Richard C. Kraaij
Delft Institute of Applied Mathematics
Delft University of Technology
r.c.kraaij@tudelft.nl

Research Stieltjes Prize 2016

Fluctuations of Markov processes and a connection to Hamilton–Jacobi equations

In 2016 Richard Kraaij has been awarded the Stieltjes Prize for one of the two best PhD theses in mathematics in the Netherlands. The prize was awarded for his thesis entitled *Semigroup Methods for Large Deviations of Markov Processes*, which he completed at Delft University of Technology. After a postdoc in Bochum he returned to Delft as Assistant Professor. In this article he describes his current research on Markov processes.

Fluctuations of simple averages

What more is there to be said about the limiting behaviour of the average of n independent and identically distributed (iid) random variables than the law of large numbers and the central limit theorem?

Consider two examples. Let X_i be a sequence of iid coin tosses:

$$X_i = \begin{cases} -1 & \text{with probability } \frac{1}{2}, \\ 1 & \text{with probability } \frac{1}{2}. \end{cases}$$

Let Y_i be a sequence of standard normal random variables $Y_i \sim \mathcal{N}(0,1)$:

$$\mathbb{P}[Y_i \leq y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{1}{2}x^2} dx.$$

In both examples $\mathbb{E}[X_i] = \mathbb{E}[Y_i] = 0$ and $\mathsf{Var}(X_i) = \mathsf{Var}(Y_i) = 1$. This means that by the law of large numbers

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 0, \quad \frac{1}{n}\sum_{i=1}^{n} Y_i \to 0,$$

almost surely, whereas by the central limit theorem the rescaled averages converge weakly to a standard normal random variable

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \to \mathcal{N}(0,1), \quad \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i \to \mathcal{N}(0,1).$$

Does this mean that both sequences behave the same?

That of course depends on which question one would like to answer. Suppose that you feel lucky and go to play some games in the casino, in which two fair games can be played. Denote by X_i and Y_i the pay-offs of these games. Clearly, the casino wants to make a profit, so they charge a price a to play a game. What is the probability that after n games, you have gained some money?

Here a difference must turn up: if $a \geq 1$, then each game of the first type will make a loss with probability one, whereas for games of the second type, there still is a chance to make a profit.

To study more precisely the asymptotic behaviour of the probability to gain some money, that is the event that $\sum_{i=1}^{n} X_i \geq na$ or similarly for the games of type Y, we next consider upper bounds for the probability

$$\mathbb{P}\left[\sum_{i=1}^{n} X_i \geq na\right] = \mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right], \quad a \geq 0.$$

This we do via Markov's inequality. For an arbitrary $\lambda > 0$, this yields

$$\begin{aligned} \mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right] &= \mathbb{P}\left[e^{\lambda \sum_{i=1}^{n} X_i} \geq e^{\lambda n a}\right] \\ &\leq e^{-na\lambda}\mathbb{E}\left[e^{\lambda \sum_{i=1}^{n} X_i}\right] \\ &= \exp\left\{-n\left(\lambda a - \log \mathbb{E}\left[e^{\lambda X_1}\right]\right)\right\}. \end{aligned} \tag{1}$$

Denote $I_X(a) = \sup_{\lambda \geq 0}\left\{\lambda a - \log \mathbb{E}\left[e^{\lambda X_1}\right]\right\}$. By optimizing over λ, we obtain

$$\mathbb{P}\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq a\right] \leq e^{-nI_X(a)}. \tag{2}$$

A direct calculation of I_X and I_Y yields

$$I_X(a) = \begin{cases} \frac{1-a}{2}\log(1-a) \\ \quad + \frac{1+a}{2}\log(1+a) & \text{if } a \in [0,1], \\ \infty & \text{if } a \geq 1, \end{cases}$$

$$I_Y(a) = \frac{1}{2}a^2.$$

Even though we have only proven an upper bound, on an exponential scale, the bound in (2) is sharp. A combination of an upper bound with a corresponding lower bound on an exponential scale is called a *large deviation principle* (LDP).

