiated at each of the slices are used to estimate the function $K_s(\tau)$. See text for more details on the calculations.

Figure 2.

The natural log of the reaction progress : $\log\left(1-\frac{P_B(t)}{P_B^{eq}}\right)$ as a function of time, $t$, for one dimensional free diffusion. Note that there is no required relaxation in the direction perpendicualr to the reaction coordinate in this case. $P_B(t)$ is the population fraction of the product, and $P_B^{eq}$ is the fraction of population of the product at equilibrium. For exponential kinetics the above expression is equal $kt$, where $k$ is the rate constant and $t$ is the time. We consider two temperatures 0.025, and 0.2 (2.a and 2.b). Data on more temperatures is available in table 1. We compute each of the observables with exact trajectories (solid lines) and with different discretization level with the QK formalism (four slices – circles, eight slices -- squares, and sixteen slices --diamonds).

Figure 3.

A sample calculation of the memory function $K_s(\tau)$ for the two-dimensional energy surface using 8 slices. Two memory functions are plotted $K_3^+(\tau)$ (dotted line) and $K_4^+(\tau)$ (solid line). Position 3 is x = -0.4286 and position 4 is x = -0.14825. The temperature is T = 0.05. See text for more details on the calculations.

Figure 4.

The equilibrium spatial distribution of the harmonic oscillator is calculated ($U(x) = (1/2)x^2$ and the temperature $T = 0.1$). Three computational approaches are shown: exact analytical result (o and solid line), binning a long Brownian trajectory (+), and asymptotic behavior of the QK equations (*). The deviations at the edges are due to the use of reflecting boundary condition in the last planes. Note the excellent agreement anywhere else.

Figure 5.

A contour plot of the two-dimensional energy surface that was used to test the QK formalism. The system was initiated at the left well and transitions were considered to the right well. The energy difference between contour lines is 0.25. See text for more details. Note that the barrier is primarily entropic.

Figure 6.

Relaxation to equilibrium of two milestones in the two-dimensional energy surface of figure 5. The relaxation induced by diffusion from the milestone at x=0.333 to the

milestone at x=-0.333. The temperature is 0.05. Snap shots in time of the spatial distribution in the milestones are shown in figure 6a. The solid line is the equilibrium distribution at the target milestone ( $x = -0.333$ ), the dotted line is the initial distribution in the previous milestone ( $x = 0.333$ ), and the circles present the distribution after time of 0.5. Note that the first passage time is 6.37. In figure 6b we show the difference between the time dependent probability density and the probability density in equilibrium: $\int \left( \rho(t,x) - \rho_{eq}(x) \right)^2 dx$, where $\rho_{eq}(x)$ is the equilibrium distribution, and $t$ is the time.

Figure 7

This plot is similar to figure 2, except that the calculations are for the two-dimensional energy surface (figure 5). As in figure 2, two temperatures are considered 0.025, and 0.2 (figures 7.a and 7.b). See legend of figure 2 for more details.

Figure 8

Temperature dependence of rate constants. We plot the log of the rate constant versus the log of the temperature (the rate constant is not exponential in the temperature). See text for more details how the calculations of the rate constants were performed. The exact results are the solid line. QK calculations are shown with 4 slices (circles), 8 slices (squares) and 16 slices (diamonds). Figure 8.a is for the simulation using the one-dimensional model and 8.b for the two dimensional model.

Figure 1.

