Acknowledgments

AD acknowledges support from the Lagrange Mathematics and Computing Research Center. AD and MN would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme *The mathematical and statistical foundation of future data-driven engineering* when work on this paper was undertaken. MN acknowledges funding from the grant SCAI (ANR-19-CHIA-0002).

References

Betancourt, M. A general metric for Riemannian manifold Hamiltonian Monte Carlo. In *International Conference on Geometric Science of Information*, pp. 327–334. Springer, 2013.

Betancourt, M. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*, 2017.

Brofos, J. and Lederman, R. R. Evaluating the implicit midpoint integrator for Riemannian Hamiltonian Monte Carlo. In *International Conference on Machine Learning*, pp. 1072–1081. PMLR, 2021a.

Brofos, J. A. and Lederman, R. R. On numerical considerations for Riemannian Manifold Hamiltonian Monte Carlo. *arXiv preprint arXiv:2111.09995*, 2021b.

Brubaker, M., Salzmann, M., and Urtasun, R. A family of MCMC methods on implicitly defined manifolds. In *Artificial Intelligence and Statistics*, pp. 161–172. PMLR, 2012.

Bubeck, S., Eldan, R., and Lehec, J. Finite-time analysis of projected Langevin Monte Carlo. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, NIPS'15, pp. 1243–1251, Cambridge, MA, USA, 2015. MIT Press. URL http://dl.acm.org/citation.cfm?id=2969239.2969378.

Byrne, S. and Girolami, M. Geodesic Monte Carlo on embedded manifolds. *Scandinavian Journal of Statistics*, 40(4):825–845, 2013.

Cobb, A. D., Baydin, A. G., Markham, A., and Roberts, S. J. Introducing an explicit symplectic integration scheme for Riemannian Manifold Hamiltonian Monte Carlo. *arXiv preprint arXiv:1910.06243*, 2019.

Cousins, B. and Vempala, S. Gaussian Cooling and O*(n^3) Algorithms for Volume and Gaussian Volume. *ArXiv e-prints*, September 2014.

Dalalyan, A. S. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *Journal of the Royal Statistical Society, Series B*, 79(3):651–676, 2017.

Douc, R., Moulines, E., Priouret, P., and Soulier, P. *Markov Chains*. Springer, 2018.

Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D. Hybrid Monte Carlo. *Physics letters B*, 195(2):216–222, 1987.

Durmus, A. and Moulines, E. Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *The Annals of Applied Probability*, 27(3):1551–1587, 06 2017. doi: 10.1214/16-AAP1238.

Dyer, M. and Frieze, A. Computing the volume of convex bodies: a case where randomness provably helps. *Probabilistic combinatorics and its applications*, 44:123–170, 1991.

Gelfand, A. E., Smith, A. F., and Lee, T.-M. Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *Journal of the American Statistical Association*, 87(418): 523–532, 1992.

Girolami, M. and Calderhead, B. Riemann Manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(2):123–214, 2011.

Hairer, E., Hochbruck, M., Iserles, A., and Lubich, C. Geometric numerical integration. *Oberwolfach Reports*, 3(1):805–882, 2006.

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., Haraldsdóttir, H. S., Wachowiak, J., Keating, S. M., Vlasov, V., et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. *Nature protocols*, 14(3):639–702, 2019.

Kannan, R. and Narayanan, H. Random walks on polytopes and an affine interior point method for linear programming. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 561–570, 2009.

Kook, Y., Lee, Y.-T., Shen, R., and Vempala, S. Sampling with Riemannian Hamiltonian Monte Carlo in a constrained space. *Advances in Neural Information Processing Systems*, 35:31684–31696, 2022a.

Kook, Y., Lee, Y. T., Shen, R., and Vempala, S. S. Condition-number-independent convergence rate of Riemannian Hamiltonian Monte Carlo with numerical integrators. *arXiv preprint arXiv:2210.07219*, 2022b.

Lan, S. and Shahbaba, B. Sampling constrained probability distributions using Spherical Augmentation. *ArXiv e-prints*, June 2015.

Lee, J. M. *Riemannian Manifolds: an introduction to curvature*, volume 176. Springer Science & Business Media, 2006.

Lee, Y. T. and Vempala, S. S. Eldan's stochastic localization and the KLS hyperplane conjecture: an improved lower bound for expansion. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 998–1007. IEEE, 2017a.

Lee, Y. T. and Vempala, S. S. Geodesic walks in polytopes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on theory of Computing*, pp. 927–940, 2017b.

Lee, Y. T. and Vempala, S. S. Convergence rate of Riemannian Hamiltonian Monte Carlo and faster polytope volume computation. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1115–1121, 2018.

Leimkuhler, B. J. and Skeel, R. D. Symplectic numerical integrators in constrained Hamiltonian systems. *Journal of Computational Physics*, 112(1):117–125, 1994.

Lelièvre, T., Rousset, M., and Stoltz, G. Hybrid Monte Carlo methods for sampling probability measures on submanifolds. *Numerische Mathematik*, 143(2):379–421, 2019.

Lelièvre, T., Stoltz, G., and Zhang, W. Multiple projection Markov Chain Monte Carlo algorithms on submanifolds. *IMA Journal of Numerical Analysis*, 2022.

Lewis, N. E., Nagarajan, H., and Palsson, B. O. Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4):291–305, 2012.

Liu, J. S. Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and computing*, 6:113–119, 1996.

Liu, J. S. and Liu, J. S. *Monte Carlo strategies in scientific computing*, volume 10. Springer, 2001.

Lovász, L. and Kannan, R. Faster mixing via average conductance. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pp. 282–287. ACM, 1999.

Lovász, L. and Simonovits, M. Random walks in a convex body and an improved volume algorithm. *Random structures & algorithms*, 4(4):359–412, 1993.

Lovász, L. and Vempala, S. Hit-and-run from a corner. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pp. 310–314, 2004.

Lovász, L. and Vempala, S. The geometry of logconcave functions and sampling algorithms. *Random Struct. Algorithms*, 30(3):307–358, May 2007. ISSN 1042-9832. doi: 10.1002/rsa.v30:3. URL http://dx.doi.org/10.1002/rsa.v30:3.

Mok, K.-P. Metrics and connections on the cotangent bundle. In *Kodai Mathematical Seminar Reports*, volume 28, pp. 226–238. Department of Mathematics, Tokyo Institute of Technology, 1977.

Morris, B. J. Improved bounds for sampling contingency tables. *Random Structures & Algorithms*, 21(2):135–146, 2002.

Neal, R. M. et al. MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, 2 (11):2, 2011.

Nesterov, Y. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2003.

Nesterov, Y. and Nemirovski, A. Multi-parameter surfaces of analytic centers and long-step surface-following interior point methods. *Mathematics of Operations Research*, 23(1):1–38, 1998.

Nesterov, Y. and Nemirovskii, A. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

Pakman, A. and Paninski, L. Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians. *Journal of Computational and Graphical Statistics*, 23(2):518–542, 2014.

Roberts, G. O. and Rosenthal, J. S. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical science*, 16(4):351–367, 2001.

Roberts, G. O. and Stramer, O. Langevin diffusions and Metropolis-Hastings algorithms. *Methodology and computing in applied probability*, 4:337–357, 2002.

Sasaki, S. On the differential geometry of tangent bundles of Riemannian manifolds. *Tohoku Mathematical Journal, Second Series*, 10(3):338–354, 1958.

Shahbaba, B., Lan, S., Johnson, W. O., and Neal, R. M. Split Hamiltonian Monte Carlo. *Statistics and Computing*, 24(3):339–349, 2014.

Strang, G. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.

Tao, M. Explicit symplectic approximation of nonseparable Hamiltonians: Algorithm and long time performance. *Physical Review E*, 94(4):043303, 2016.

Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., Haraldsdottir, H., Mo, M. L., Rolfsson, O., Stobbe, M. D., et al. A community-driven global reconstruction of human metabolism. *Nature biotechnology*, 31(5):419–425, 2013.

Zappa, E., Holmes-Cerfon, M., and Goodman, J. Monte Carlo on manifolds: sampling densities and integrating functions. *Communications on Pure and Applied Mathematics*, 71(12):2609–2647, 2018.

Organisation of the supplementary

This appendix is organized as follows. Appendix A summarizes general facts that will be useful for proofs. Appendix B and Appendix C provide additional details on Riemannian metrics and technicalities on self-concordance, respectively. The c-BHMC algorithm is presented in Appendix D with some results whose proofs are given in Appendix E. Appendix F presents general facts about numerical integration and specific facts on the integrators used in n-BHMC. Appendix G dispenses the proof of the result on implicit integrators of n-BHMC, stated in Section 4.1. Appendix H presents the modified version of n-BHMC, incorporating a step-size condition, for which we state the reversibility with respect to the target distribution in Section 4.2. Proof of this last result is given in Appendix I. A detailed comparison between n-BHMC and CRHMC (Kook et al., 2022a) is given in Appendix J. Finally, Appendix K provides more details on the experiments of Section 6.

Notation. For any Riemannian manifold $(\mathsf{M}, \mathfrak{g})$ and any $z = (x, p) \in \mathrm{T}^\star\mathsf{M}$, we denote by $\|\cdot\|_z$ the norm (7) defined on $\mathrm{T}^\star\mathsf{M}$ by $\|(x', p')\|_z = \|x'\|_{\mathfrak{g}(x)} + \|p'\|_{\mathfrak{g}(x)^{-1}}$ for any $(x', p') \in \mathrm{T}^\star\mathsf{M}$. For any open subset $\mathsf{U} \subset \mathbb{R}^d$ and any $k \in \mathbb{N}$, we denote by $\mathrm{C}^k(\mathsf{U}, \mathbb{R}^{d'})$ the set of functions $f : \mathsf{U} \to \mathbb{R}^{d'}$ such that f is k times continuously differentiable.

A Useful facts and lemmas

We recall here some basic knowledge on linear algebra and probability, and state useful inequalities for our proofs.

Linear algebra reminders. For any matrices $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{d\times d}$, we write $\mathrm{A} \preceq \mathrm{B}$ if for any $x \in \mathbb{R}^d$, $x^\top(\mathrm{B} - \mathrm{A})x \geq 0$. Any positive-definite matrix $\mathrm{A} \in \mathbb{R}^{d\times d}$ induces a scalar product $\langle\cdot,\cdot\rangle_\mathrm{A}$ on $\mathbb{R}^d$, defined by $\langle x, y\rangle_\mathrm{A} = \langle x, \mathrm{A}y\rangle$. This scalar product induces the norm $\|\cdot\|_\mathrm{A}$ on $\mathbb{R}^d$, defined by $\|x\|_\mathrm{A} = \sqrt{\langle x, x\rangle_\mathrm{A}} = \|\mathrm{A}^{1/2}x\|_2$. For any positive-definite matrices $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{d\times d}$ and for any $\alpha \geq 0$, $\mathrm{A} \preceq \alpha\mathrm{B}$ is equivalent to $\|\cdot\|_\mathrm{A} \leq \sqrt{\alpha}\|\cdot\|_\mathrm{B}$. The canonical norm of $\mathbb{R}^d$ induces the norm $\|\cdot\|_2$ on $\mathbb{R}^{d\times d}$, defined by $\|\mathrm{A}\|_2 = \sup_{\|u\|_2=1}\|\mathrm{A}u\|_2$. In particular, for any matrices $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{d\times d}$ and any vector $x \in \mathbb{R}^d$, $\|\mathrm{A}x\|_2 \leq \|\mathrm{A}\|_2\|x\|_2$ and $\|\mathrm{AB}\|_2 \leq \|\mathrm{A}\|_2\|\mathrm{B}\|_2$. Moreover, if A is non-negative-definite, then $\|\mathrm{A}\|_2 = \lambda(\mathrm{A})$, where $\lambda(\mathrm{A})$ is the largest eigenvalue of A, and $\|\mathrm{A}^\alpha\|_2 = \lambda(\mathrm{A})^\alpha$ for any $\alpha > 0$.

Probability reminders. In this section, we consider a smooth manifold $\mathsf{M} \subset \mathbb{R}^d$. We first recall the definition of reversibility (Douc et al., 2018, Definition 1.5.1) before stating general results on reversibility.

Definition 4. *Let $\mathrm{Q} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$ be a transition probability kernel and let $\bar{\pi}$ be a probability distribution on $\mathrm{T}^\star\mathsf{M}$. Then,*

(a) Q is said to be reversible with respect to $\bar{\pi}$ if for any $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support

$$\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z, z')\mathrm{Q}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z', z)\mathrm{Q}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) .$$

(b) Q is said to be reversible up to momentum reversal with respect to $\bar{\pi}$ if for any $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support

$$\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z, z')\mathrm{Q}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'), s(z))\mathrm{Q}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) ,$$

where we recall that for any $(x, p) \in \mathbb{R}^d \times \mathbb{R}^d$, $s(x, p) = (x, -p)$.

Lemma 5. *Let $\mathrm{Q} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$ be a transition probability kernel and let $\bar{\pi}$ be a probability distribution on $\mathrm{T}^\star\mathsf{M}$. Assume that $s_\#\bar{\pi} = \bar{\pi}$ and that Q is reversible up to momentum reversal with respect to $\bar{\pi}$. Then $\bar{\pi}$ is an invariant measure for Q, i.e., for any $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support*

$$\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z')\mathrm{Q}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z)\bar{\pi}(\mathrm{d}z) .$$

Proof. Let $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support. We have

$$\begin{aligned}
&\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z')\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z))\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Definition 4)} \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z)\mathrm{Q}(s(z),\mathrm{d}z')(s_\#\bar{\pi})(\mathrm{d}z) = \int_{\mathrm{T}^\star\mathsf{M}} f(z)\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z) \;,
\end{aligned}$$

which concludes the proof. ☐

Lemma 6. *Let* $\mathrm{Q} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0,1]$ *be a transition probability kernel and let* $\bar{\pi}$ *be a probability distribution on* $\mathrm{T}^\star\mathsf{M}$*. Assume that* $s_\#\mathrm{Q} = \mathrm{Q}$ *and that* Q *is reversible with respect to* $\bar{\pi}$*. Then,* Q *is reversible up to momentum reversal with respect to* $\bar{\pi}$*.*

Proof. Let $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support. We have

$$\begin{aligned}
&\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,z')\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,s(z'))(s_\#\mathrm{Q})(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z') \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z',s(z))\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Definition 4)} \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'),s(z))(s_\#\mathrm{Q})(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z') \\
&\quad= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'),s(z))\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \;,
\end{aligned}$$

which concludes the proof. ☐

Useful inequalities. The following inequalities hold:

(a) For any $u \in [0, 1/5]$, $(1-u)^{-2} \leq 1 + 3u$ and $(1-u)^{-3} \leq 1 + 5u$.

(b) For any $u \in [0, 1/2]$, $(1-u)^{-1} \leq 1 + 2u$.

(c) For any $u \in [0, 1]$, $|(1-u)^2 - 1| \leq 3u$.

(d) For any $u \geq 0$, we have $1 - (1+u)^{-1} \leq u$ and $(1-u)^2 - 1 \geq -2u$.

B Details on Riemannian metrics

Let M be a smooth d-dimensional manifold, endowed with a metric $\mathfrak{g}$. We recall that the *Riemannian volume element* corresponding to $(\mathsf{M}, \mathfrak{g})$ is given in local coordinates by $\mathrm{dvol}_\mathsf{M}(x) = \sqrt{\det(\mathfrak{g})}\mathrm{d}x$, where $\mathrm{d}x$ is a dual coframe, (Lee, 2006, Lemma 3.2.). For any $x \in \mathsf{M}$, we respectively denote by $\mathrm{T}_x\mathsf{M}$ and $\mathrm{T}^\star_x\mathsf{M}$, the tangent space at x and its dual space, *i.e.*, the cotangent space at x. Note that $\mathrm{T}_x\mathsf{M}$ and $\mathrm{T}^\star_x\mathsf{M}$ are space vectors, and $\mathrm{T}_x\mathsf{M}$ is endowed with the scalar product $\langle\cdot,\cdot\rangle_{\mathfrak{g}(x)}$ by definition of the Riemannian metric. For clarity sake, we denote by v (resp. p) an element of the tangent (resp. cotangent) space. We recall that the tangent bundle $\mathrm{T}\mathsf{M}$ and the cotangent bundle $\mathrm{T}^\star\mathsf{M}$ are respectively defined by $\mathrm{T}\mathsf{M} = \sqcup_{x\in\mathsf{M}}\{x\} \cup \mathrm{T}_x\mathsf{M}$ and $\mathrm{T}^\star\mathsf{M} = \sqcup_{x\in\mathsf{M}}\{x\} \cup \mathrm{T}^\star_x\mathsf{M}$. These two sets are $2d$-dimensional manifolds.

Metric on the tangent bundle $\mathrm{T}\mathsf{M}$**.** Sasaki (1958) originally introduced on $\mathrm{T}\mathsf{M}$ a Riemannian metric $\hat{\mathfrak{g}}$, which, among other properties, preserves the Euclidean metric induced by $\mathfrak{g}$ on each tangent space. This metric is defined by

$$\hat{\mathfrak{g}} = \begin{pmatrix} \mathfrak{g}_{ij} + v^k v^l \Gamma^k_{is}\Gamma^l_{tj}\mathfrak{g}^{st} & -\mathfrak{g}^{js}v^k\Gamma^k_{is} \\ -\mathfrak{g}^{is}v^k\Gamma^k_{js} & \mathfrak{g}_{ij} \end{pmatrix} \;,$$

where $\mathfrak{g}_{ij}$ and $\mathfrak{g}^{ij}$ respectively refer to $\mathfrak{g}$ and $\mathfrak{g}^{-1}$, and Γ corresponds to the Christoffel symbol. Since $(\mathrm{T}\mathsf{M}, \hat{\mathfrak{g}})$ is a Riemannian manifold, the volume form on $\mathrm{T}\mathsf{M}$ satisfies for any $(x, v) \in \mathrm{T}\mathsf{M}$

$$\mathrm{dvol}_{\mathrm{T}\mathsf{M}}(x,v) = \sqrt{\det(\hat{\mathfrak{g}}(x,v))}\mathrm{d}x\mathrm{d}v = \det(\mathfrak{g}(x))\mathrm{d}x\mathrm{d}v \;.$$

Metric on the cotangent bundle $\mathrm{T}^\star\mathsf{M}$. Elaborating on the metric defined by Sasaki (1958), Mok (1977) showed that for any $x \in \mathsf{M}$, $\mathrm{T}_x^\star\mathsf{M}$ is naturally endowed with the scalar product $\langle\cdot,\cdot\rangle_{\mathfrak{g}(x)^{-1}}$, and proposed a Riemannian metric $\mathfrak{g}^\star$ on $\mathrm{T}^\star\mathsf{M}$, which notably preserves this result on each cotangent space. This metric is closely related to $\hat{\mathfrak{g}}$ and is defined by

$$\mathfrak{g}^\star = \begin{pmatrix} \mathfrak{g}_{ij} + p^k p^l \Gamma^k_{is}\Gamma^l_{tj}\mathfrak{g}^{st} & -\mathfrak{g}^{js}p^k\Gamma^k_{is} \\ -\mathfrak{g}^{is}p^k\Gamma^k_{js} & \mathfrak{g}^{ij} \end{pmatrix} .$$

Since $(\mathrm{T}^\star\mathsf{M}, \mathfrak{g}^\star)$ is a Riemannian manifold, the volume form on $\mathrm{T}^\star\mathsf{M}$ satisfies for any $(x,p) \in \mathrm{T}^\star\mathsf{M}$

$$\mathrm{dvol}_{\mathrm{T}^\star\mathsf{M}}(x,p) = \sqrt{\det(\mathfrak{g}^\star(x,p))}\mathrm{d}x\mathrm{d}p = \mathrm{d}x\mathrm{d}p .$$

In contrast to the tangent bundle, the volume form on the cotangent bundle does not depend on $\mathfrak{g}$, which motivates to augment on $\mathrm{T}^\star\mathsf{M}$ (instead of $\mathrm{T}\mathsf{M}$) any target measure π defined on M.

C Properties of self-concordance

We first recall the definition of self-concordance and derive some of its properties in Lemmas 8 and 9.

Definition 7 (Nesterov & Nemirovskii (1994))**.** *Let U be a non-empty open convex domain in $\mathbb{R}^d$. A function $\phi : \mathsf{U} \to \mathbb{R}$ is said to be a ν-self-concordant barrier (with $\nu \geq 1$) on U if it satisfies:*

(a) $\phi \in \mathrm{C}^3(\mathsf{U}, \mathbb{R})$ *and ϕ is convex,*

(b) $\phi(x) \to +\infty$ *as $x \to \partial\mathsf{U}$,*

(c) $|D^3\phi(x)[h,h,h]| \leq 2\|h\|^3_{\mathfrak{g}(x)}$, *for any $x \in \mathsf{M}$, $h \in \mathbb{R}^d$,*

(d) $|D\phi(x)[h]|^2 \leq \nu\|h\|^2_{\mathfrak{g}(x)}$, *for any $x \in \mathsf{M}$, $h \in \mathbb{R}^d$,*

where $\mathfrak{g}(x) = D^2\phi(x)$.

Self-concordance can be thought as a certain kind of regularity. Indeed, if ϕ is a ν-self-concordant barrier on a convex body U, then $D^2\phi$ is 2-Lipschitz continuous on U with respect to the local norm induced by $\mathfrak{g}$ (see Property (c)), and more restrictively, ϕ is ν-Lipschitz continuous with respect to the same norm (see Property (d)).

Lemma 8. *Let $\phi : \mathsf{U} \to \mathbb{R}$ be a ν-self-concordant barrier with $\mathfrak{g} = D^2\phi$. Assume that U is bounded. Then, $\|\nabla\phi(x)\|_2 \to +\infty$ and $\|\mathfrak{g}(x)\|_2 \to +\infty$ as $x \to \partial\mathsf{U}$.*

Proof. Since ϕ is convex, we have, for any $(x,y) \in \mathsf{U}^2$

$$\begin{aligned} \phi(x) &\leq \phi(y) + \nabla\phi(x)(x-y) , \\ |\phi(x)| &\leq |\phi(y)| + \|\nabla\phi(x)\|_2\|x-y\|_2 \\ &\leq |\phi(y)| + \|\nabla\phi(x)\|_2 \operatorname{diam}(\mathsf{U}) , \end{aligned}$$

where we used Cauchy-Schwartz inequality in the second line. Using Property (b) of ϕ, we obtain that $\|\nabla\phi(x)\|_2 \to +\infty$ as $x \to \partial\mathsf{U}$. By combining this result with Property (d) of ϕ, we obtain that $\|\mathfrak{g}(x)\|_2 \to +\infty$ as $x \to \partial\mathsf{U}$. □

According to Nesterov (2003) (see Lemma 4.1.2), Property (c) of self-concordance is equivalent to

$$|D^3\phi(x)[h_1,h_2,h_3]| \leq 2\textstyle\prod_{i=1}^3 \|h_i\|_{\mathfrak{g}(x)} ,$$

for any $x \in \mathsf{M}$ and any $(h_1,h_2,h_3) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$.