The law of large numbers is reflected by $I_X(0) = I_Y(0) = 0$. That the central limit theorem holds for both sequences of averages is reflected by the fact that I_X and I_Y have the same second order Taylor expansion around 0, see Bryc [2]. Finally, note that $I_Y(a) < I_X(a)$ for all $a > 0$, with the difference becoming larger for large values of a. This means that we have a much larger probability (even though both exponentially small) to win some money in the long run by playing games of type Y rather than by playing games of type X.

A general framework

To prepare for some more involved large deviation principles, we extend our set-up. A general framework for the study of large deviations was introduced by Varadhan [13], who was awarded the Abel prize for his contributions to large deviation theory. We say that random variables Z_n on a space E satisfy an LDP with rate function $I : E \to [0,\infty]$ if

$$\limsup_n \frac{1}{n} \log \mathbb{P}[Z_n \in G] \leq -\inf_{a\in G} I(a) \quad \text{for all closed sets } G,$$

$$\liminf_n \frac{1}{n} \log \mathbb{P}[Z_n \in U] \geq -\inf_{a\in U} I(a) \quad \text{for all open sets } U.$$

Indeed, the large deviation principle can be interpreted as

$$\mathbb{P}[Z_n \approx a] \asymp e^{-nI(a)}$$

by considering the large deviation upper and lower bound for open and closed balls of small radius. In the case of previous section, playing the game of type X, leads to $Z_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $G = [a,\infty)$.

The procedure to obtain an LDP in terms of the logarithm of the Laplace transform in (1) and (2) was generalized to this general setting by Bryc [3].[1] He established that the random variables Z_n satisfy an LDP if[2] for all continuous and bounded functions f the limit

$$\Lambda_f := \lim_n \frac{1}{n} \log \mathbb{E}[e^{nf(Z_n)}] \tag{3}$$

exists. In that case the rate function is given by

$$I(a) = \sup_f f(a) - \Lambda_f. \tag{4}$$

Having introduced this more general machinery, we can consider large deviations for more difficult objects. We proceed by considering the large deviations of *processes*.

Adding a temporal component

The question of making a profit after n games is not natural. We would rather ask for the probability that at no single time we have made a loss, that is: the probability of not going bankrupt. After all, having gone bankrupt, there is no way to recover.

Say that we play n games per hour. Then the rescaled total pay-off (if we play with games of type X) at time $t \geq 0$ (time in hours) is

$$t \mapsto Z_{X,n}(t) := \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} X_i. \tag{5}$$

We can rephrase our large deviation principle of the previous section as

$$\mathbb{P}\big[Z_{X,n}(1) \geq a\big] \asymp e^{-nI_X(a)}.$$

The probability of not going bankrupt in the first hour translates into

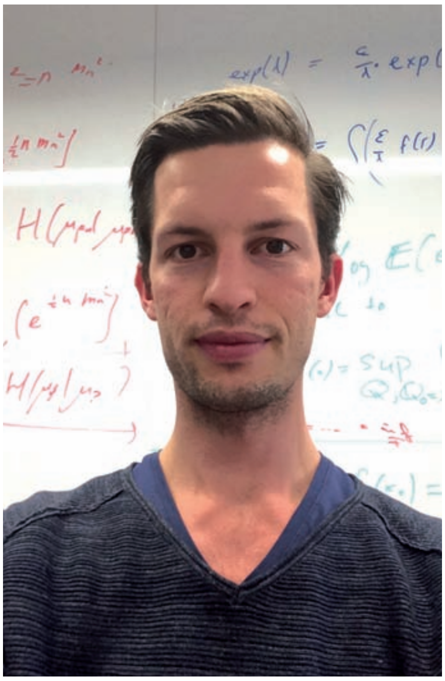
$$\mathbb{P}\big[\forall t \leq 1 \colon Z_{X,n}(t) \geq at\big]. \tag{6}$$

Even though this probability is evidently smaller, the exponential asymptotic rate of decay is also $I_X(a)$!

