

## Dynamical moderate deviations for the Curie-Weiss model

Collet, Francesca; Kraaij, Richard

**DOI**

[10.1016/j.spa.2017.01.002](https://doi.org/10.1016/j.spa.2017.01.002)

**Publication date**

2017

**Document Version**

Accepted author manuscript

**Published in**

Stochastic Processes and their Applications

**Citation (APA)**

Collet, F., & Kraaij, R. (2017). Dynamical moderate deviations for the Curie-Weiss model. *Stochastic Processes and their Applications*, 127(9), 2900-2925. <https://doi.org/10.1016/j.spa.2017.01.002>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Dynamical moderate deviations for the Curie-Weiss model

Francesca Collet\*      Richard C. Kraaij†

## Abstract

We derive moderate deviation principles for the trajectory of the empirical magnetization of the standard Curie-Weiss model via a general analytic approach based on convergence of generators and uniqueness of viscosity solutions for associated Hamilton-Jacobi equations. The moderate asymptotics depend crucially on the phase under consideration.

*Keywords:* Moderate deviations; Interacting particle systems; Mean-field interaction; Viscosity solutions; Hamilton-Jacobi equation

*MSC[2010]:* primary 60F10; 60J99

## 1 Introduction

The study of the normalized sum of random variables and its asymptotic behavior plays a central role in probability and statistical mechanics. Whenever the variables are independent and have finite variance, the central limit theorem ensures that the sum with square-root normalization converges to a Gaussian distribution. The generalization of this result to dependent variables is particularly interesting in statistical mechanics where the random variables are correlated through an interaction Hamiltonian. Ellis and Newman characterized the distribution of the normalized sum of spins (*empirical magnetization*) for a wide class of mean-field Hamiltonian of Curie-Weiss type [8, 9, 10]. They found conditions, in terms of thermodynamic properties, that lead in the infinite volume limit to a Gaussian behavior and those which lead to a higher order exponential probability distribution.

A natural further step was the investigation of large and moderate fluctuations for the magnetization. The large deviation principle is due to Ellis [7]. Moderate deviation properties have been treated by Eichelsbacher and Löwe in [6]. A moderate deviation principle is technically a large deviation principle and consists in a refinement of a (standard or non-standard) central limit theorem, in the sense that it characterizes the exponential decay of deviations from the average on a smaller scale. In [6], it was shown that the physical phase transition in Curie-Weiss type models is reflected by a radical change in the asymptotic behavior of moderate deviations. Indeed, whereas the rate function is quadratic at non-critical temperatures, it becomes non-quadratic at criticality.

All the results mentioned so far have been derived at equilibrium; on the contrary, we are interested in describing the time evolution of fluctuations, obtaining non-equilibrium properties. Fluctuations for the standard Curie-Weiss model were studied on the level of a path-space large deviation principle by Comets [2] and

---

\*Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft (The Netherlands). *E-mail address:* f.collet-1@tudelft.nl

†Faculty of Mathematics, Ruhr-University of Bochum, Postfach 102148, 44721 Bochum (Germany). *E-mail address:* richard.kraaij@rub.de

Kraaij [15] and on the level of a path-space central limit theorem by Collet and Dai Pra in [1]. The purpose of the present paper is to study dynamical moderate deviations to complete the analysis of fluctuations of the empirical magnetization. We apply the generator convergence approach to large deviations by Feng-Kurtz [13] to characterize the most likely behaviour for the trajectories of fluctuations around the stationary solution(s) in the various regimes. The moderate asymptotics depend crucially on the phase we are considering. The criticality of the inverse temperature  $\beta = 1$  shows up at this level via a sudden change in the speed and rate function of the moderate deviation principle for the magnetization. In particular, our findings indicate that fluctuations are Gaussian-like in the sub- and super-critical regimes, while they are not at the critical point.

Besides, we analyze the deviation behaviour when the temperature is size-dependent and is increasing to the critical point. In this case, the rate function inherits features of both the uniqueness and multiple phases: it is the combination of the critical and non-critical rate functions. To conclude, it is worth to mention that our statements are in agreement with the results found in [6].

The outline of the paper is as follows: in Section 2 we formally introduce the Curie-Weiss model and we state our main results. All the proofs, if not immediate, are postponed to Section 3. Appendix A is devoted to the derivation of a large deviation principle via solution of Hamilton-Jacobi equation and it is included to make the paper as much self-contained as possible.

## 2 Model and main results

### 2.1 Notation and definitions

Before we give our main results, we introduce some notation. We start with the definition of good rate-functions and what it means for random variables to satisfy a large deviation principle.

**Definition 2.1.** Let  $X_1, X_2, \dots$  be random variables on a Polish space  $F$ . Furthermore let  $I : F \rightarrow [0, \infty]$ .

1. We say that  $I$  is a *good rate-function* if for every  $c \geq 0$ , the set  $\{x \mid I(x) \leq c\}$  is compact.
2. We say that the sequence  $\{X_n\}_{n \geq 1}$  is *exponentially tight* if for every  $a \geq 0$  there is a compact set  $K_a \subseteq X$  such that  $\limsup_n \mathbb{P}[X \in K_a^c] \leq -a$ .
3. We say that the sequence  $\{X_n\}_{n \geq 1}$  satisfies the *large deviation principle* with rate  $r(n)$  and good rate-function  $I$ , denoted by

$$\mathbb{P}[X_n \approx a] \sim e^{-r(n)I(a)},$$

if we have for every closed set  $A \subseteq X$

$$\limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P}[X_n \in A] \leq - \inf_{x \in A} I(x),$$

and for every open set  $U \subseteq X$

$$\liminf_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P}[X_n \in U] \geq - \inf_{x \in U} I(x).$$

Throughout the whole paper  $\mathcal{AC}$  will denote the set of absolutely continuous curves in  $\mathbb{R}$ .

**Definition 2.2.** A curve  $\gamma : [0, T] \rightarrow \mathbb{R}$  is absolutely continuous if there exists a function  $g \in L^1[0, T]$  such that for  $t \in [0, T]$  we have  $\gamma(t) = \gamma(0) + \int_0^t g(s)ds$ . We write  $g = \dot{\gamma}$ .

A curve  $\gamma : \mathbb{R}^+ \rightarrow \mathbb{R}$  is absolutely continuous if the restriction to  $[0, T]$  is absolutely continuous for every  $T \geq 0$ .

## 2.2 Glauber dynamics for the Curie-Weiss model

Let  $\sigma = (\sigma_i)_{i=1}^n \in \{-1, +1\}^n$  be a configuration of  $n$  spins and denote by

$$m_n(\sigma) = n^{-1} \sum_{i=1}^n \sigma_i$$

the empirical magnetization. The stochastic process  $\{\sigma(t)\}_{t \geq 0}$  is described as follows. For  $\sigma \in \{-1, +1\}^n$ , let us define  $\sigma^j$  the configuration obtained from  $\sigma$  by flipping the  $j$ -th spin. The spins will be assumed to evolve with Glauber spin-flip dynamics: at any time  $t$ , the system may experience a transition  $\sigma \rightarrow \sigma^j$  at rate  $\exp[-\beta \sigma_j(t) m_n(t)]$ , where  $\beta > 0$  represents the inverse temperature and where by abuse of notation  $m_n(t) := m_n(\sigma(t))$ . More formally, we can say that  $\{\sigma(t)\}_{t \geq 0}$  is a Markov process on  $\{-1, +1\}^n$ , with infinitesimal generator

$$\mathcal{G}_n f(\sigma) = \sum_{i=1}^n e^{-\beta \sigma_i m_n(\sigma)} [f(\sigma^i) - f(\sigma)]. \quad (1)$$

Let

$$E_n := m_n(\{-1, +1\}^n) = \left\{ -1, -1 + \frac{2}{n}, \dots, 1 - \frac{2}{n}, 1 \right\} \subseteq [-1, 1]$$

be the set of possible values taken by the magnetization (we will keep using this notation for the state space of the magnetization). The Glauber dynamics (1) on the configurations induce Markovian dynamics for the process  $\{m_n(t)\}_{t \geq 0}$  on  $E_n$ , that in turn evolves with generator

$$\mathcal{A}_n f(x) = n \frac{1-x}{2} e^{\beta x} \left[ f\left(x + \frac{2}{n}\right) - f(x) \right] + n \frac{1+x}{2} e^{-\beta x} \left[ f\left(x - \frac{2}{n}\right) - f(x) \right].$$

This generator can be derived in two ways from (1). First of all, the microscopic jumps induce a change of size  $\frac{2}{n}$  on the empirical magnetization. The jump rate of  $x$  to  $x + \frac{2}{n}$  corresponds to any  $-1$  spin switching to  $+1$  with rate  $e^{\beta x}$ . The total number of  $-1$  spins can be computed from the empirical magnetization  $x$  and equals  $n \frac{1-x}{2}$ . A similar computation yields the jump rate of  $x$  to  $x - \frac{2}{n}$ . A second way to see that  $\mathcal{A}_n$  is the generator of the empirical magnetization is via the martingale problem and the property that  $\mathcal{A}_n f(m_n(\sigma)) = \mathcal{G}_n(f \circ m_n)(\sigma)$ .

Assume the initial condition  $m_n(0)$  obeys a large deviation principle, then it can be shown that  $\{m_n(t)\}_{t \geq 0}$  obeys a large deviation principle on the Skorohod space of càdlàg functions  $D_{\mathbb{R}}(\mathbb{R}^+)$ . We refer to [12] for definition and properties of Skorohod spaces and to [2, 5] for the proof of the large deviation principle. Moreover, see [15, Theorem 1] for a LDP obtained by using similar techniques as in this paper. This path space large deviation principle allows to derive the infinite volume dynamics for our model: if  $m_n(0)$  converges weakly to the constant  $m_0$ , then the empirical magnetization process  $\{m_n(t)\}_{t \geq 0}$  converges weakly, as  $n \rightarrow \infty$ , to the solution of

$$\dot{m}(t) = -2 m(t) \cosh(\beta m(t)) + 2 \sinh(\beta m(t)) \quad (2)$$

with initial condition  $m_0$ . It is well known that the dynamical system (2) exhibits a phase transition at the critical value  $\beta = 1$ . The solution  $m = 0$  is an equilibrium

of (2) for all values of the parameters. For  $\beta \leq 1$ , it is globally stable; whereas, for  $\beta > 1$ , it loses stability and two new stable fixed point  $\mathbf{m} = \pm \mathbf{m}_\beta$ ,  $\mathbf{m}_\beta > 0$ , bifurcate. We refer the reader to [7]. For later convenience, let us introduce the notation

$$\mathbf{G}_{1,\beta}(\mathbf{x}) = \cosh(\beta \mathbf{x}) - \mathbf{x} \sinh(\beta \mathbf{x}) \quad \text{and} \quad \mathbf{G}_{2,\beta}(\mathbf{x}) = \sinh(\beta \mathbf{x}) - \mathbf{x} \cosh(\beta \mathbf{x}).$$

Observe that the equilibria of (2) are solutions to  $\mathbf{G}_{2,\beta}(\mathbf{x}) = 0$ .

### 2.3 Main results

We want to discuss the moderate deviations behavior of the magnetization around its limiting stationary points in the various regimes. We have the following three results that can be obtained as particular cases of the more general Theorem 3.7 stated and proven in Section 3.2. The first of our statements is mainly of interest for sub-critical inverse temperatures  $\beta < 1$ , but is indeed valid for all  $\beta \geq 0$ . The results for the critical and super-critical regimes follow afterwards.