Let $u : \mathbb{R}^d \to \mathbb{R}^d$ be a linear operator. For any $x \in \mathbb{R}^d$, we define the operator norms $\|u\|_{\mathfrak{g}(x)}$ and $\|u\|_{\mathfrak{g}(x)^{-1}}$ by

$$\begin{aligned} \|u\|_{\mathfrak{g}(x)} &= \sup\{u(\mathfrak{g}(x)h) \ : \ \|h\|_{\mathfrak{g}(x)} = 1\} , \\ \|u\|_{\mathfrak{g}(x)^{-1}} &= \sup\{u(\mathfrak{g}(x)^{-1}h) \ : \ \|h\|_{\mathfrak{g}(x)^{-1}} = 1\} . \end{aligned}$$

In addition, a set $\mathsf{V} \subset \mathbb{R}^d$ contains no straight line if for any $\mathsf{D}_{x_0} = \{tx_0 \ : \ t \in \mathbb{R}\}$ with $x_0 \neq 0$, we have $\mathsf{V}^c \cap \mathsf{D}_{x_0} \neq \emptyset$.

Lemma 9. *Let $\phi : \mathsf{U} \to \mathbb{R}$ be a ν-self-concordant barrier with $\mathfrak{g} = D^2\phi$. Assume that U contains no straight line. For any $x \in \mathsf{U}$ and any $r > 0$, we denote by $\mathsf{W}^0(x, r)$ the open Dikin ellipsoid of ϕ at x, given by $\mathsf{W}^0(x, r) = \{y \in \mathbb{R}^d \mid \|y - x\|_{\mathfrak{g}(x)} < r\}$. Then, the following properties hold:*

(a) For any $x \in \mathsf{U}$ and any $(h_1, h_2) \in \mathbb{R}^d \times \mathbb{R}^d$

$$\|D^3\phi(x)[h_1, h_2]\|_{\mathfrak{g}(x)^{-1}} \leq 2\|h_1\|_{\mathfrak{g}(x)}\|h_2\|_{\mathfrak{g}(x)} \ .$$

(b) For any $x \in \mathsf{U}$, $\mathsf{W}^0(x, 1) \subset \mathsf{U}$, and for any $y \in \mathsf{W}^0(x, 1)$

$$(1 - \|y - x\|_{\mathfrak{g}(x)})^2\mathfrak{g}(x) \preceq \mathfrak{g}(y) \preceq (1 - \|y - x\|_{\mathfrak{g}(x)})^{-2}\mathfrak{g}(x) \ .$$

(c) For any $x \in \mathsf{U}$, any $y \in \mathsf{W}^0(x, 1)$ and any $h \in \mathbb{R}^d$, the following hold

$$\begin{aligned}
\|h\|_{\mathfrak{g}(x)} &\leq (1 - \|y - x\|_{\mathfrak{g}(x)})^{-1}\|h\|_{\mathfrak{g}(y)} \ , \\
\|h\|_{\mathfrak{g}(y)} &\leq (1 - \|y - x\|_{\mathfrak{g}(x)})^{-1}\|h\|_{\mathfrak{g}(x)} \ , \\
\|h\|_{\mathfrak{g}(x)^{-1}} &\leq (1 - \|y - x\|_{\mathfrak{g}(x)})^{-1}\|h\|_{\mathfrak{g}(y)^{-1}} \ , \\
\|h\|_{\mathfrak{g}(y)^{-1}} &\leq (1 - \|y - x\|_{\mathfrak{g}(x)})^{-1}\|h\|_{\mathfrak{g}(x)^{-1}} \ .
\end{aligned}$$

Proof. The result is a direct consequence of (Nesterov, 2003, Theorem 4.1.3, Theorem 4.1.5, Theorem 4.1.6). □

We now introduce α-regularity, which slightly strengthens self-concordance, by ensuring that $D^3\phi$ is $\alpha(\alpha + 1)$-Lipschitz continuous with respect to the local norm induced by $\mathfrak{g}$ (see (11)). We state below the definition of α-regularity as well as some of its properties in Lemma 11.

Definition 10 (Nesterov & Nemirovski (1998))**.** *Let $\alpha \geq 1$ and U be non-empty open convex domain in $\mathbb{R}^d$. A self-concordant function ϕ on U is said α-regular, if $\phi \in \mathrm{C}^4(\mathsf{U}, \mathbb{R})$ and for any $x \in \mathsf{U}$ and any $h \in \mathbb{R}^d$*

$$|D^4\phi(x)[h, h, h, h]| \leq \alpha(\alpha + 1)\|h\|^2_{\mathfrak{g}(x)}\|h\|^2_{\mathsf{U},x} \ ,$$

where $\|h\|_{\mathsf{U},x} := \inf\{t^{-1} \mid t > 0, x \pm th \in \mathsf{U}\}$.

Lemma 11. *Let $\phi : \mathsf{U} \to \mathbb{R}$ be a α-regular function with $\mathfrak{g} = D^2\phi$. Assume that U contains no straight line.*

(a) For any $x \in \mathsf{U}$ and any $h \in \mathbb{R}^d$

$$|D^4\phi(x)[h, h, h, h]| \leq \alpha(\alpha + 1)\|h\|^4_{\mathfrak{g}(x)} \ . \tag{11}$$

(b) For any $x \in \mathsf{U}$, any $(h_1, h_2, h_3, h_4) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$

$$|D^4\phi(x)[h_1, h_2, h_3, h_4]| \leq \alpha(\alpha + 1)\prod_{i=1}^{4}\|h_i\|_{\mathfrak{g}(x)} \ . \tag{12}$$

(c) For any $x \in \mathsf{U}$, any $y \in \mathsf{W}^0(x, 1)$ and any $(h_1, h_2) \in \mathbb{R}^d \times \mathbb{R}^d$

$$\|D^3\phi(x)[h_1, h_2] - D^3\phi(y)[h_1, h_2]\|_{\mathfrak{g}(x)^{-1}} \leq \alpha(\alpha + 1)/3\|h_1\|_{\mathfrak{g}(x)}\|h_2\|_{\mathfrak{g}(x)}((1 - \|y - x\|_{\mathfrak{g}(x)})^{-3} - 1) \ . \tag{13}$$

Proof. We first prove (11). First remark that the result is true for $h = 0$. Consider now $h \neq 0$. Let $\varepsilon > 0$. We define $t_\varepsilon = (1 + \varepsilon)^{-1}\|h\|^{-1}_{\mathfrak{g}(x)} > 0$, $y^+_\varepsilon = x + t_\varepsilon h$ and $y^-_\varepsilon = x - t_\varepsilon h$. We have $\|x - y^+_\varepsilon\|_{\mathfrak{g}(x)} = \|x - y^-_\varepsilon\|_{\mathfrak{g}(x)} = 1/(1 + \varepsilon) < 1$. Hence, $y^-_\varepsilon \in \mathsf{U}$ and $y^+_\varepsilon \in \mathsf{U}$ by Lemma 9. Therefore, $\|h\|_{\mathsf{U},x} \leq t^{-1}_\varepsilon$ and

$$|D^4\phi(x)[h, h, h, h]| \leq \alpha(\alpha + 1)\|h\|^4_{\mathfrak{g}(x)}(1 + \varepsilon)^2 \ .$$

We conclude upon taking $\varepsilon \to 0$. Then, (12) is an immediate consequence of (11) using (Nesterov & Nemirovskii, 1994, Proposition 9.1.1). We now prove (13). Using Lemma 9, we have

$$
\begin{aligned}
&\|D^3\phi(x)[h_1, h_2] - D^3\phi(y)[h_1, h_2]\|_{\mathfrak{g}(x)^{-1}} \\
&\quad \leq \int_0^1 \|D^4\phi(x + t(y-x))[h_1, h_2, y-x]\|_{\mathfrak{g}(x)^{-1}}\mathrm{d}t \\
&\quad \leq \int_0^1 (1 - t\|y-x\|_{\mathfrak{g}(x)})^{-1}\|D^4\phi(x + t(y-x))[h_1, h_2, y-x]\|_{\mathfrak{g}(x+t(y-x))^{-1}}\mathrm{d}t \\
&\quad \leq \int_0^1 \alpha(\alpha+1)/(1 - t\|y-x\|_{\mathfrak{g}(x)})\|h_1\|_{\mathfrak{g}(x+t(y-x))}\|h_2\|_{\mathfrak{g}(x+t(y-x))}\|y-x\|_{\mathfrak{g}(x+t(y-x))}\mathrm{d}t \\
&\quad \leq \alpha(\alpha+1)\|h_1\|_{\mathfrak{g}(x)}\|h_2\|_{\mathfrak{g}(x)} \int_0^1 \|y-x\|_{\mathfrak{g}(x)}/(1 - t\|y-x\|_{\mathfrak{g}(x)})^4\mathrm{d}t \\
&\quad \leq \alpha(\alpha+1)/3\|h_1\|_{\mathfrak{g}(x)}\|h_2\|_{\mathfrak{g}(x)}((1 - \|y-x\|_{\mathfrak{g}(x)})^{-3} - 1) \; ,
\end{aligned}
$$

which concludes the proof. □

We now prove the equivalence between the Euclidean norm and the norms induced by $\mathfrak{g}$ and $\mathfrak{g}^{-1}$.

Lemma 12. *Let* $(\mathsf{U}, \mathfrak{g})$ *be non-empty Riemannian manifold in* $\mathbb{R}^d$*. For any* $x_0 \in \mathsf{U}$*, any* $(x, p) \in \mathsf{U} \times \mathbb{R}^d$*, we have*

$$
\begin{aligned}
\|\mathfrak{g}(x_0)^{-1}\|_2^{-1/2}\|x\|_2 \leq \|x\|_{\mathfrak{g}(x_0)} \quad &\leq \|\mathfrak{g}(x_0)\|_2^{1/2}\|x\|_2 \; , \\
\|\mathfrak{g}(x_0)\|_2^{-1/2}\|p\|_2 \leq \|p\|_{\mathfrak{g}(x_0)^{-1}} &\leq \|\mathfrak{g}(x_0)^{-1}\|_2^{1/2}\|p\|_2 \; .
\end{aligned}
$$

In addition, let

$$
C_{x_0} = \max(\|\mathfrak{g}(x_0)\|_2^{1/2}, \|\mathfrak{g}(x_0)^{-1}\|_2^{1/2}) > 0 \; .
$$

Then, we have

$$
1/C_{x_0}(\|x\|_2 + \|p\|_2) \leq \|x\|_{\mathfrak{g}(x_0)} + \|p\|_{\mathfrak{g}(x_0)^{-1}} \leq C_{x_0}(\|x\|_2 + \|p\|_2) \; .
$$

D C-BHMC: Algorithm and Results

In this section, we first state general results on the Hamiltonian dynamics (3) under our main assumptions. It allows us to introduce the *continuous* version of BHMC (c-BHMC), for which we derive the reversibility with respect to π. Full proofs of this section are provided in Appendix E.

Results on the Hamiltonian dynamics. Using Cauchy-Lipschitz's theorem, we obtain in Proposition 13 the existence and uniqueness of the Hamiltonian dynamics (3), starting from any $z_0 \in \mathrm{T}^\star\mathsf{M}$.

Proposition 13. *Assume* **A**1*,* **A**2*. Let* $z_0 \in \mathrm{T}^\star\mathsf{M}$*. Then* (3) *admits a unique maximal solution* (J_{z_0}, z)*, where (i)* $J_{z_0} \subset \mathbb{R}$ *is an open neighbourhood of* 0*, (ii)* $z \in \mathrm{C}^1(J_{z_0}, \mathrm{T}^\star\mathsf{M})$*, (iii)* $z(0) = z_0$ *and (iv) for any* $t \in J_{z_0}$*,* $H(z(t)) = H(z_0)$*.*

In contrast to the Euclidean case (*i.e.*, $\mathfrak{g}(x) = \mathrm{I}_d$), the Hamiltonian dynamics (3) relies on a metric derived from a self-concordant barrier, and thus is defined up to an explosion time. In the next result, we characterize the behaviour of the solutions when this explosion time is finite.

Proposition 14. *Assume* **A**1*,* **A**2*. Let* $z_0 \in \mathrm{T}^\star\mathsf{M}$*,* (J_{z_0}, z) *as in Proposition 13. Assume* $T_{z_0} = \sup J_{z_0} < +\infty$*, then we have*

(a) $\lim_{t \to T_{z_0}} \|p(t)\|_2 = +\infty$.

(b) There exists $x_\mathsf{M} \in \partial\mathsf{M}$ *such that* $\lim_{t \to T_{z_0}} x(t) = x_\mathsf{M}$*.*

In the rest of this section, we define the *Hamiltonian flow* as the map $\mathrm{T} : \mathsf{D} \to \mathrm{T}^\star\mathsf{M}$, where $\mathsf{D} = \{(h, z_0) : z_0 \in \mathrm{T}^\star\mathsf{M}, h < \sup J_{z_0}\}$, such that $\mathrm{T}_h(z_0) = z(h)$ for any $(h, z_0) \in \mathsf{D}$, z being uniquely defined from z_0 in Proposition 13.

Introduction to c-BHMC. In the case where $\mathsf{D} = \mathrm{T}^\star\mathsf{M} \times \mathbb{R}$, the definition of the Hamiltonian dynamics ensures that T_h is involutive (up to momentum reversal) for any $h > 0$, which is key to derive reversibility in BHMC. However, in our manifold setting, there is one subtlety that needs to be checked, namely that the Hamiltonian flow does not leave the cotangent bundle in finite time. As detailed in Proposition 14, there is indeed no guarantee that this flow is defined for all times, *i.e.*, that we have $\mathsf{D} = \mathrm{T}^\star\mathsf{M} \times \mathbb{R}$, due to the ill-conditioned behavior of the dynamics near the boundary of the manifold. Therefore, given a time horizon $h > 0$, we restrict the cotangent bundle to the points where the Hamiltonian flow is defined up to time h. Hence, we define, for any $h > 0$, the sets

$$\mathsf{E}_h = \{z_0 \in \mathrm{T}^\star\mathsf{M} \;:\; h < \sup J_{z_0}\} \;, \quad \bar{\mathsf{E}}_h = \{z_0 \in \mathsf{E}_h \;:\; h < \sup J_{(s\circ \mathrm{T}_h)(z_0)}\} \;.$$

Note that elements of $\bar{\mathsf{E}}_h$ correspond to points z_0 for which:

(a) we are able to compute *exactly* the Hamiltonian dynamics (3) starting from z_0 on the time interval $[0, h]$, by considering $\{\mathrm{T}_t(z_0)\}_{t\in[0,h]}$ (since $\bar{\mathsf{E}}_h \subset \mathsf{E}_h$),

(b) we are also able to compute *exactly* the **reversed** dynamics - up to momentum reversal - starting from $(s \circ \mathrm{T}_h)(z_0)$ on the time interval $[0, h]$.

In Appendix E, we prove that T_h is symplectic (up to momentum reversal) on $\bar{\mathsf{E}}_h$ for any $h > 0$, which thus ensures that the Hamiltonian dynamics is *reversible* with respect to $\bar{\pi}$, see Lemma 16.

Let $h > 0$. Our algorithm c-BHMC, whose pseudo-code is provided in Algorithm 2, then proceeds as follows at stage $n \geq 1$. Assume that the current state is $(X_{n-1}, P_{n-1}) \in \mathrm{T}^\star\mathsf{M}$. First define a partial refreshment of the momentum $\tilde{P}_n \leftarrow \sqrt{1-\beta}P_{n-1} + \sqrt{\beta}G_n$, where $G_n \sim \mathrm{N}_x(0, \mathrm{I}_d)$ is independent from P_{n-1}. Then, verify that $(X_{n-1}, \tilde{P}_n) \in \bar{\mathsf{E}}_h$, which ensures that the Hamiltonian flow is involutive on the time interval $[0, h]$, starting at this state. In the same spirit as in Algorithm 1, we refer to this step as an "involution checking step", which is highlighted in yellow in Algorithm 2, although the involution property directly derives from the Hamiltonian dynamics. If this step is valid, define $(X_n, \bar{P}_n) = \mathrm{T}_h(X_{n-1}, \tilde{P}_n)$; otherwise, set $(X_n, \bar{P}_n) \leftarrow s(X_{n-1}, \tilde{P}_n)$. Finally, update the momentum as in the beginning of the iteration to obtain the new state (X_n, P_n).

Algorithm 2: c-BHMC with Momentum Refresh

Input: $(X_0, P_0) \in \mathrm{T}^\star\mathsf{M}$, $\beta \in (0, 1]$, $N \in \mathbb{N}$, $h > 0$
Output: $(X_n, P_n)_{n\in[N]}$

for $n = 1, \ldots, N$ **do**
  `Step 1:` $G_n \sim \mathrm{N}_x(0, \mathrm{I}_d)$, $\tilde{P}_n \leftarrow \sqrt{1-\beta}P_{n-1} + \sqrt{\beta}G_n$
  `Step 2: solving continuous ODE (3)`
  if $(X_{n-1}, \tilde{P}_n) \in \bar{\mathsf{E}}_h$ **then** $(X_n, \bar{P}_n) \leftarrow \mathrm{T}_h(X_{n-1}, \tilde{P}_n)$
  else $(X_n, \bar{P}_n) \leftarrow s(X_{n-1}, \tilde{P}_n)$
  `Step 3:` $G'_n \sim \mathrm{N}_x(0, \mathrm{I}_d)$, $P_n \leftarrow \sqrt{1-\beta}\bar{P}_n + \sqrt{\beta}G'_n$
end

Denote by $\mathrm{Q}_c : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel of the (homogeneous) Markov chain $(X_n, P_n)_{n\in[N]}$ obtained with Algorithm 2. Using the properties of the Hamiltonian dynamics, we get the following result.

Theorem 15. *Assume* **A**1, **A**2*. Then,* Q_c *is reversible up to momentum reversal, i.e., we have for any* $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ *with compact support*

$$\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z, z')\mathrm{d}\bar{\pi}(z)\mathrm{Q}_c(z, \mathrm{d}z') = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'), s(z))\mathrm{d}\bar{\pi}(z)\mathrm{Q}_c(z, \mathrm{d}z') \;.$$

In particular, $\bar{\pi}$ *is an invariant measure for* Q_c*.*

E Proofs of Appendix D

E.1 Existence, uniqueness and explosion time of Hamiltonian dynamics

We prove below Proposition 13 and Proposition 14.

Proof of Proposition 13. Assume **A**1, **A**2. In particular, M contains no straight line, and Lemmas 9 and 11 apply here. We rewrite the ODE (3) as a Cauchy problem defined on Banach space $\mathsf{E} = (\mathbb{R}^d \times \mathbb{R}^d, \|\cdot\|_{\mathsf{E}})$ where $\|(a, b)\|_{\mathsf{E}} = \|a\|_2 + \|b\|_2$ for any pair $(a, b) \in \mathbb{R}^d \times \mathbb{R}^d$. We define $\Omega = \mathsf{M} \times \mathbb{R}^d$, which is an open set in E. A solution (J, z) to this Cauchy problem is defined for all $t \in J$ by

$$\dot{z}(t) = F(z(t)) \;, \; z(0) = z_0 \;, \quad \text{where} \quad F(x, p) = (\partial_p H(x, p), -\partial_x H(x, p)) \;,$$

such that J is an open interval of $\mathbb{R}$ with $0 \in J$, and $z = (x, p) : J \to \Omega$ is differentiable on J. We now prove two results on F to establish existence and uniqueness of a maximal solution (J, z) for this Cauchy problem:

(a) F is continuously differentiable on Ω.

(b) F is locally Lipschitz-continuous on Ω with respect to $\|\cdot\|_{\mathsf{E}}$.

Since $\nabla V, D\mathfrak{g} \in \mathrm{C}^1(\mathsf{M}, \mathbb{R})$, we directly obtain Item (a). We now prove Item (b). Let $\bar{z} = (\bar{x}, \bar{p}) \in \Omega$. We endow E with the norm $\|\cdot\|_{\bar{z}}$. Since M is an open subset of $\mathbb{R}^d$, there exists $0 < r \leq 1/(11 C_{\bar{x}})$, where we recall that $C_{\bar{x}}$ is defined in Lemma 12, such that $\mathsf{B}_{\|\cdot\|_2}(\bar{x}, r) \subset \mathsf{M}$. We consider such r, and we define $\mathsf{B} = \mathsf{B}_{\|\cdot\|_{\mathsf{E}}}(\bar{z}, r)$ and $\mathsf{B}_{\mathfrak{g}} = \mathsf{B}_{\|\cdot\|_{\bar{z}}}(\bar{z}, \bar{r})$, where $\bar{r} = C_{\bar{x}} r \leq 1/11$. In particular, we have by Lemma 9-(b) that $\mathsf{B}_{\mathfrak{g}} \subset \mathsf{B}_{\|\cdot\|_{\mathfrak{g}(\bar{x})}}(\bar{x}, \bar{r}) \times \mathsf{B}_{\|\cdot\|_{\mathfrak{g}(\bar{x})^{-1}}}(\bar{p}, \bar{r}) \subset \Omega$ since $\bar{r} < 1$. Moreover, using Lemma 12, we have $\mathsf{B} \subset \mathsf{B}_{\mathfrak{g}} \subset \Omega$.