We explore where this decay rate comes from. Consider a large deviation at time $s \in (0,1)$ of size sa, as well as a large deviation of size a at time $t = 1$:

$$\mathbb{P}\big[Z_{X,n}(s) \approx sa, Z_{X,n}(1) \approx a\big].$$

By the Markov property, we can condition on the state at time s, and find that this probability equals



Richard Kraaij

$$\begin{aligned}&\mathbb{P}\big[Z_{X,n}(s) \approx as\big]\mathbb{P}\big[Z_{X,n}(1) \approx a \,|\, Z_{X,n}(s) \approx as\big] \\ &= \mathbb{P}\big[Z_{X,n}(s) \approx as\big]\mathbb{P}\big[Z_{X,n}(1-s) \approx a(1-s)\big].\end{aligned} \tag{7}$$

By appropriate rescaling in the calculation that lead to (2), we find

$$\begin{aligned}&\mathbb{P}\big[Z_{X,n}(s) \approx sa, Z_{X,n}(1) \approx a\big] \\ &\asymp e^{-n(sI_X(a) + (1-s)I_X(a))} \\ &= e^{-nI_X(a)},\end{aligned} \tag{8}$$

which indeed gives us the same rate $I_X(a)$. On the other hand, a general structure can be recognized from the middle term by noting that if γ is a the function $\gamma(t) = at$, then $a = \frac{\gamma(s)-\gamma(0)}{s}$ and $a = \frac{\gamma(1)-\gamma(s)}{1-s}$. Thus, the term in the exponent equals

$$-n\Big(sI_x\Big(\frac{\gamma(s)-\gamma(0)}{s}\Big) + (1-s)I_X\Big(\frac{\gamma(1)-\gamma(s)}{1-s}\Big)\Big).$$

For more general trajectories γ, one finds similarly that the probability of Z_n being close to γ at times t_i equals

$$\begin{aligned}&\mathbb{P}\big[Z_{X,n}(t_i) \approx \gamma(t_i), i \in \{0,\ldots,k\}\big] \\ &\asymp e^{-n\sum_{i=1}^k (t_i - t_{i-1}) I_X\left(\frac{\gamma(t_i)-\gamma(t_{i-1})}{t_i - t_{i-1}}\right)}.\end{aligned} \tag{9}$$

Thus, the exponent contains a Riemann sum of terms I_X that are evaluated in difference quotients along the trajectory γ. By taking more and more times at which we consider a large deviation, a general principle appears, from which the asymptotic decay rate for (6) can be derived.[3]

Theorem 1 (Mogulskii [12]). *Consider iid random variables U_i and consider $t \mapsto Z_{U,n}(t)$ as in (5).*[4] *Denote by*

$$\Lambda_U(\lambda) := \log \mathbb{E}[e^{\lambda U}],$$
$$I_U(a) := \sup_\lambda \{\lambda a - \Lambda_U(\lambda)\}.$$

Then

$$\mathbb{P}\big[\{Z_{U,n}(t)\}_{t\in[0,1]} \approx \{\gamma(t)\}_{t\in[0,1]}\big] \asymp e^{-nJ_U(\gamma)}.$$

The function J_U is given by

$$J_U(\gamma) = \begin{cases} \int_0^1 I_U(\dot\gamma(s))\, ds & \text{if } \gamma \text{ is absolutely continuous,} \\ \infty & \text{otherwise} \end{cases}$$

As an unavoidable consequence of the more complex set-up of our problem, we end up with a more complex rate function J_U. Note, however, the similarities of Mogulskii result to (1) and (2) or (3) and (4). Indeed, the optimization procedure that lead to (2) and (4) is replaced by the integral over an expression that is obtained from a similar optimization procedure that is carried out for each infinitesimal interval of time.