**Theorem 2.3** (Moderate deviations around 0). *Let  $\{\mathbf{b}_n\}_{n \geq 1}$  be a sequence of positive real numbers such that  $\mathbf{b}_n \rightarrow \infty$  and  $\mathbf{b}_n^2 n^{-1} \rightarrow 0$ . Suppose that  $\mathbf{b}_n \mathbf{m}_n(0)$  satisfies the large deviation principle with speed  $n \mathbf{b}_n^{-2}$  on  $\mathbb{R}$  with rate function  $I_0$ . Then the trajectories  $\{\mathbf{b}_n \mathbf{m}_n(t)\}_{t \geq 0}$  satisfy the large deviation principle on  $\mathbf{D}_{\mathbb{R}}(\mathbb{R}^+)$ :*

$$\mathbb{P} \left[ \{\mathbf{b}_n \mathbf{m}_n(t)\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \right] \sim e^{-n \mathbf{b}_n^{-2} I(\gamma)},$$

where  $I$  is the good rate function

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases} \quad (3)$$

with

$$\mathcal{L}(x, v) = \frac{1}{8} |v + 2x(1 - \beta)|^2.$$

**Theorem 2.4** (Moderate deviations: critical temperature  $\beta = 1$ ). *Let  $\{\mathbf{b}_n\}_{n \geq 1}$  be a sequence of positive real numbers such that  $\mathbf{b}_n \rightarrow \infty$  and  $\mathbf{b}_n^4 n^{-1} \rightarrow 0$ . Suppose that  $\mathbf{b}_n \mathbf{m}_n(0)$  satisfies the large deviation principle with speed  $n \mathbf{b}_n^{-4}$  on  $\mathbb{R}$  with rate function  $I_0$ . Then the trajectories  $\{\mathbf{b}_n \mathbf{m}_n(\mathbf{b}_n^2 t)\}_{t \geq 0}$  satisfy the large deviation principle on  $\mathbf{D}_{\mathbb{R}}(\mathbb{R}^+)$ :*

$$\mathbb{P} \left[ \{\mathbf{b}_n \mathbf{m}_n(\mathbf{b}_n^2 t)\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \right] \sim e^{-n \mathbf{b}_n^{-4} I(\gamma)},$$

where  $I$  is the good rate function

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases} \quad (4)$$

with

$$\mathcal{L}(x, v) = \frac{1}{8} \left| v + \frac{2}{3} x^3 \right|^2.$$

**Theorem 2.5** (Moderate deviations: super-critical temperatures  $\beta > 1$ ). *Let  $\mathbf{m} \in \{-\mathbf{m}_\beta, +\mathbf{m}_\beta\}$  be a non-zero solution of  $\mathbf{G}_{2,\beta}(\mathbf{x}) = 0$ . Moreover, let  $\{\mathbf{b}_n\}_{n \geq 1}$  be a sequence of positive real numbers such that  $\mathbf{b}_n \rightarrow \infty$  and  $\mathbf{b}_n^2 n^{-1} \rightarrow 0$ . Suppose that  $\mathbf{b}_n(\mathbf{m}_n(0) - \mathbf{m})$  satisfies the large deviation principle with speed  $n \mathbf{b}_n^{-2}$  on  $\mathbb{R}$  with rate function  $I_0$ . Then the trajectories  $\{\mathbf{b}_n(\mathbf{m}_n(t) - \mathbf{m})\}_{t \geq 0}$  satisfy the large deviation principle on  $\mathbf{D}_{\mathbb{R}}(\mathbb{R}^+)$ :*

$$\mathbb{P} \left[ \{\mathbf{b}_n(\mathbf{m}_n(t) - \mathbf{m})\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \right] \sim e^{-n \mathbf{b}_n^{-2} I(\gamma)},$$

where  $I$  is the good rate function

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases} \quad (5)$$

with

$$\mathcal{L}(x, v) = \frac{(v - 2xG'_{2,\beta}(m))^2}{8G_{1,\beta}(m)}.$$

The rate functions (3) and (5) have a similar structure. Indeed, whenever  $\beta < 1$ ,  $m = 0$  is the unique solution of  $G_{2,\beta}(x) = 0$ ; moreover, it yields  $G_{1,\beta}(0) = 1$  and  $G'_{2,\beta}(0) = \beta - 1$ .

By choosing the sequence  $b_n = n^\alpha$ , with  $\alpha > 0$ , we can rephrase Theorems 2.3, 2.4 and 2.5 in terms of more familiar “moderate” scalings involving powers of the volume. We therefore get estimates for the probability of a typical trajectory on a scale that is between a law of large numbers and a central limit theorem. We give a schematic summary of these special results in Table 1 below.

For any of the three cases above, we define the *Hamiltonian*  $H : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  by taking the Legendre transform of  $\mathcal{L}$ :  $H(x, p) = \sup_v pv - \mathcal{L}(x, v)$ . It is well known that the rate function  $S$  of the stationary measures of the Markov processes, also known as the quasi-potential, solves the equation  $H(x, S'(x)) = 0$ , cf. Theorem 5.4.3 in [14]. We use this property to show that our results are consistent with the moderate deviation principles obtained for the stationary measures in [6]. We give the Hamiltonian of the three cases above

- (a) *sub-critical temperatures*, Theorem 2.3  $H(x, p) = 2x(\beta - 1)p + 2p^2$ ,
- (b) *critical temperature*, Theorem 2.4  $H(x, p) = -\frac{2}{3}x^3p + 2p^2$ ,
- (c) *super-critical temperatures*, Theorem 2.5  $H(x, p) = 2xG'_2(m)p + 2G_1(m)p^2$ .

The stationary rate function  $S$  in each of these three cases obtained in [6, Theorem 1.18] is given by

- (a) *sub-critical temperatures*,  $S(x) = \frac{1}{2}(1 - \beta)x^2$ ,
- (b) *critical temperature*,  $S(x) = \frac{1}{12}x^4$ ,
- (c) *super-critical temperatures*,  $S(x) = \frac{1}{2}cx^2$ , where  $c := (\phi''(\beta m))^{-1} - \beta$ , and  $m$  is a solution of  $G_{2,\beta}(x) = 0$  and  $\phi(x) = \log(\cosh(x))$ .

For (a) and (b), it is clear that  $H(x, S'(x)) = 0$  for all  $x$ . For (c), since  $m = \tanh(\beta m)$ , by inverse function theorem we obtain  $\phi''(\beta m) = 1 - m^2$ . Therefore, we have

$$\begin{aligned} c &= (\phi''(\beta m))^{-1} - \beta = (1 - m \tanh(\beta m))^{-1} - \beta \\ &= - \left[ \frac{\beta G_{1,\beta}(m) - \cosh(\beta m)}{G_{1,\beta}(m)} \right] = - \frac{G'_{2,\beta}(m)}{G_{1,\beta}(m)}, \end{aligned}$$

which implies that also in this case  $H(x, S'(x)) = 0$  for all  $x$ .

The next theorem complements the results in [1, Proposition 2.2] for the subcritical regime and shows that also in the supercritical case the fluctuations around a ferromagnetic stationary point converge to a diffusion process. As previously, the statement is a direct consequence of a more general central limit theorem given in Theorem 3.8 below.

**Theorem 2.6** (Central limit theorem: super-critical temperatures  $\beta > 1$ ). *Let  $\mathbf{m} \in \{-\mathbf{m}_\beta, +\mathbf{m}_\beta\}$  be a non-zero solution of  $G_{2,\beta}(x) = 0$ . Suppose that  $\mathbf{n}^{1/2}(\mathbf{m}_\mathbf{n}(0) - \mathbf{m})$  converges in law to  $\nu$ . Then the process  $\mathbf{n}^{1/2}(\mathbf{m}_\mathbf{n}(t) - \mathbf{m})$  converges weakly in law on  $D_{\mathbb{R}}(\mathbb{R}^+)$  to the unique solution of:*

$$\begin{cases} dY(t) = 2Y(t)G'_{2,\beta}(\mathbf{m})dt + 2\sqrt{G_{1,\beta}(\mathbf{m})}dW(t) \\ Y(0) \sim \nu, \end{cases} \quad (6)$$

where  $W(t)$  is a standard Brownian motion on  $\mathbb{R}$ .

We want to conclude the analysis by considering moderate deviations and non-standard central limit theorem for volume-dependent temperatures decreasing to the critical point. In the sequel let  $\{\mathbf{m}_\mathbf{n}^\beta(t)\}_{t \geq 0}$  denote the process evolving at temperature  $\beta$ .

**Theorem 2.7** (Moderate deviations: critical temperature  $\beta = 1$ , temperature rescaling). *Let  $\kappa \geq 0$  and let  $\{\mathbf{b}_\mathbf{n}\}_{\mathbf{n} \geq 1}$  be a sequence of positive real numbers such that  $\mathbf{b}_\mathbf{n} \rightarrow \infty$  and  $\mathbf{b}_\mathbf{n}^4 \mathbf{n}^{-1} \rightarrow 0$ . Suppose that  $\mathbf{b}_\mathbf{n} \mathbf{m}_\mathbf{n}^{1+\kappa \mathbf{b}_\mathbf{n}^{-2}}(0)$  satisfies the large deviation principle with speed  $\mathbf{n} \mathbf{b}_\mathbf{n}^{-4}$  on  $\mathbb{R}$  with rate function  $I_0$ . Then the trajectories  $\left\{ \mathbf{b}_\mathbf{n} \mathbf{m}_\mathbf{n}^{1+\kappa \mathbf{b}_\mathbf{n}^{-2}}(\mathbf{b}_\mathbf{n}^2 t) \right\}_{t \geq 0}$  satisfy the large deviation principle on  $D_{\mathbb{R}}(\mathbb{R}^+)$ :*

$$\mathbb{P} \left[ \left\{ \mathbf{b}_\mathbf{n} \mathbf{m}_\mathbf{n}^{1+\kappa \mathbf{b}_\mathbf{n}^{-2}}(\mathbf{b}_\mathbf{n}^2 t) \right\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \right] \sim e^{-\mathbf{n} \mathbf{b}_\mathbf{n}^{-4} I(\gamma)},$$

where  $I$  is the good rate function

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases} \quad (7)$$

with

$$\mathcal{L}(x, v) = \frac{1}{8} \left| v - 2 \left( \kappa x - \frac{1}{3} x^3 \right) \right|^2.$$

Notice that in this borderline case the moderate deviations rate function (7) is a sort of mixture of the one at the criticality (4) and the Gaussian rate function (3).

**Theorem 2.8** (Critical fluctuations: critical temperature  $\beta = 1$ , temperature rescaling). *Let  $\kappa \geq 0$  and suppose that  $\mathbf{n}^{1/4} \mathbf{m}_\mathbf{n}^{1+\kappa \mathbf{n}^{-1/2}}(0)$  converges in law to  $\nu$ . Then the process  $\mathbf{n}^{1/4} \mathbf{m}_\mathbf{n}^{1+\kappa \mathbf{n}^{-1/2}}(\mathbf{n}^{1/2} t)$  converges weakly in law on  $D_{\mathbb{R}}(\mathbb{R}^+)$  to the unique solution of:*

$$\begin{cases} dY(t) = 2 \left[ \kappa Y(t) - \frac{1}{3} Y(t)^3 \right] dt + 2 dW(t) \\ Y(0) \sim \nu, \end{cases} \quad (8)$$

where  $W(t)$  is a standard Brownian motion on  $\mathbb{R}$ .

The proof of Theorem 2.8 is a simple adaptation of the proofs of Theorems 3.8 and 2.7 and therefore is omitted.

The results in the present section, together with the large deviation principle in [15, Theorem 1] and the study of fluctuations at  $\beta = 1$  in [1, Theorem 2.10], give a complete picture of the behaviour of fluctuations for the Curie-Weiss model. Indeed all the possible scales are covered. We summarize our findings in Table 1. For completeness, we give also the Hamiltonian of the large deviation principle for the dynamics of  $\mathbf{m}_\mathbf{n}(t)$  around its limiting trajectory (2):

$$\begin{aligned} H(x, p) &= \frac{1-x}{2} e^{\beta x} [e^{2p} - 1] + \frac{1+x}{2} e^{-\beta x} [e^{-2p} - 1] \\ &= [\cosh(2p) - 1] G_{1,\beta}(x) + \sinh(2p) G_{2,\beta}(x). \end{aligned} \quad (9)$$

The displayed conclusions are drawn under the assumption that in each case either the initial condition satisfies a large deviation principle at the correct speed or the initial measure converges weakly.