Since $V \in \mathrm{C}^2(\mathsf{M}, \mathbb{R})$, ∇V is Lipschitz-continuous on $\mathsf{B}_{\mathfrak{g}}$, *i.e.*, there exists $C_V > 0$ such that for any $(z, z') \in \mathsf{B}_{\mathfrak{g}} \times \mathsf{B}_{\mathfrak{g}}$,

$$\|\nabla V(x) - \nabla V(x')\|_{\mathfrak{g}(\bar{x})^{-1}} \leq C_V \|x - x'\|_{\mathfrak{g}(\bar{x})} \; . \tag{14}$$

We first show that F is Lipschitz-continuous on $\mathsf{B}_{\mathfrak{g}}$ with respect to $\|\cdot\|_{\bar{z}}$. Consider $(z, z') \in \mathsf{B}_{\mathfrak{g}} \times \mathsf{B}_{\mathfrak{g}}$ denoted by $z = (x, p)$ and $z' = (x', p')$. Note that $(x, x') \in W^0(\bar{x}, 1) \times W^0(\bar{x}, 1)$, where $W^0(\bar{x}, 1)$ is defined in Lemma 9. We first bound $\|F^{(1)}(z) - F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} = \|\partial_p H(z) - \partial_p H(z')\|_{\mathfrak{g}(\bar{x})}$. We have

$$\partial_p H(z) - \partial_p H(z') = \mathfrak{g}(x)^{-1} p - \mathfrak{g}(x')^{-1} p' = \mathfrak{g}(x)^{-1}(p - p') + (\mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1}) p' \; ,$$

then,

$$\begin{aligned}\|\partial_p H(z) - \partial_p H(z')\|_{\mathfrak{g}(\bar{x})} \leq \; & \|\mathfrak{g}(\bar{x})^{1/2} \mathfrak{g}(x)^{-1} \mathfrak{g}(\bar{x})^{1/2}\|_2 \|p - p'\|_{\mathfrak{g}(\bar{x})^{-1}} \\ & + \|\mathfrak{g}(\bar{x})^{1/2} (\mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1}) \mathfrak{g}(\bar{x})^{1/2}\|_2 \|p'\|_{\mathfrak{g}(\bar{x})^{-1}} \; .\end{aligned} \tag{15}$$

Using Lemma 9-(b), we have

$$(1 - \|\bar{x} - x\|_{\mathfrak{g}(x)})^2 \mathfrak{g}(\bar{x})^{-1} \preceq \mathfrak{g}(x)^{-1} \preceq (1 - \|\bar{x} - x\|_{\mathfrak{g}(x)})^{-2} \mathfrak{g}(\bar{x})^{-1} \; ,$$

and thus,

$$\begin{aligned}(1 - \|\bar{x} - x\|_{\mathfrak{g}(x)})^2 \mathrm{I}_d \preceq \mathfrak{g}(\bar{x})^{1/2} \mathfrak{g}(x)^{-1} \mathfrak{g}(\bar{x})^{1/2} &\preceq (1 - \|\bar{x} - x\|_{\mathfrak{g}(x)})^{-2} \mathrm{I}_d \\ (1 - \bar{r})^2 \mathrm{I}_d \prec \mathfrak{g}(\bar{x})^{1/2} \mathfrak{g}(x)^{-1} \mathfrak{g}(\bar{x})^{1/2} &\prec (1 - \bar{r})^{-2} \mathrm{I}_d \; .\end{aligned}$$

Therefore, we have

$$\|\mathfrak{g}(\bar{x})^{1/2} \mathfrak{g}(x)^{-1} \mathfrak{g}(\bar{x})^{1/2}\|_2 \leq (1 - \bar{r})^{-2} \; . \tag{16}$$

In addition, we have

$$\begin{aligned}\|x - x'\|_{\mathfrak{g}(x)} &\leq (1 - \|\bar{x} - x'\|_{\mathfrak{g}(\bar{x})})^{-1} \|x - x\|_{\mathfrak{g}(\bar{x})} \\ &< (1 - \bar{r})^{-1} 2\bar{r} \\ &< 1/5,\end{aligned} \tag{17}$$

where we used Lemma 9-(c) in the first inequality and $\bar{r} \leq 1/11$ in the last inequality. Therefore, $x' \in \mathrm{W}^0(x, 1)$ and we have using Lemma 9-(b)

$$\begin{aligned}\{(1 - \|x' - x\|_{\mathfrak{g}(x)})^2 - 1\} \mathfrak{g}(x')^{-1} &\preceq \mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1} \\ &\preceq \{(1 - \|x' - x\|_{\mathfrak{g}(x)})^{-2} - 1\} \mathfrak{g}(x')^{-1} \\ \{(1 - \|x' - x\|_{\mathfrak{g}(x)})^2 - 1\}(1 - \bar{r})^2 \mathrm{I}_d &\preceq \mathfrak{g}(\bar{x})^{1/2} (\mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1}) \mathfrak{g}(\bar{x})^{1/2} \\ &\preceq \{(1 - \|x' - x\|_{\mathfrak{g}(x)})^{-2} - 1\}(1 - \bar{r})^{-2} \mathrm{I}_d \; .\end{aligned}$$

Then,

$$\begin{aligned}&\|\mathfrak{g}(\bar{x})^{1/2} (\mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1}) \mathfrak{g}(\bar{x})^{1/2}\|_2 \\ &\quad \leq \max(|(1 - \|x' - x\|_{\mathfrak{g}(x)})^2 - 1|(1 - \bar{r})^2, \{(1 - \|x' - x\|_{\mathfrak{g}(x)})^{-2} - 1\}(1 - \bar{r})^{-2}) \\ &\quad \leq (1 - \bar{r})^{-2} \max(|(1 - \|x' - x\|_{\mathfrak{g}(x)})^2 - 1|, (1 - \|x' - x\|_{\mathfrak{g}(x)})^{-2} - 1) \; .\end{aligned} \tag{18}$$

Using Inequalities (a) and (c) with $u = \|x' - x\|_{\mathfrak{g}(x)}$, where $u \leq 1/5$ by (17), we obtain

(a) $|(1-\|x'-x\|_{\mathfrak{g}(x)})^2 - 1| \leq 3\|x'-x\|_{\mathfrak{g}(x)} < 3(1-\bar{r})^{-1}\|x-x'\|_{\mathfrak{g}(\bar{x})}$,

(b) $(1-\|x'-x\|_{\mathfrak{g}(x)})^{-2} - 1 \leq 3\|x'-x\|_{\mathfrak{g}(x)} < 3(1-\bar{r})^{-1}\|x-x'\|_{\mathfrak{g}(\bar{x})}$.

Moreover, we have

$$\|p'\|_{\mathfrak{g}(\bar{x})^{-1}} = \|p'-\bar{p}+\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}} \leq \bar{r} + \|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}} .$$

Combining (15), (16) and (18), it comes that

$$\|F^{(1)}(z)-F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} \leq (1-\bar{r})^{-2}\|p-p'\|_{\mathfrak{g}(\bar{x})^{-1}} + 3(1-\bar{r})^{-3}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})\|x-x'\|_{\mathfrak{g}(\bar{x})} . \quad (19)$$

We define $A_{\bar{x},1} = (1-\bar{r})^{-2} + 3(1-\bar{r})^{-3}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})$ such that $\|F^{(1)}(z)-F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} \leq A_{\bar{x},1}\|z-z'\|_{\bar{z}}$ and we aim to bound $\|F^{(2)}(z)-F^{(2)}(x')\|_{\mathfrak{g}(\bar{x})^{-1}} = \|\partial_x H(z) - \partial_x H(z')\|_{\mathfrak{g}(\bar{x})^{-1}}$. We have

$$\begin{aligned}\partial_x H(z) - \partial_x H(z') =& \nabla V(x) - \nabla V(x') \\ &+ \tfrac{1}{2}D\mathfrak{g}(x')[\mathfrak{g}(x')^{-1}p', \mathfrak{g}(x')^{-1}p'] - \tfrac{1}{2}D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p, \mathfrak{g}(x)^{-1}p] \\ &+ \tfrac{1}{2}\mathfrak{g}(x)^{-1} : D\mathfrak{g}(x) - \tfrac{1}{2}\mathfrak{g}(x')^{-1} : D\mathfrak{g}(x')\end{aligned} \quad (20)$$

We have

$$\begin{aligned}&\|D\mathfrak{g}(x')[\mathfrak{g}(x')^{-1}p', \mathfrak{g}(x')^{-1}p'] - D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p, \mathfrak{g}(x)^{-1}p]\|_{\mathfrak{g}(\bar{x})^{-1}} \\ &\quad \leq \|D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p, \mathfrak{g}(x)^{-1}p] - D\mathfrak{g}(x)[\mathfrak{g}(x')^{-1}p', \mathfrak{g}(x')^{-1}p']\|_{\mathfrak{g}(\bar{x})^{-1}} \\ &\qquad + \|D\mathfrak{g}(x)[\mathfrak{g}(x')^{-1}p', \mathfrak{g}(x')^{-1}p'] - D\mathfrak{g}(x')[\mathfrak{g}(x')^{-1}p', \mathfrak{g}(x')^{-1}p']\|_{\mathfrak{g}(\bar{x})^{-1}}\end{aligned} \quad (21)$$

Note that we have $\mathfrak{g}(x)^{-1}p = F^{(1)}(z)$ (resp. $\mathfrak{g}(x')^{-1}p' = F^{(1)}(z')$). Using Lemma 9-(c), the first upper bound in (21) is bounded by

$$\begin{aligned}&(1-\bar{r})^{-1}\|D\mathfrak{g}(x)[F^{(1)}(z), F^{(1)}(z)] - D\mathfrak{g}(x)[F^{(1)}(z'), F^{(1)}(z')]\|_{\mathfrak{g}(x)^{-1}} \\ &\quad \leq 2(1-\bar{r})^{-1}\|D\mathfrak{g}(x)[F^{(1)}(z') - F^{(1)}(z), F^{(1)}(z')]\|_{\mathfrak{g}(x)^{-1}} \\ &\qquad + (1-\bar{r})^{-1}\|D\mathfrak{g}(x)[F^{(1)}(z') - F^{(1)}(z), F^{(1)}(z') - F^{(1)}(z)]\|_{\mathfrak{g}(x)^{-1}} \\ &\quad \leq 4(1-\bar{r})^{-1}\|F^{(1)}(z') - F^{(1)}(z)\|_{\mathfrak{g}(x)}\|F^{(1)}(z')\|_{\mathfrak{g}(x)} \\ &\qquad + 2(1-\bar{r})^{-1}\|F^{(1)}(z') - F^{(1)}(z)\|^2_{\mathfrak{g}(x)} && \text{(Lemma 9-(a))} \\ &\quad \leq 4(1-\bar{r})^{-3}\|F^{(1)}(z') - F^{(1)}(z)\|_{\mathfrak{g}(\bar{x})}\|F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} \\ &\qquad + 2(1-\bar{r})^{-3}\|F^{(1)}(z') - F^{(1)}(z)\|^2_{\mathfrak{g}(\bar{x})} && \text{(Lemma 9-(c))} \\ &\quad \leq 4(1-\bar{r})^{-3}A_{\bar{x},1}\|z-z'\|_{\bar{z}}\|F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} + 2(1-\bar{r})^{-3}A^2_{\bar{x},1}\|z-z'\|^2_{\bar{z}} && \text{(see (19))} \\ &\quad \leq \{4A_{\bar{x},1}(1-\bar{r})^{-5}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}}) + 4(1-\bar{r})^{-3}A^2_{\bar{x},1}\bar{r}\}\|z-z'\|_{\bar{z}} ,\end{aligned}$$

where we used that $\|F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} \leq (1-\bar{r})^{-2}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})$ in the last inequality. Combining Lemma 9-(c) and Lemma 11-(c), the second upper bound in (21) is bounded by

$$\begin{aligned}&(1-\bar{r})^{-1}\|D\mathfrak{g}(x)[F^{(1)}(z'), F^{(1)}(z')] - D\mathfrak{g}(x')[F^{(1)}(z'), F^{(1)}(z')]\|_{\mathfrak{g}(x)^{-1}} \\ &\quad \leq \alpha(\alpha+1)(1-\bar{r})^{-1}/3\left\{(1-\|x'-x\|_{\mathfrak{g}(x)})^{-3} - 1\right\}\|F^{(1)}(z')\|^2_{\mathfrak{g}(x)} \\ &\quad \leq \alpha(\alpha+1)(1-\bar{r})^{-3}/3\left\{(1-(1-\bar{r})^{-1}\|x'-x\|_{\mathfrak{g}(\bar{x})})^{-3} - 1\right\}\|F^{(1)}(z')\|^2_{\mathfrak{g}(\bar{x})} \\ &\quad \leq \alpha(\alpha+1)(1-\bar{r})^{-7}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})^2/3\left\{(1-(1-\bar{r})^{-1}\|x'-x\|_{\mathfrak{g}(\bar{x})})^{-3} - 1\right\} .\end{aligned}$$

Using Inequality (a) with $u = (1-\bar{r})^{-1}\|x'-x\|_{\mathfrak{g}(\bar{x})}$, where $u \leq 1/5$ by (17), we obtain

$$(1-(1-\bar{r})^{-1}\|x'-x\|_{\mathfrak{g}(\bar{x})})^{-3} - 1 \leq 5(1-\bar{r})^{-1}\|x'-x\|_{\mathfrak{g}(\bar{x})} . \quad (22)$$

Then, the second upper bound in (21) is bounded by

$$(5/3)\alpha(\alpha+1)(1-\bar{r})^{-8}(\bar{r}+\|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})^2\|x'-x\|_{\mathfrak{g}(\bar{x})} .$$

We now define $h : y \in \mathsf{M} \to \mathfrak{g}(y)^{-1} : D\mathfrak{g}(y)$. Since $\phi \in \mathrm{C}^4(\mathsf{M})$, $h \in \mathrm{C}^1(\mathsf{M})$. Moreover, $[x, x'] \in \mathsf{M}$ by convexity of M and we can define $A_\phi = \sup_{y\in[x,x']} \|Dh(y)\|$, where $\|Dh(y)\| = \sup_{\|u\|_{\mathfrak{g}(\bar{x})}=1} \|Dh(y)[u]\|_{\mathfrak{g}(\bar{x})^{-1}}$. Using the mean value theorem on h, we have

$$\|h(x) - h(x')\|_{\mathfrak{g}(\bar{x})^{-1}} \le A_\phi \|x' - x\|_{\mathfrak{g}(\bar{x})} \ . \tag{23}$$

By combining (14), (20), (21) and (23), we have

$$\begin{aligned} &\|F^{(2)}(z) - F^{(2)}(z')\|_{\mathfrak{g}(\bar{x})^{-1}} \\ &\quad \le C_V \|x - x'\|_{\mathfrak{g}(\bar{x})} \\ &\qquad + (1/2)\{4A_{\bar{x},1}(1-\bar{r})^{-5}(\bar{r} + \|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}}) + 4A_{\bar{x},1}^2(1-\bar{r})^{-3}\bar{r}\}\|z - z'\|_{\bar{z}} \\ &\qquad + (1/2)\cdot(5/3)\alpha(\alpha+1)(1-\bar{r})^{-8}(\bar{r} + \|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})^2\|x' - x\|_{\mathfrak{g}(\bar{x})} \\ &\qquad + (1/2)A_\phi\|x' - x\|_{\mathfrak{g}(\bar{x})} \ . \end{aligned}$$

We define $A_{\bar{x},2} = C_V + 2A_{\bar{x},1}(1-\bar{r})^{-5}(\bar{r} + \|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}}) + 2A_{\bar{x},1}^2(1-\bar{r})^{-3}\bar{r} + (5/6)\alpha(\alpha+1)(1-\bar{r})^{-8}(\bar{r} + \|\bar{p}\|_{\mathfrak{g}(\bar{x})^{-1}})^2 + A_\phi/2$ such that

$$\|F^{(2)}(z) - F^{(2)}(z')\|_{\mathfrak{g}(\bar{x})} \le A_{\bar{x},2}\|z - z'\|_{\bar{z}} \ .$$

Finally, we have

$$\begin{aligned} \|F(z) - F(z')\|_{\bar{z}} &= \|F^{(1)}(z) - F^{(1)}(z')\|_{\mathfrak{g}(\bar{x})} + \|F^{(2)}(z) - F^{(2)}(z')\|_{\mathfrak{g}(\bar{x})^{-1}} \\ &\le (A_{\bar{x},1} + A_{\bar{x},2})\|z - z'\|_{\bar{z}} \ , \end{aligned}$$

i.e., F is $(A_{\bar{x},1} + A_{\bar{x},2})$-Lipschitz-continuous on $\mathsf{B}_{\mathfrak{g}}$ with respect to $\|\cdot\|_{\bar{z}}$.

We now show that F is Lipschitz-continuous on B with respect to $\|\cdot\|_{\mathsf{E}}$. Consider $(z, z') \in \mathsf{B} \times \mathsf{B}$ denoted by $z = (x, p)$ and $z' = (x', p')$. In particular, $(z, z') \in \mathsf{B}_{\mathfrak{g}} \times \mathsf{B}_{\mathfrak{g}}$. Using the previous result and Lemma 12, we have

$$\|F(z) - F(z')\|_{\mathsf{E}} \le C_{\bar{x}}\|F(z) - F(z')\|_{\bar{z}} \le C_{\bar{x}}(A_{\bar{x},1} + A_{\bar{x},2})\|z - z'\|_{\bar{z}} \le C_{\bar{x}}^2(A_{\bar{x},1} + A_{\bar{x},2})\|z - z'\|_{\mathsf{E}} \ ,$$

which proves that F is $C_{\bar{x}}^2(A_{\bar{x},1} + A_{\bar{x},2})$-Lipschitz-continuous on B with respect to $\|\cdot\|_{\mathsf{E}}$.

Conclusion. Combining Items (a) and (b), we obtain the results (i), (ii) and (iii) of Proposition 13 upon using Cauchy-Lipschitz's theorem. Moreover, for any $t \in J_{z_0}$, $\partial_t H(z(t)) = \partial_x H(z(t))^\top \dot{x}(t) + \partial_p H(z(t))^\top \dot{p}(t) = -\dot{p}(t)^\top \dot{x}(t) + \dot{p}(t)^\top \dot{x}(t) = 0$, which proves the result (iv) of Proposition 13. □

Proof of Proposition 14. Assume **A**1, **A**2. Let $z_0 \in \mathrm{T}^\star\mathsf{M}$, (J_{z_0}, z) as in Proposition 13. Assume that $T_{z_0} = \sup J_{z_0} < +\infty$. Then, by Cauchy-Lipschitz's theorem, z leaves any compact set of E at the neighbourhood of T_{z_0}. By construction of $\|\cdot\|_{\mathsf{E}}$, this property is induced on both variables $x \in \mathsf{M}$ and $p \in \mathbb{R}^d$, each with respect to $\|\cdot\|_2$. We directly obtain the result of Proposition 14 by continuity of (x, p). □

E.2 Proof of reversibility in Algorithm 2

In (RM)HMC, one often sets the time horizon of the Hamiltonian dynamics with a hyperparameter h. However, under **A**1 and **A**2, we are not able to prove that the Hamiltonian dynamics (3) is defined at time $h > 0$, given *any* starting point $z_0 \in \mathrm{T}^\star\mathsf{M}$. Actually, this dynamics explodes if the time horizon is finite (see Proposition 14). This theoretical limitation requires to implement some sort of "involution checking step" (see Line 4 in Algorithm 2), which is theoretically explained below.

In the rest of this section, for any $z_0 \in \mathrm{T}^\star\mathsf{M}$, we will denote by $z(\cdot, z_0) \in \mathrm{C}^1(J_{z_0}, \mathrm{T}^\star\mathsf{M})$ the maximal solution to (3) with starting point z_0, uniquely defined in Proposition 13. For any $h > 0$, we recall below the definition of the sets E_h and $\bar{\mathsf{E}}_h$ the map T_h introduced in Appendix D:

(a) $\mathsf{E}_h = \{z_0 \in \mathrm{T}^\star\mathsf{M} \ : \ h < \sup J_{z_0}\}$,

(b) $\bar{\mathsf{E}}_h = \{z_0 \in \mathsf{E}_h \ : \ h < \sup J_{(s\circ\mathrm{T}_h)(z_0)}\}$,

(c) the map $\mathrm{T}_h : \mathsf{E}_h \to \mathrm{T}^\star\mathsf{M}$ is such that $\mathrm{T}_h(z_0) = z(h, z_0)$ for any $z_0 \in \mathsf{E}_h$.

In particular, we have

$$\bar{\mathsf{E}}_h = \mathsf{E}_h \cap (s \circ \mathrm{T}_h)^{-1}(\mathsf{E}_h) \ . \tag{24}$$

We first prove that the restriction of $s \circ \mathrm{T}_h$ to $\bar{\mathsf{E}}_h$ is an involutive integrator, which is crucial to derive the "reversibility" of Algorithm 2 in Theorem 15.

Lemma 16. *Assume* **A**1, **A**2*. Then, for any $h > 0$, $s \circ \mathrm{T}_h$ is a C^1-diffeomorphism on $\bar{\mathsf{E}}_h$ such that $|\det\mathrm{Jac}[s \circ \mathrm{T}_h]| \equiv 1$.*

Proof. Assume **A**1, **A**2. Let $h > 0$. It is clear that T_h (and thus $s \circ \mathrm{T}_h$) is well defined and continuously differentiable on E_h (in particular on $\bar{\mathsf{E}}_h$), by elaborating on the proof of Proposition 13 combined with Cauchy-Lipschitz's theorem. Let $z_0 \in \bar{\mathsf{E}}_h$. By definition of $\bar{\mathsf{E}}_h$, $\mathrm{T}_h(z_0)$, resp. J_{z_0}, and $(\mathrm{T}_h \circ s \circ \mathrm{T}_h)(z_0)$, resp. $J_{(s\circ \mathrm{T}_h)(z_0)}$, are well defined. We now aim to prove that $(\mathrm{T}_h \circ s \circ \mathrm{T}_h)(z_0) = s(z_0)$.

We define $z' : t \in [0, h] \mapsto s(z(h-t, z_0))$, which existence is straightforward since $(h-t) \in J_{z_0}$ for any $t \in [0, h]$. In particular, $z'(0) = (s \circ \mathrm{T}_h)(z_0)$ and z' is a solution of the ODE (3) on the interval $[0, h]$. Yet, $[0, h] \subset J_{(s\circ \mathrm{T}_h)(z_0)}$, and $z'(0) = z(0, (s\circ \mathrm{T}_h)(z_0))$. Then, by unicity of $z(\cdot, (s\circ \mathrm{T}_h)(z_0))$ (see Proposition 13), $z(\cdot, (s \circ \mathrm{T}_h)(z_0))$ and z' coincide on $[0, h]$. In particular, we have at time h

$$s(z_0) = z'(h) = z(h, (s \circ \mathrm{T}_h)(z_0)) = (\mathrm{T}_h \circ s \circ \mathrm{T}_h)(z_0) \ .$$

Then for any $z_0 \in \bar{\mathsf{E}}_h$, $(s \circ \mathrm{T}_h \circ s \circ \mathrm{T}_h)(z_0) = z_0$, *i.e.*, $s \circ \mathrm{T}_h$ is an involution on $\bar{\mathsf{E}}_h$. Since $s \circ \mathrm{T}_h \in \mathrm{C}^1(\bar{\mathsf{E}}_h, \mathrm{T}^\star\mathsf{M})$, $s \circ \mathrm{T}_h$ is a C^1-diffeomorphism on $\bar{\mathsf{E}}_h$ such that $|\det\mathrm{Jac}[s \circ \mathrm{T}_h]| \equiv 1$. □

We recall that we denote by $\mathrm{Q}_c : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel of the (homogeneous) Markov chain $(X_n, P_n)_{n\in[N]}$ obtained with Algorithm 2. We also denote by:

(a) $\mathrm{Q}_0 : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel referring to Step 1 (also Step 3) in Algorithm 2.

(b) $\mathrm{Q}_{c,1} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel referring to Step 2 in Algorithm 2.

We provide below details on Markov kernels Q_0, $\mathrm{Q}_{c,1}$ and Q_c.