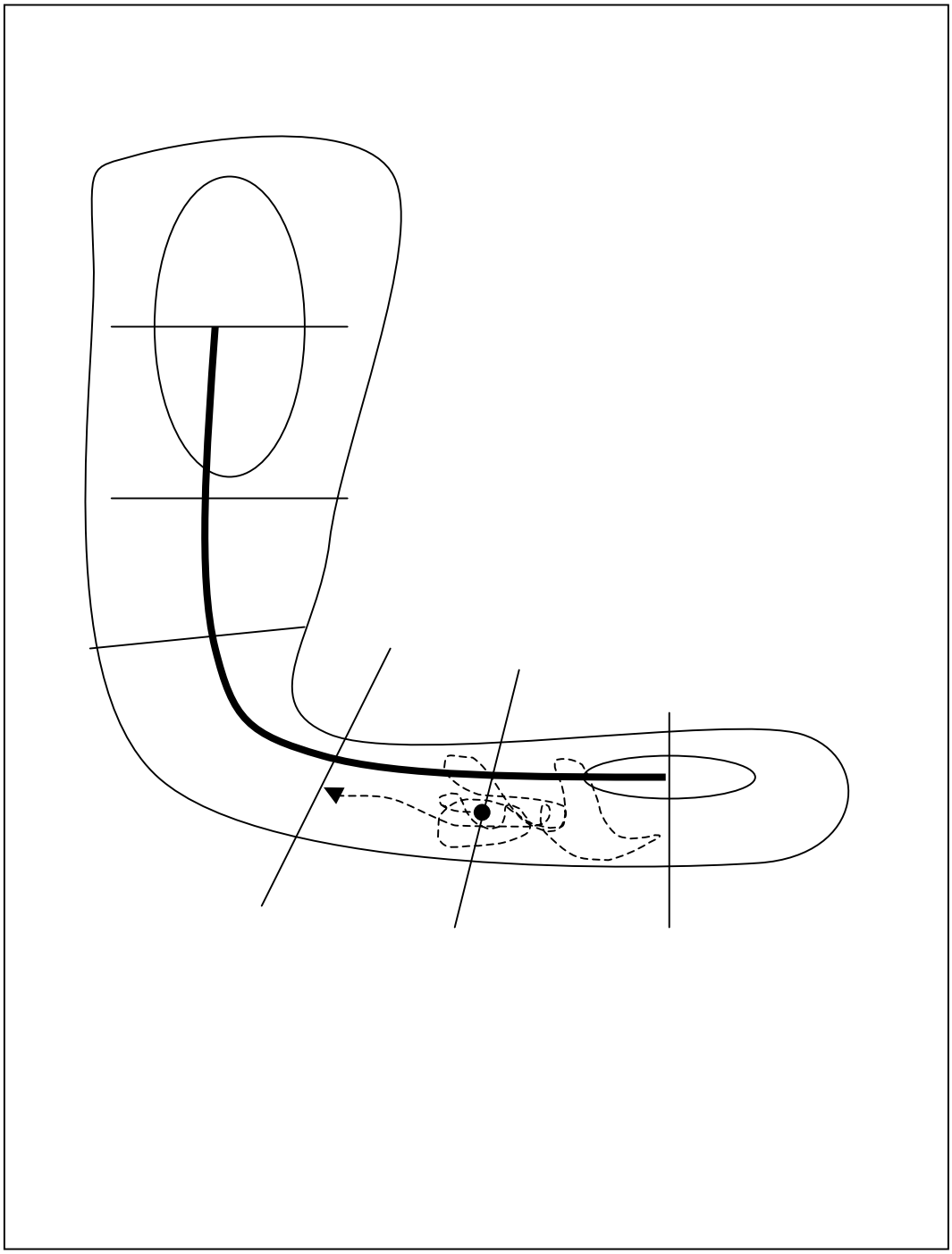


Figure 2.
Figure 2a

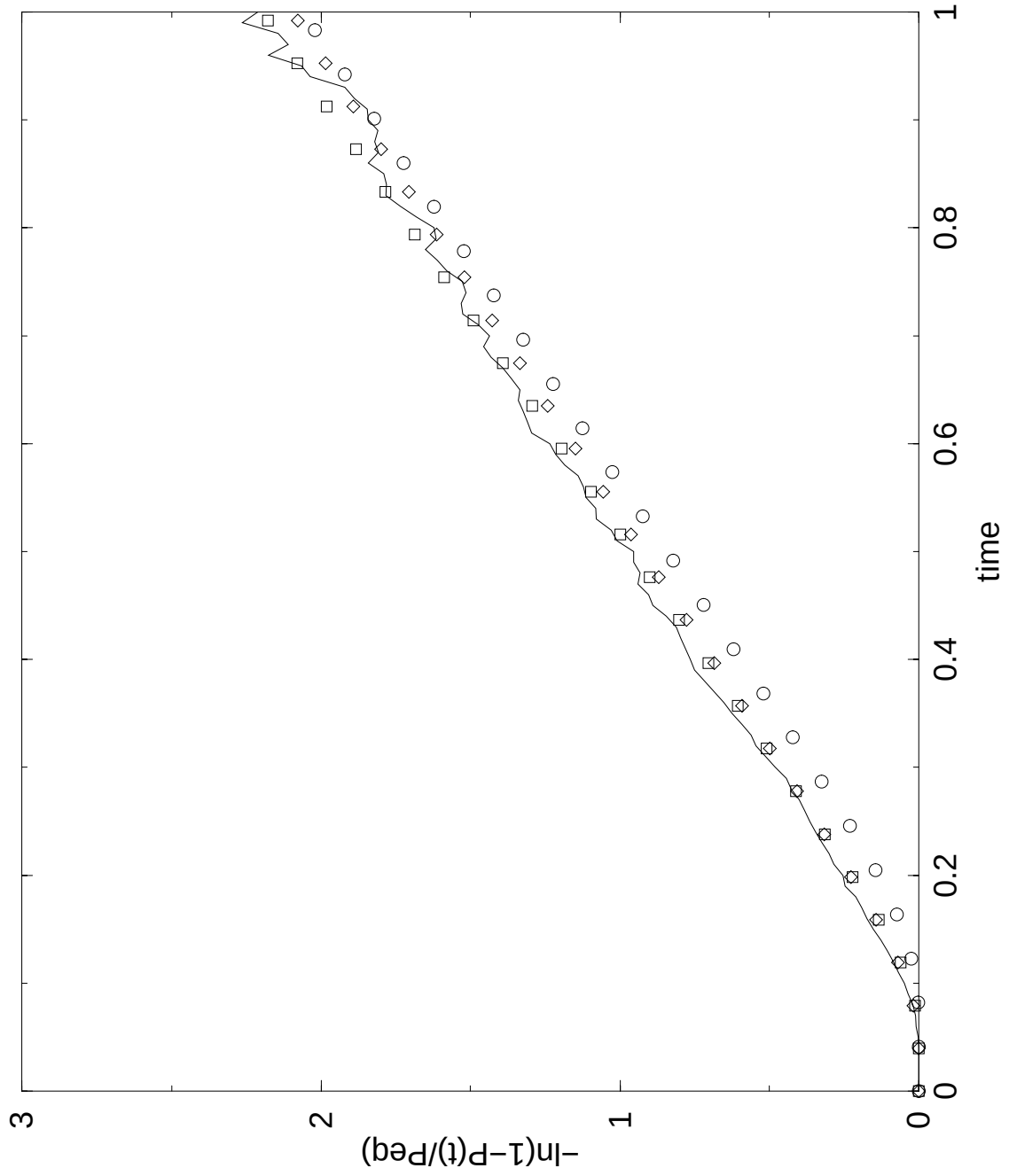


Figure 2b

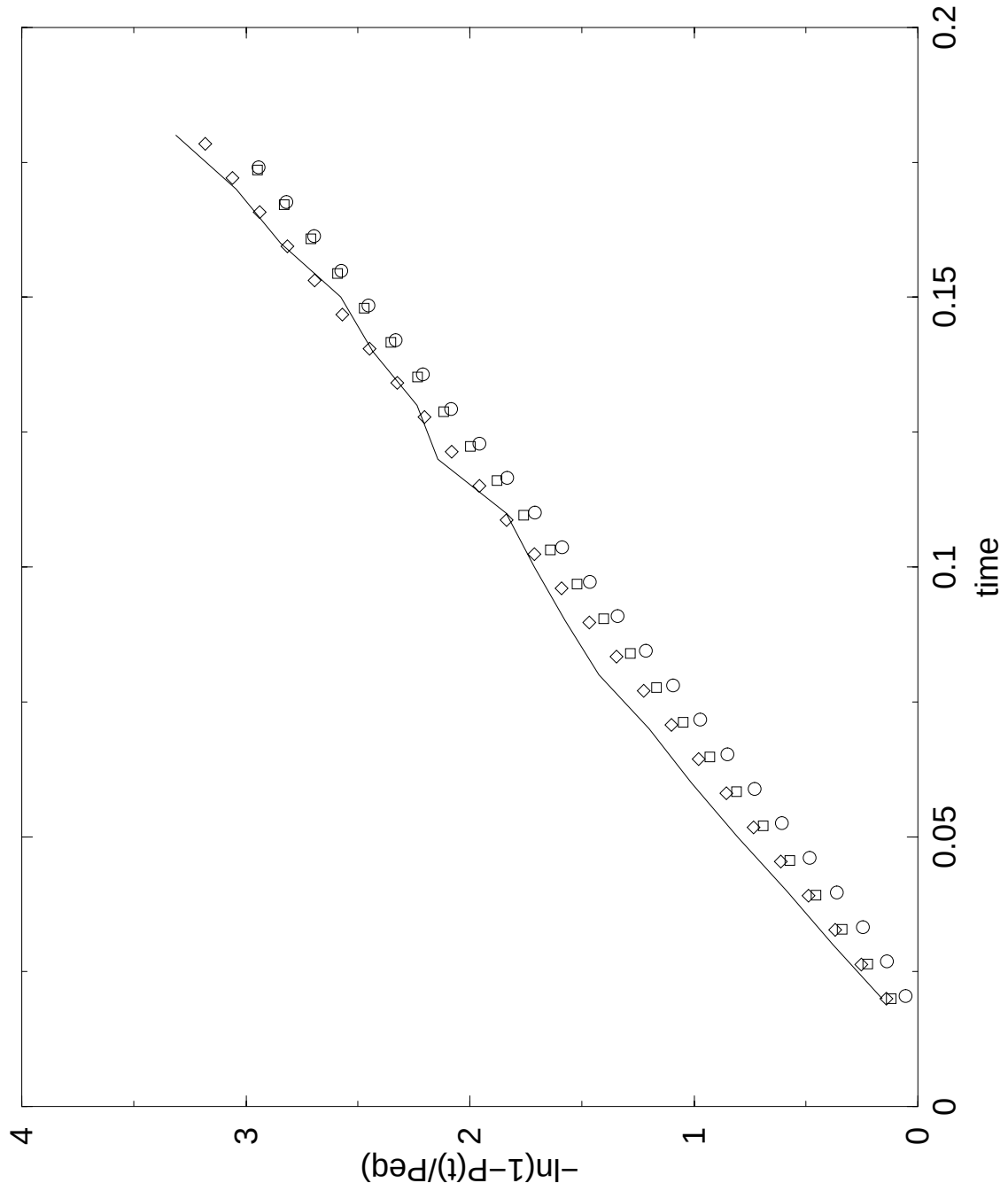


Figure 3

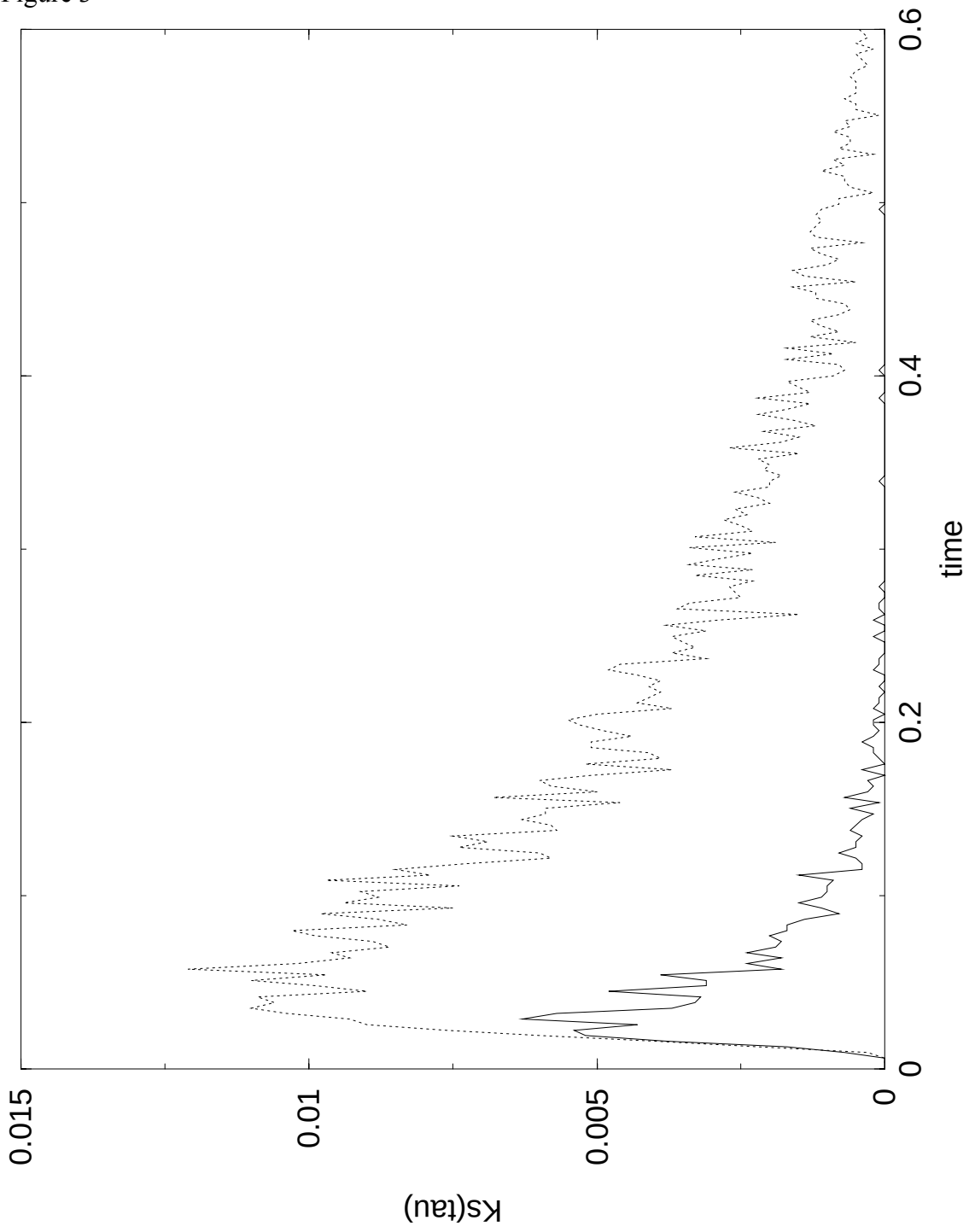


Figure 4

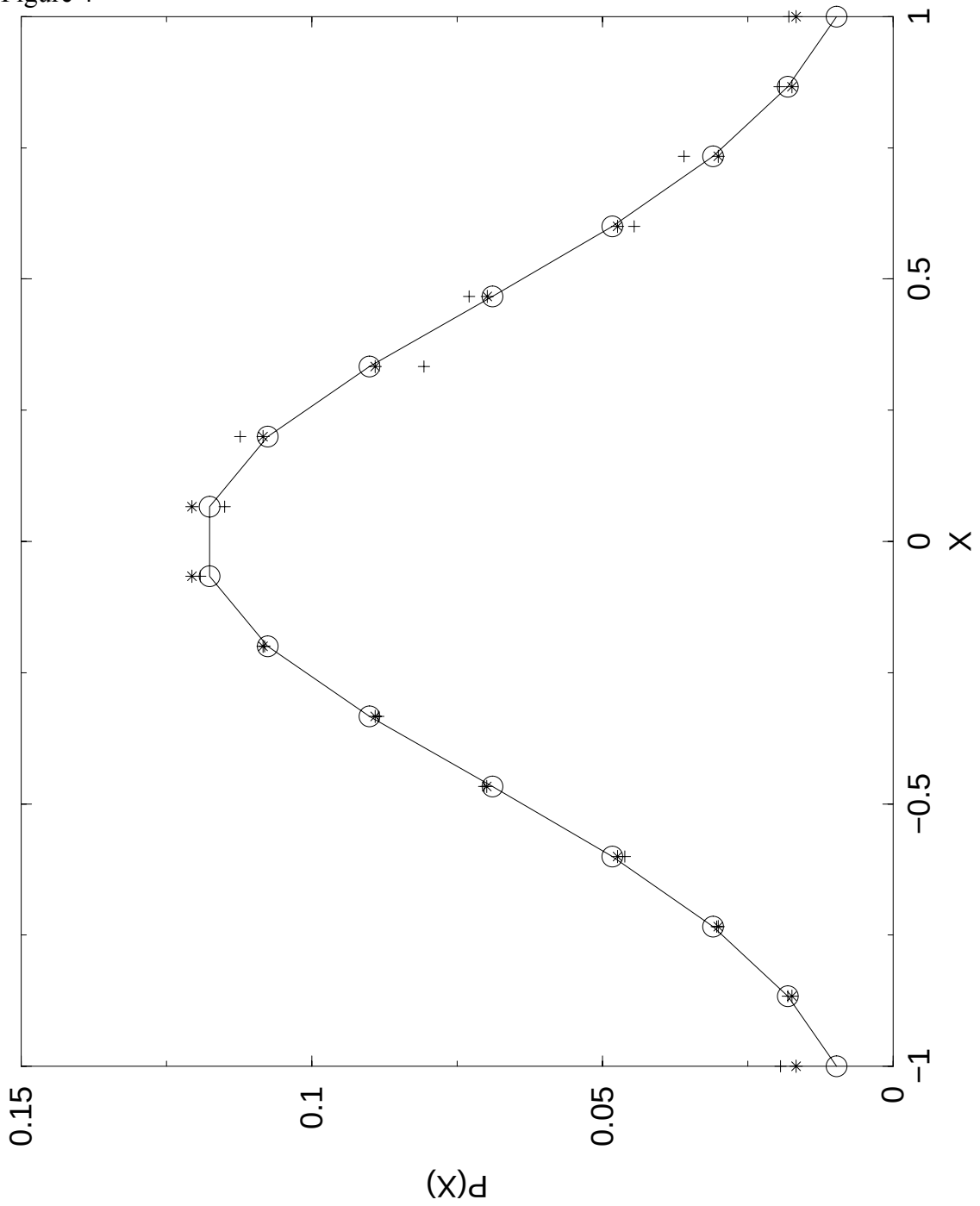