What we have gained from Mogulskii's result is that in addition to the asymptotics corresponding to never going bankrupt, we can now calculate the asymptotics of any possible trajectory of capital gains. For example, the casino might decide that instead of having a fixed price to play a game, the price is dependent on the time of the day that the game is played. Playing during the morning, in off peak hours is cheap, whereas playing during the evening is expensive. By Mogulskii's theorem we can still calculate the asymptotics of not going bankrupt.

The key components in Mogulskii's theorem

The structure of the question above becomes more complex if the price of playing a game becomes dependent on how much money you have earned on average in the past (e.g. a progressive policy in which rich people pay more). To compute the asymptotic probability of not going bankrupt, we now have to take into account the price we pay for each game in the stochastic process $Z_n(t)$. In addition, to make the analysis a little easier, we assume that instead [5] of playing at times k/n, we assume that the time between the played games has an exponential distribution with mean $1/n$. If a game is played at time t, your rescaled wealth changes at time t by

$$\frac{1}{n}U_t - \frac{1}{n}a(Z_n(t)).$$

Here U_t (e.g. a game of type X or Y) is the pay-off of the game played at time t and $a(Z_n(t))$ is the cost of this game. The probability of not going bankrupt in the first hour translates into

$$\mathbb{P}[\forall t \leq 1\colon Z_n(t) \geq 0].$$

To extend the analysis to this more complex process, we identify the three crucial components of the derivation above:

1. We chose a finer and finer mesh of times in the interval $[0,1]$.
2. The rate function for these meshes had an additive structure due to the Markov property of the process in (5), see (7) and (9).
3. We were able to explicitly calculate what was the limiting rate function due to the simple form of the rate function for a finite set of times.

For our more general Markov processes $t \mapsto Z_n(t)$, clearly step 1 does not lead to any issues. By using the Markov property, we can reproduce an analogue of the first line of (7), but not the second, and it is not clear that on an exponential scale, these probabilities have suitable asymptotic behaviour. In particular, taking two times $0 \leq s < t$ and a point x it is not clear that there is a function $J_{t-s}(\cdot \mid x)$ such that

$$\mathbb{P}[Z_n(t) \approx y \mid Z_n(s) \approx x] \asymp e^{-nJ_{t-s}(y \mid x)}. \quad (10)$$

If there is a family of such functions, we do not necessarily know what they look like. Thus, we are not immediately able to derive a rate function for the full process, as we were able to do for Mogulskii's theorem.

Feng and Kurtz [9] made a major step towards a general resolution of these issues. Let us follow their first step towards establishing the asymptotic behaviour in (10). The resolution of this issue will lead to a key concept that allows us to solve the second problem as well.

To compute the large deviation asymptotics for (10), we turn to Bryc's result. Based on (3), we aim to study the logarithm of the exponential moments of the process at some t, given that at some earlier time we are at x. This leads us to define a collection of (non-linear) operators $\{V_n(t)\}_{t \geq 0}$ on the space of bounded continuous functions

$$V_n(t)f(x) = \frac{1}{n}\log \mathbb{E}\left[e^{nf(Z_n(t))} \mid Z_n(0) = x\right].$$

Note that the Markov property of our process yields that V_n is a semigroup with respect to time: $V_n(t)V_n(s)f = V_n(t+s)f$ and $V_n(0) = \mathbb{1}$. By Bryc's theorem and the Markov property, it follows that the LDP for the vector $(Z_n(t_0), Z_n(t_1), Z_n(t_2), \ldots, Z_n(t_k))$ with $0 = t_0 < t_1 < \cdots < t_k$ holds if there is a operator semigroup $V(t)$, such that

$$V_n(t)f \to V(t)f. \quad (11)$$

The function $J_{t-s}(y \mid x)$ of (10) is then given, see (4), by

$$J_{t-s}(y \mid x) = \sup_f f(y) - V(t-s)f(x).$$

Putting all increments together, we see that the vector $(Z_n(t_1), Z_n(t_2), \ldots, Z_n(t_k))$ with $0 = t_0 < t_1 < \cdots < t_k$ satisfies an LDP:

$$\mathbb{P}[\forall i \in \{1,\ldots,k\}\colon Z_n(t_i) \approx \gamma(t_i)] \asymp \exp\left\{-n\sum_{i=1}^{k} J_{t_i - t_{i-1}}(\gamma(t_i) \mid \gamma(t_{i-1}))\right\}. \quad (12)$$