Kernel Q_0. This kernel corresponds to the Gaussian momentum update with refreshing rate β. For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\begin{aligned} \mathrm{Q}_0(z, \mathrm{d}z') &= \mathrm{N}_x(p'; \sqrt{1-\beta}p, \beta \mathrm{Id})\mathrm{d}p' \updelta_x(\mathrm{d}x') \\ &= (2\pi\beta)^{-d/2} \det(\mathfrak{g}(x))^{-1/2} \exp[-(2\beta)^{-1}\|p' - \sqrt{1-\beta}p\|^2_{\mathfrak{g}(x)^{-1}}]\mathrm{d}p' \updelta_x(\mathrm{d}x') \ . \end{aligned} \tag{25}$$

Kernel $\mathrm{Q}_{c,1}$. This kernel is deterministic and corresponds to the *exact* integration of the Hamiltonian up until time h. For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\mathrm{Q}_{c,1}(z, \mathrm{d}z') = \mathbb{1}_{\bar{\mathsf{E}}_h}(z)\updelta_{\mathrm{T}_h(z)}(\mathrm{d}z') + \mathbb{1}_{\bar{\mathsf{E}}_h^c}(z)\updelta_{s(z)}(\mathrm{d}z') \ . \tag{26}$$

Kernel Q_c. This kernel corresponds to one step of Algorithm 2 (*i.e.*, comprising Steps 1 to 3). For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\mathrm{Q}_c(z, \mathrm{d}z') = \textstyle\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} \mathrm{Q}_0(z, \mathrm{d}z_1)\mathrm{Q}_{c,1}(z_1, \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z') \ . \tag{27}$$

We recall that $\bar{\pi}$, as defined in (2), admits a density with respect to the product Lebesgue measure given for any $z = (x, p) \in \mathrm{T}^\star\mathsf{M}$ by

$$(\mathrm{d}\bar{\pi}/(\mathrm{d}x\mathrm{d}p))(x, p) = (1/Z) \exp[-(1/2)\|p\|^2_{\mathfrak{g}(x)^{-1}}] \det(\mathfrak{g}(x))^{-1/2} \exp[-V(x)] \ .$$

Lemma 17. *Assume* **A**1*. Then, the Markov kernel Q_0, defined in* (25)*, is reversible (up to momentum reversal) with respect to $\bar{\pi}$.*

Proof. Assume **A**1. For any $(z, z') \in \mathrm{T^\star M} \times \mathrm{T^\star M}$, we have

$$\begin{aligned}
\mathrm{Q}_0(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) &= (1/Z)(2\pi)^{-d/2}\beta^{-d/2}\det(\mathfrak{g}(x))^{-1}\exp[-V(x)]\mathrm{d}x\mathrm{d}p\mathrm{d}p'\delta_x(\mathrm{d}x') \\
&\quad \exp[-(2\beta)^{-1}\|p' - \sqrt{1-\beta}p\|^2_{\mathfrak{g}(x)^{-1}}]\exp[-(1/2)\|p\|^2_{\mathfrak{g}(x)^{-1}}] \\
&= (1/Z)(2\pi)^{-d/2}\beta^{-d/2}\det(\mathfrak{g}(x))^{-1}\exp[-V(x)]\mathrm{d}x\mathrm{d}p\mathrm{d}p'\delta_x(\mathrm{d}x') \\
&\quad \exp[-(2\beta)^{-1}\{\|p\|^2_{g(x)^{-1}} - 2\sqrt{1-\beta}\langle p', p\rangle_{g(x)^{-1}} + \|p'\|^2_{g(x)^{-1}}\}] \\
&= (1/Z)(2\pi)^{-d/2}\beta^{-d/2}\det(\mathfrak{g}(x'))^{-1}\exp[-V(x')]\mathrm{d}x'\mathrm{d}p'\mathrm{d}p\delta_x(\mathrm{d}x) \\
&\quad \exp[-(2\beta)^{-1}\{\|p'\|^2_{g(x')^{-1}} - 2\sqrt{1-\beta}\langle p, p'\rangle_{g(x')^{-1}} + \|p\|^2_{g(x')^{-1}}\}] \\
&= \mathrm{Q}_0(z', \mathrm{d}z)\bar{\pi}(\mathrm{d}z') \; ,
\end{aligned}$$

which proves the reversibility of Q_0 with respect to $\bar{\pi}$. Since $s_{\#}\mathrm{Q}_0 = \mathrm{Q}_0$, we conclude the proof with Lemma 6. □

Lemma 18. *Assume* **A**1*,* **A**2*. Then, for any $h > 0$, the Markov kernel $\mathrm{Q}_{c,1}$ with step-size h, defined in (26), is reversible up to momentum reversal with respect to $\bar{\pi}$.*

Proof. Assume **A**1, **A**2. We recall that the set $\bar{\mathsf{E}}_h$ is defined in (24). Let $f \in \mathrm{C}(\mathrm{T^\star M} \times \mathrm{T^\star M}, \mathbb{R})$ with compact support. According to Definition 4, we aim to show that

$$\int_{\mathrm{T^\star M}\times\mathrm{T^\star M}} f(z, z')\mathrm{Q}_{c,1}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) = \int_{\mathrm{T^\star M}\times\mathrm{T^\star M}} f(s(z'), s(z))\mathrm{Q}_{c,1}(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) \; . \tag{28}$$

We denote by I the left integral of (28). We have $I = I_1 + I_2$ where

$$I_1 = \int_{\bar{\mathsf{E}}_h} \bar{\pi}(z) f(z, \mathrm{T}_h(z))\mathrm{d}z \; , \qquad I_2 = \int_{\bar{\mathsf{E}}_h^c} \bar{\pi}(z) f(z, s(z))\mathrm{d}z \; .$$

We denote by J the right integral of (28). By symmetry, we have $J = J_1 + J_2$ where

$$J_1 = \int_{\bar{\mathsf{E}}_h} \bar{\pi}(z) f((s\circ\mathrm{T}_h)(z), s(z))\mathrm{d}z \; , \qquad J_2 = \int_{\bar{\mathsf{E}}_h^c} \bar{\pi}(z) f(z, s(z))\mathrm{d}z \; .$$

We directly have $I_2 = J_2$. Let us now prove that $I_1 = J_1$. By change of variable $z \mapsto (s\circ\mathrm{T}_h)(z)$ in I_1, we have

$$\begin{aligned}
I_1 &= \int_{\bar{\mathsf{E}}_h} \bar{\pi}(z) f(z, \mathrm{T}_h(z))\mathrm{d}z \\
&= \int_{(s\circ\mathrm{T}_h)(\bar{\mathsf{E}}_h)} \bar{\pi}((s\circ\mathrm{T}_h)(z)) f((s\circ\mathrm{T}_h)(z), s(z))\mathrm{d}z && \text{(Lemma 16)} \\
&= \int_{\bar{\mathsf{E}}_h} \bar{\pi}(z) f((s\circ\mathrm{T}_h)(z), s(z))\mathrm{d}z && \text{(Proposition 13-(iv) \& } (s\circ\mathrm{T}_h)(\bar{\mathsf{E}}_h) = \bar{\mathsf{E}}_h) \\
&= J_1 \; .
\end{aligned}$$

Finally, we obtain $I = J$ and thus prove (28) for any continuous function with compact support, which concludes the proof.

□

We are now ready to prove Theorem 15, which states that Q_c is reversible up to momentum reversal with respect to $\bar{\pi}$.

Proof of Theorem 15. Assume **A**1, **A**2. Let $f : \mathrm{T^\star M} \times \mathrm{T^\star M} \to \mathbb{R}$ be a continuous function with compact support. We have

$$\begin{aligned}
&\int_{\mathrm{T^\star M}\times\mathrm{T^\star M}} f(z, z')\mathrm{Q}_c(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\
&= \int_{(\mathrm{T^\star M})^4} f(z, z')\mathrm{Q}_0(z, \mathrm{d}z_1)\mathrm{Q}_{c,1}(z_1, \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(see (27))} \\
&= \int_{(\mathrm{T^\star M})^4} f(s(z_1), z')\mathrm{Q}_0(z, \mathrm{d}z_1)\mathrm{Q}_{c,1}(s(z), \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 17)} \\
&= \int_{(\mathrm{T^\star M})^4} f(s(z_1), z')\mathrm{Q}_0(s(z), \mathrm{d}z_1)\mathrm{Q}_{c,1}(z, \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z) \\
&= \int_{(\mathrm{T^\star M})^4} f(s(z_1), z')\mathrm{Q}_0(z_2, \mathrm{d}z_1)\mathrm{Q}_{c,1}(z, \mathrm{d}z_2)\mathrm{Q}_0(s(z), \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 18)} \\
&= \int_{(\mathrm{T^\star M})^4} f(s(z_1), z')\mathrm{Q}_0(z_2, \mathrm{d}z_1)\mathrm{Q}_{c,1}(s(z), \mathrm{d}z_2)\mathrm{Q}_0(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z) \\
&= \int_{(\mathrm{T^\star M})^4} f(s(z_1), s(z))\mathrm{Q}_0(z_2, \mathrm{d}z_1)\mathrm{Q}_{c,1}(z', \mathrm{d}z_2)\mathrm{Q}_0(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 17)} \\
&= \int_{\mathrm{T^\star M}\times\mathrm{T^\star M}} f(s(z'), s(z))\mathrm{Q}_c(z, \mathrm{d}z')\bar{\pi}(\mathrm{d}z) \; .
\end{aligned}$$

Moreover, $s_{\#}\bar{\pi} = \bar{\pi}$. Hence, by combining Definition 4 and Lemma 5, we obtain the result of Theorem 15. □

F Numerical integration in HMC

In this section, we first recall the definition of symplectic maps, and then discuss the choice of integrators in RMHMC. Finally, we focus on the implicit integrator defined in (6) and provide some notations, which will be notably used in appendix G.

Reminders on symplecticity. We define $J_d = \begin{pmatrix} 0 & \mathrm{I}_d \\ -\mathrm{I}_d & 0 \end{pmatrix} \in \mathbb{R}^{2d \times 2d}$.

Definition 19 (Hairer et al. (2006))**.** *A linear mapping* $A : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ *is called symplectic if* $A^\top J_d A = J_d$.

Definition 20 (Hairer et al. (2006))**.** *A differentiable map* $F : \mathsf{U} \to \mathbb{R}^{2d}$*, where* $\mathsf{U} \subset \mathbb{R}^{2d}$ *is an open set, is called symplectic if the Jacobian matrix* $\mathrm{Jac}[F]$ *is symplectic, i.e., if* $\mathrm{Jac}[F](z)^\top J_d \, \mathrm{Jac}[F](z) = J_d$ *for any* $z \in \mathsf{U}$*. In particular, if* F *is symplectic, then* $|\det \mathrm{Jac}[F]| \equiv 1$.

Choice of the integrators in RMHMC. Introducing a Riemannian metric $\mathfrak{g}$ in the Hamiltonian *a priori* makes it non separable. Thus, designing an integrator which is (i) symplectic, (ii) reversible and (iii) not too computationally heavy is challenging. The *generalized Leapfrog* integrator (GLI), or equivalently the Störmer-Verlet integrator, combined with fixed point iterations, as chosen by Girolami & Calderhead (2011), is considered as the standard scheme for RMHMC. Brofos & Lederman (2021b) analyze the impact of GLI on the ergodicity of RMHMC under a metric which is designed in the same manner as Cobb et al. (2019). In particular, they show that the convergence threshold used to compute the fixed-point process only matters to an extent, and identify a diminishing return to using smaller convergence thresholds. They compare the fixed-point approach with Newton's method, which gives less iterations to converge. On the other hand, the *implicit midpoint* integrator (IMI) enjoys the same theoretical properties as GLI, when these two integrators are combined with fixed-point iterations. As shown by Brofos & Lederman (2021a), IMI may also show numerical advantages in terms of reversibility and volume preservation for specific target distributions. However, this integrator is hard to use for reversibility proofs due to its non-separability. Besides this, Cobb et al. (2019) combine a 2-state augmented Hamiltonian, based on the work of Tao (2016), with Strang splitting (Strang, 1968) to propose a symplectic *explicit* integrator which thus has the advantage not to rely on fixed-point iterations. Yet, this integrator does not satisfy reversibility in theory.

Details on the Störmer-Verlet integrator. The scheme defined in (6) can be written as $\mathrm{G}_h = \overline{\mathrm{G}}_h \circ \underline{\mathrm{G}}_h$ where:

(a) $\underline{\mathrm{G}}_h : \mathrm{T}^\star\mathsf{M} \to 2^{\mathrm{T}^\star\mathsf{M}}$ is the set-valued map *implicitly* defined by the first two aligns of (6), *i.e.*, for any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, $z' \in \underline{\mathrm{G}}_h(z)$ if and only if

$$p' = p - \tfrac{h}{2}\partial_x H_2(x, p'), \qquad x' = x + \tfrac{h}{2}[\partial_p H_2(x, p') + \partial_p H_2(x', p')] \;, \tag{29}$$

We also define the map $\underline{\mathrm{g}}_h : \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M} \to \mathbb{R}^n \times \mathbb{R}^n$ by

$$\underline{\mathrm{g}}_h(z, z') = (x' - x - \tfrac{h}{2}[\partial_p H_2(x, p') + \partial_p H_2(x', p')], p' - p + \tfrac{h}{2}\partial_x H_2(x, p')) \;. \tag{30}$$

such that $\underline{\mathrm{g}}_h(z, z') = 0$ if and only if $z' \in \underline{\mathrm{G}}_h(z)$.

(b) $\overline{\mathrm{G}}_h : \mathrm{T}^\star\mathsf{M} \to \mathrm{T}^\star\mathsf{M}$ is *explicitly* defined by the last align of (6), *i.e.*, for any $z \in \mathrm{T}^\star\mathsf{M}$,

$$\overline{\mathrm{G}}_h(z) = (x, p - \tfrac{h}{2}\partial_x H_2(x, p)) \;. \tag{31}$$

G Proofs of Section 4.1

In this section, we state several results on the maps defined by the implicit integrators of the Hamiltonian (see Section 3.1), in order to prove Proposition 2.

Lemma 21. *Let* $\mathsf{U} \subset \mathbb{R}^d$ *a non-empty open set and* $z \in \mathsf{U}$. *Assume there exist a neighbourhood* $\mathsf{U}_z \subset \mathsf{U}$ *of* z, $\mathtt{L} > 0$ *and a* $\mathtt{L}$*-Lipschitz-continuous map* $\psi : \mathsf{U}_z \to \mathbb{R}^d$ *such that* $\psi \in \mathrm{C}^1(\mathsf{U}_z, \mathbb{R}^d)$. *Then, for any* $h \in (0, 1/\mathtt{L})$ *and any* $a \in \mathbb{R}^d$, *the map* $\psi_h : \mathsf{U}_z \to \mathbb{R}^d$ *defined for any* $z' \in \mathsf{U}_z$ *by* $\psi_h(z') = z' + h\psi(z') + a$ *is a* C^1*-diffeomorphism on a neighbourhood* $\mathsf{U}'_z \subset \mathsf{U}_z$ *of* z.

Proof. Assume we are provided with $\mathsf{U}, z, \mathsf{U}_z, \mathtt{L}$ and ψ as described in Lemma 21. Let $h \in (0, 1/\mathtt{L})$ and $a \in \mathbb{R}^d$. We define the map $\psi_h : \mathsf{U}_z \to \mathbb{R}^d$ by $\psi_h(z') = z' + h\psi(z') + a$, for any $z' \in \mathsf{U}_z$. Since $\psi \in \mathrm{C}^1(\mathsf{U}_z, \mathbb{R}^d)$, it is clear that $\psi_h \in \mathrm{C}^1(\mathsf{U}_z, \mathbb{R}^d)$. We have, for any $z' \in \mathsf{U}_z$, $D\psi_h(z') = \mathrm{Id} + hD\psi(z')$. In particular, for any $w \in \mathbb{R}^d$, it comes that

$$\|D\psi_h(z)w\| \geq \|w\| - h\mathtt{L}\|w\| = (1 - hL)\|w\| \; .$$

Since $1 - h\mathtt{L} > 0$, $\mathrm{Jac}[\psi_h](z)$ is invertible. We conclude the proof by using the local inverse function theorem. □

We recall that the set-valued map G_h, defined in (6), is the *generalised Leapfrog* integrator, or equivalently the *Störmer-Verlet* integrator, of the non-separable Hamiltonian H_2 defined in Section 3.1. We state below a result of local smoothness of this integrator.

Lemma 22. *Let* $(\mathsf{M}, \mathfrak{g})$ *be a a smooth manifold of* $\mathbb{R}^d$ *and* $h > 0$. *Then, for any* $(x^{(0)}, x^{(1)}) \in \mathsf{M} \times \mathsf{M}$, *the vector* $p^{(0)} \in \mathrm{T}^\star_{x^{(0)}}\mathsf{M}$ *such that* $(x^{(1)}, p^{(1)}) \in \mathrm{G}_h(x^{(0)}, p^{(0)})$ *for any* $p^{(1)} \in \mathrm{T}^\star_{x^{(1)}}\mathsf{M}$ *is uniquely determined by* $p^{(0)} = \mathrm{G}_{h,x^{(0)}}(x^{(1)})$, *where the map* $\mathrm{G}_{h,x^{(0)}} : \mathsf{M} \to \mathrm{T}^\star_{x^{(0)}}\mathsf{M}$ *is defined by*

$$\mathrm{G}_{h,x^{(0)}}(y) = p^{(1/2)}(y) - \frac{h}{4} D\mathfrak{g}(x^{(0)})[\mathfrak{g}(x^{(0)})^{-1}p^{(1/2)}(y), \mathfrak{g}(x^{(0)})^{-1}p^{(1/2)}(y)] \; , \tag{32}$$

with $p^{(1/2)}(y) = \frac{2}{h}(\mathfrak{g}(x^{(0)})^{-1} + \mathfrak{g}(y)^{-1})^{-1}(y - x^{(0)})$. *In particular, if* $\mathfrak{g} \in \mathrm{C}^2(\mathsf{M}, \mathbb{R}^{d\times d})$, *then, for any* $x^{(0)} \in \mathsf{M}$, $\mathrm{G}_{h,x^{(0)}} \in \mathrm{C}^1(\mathsf{M}, \mathrm{T}^\star_{x^{(0)}}\mathsf{M})$.

For any $\alpha \geq 1$, *we define* $r^\star_\alpha > 0$ *by*

$$r^\star_\alpha = \min(1/50000, 1/\{1000\alpha(\alpha+1)\}) \; . \tag{33}$$

We recall that we denote by $\mathsf{W}^0(x, r)$ *the open Dikin ellipsoid (w.r.t.* $\mathfrak{g}$*) at* $x \in \mathsf{M}$ *of radius* $r > 0$, *given by* $\mathsf{W}^0(x, r) = \{y \in \mathbb{R}^d \; : \; \|y - x\|_{\mathfrak{g}(x)} < r\}$. *Assume* **A***1*, **A***2*. *Let* $x^{(0)} \in \mathsf{M}$ *and* $r \in (0, r^\star_\alpha]$, *where* α *is given by* **A***2*. *Then, for any* $x \in \mathsf{W}^0(x^{(0)}, r)$, $x \in \mathsf{M}$ *and* $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x)$ *is invertible for any* $h > 0$.

Proof. Let $(\mathsf{M}, \mathfrak{g})$ be a a smooth manifold of $\mathbb{R}^d$ and $h > 0$. We first prove the existence and uniqueness of the map defined in (32). Let $(x^{(0)}, x^{(1)}) \in \mathsf{M} \times \mathsf{M}$. We define for any $y \in \mathsf{M}$

$$p^{(1/2)}(y) = \frac{2}{h}(\mathfrak{g}(x^{(0)})^{-1} + \mathfrak{g}(y)^{-1})^{-1}(y - x^{(0)}) \; , \qquad \tilde{p}(y) = p^{(1/2)}(y) + \frac{h}{2}\partial_x H_2(x^{(0)}, p^{(1/2)}(y)) \; .$$

Note that these expressions are obtained by simply inverting the first two aligns of (6). Then, by definition of G_h, the vector $p \in \mathrm{T}^\star_{x^{(0)}}\mathsf{M}$ such that $(x^{(1)}, p^{(1)}) \in \mathrm{G}_h(x^{(0)}, p)$ for any $p^{(1)} \in \mathrm{T}^\star_{x^{(1)}}\mathsf{M}$ is uniquely determined by $p = \tilde{p}(x^{(1)})$. We thus obtain (32), using the expression of $\partial_x H_2$ given in Section 3.1. Moreover, it is clear that $\mathrm{G}_{h,x^{(0)}}$ is continuously differentiable if $\mathfrak{g} \in \mathrm{C}^2(\mathsf{M}, \mathbb{R}^{d\times d})$.

We are now going to prove the second result of Lemma 22. Assume **A**1, **A**2. Let $x^{(0)} \in \mathsf{M}$ and $r \in (0, r^\star_\alpha]$, where $r^\star_\alpha$ is defined in (33) and α is given by **A**2. In particular $r^\star_\alpha \leq 1/11$. For the sake of clarity, we will now denote $\mathfrak{g}(x^{(0)})$ by $\mathfrak{g}_0$ for the rest of the proof.

We define the open Dikin ellipsoid $\mathsf{B}_0 = \mathsf{W}^0(x^{(0)}, r)$. Since $r < 1$, $\mathsf{B}_0 \subset \mathsf{M}$ by Lemma 9-(b). Let $h > 0$. We now define the following maps on B_0

$$\mathrm{M}_0 = \left\{ \begin{array}{ll} \mathsf{B}_0 & \to S^{++}_d(\mathbb{R}) \\ y & \mapsto \mathfrak{g}_0^{-1} + \mathfrak{g}(y)^{-1} \end{array} \right. , \qquad \mathrm{K}_0 = \left\{ \begin{array}{ll} \mathsf{B}_0 & \to \mathbb{R}^d \\ y & \mapsto \mathfrak{g}_0^{-1}\mathrm{M}_0(y)^{-1}(y - x^{(0)}) \end{array} \right. ,$$

$$\phi_0 = \left\{ \begin{array}{ll} \mathsf{B}_0 & \to \mathbb{R}^d \\ y & \mapsto \mathrm{M}_0(y)D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] \end{array} \right. , \qquad \tilde{\phi}_0 = \left\{ \begin{array}{ll} \mathsf{B}_0 & \to \mathbb{R}^d \\ y & \mapsto \frac{h}{2}\mathrm{M}_0(y)\mathrm{G}_{h,x^{(0)}}(y) \end{array} \right. .$$

Note that we have, for any $y \in \mathsf{B}_0$, $\tilde{\phi}(y) = (y - x^{(0)}) - \frac{1}{2}\phi_0(y)$. Since $\mathfrak{g}$ is twice continuously differentiable on B_0, the maps previously defined are continuously differentiable. We now compute the derivatives of these maps. Let $y \in \mathsf{B}_0$ and $w \in \mathbb{R}^d$. We have

$$\begin{aligned}
D\mathrm{M}_0(y)[w] &= -\mathfrak{g}(y)^{-1} D\mathfrak{g}(y)[w]\mathfrak{g}(y)^{-1} \,, \\
D\mathrm{K}_0(y)[w] &= -\mathfrak{g}_0^{-1}\mathrm{M}_0(y)^{-1} D\mathrm{M}_0(y)[w]\mathrm{M}_0(y)^{-1}(y - x^{(0)}) + \mathfrak{g}_0^{-1}\mathrm{M}_0(y)^{-1} w \\
&= -\mathfrak{g}_0^{-1}\{\mathfrak{g}(y)\mathfrak{g}_0^{-1} + \mathrm{I}_d\}^{-1}\mathfrak{g}(y) D\mathrm{M}_0(y)[w]\mathfrak{g}(y)\{\mathfrak{g}_0^{-1}\mathfrak{g}(y) + \mathrm{I}_d\}^{-1}(y - x^{(0)}) \\
&\qquad + \mathfrak{g}_0^{-1}\mathrm{M}_0(y)^{-1} w \\
&= \mathfrak{g}_0^{-1}\{\mathfrak{g}(y)\mathfrak{g}_0^{-1} + \mathrm{I}_d\}^{-1} D\mathfrak{g}(y)[w]\{\mathfrak{g}_0^{-1}\mathfrak{g}(y) + \mathrm{I}_d\}^{-1}(y - x^{(0)}) + \mathfrak{g}_0^{-1}\mathrm{M}_0(y)^{-1} w \,, \\
D\phi_0(y)[w] &= D\mathrm{M}_0(y)[w] D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] + \mathrm{M}_0(y) D\mathfrak{g}_0[D\mathrm{K}_0(y)[w], \mathrm{K}_0(y)] \\
&\qquad + \mathrm{M}_0(y) D\mathfrak{g}_0[\mathrm{K}_0(y), D\mathrm{K}_0(y)[w]] \,, \\
D\tilde{\phi}_0(y)[w] &= \mathrm{I}_d - \frac{1}{2} D\phi_0(y)[w] \,.
\end{aligned}$$

We have in particular $\tilde{\phi}_0(x^{(0)}) = 0$ and $D\phi_0(x^{(0)}) = 0$. Let us first study the smoothness of $D\phi_0$ on B_0.