**Table 1:** Fluctuations for the empirical magnetization of the Curie-Weiss spin-flip dynamics

| SCALING EXPONENT              | TEMPERATURE  | RESCALED PROCESS               | LIMITING THEOREM   |
|-------------------------------|--|--------------------------------|--|
| $\alpha = 0$                  | all $\beta$  | $m_n(t)$                       | LDP at speed $n$ with Hamiltonian as in (9) (see [2, 15])                                      |
| <b>NON-CRITICAL CASES</b>     |  |                                |  |
| $\alpha \in (0, \frac{1}{2})$ | all $\beta$  | $n^\alpha m_n(t)$              | LDP at speed $n^{1-2\alpha}$ with rate function (3)  |
|                               | $\beta > 1$  | $n^\alpha(m_n(t) \pm m_\beta)$ | LDP at speed $n^{1-2\alpha}$ with rate function (5)  |
| $\alpha = \frac{1}{2}$        | all $\beta$  | $n^{1/2}m_n(t)$                | CLT weak convergence to the unique solution of $dY(t) = 2(\beta - 1)Y(t)dt + 2dW(t)$ (see [1]) |
|                               | $\beta > 1$  | $n^{1/2}(m_n(t) \pm m_\beta)$  | CLT weak convergence to the unique solution of (6)   |
| <b>CRITICAL CASES</b>         |  |                                |  |
| $\alpha \in (0, \frac{1}{4})$ | $\beta = 1$  | $n^\alpha m_n(n^{2\alpha}t)$   | LDP at speed $n^{1-4\alpha}$ with rate function (4)  |
|                               | $\beta = 1 + \kappa n^{-2\alpha}$<br>(with $\kappa \geq 0$ ) | $n^\alpha m_n(n^{2\alpha}t)$   | LDP at speed $n^{1-4\alpha}$ with rate function (7)  |
| $\alpha = \frac{1}{4}$        | $\beta = 1$  | $n^{1/4}m_n(n^{1/2}t)$         | weak convergence to the unique solution of $dY(t) = -\frac{2}{3}Y(t)^3 dt + 2dW(t)$ (see [1])  |
|                               | $\beta = 1 + \kappa n^{-1/2}$<br>(with $\kappa \geq 0$ )     | $n^{1/4}m_n(n^{1/2}t)$         | weak convergence to the unique solution of (8)   |



## 3 Proofs

### 3.1 Strategy of the proof

We will analyze large/moderate deviations behaviour following the Feng-Kurtz approach to large deviations [13]. This method is based on three observations:

1. If the processes are exponentially tight, it suffices to establish the large deviation principle for finite dimensional distributions.
2. The large deviation principle for finite dimensional distributions can be established by proving that the semigroup of log Laplace-transforms of the conditional probabilities converges to a limiting semigroup.
3. One can often rewrite the limiting semigroup as a variational semigroup, which allows to rewrite the rate-function on the Skorohod space in Lagrangian form.

The strategy to prove a large deviation principle with speed  $r(\mathfrak{n})$  for a sequence of Markov processes  $\{X_{\mathfrak{n}}\}_{\mathfrak{n} \geq 1}$ , having generators  $\{A_{\mathfrak{n}}\}_{\mathfrak{n} \geq 1}$ , formally works as follows:

1. *Identification of a limiting Hamiltonian*  $H$ . The semigroups of log-Laplace transforms of the conditional probabilities

$$V_{\mathfrak{n}}(t)f(x) = \frac{1}{r(\mathfrak{n})} \log \mathbb{E} \left[ e^{r(\mathfrak{n})f(X_{\mathfrak{n}}(t))} \mid X_{\mathfrak{n}}(0) = x \right]$$

formally have generators  $H_{\mathfrak{n}}f = r(\mathfrak{n})^{-1} e^{-r(\mathfrak{n})f} A_{\mathfrak{n}} e^{r(\mathfrak{n})f}$ . Then one verifies that the sequence  $\{H_{\mathfrak{n}}\}_{\mathfrak{n} \geq 1}$  converges to a limiting operator  $H$ ; i.e. one shows that, for any  $f \in \mathcal{D}(H)$ , there exists  $f_{\mathfrak{n}} \in \mathcal{D}(H_{\mathfrak{n}})$  such that  $f_{\mathfrak{n}} \rightarrow f$  and  $H_{\mathfrak{n}}f_{\mathfrak{n}} \rightarrow Hf$ , as  $\mathfrak{n} \rightarrow \infty$ .

2. *Exponential tightness*. Provided one can verify the exponential compact containment condition, the convergence of the sequence  $\{H_{\mathfrak{n}}\}_{\mathfrak{n} \geq 1}$  gives exponential tightness.
3. *Verification of a comparison principle*. The theory of viscosity solutions gives applicable conditions for proving that the limiting Hamiltonian generates a semigroup. If for all  $\lambda > 0$  and bounded continuous functions  $h$ , the Hamilton-Jacobi equation  $f - \lambda Hf = h$  admits a unique solution, one can extend the generator  $H$  so that the extension satisfies the conditions of Crandall-Liggett theorem and thus generates a semigroup  $V(t)$ . Additionally, it follows that the semigroups  $V_{\mathfrak{n}}(t)$  converge to  $V(t)$ , giving the large deviation principle. Uniqueness of the solution of the Hamilton-Jacobi equation can be established via comparison principle for sub- and super-solutions.
4. *Variational representation of the limiting semigroup*. By Legendre transforming the limiting Hamiltonian  $H$ , one can define a ‘‘Lagrangian’’ which can be used to define a variational semigroup and a variational resolvent. It can be shown that the variational resolvent provides a solution of the Hamilton-Jacobi equation and therefore, by uniqueness of the solution, identifies the resolvent of  $H$ . As a consequence, an approximation procedure yields that the variational semigroup and the limiting semigroup  $V(t)$  agree. A standard argument is then sufficient to give a Lagrangian form of the path-space rate function.

We refer to Appendix A for an overview of the derivation of a large deviation principle via solution of Hamilton-Jacobi equation.

We proceed with the verification of the model-specific conditions needed to apply

the results from the appendix. The treatment is carried out in a more abstract way than strictly necessary to single out important arguments that might get lost if all the objects are explicit. The appendix and the following definitions are written to prove a path-space large deviation principle on  $D_E(\mathbb{R}^+)$ , the Skorohod space of paths taking values in the closed set  $E \subseteq \mathbb{R}^d$  which is contained in the  $\mathbb{R}^d$ -closure of its  $\mathbb{R}^d$ -interior. Two types of functions are of importance to this purpose: good penalization and good containment functions.

**Definition 3.1.** We say that  $\{\Psi_\alpha\}_{\alpha>0}$ , with  $\Psi_\alpha : E^2 \rightarrow \mathbb{R}$ , are *good penalization functions* if there are extensions of  $\Psi_\alpha$  to an open neighbourhood of  $E^2$  in  $\mathbb{R}^d \times \mathbb{R}^d$  (also denoted by  $\Psi_\alpha$ ) so that

(Ψa) For all  $\alpha > 0$ , we have  $\Psi_\alpha \geq 0$  and  $\Psi_\alpha(x, y) = 0$  if and only if  $x = y$ . Additionally,  $\alpha \mapsto \Psi_\alpha$  is increasing and

$$\lim_{\alpha \rightarrow \infty} \Psi_\alpha(x, y) = \begin{cases} 0 & \text{if } x = y \\ \infty & \text{if } x \neq y. \end{cases}$$

(Ψb)  $\Psi_\alpha$  is twice continuously differentiable in both coordinates for all  $\alpha > 0$ ,

(Ψc)  $(\nabla \Psi_\alpha(\cdot, y))(x) = -(\nabla \Psi_\alpha(x, \cdot))(y)$  for all  $\alpha > 0$ .

**Definition 3.2.** We say that  $\Upsilon : E \rightarrow \mathbb{R}$  is a *good containment function* (for H) if there is an extension of  $\Upsilon$  to an open neighbourhood of  $E$  in  $\mathbb{R}^d$  (also denoted by  $\Upsilon$ ) so that

(Υa)  $\Upsilon \geq 0$  and there exists a point  $x_0 \in E$  such that  $\Upsilon(x_0) = 0$ ,

(Υb)  $\Upsilon$  is twice continuously differentiable,

(Υc) for every  $c \geq 0$ , the set  $\{x \in E \mid \Upsilon(x) \leq c\}$  is compact,

(Υd) we have  $\sup_{z \in E} H(z, \nabla \Upsilon(z)) < \infty$ .

In the rest of the paper, we will denote by  $C_c^2(E)$  the set of functions that are constant outside some compact set in  $E$  and twice continuously differentiable on a neighbourhood of  $E$  in  $\mathbb{R}^d$ .

Let us denote by  $E_n$ , closed subset of  $E \subseteq \mathbb{R}^d$ , the set where the finite  $n$  process takes values. Our large or moderate deviation principles will all follow from the application of Theorem A.17 after having checked the following conditions:

(a) For all  $f \in C_c^2(E)$  and compact sets  $K \subseteq E$ , we have

$$\lim_{n \rightarrow \infty} \sup_{x \in K \cap E_n} |H_n f(x) - H f(x)| = 0.$$

(b) There exists a good containment function  $\Upsilon$  for H.

(c) For all  $\lambda > 0$  and  $h \in C_b(E)$ , the comparison principle holds for  $f - \lambda H f = h$ .

Rigorous definitions of Hamilton-Jacobi equation, viscosity solutions and comparison principle will be given in Appendix A.

The limiting Hamiltonians we will encounter are all of quadratic type and the space  $E$  will always equal  $\mathbb{R}$ . The following two known results establish (b) and (c) for Hamiltonians of this type, and are given for completeness. We postpone the verification of condition (a) for our various cases to Subsections 3.2 and 3.3.

**Definition 3.3.** Let  $E \subseteq \mathbb{R}^d$  be a closed set. We say that a vector field  $\mathbf{F} : E \rightarrow \mathbb{R}^d$  is one-sided Lipschitz if there exists a constant  $M \geq 0$  such that, for all  $x, y \in E$ , it holds

$$\langle x - y, \mathbf{F}(x) - \mathbf{F}(y) \rangle \leq M|x - y|^2.$$

**Lemma 3.4.** Let  $\mathbf{F} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be one-sided Lipschitz,  $A$  a positive-definite matrix and  $c \geq 0$ . Assume  $\mathbf{F}(0) = 0$ . Consider the Hamiltonian  $H$  with domain  $C_c^2(\mathbb{R}^d)$  and of the form  $Hf(x) = H(x, \nabla f(x))$ , where

$$H(x, p) = \langle p, \mathbf{F}(x) \rangle + c \langle Ax, p \rangle.$$

Then  $\Upsilon(x) = \log(1 + \frac{1}{2}|x|^2)$  is a good containment function for  $H$ .

*Proof.* The proof is based on a simplification of Example 4.23 in [13]. Since the  $i$ -th component of the gradient of  $\Upsilon$  is given by

$$(\nabla \Upsilon(x))_i = \frac{x_i}{1 + \frac{1}{2}|x|^2},$$

we obtain

$$H(x, \nabla \Upsilon(x)) = \frac{\langle x, \mathbf{F}(x) \rangle}{1 + \frac{1}{2}|x|^2} + \frac{c \langle Ax, x \rangle}{(1 + \frac{1}{2}|x|^2)^2}.$$

Notice that, by the one-sided Lipschitz property of  $\mathbf{F}$ , there exists a constant  $M$  such that  $\langle x, \mathbf{F}(x) \rangle = \langle x - 0, \mathbf{F}(x) - \mathbf{F}(0) \rangle \leq M|x|^2$ . Moreover, by Cauchy-Schwarz inequality, we have  $\langle Ax, x \rangle \leq |Ax||x| \leq \|A\||x|^2$ , where  $\|A\| := \sup_{|x|=1} |Ax|$ . Therefore, we get the estimate

$$H(x, \nabla \Upsilon(x)) \leq \frac{M|x|^2}{1 + \frac{1}{2}|x|^2} + \frac{c\|A\||x|^2}{(1 + \frac{1}{2}|x|^2)^2} \leq 4(M + c\|A\|),$$

which gives  $\sup_x H(x, \nabla \Upsilon(x)) < \infty$ , implying that  $\Upsilon$  is a good containment function.  $\blacksquare$

**Proposition 3.5.** Let  $\mathbf{F} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be one-sided Lipschitz,  $A$  a positive-definite matrix and  $c \in \mathbb{R}$ . Assume  $\mathbf{F}(0) = 0$ . Consider the Hamiltonian  $H$  with domain  $C_c^2(\mathbb{R}^d)$  and of the form  $Hf(x) = H(x, \nabla f(x))$ , where

$$H(x, p) = \langle p, \mathbf{F}(x) \rangle + c \langle Ap, p \rangle.$$

Then, for every  $\lambda > 0$  and  $h \in C_b(\mathbb{R}^d)$ , the comparison principle is satisfied for  $f - \lambda Hf = h$ .