This rate function is the analogue of (9) in the setting of playing games in the casino with a progressive pricing policy. Note, however, that all steps in this derivation can be carried out for a general Markov process. We conclude that the large deviation principle of the trajectories of a sequence of Markov processes can be traced back to the convergence of non-linear operator semigroups.

At this moment, it is not clear whether there is an explicit form for the functions J_t, nor of the limiting rate function if we take a finer and finer mesh of times. This problem has to wait until we have better grip on the concepts that can be used to study the convergence of non-linear semigroups, which is an analytic problem that we treat next.

Operator semigroups and their generators

Consider the following problem posed by Cauchy [4]. Find all maps $f\colon \mathbb{R}^+ \to \mathbb{C}$ satisfying

$$\begin{cases} \phi(t+s) = \phi(t)\phi(s) & \text{for all } s,t \geq 0, \\ \phi(0) = 1. \end{cases}$$

Assuming that ϕ is continuous [6], it can be shown that all maps of this type are of the type $\phi_a(t) := e^{ta}$ with $a \in \mathbb{C}$.

The factor a, which can be found by $a = \partial_t \phi_a(t) \mid_{t=0}$, captures all essential information of the semigroup ϕ_a. In addition, the dependence of ϕ_a on a is robust under convergence: for a sequence of $a_n \in \mathbb{C}$ with $a_n \to a$, it holds that $\phi_{a_n} \to \phi_a$ uniformly on compacts.

We use this insight to tackle the convergence question posed in (11).

Consider any of the non-linear semigroups $\{V_n(t)\}_{t \geq 0}$. We formally define the (non-linear and unbounded) operators H_n by

$$H_n f := \partial_t V_n(t) f \mid_{t=0} \quad (13)$$

for f for which the right-hand side exists.[7]

In analogy to the setting that Cauchy considered, we aim to find an operator H such that $H_n \to H$, to write $V_n(t) = e^{tH_n}$ and $V(t) = e^{tH}$, and to conclude that $V_n(t)f \to V(t)f$.

This statement is easier said than done. Taking the exponent of an unbounded non-linear operator needs some care. In particular, defining the exponent in terms of a power series gives trouble as we are taking a sum over compositions of unbounded operators. Following Hille [10], we turn to Euler's procedure to obtain the exponential, namely, we aim to write

$$V_n(t)f = \lim_{n\to\infty}\left(\mathbb{1} - \frac{t}{n}H_n\right)^{-n} f. \quad (14)$$

Suppose that for all $\lambda > 0$ the resolvent $R_n(\lambda) := (\mathbb{1} - \lambda H_n)^{-1}$ exists, then by general theory $R_n(\lambda)$ is a continuous operator.[8] Thus, the formula (14) involves only bounded operators, which are much easier to handle than a definition of e^{tH_n} in terms of a power series.

To answer the question whether for all $\lambda > 0$ the resolvent $R_n(\lambda) := (\mathbb{1} - \lambda H_n)^{-1}$ exists as a non-linear operator, we need to be able to solve for all functions h and $\lambda > 0$ the equation

$$f - \lambda H_n f = h.$$

This equation is called the Hamilton–Jacobi equation for the operator H_n. Using the Markovian structure of the process, we can always find such a solution, and therefore we have a resolvent.

Finally, to come back to the convergence question, the following result (glossing over technical definitions and issues) was proven by Feng and Kurtz [9].