Smoothness of $D\phi_0$. We prove here that $D\phi_0$ is $(r_\alpha^\star)^{-1}$-Lipschitz-continuous on B_0 with respect to $\|\cdot\|_{\mathfrak{g}_0}$. Let $(y, y') \in \mathsf{B}_0$ and $w \in \mathbb{R}^d$. We have

$$\begin{aligned}
&\|D\phi_0(y)[w] - D\phi_0(y')[w]\|_{\mathfrak{g}_0} \\
&\quad \le \|D\mathrm{M}_0(y)[w] D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathrm{M}_0(y')[w] D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')]\|_{\mathfrak{g}_0} \\
&\qquad + \|\mathrm{M}_0(y) D\mathfrak{g}_0[D\mathrm{K}_0(y)[w], \mathrm{K}_0(y)] - \mathrm{M}_0(y') D\mathfrak{g}_0[D\mathrm{K}_0(y')[w], \mathrm{K}_0(y')]\|_{\mathfrak{g}_0} \\
&\qquad + \|\mathrm{M}_0(y) D\mathfrak{g}_0[\mathrm{K}_0(y), D\mathrm{K}_0(y)[w]] - \mathrm{M}_0(y') D\mathfrak{g}_0[\mathrm{K}_0(y'), D\mathrm{K}_0(y')[w]]\|_{\mathfrak{g}_0} \,.
\end{aligned} \tag{34}$$

Remark that the two last terms in (34) are very similar and can be bounded in the same way. The rest of the proof on the Lipschitz-continuity of $D\phi_0$ is separated in two steps, each of them consisting of bounding from above a term of (34).

Step 1. Let us first bound the first term in (34). For any $x \in \mathsf{B}_0$, we denote $D\mathrm{M}_0(x)[w] D\mathfrak{g}_0[\mathrm{K}_0(x), \mathrm{K}_0(x)]$ by $a_1(x)$. We have

$$\begin{aligned}
&\mathfrak{g}_0^{1/2}(a_1(y) - a_1(y')) \\
&\quad = \{\mathfrak{g}_0^{1/2}(D\mathrm{M}_0(y)[w] - D\mathrm{M}_0(y')[w])\mathfrak{g}_0^{1/2}\}\{\mathfrak{g}_0^{-1/2} D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\} \\
&\qquad + \{\mathfrak{g}_0^{1/2} D\mathrm{M}_0(y')[w]\mathfrak{g}_0^{1/2}\}\{\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')])\} \,.
\end{aligned} \tag{35}$$

We now aim to bound each one of the four terms that appear in (35).

Step 1.1. First, we have

$$\begin{aligned}
\mathfrak{g}_0^{1/2}(D\mathrm{M}_0(y)[w] - D\mathrm{M}_0(y')[w])\mathfrak{g}_0^{1/2} &= \mathfrak{g}_0^{1/2}(\mathfrak{g}(y')^{-1} D\mathfrak{g}(y')[w]\mathfrak{g}(y')^{-1} - \mathfrak{g}(y)^{-1} D\mathfrak{g}(y)[w]\mathfrak{g}(y)^{-1})\mathfrak{g}_0^{1/2} \\
&= \mathfrak{g}_0^{1/2}\{\mathfrak{g}(y')^{-1} - \mathfrak{g}(y)^{-1}\} D\mathfrak{g}(y')[w]\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2} \\
&\qquad + \mathfrak{g}_0^{1/2}\mathfrak{g}(y)^{-1}\{D\mathfrak{g}(y')[w] - D\mathfrak{g}(y)[w]\}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2} \\
&\qquad + \mathfrak{g}_0^{1/2}\mathfrak{g}(y)^{-1} D\mathfrak{g}(y)[w]\{\mathfrak{g}(y')^{-1} - \mathfrak{g}(y)^{-1}\}\mathfrak{g}_0^{1/2} \,.
\end{aligned}$$

We recall that $r \le r_\alpha^\star \le 1/11$, and thus, we have for any $x \in \{y, y'\}$ the following inequalities

$$\begin{aligned}
\|\mathfrak{g}_0^{1/2}\{\mathfrak{g}(y')^{-1} - \mathfrak{g}(y)^{-1}\}\mathfrak{g}_0^{1/2}\|_2 &\le 3(1-r)^{-3}\|y - y'\|_{\mathfrak{g}_0} \,, && \text{(on the model of (18))} \\
\|\mathfrak{g}_0^{1/2}\mathfrak{g}(x)^{-1}\mathfrak{g}_0^{1/2}\|_2 &\le (1-r)^{-2} \,, && \text{(Lemma 9-(b))} \\
\|\mathfrak{g}_0^{-1/2} D\mathfrak{g}(x)[w]\mathfrak{g}_0^{-1/2}\|_2 &\le (1-r)^{-2}\|\mathfrak{g}(x)^{-1/2} D\mathfrak{g}(x)[w]\mathfrak{g}(x)^{-1/2}\|_2 \,.
\end{aligned}$$

In particular, we have for any $x \in \{y, y'\}$

$$\|\mathfrak{g}(x)^{-1/2} D\mathfrak{g}(x)[w]\mathfrak{g}(x)^{-1/2}\|_2 = \sup_{\{w' \in \mathbb{R}^d : \|w'\|_2 = 1\}} \|D\mathfrak{g}(x)[w, \mathfrak{g}(x)^{-1/2} w']\|_{\mathfrak{g}(x)^{-1}} \tag{36}$$

$$\leq 2\|w\|_{\mathfrak{g}(x)} \sup_{\{w' \in \mathbb{R}^d : \|w'\|_2 = 1\}} \|\mathfrak{g}(x)^{-1/2} w'\|_{\mathfrak{g}(x)} \quad \text{(Lemma 9-(a))}$$

$$\leq 2(1-r)^{-1}\|w\|_{\mathfrak{g}_0} , \quad \text{(Lemma 9-(c))}$$

and using again that $r \leq r_\alpha^\star \leq 1/11$, we have

$$\|\mathfrak{g}(y)^{-1/2}\{D\mathfrak{g}(y')[w] - D\mathfrak{g}(y)[w]\}\mathfrak{g}(y')^{-1/2}\|_2 \tag{37}$$

$$= \sup_{\{w' \in \mathbb{R}^d : \|w'\|_2 = 1\}} \|D\mathfrak{g}(y')[w, \mathfrak{g}(y')^{-1/2} w'] - D\mathfrak{g}(y)[w, \mathfrak{g}(y')^{-1/2} w']\}\|_{\mathfrak{g}(y)^{-1}}$$

$$\leq \alpha(\alpha+1)/3\|w\|_{\mathfrak{g}(y)}((1 - \|y - y'\|_{\mathfrak{g}(y)})^{-3} - 1) \quad \text{(Lemma 11-(c))}$$

$$\leq 5\alpha(\alpha+1)/3(1-r)^{-2}\|w\|_{\mathfrak{g}_0}\|y - y'\|_{\mathfrak{g}_0} . \quad \text{(on the model of (22))}$$

Thus, by using (36) and (37), we have

$$\|\mathfrak{g}_0^{1/2}(D\mathrm{M}_0(y)[w] - D\mathrm{M}_0(y')[w])\mathfrak{g}_0^{1/2}\|_2 \leq (12(1-r)^{-8} + 5\alpha(\alpha+1)/3(1-r)^{-6})\|y - y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} .$$

Step 1.2. Secondly, we have

$$\|\mathfrak{g}_0^{-1/2} D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\|_2 \leq 2\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0}^2 \quad \text{(Lemma 9-(a))}$$

$$\leq 2\|\mathfrak{g}_0^{-1/2}\mathrm{M}_0(y)^{-1}(y - x^{(0)})\|_2^2 \quad \text{(definition of } \mathrm{K}_0\text{)}$$

$$\leq 2\|\mathfrak{g}_0^{-1/2}\mathrm{M}_0(y)^{-1}\mathfrak{g}_0^{-1/2}\|_2^2\|y - x^{(0)}\|_{\mathfrak{g}_0}^2$$

$$\leq 2r^2\|\mathfrak{g}_0^{-1/2}\mathrm{M}_0(y)^{-1}\mathfrak{g}_0^{-1/2}\|_2^2 ,$$

where we have by Lemma 9-(b)

$$\mathfrak{g}_0^{-1/2}\mathrm{M}_0(y)^{-1}\mathfrak{g}_0^{-1/2} = (\mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y)^{-1}\mathfrak{g}_0^{1/2})^{-1} \preceq (1 + (1-r)^2)^{-1}\mathrm{I}_d . \tag{38}$$

Hence, we obtain

$$\|\mathfrak{g}_0^{-1/2} D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\|_2 \leq 2r^2(1 + (1-r)^2)^{-2} . \tag{39}$$

Step 1.3. Thirdly, we have

$$\|\mathfrak{g}_0^{1/2} D\mathrm{M}_0(y')[w]\mathfrak{g}_0^{1/2}\|_2 \leq (1-r)^{-2}\|\mathfrak{g}(x)^{-1/2} D\mathfrak{g}(y')[w]\mathfrak{g}(x)^{-1/2}\|_2 \tag{40}$$

$$\leq 2(1-r)^{-3}\|w\|_{\mathfrak{g}_0} . \quad \text{(using (36))}$$

Step 1.4. Fourthly, we have

$$\|\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')])\|_2 \tag{41}$$

$$\leq 2\|D\mathfrak{g}_0[\mathrm{K}_0(y) - \mathrm{K}_0(y'), \mathrm{K}_0(y)]\|_{\mathfrak{g}_0^{-1}} + \|D\mathfrak{g}_0[\mathrm{K}_0(y) - \mathrm{K}_0(y'), \mathrm{K}_0(y) - \mathrm{K}_0(y')]\|_{\mathfrak{g}_0^{-1}}$$

$$\leq 2\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0} + \|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0} \quad \text{(Lemma 9-(a))}$$

$$\leq 4\|\mathrm{M}_0(y)^{-1}(y - x^{(0)}) - \mathrm{M}_0(y')^{-1}(y' - x^{(0)})\|_{\mathfrak{g}_0^{-1}}\|\mathrm{M}_0(y)^{-1}(y - x^{(0)})\|_{\mathfrak{g}_0^{-1}}$$

$$+ 2\|\mathrm{M}_0(y)^{-1}(y - x^{(0)}) - \mathrm{M}_0(y')^{-1}(y' - x^{(0)})\|_{\mathfrak{g}_0^{-1}}^2 .$$

In particular, $\mathrm{M}_0(y)^{-1}(y - x^{(0)}) - \mathrm{M}_0(y')^{-1}(y' - x^{(0)}) = \mathrm{M}_0(y)^{-1}(y - y') + (\mathrm{M}_0(y)^{-1} - \mathrm{M}_0(y')^{-1})(y' - x^{(0)})$, and thus we have

$$\|\mathrm{M}_0(y)^{-1}(y - x^{(0)}) - \mathrm{M}_0(y')^{-1}(y' - x^{(0)})\|_{\mathfrak{g}_0^{-1}} \tag{42}$$

$$\leq \|\mathfrak{g}_0^{-1/2}\mathrm{M}_0(y)^{-1}\mathfrak{g}_0^{-1/2}\|_2\|y - y'\|_{\mathfrak{g}_0} + r\|\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y)^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2}\|_2$$

$$\leq (1 + (1-r)^2)^{-1}\|y - y'\|_{\mathfrak{g}_0} + r\|\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y')^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2}\|_2 . \quad \text{(using (38))}$$

We now aim to find an upper bound for $\|\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y)^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2}\|_2$. We have

$$\begin{aligned}
&\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y)^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2} \\
&\quad = (\mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y)^{-1}\mathfrak{g}_0^{1/2})^{-1} - (\mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2})^{-1} \\
&\quad = (\mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2} + \mathfrak{g}_0^{1/2}\{\mathfrak{g}(y)^{-1} - \mathfrak{g}(y')^{-1}\}\mathfrak{g}_0^{1/2})^{-1} - (\mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2})^{-1} \\
&\quad = \mathrm{B}(y')^{-1/2}[(\mathrm{I}_d + \mathrm{B}(y')^{-1/2}\mathfrak{g}_0^{1/2}\{\mathfrak{g}(y)^{-1} - \mathfrak{g}(y')^{-1}\}\mathfrak{g}_0^{1/2}\mathrm{B}(y')^{-1/2})^{-1} - \mathrm{I}_d]\mathrm{B}(y')^{-1/2} ,
\end{aligned}$$

where $\mathrm{B}(y') = \mathrm{I}_d + \mathfrak{g}_0^{1/2}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2}$. In particular, we have

$$(1 + (1-r)^{-2})^{-1/2}\mathrm{I}_d \preceq \mathrm{B}(y')^{-1/2} \preceq (1 + (1-r)^2)^{-1/2}\mathrm{I}_d ,$$

by Lemma 9-(b). Note that (17) still holds, where x, x' and $\bar{r}$ are respectively replaced by y, y' and r. Thus, on the model on (18), we have

$$\{(1 - \|y'-y\|_{\mathfrak{g}(y)})^2 - 1\}(1-r)^2\mathrm{I}_d \preceq \mathfrak{g}_0^{1/2}\{\mathfrak{g}(y)^{-1} - \mathfrak{g}(y')^{-1}\}\mathfrak{g}_0^{1/2} \preceq \{(1 - \|y'-y\|_{\mathfrak{g}(y)})^{-2} - 1\}(1-r)^{-2}\mathrm{I}_d ,$$

and then

$$\begin{aligned}
&(1 + \{(1 - \|y'-y\|_{\mathfrak{g}(y)})^2 - 1\}(1-r)^2(1 + (1-r)^{-2})^{-1})\mathrm{I}_d \\
&\qquad \preceq \mathrm{I}_d + \mathrm{B}(y')^{-1/2}\mathfrak{g}_0^{1/2}\{\mathfrak{g}(y)^{-1} - \mathfrak{g}(y')^{-1}\}\mathfrak{g}_0^{1/2}\mathrm{B}(y')^{-1/2} \\
&\qquad\qquad \preceq (1 + \{(1 - \|y'-y\|_{\mathfrak{g}(y)})^{-2} - 1\}(1-r)^{-2}(1 + (1-r)^2)^{-1})\mathrm{I}_d .
\end{aligned}$$

Therefore, we have $\|\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y')^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2}\|_2 \leq (1 + (1-r)^2)^{-1}\max(|f_1(r)|, f_2(r))$ where

$$\begin{aligned}
f_1(r) &= (1 + \{(1 - \|y'-y\|_{\mathfrak{g}(y)})^{-2} - 1\}(1-r)^{-2}(1 + (1-r)^2)^{-1})^{-1} - 1, \\
f_2(r) &= (1 + \{(1 - \|y'-y\|_{\mathfrak{g}(y)})^{2} - 1\}(1-r)^{2}(1 + (1-r)^{-2})^{-1})^{-1} - 1 .
\end{aligned}$$

We now aim to control the upper bound $\max(|f_1(r)|, f_2(r))$.

(a) We first bound $|f_1(r)|$. Since $\|y'-y\|_{\mathfrak{g}(y)} \geq 0$, it is clear that $f_1(r) \leq 0$. Using Inequality (a) with $u = \|y'-y\|_{\mathfrak{g}(y)}$, where $u \leq 1/5$ by (17), we obtain

$$\begin{aligned}
|f_1(r)| &= -f_1(r) \\
&\leq 1 - (1 + 3(1-r)^{-2}(1 + (1-r)^2)^{-1}\|y'-y\|_{\mathfrak{g}(y)})^{-1} \\
&\leq 3(1-r)^{-2}(1 + (1-r)^2)^{-1}\|y'-y\|_{\mathfrak{g}(y)} && \text{(Inequality (d))} \\
&\leq 3(1-r)^{-3}(1 + (1-r)^2)^{-1}\|y'-y\|_{\mathfrak{g}_0} . && \text{(Lemma 9-(c))}
\end{aligned}$$

(b) We now bound $f_2(r)$. We have

$$\begin{aligned}
f_2(r) &\leq (1 - 2(1-r)^2(1 + (1-r)^{-2})^{-1}\|y'-y\|_{\mathfrak{g}(y)})^{-1} - 1 \\
&\leq 4(1-r)^2(1 + (1-r)^{-2})^{-1}\|y'-y\|_{\mathfrak{g}(y)} \\
&\leq 4(1-r)(1 + (1-r)^{-2})^{-1}\|y'-y\|_{\mathfrak{g}_0} , && \text{(Lemma 9-(c))}
\end{aligned}$$

where we used (i) Inequality (d) in the first line with $u = \|y'-y\|_{\mathfrak{g}(y)}$ and (ii) Inequality (b) in the second line with $u = 2(1-r)^2(1 + (1-r)^{-2})^{-1}\|y'-y\|_{\mathfrak{g}(y)} \leq 4r(1-r)(1 + (1-r)^{-2})^{-1} \leq 1/2$ with $r \leq 1/11$.

We recall that $r \in (0, 1)$, and thus, we have respectively $(1-r)^{-3} \geq (1-r)$ and $(1 + (1-r)^2)^{-1} \geq (1 + (1-r)^{-2})^{-1}$. Therefore, $\max(|f_1(r)|, f_2(r)) \leq (4/3)|f_1(r)|$ and

$$\|\mathfrak{g}_0^{-1/2}(\mathrm{M}_0(y')^{-1} - \mathrm{M}_0(y')^{-1})\mathfrak{g}_0^{-1/2}\|_2 \leq 4(1-r)^{-3}(1 + (1-r)^2)^{-2}\|y'-y\|_{\mathfrak{g}_0} , \tag{43}$$

$$\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0} \leq (1 + (1-r)^2)^{-1}\{1 + 4r(1-r)^{-3}(1 + (1-r)^2)^{-1}\}\|y'-y\|_{\mathfrak{g}_0} .$$

By combining (41), (42) and (43), we finally have

$$
\begin{aligned}
&\|\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[\mathrm{K}_0(y),\mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'),\mathrm{K}_0(y')])\|_2 \\
&\quad \le 4r(1+(1-r)^2)^{-2}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}\|y-y'\|_{\mathfrak{g}_0} \\
&\qquad + 4r(1+(1-r)^2)^{-2}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}^2\|y-y'\|_{\mathfrak{g}_0} \\
&\quad \le 4r(1+(1-r)^2)^{-2}\{2+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}^2\|y-y'\|_{\mathfrak{g}_0}\ .
\end{aligned}
$$

Conclusion of Step 1. By combining the results of Steps 1.1 to 1.4, it comes that

$$
\begin{aligned}
&\|D\mathrm{M}_0(y)[w]D\mathfrak{g}_0[\mathrm{K}_0(y),\mathrm{K}_0(y)] - D\mathrm{M}_0(y')[w]D\mathfrak{g}_0[\mathrm{K}_0(y'),\mathrm{K}_0(y')]\|_{\mathfrak{g}_0} \\
&\quad = \|\mathfrak{g}_0^{1/2}(a_1(y)-a_1(y'))\|_2 \\
&\quad \le \{2r^2(1+(1-r)^2)^{-2}(12(1-r)^{-8}+5\alpha(\alpha+1)/3(1-r)^{-6}) \\
&\qquad + 8r(1-r)^{-3}(1+(1-r)^2)^{-2}\{2+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}^2\}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0}\ .
\end{aligned}
$$

Step 2. Let us now bound the second term in (34). For any $x \in \mathsf{B}_0$, we denote $\mathrm{M}_0(y)D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)]$ by $a_2(x)$. We have

$$
\begin{aligned}
\mathfrak{g}_0^{1/2}(a_2(y)-a_2(y')) &= \{\mathfrak{g}_0^{1/2}(\mathrm{M}_0(y)-\mathrm{M}_0(y'))\mathfrak{g}_0^{1/2}\}\{\mathfrak{g}_0^{-1/2}D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)]\} \\
&\quad + \{\mathfrak{g}_0^{1/2}\mathrm{M}_0(y')\mathfrak{g}_0^{1/2}\}\{\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)] - D\mathfrak{g}_0[D\mathrm{K}_0(y')[w],\mathrm{K}_0(y')])\}\ .
\end{aligned} \tag{44}
$$

We now aim to bound each of the four terms that appear in (44).