*Proof.* We apply Proposition A.11. We have to check (22). We use the good containment function introduced in Lemma 3.4 and the collection of good penalization functions  $\{\Psi_\alpha\}_{\alpha>0}$ , with  $\Psi_\alpha(x, y) = \frac{\alpha}{2}|x - y|^2$ . We fix  $\varepsilon > 0$  and write  $x_\alpha, y_\alpha$  instead of  $x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon}$  to lighten the notation. As the term  $\langle Ap, p \rangle$  does not depend on  $x$  or  $y$ , we have

$$\begin{aligned} & H(x_\alpha, \alpha(x_\alpha - y_\alpha)) - H(y_\alpha, \alpha(x_\alpha - y_\alpha)) \\ &= \langle \alpha(x_\alpha - y_\alpha), \mathbf{F}(x_\alpha) \rangle - \langle \alpha(x_\alpha - y_\alpha), \mathbf{F}(y_\alpha) \rangle \leq M\Psi_\alpha(x_\alpha, y_\alpha). \end{aligned}$$

By Lemma A.10, we obtain  $\lim_{\alpha \rightarrow \infty} \Psi_\alpha(x_\alpha, y_\alpha) = 0$  and the conclusion follows.  $\blacksquare$

**Remark 3.6.** Lemma 3.4 and Proposition 3.5 can be suitably adapted (by centering  $\Upsilon$ ) to deal with the case when  $\mathbf{F}(x_s) = 0$  at some  $x_s \neq 0$ .

### 3.2 Proof of Theorems 2.3, 2.4, 2.5 and 2.6

We introduce a general version of dynamics à la Curie-Weiss where the evolution of the magnetization is driven by a sufficiently smooth arbitrary potential. We characterize moderate deviations and central limit theorem for this generalization getting then as corollaries the statements of Theorems 2.3–2.6.

Let  $U$  be a continuously differentiable potential and consider the dynamics such that the empirical magnetization  $\{m_n(t)\}_{t \geq 0}$  is a Markov process on  $E_n$ , with generator

$$\begin{aligned} \mathcal{A}_n f(x) = & \frac{n(1-x)}{2} e^{U'(x)} \left[ f\left(x + \frac{2}{n}\right) - f(x) \right] \\ & + \frac{n(1+x)}{2} e^{-U'(x)} \left[ f\left(x - \frac{2}{n}\right) - f(x) \right]. \end{aligned} \quad (10)$$

The infinite volume dynamics corresponding to the Markov process (10) may be derived from the large deviation principle in [15, Theorem 1]. In particular, the stationary points for the limiting dynamics are the solutions of  $G_2(x) = 0$ , where

$$G_2(x) := \sinh(U'(x)) - x \cosh(U'(x)).$$

For later convenience we define also  $G_1(x) := \cosh(U'(x)) - x \sinh(U'(x))$ . In this setting, we have a moderate deviation principle and a weak convergence result.

**Theorem 3.7** (Moderate deviations, arbitrary potential and stationary point). *Let  $m$  be a solution of  $G_2(x) = 0$ . Let  $k \in \mathbb{N} \cup \{0\}$  and suppose that  $U'$  is  $2k+1$  times continuously differentiable. Additionally, suppose that*

- (a)  $G_2^{(l)}(m) = 0$  for  $l \leq 2k$ ,
- (b) if  $k > 0$ , then  $G_2^{(2k+1)}(m) \leq 0$ .

Let  $\{b_n\}_{n \geq 1}$  be a sequence of positive real numbers such that

$$b_n \rightarrow \infty, \quad \frac{b_n^{2(k+1)}}{n} \rightarrow 0.$$

Suppose that  $b_n(m_n(0) - m)$  satisfies the large deviation principle with speed  $nb_n^{-2(k+1)}$  on  $\mathbb{R}$  with rate function  $I_0$ . Then the trajectories  $\{b_n(m_n(b_n^{2k}t) - m)\}_{t \geq 0}$  satisfy the large deviation principle on  $D_{\mathbb{R}}(\mathbb{R}^+)$ :

$$\mathbb{P} \left[ \{b_n(m_n(b_n^{2k}t) - m)\}_{t \geq 0} \approx \{\gamma(t)\}_{t \geq 0} \right] \sim e^{-nb_n^{-2(k+1)}I(\gamma)},$$

where  $I$  is the good rate function

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases}$$

and

$$\mathcal{L}(x, v) = \frac{\left(v - \frac{2x^{2k+1}}{(2k+1)!} G_2^{(2k+1)}(m)\right)^2}{8G_1(m)}.$$

*Proof.* The generator  $\mathcal{A}_n$  of the process  $b_n(m_n(b_n^{2k}t) - m)$  can be deduced from (10) and is given by

$$\begin{aligned} \mathcal{A}_n f(x) = & b_n^{2k} n \frac{1 - m - xb_n^{-1}}{2} e^{U'(m+xb_n^{-1})} \left[ f\left(x + 2b_n n^{-1}\right) - f(x) \right] \\ & + b_n^{2k} n \frac{1 + m + xb_n^{-1}}{2} e^{-U'(m+xb_n^{-1})} \left[ f\left(x - 2b_n n^{-1}\right) - f(x) \right]. \end{aligned}$$

Therefore the Hamiltonian

$$H_n f = b_n^{2(k+1)} n^{-1} e^{-n b_n^{-2(k+1)} f} A_n e^{n b_n^{-2(k+1)} f},$$

results in

$$\begin{aligned} H_n f(x) &= b_n^{4k+2} \frac{1 - m - x b_n^{-1}}{2} e^{u'(m + x b_n^{-1})} \left[ e^{n b_n^{-2(k+1)} (f(x + 2b_n n^{-1}) - f(x))} - 1 \right] \\ &\quad + b_n^{4k+2} \frac{1 + m + x b_n^{-1}}{2} e^{-u'(m + x b_n^{-1})} \left[ e^{n b_n^{-2(k+1)} (f(x - 2b_n n^{-1}) - f(x))} - 1 \right]. \end{aligned}$$

We now prove the convergence of the sequence  $H_n f$  for  $f \in C_c^2(\mathbb{R})$ . To compensate for the  $b_n^{4k+2}$  up front, we Taylor expand the exponential containing  $f$  up terms of  $O(b_n^{-4k-2})$ :

$$\begin{aligned} \exp \left\{ n b_n^{-2(k+1)} (f(x \pm 2b_n n^{-1}) - f(x)) \right\} - 1 \\ = \pm 2b_n^{-2k-1} f'(x) + 2b_n^{-4k-2} (f'(x))^2 + o(b_n^{-4k-2}). \end{aligned}$$

Observe that the above expansion holds since  $b_n^{-2k} n^{-1} = o(b_n^{-4k-2})$  by hypothesis. Thus, combining the terms with  $f'$  and the terms with  $(f')^2$ , we find that

$$H_n f(x) = 2b_n^{2k+1} G_2(m + x b_n^{-1}) f'(x) + 2G_1(m + x b_n^{-1}) (f'(x))^2 + o(1).$$

Next, we Taylor expand  $G_1, G_2$  around  $m$ . For  $G_1$  it is clear that only the zero'th order term remains. For  $G_2$ , we use that the first  $2k$  terms disappear. This gives

$$H_n f(x) = \frac{2x^{2k+1}}{(2k+1)!} G_2^{(2k+1)}(m) f'(x) + 2G_1(m) (f'(x))^2 + o(1),$$

where the  $o(1)$  is uniform on compact sets. Thus, for  $f \in C_c^2(\mathbb{R})$ ,  $H_n f$  converges uniformly to  $Hf(x) = H(x, f'(x))$  where

$$H(x, p) = \frac{2x^{2k+1}}{(2k+1)!} G_2^{(2k+1)}(m) p + 2G_1(m) p^2.$$

The large deviation result follows by Theorem A.17, Lemma 3.4 and Proposition 3.5. Note that condition (b) in the statement guarantees that the vector field is one-sided Lipschitz. The Lagrangian is found by taking a Legendre transform of  $H$ .  $\blacksquare$

**Theorem 3.8** (Central limit theorem, general potential, arbitrary stable stationary point). *Let  $m$  be a solution of  $G_2(x) = 0$ . Let  $k \in \mathbb{N} \cup \{0\}$  and suppose that  $U'$  is  $2k+1$  times continuously differentiable. Additionally, suppose that*

(a)  $G_2^{(l)}(m) = 0$  for  $l \leq 2k$ ,

(b) if  $k > 0$ , then  $G_2^{(2k+1)}(m) \leq 0$ .

Suppose that  $n^{\frac{1}{2k+2}}(m_n(0) - m)$  converges in law to  $\nu$ . Then the process

$$n^{\frac{1}{2k+2}} \left( m_n \left( n^{\frac{k}{k+1}} t \right) - m \right)$$

converges weakly in law on  $D_{\mathbb{R}}(\mathbb{R}^+)$  to the unique solution of:

$$\begin{cases} dY(t) = \frac{2}{(2k+1)!} Y(t)^{2k+1} G_2^{(2k+1)}(m) dt + 2\sqrt{G_1(m)} dW(t) \\ Y(0) \sim \nu, \end{cases} \quad (11)$$

where  $W(t)$  is a standard Brownian motion on  $\mathbb{R}$ .

The limiting diffusion (11) admits a unique invariant measure with density proportional to  $\exp\{-\mathbf{c}y^{2k+2}/(2k+2)!\}$ , with  $\mathbf{c} = 4|\mathbf{G}_2^{(2k+1)}(\mathbf{m})|$ . Observe that it is precisely the limiting distribution prescribed by the analysis of equilibrium fluctuations performed in [9].

The proof of the theorem given below is in the spirit of the proofs of the moderate deviation principles and based on a combination of proving the compact containment condition and the convergence of the generators.

We first prove the compact containment condition. Let

$$X_n(t) := n^{\frac{1}{2k+2}} \left( m_n \left( n^{\frac{k}{k+1}} t \right) - m \right)$$

be the space-time rescaled fluctuation process. We introduce the family  $\{\tau_n^C\}_{n \geq 1}$  of stopping times, defined by

$$\tau_n^C := \inf_{t \geq 0} \{|X_n(t)| \geq C\}.$$

We start by studying the asymptotic behavior of the sequence  $\{\tau_n^C\}_{n \geq 1}$ .

**Lemma 3.9.** *For any  $T \geq 0$  and  $\varepsilon > 0$ , there exist  $n_\varepsilon \geq 1$  and  $C_\varepsilon > 0$  such that*

$$\sup_{n \geq n_\varepsilon} \mathbb{P}(\tau_n^{C_\varepsilon} \leq T) \leq \varepsilon.$$

*Proof.* Let  $C$  be a strictly positive constant. First observe that

$$\mathbb{P}(\tau_n^C \leq T) \leq \mathbb{P}\left(\sup_{0 \leq t \leq T \wedge \tau_n^C} |X_n(t)| \geq C\right). \quad (12)$$

We will obtain bounds for (12) and show that it can be made arbitrarily small whenever  $n$  is large enough. The idea is to get the estimate by considering a martingale decomposition for  $\tilde{f}(X_n)$ , where  $\tilde{f} \in C^3(\mathbb{R})$  has bounded partial derivatives and is such that  $\tilde{f}(x) = \tilde{f}(-x)$  and  $\lim_{x \rightarrow \infty} \tilde{f}(x) = \infty$ .