Theorem 2 (Feng–Kurtz extension of the Trotter–Kato theorem). *Put* $H_n = \partial_t V_n(t)\,|_{t=0}$. *Suppose that* H *is an operator such that*

- $H \subseteq \mathrm{LIM}\, H_n$ *in the sense that if* $(f,g) \in H$ *(where we interpret* H *as a graph), then there are* $(f_n, g_n) \in H_n$ *such that* $\lim_n f_n = f$ *and* $\lim_n g_n = g$;
- *for all* h *and* $\lambda > 0$ *there is a unique viscosity solution to the Hamilton-Jacobi equation* $f - \lambda H f = h$.[9]

Then there is a semigroup $V(t) = \lim_n (\mathbb{1} - \frac{t}{n}H)^{-n}$ *and we have for all* f *and* $t \geq 0$ that $\lim_n V_n(t) f = V(t) f$.

The key component idea in the proof of this theorem is to show that the convergence of Hamiltonians allows one to show that solutions to $f - \lambda H_n f = h$ give a candidate solution for the limiting equation $f - \lambda H f = h$. Uniqueness of solutions makes this procedure robust, i.e. the resolvents converge in a 'strong' topology, and this can be used to derive the convergence of semigroups.

We apply this in the setting of our casino game with wealth-dependent cost for each game. In the case that we play games of type X, then the operator H_n is given by

$$H_{n,X} f(x) = \frac{1}{2}\left[e^{n\left(f\left(x + \frac{1-a(x)}{n}\right) - f(x)\right)} - 1\right] + \frac{1}{2}\left[e^{n\left(f\left(x - \frac{1-a(x)}{n}\right) - f(x)\right)} - 1\right].$$

The operator carries in its structure the 50% chance of winning a game, after which the wealth changes by $1 - a(x)$, and the 50% chance of losing $1 + a(x)$. This operator has a limit H given by

$$H_X f(x) = \frac{1}{2}\left[e^{f'(x)(1-a(x))} - 1\right] + \frac{1}{2}\left[e^{-f'(x)(1+a(x))} - 1\right]. \quad (15)$$

In the case that we play with games of type Y, we similarly obtain

$$H_{n,Y} f(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\left[e^{n\left(f\left(x + \frac{y - a(x)}{n}\right) - f(x)\right)} - 1\right]$$

and a limiting operator

$$H_Y f(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\left[e^{f'(x)(y - a(x))} - 1\right]. \quad (16)$$

In both cases,[10] there are unique viscosity solutions for the Hamilton-Jacobi equation and we obtain that $V_n(t) \to V(t)$.

Large deviations

As we have seen, Theorem 2, combined with the discussion on (11), implies that large deviation principles for the trajectories of a sequence of Markov processes can be traced back to the two conditions of Theorem 2 and the verification of a compactness property.[11] The rate function is in terms of a limit of a Riemann-sum like the expression in (12).

To conclude our analysis, we re-express the large deviation rate function in a more accessible form. This can be done using control theory. As in (15) and (16), the operator H can usually be expressed in the form $Hf(x) = \mathcal{H}(x, f'(x))$ where $\mathcal{H}$ is a function of space and 'momentum'. In analogy with the basic notions in mechanics, we call $\mathcal{H}$ a *Hamiltonian*.[12] As in Mogulskii's theorem, Theorem 1, one obtains from the Hamiltonian $\mathcal{H}$ a *Lagrangian* $\mathcal{L}(x,v) = \sup_p \langle p, v\rangle - \mathcal{H}(x,p)$. Using this Lagrangian one can use control theory to define an operator $\hat{R}(\lambda)$ such that the function $\hat{R}(\lambda)h$ is a viscosity solution to $f - \lambda H f = h$. If the Hamilton–Jacobi equation has a unique viscosity solution, both our solutions must be the same. Thus, we have found an explicit form for the resolvent, that can be used to give an explicit form for the semigroup $V(t)$ and the conditional rate functions J_t. This leads to a generalization of Theorem 1.