Step 2.1. First, we have

$$
\begin{aligned}
\|\mathfrak{g}_0^{1/2}(\mathrm{M}_0(y)-\mathrm{M}_0(y'))\mathfrak{g}_0^{1/2}\|_2 &= \|\mathfrak{g}_0^{1/2}(\mathfrak{g}(y)^{-1}-\mathfrak{g}(y')^{-1})\mathfrak{g}_0^{1/2}\|_2 \\
&\le 3(1-r)^{-3}\|y-y'\|_{\mathfrak{g}_0}\ . \qquad \text{(on the model of (18))}
\end{aligned}
$$

Step 2.2. Secondly, we have

$$
\begin{aligned}
\|\mathfrak{g}_0^{-1/2}D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)]\|_2 &\le 2\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0}\|\mathrm{M}_0(y)^{-1}(y-x^{(0)})\|_{\mathfrak{g}_0^{-1}} \qquad \text{(Lemma 9-(a))} \\
&\le 2r(1+(1-r)^2)^{-1}\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0}\ , \qquad \text{(see Step 1.2.)}
\end{aligned}
$$

where

$$
\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0} \le r\|\mathfrak{g}_0^{-1/2}(\mathfrak{g}(y)\mathfrak{g}_0^{-1}+\mathrm{I}_d)^{-1}\mathfrak{g}_0^{1/2}\|_2^2\|\mathfrak{g}_0^{-1/2}\mathfrak{g}(y)^{1/2}\|_2^2\|\mathfrak{g}(y)^{-1/2}D\mathfrak{g}(y)[w]\mathfrak{g}(y)^{-1/2}\|_2 + \|\mathrm{M}_0(y)^{-1}w\|_{\mathfrak{g}_0^{-1}}\ .
$$

Using (36) and (38), it comes that

$$
\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0} \le \{2r(1-r)^{-3}(1+(1-r)^2)^{-2}+(1+(1-r)^2)^{-1}\}\|w\|_{\mathfrak{g}_0}\ .
$$

Then, we have

$$
\|\mathfrak{g}_0^{-1/2}D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)]\|_2 \le 2r(1+(1-r)^2)^{-2}\{1+2r(1-r)^{-3}(1+(1-r)^2)^{-1}\}\|w\|_{\mathfrak{g}_0}\ .
$$

Step 2.3. Thirdly, we have $\|\mathfrak{g}_0^{1/2}\mathrm{M}_0(y')\mathfrak{g}_0^{1/2}\|_2 = \|\mathfrak{g}_0^{1/2}\mathfrak{g}(y')^{-1}\mathfrak{g}_0^{1/2}+\mathrm{I}_d\|_2 \le 1+(1-r)^{-2}$.

Step 2.4. Fourthly, using Lemma 9-(a), we have

$$
\begin{aligned}
&\|\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)] - D\mathfrak{g}_0[D\mathrm{K}_0(y')[w],\mathrm{K}_0(y')])\|_2 \\
&\quad \le \|D\mathfrak{g}_0[D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w],\mathrm{K}_0(y)]\|_{\mathfrak{g}_0^{-1}} + \|D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)-\mathrm{K}_0(y')]\|_{\mathfrak{g}_0^{-1}} \\
&\qquad + \|D\mathfrak{g}_0[D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w],\mathrm{K}_0(y)-\mathrm{K}_0(y')]\|_{\mathfrak{g}_0^{-1}} \\
&\quad \le 2\|D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w]\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0} + 2\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)-\mathrm{K}_0(y')\|_{\mathfrak{g}_0} \\
&\qquad + 2\|D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w]\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)-\mathrm{K}_0(y')\|_{\mathfrak{g}_0} \\
&\quad \le 2\{\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0}+\|\mathrm{K}_0(y)-\mathrm{K}_0(y')\|_{\mathfrak{g}_0}\}\|D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w]\|_{\mathfrak{g}_0} \\
&\qquad + 2\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)-\mathrm{K}_0(y')\|_{\mathfrak{g}_0}\ .
\end{aligned}
$$

We recall that the following inequalities hold

$$\|\mathrm{K}_0(y)]\|_{\mathfrak{g}_0} \le r(1+(1-r)^2)^{-1} \ ,$$

$$\|D\mathrm{K}_0(y)[w]\|_{\mathfrak{g}_0} \le (1+(1-r)^2)^{-1}\{1+2r(1-r)^{-3}(1+(1-r)^2)^{-1}\}\|w\|_{\mathfrak{g}_0} \ ,$$

$$\|\mathrm{K}_0(y)-\mathrm{K}_0(y')\|_{\mathfrak{g}_0} \le (1+(1-r)^2)^{-1}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}\|y'-y\|_{\mathfrak{g}_0} \ .$$

We are now going to bound $\|D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w]\|_{\mathfrak{g}_0}$. We have

$$\|D\mathrm{K}_0(y)[w]-D\mathrm{K}_0(y')[w]\|_{\mathfrak{g}_0} \le \|(\mathrm{M}_0(y)^{-1}-\mathrm{M}_0(y')^{-1})w\|_{\mathfrak{g}_0^{-1}} + A_1 + A_2 + A_3 + A_4 \ ,$$

where

$$A_1 = \|\{(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y)\mathfrak{g}_0^{-1/2})^{-1} - (\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\}$$
$$\times \mathfrak{g}_0^{-1/2}D\mathfrak{g}(y)[w]\mathfrak{g}_0^{-1/2}(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y)\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{1/2}(y-x^{(0)})\|_2 \ ,$$

$$A_2 = \|(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{-1/2}\{D\mathfrak{g}(y)[w]-D\mathfrak{g}(y')[w]\}\mathfrak{g}_0^{-1/2}(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y)\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{1/2}(y-x^{(0)})\|_2 \ ,$$

$$A_3 = \|(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{-1/2}D\mathfrak{g}(y)[w]\mathfrak{g}_0^{-1/2}$$
$$\times \{(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y)\mathfrak{g}_0^{-1/2})^{-1} - (\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\}\mathfrak{g}_0^{1/2}(y-x^{(0)})\|_2 \ ,$$

$$A_3 = \|(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{-1/2}D\mathfrak{g}(y)[w]\mathfrak{g}_0^{-1/2}(\mathrm{I}_d+\mathfrak{g}_0^{-1/2}\mathfrak{g}(y')\mathfrak{g}_0^{-1/2})^{-1}\mathfrak{g}_0^{1/2}(y-y')\|_2 \ .$$

In particular, we have

$$\|(\mathrm{M}_0(y')^{-1}-\mathrm{M}_0(y')^{-1})w\|_{\mathfrak{g}_0^{-1}} \le 4(1-r)^{-3}(1+(1-r)^2)^{-2}\|y'-y\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ , \qquad ((43))$$

$$\max(A_1,A_3) \le 8r(1-r)^{-6}(1+(1-r)^2)^{-3}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ , \qquad ((36), (38), (43))$$

$$A_2 \le 5\alpha(\alpha+1)/3r(1-r)^{-4}(1+(1-r)^2)^{-2}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ , \qquad ((37), (38))$$

$$A_4 \le 2(1-r)^{-3}(1+(1-r)^2)^{-2}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ . \qquad ((36) \text{ and } (38))$$

Therefore, we have

$$\begin{aligned}
&\|\mathfrak{g}_0^{-1/2}(D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)] - D\mathfrak{g}_0[D\mathrm{K}_0(y')[w],\mathrm{K}_0(y')])\|_2 \\
&\quad \le 2r(1+(1-r)^2)^{-1}\{3+4r(1-r)^{-3}(1+(1-r)^2)^{-1})\} \\
&\qquad \times \{4(1-r)^{-3}(1+(1-r)^2)^{-2} + 16r(1-r)^{-6}(1+(1-r)^2)^{-3} \\
&\qquad\quad + 5\alpha(\alpha+1)/3r(1-r)^{-4}(1+(1-r)^2)^{-2} + 2(1-r)^{-3}(1+(1-r)^2)^{-2}\}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \\
&\qquad + 2\{1+2r(1-r)^{-3}(1+(1-r)^2)^{-1}\} \\
&\qquad\quad \times (1+(1-r)^2)^{-2}(1+4r(1-r)^{-3}(1+(1-r)^2)^{-1})\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ .
\end{aligned}$$

Conclusion of Step 2. By combining the results of Steps 2.1 to 2.4, it comes that

$$\begin{aligned}
&\|\mathrm{M}_0(y)D\mathfrak{g}_0[D\mathrm{K}_0(y)[w],\mathrm{K}_0(y)] - \mathrm{M}_0(y')D\mathfrak{g}_0[D\mathrm{K}_0(y')[w],\mathrm{K}_0(y')]\|_{\mathfrak{g}_0} \\
&\quad = \|\mathfrak{g}_0^{1/2}(a_2(y)-a_2(y'))\|_2 \\
&\quad \le \{6r(1-r)^{-3}(1+(1-r)^2)^{-2}[1+2r(1-r)^{-3}(1+(1-r)^2)^{-1}] \\
&\qquad + 2r[1+(1-r)^{-2}](1+(1-r)^2)^{-1}\{3+4r(1-r)^{-3}(1+(1-r)^2)^{-1})\} \\
&\qquad\quad \times \{4(1-r)^{-3}(1+(1-r)^2)^{-2} + 16r(1-r)^{-6}(1+(1-r)^2)^{-3} \\
&\qquad\qquad + 5\alpha(\alpha+1)/3r(1-r)^{-4}(1+(1-r)^2)^{-2} + 2(1-r)^{-3}(1+(1-r)^2)^{-2}\} \\
&\qquad + 2[1+(1-r)^{-2}]\{1+2r(1-r)^{-3}(1+(1-r)^2)^{-1}\} \\
&\qquad\quad \times (1+(1-r)^2)^{-2}(1+4r(1-r)^{-3}(1+(1-r)^2)^{-1})\}\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0} \ .
\end{aligned}$$

Conclusion. Finally, we have $\|D\phi_0(y)[w]-D\phi_0(y')[w]\|_{\mathfrak{g}_0} \le c(r)\|y-y'\|_{\mathfrak{g}_0}\|w\|_{\mathfrak{g}_0}$, where

$$\begin{aligned}
c(r) &\le 24r^2(1-r)^{-12} + 10\alpha(\alpha+1)/3r^2(1-r)^{-10} \\
&\quad + 32r(1-r)^{-7} + 128r^2(1-r)^{-12} + 128r^3(1-r)^{-17} \\
&\quad + 12r(1-r)^{-7} + 24r^2(1-r)^{-12} \\
&\quad + \{48r(1-r)^{-9} + 192r^2(1-r)^{-14} + 20\alpha(\alpha+1)r^2(1-r)^{-10} + 24r(1-r)^{-9}\}\{1+(1-r)^{-2}\} \\
&\quad + \{64r^2(1-r)^{-14} + 256r^3(1-r)^{-19} + 80\alpha(\alpha+1)/3r^3(1-r)^{-15} + 32r^2(1-r)^{-14}\}\{1+(1-r)^{-2}\} \\
&\quad + \{4(1-r)^{-4} + 24r(1-r)^{-9} + 32r^2(1-r)^{-14}\}\{1+(1-r)^{-2}\} \ ,
\end{aligned}$$

by combining the results of Step 1 (two first lines) and Step 2 (following lines). Hence, using that $r \leq 1/11$, we have

$$\begin{aligned} c(r) &\leq \max(24 \times 256, 5 \times 80\alpha(\alpha+1)/3)(1-r)^{-21} \\ &\leq \max(6144, 400\alpha(\alpha+1)/3)15/2 \\ &\leq \max(46080, 1000\alpha(\alpha+1)) \\ &\leq 1/r_\alpha^\star \,, && \text{(see (33))} \end{aligned}$$

i.e., for any $(y, y') \in \mathsf{B}_0 \times \mathsf{B}_0$ and $w \in \mathbb{R}^d$, we have

$$\|D\phi_0(y)[w] - D\phi_0(y)[w]\|_{\mathfrak{g}_0} \leq (r_\alpha^\star)^{-1} \|y - y'\|_{\mathfrak{g}_0} \|w\|_{\mathfrak{g}_0} \,,$$

and thus

$$\|D\phi_0(y) - D\phi_0(y)\|_{\mathfrak{g}_0} \leq (r_\alpha^\star)^{-1} \|y - y'\|_{\mathfrak{g}_0} \,.$$

This last inequality finally proves that $D\phi_0$ is $(r_\alpha^\star)^{-1}$-Lipschitz-continuous on B_0 with respect to $\| \cdot \|_{\mathfrak{g}_0}$.

Inequality on $D\tilde{\phi}_0$. Elaborating on the smoothness of $D\phi_0$, we prove here that $\|D\tilde{\phi}_0(y)\|_{\mathfrak{g}_0} > 1/2$ for any $y \in \mathsf{B}_0$. Let $y \in \mathsf{B}_0$, we have

$$\begin{aligned} \|\mathrm{I}_d\|_{\mathfrak{g}_0} &= \|D\tilde{\phi}_0(y) + 1/2 D\phi_0(y)\|_{\mathfrak{g}_0} \,, \\ 1 &\leq \|D\tilde{\phi}_0(y)\|_{\mathfrak{g}_0} + 1/2 \|D\phi_0(y)\|_{\mathfrak{g}_0} \,, \\ 1 - 1/2\|D\phi_0(y)\|_{\mathfrak{g}_0} &\leq \|D\tilde{\phi}_0(y)\|_{\mathfrak{g}_0} \,. \end{aligned}$$

We recall that $D\phi_0(x^{(0)}) = 0$ and that $D\phi_0$ is $(r_\alpha^\star)^{-1}$-Lipschitz-continuous on B_0. Thus, we obtain

$$\|D\phi_0(y)\|_{\mathfrak{g}_0} = \|D\phi_0(y) - D\phi_0(x^{(0)})\|_{\mathfrak{g}_0} \leq (r_\alpha^\star)^{-1} \|y - x^{(0)}\|_{\mathfrak{g}_0} < (r_\alpha^\star)^{-1} r \leq 1 \,,$$

and then $\|D\tilde{\phi}_0(y)\|_{\mathfrak{g}_0} > 1/2$, which proves the result.

Smoothness of ϕ_0 and $\tilde{\phi}_0$. We prove here that ϕ_0 (and thus $\tilde{\phi}_0$) is Lipschitz-continuous on B_0 with respect to $\| \cdot \|_{\mathfrak{g}_0}$. Let $(y, y') \in \mathsf{B}_0 \times \mathsf{B}_0$. We have

$$\begin{aligned} \|\phi_0(y) - \phi_0(y')\|_{\mathfrak{g}_0} \leq \|\{\mathrm{M}_0(y) - \mathrm{M}_0(y')\} D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\|_{\mathfrak{g}_0} \\ + \|\mathrm{M}_0(y')\{D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')]\}\|_{\mathfrak{g}_0} \,. \end{aligned} \tag{45}$$

To prove the Lipschitz-continuity of ϕ_0, we are going to proceed in two steps, each of them consisting of bounding from above a term of (45).

<u>*Step 1.*</u> Let us first bound the first term in (45). Since $r \leq 1/11$, we have

$$\begin{aligned} &\|\{\mathrm{M}_0(y) - \mathrm{M}_0(y')\} D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\|_{\mathfrak{g}_0} \\ &\quad \leq \|\mathfrak{g}_0^{1/2}\{\mathrm{M}_0(y) - \mathrm{M}_0(y')\}\mathfrak{g}_0^{1/2}\|_2 \|D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)]\|_{\mathfrak{g}_0^{-1}} \\ &\quad \leq 3(1-r)^{-3}\|y - y'\|_{\mathfrak{g}_0} \times 2\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0}^2 && \text{(see (18) and Lemma 9-(a))} \\ &\quad \leq 6r^2(1-r)^{-3}(1 + (1-r)^2)^{-2}\|y - y'\|_{\mathfrak{g}_0} \,. && \text{(see (39))} \end{aligned}$$

<u>*Step 2.*</u> Let us now bound the second term in (45). We have

$$\begin{aligned} &\|\mathrm{M}_0(y')\{D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')]\}\|_{\mathfrak{g}_0} \\ &\quad \leq \|\mathfrak{g}_0^{1/2}\mathrm{M}_0(y')\mathfrak{g}_0^{1/2}\|_2 \|D\mathfrak{g}_0[\mathrm{K}_0(y), \mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'), \mathrm{K}_0(y')]\|_{\mathfrak{g}_0^{-1}} \\ &\quad \leq \{1 + (1-r)^{-2}\}\{2\|D\mathfrak{g}_0[\mathrm{K}_0(y) - \mathrm{K}_0(y'), \mathrm{K}_0(y)]\|_{\mathfrak{g}_0^{-1}} + \|D\mathfrak{g}_0[\mathrm{K}_0(y) - \mathrm{K}_0(y'), \mathrm{K}_0(y) - \mathrm{K}_0(y')]\|_{\mathfrak{g}_0^{-1}}\} \\ &\quad \leq \{1 + (1-r)^{-2}\}\{4\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0} + 2\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}^2\} \\ &\quad \leq 1\{1 + (1-r)^{-2}\}\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}\{2\|\mathrm{K}_0(y)\|_{\mathfrak{g}_0} + \|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0}\} \,, \end{aligned}$$

where we recall that the following inequalities hold

$$\|\mathrm{K}_0(y)]\|_{\mathfrak{g}_0} \leq r(1+(1-r)^2)^{-1} ,$$
$$\|\mathrm{K}_0(y) - \mathrm{K}_0(y')\|_{\mathfrak{g}_0} \leq (1+(1-r)^2)^{-1}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}\|y'-y\|_{\mathfrak{g}_0} .$$

Thus, we have

$$\begin{aligned}&\|\mathrm{M}_0(y')\{D\mathfrak{g}_0[\mathrm{K}_0(y),\mathrm{K}_0(y)] - D\mathfrak{g}_0[\mathrm{K}_0(y'),\mathrm{K}_0(y')]\}\|_{\mathfrak{g}_0}\\ &\quad \leq 4r\{1+(1-r)^2\}^{-2}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}^2\|y-y'\|_{\mathfrak{g}_0} .\end{aligned}$$

<u>*Conclusion.*</u> Since $r \leq 1/11$, we finally have

$$\begin{aligned}\|\phi_0(y)-\phi_0(y')\|_{\mathfrak{g}_0} &\leq \{6r^2(1-r)^{-3}(1+(1-r)^2)^{-2}\\ &\qquad 4r\{1+(1-r)^2\}^{-2}\{1+4r(1-r)^{-3}(1+(1-r)^2)^{-1}\}^2\}\|y-y'\|_{\mathfrak{g}_0}\\ &\leq \{6r^2(1-r)^{-7}+4r(1-r)^{-4}+32r^2(1-r)^{-9}+64r^3(1-r)^{-14}\}\|y-y'\|_{\mathfrak{g}_0}\\ &\leq 4\times 64r(1-r)^{-14}\|y-y'\|_{\mathfrak{g}_0}\\ &\leq 4864/5r\|y-y'\|_{\mathfrak{g}_0} .\end{aligned}$$

Therefore, ϕ_0 and $\tilde{\phi}_0$ are respectively $(4864/5r)$ and $(1+2432/5r)$-Lipschitz-continuous on B_0 with respect to $\|\cdot\|_{\mathfrak{g}_0}$.

We are now ready to prove the second result of Lemma 22.

Invertibility of $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}]$. Let $x \in \mathsf{B}_0$ and $w \in \mathbb{R}^d$. We have $(h/2)\mathrm{G}_{h,x^{(0)}} = \mathrm{M}_0(y)^{-1}\tilde{\phi}_0(y)$ and thus,

$$\begin{aligned}(h/2)D\mathrm{G}_{h,x^{(0)}}(x)[w] &= \mathrm{M}_0(x)^{-1}D\tilde{\phi}_0(x)[w] + D\mathrm{M}_0(x)^{-1}[w]\tilde{\phi}_0(x) ,\\ D\tilde{\phi}_0(x)[w] &= (h/2)\mathrm{M}_0(x)D\mathrm{G}_{h,x^{(0)}}(x)[w] - \mathrm{M}_0(x)D\mathrm{M}_0(x)^{-1}[w]\tilde{\phi}_0(x) .\end{aligned}$$

Then, we have

$$\begin{aligned}\|D\tilde{\phi}_0(x)[w]\|_{\mathfrak{g}_0} &\leq (h/2)\|\mathfrak{g}_0^{1/2}\mathrm{M}_0(x)\mathfrak{g}_0^{1/2}\|_2\|D\mathrm{G}_{h,x^{(0)}}(x)[w]\|_{\mathfrak{g}_0^{-1}} \\ &\quad + \|\mathfrak{g}_0^{1/2}\mathrm{M}_0(x)\mathfrak{g}_0^{1/2}\|_2\|\mathfrak{g}_0^{-1/2}D\mathrm{M}_0(x)^{-1}[w]\mathfrak{g}_0^{-1/2}\|_2\|\tilde{\phi}_0(x)\|_{\mathfrak{g}_0}\\ &\leq \{1+(1-r)^{-2}\}\{(h/2)\|D\mathrm{G}_{h,x^{(0)}}(x)[w]\|_{\mathfrak{g}_0^{-1}} + \|\mathfrak{g}_0^{-1/2}D\mathrm{M}_0(x)^{-1}[w]\mathfrak{g}_0^{-1/2}\|_2\|\tilde{\phi}_0(x)\|_{\mathfrak{g}_0}\} ,\end{aligned} \tag{46}$$

where $D\mathrm{M}_0(x)^{-1}[w] = -\mathrm{M}_0(x)^{-1}D\mathrm{M}_0(x)[w]\mathrm{M}_0(x)^{-1}$ and

$$\begin{aligned}\|\mathfrak{g}_0^{-1/2}D\mathrm{M}_0(x)^{-1}[w]\mathfrak{g}_0^{-1/2}\|_2 &\leq \|\mathfrak{g}_0^{-1/2}\mathrm{M}_0(x)^{-1}\mathfrak{g}_0^{-1/2}\|_2^2\|\mathfrak{g}_0^{-1/2}D\mathrm{M}_0(x)[w]\mathfrak{g}_0^{-1/2}\|_2 \\ &\leq 2(1-r)^{-3}(1+(1-r)^2)^{-1}\|w\|_{\mathfrak{g}_0} \qquad \text{(see (40))}\\ &\leq 2(1-r)^{-7}\|w\|_{\mathfrak{g}_0} .\end{aligned} \tag{47}$$

Moreover, using the Lipschitz-continuity of $\tilde{\phi}_0$ on B_0 and $\tilde{\phi}_0(x^{(0)}) = 0$, we have

$$\|\tilde{\phi}_0(x)\|_{\mathfrak{g}_0} \leq \|\tilde{\phi}_0(x)-\tilde{\phi}_0(x^{(0)})\|_{\mathfrak{g}_0} + \|\tilde{\phi}_0(x^{(0)})\|_{\mathfrak{g}_0} \leq (1+2432/5r)\|x-x^{(0)}\|_{\mathfrak{g}_0} < (1+2432/5r)r . \tag{48}$$

Since $r \leq 1/11$, we have $1+(1-r)^{-2} \leq 5/2$, and by combining (46),(47) and(48), we obtain

$$\|D\tilde{\phi}_0(x)[w]\| \leq (5h/4)\|D\mathrm{G}_{h,x^{(0)}}(x)[w]\|_{\mathfrak{g}_0^{-1}} + 4874(1-r)^{-7}r\|w\|_{\mathfrak{g}_0} .$$

Since $\|D\tilde{\phi}_0(x)[w]\|_{\mathfrak{g}_0} > \|w\|_{\mathfrak{g}_0}/2$ and $(1-r)^{-7} \leq 2$ with $r \leq 1/11$, the last inequality becomes

$$\begin{aligned}(5h/4)\|D\mathrm{G}_{h,x^{(0)}}(x)[w]\|_{\mathfrak{g}_0^{-1}} &\geq \{1/2 - 9748r\}\|w\|_{\mathfrak{g}_0} ,\\ &\geq 1/2 - r/(4r_\alpha^\star) \geq 1/4 . \qquad \text{(see (33))}\end{aligned}$$

In particular, $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x)$ is invertible, which concludes the proof. □

We prove below that, for any $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$, the set $\underline{\mathrm{G}}_h(z^{(0)})$ is reduced to a single point, for h small enough depending on $z^{(0)}$.