For any  $n \geq 1$  and  $j \in \{-1, +1\}$ , let  $N_n(j, dt)$  be the Poisson process counting the number of flips of spins with value  $j$  up to time  $n^{\frac{k}{k+1}} t$ . The intensity of  $N_n(j, dt)$  is  $R_n(j, t)dt$  with

$$R_n(j, t) = \frac{n^{\frac{2k+1}{2k+2}}}{2} \left[ 1 + j \left( m + xn^{-\frac{1}{2k+2}} \right) \right] e^{-jU' \left( m + xn^{-\frac{1}{2k+2}} \right)}.$$

Moreover, we define

$$\tilde{N}_n(j, dt) := N_n(j, dt) - R_n(j, t) dt. \quad (13)$$

Let  $\tilde{f}(x) = \log \sqrt{1+x^2}$  and consider the semi-martingale decomposition

$$\tilde{f}(X_n(t)) = \tilde{f}(X_n(0)) + \int_0^t A_n \tilde{f}(X_n(s)) ds + M_n^{(1)}(t),$$

where  $M_n^{(1)}$  is the local martingale given by

$$M_n^{(1)}(t) := \int_0^t \sum_{j=\pm 1} (\bar{\nabla}_j \tilde{f}(X_n(s)))^2 \tilde{N}_n(j, ds)$$

and

$$\bar{\nabla}_j \tilde{f}(X_n(t)) := \log \sqrt{1 + \left( X_n(t) - 2jn^{-\frac{2k+1}{2k+2}} \right)^2} - \log \sqrt{1 + X_n^2(t)}.$$

We have

$$\begin{aligned} \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \tau_n^C} |X_n(t)| \geq C \right) &\leq \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \tau_n^C} \tilde{f}(X_n(t)) \geq \tilde{f}(C) \right) \\ &\leq \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \tau_n^C} \tilde{f}(X_n(0)) \geq \frac{\tilde{f}(C)}{3} \right) + \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \tau_n^C} A_n \tilde{f}(X_n(t)) \geq \frac{\tilde{f}(C)}{3T} \right) \\ &\quad + \mathbb{P} \left( \sup_{0 \leq t \leq T \wedge \tau_n^C} M_n^{(1)}(t) \geq \frac{\tilde{f}(C)}{3} \right). \end{aligned}$$

We estimate the three terms in the right hand-side of previous inequality. All the constants appearing in the bounds below are *independent of n*.

- Convergence in law of the initial condition implies  $\mathbb{P}(\tilde{f}(X_n(0)) \geq c_0(\varepsilon)) \leq \frac{\varepsilon}{3}$  for a sufficiently large  $c_0(\varepsilon)$  and for all  $n$ .
- Since we are stopping the process  $X_n(t)$  whenever it leaves the compact set  $[-C, C]$ , we find that  $f(X_n(t))$  is bounded on the set of interest. Therefore, we can apply (19) proven below to obtain

$$A_n \tilde{f}(x) = \frac{2G_2^{(2k+1)}(m)}{(2k+1)!} \frac{x^{2k+2}}{1+x^2} + 4G_1(m) \frac{1-x^2}{(1+x^2)^2} + o(1), \quad (14)$$

that, for  $t \leq \tau_n^C$ , implies

$$A_n \tilde{f}(X_n(t)) \leq c_1, \quad (15)$$

for a sufficiently large  $c_1 > 0$ , independent of  $C$ . Note that the first term in the right hand-side of (14) is bounded if  $k = 0$  and negative if  $k > 0$ .

- Since

$$\begin{aligned} |R_n(j, t)| &\leq n^{\frac{2k+1}{2k+2}} e^{-jU'(m+xn^{-\frac{1}{2k+2}})} \\ &\leq n^{\frac{2k+1}{2k+2}} \left( e^{U'(m)} + O\left(n^{-\frac{1}{2k+2}}\right) \right) \end{aligned} \quad (16)$$

and

$$\left( \bar{\nabla}_j \tilde{f}(X_n(t)) \right)^2 = \left( \int_{X_n(t)}^{X_n(t) - 2jn^{-\frac{2k+1}{2k+2}}} \frac{y}{1+y^2} dy \right)^2 \leq c_2 n^{-\frac{4k+2}{2k+2}}, \quad (17)$$

we obtain

$$\mathbb{E} \left[ \left( M_n^{(1)}(T \wedge \tau_n^C) \right)^2 \right] = \mathbb{E} \left[ \int_0^{T \wedge \tau_n^C} \sum_{j=\pm 1} \left( \bar{\nabla}_j \tilde{f}(X_n(s)) \right)^2 R_n(j, s) ds \right] \leq c_3 T \quad (18)$$

for sufficiently large  $n$  and suitable  $c_2, c_3 > 0$ , independent of  $C$ . By Doob's maximal inequality, we can conclude

$$\mathbb{P} \left( \sup_{0 < t < T \wedge \tau_n^C} M_n^{(1)}(t) \geq \frac{\tilde{f}(C)}{3} \right) \leq \frac{9Tc_3}{\tilde{f}(C)^2}.$$

Therefore, for any  $\varepsilon > 0$  and for sufficiently large  $n$ , by choosing the constant  $C_\varepsilon \geq \max \{c_0(\varepsilon), \sqrt{3c_1}, \tilde{f}^{-1}(\frac{27Tc_3}{\varepsilon})\}$ , we obtain  $\sup_{n \geq n_\varepsilon} \mathbb{P}(\tau_n^{C_\varepsilon} \leq T) \leq \varepsilon$  as wanted. ■

*Proof of Theorem 3.8.* As introduced above, let  $X_n(t) := n^{\frac{1}{2k+2}} \left( m_n \left( n^{\frac{k}{k+1}} t \right) - m \right)$  be the space-time rescaled process. The infinitesimal generator of the process  $X_n(t)$  can be easily deduced from (10). It yields

$$\begin{aligned} A_n f(x) &= n^{\frac{2k+1}{k+1}} \frac{1 - m - xn^{-\frac{1}{2k+2}}}{2} e^{U'(m+xn^{-\frac{1}{2k+2}})} \left[ f \left( x + 2n^{-\frac{2k+1}{2k+2}} \right) - f(x) \right] \\ &\quad + n^{\frac{2k+1}{k+1}} \frac{1 + m + xn^{-\frac{1}{2k+2}}}{2} e^{-U'(m+xn^{-\frac{1}{2k+2}})} \left[ f \left( x - 2n^{-\frac{2k+1}{2k+2}} \right) - f(x) \right]. \end{aligned}$$

We want to characterize the limit of the sequence  $A_n f$  for  $f \in C_c^3(\mathbb{R})$ , the set of three times continuously differentiable functions that are constant outside of a compact set. We first Taylor expand  $f$  up to the second order

$$f \left( x \pm 2n^{-\frac{2k+1}{2k+2}} \right) - f(x) = \pm 2n^{-\frac{2k+1}{2k+2}} f'(x) + 2n^{-\frac{2k+1}{k+1}} f''(x) + o \left( n^{-\frac{2k+1}{k+1}} \right)$$

and then, by combining the terms with  $f'$  and the terms with  $f''$ , we obtain

$$A_n f(x) = 2n^{\frac{2k+1}{2k+2}} G_2 \left( m + xn^{-\frac{1}{2k+2}} \right) f'(x) + 2 G_1 \left( m + xn^{-\frac{1}{2k+2}} \right) f''(x) + o(1).$$

Now we Taylor expand  $G_1$  and  $G_2$  around  $m$ . For  $G_2$  we use that the first  $2k$  terms in the expansion vanish. As far as  $G_1$ , only the zero-th order term matters. Therefore it yields

$$A_n f(x) = \frac{2}{(2k+1)!} G_2^{(2k+1)}(m) x^{2k+1} f'(x) + 2 G_1(m) f''(x) + o(1). \quad (19)$$

In other words, we conclude that for  $f \in C_c^3(\mathbb{R})$  we have  $\lim_n |A_n f - Af| = 0$ .

To prove the weak convergence result, we verify the conditions for Corollary 4.8.16 in [12]. The martingale problem for the operator  $(A, C_c^3(\mathbb{R}))$  has a unique solution by Theorem 8.2.6 in [12]. Additionally, the set  $C_c^3(\mathbb{R})$  is an algebra that separates points. Finally, by Lemma 3.9 the sequence  $\{X_n\}_{n \geq 1}$  satisfies the compact containment condition. Thus the result follows by an application of Corollary 4.8.16 in [12]. ■

The results in Section 2 are recovered by setting  $U(x) = \frac{\beta x^2}{2}$ , with  $\beta > 0$ . Note that simply by choosing  $U(x) = \frac{\beta x^2}{2} + Bx$ , with  $\beta, B > 0$ , we get the corresponding results for the Curie-Weiss with magnetic field.

### 3.3 Proof of Theorem 2.7

The infinitesimal generator  $A_n$  of the process  $b_n m_n^{1+\kappa b_n^{-2}}(b_n^2 t)$  can be easily deduced from (10) by using  $U(x) = \frac{(1+\kappa b_n^{-2})x^2}{2}$ . The Hamiltonian

$$H_n f = b_n^4 n^{-1} e^{-nb_n^{-4} f} A_n e^{nb_n^{-4} f},$$

is given by

$$\begin{aligned} H_n f(x) &= b_n^6 \frac{1 - \kappa b_n^{-1}}{2} e^{(1+\kappa b_n^{-2})x b_n^{-1}} \left[ e^{nb_n^{-4}(f(x+2b_n n^{-1})-f(x))} - 1 \right] \\ &\quad + b_n^6 \frac{1 + \kappa b_n^{-1}}{2} e^{-(1+\kappa b_n^{-2})x b_n^{-1}} \left[ e^{nb_n^{-4}(f(x-2b_n n^{-1})-f(x))} - 1 \right]. \end{aligned}$$

We start by studying the limiting behaviour of the sequence  $H_n f$  for  $f \in C_c^2(\mathbb{R})$ . Let  $A_n$  be the generator of the process that has been speeded up by a factor  $b_n^2$ ,



i.e.  $A_n f = b_n^2 \mathcal{A}_n f$ . To compensate for the  $b_n^6$  up front, we Taylor expand the exponential containing  $f$  up terms of  $O(b_n^{-6})$ :

$$\exp \{ n b_n^{-4} (f(x \pm 2b_n n^{-1}) - f(x)) \} - 1 = \pm 2b_n^{-3} f'(x) + 2b_n^{-6} (f'(x))^2 + o(b_n^{-6}).$$

Thus, combining the terms with  $f'$  and the terms with  $(f')^2$ , we find that

$$\begin{aligned} H_n f(x) &= 2 [b_n^3 \sinh(xb_n^{-1} + \kappa b_n^{-3}) - xb_n^2 \cosh(xb_n^{-1} + \kappa b_n^{-3})] f'(x) \\ &\quad + 2(f'(x))^2 + o(1). \end{aligned}$$

By Taylor expanding the hyperbolic functions

$$\begin{aligned} \sinh(xb_n^{-1} + \kappa b_n^{-3}) &= xb_n^{-1} + \kappa b_n^{-3} + \frac{1}{6}(xb_n^{-1} + \kappa b_n^{-3})^3 + o(b_n^{-3}) \\ \cosh(xb_n^{-1} + \kappa b_n^{-3}) &= 1 + \frac{1}{2}(xb_n^{-1} + \kappa b_n^{-3})^2 + o(b_n^{-2}), \end{aligned}$$

we get

$$H_n f(x) = 2 \left[ \kappa x - \frac{1}{3} x^3 \right] f'(x) + 2(f'(x))^2 + o(1),$$

where the  $o(1)$  is uniform on compact sets. Thus, for  $f \in C_c^2(\mathbb{R})$ ,  $H_n f$  converges uniformly to  $Hf(x) = H(x, f'(x))$  where

$$H(x, p) = 2 \left[ \kappa x - \frac{1}{3} x^3 \right] p + 2p^2$$

The large deviation result follows by Theorem A.17, Lemma 3.4 and Proposition 3.5. The Lagrangian is found by taking a Legendre transform of  $H$ .

## A Appendix: Large deviation principle via the Hamilton-Jacobi equation

In the Appendix, we will explain the basic steps to prove the path-space large deviation principle via uniqueness of solutions to the Hamilton-Jacobi equation. These steps follow the proofs in [3, 13, 4] and have also been used in the proof of the large deviation principle for the dynamics of variants of the Curie-Weiss model via well-posedness of the Hamilton-Jacobi equation in [15].

First, we prove an abstract result on how to obtain uniqueness of viscosity solution of the Hamilton-Jacobi equation via the comparison principle. Then, we will state a result on how uniqueness, together with an exponential compact containment condition, yields the large deviation principle. The verification of the conditions for this result have already been carried out in Section 3.1.

We make the remark that the requirements on our space  $E$  in Section A.3 and are more stringent than the ones in Sections A.1 and A.2. The definitions of good penalization functions and of a good containment function are unchanged for the next two sections.