Figure 5

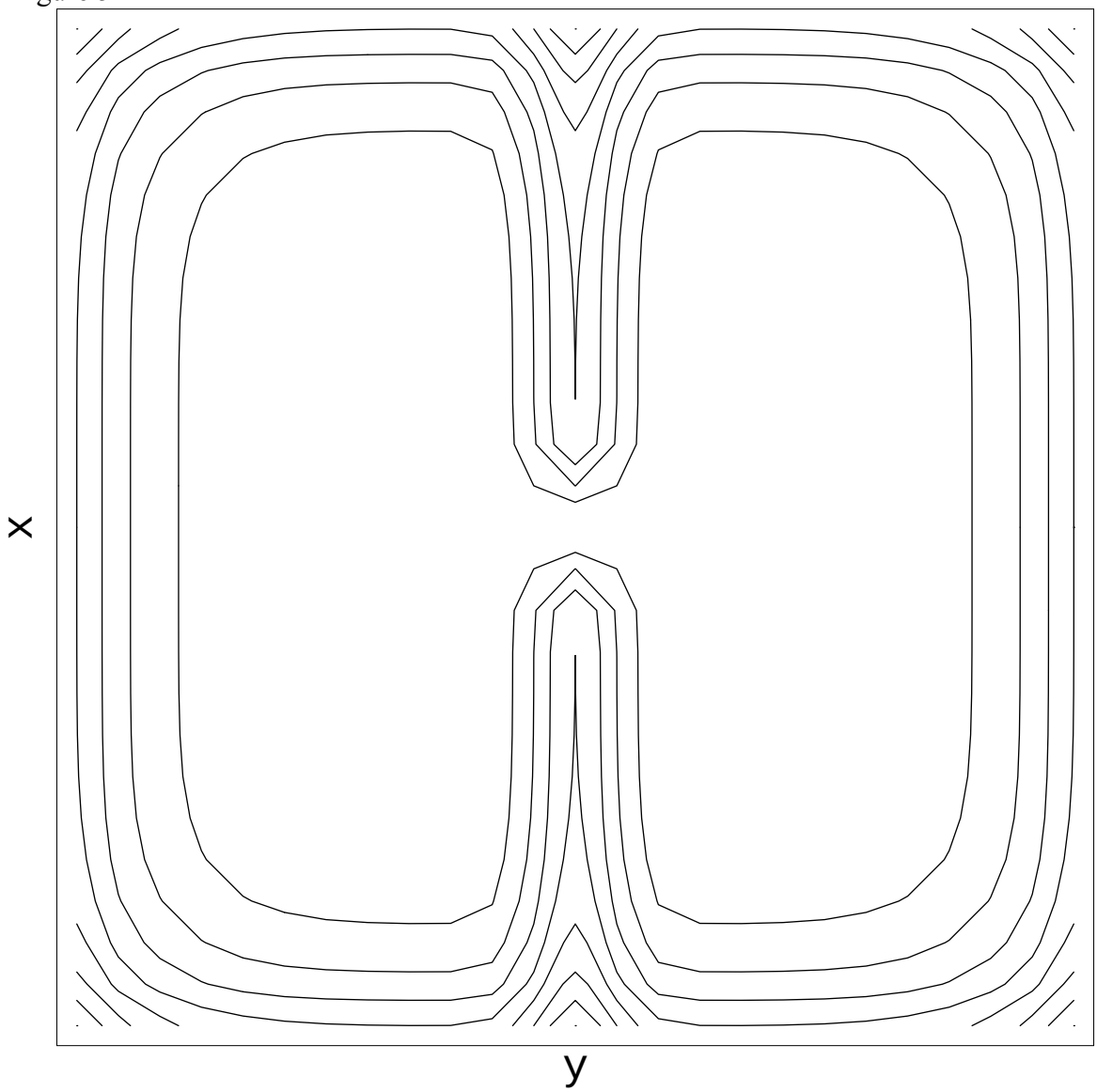


Figure 6
Figure 6a

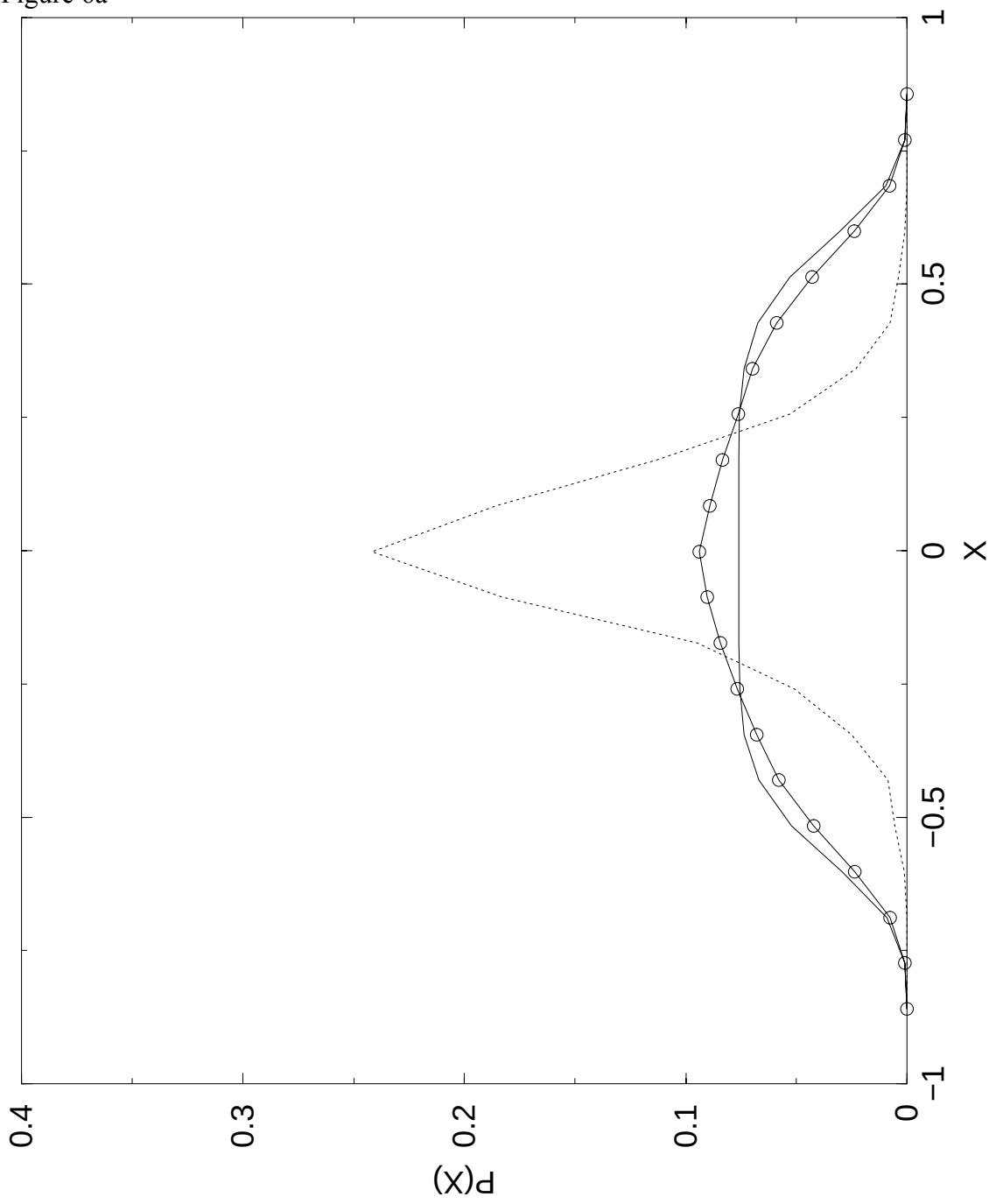


Figure 6b

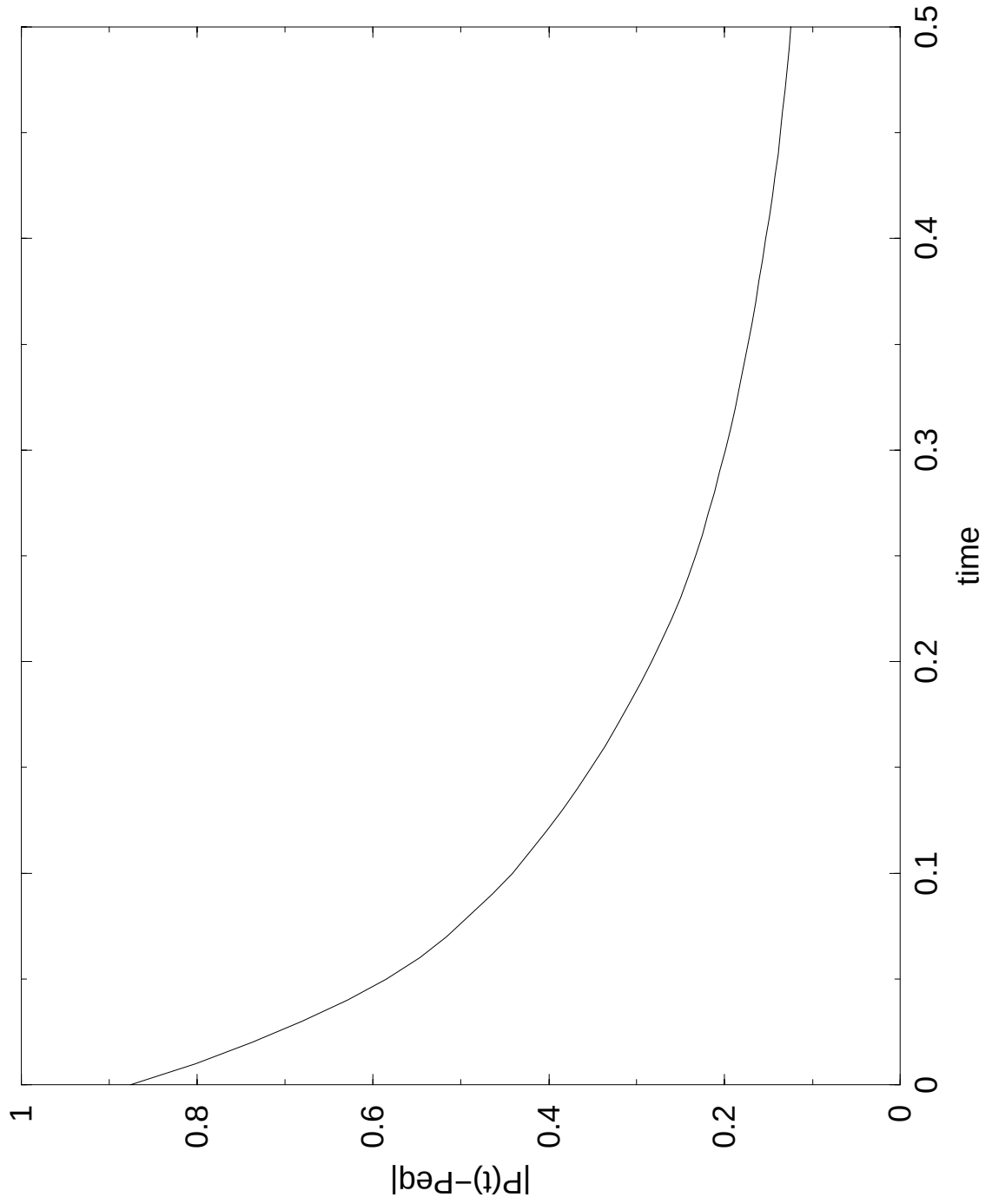


Figure 7
Figure 7a

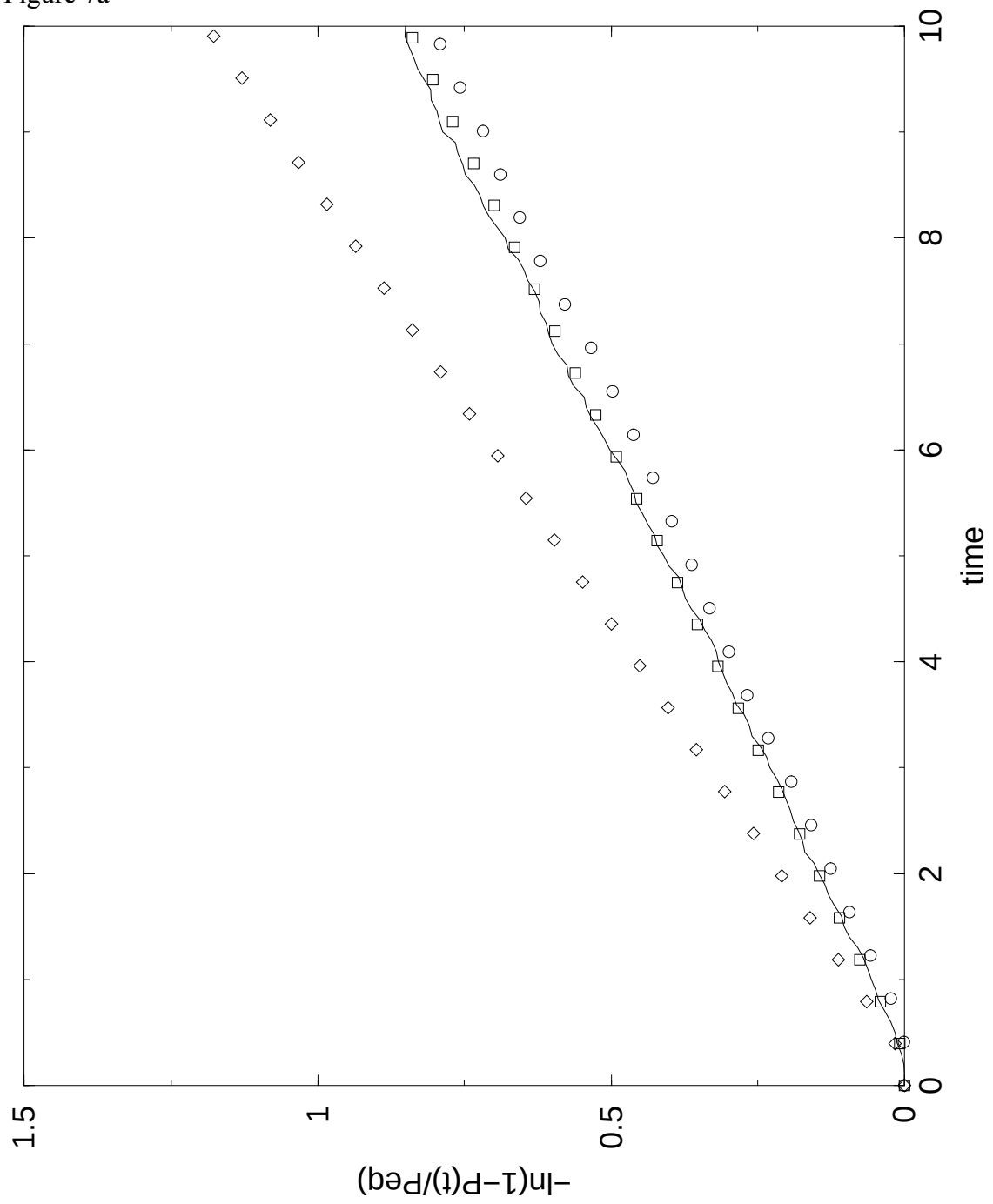