Lemma 23. *For any $w \geq 0$, any $r \in (0,1)$ and any $\alpha \geq 1$, we define*

$$h_1(w,r,\alpha) = \min\left(\frac{1}{1+(1-r)^{-2}+2(2r+w)}, \frac{(1-r)^3}{3(r+w)}, \frac{r}{(r+w)(1+3r/2)+1/2(r+w)^2}\right) . \tag{49}$$

*We recall that the set-valued map $\underline{\mathrm{G}}_h$ is defined in (29). Assume **A**1, **A**2. Then, for any $r \in (0, 1/11]$, any $z = (x,p) \in \mathrm{T}^\star\mathsf{M}$ and any $h \in (0, h_1(\|p\|_{\mathfrak{g}(x)^{-1}}, r, \alpha))$, there exists a unique $z' \in \mathsf{B}_{\|\cdot\|_z}(z,r) \subset \mathrm{T}^\star\mathsf{M}$ such that $z' \in \underline{\mathrm{G}}_h(z)$.*

Proof. Assume **A**1, **A**2. Let $r \in (0,1/11]$, $z = (x,p) \in \mathrm{T}^\star\mathsf{M}$ and $h \in (0, h_1(\|p\|_{\mathfrak{g}(x)^{-1}}, r, \alpha))$, where h_1 is defined in (49). We recall that the map $\underline{\mathrm{g}}_h$ is defined in (30). We define $\mathsf{B} = \mathsf{B}_{\|\cdot\|_z}(z,r)$ and the map $\underline{\mathrm{g}}_{h,z} : \mathrm{T}^\star\mathsf{M} \to \mathbb{R}^n \times \mathbb{R}^n$ by $\underline{\mathrm{g}}_{h,z}(z') = z' - \underline{\mathrm{g}}_h(z,z')$ for any $z' \in \mathrm{T}^\star\mathsf{M}$. Then, $\underline{\mathrm{g}}_{h,z}(z') = z'$ if and only if $z' \in \underline{\mathrm{G}}_h(z)$. The proof is divided in two steps.

Step 1. We first prove that $\underline{\mathrm{g}}_{h,z}(\mathsf{B}) \subset \mathsf{B}$. Let $z' \in \mathsf{B}$, we have by Lemma 9-(a)

$$\begin{aligned}
\|\underline{\mathrm{g}}_{h,z}(z') - z\|_z &= \tfrac{h}{2}\|\{\mathfrak{g}(x)^{-1} + \mathfrak{g}(x')^{-1}\}p'\|_{\mathfrak{g}(x)} + \tfrac{h}{4}\|D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p', \mathfrak{g}(x)^{-1}p']\|_{\mathfrak{g}(x)^{-1}} \\
&= \tfrac{h}{2}\|2\mathfrak{g}(x)^{-1}p' + \{\mathfrak{g}(x')^{-1} - \mathfrak{g}(x)^{-1}\}p'\|_{\mathfrak{g}(x)} + \tfrac{h}{4}\|D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p', \mathfrak{g}(x)^{-1}p']\|_{\mathfrak{g}(x)^{-1}} \\
&\leq \tfrac{h}{2}(2\|p'\|_{\mathfrak{g}(x)^{-1}} + \|\mathfrak{g}(x)^{1/2}(\mathfrak{g}(x')^{-1} - \mathfrak{g}(x)^{-1})\mathfrak{g}(x)^{1/2}\|_2\|p'\|_{\mathfrak{g}(x)^{-1}}) + \tfrac{h}{2}\|p'\|^2_{\mathfrak{g}(x)^{-1}} ,
\end{aligned}$$

where the following inequalities hold

(a) $\|p'\|_{\mathfrak{g}(x)^{-1}} = \|p' - p + p\|_{\mathfrak{g}(x)^{-1}} \leq r + \|p\|_{\mathfrak{g}(x)^{-1}}$.

(b) $\|\mathfrak{g}(x)^{1/2}(\mathfrak{g}(x')^{-1} - \mathfrak{g}(x)^{-1})\mathfrak{g}(x)^{1/2}\|_2 \leq 3\|x - x'\|_{\mathfrak{g}(x)} \leq 3r$, on the model of (18), using $r \leq 1/11$.

Therefore, we have

$$\begin{aligned}
\|\underline{\mathrm{g}}_{h,z}(z') - z\|_z &\leq h((r + \|p\|_{\mathfrak{g}(x)^{-1}}) + \tfrac{3}{2}r(r + \|p\|_{\mathfrak{g}(x)^{-1}}) + \tfrac{1}{2}(r + \|p\|_{\mathfrak{g}(x)^{-1}})^2) \\
&\leq h((r + \|p\|_{\mathfrak{g}(x)^{-1}})(1 + \tfrac{3}{2}r) + \tfrac{1}{2}(r + \|p\|_{\mathfrak{g}(x)^{-1}})^2) \\
&\leq hr/h_1(\|p\|_{\mathfrak{g}(x)^{-1}}, r, \alpha) < r ,
\end{aligned}$$

which proves the statement.

Step 2. We now prove that $\underline{\mathrm{g}}_{h,z}$ is a contraction on B. This proof notably recovers some elements of the proof of Proposition 13 (see Appendix E). Let $(z_1, z_2) \in \mathsf{B} \times \mathsf{B}$ with $z_1 = (x_1, p_1)$ and $z_2 = (x_2, p_2)$. Remark that $x_1, x_2 \in \mathsf{W}^0(x,1) \times \mathsf{W}^0(x,1)$. Let us first bound $\|\underline{\mathrm{g}}^{(1)}_{h,z}(z_1) - \underline{\mathrm{g}}^{(1)}_{h,z}(z_2)\|_{\mathfrak{g}(x)}$. We have

$$\begin{aligned}
\underline{\mathrm{g}}^{(1)}_{h,z}(z_1) - \underline{\mathrm{g}}^{(1)}_{h,z}(z_2) &= \tfrac{h}{2}\{\mathfrak{g}(x)^{-1}(p_1 - p_2) + \mathfrak{g}(x_1)^{-1}p_1 - \mathfrak{g}(x_2)^{-1}p_2\} \\
&= \tfrac{h}{2}\{\mathfrak{g}(x)^{-1}(p_1 - p_2) + \mathfrak{g}(x_1)^{-1}(p_1 - p_2) + (\mathfrak{g}(x_1)^{-1} - \mathfrak{g}(x_2)^{-1})p_2\} ,
\end{aligned}$$

and then, since $r \leq 1/11$, we have on the model of (18)

$$\begin{aligned}
\|\underline{\mathrm{g}}^{(1)}_{h,z}(z_1) - \underline{\mathrm{g}}^{(1)}_{h,z}(z_2)\|_{\mathfrak{g}(x)} &\leq \tfrac{h}{2}\{1 + \|\mathfrak{g}(x)^{1/2}\mathfrak{g}(x_1)^{-1}\mathfrak{g}(x)^{1/2}\|_2\}\|p_1 - p_2\|_{\mathfrak{g}(x)^{-1}} \\
&\quad + \tfrac{h}{2}\|\mathfrak{g}(x)^{1/2}(\mathfrak{g}(x_1)^{-1} - \mathfrak{g}(x_2)^{-1})\mathfrak{g}(x)^{1/2}\|_2\|p_2\|_{\mathfrak{g}(x)^{-1}} \\
&\leq \tfrac{h}{2}\{1 + (1-r)^{-2}\}\|p_1 - p_2\|_{\mathfrak{g}(x)^{-1}} + \tfrac{3h}{2}(1-r)^{-3}\|x_1 - x_2\|_{\mathfrak{g}(x)}(r + \|p\|_{\mathfrak{g}(x)^{-1}}) .
\end{aligned}$$

Let us now bound $\|\underline{\mathrm{g}}^{(2)}_{h,z}(z_1) - \underline{\mathrm{g}}^{(2)}_{h,z}(z_2)\|_{\mathfrak{g}(x)^{-1}}$. We have

$$\underline{\mathrm{g}}^{(2)}_{h,z}(z_1) - \underline{\mathrm{g}}^{(2)}_{h,z}(z_2) = \tfrac{h}{4}\{D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p_1, \mathfrak{g}(x)^{-1}p_1] - D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p_2, \mathfrak{g}(x)^{-1}p_2]\} ,$$

which gives using Lemma 9-(a)

$$\begin{aligned}\|\underline{\mathrm{g}}^{(2)}_{h,z}(z_1)-\underline{\mathrm{g}}^{(2)}_{h,z}(z_2)\|_{\mathfrak{g}(x)^{-1}} &= \tfrac{h}{2}\|D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}(p_1-p_2),\mathfrak{g}(x)^{-1}p_2]\|_{\mathfrak{g}(x)^{-1}}\\ &\quad+\tfrac{h}{4}\|D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}(p_1-p_2),\mathfrak{g}(x)^{-1}(p_1-p_2)]\|_{\mathfrak{g}(x)^{-1}}\\ &\le h\|\mathfrak{g}(x)^{-1}(p_1-p_2)\|_{\mathfrak{g}(x)}\|\mathfrak{g}(x)^{-1}p_2\|_{\mathfrak{g}(x)}+\tfrac{h}{2}\|\mathfrak{g}(x)^{-1}(p_1-p_2)\|^2_{\mathfrak{g}(x)}\\ &\le h(r+\|p\|_{\mathfrak{g}(x)^{-1}})\|p_1-p_2\|_{\mathfrak{g}(x)^{-1}}+hr\|p_1-p_2\|_{\mathfrak{g}(x)^{-1}}\ .\end{aligned}$$

Finally, it comes that

$$\begin{aligned}\|\underline{\mathrm{g}}_{h,z}(z_1)-\underline{\mathrm{g}}_{h,z}(z_2)\|_z &\le \tfrac{h}{2}\{1+(1-r)^{-2}+4r+2\|p\|_{\mathfrak{g}(x)^{-1}}\}\|p_1-p_2\|_{\mathfrak{g}(x)^{-1}}\\ &\quad+\tfrac{3h}{2}(1-r)^{-3}(r+\|p\|_{\mathfrak{g}(x)^{-1}})\|x_1-x_2\|_{\mathfrak{g}(x)}\\ &\le (1/2)h\{\|x_1-x_2\|_{\mathfrak{g}(x)}+\|p_1-p_2\|_{\mathfrak{g}(x)^{-1}}\}/h_1(\|p\|_{\mathfrak{g}(x)^{-1}},r,\alpha)\\ &< (1/2)\|z_1-z_2\|_z,\end{aligned}$$

which proves that $\underline{\mathrm{g}}_{h,z}$ is a contraction on B.

Conclusion. We obtain the result of Lemma 23 by applying the fixed point theorem on $\underline{\mathrm{g}}_{h,z}$ and B. □

Elaborating on Lemma 23, we prove in Lemma 24 that the only element of $\underline{\mathrm{G}}_h(z^{(0)})$ verifies smoothness properties if h is chosen small enough, depending on $z^{(0)}$.

Lemma 24. *For any $z=(x,p)\in\mathrm{T}^\star\mathsf{M}$, any $r\in(0,1)$ and any $\alpha\ge 1$, we define*

$$\bar{h}(z,r,\alpha)=\min\left(h_1(\|p\|_{\mathfrak{g}(x)^{-1}},r,\alpha),h_2(z),h_3(z,r)\right) \tag{50}$$

where h_1 is defined in (49), and h_2 and h_3 are defined by

$$h_2(z)=\frac{1}{3/2+3\|p\|_{\mathfrak{g}(x)^{-1}}}\ ,\qquad h_3(z,r)=\frac{1}{1+(1-r)^{-1}(r+\|p\|_{\mathfrak{g}(x)^{-1}})}\ .$$

*We recall that $r^\star_\alpha$ is defined in (33). We also recall that the maps $\underline{\mathrm{G}}_h$, $\underline{\mathrm{g}}_h$, $\overline{\mathrm{G}}_h$ and $\mathrm{G}_{h,x^{(0)}}$ are respectively defined in (29), (30), (31) and (32). Assume **A**1, **A**2. Then, for any $z^{(0)}\in\mathrm{T}^\star\mathsf{M}$ and any $h\in(0,\bar{h}(z^{(0)},r^\star_\alpha,\alpha))$, there exists a unique element $z_h^{(1/2)}\in\mathrm{T}^\star\mathsf{M}$ such that $z_h^{(1/2)}\in\underline{\mathrm{G}}_h(z^{(0)})\cap\mathsf{B}_{\|\cdot\|_{z^{(0)}}}(z^{(0)},r^\star_\alpha)$. Moreover, we have*

(a) $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)},z_h^{(1/2)})$ *and* $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ *are invertible,*

(b) $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x_h^{(1/2)})$ *is invertible.*

Proof. Assume **A**1, **A**2. Let $z^{(0)}\in\mathrm{T}^\star\mathsf{M}$ and $h\in(0,\bar{h}(z^{(0)},r^\star_\alpha,\alpha))$, where $\bar{h}$ is defined in (50). Since $r^\star_\alpha\le 1/11$, Lemma 23 ensures the existence and uniqueness of $z_h^{(1/2)}=(x_h^{(1/2)},p_h^{(1/2)})\in\mathrm{T}^\star\mathsf{M}$ such that $z_h^{(1/2)}\in\underline{\mathrm{G}}_h(z^{(0)})\cap\mathsf{B}_{\|\cdot\|_{z^{(0)}}}(z^{(0)},r^\star_\alpha)$. Moreover, we have $x_h^{(1/2)}\in\mathsf{W}^0(x^{(0)},r^\star_\alpha)$, and thus $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x_h^{(1/2)})$ is invertible by Lemma 22. Let us prove that $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)},z_h^{(1/2)})$ and $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ are invertible.

Invertibility of $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)},z_h^{(1/2)})$. We first remark that $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)},z_h^{(1/2)})=\mathrm{Jac}[\psi_{0,h}](z_h^{(1/2)})$ where $\psi_{0,h}=\mathrm{Id}+\frac{h}{2}\psi_0$ and ψ_0 is defined on $\mathrm{T}^\star\mathsf{M}$ by

$$\psi_0(z)=(-\{\mathfrak{g}(x^{(0)})^{-1}+\mathfrak{g}(y)^{-1}\}^{-1}p,-\tfrac{1}{2}D\mathfrak{g}(x^{(0)})[\mathfrak{g}(x^{(0)})^{-1}p,\mathfrak{g}(x^{(0)})^{-1}p])\ .$$

We recall that $\|z_h^{(1/2)}-z^{(0)}\|_{z^{(0)}}<r^\star_\alpha\le 1/11$ and thus define $r_1=1/11-\|z_h^{(1/2)}-z^{(0)}\|_{z^{(0)}}>0$ and $\mathsf{B}_1=\mathsf{B}_{\|\cdot\|_{z^{(0)}}}(z_h^{(1/2)},r_1)$. We set $\tilde{r}=1/11$. Note that $\mathsf{B}_1\subset\mathsf{B}_{\|\cdot\|_{z^{(0)}}}(z^{(0)},1/11)$. We show below that ψ_0 is Lipschitz-continuous on B_1 with respect to $\|\cdot\|_{z^{(0)}}$. Let $(z, z') \in \mathsf{B}_1 \times \mathsf{B}_1$. Similarly to the proof of Lemma 23, we have

$$\begin{aligned}\|\psi_0^{(1)}(z) - \psi_0^{(1)}(z')\|_{\mathfrak{g}(x^{(0)})} &\leq \|p - p'\|_{\mathfrak{g}(x^{(0)})^{-1}} \\ &\quad + \|\mathfrak{g}(x^{(0)})^{1/2}\mathfrak{g}(x)^{-1}\mathfrak{g}(x^{(0)})^{1/2}\|_2\|p - p'\|_{\mathfrak{g}(x^{(0)})^{-1}} \\ &\quad + \|\mathfrak{g}(x^{(0)})^{1/2}\{\mathfrak{g}(x)^{-1} - \mathfrak{g}(x')^{-1}\}\mathfrak{g}(x^{(0)})^{1/2}\|_2\|p\|_{\mathfrak{g}(x^{(0)})^{-1}} \\ &\leq (1 + (1 - \tilde{r})^{-2})\|p - p'\|_{\mathfrak{g}(x^{(0)})^{-1}} \\ &\quad + 3(1 - \tilde{r})^{-3}(\tilde{r} + \|p^{(0)}\|_{\mathfrak{g}(x^{(0)})^{-1}})\|x - x'\|_{\mathfrak{g}(x^{(0)})} .\end{aligned}$$

In the same manner, we have

$$\|\psi_0^{(2)}(z) - \psi_0^{(2)}(z')\|_{\mathfrak{g}(x^{(0)})^{-1}} \leq 2(2\tilde{r} + \|p^{(0)}\|_{\mathfrak{g}(x^{(0)})^{-1}})\|p - p'\|_{\mathfrak{g}(x^{(0)})^{-1}} .$$

Therefore, recalling that $\tilde{r} = 1/11$, we obtain with the previous inequalities

$$\begin{aligned}\|\psi_0(z) - \psi_0(z')\|_{z^{(0)}} &\leq \{3 + 6\|p^{(0)}\|_{\mathfrak{g}(x^{(0)})^{-1}}\}\|z - z'\|_{z^{(0)}} \\ &\leq 2\|z - z'\|_{z^{(0)}}/h_2(z^{(0)}) .\end{aligned}$$

Hence, ψ_0 is $(2/h_2(z^{(0)}))$-Lipschitz-continuous on B_1 with respect to $\|\cdot\|_{z^{(0)}}$. Then, since $h < h_2(z^{(0)})$, $\mathrm{Jac}[\psi_{0,h}](z)$ is invertible for any $z \in \mathsf{B}_1$ by Lemma 21. In particular, $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)}, z_h^{(1/2)}) = \mathrm{Jac}[\psi_{0,h}](z_h^{(1/2)})$ is invertible.

Invertibility of $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$. Let us remark that $\overline{\mathrm{G}}_h = \mathrm{Id} + h\zeta$ where ζ is defined on $\mathrm{T}^\star\mathsf{M}$ by

$$\zeta(z) = (0, \frac{1}{4}D\mathfrak{g}(x)[\mathfrak{g}(x)^{-1}p, \mathfrak{g}(x)^{-1}p]).$$

We define $r_2 = 1/2$ and $\mathsf{B}_2 = \mathsf{B}_{\|\cdot\|_{z_h^{(1/2)}}}(z_h^{(1/2)}, r_2)$. We show below that ζ is Lipschitz-continuous on B_2 with respect to $\|\cdot\|_{z_h^{(1/2)}}$, with a Lipschitz constant which does not depend on $z_h^{(1/2)}$. Let $(z, z') \in \mathsf{B}_2 \times \mathsf{B}_2$. Using Lemma 9-(a) with $r_2 = 1/2$, it is clear that

$$\|\zeta(z) - \zeta(z')\|_{z_h^{(1/2)}} \leq (1 + \|p_h^{(1/2)}\|_{\mathfrak{g}(x_h^{(1/2)})^{-1}})\|z - z'\|_{z_h^{(1/2)}} .$$

Moreover, since $\|z_h^{(1/2)} - z^{(0)}\|_{z^{(0)}} < r_\alpha^\star$, we have by Lemma 12

$$\begin{aligned}\|p_h^{(1/2)}\|_{\mathfrak{g}(x_h^{(1/2)})^{-1}} &\leq (1 - r_\alpha^\star)^{-1}\|p_h^{(1/2)}\|_{\mathfrak{g}(x^{(0)})^{-1}} \\ &\leq (1 - r_\alpha^\star)^{-1}(r_\alpha^\star + \|p^{(0)}\|_{\mathfrak{g}(x^{(0)})^{-1}}) ,\end{aligned}$$

and thus

$$\begin{aligned}\|\zeta(z) - \zeta(z')\|_{z_h^{(1/2)}} &\leq \{1 + (1 - r_\alpha^\star)^{-1}(r_\alpha^\star + \|p^{(0)}\|_{\mathfrak{g}(x^{(0)})^{-1}})\}\|z - z'\|_{z_h^{(1/2)}} \\ &\leq \|z - z'\|_{z_h^{(1/2)}}/h_3(z^{(0)}, r_\alpha^\star) .\end{aligned}$$

Hence, ζ is $(1/h_3(z^{(0)}, r_\alpha^\star))$-Lipschitz-continuous on B_2 with respect to $\|\cdot\|_{z_h^{(1/2)}}$. Since $h < h_3(z^{(0)}, r_\alpha^\star)$, $\mathrm{Jac}[\overline{\mathrm{G}}_h](z)$ is invertible for any $z \in \mathsf{B}_2$ by Lemma 21. In particular, $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ is invertible, which concludes the proof. □

The following lemma states the existence of a diffeomorphism between $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$ and $z^{(1)} \in \mathrm{G}_h(z^{(0)})$ under smoothness conditions verified by $z^{(1)}$.

Lemma 25. *Let $(z^{(0)}, z^{(1)}) \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$. We recall that the maps $\underline{\mathrm{g}}_h$, $\overline{\mathrm{G}}_h$ and G_h are respectively defined in (30), (31) and (6). Assume that $\mathfrak{g} \in \mathrm{C}^2(\mathsf{M}, \mathbb{R}^{d\times d})$ and that there exist $h > 0$ and $z_h^{(1/2)} \in \mathrm{T}^\star\mathsf{M}$ with the following properties:*

(a) $\underline{\mathrm{g}}_h(z^{(0)}, z_h^{(1/2)}) = 0$ *and* $z^{(1)} = \overline{\mathrm{G}}_h(z_h^{(1/2)})$.

(b) $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)}, z_h^{(1/2)})$ *and* $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ *are invertible.*

Then, there exists a neighbourhood $\mathsf{U} \subset \mathrm{T}^\star\mathsf{M}$ *of* $z^{(0)}$ *and a* C^1*-diffeomorphism* $\xi : \mathsf{U} \to \xi(\mathsf{U}) \subset \mathrm{T}^\star\mathsf{M}$ *such that (i)* $\xi(z^{(0)}) = z^{(1)}$*, (ii) for any* $z \in \mathsf{U}$*,* $\xi(z) \in \mathrm{G}_h(z)$ *and (iii)* $|\det\mathrm{Jac}\,\xi| \equiv 1$*.*

Proof. Let $(z^{(0)}, z^{(1)}) \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$. Assume that $\mathfrak{g} \in \mathrm{C}(\mathsf{M}, \mathbb{R}^{d\times d})$ and consider $h > 0$ and $z_h^{(1/2)} \in \mathrm{T}^\star\mathsf{M}$ as described in Lemma 25. Under the assumption on $\mathfrak{g}$, $\underline{\mathrm{g}}_h$ and $\overline{\mathrm{G}}_h$ are continuously differentiable respectively on $\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$ and $\mathrm{T}^\star\mathsf{M}$. We are going to prove Lemma 25, by first deriving intermediary results on $\underline{\mathrm{g}}_h$ and $\overline{\mathrm{G}}_h$.