### A.1 Viscosity solutions for the Hamilton-Jacobi equation

Fix some  $d \geq 1$ . In this section, and in Section A.2, let  $E$  be a subset of  $\mathbb{R}^d$  that is contained in the  $\mathbb{R}^d$ -closure of its  $\mathbb{R}^d$ -interior. Additionally, assume that  $E$  is a Polish space when equipped with its subspace-topology.

**Remark A.1.** The assumption that  $E$  is contained in the  $\mathbb{R}^d$ -closure of its  $\mathbb{R}^d$ -interior is needed to ensure that the gradient  $\nabla f(x)$  of a function  $f \in C_b^1(\mathbb{R}^d)$  for  $x \in E$  is determined by its values in  $E$ .

**Remark A.2.** We say that  $A \subseteq \mathbb{R}^d$  is a  $G_\delta$  set if it is the countable intersection of open sets in  $\mathbb{R}^d$ . By Theorems 4.3.23 and 4.3.24 in [11], we find that  $E$  is Polish if and only if it is a  $G_\delta$  set in  $\mathbb{R}^d$ .

Let  $H : E \times \mathbb{R}^d \rightarrow \mathbb{R}$  be a continuous map. For  $\lambda > 0$  and  $h \in C_b(E)$ , we will solve the *Hamilton-Jacobi* equation

$$f(x) - \lambda H(x, \nabla f(x)) = h(x) \quad x \in E, \quad (20)$$

in the *viscosity* sense.

**Definition A.3.** We say that  $u$  is a (*viscosity*) *subsolution* of equation (20) if  $u$  is bounded, upper semi-continuous and if, for every  $f \in \mathcal{D}(H)$  such that  $\sup_x u(x) - f(x) < \infty$  and every sequence  $x_n \in E$  such that

$$\lim_{n \rightarrow \infty} u(x_n) - f(x_n) = \sup_x u(x) - f(x),$$

we have

$$\lim_{n \rightarrow \infty} u(x_n) - \lambda Hf(x_n) - h(x_n) \leq 0.$$

We say that  $v$  is a (*viscosity*) *supersolution* of equation (20) if  $v$  is bounded, lower semi-continuous and if, for every  $f \in \mathcal{D}(H)$  such that  $\inf_x v(x) - f(x) > -\infty$  and every sequence  $x_n \in E$  such that

$$\lim_{n \rightarrow \infty} v(x_n) - f(x_n) = \inf_x v(x) - f(x),$$

we have

$$\lim_{n \rightarrow \infty} v(x_n) - \lambda Hf(x_n) - h(x_n) \geq 0.$$

At various points, we will refer to [13]. The notion of viscosity solution used here corresponds to the notion of *strong viscosity solution* in [13]. For operators of the form  $Hf(x) = H(x, \nabla f(x))$  these two notions are equivalent. See also Lemma 9.9 in [13].

**Definition A.4.** We say that equation (20) satisfies the *comparison principle* if for a subsolution  $u$  and supersolution  $v$  we have  $u \leq v$ .

Note that if the comparison principle is satisfied, then a viscosity solution is unique. To prove the comparison principle, we extend our scope by considering viscosity sub- and supersolutions to the Hamilton-Jacobi equation with two different operators that extend the original Hamiltonian in a suitable way.

Let  $M(E, \overline{\mathbb{R}})$  be the set of measurable functions from  $E$  to  $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ .

**Definition A.5.** We say that  $H_\dagger \subseteq M(E, \overline{\mathbb{R}}) \times M(E, \overline{\mathbb{R}})$  is a *viscosity sub-extension* of  $H$  if  $H \subseteq H_\dagger$  and if for every  $\lambda > 0$  and  $h \in C_b(E)$  a viscosity subsolution to  $f - \lambda Hf = h$  is also a viscosity subsolution to  $f - \lambda H_\dagger f = h$ . Similarly, we define a *viscosity super-extension*.

**Definition A.6.** Consider two operators  $H_\dagger, H_\ddagger \subseteq M(E, \overline{\mathbb{R}}) \times M(E, \overline{\mathbb{R}})$  and pick  $h \in C_b(E)$  and  $\lambda > 0$ . We say that the equations

$$f - \lambda H_\dagger f = h, \quad f - \lambda H_\ddagger f = h$$

satisfy the comparison principle if any subsolution  $u$  to the first and any supersolution  $v$  to the second equation satisfy  $u \leq v$ .

We have the following straightforward result.

**Lemma A.7.** *Suppose that  $H_{\dagger}$  and  $H_{\ddagger}$  are a sub- and superextension of  $H$  respectively. Fix  $\lambda > 0$  and  $h \in C_b(E)$ . If the comparison principle is satisfied for the equations*

$$f - \lambda H_{\dagger} f = h, \quad f - \lambda H_{\ddagger} f = h,$$

*then the comparison principle is satisfied for*

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

## A.2 Abstract proof of the comparison principle

We introduce two convenient viscosity extensions of a particular Hamiltonian  $H$  in terms of good penalization functions  $\{\Psi_{\alpha}\}_{\alpha \geq 0}$  and containment function  $\Upsilon$ .

$$\begin{aligned} \mathcal{D}(H_{\dagger}) &:= C_b^1(E) \cup \{x \mapsto (1 - \varepsilon)\Psi_{\alpha}(x, y) + \varepsilon\Upsilon(x) + c \mid \alpha, \varepsilon > 0, c \in \mathbb{R}\}, \\ \mathcal{D}(H_{\ddagger}) &:= C_b^1(E) \cup \{y \mapsto -(1 + \varepsilon)\Psi_{\alpha}(x, y) - \varepsilon\Upsilon(y) + c \mid \alpha, \varepsilon > 0, c \in \mathbb{R}\}. \end{aligned}$$

For  $f \in \mathcal{D}(H_{\dagger})$ , set  $H_{\dagger}f(x) = H(x, \nabla f(x))$  and for  $f \in \mathcal{D}(H_{\ddagger})$ , set  $H_{\ddagger}f(x) = H(x, \nabla f(x))$ .

**Lemma A.8.** *The operator  $(H_{\dagger}, \mathcal{D}(H_{\dagger}))$  is a viscosity sub-extension of  $H$  and  $(H_{\ddagger}, \mathcal{D}(H_{\ddagger}))$  is a viscosity super-extension of  $H$ .*

In the proof we need Lemma 7.7 from [13]. We recall it here for the sake of readability. Let  $M_{\infty}(E, \overline{\mathbb{R}})$  denote the set of measurable functions  $f : E \rightarrow \mathbb{R} \cup \{\infty\}$  that are bounded from below.

**Lemma A.9** (Lemma 7.7 in [13]). *Let  $H$  and  $H_{\dagger} \subseteq M_{\infty}(E, \overline{\mathbb{R}}) \times M(E, \overline{\mathbb{R}})$  be two operators. Suppose that for all  $(f, g) \in H_{\dagger}$  there exist  $\{(f_n, g_n)\} \subseteq H_{\dagger}$  that satisfy the following conditions:*

- (a) *For all  $n$ , the function  $f_n$  is lower semi-continuous.*
- (b) *For all  $n$ , we have  $f_n \leq f_{n+1}$  and  $f_n \rightarrow f$  point-wise.*
- (c) *Suppose  $x_n \in E$  is a sequence such that  $\sup_n f_n(x_n) < \infty$  and  $\inf_n g_n(x_n) > -\infty$ , then  $\{x_n\}_{n \geq 1}$  is relatively compact and if a subsequence  $x_{n(k)}$  converges to  $x \in E$ , then*

$$\limsup_{k \rightarrow \infty} g_{n(k)}(x_{n(k)}) \leq g(x).$$

*Then  $H_{\dagger}$  is a viscosity sub-extension of  $H$ .*

*An analogous result holds for super-extensions  $H_{\ddagger}$  by taking  $f_n$  a decreasing sequence of upper semi-continuous functions and by replacing requirement (c) with*

- (c') *Suppose  $x_n \in E$  is a sequence such that  $\inf_n f_n(x_n) > -\infty$  and  $\sup_n g_n(x_n) < \infty$ , then  $\{x_n\}_{n \geq 1}$  is relatively compact and if a subsequence  $x_{n(k)}$  converges to  $x \in E$ , then*

$$\liminf_{k \rightarrow \infty} g_{n(k)}(x_{n(k)}) \geq g(x).$$

*Proof of Lemma A.8.* We only prove the sub-extension part.

Consider a collection of smooth functions  $\phi_n : \mathbb{R} \rightarrow \mathbb{R}$  defined as  $\phi_n(x) = x$  if  $x \leq n$  and  $\phi_n(x) = n + 1$  for  $x \geq n + 1$ . Note that  $\phi_{n+1} \geq \phi_n$  for all  $n$ .

Fix a function  $f \in \mathcal{D}(H_{\dagger})$  of the type  $f(x) = (1 - \varepsilon)\Psi_{\alpha}(x, y) + \varepsilon\Upsilon(x) + c$  and write  $g = H_{\dagger}f$ .

Because  $\Psi_{\alpha}$  and  $\Upsilon$  are good penalization and good containment functions,  $f$  has a twice continuously differentiable extensions to a neighbourhood of  $E$  in  $\mathbb{R}^d$ . We will denote this extension also by  $f$ . Set  $f_n = \phi_n \circ f$ . Since  $f$  is bounded from below and  $\Upsilon$  has compact level sets, we find that  $f_n$  is constant outside a compact set

in  $E$ . Furthermore,  $f$  and  $\phi_n$  are twice continuously differentiable, so that we find  $f_n \in C_c^2(E)$ . We obtain  $f_n \in \mathcal{D}(H)$  and write  $g_n = Hf_n$ .

We verify conditions (a)-(c) of Lemma A.9. (a) has been verified above. (b) is a consequence of the fact that  $n \mapsto \phi_n$  is increasing. For (c), let  $\{x_n\}_{n \geq 1}$  be a sequence such that  $\sup_n f_n(x_n) = M < \infty$ . It follows by the compactness of the level sets of  $\Upsilon$  and the positivity of  $\Psi_\alpha$  that the set

$$K := \{z \in E \mid f(z) \leq M\}$$

is compact. Thus the sequence  $\{x_n\}$  is relatively compact, and in particular, there exist converging subsequences  $x_{n(k)}$  with limits  $x \in K$ . For any such subsequence, we show that  $\limsup_k g_{n(k)}(x_{n(k)}) \leq g(x)$ .

As  $\Psi_\alpha$  and  $\Upsilon$ , and thus  $f$ , are twice continuously differentiable up to a neighbourhood  $U$  of  $E$  in  $\mathbb{R}^d$ , we find that the set

$$V := \{z \in U \mid f(z) < M + 1\}$$

is open in  $\mathbb{R}^d$  and contains  $K$ . For two arbitrary continuously differentiable functions  $h_1, h_2$  on  $U$ , if  $h_1(z) = h_2(z)$  for all  $z \in V$ , then  $\nabla h_1(z) = \nabla h_2(z)$  for all  $z \in V$ .

Now suppose  $x_{n(k)}$  is a subsequence in  $K$  converging to some point  $x \in K$ . As  $f$  is bounded on  $V$ , there exists a sufficiently large  $N$  such that for all  $n \geq N$  and  $y \in V$ , we have  $f_n(y) = f(y)$ . We conclude  $\nabla f_n(y) = \nabla f(y)$  for  $y \in K \subseteq V$  and hence

$$g_n(y) = H(y, \nabla f_n(y)) = H(y, \nabla f(y)) = g(y).$$

In particular, we find  $\limsup_k g_{n(k)}(x_{n(k)}) \leq g(x)$ . ■

We have the following variants of Lemma 9.2 in [13] and Proposition 3.7 in [3]. Note that the presence of the containment function  $\Upsilon$  makes sure that the suprema are attained. This motivates the name containment function:  $\Upsilon$  forces the maxima to be in some compact set.