Theorem 3 (Feng and Kurtz). *Let* $Z_n(t)$ *be Markov processes with semigroups* $V_n(t)$ *and operators* H_n *that satisfy the conditions of Theorem 2.*[13] *Then a large deviation principle holds with rate function* J *given by*

$$\begin{cases} \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s))\, ds & \text{if } \gamma \text{ is absolutely continuous,} \\ \infty & \text{otherwise.} \end{cases}$$

By evaluating the Lagrangians $\mathcal{L}_X$ and $\mathcal{L}_Y$ that correspond to (15) and (16), we have established the asymptotic form of the probability that one does not go bankrupt in a casino with a progressive cost for playing games.

We conclude that this method leads to an effective way of establishing large deviation principles for Markov processes by turning probabilistic questions into analytic ones. In particular, the main challenge is to establish uniqueness of viscosity solutions of Hamilton–Jacobi equations. Following the introduction of viscosity solutions in the 1980's by Crandall and Lions [6], see also Crandall, Ishii and Lions [5] for an overview of the early results, well-posedness was established for a class of Hamilton–Jacobi equations on open domains for which the operator H was either Lipschitz or coercive, see Bardi and Capuzzo-Dolcetta [1].

Application of these methods in statistical physics, even for the simplest model like the Curie–Weiss model which tries to explain phase-transitions in ferromagnetic materials like iron, leads to Hamiltonians of the type ($\beta \geq 0$)

$$\begin{aligned} \hat{H} f(x) &= \hat{\mathcal{H}}(x, f'(x)) \\ &= \frac{1-x}{2} e^{\beta x}\left[e^{2f'(x)} - 1\right] \\ &\quad + \frac{1+x}{2} e^{-\beta x}\left[e^{-2f'(x)} - 1\right], \\ &\quad x \in [-1, 1]. \end{aligned} \quad (17)$$

In this case the domain is the closed interval $[-1,1]$, the exponentials make the operator non-Lipschitz, and the pre-factors $1 - x$ and $1 + x$ make the operator non-coercive.

Theorem 4 (Kraaij [11]). *For all $\lambda > 0$ and continuous and bounded h, uniqueness holds for the Hamilton-Jacobi equation $f - \lambda \hat{H} f = h$.*

We observe that already the simple class of mean-field models like the Curie–Weiss model leads to non-trivial Hamiltonians that fall outside of the 'classical' setting. Extensions of this setting are considered in the study of chemical reaction networks where similar non-Lipschitz and non-coercive behaviour occurs. Well-posedness in this setting remains an open problem. More generally, the large deviation analysis in statistical physics or mathematical biology introduces various kinds of new Hamiltonians that all fall outside the 'classical' setting and for which well-posedness issues are unsolved. Issues arise from boundary conditions, singular behaviour or infinite dimensional contexts.

Theorem 4 and more recent works by the author, as well as work by Dupuis, Ishii, and Soner [7], Feng and Katsoulakis [8], indicate that these methods are robust also in these more difficult settings as well as in infinite dimensional contexts. But in addition, they indicate that this field has more non-trivial open problems then problems that we are able to solve at the moment.

Biography

- Assistant professor, 2018 onward at the Delft University of Technology.
- Visiting researcher at École Polytechnique of Paris, in September 2017 and May and December of 2018.
- Postdoc, 2016–2018 at the Ruhr University of Bochum, Germany.
- PhD student of Mathematics (cum laude), 2012–2016 at the Delft University of Technology.
- Master of Mathematics (cum laude), 2010–2012 at Free University of Amsterdam,
- Bachelor of Mathematics (cum laude), 2007–2010 at Radboud university of Nijmegen.

Notes

1 Preceding Bryc's work, Varadhan [13] proved that the LDP for the variables Z_n with rate function I implies that (3) holds with $\Lambda_f = \sup_a f(a) - I(a)$. Thus, the definition of the large deviation principle, Bryc's result and Varadhan's result can be considered a non-linear generalization of the Portmanteau theorem. In fact, most large deviation results have weak-convergence counterparts.

2 To be complete: one has to assume exponential tightness as well. This can be interpreted as a sort of compactness property for the sequence of random variables Z_n.