Figure 7b

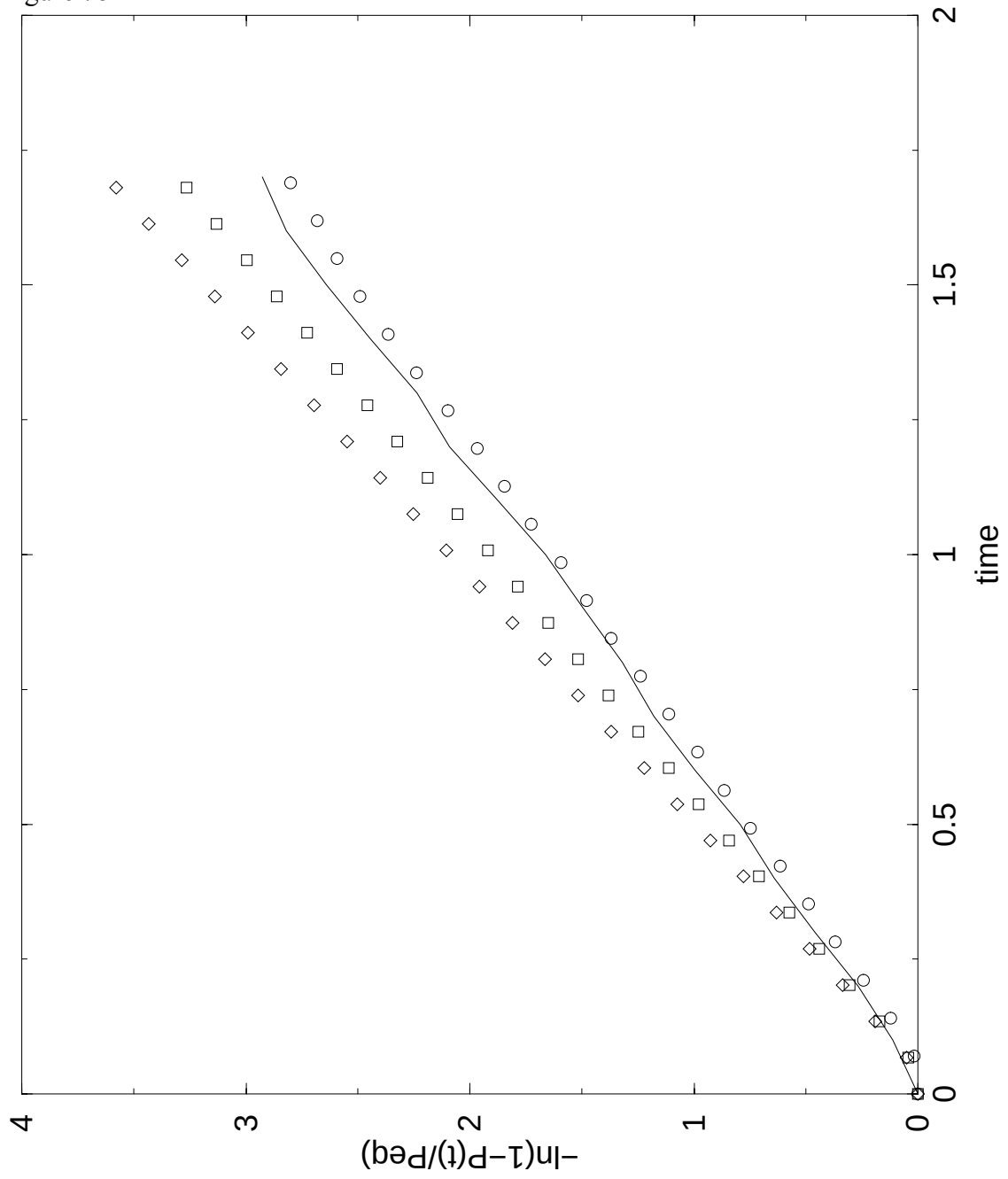


Figure 8

Figure 8a

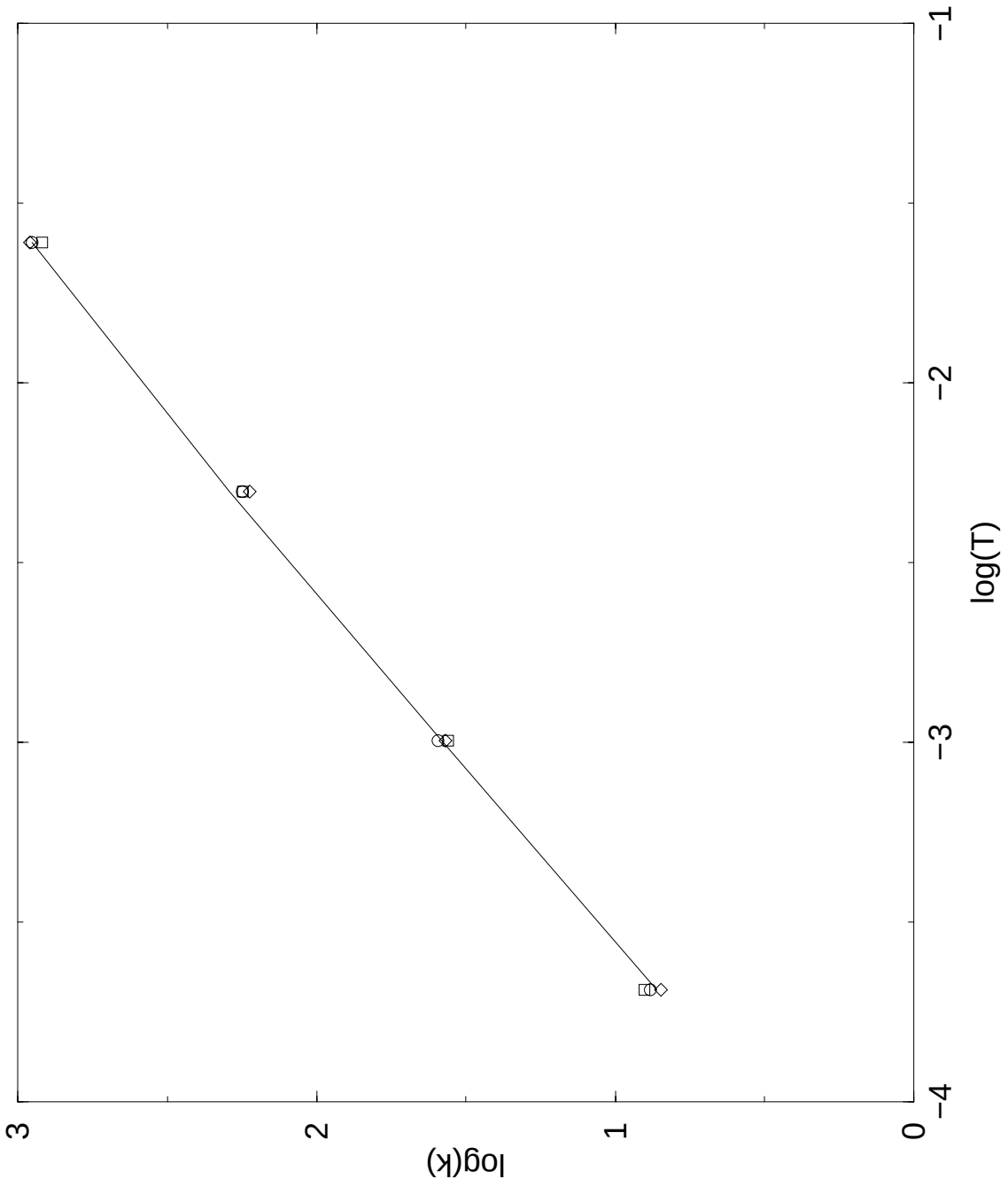
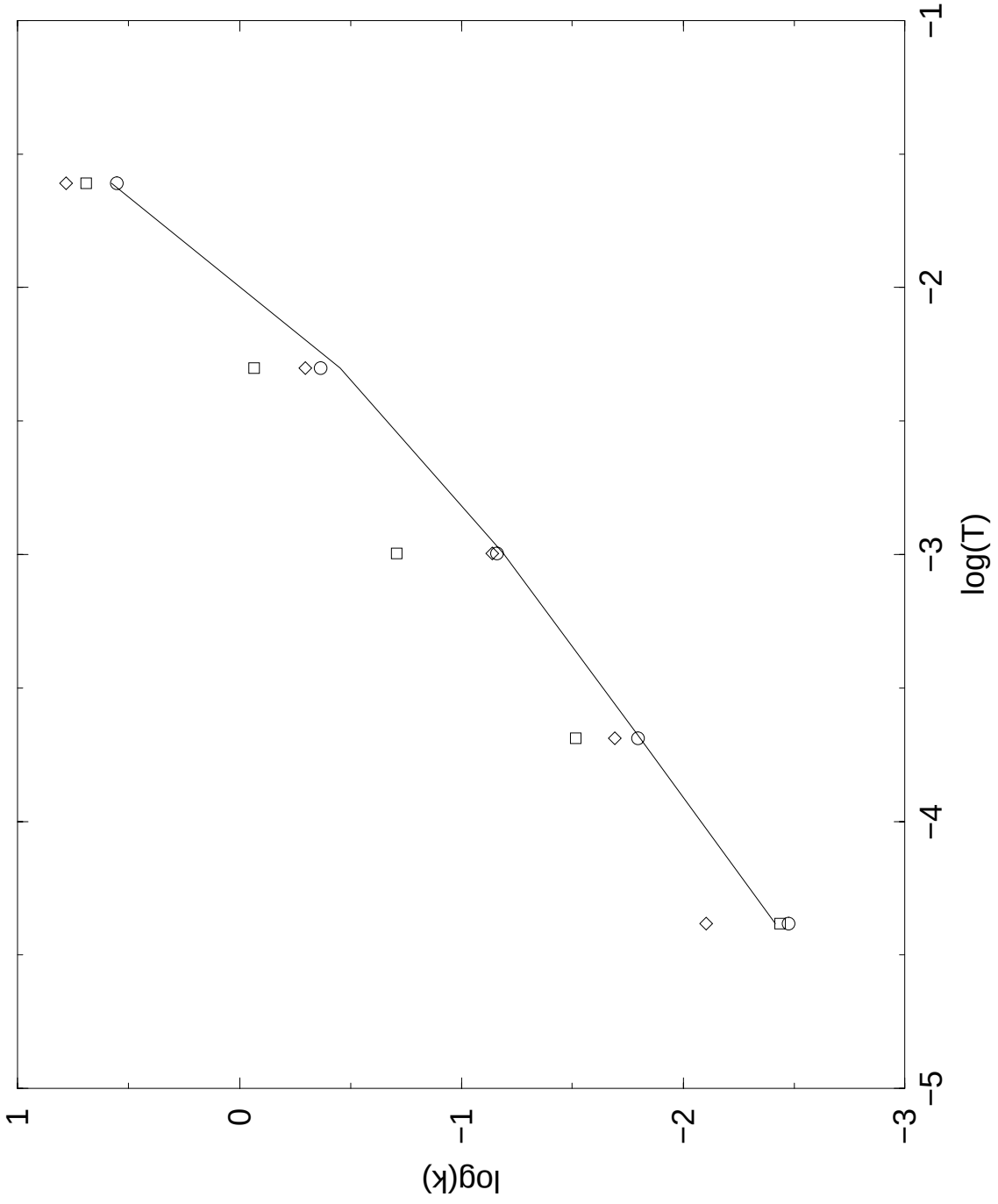


Figure 8b



## References


[1] P. G. Bolhuis, D. Chandler, C. Dellago, et al., Annual Review of Physical Chemistry **53**, 291 (2002).

[2] W. Y. Yang and M. Gruebele, Nature **423**, 193 (2003).

[3] R. Elber, A. Ghosh, A. Cardenas, et al., Adv. Chem. Phys. **126**, 93 (2003).

[4] W. Nowak, R. Czerminski, and E. R., Journal of the American Chemical Society **113**, 5627 (1991).

[5] D. Moroni, P. G. Bolhuis, and T. S. van Erp, submitted (2003).

[6] T. S. van Erp, D. Moroni, and P. G. Bolhuis, Journal of Chemical Physics **118**, 7762 (2003).

[7] D. Zuckerman and T. Woolf, Journal of Chemical Physics **116**, 2586 (2002).

[8] C. W. Gardiner, *Handbook of stochastic methods for physics chemistry and natural sciences* (Springer-Verlag, Berlin, New York, 1942).

[9] Zwanzig R., *Non-equilibrium Statistical Mechanics* (Oxford University Press, Oxford UK, 2002).

[10] V. M. Kenkre, E. W. Montroll, and M. F. Shlesinger, Journal of Statistical Physics **9**, 45 (1973).

[11] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Clarendon Press, Oxford, 1987).