**Lemma A.10.** *Let  $u$  be bounded and upper semi-continuous, let  $v$  be bounded and lower semi-continuous, let  $\Psi_\alpha : E^2 \rightarrow \mathbb{R}^+$  be good penalization functions and let  $\Upsilon$  be a good containment function.*

*Fix  $\varepsilon > 0$ . For every  $\alpha > 0$  there exist points  $x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon} \in E$ , such that*

$$\begin{aligned} & \frac{u(x_{\alpha, \varepsilon})}{1 - \varepsilon} - \frac{v(y_{\alpha, \varepsilon})}{1 + \varepsilon} - \Psi_\alpha(x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon}) - \frac{\varepsilon}{1 - \varepsilon} \Upsilon(x_{\alpha, \varepsilon}) - \frac{\varepsilon}{1 + \varepsilon} \Upsilon(y_{\alpha, \varepsilon}) \\ &= \sup_{x, y \in E} \left\{ \frac{u(x)}{1 - \varepsilon} - \frac{v(y)}{1 + \varepsilon} - \Psi_\alpha(x, y) - \frac{\varepsilon}{1 - \varepsilon} \Upsilon(x) - \frac{\varepsilon}{1 + \varepsilon} \Upsilon(y) \right\}. \end{aligned}$$

*Additionally, for every  $\varepsilon > 0$  we have that*

- (a) *The set  $\{x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon} \mid \alpha > 0\}$  is relatively compact in  $E$ .*
- (b) *All limit points of  $\{(x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon})\}_{\alpha > 0}$  are of the form  $(z, z)$  and for these limit points we have  $u(z) - v(z) = \sup_{x \in E} \{u(x) - v(x)\}$ .*
- (c) *Suppose  $\Psi_\alpha$  can be written as  $\Psi_\alpha = \alpha\Psi$ , where  $\Psi \geq 0$ . Then we have*

$$\lim_{\alpha \rightarrow \infty} \alpha\Psi(x_{\alpha, \varepsilon}, y_{\alpha, \varepsilon}) = 0.$$

*Proof.* The proof essentially follows the one of Proposition 3.7 in [3]. Fix  $\varepsilon > 0$ . As  $\Upsilon$  is a good containment function, its level sets are compact. This property, combined with the boundedness of  $u$  and  $v$  and the non-negativity of  $\Psi_\alpha$ , implies

that the supremum can be restricted to a compact set  $K_\varepsilon \subseteq E \times E$  that is independent of  $\alpha > 0$ . As  $u$  is upper semi-continuous, and  $v, \Psi_\alpha$  and  $\Upsilon$  are lower semi-continuous, the supremum is attained for some pair  $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \in K_\varepsilon$ . This proves (a).

We proceed with the proof of (b). Let  $(x_0, y_0)$  be a limit point of  $\{(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})\}_{\alpha>0}$  such that  $x_0 \neq y_0$ . Without loss of generality, assume that  $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \rightarrow (x_0, y_0)$ . By property  $(\Psi a)$ , the map  $\alpha \mapsto \Psi_\alpha$  is increasing. Thus, for all  $\alpha_0$  we have that

$$\liminf_{\alpha \rightarrow \infty} \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \geq \liminf_{\alpha \rightarrow \infty} \Psi_{\alpha_0}(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \geq \Psi_{\alpha_0}(x_0, y_0)$$

by the lower semi-continuity of  $\Psi_{\alpha_0}$ . Thus, we conclude that

$$\liminf_{\alpha \rightarrow \infty} \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) = \infty$$

as  $\lim_{\alpha \rightarrow \infty} \Psi_\alpha(x, y) = \infty$  for all  $x \neq y$ . This contradicts the boundedness of  $u$  and  $v$ . We now prove (c). Let us define the constants

$$\begin{aligned} M_\alpha &:= \frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y_{\alpha,\varepsilon}) \\ &= \sup_{x,y \in E} \left\{ \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \Psi_\alpha(x, y) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y) \right\}. \end{aligned}$$

Observe that the sequence  $M_\alpha$  is decreasing as  $\alpha \mapsto \Psi_\alpha$  is increasing point-wise. Moreover, the limit  $\lim_{\alpha \rightarrow \infty} M_\alpha$  exists, since functions  $u$  and  $v$  are bounded from below. For any  $\alpha > 0$ , we obtain

$$\begin{aligned} M_{\alpha/2} &\geq \frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_{\alpha/2}(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y_{\alpha,\varepsilon}) \\ &\geq M_\alpha + \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \Psi_{\alpha/2}(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \\ &\geq M_\alpha + \frac{\alpha}{2} \Psi(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \\ &\geq M_\alpha, \end{aligned}$$

that implies  $\frac{\alpha}{2} \Psi(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) \rightarrow 0$ , as  $M_{\alpha/2}$  and  $M_\alpha$  converge to the same limit.  $\blacksquare$

**Proposition A.11.** *Fix  $\lambda > 0$ ,  $h \in C_b(E)$  and consider  $u$  and  $v$  sub- and super-solution to  $f - \lambda Hf = h$ .*

*Let  $\{\Psi_\alpha\}_{\alpha>0}$  be a family of good penalization functions and  $\Upsilon$  be a good containment function. Moreover, for every  $\alpha, \varepsilon > 0$  let  $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$  be such that*

$$\begin{aligned} &\frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y_{\alpha,\varepsilon}) \\ &= \sup_{x,y \in E} \left\{ \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \Psi_\alpha(x, y) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y) \right\}. \end{aligned} \quad (21)$$

*Suppose that*

$$\liminf_{\varepsilon \rightarrow 0} \liminf_{\alpha \rightarrow \infty} H(x_{\alpha,\varepsilon}, \nabla \Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) - H(y_{\alpha,\varepsilon}, \nabla \Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) \leq 0, \quad (22)$$

*then  $u \leq v$ . In other words:  $f - \lambda Hf = h$  satisfies the comparison principle.*

*Proof.* By Lemma A.8 we get immediately that  $u$  is a sub-solution to  $f - \lambda H_\dagger f = h$  and  $v$  is a super-solution to  $f - \lambda H_\ddagger f = h$ . Thus, it suffices to verify the comparison principle for the equations involving the extensions  $H_\dagger$  and  $H_\ddagger$ .

Let  $x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon} \in E$  such that (21) is satisfied. Then, for all  $\alpha$  we obtain that

$$\begin{aligned}
& \sup_x u(x) - v(x) \\
&= \lim_{\varepsilon \rightarrow 0} \sup_x \frac{u(x)}{1-\varepsilon} - \frac{v(x)}{1+\varepsilon} \\
&\leq \liminf_{\varepsilon \rightarrow 0} \sup_{x,y} \frac{u(x)}{1-\varepsilon} - \frac{v(y)}{1+\varepsilon} - \Psi_\alpha(x,y) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y) \\
&= \liminf_{\varepsilon \rightarrow 0} \frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon} - \Psi_\alpha(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon}) - \frac{\varepsilon}{1-\varepsilon} \Upsilon(x_{\alpha,\varepsilon}) - \frac{\varepsilon}{1+\varepsilon} \Upsilon(y_{\alpha,\varepsilon}) \\
&\leq \liminf_{\varepsilon \rightarrow 0} \frac{u(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{v(y_{\alpha,\varepsilon})}{1+\varepsilon}, \tag{23}
\end{aligned}$$

as  $\Upsilon$  and  $\Psi_\alpha$  are non-negative functions. Since  $u$  is a sub-solution to  $f - \lambda H_\dagger f = h$  and  $v$  is a super-solution to  $f - \lambda H_\ddagger f = h$ , we find by our particular choice of  $x_{\alpha,\varepsilon}$  and  $y_{\alpha,\varepsilon}$  that

$$u(x_{\alpha,\varepsilon}) - \lambda H(x_{\alpha,\varepsilon}, (1-\varepsilon)\nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) + \varepsilon\nabla\Upsilon(x_{\alpha,\varepsilon})) \leq h(x_{\alpha,\varepsilon}), \tag{24}$$

$$v(y_{\alpha,\varepsilon}) - \lambda H(y_{\alpha,\varepsilon}, -(1+\varepsilon)\nabla\Psi_\alpha(x_{\alpha,\varepsilon}, \cdot)(y_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) \geq h(y_{\alpha,\varepsilon}). \tag{25}$$

For all  $z \in E$ , the map  $p \mapsto H(z, p)$  is convex. Thus, (24) implies that

$$\begin{aligned}
u(x_{\alpha,\varepsilon}) &\leq h(x_{\alpha,\varepsilon}) + (1-\varepsilon)\lambda H(x_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) \\
&\quad + \varepsilon\lambda H(x_{\alpha,\varepsilon}, \nabla\Upsilon(x_{\alpha,\varepsilon})). \tag{26}
\end{aligned}$$

For the second inequality, first note that because  $\Psi_\alpha$  are good penalization functions, we have  $-(\nabla\Psi_\alpha(x_{\alpha,\varepsilon}, \cdot))(y_{\alpha,\varepsilon}) = \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})$ . Next, we need a more sophisticated bound using the convexity of  $H$ :

$$\begin{aligned}
& H(y_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) \\
&\leq \frac{1}{1+\varepsilon} H(y_{\alpha,\varepsilon}, (1+\varepsilon)\nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}) - \varepsilon\nabla\Upsilon(y_{\alpha,\varepsilon})) + \frac{\varepsilon}{1+\varepsilon} H(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon})).
\end{aligned}$$

Thus, (25) gives us

$$v(y_{\alpha,\varepsilon}) \geq h(y_{\alpha,\varepsilon}) + \lambda(1+\varepsilon)H(y_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) - \varepsilon\lambda H(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon})). \tag{27}$$

By combining (23) with (26) and (27), we find

$$\begin{aligned}
& \sup_x u(x) - v(x) \\
&\leq \liminf_{\varepsilon \rightarrow 0} \liminf_{\alpha \rightarrow \infty} \left\{ \frac{h(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{h(y_{\alpha,\varepsilon})}{1+\varepsilon} \right. \tag{28}
\end{aligned}$$

$$+ \frac{\varepsilon}{1-\varepsilon} H(x_{\alpha,\varepsilon}, \nabla\Upsilon(x_{\alpha,\varepsilon})) + \frac{\varepsilon}{1+\varepsilon} H(y_{\alpha,\varepsilon}, \nabla\Upsilon(y_{\alpha,\varepsilon})) \tag{29}$$

$$\left. + \lambda [H(x_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon})) - H(y_{\alpha,\varepsilon}, \nabla\Psi_\alpha(\cdot, y_{\alpha,\varepsilon})(x_{\alpha,\varepsilon}))] \right\}. \tag{30}$$

The term (30) vanishes by assumption. Now observe that, for fixed  $\varepsilon$  and varying  $\alpha$ , the sequence  $(x_{\alpha,\varepsilon}, y_{\alpha,\varepsilon})$  takes its values in a compact set and, hence, admits converging subsequences. All these subsequences converge to points of the form  $(z, z)$ . Therefore, as  $\alpha \rightarrow \infty$ , we find

$$\liminf_{\varepsilon \rightarrow 0} \liminf_{\alpha \rightarrow \infty} \frac{h(x_{\alpha,\varepsilon})}{1-\varepsilon} - \frac{h(y_{\alpha,\varepsilon})}{1+\varepsilon} \leq \liminf_{\varepsilon \rightarrow 0} |h| \frac{2\varepsilon}{1-\varepsilon^2} = 0,$$

giving that also the term in (28) converges to zero. The term in (29) vanishes as well, due to the uniform bounds on  $H(z, \nabla\Upsilon(z))$  by property  $(\Upsilon d)$ .

We conclude that the comparison principle holds for  $f - \lambda Hf = h$ . ■

### A.3 Compact containment and the large deviation principle

To connect the Hamilton-Jacobi equation to the large deviation principle, we introduce some additional concepts. Fix some  $d \geq 1$ . In this section, we assume that  $E$  is a *closed* subset of  $\mathbb{R}^d$  that is contained in the  $\mathbb{R}^d$ -closure of its  $\mathbb{R}^d$ -interior. Additionally, we have closed subspaces  $E_n \subseteq E$  for all  $n$  and assume that  $E = \lim_{n \rightarrow \infty} E_n$ , i.e. for every  $x \in E$  there exist  $x_n \in E_n$  such that  $x_n \rightarrow x$ . We consider the following notion of operator convergence.