3 Set $\gamma^*(t) = at$. Jensen's inequality can be used to show that indeed $\inf_{\hat{\gamma} \in \{\gamma \mid \gamma(t) \geq at\}} J_X(\hat{\gamma}) = J_X(\gamma^*) = I_X(a)$, so the easiest way not to go bankrupt, is to only barely stay solvent at all times.

4 We also assume that the random variables have exponential moments.

5 The analysis can be carried out for the Markov chain as well, but the formula's that follow are less intuitive.

6 The conclusion holds under much weaker assumptions, e.g. local integrability suffices.

7 If A_n is the infinitesimal generator of the continuous time Markov process Z_n, then $H_n f = n^{-1} e^{-nf} A_n e^{nf}$. If we work with a discrete time Markov chain instead, the analysis can be carried out in similar manner at the cost of some relatively straightforward approximation procedures by taking $H_n f = \frac{1}{n} \log\left(e^{-nf}(P_n - \mathbb{1})e^{nf}\right)$, where P_n is the transition matrix of the Markov chain. This operator is found by replacing the derivative in (13) by the corresponding difference quotient.

8 To establish continuity one uses that the operators H_n satisfy the *maximum principle*, which is satisfied for all operators obtained from Markov processes using this procedure.

9 The notion of viscosity solutions for Hamilton-Jacobi equations is based on the maximum principle of H. This notion is robust enough to effectively treat non-linear equations. For a early reference work, see Crandall, Ishii, and Lions [5].

10 Assuming, for simplicity, that a is a bounded continuous function.

11 Arising from the footnote going with Bryc's theorem, see the discussion around accompanying equation (3).

12 More generally, we interpret the underlying space as a manifold and write $Hf(x) = \mathcal{H}(x, df(x))$, where $\mathcal{H}$ is a function on the co-tangent bundle. $\mathcal{L}$ is then a function defined on the tangent bundle.

13 Assuming that the processes Z_n also satisfy some compactness property.

References

1 M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*, Birkhäuser, 1997.

2 W. Bryc, A remark on the connection between the large deviation principle and the central limit theorem, *Statistics & Probability Letters* 18(4) (1993), 253–256.

3 W. Bryc, Large deviations by the asymptotic value method, in *Diffusion Processes and Related Problems in Analysis, Vol. I (Evanston, IL, 1989)*, Progr. Probab., Vol. 22, Birkhäuser, 1990, pp. 447–472.

4 A.-L. Cauchy, *Cours d'analyse de l'École Royale Polytechnique,* 1821.

5 M.G. Crandall, H. Ishii and P.L. Lions, User's guide to viscosity solutions of second order partial differential equations, *Bull. Am. Math. Soc., New Ser.* 27(1) (1992), 1–67.

6 M.G. Crandall and P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* 277(1) (1983), 1–42.

7 P. Dupuis, H. Ishii and H.M. Soner, A viscosity solution approach to the asymptotic analysis of queueing systems, *The Annals of Probability* 18(1) (1990), 226–255.

8 J. Feng and M. Katsoulakis, A comparison principle for Hamilton–Jacobi equations related to controlled gradient flows in infinite dimensions, *Archive for Rational Mechanics and Analysis* 192(2) (2009), 275–310.

9 J. Feng and T.G. Kurtz, *Large Deviations for Stochastic Processes*, American Mathematical Society, 2006.

10 E. Hille, *Functional Analysis and Semi-Groups*, American Mathematical Society Colloquium Publications, Vol. 31, American Mathematical Society, 1948.

11 R.C. Kraaij, Large deviations for finite state Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton–Jacobi equation, *Journal of Statistical Physics* 164(2) (2016), 321–345.

12 A.A. Mogulskii, Large deviations for processes with independent increments, *Ann. Probab.* 21(1) (1993), 202–215.

13 S. Varadhan, Asymptotic probabilities and differential equations, *Communications on Pure and Applied Mathematics* 19(3) (1966), 261–286.