**Definition A.12.** Suppose that for each  $n$  we have an operator  $(B_n, \mathcal{D}(B_n))$ ,  $B_n : \mathcal{D}(B_n) \subseteq C_b(E_n) \rightarrow C_b(E_n)$ . The *extended limit*  $\text{ex} - \lim_n B_n$  is defined by the collection  $(f, g) \in C_b(E) \times C_b(E)$  such that there exist  $f_n \in \mathcal{D}(B_n)$  satisfying

$$\lim_{n \rightarrow \infty} \sup_{x \in K \cap E_n} |f_n(x) - f(x)| + |B_n f_n(x) - g(x)| = 0. \quad (31)$$

For an operator  $(B, \mathcal{D}(B))$ , we write  $B \subseteq \text{ex} - \lim_n B_n$  if the graph  $\{(f, Bf) \mid f \in \mathcal{D}(B)\}$  of  $B$  is a subset of  $\text{ex} - \lim_n B_n$ .

**Remark A.13.** The notion of extended limit can be generalized further, e.g. for limiting spaces like in the previous two sections. Such abstract generalizations are carried out in Definition 2.5 and Condition 7.11 of [13]. Our definition for a *closed* limiting space  $E$  is the simplest version of this abstract machinery.

**Assumption A.14.** Fix some  $d \geq 1$ . Let  $E$  be a closed subset of  $\mathbb{R}^d$  that is contained in the  $\mathbb{R}^d$ -closure of its  $\mathbb{R}^d$ -interior and let  $E_n$  be closed subsets of  $E$  such that  $E = \lim_{n \rightarrow \infty} E_n$ .

Assume that for each  $n \geq 1$ , we have  $A_n \subseteq C_b(E_n) \times C_b(E_n)$  and existence and uniqueness holds for the  $D_{E_n}(\mathbb{R}^+)$  martingale problem for  $(A_n, \mu)$  for each initial distribution  $\mu \in \mathcal{P}(E_n)$ . Letting  $\mathbb{P}_y^n \in \mathcal{P}(D_{E_n}(\mathbb{R}^+))$  be the solution to  $(A_n, \delta_y)$ , the mapping  $y \mapsto \mathbb{P}_y^n$  is measurable for the weak topology on  $\mathcal{P}(D_{E_n}(\mathbb{R}^+))$ . Let  $X_n$  be the solution to the martingale problem for  $A_n$  and set

$$H_n f = \frac{1}{r(n)} e^{-r(n)f} A_n e^{r(n)f} \quad e^{r(n)f} \in \mathcal{D}(A_n),$$

for some sequence of speeds  $\{r(n)\}_{n \geq 1}$ , with  $\lim_{n \rightarrow \infty} r(n) = \infty$ .

Suppose that we have an operator  $H : \mathcal{D}(H) \subseteq C_b(E) \rightarrow C_b(E)$  with  $\mathcal{D}(H) = C_c^2(E)$  of the form  $Hf(x) = H(x, \nabla f(x))$  which satisfies  $H \subseteq \text{ex} - \lim H_n$ .

**Proposition A.15.** *Suppose Assumption A.14 is satisfied and assume that  $\Upsilon$  is a good containment function. Suppose that the sequence  $\{X_n(0)\}_{n \geq 1}$  is exponentially tight with speed  $\{r(n)\}_{n \geq 1}$ .*

*Then the sequence  $\{X_n\}_{n \geq 1}$  satisfies the exponential compact containment condition with speed  $\{r(n)\}_{n \geq 1}$ : for every  $T > 0$  and  $a \geq 0$ , there exists a compact set  $K_{a,T} \subseteq E$  such that*

$$\limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P}[X_n(t) \notin K_{a,T} \text{ for some } t \leq T] \leq -a.$$

In the proof of this proposition, we apply Lemma 4.22 from [13]. We recall it here for the sake of readability.

**Lemma A.16** (Lemma 4.22 in [13]). *Let  $X_n$  be solutions of the martingale problem for  $A_n$  and suppose that  $\{X_n(0)\}_{n \geq 1}$  is exponentially tight with speed  $\{r(n)\}_{n \geq 1}$ . Let  $K$  be compact and let  $G \supseteq K$  be open. For each  $n$ , suppose we have  $(f_n, g_n) \in H_n$ . Define*

$$\beta(K, G) := \liminf_{n \rightarrow \infty} \left( \inf_{x \in G^c} f_n(x) - \sup_{x \in K} f_n(x) \right) \quad \text{and} \quad \gamma(G) := \limsup_{n \rightarrow \infty} \sup_{x \in G} g_n(x).$$

Then

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P} [X_n(t) \notin G \text{ for some } t \leq T] \\ \leq \max \left\{ -\beta(K, G) + T\gamma(G), \limsup_{n \rightarrow \infty} \mathbb{P} [X_n(0) \notin K] \right\}. \end{aligned} \quad (32)$$

Note that in the case the closure of  $G$  is compact, this result is suitable for proving the compact containment condition. This is what we will use below.

*Proof of Proposition A.15.* Fix  $\alpha \geq 0$  and  $T > 0$ . We will construct a compact set  $K'$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P} [X_n(t) \notin K' \text{ for some } t \leq T] \leq -\alpha.$$

As  $X_n(0)$  is exponentially tight with speed  $\{r(n)\}_{n \geq 1}$ , we can find a sufficiently large  $R \geq 0$  so that

$$\limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P} [X_n(0) \notin \overline{B(x_0, R)}] \leq -\alpha,$$

where  $x_0$  is a point such that  $\Upsilon(x_0) = 0$  and  $\overline{B(x_0, R)}$  is the closed ball with radius  $R$  and center  $x_0$ . Thus, by Lemma A.16 it suffices to find  $(f_n, g_n) \in H_n$ , a compact  $K$  and an open set  $G$  such that  $-\beta(K, G) + T\gamma(G) \leq -\alpha$ .

Set  $\gamma := \sup_z H(z, \nabla \Upsilon(z))$  and  $c_1 := \sup_{z \in \overline{B(x_0, R)}} \Upsilon(z)$ . Observe that  $\gamma < \infty$  by assumption  $(\Upsilon d)$  and  $c_1 < \infty$  by compactness. Now choose  $c_2$  such that

$$-[c_2 - c_1] + T\gamma = -\alpha \quad (33)$$

and take  $K = \{z \in E \mid \Upsilon(z) \leq c_1\}$  and  $G = \{z \in E \mid \Upsilon(z) < c_2\}$ .

Let  $\theta : [0, \infty) \rightarrow [0, \infty)$  be a compactly supported smooth function with the property that  $\theta(z) = z$  for  $z \leq c_2$ . For each  $n$ , define  $f_n := \theta \circ \Upsilon$  and  $g_n := H_n f_n$ . By Assumption A.14,  $g_n \rightarrow Hf$  in the sense of (31) and moreover, by construction  $\beta(K, G) = c_2 - c_1$  and  $\gamma(G) = \gamma$ . Thus by (33) and Lemma A.16 we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{r(n)} \log \mathbb{P} [X_n(t) \notin G \text{ for some } t \leq T] \leq -\alpha$$

and the compact containment condition holds with  $K_{\alpha, T} = \overline{G}$ .  $\blacksquare$

**Theorem A.17** (Large deviation principle). *Suppose Assumption A.14 is satisfied and assume that  $\Upsilon$  is a good containment function for  $H$ . Then we have the following result.*

*Suppose that for all  $\lambda > 0$  and  $h \in C_b(E)$  the comparison principle holds for  $f - \lambda Hf = h$ . And suppose that the sequence  $\{X_n(0)\}_{n \geq 1}$  satisfies the large deviation principle with speed  $\{r(n)\}_{n \geq 1}$  on  $E$  with good rate function  $I_0$ .*

*Then the large deviation principle with speed  $\{r(n)\}_{n \geq 1}$  holds for  $\{X_n\}_{n \geq 1}$  on  $D_E(\mathbb{R}^+)$  with good rate function  $I$ . Additionally, suppose that the map  $p \mapsto H(x, p)$  is convex and differentiable for every  $x$  and that the map  $(x, p) \mapsto \frac{d}{dp} H(x, p)$  is continuous. Then the rate function  $I$  is given by*

$$I(\gamma) = \begin{cases} I_0(\gamma(0)) + \int_0^\infty \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds & \text{if } \gamma \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases}$$

where  $\mathcal{L} : E \times \mathbb{R}^d \rightarrow \mathbb{R}$  is defined by  $\mathcal{L}(x, v) = \sup_p \langle p, v \rangle - H(x, p)$ .



*Proof.* The large deviation result follows from Theorem 7.18 in [13]. Referring to the notation therein, we are using  $H_{\dagger} = H_{\ddagger} = H$ .

The representation of the rate function in terms of the Lagrangian can be carried out by using Theorem 8.27 and Corollary 8.28 in [13]. See for example the Section 10.3 in [13] for this representation in the setting of Freidlin-Wentzell theory. Alternatively, an application of this result in a compact setting has been carried out also in [15]. The only extra condition compared to Theorem 6 in [15] due to the non-compactness is the verification of Condition 8.9.(4) in [13]:

For each compact  $K \subseteq E$ ,  $T > 0$  and  $0 \leq M < \infty$ , there exists a compact set  $\hat{K} = \hat{K}(K, T, M) \subseteq E$  such that if  $\gamma \in \mathcal{AC}$  with  $\gamma(0) \in K$  and

$$\int_0^T \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds \leq M, \quad (34)$$

then  $\gamma(t) \in \hat{K}$  for all  $t \leq T$ .

This condition can be verified as follows. Recall that the level sets of  $\Upsilon$  are compact. Thus, we control the growth of  $\Upsilon$ . Let  $\gamma \in \mathcal{AC}$  satisfy the conditions given above. Then

$$\begin{aligned} \Upsilon(\gamma(t)) &= \Upsilon(\gamma(0)) + \int_0^t \langle \nabla \Upsilon(\gamma(s)), \dot{\gamma}(s) \rangle ds \\ &\leq \Upsilon(\gamma(0)) + \int_0^t \mathcal{L}(\gamma(s), \dot{\gamma}(s)) + H(\gamma(s), \nabla \Upsilon(\gamma(s))) ds \\ &\leq \sup_{y \in K} \Upsilon(y) + M + \int_0^T \sup_z H(z, \nabla \Upsilon(z)) ds \\ &=: C < \infty. \end{aligned}$$

Thus, we can take  $\hat{K} = \{z \in E \mid \Upsilon(z) \leq C\}$ . ■

## Acknowledgement

The authors are supported by The Netherlands Organisation for Scientific Research (NWO): RK via grant 600.065.130.12N109 and FC via TOP-1 grant 613.001.552.

## References

- [1] F. Collet and P. Dai Pra. The role of disorder in the dynamics of critical fluctuations of mean field models. *Electron. J. Probab.*, 17(26):1–40, 2012.
- [2] F. Comets. Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probability Theory and Related Fields*, 80(3):407–432, 1989.
- [3] 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):1–67, 1992.
- [4] X. Deng, J. Feng, and Y. Liu. A singular 1-D Hamilton-Jacobi equation, with application to large deviation of diffusions. *Communications in Mathematical Sciences*, 9(1), 2011.
- [5] P. Dupuis, K. Ramanan, and W. Wu. Large deviation principle for finite-state mean field interacting particle systems. *preprint; ArXiv:1601.06219*, 2016.

- [6] P. Eichelsbacher and M. Löwe. Moderate deviations for a class of mean-field models. *Markov Process. Related Fields*, 10(2):345–366, 2004.
- [7] R. S. Ellis. *Entropy, large deviations, and statistical mechanics*. Springer-Verlag, New York, 1985.
- [8] R. S. Ellis and C. M. Newman. Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete*, 44(2):117–139, 1978.
- [9] R. S. Ellis and C. M. Newman. The statistics of Curie-Weiss models. *J. Stat. Phys.*, 19(2):149–161, 1978.
- [10] R. S. Ellis, C. M. Newman, and J. S. Rosen. Limit theorems for sums of dependent random variables occurring in statistical mechanics, ii. *Z. Wahrsch. Verw. Gebiete*, 51(2):153–169, 1980.
- [11] R. Engelking. *General topology*. Heldermann Verlag, Berlin, second edition, 1989.
- [12] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and Convergence*. Wiley, 1986.
- [13] J. Feng and T. G. Kurtz. *Large Deviations for Stochastic Processes*. American Mathematical Society, 2006.
- [14] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*. Springer-Verlag, second edition, 1998.
- [15] R. 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):321–345, 2016.