Result on $\underline{\mathrm{g}}_h$. We recall that $\underline{\mathrm{g}}_h(z^{(0)}, z_h^{(1/2)}) = 0$ and $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z^{(0)}, z_h^{(1/2)})$ is invertible. Then, by applying the implicit function theorem on $\underline{\mathrm{g}}_h$ at $(z^{(0)}, z_h^{(1/2)})$, we obtain the existence of a neighbourhood $\mathsf{U}_0 \subset \mathrm{T}^\star\mathsf{M}$ of $z^{(0)}$ and a C^1-diffeomorphism $\xi_0 : \mathsf{U}_0 \to \xi_0(\mathsf{U}_0) \subset \mathrm{T}^\star\mathsf{M}$ such that $\xi_0(z^{(0)}) = z_h^{(1/2)}$ and for any $z \in \mathsf{U}_0$, $\underline{\mathrm{g}}_h(z, \xi_0(z)) = 0$, *i.e.*, $\xi_0(z) \in \underline{\mathrm{G}}_h(z)$. Moreover, for any $z \in \mathsf{U}_0$, the Jacobian of ξ_0 at z is given by

$$\mathrm{Jac}[\xi_0](z) = -\,\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z, \xi_0(z))^{-1}\,\mathrm{Jac}_z[\underline{\mathrm{g}}_h](z, \xi_0(z))\ .$$

Result on $\overline{\mathrm{G}}_h$. We recall that $z^{(1)} = \overline{\mathrm{G}}_h(z_h^{(1/2)})$ and $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ is invertible. Then, we apply the inverse function theorem on $\overline{\mathrm{G}}_h$ at $z_h^{(1/2)}$ and obtain the existence of a neighbourhood $\mathsf{U}_1 \subset \mathrm{T}^\star\mathsf{M}$ of $z_h^{(1/2)}$ such that $\xi_1 = \overline{\mathrm{G}}_h|_{\mathsf{U}_1}$ is a C^1-diffeomorphism on U_1.

Final result. We now consider the subset U and the map ξ, respectively defined by

(a) $\mathsf{U} = \xi_0^{-1}(\mathsf{U}_1 \cap \xi_0(\mathsf{U}_0)) \subset \mathrm{T}^\star\mathsf{M}$, neighbourhood of $z^{(0)}$.

(b) $\xi = \xi_1|_{\xi_0(\mathsf{U})} \circ \xi_0$, C^1-diffeomorphism on U such that $\xi(z^{(0)}) = z^{(1)}$, and for any $z \in \mathsf{U}$, $\xi(z) \in \mathrm{G}_h(z)$.

Let us now prove that $|\det\mathrm{Jac}\,\xi| \equiv 1$. Let $z \in \mathsf{U}$. We define $z_0 = \xi_0(z)$. By the chain rule, we have

$$\begin{aligned}\mathrm{Jac}[\xi](z) &= \mathrm{Jac}[\xi_1 \circ \xi_0](z)\\ &= \mathrm{Jac}[\xi_1](\xi_0(z))\,\mathrm{Jac}[\xi_0](z)\\ &= -\,\mathrm{Jac}[\xi_1](\xi_0(z))\,\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z, \xi_0(z))^{-1}\,\mathrm{Jac}_z[\underline{\mathrm{g}}_h](z, \xi_0(z))\ ,\end{aligned}$$

where we have

(a) $\mathrm{Jac}[\xi_1](z') = \begin{pmatrix} \mathrm{I}_d & 0 \\ -\frac{h}{2}\partial_{x,x}H_2(z') & \mathrm{I}_d - \frac{h}{2}\partial_{x,p}H_2(z') \end{pmatrix}$,

(b) $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](z, z') = \begin{pmatrix} \mathrm{I}_d - \frac{h}{2}\partial_{x,p}H_2(z') & -\frac{h}{2}\{\partial_{p,p}H_2(x,p') + \partial_{p,p}H_2(x',p')\} \\ 0 & \mathrm{I}_d + \frac{h}{2}\partial_{x,p}H_2(x,p') \end{pmatrix}$,

(c) $\mathrm{Jac}_z[\underline{\mathrm{g}}_h](z, z') = \begin{pmatrix} -\mathrm{I}_d - \frac{h}{2}\partial_{x,p}H_2(x,p') & 0 \\ \frac{h}{2}\partial_{x,x}H_2(x,p') & -\mathrm{I}_d \end{pmatrix}$.

Therefore, we obtain

$$\begin{aligned}&|\det\mathrm{Jac}[\xi](z)|\\ &\quad= \det(\mathrm{I}_d - \frac{h}{2}\partial_{x,p}H_2(z_0))\det(\{\mathrm{I}_d - \frac{h}{2}\partial_{x,p}H_2(z_0)\}\{\mathrm{I}_d + \frac{h}{2}\partial_{x,p}H_2(x,p_0)\})^{-1}\det(\mathrm{I}_d + \frac{h}{2}\partial_{x,p}H_2(x,p_0)) = 1\ ,\end{aligned}$$

which concludes the proof. □

Lemma 26. *Let* $h > 0$*. We recall that the set-valued map* F_h *is defined in Section 3.1. Let* $z \in \mathrm{T}^\star\mathsf{M}$*. Assume that there exist* $(z', z'') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$ *such that* *(a)* $z' \in \mathrm{F}_h(z)$*,*

(b) $z'' \in \mathrm{F}_h(z)$*, and*

(c) $x' = x''$*.*

Then, we have $p' = p''$*, and thus* $z' = z''$*.*

Proof. Let $h > 0$. Let $z \in \mathrm{T}^\star\mathsf{M}$. We consider $(z', z'') \in \mathrm{F}_h(z) \times \mathrm{F}_h(z)$ such that $x' = x'' = \tilde{x}$. By using the last two aligns from (6) with z' and z'', we have $p' = -p^{(1/2)} + \frac{h}{2}\mathfrak{g}(\tilde{x})^{-1}p^{(1/2)} = p''$, where $p^{(1/2)} = \frac{2}{h}(\mathfrak{g}(x)^{-1} + \mathfrak{g}(\tilde{x})^{-1})^{-1}(\tilde{x} - x)$, which concludes the proof. □

Under the assumption **A**2, we define, for any $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$, $h^\star(z^{(0)}) = \bar{h}(s(z^{(0)}), r_\alpha^\star, \alpha)$, where α is given by **A**2, and $r_\alpha^\star$ and $\bar{h}$ are respectively defined in (33) and (50). Note that $h^\star$ appears in Proposition 2, for which we derive the proof below.

Proof of Proposition 2. We recall that maps $\underline{\mathrm{G}}_h$, $\underline{\mathrm{g}}_h$ and $\overline{\mathrm{G}}_h$ are respectively defined in (29), (30) and (31). Assume **A**1, **A**2. Let $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$. We define $\tilde{z}^{(0)} = s(z^{(0)})$. Let $h \in (0, h^\star(z^{(0)}))$. By using Lemma 24 on $\tilde{z}^{(0)}$, we obtain the existence of $z_h^{(1/2)} \in \mathrm{T}^\star\mathsf{M}$ with the following properties:

(a) $z_h^{(1/2)} \in \underline{\mathrm{G}}_h(\tilde{z}^{(0)})$, *i.e.*, $\underline{\mathrm{g}}_h(\tilde{z}^{(0)}, z_h^{(1/2)}) = 0$,

(b) $\mathrm{Jac}_{z'}[\underline{\mathrm{g}}_h](\tilde{z}^{(0)}, z_h^{(1/2)})$ and $\mathrm{Jac}[\overline{\mathrm{G}}_h](z_h^{(1/2)})$ are invertible,

(c) $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x_h^{(1/2)})$ is invertible.

We then define $z_h^{(1)} = \overline{\mathrm{G}}_h(z_h^{(1/2)})$. In particular, $z_h^{(1)} \in \mathrm{G}_h(\tilde{z}^{(0)})$, *i.e.*, $z_h^{(1)} \in \mathrm{F}_h(z^{(0)})$ and $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x_h^{(1)}) = \mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x_h^{(1/2)})$ is invertible. By combining the properties of $z_h^{(1/2)}$ and with Lemma 25, we also obtain the existence of a neighbourhood $\mathsf{U}' \subset \mathrm{T}^\star\mathsf{M}$ of $\tilde{z}^{(0)}$ and a C^1-diffeomorphism $\xi_h : \mathsf{U}' \to \xi_h(\mathsf{U}') \subset \mathrm{T}^\star\mathsf{M}$ such that (i) $\xi_h(\tilde{z}^{(0)}) = z_h^{(1)}$, (ii) for any $z \in \mathsf{U}'$, $\xi_h(z) \in \mathrm{G}_h(z)$ and (iii) $|\det \mathrm{Jac}\, \xi_h| \equiv 1$. Therefore, we can define the subset U and the map γ_h of Proposition 2 by

(a) $\mathsf{U} = s(\mathsf{U}')$, neighbourhood of $z^{(0)}$ in $\mathrm{T}^\star\mathsf{M}$,

(b) $\gamma_h = \xi_h \circ s$, such that (i) $\gamma_h(z^{(0)}) = z_h^{(1)}$, (ii) for any $z \in \mathsf{U}$, $\gamma_h(z) \in \mathrm{F}_h(z)$ and (iii) $\det \mathrm{Jac}[\gamma_h] \equiv 1$.

We now prove that U can be reduced to a smaller subset such that $\gamma_h(z)$ is the only element of $\mathrm{F}_h(z)$ in $\gamma_h(\mathsf{U})$ for any $z \in \mathsf{U}$. Motivated by Lemma 22, we first define, for any $(x, x') \in \mathsf{M} \times \mathsf{M}$, $\mathrm{F}_{h,x}(x')$ as the only element $p \in \mathrm{T}_x^\star\mathsf{M}$ such that $(x', p') \in \mathrm{F}_h(x, p)$ for any $p' \in \mathrm{T}_{x'}^\star\mathsf{M}$. It is clear that $\mathrm{F}_{h,x}(x') = -\mathrm{G}_{h,x}(x')$, where $\mathrm{G}_{h,x}$ is defined in (32). We also define $\mathrm{Z}_h : \mathsf{M} \times \mathsf{M} \to \mathrm{T}^\star\mathsf{M}$ by

$$\mathrm{Z}_h(x, x') = (x, \mathrm{F}_{h,x}(x')) = (x, -\mathrm{G}_{h,x}(x')) \ ,$$

which is continuously differentiable since $\mathfrak{g} \in \mathrm{C}^2(\mathsf{M}, \mathbb{R}^{d\times d})$. Besides this, we obtain by Lemma 22 that $\mathrm{Jac}[\mathrm{G}_{h,x^{(0)}}](x^{(0)}, x_h^{(1)})$ is invertible, and then $\mathrm{Jac}[\mathrm{Z}_h](x_h^{(1)})$ is invertible. Therefore, by applying the inverse function theorem on Z_h at $(x^{(0)}, x_h^{(1)})$, it comes that Z_h is a C^1-diffeomorphism in a neighbourhood of $(x^{(0)}, x_h^{(1)})$. We are now going to prove the result by contradiction. Assume for now that there is no neighbourhood U of $z^{(0)}$ such that $\gamma_h(z)$ is the only element of $\mathrm{F}_h(z)$ in $\gamma_h(\mathsf{U})$, for any $z \in \mathsf{U}$. Then, we can find a sequence $(z_i)_{i\in\mathbb{N}}$ which converges to $z^{(0)}$ such that for any $i \in \mathbb{N}$, there exist two different elements $z_{i,1} \in \mathrm{F}_h(z_i)$ and $z_{i,2} \in \mathrm{F}_h(z_i)$. Therefore, for any $i \in \mathbb{N}$, $x_{i,1} \neq x_{i,2}$ and by Lemma 26, we have

$$\mathrm{F}_{h,x_i}(x_{i,1}) = \mathrm{F}_{h,x_i}(x_{i,2}) = p_i \ ,$$

and thus

$$Z_h(x_i, x_{i,1}) = Z_h(x_i, x_{i,2}) = z_i \ . \tag{51}$$

Moreover, by continuity of γ_h, the sequences $(z_{i,1})_{i\in\mathbb{N}}$ and $(z_{i,2})_{i\in\mathbb{N}}$ converge to $z_h^{(1)}$, and therefore, $(x_{i,1})_{i\in\mathbb{N}}$ and $(x_{i,2})_{i\in\mathbb{N}}$ also converge to $x^{(1)}$. Combined with (51), this result of convergence is in contradiction with the fact that Z_h is a diffeomorphism in a neighbourhood of $(x^{(0)}, x_h^{(1)})$. Therefore, we can reduce U to a smaller subset such that $\gamma_h(z)$ is the only element of $\mathrm{F}_h(z)$ in $\gamma_h(\mathsf{U})$ for any $z \in \mathsf{U}$, which concludes the proof of Proposition 2. □

H Modification of n-BHMC algorithm with step-size conditioning

In the rest of the paper, for any $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$, we will denote by $h_\star(z^{(0)})$ the value of $h_\star$ given by **A**3.

Beyond n-BHMC. A crucial part of the proof of reversibility in BHMC, as much for c-BHMC as for n-BHMC, relies on local *symplectic* properties of the integrator of the Hamiltonian dynamics. Although Algorithm 1 can be implemented without any practical limitation, it is hard to state such properties for its numerical integrator Φ_h under **A**1, **A**2 and **A**3, given *any* value of h. Indeed, we know from **A**3 that Φ_h is a local involution around $z^{(0)} \in \mathrm{T}^\star\mathsf{M}$ when $h < h_\star(z^{(0)})$; however, we cannot ensure this result when $h > h_\star(z^{(0)})$... To circumvent this issue, we propose to study Theoretical n-BHMC (Tn-BHMC), presented in Algorithm 3. In this modified version of n-BHMC, we actually enforce a condition on h to be small enough. We now get into the details of this new algorithm and assume **A**3 for the rest of this section.

Theoretical motivations. Let $h > 0$. We recall the definition of the set A_h introduced in Section 4.2

$$\mathsf{A}_h = \{z \in \mathrm{T}^\star\mathsf{M} \ : \ h < \min_{\tilde{z}\in \mathsf{B}_{\|\cdot\|_z}(z,1)} h_\star(\tilde{z})\} \ .$$

It is clear that $\mathsf{A}_h \subset \mathrm{dom}_{\Phi_h}$: indeed, if $z^{(0)} \in \mathsf{A}_h$, we have in particular that $h < h_\star(z^{(0)})$, and therefore $z^{(0)} \in \mathrm{dom}_{\Phi_h}$ by **A**3. Let $z^{(0)} \in \mathsf{A}_h$. By **A**3, we know that Φ_h is an involution on a neighbourhood of $z^{(0)}$; in particular, it comes that $(\Phi_h \circ \Phi_h)(z^{(0)}) = z^{(0)}$. Hence, the condition (b) of the "involution checking step" in Algorithm 1 is *de facto* satisfied. This naturally leads to replace the condition "$z \in \mathrm{dom}_{\Phi_h}$" by the more restrictive condition "$z \in \mathsf{A}_h$".

Algorithm 3: Tn-BHMC with Momentum Refreshment

Input: $(x_0, p_0) \in \mathrm{T}^\star\mathsf{M}$, $\beta \in (0, 1]$, $N \in \mathbb{N}$, $h > 0$, $\eta > 0$, Φ_h with domain dom_{Φ_h}, A_h
Output: $(X_n, P_n)_{n\in[N]}$

for $n = 1, ..., N$ **do**
  Step 1: $G_n \sim \mathrm{N}_x(0, \mathrm{I}_d)$, $\tilde{P}_n \leftarrow \sqrt{1-\beta}P_{n-1} + \sqrt{\beta}G_n$
  Step 2: solving a discretized version of ODE (3)
  $X'_n, P'_n \leftarrow X_{n-1}, \tilde{P}_n$; $X_n^{(0)}, P_n^{(0)} \leftarrow (s \circ \mathrm{S}_{h/2})(X_n, \tilde{P}_n)$
  if $Z_n^{(0)} = (X_n^{(0)}, P_n^{(0)}) \in \mathsf{A}_h$ **then**
    $Z_n^{(1)} = \Phi_h(Z_n^{(0)})$
    if $Z_n^{(1)} \in \mathsf{A}_h$ **then**
      $X'_n, P'_n \leftarrow (s \circ \mathrm{S}_{h/2})(Z_n^{(1)})$
    end
  end
  Step 3: $A_n \leftarrow \min(1, \exp[-H(X'_n, P'_n) + H(X_{n-1}, \tilde{P}_n)])$
  $U_n \sim U[0, 1]$
  if $U_n \leq A_n$ **then** $\bar{X}_n, \bar{P}_n \leftarrow X'_n, P'_n$
  else $\bar{X}_n, \bar{P}_n \leftarrow X_{n-1}, \tilde{P}_n$
  Step 4: $X_n, \hat{P}_n \leftarrow s(\bar{X}_n, \bar{P}_n)$
  Step 5: $G'_n \sim \mathrm{N}_x(0, \mathrm{I}_d)$, $P_n \leftarrow \sqrt{1-\beta}\hat{P}_n + \sqrt{\beta}G'_n$
end

Implementation of Tn-BHMC. In Algorithm 3, we highlight in yellow the modifications of Tn-BHMC in contrast to n-BHMC. Namely, we replace

(a) "$Z_n^{(0)} \in \mathrm{dom}_{\Phi_h}$" (Line 5 in Algorithm 1) by "$Z_n^{(0)} \in \mathsf{A}_h$" (Line 5 in Algorithm 3).

(b) "$Z_n^{(1)} \in \mathrm{dom}_{\Phi_h}$" (Line 8 in Algorithm 1) by "$Z_n^{(1)} \in \mathsf{A}_h$" (Line 7 in Algorithm 3).

Note that there is no need to maintain the "involution checking step" of Algorithm 1, since it is automatically verified once $Z_n^{(0)} \in \mathsf{A}_h$. On the whole, these new conditions are more *restrictive* than the conditions of Algorithm 1 since $\mathsf{A}_h \subset \mathrm{dom}_{\Phi_h}$; moreover, they can be thought as conditions directly applied on the step-size h. The specific choice of A_h, instead of another subset of dom_{Φ_h}, is actually sufficient to derive the proof of reversibility of Tn-BHMC (see Section 4.2).

I Proofs of Section 4.2

I.1 Expression and properties of $r^\star$

Given a Riemannian manifold $(\mathsf{M}, \mathfrak{g})$, we define $r^\star(x)$ for any $x \in \mathsf{M}$ by

$$r^\star(x) = \min(\|\mathfrak{g}(x)\|_2^{-1/2}, \|\mathfrak{g}(x)^{-1}\|_2^{-1/2}) \, . \tag{52}$$

Note that $r^\star$ is used in **A**3 and that $r^\star(x) = 1/C_x$, where C_x is defined in Lemma 12. We prove below that $r^\star$ has a smooth behaviour on M under our main assumptions.

Lemma 27. *Assume* **A***1,* **A***2. Also assume that $x \in \mathsf{M} \mapsto \|\mathfrak{g}^{-1}(x)\|_2$ is bounded from above. As defined in (52), $r^\star : \mathsf{M} \to (0, +\infty)$ satisfies the following properties:*

(a) $r^\star(x) \to 0$ as $x \to \partial M$.

(b) There exists $\mathtt{L} > 0$ such that $r^\star$ is $\mathtt{L}$-Lipschitz-continuous on M with respect to $\|\cdot\|_2$.

(c) There exists $\mathtt{M} > 0$ such that $r^\star(x) \le \mathtt{M}$ for any $x \in \mathsf{M}$.

Proof. Assume **A**1, **A**2. Also assume that $x \in \mathsf{M} \mapsto \|\mathfrak{g}^{-1}(x)\|_2$ is bounded from above. We first define $r_1 : x \in \mathsf{M} \mapsto \|\mathfrak{g}(x)\|_2^{-1/2}$ and $r_2 : x \in \mathsf{M} \mapsto \|\mathfrak{g}(x)^{-1}\|_2^{-1/2}$, such that $r^\star = \min(r_1, r_2)$. Since $\mathfrak{g} \in \mathrm{C}^2(\mathsf{M}, \mathbb{R}^{d\times d})$, it is clear that r_1 and r_2 are continuously differentiable on M. We have:

(a) $r_1(x) \to 0$ as $x \to \partial\mathsf{M}$ by Lemma 8, and

(b) $r_2(x) \not\to 0$ as $x \to \partial\mathsf{M}$, since $1/r_2^2 : x \mapsto \|\mathfrak{g}(x)^{-1}\|_2$ is bounded on M.

Combining the fact that $r_2(x) > 0$ for any $x \in \mathsf{M}$ and item (a), we obtain item (a) of Lemma 27. We denote by d the distance induced by $\|\cdot\|_2$ and now define, for any $\varepsilon > 0$,

(i) $\mu_\varepsilon = \inf_{\{y\in\mathsf{M}\,:\, d(y,\partial\mathsf{M})\le\epsilon\}} r_2(y)$.

(ii) $\mathsf{M}_\varepsilon = \mathrm{Int}(\{x \in \mathsf{M} \;:\; d(x, \partial\mathsf{M}) \le \varepsilon, r_1(x) \le \mu_\varepsilon\})$, open set in M.

(iii) $\mathsf{M}_{-\varepsilon} = \mathsf{M}\backslash\mathsf{M}_\varepsilon$, closed and bounded (and thus, compact) set in M.

Using items (a) and (b), we can ensure the existence of some $\varepsilon \in (0, \mathrm{diam}(\mathsf{M}))$ such that (i) $\mu_\varepsilon > 0$ and (ii) M_ε and $\mathsf{M}_{-\varepsilon}$ are not empty. We consider such ε for the rest of the proof.

Smoothness of r_2. We define $\delta = d(\mathsf{M}^c, \mathsf{M}_{-\varepsilon})$ and $\mathsf{M}_\delta = \mathsf{M}^c + \bar{\mathsf{B}}(0, \delta/4)$. Note that

(i) $\delta > 0$ since $\mathsf{M}_{-\varepsilon} \subset \mathsf{M}$,

(ii) M_δ is closed since the ball $\bar{\mathsf{B}}(0, \delta/4)$ is compact, and

(iii) $\mathsf{M}_\delta \cap \mathsf{M}_{-\varepsilon} = \emptyset$.

According to the smooth version Urysohn's lemma applied to M_δ and $\mathsf{M}_{-\varepsilon}$, there exists $\chi \in \mathrm{C}^1(\mathbb{R}^d, [0,1])$ such that $\chi(\mathsf{M}_{-\varepsilon}) = 1$ and $\chi(\mathsf{M}_\delta) = 0$. We then define $\tilde{r}_2 : \mathbb{R}^d \to (0, +\infty)$ by $\tilde{r}_2 = \chi r_2 + (1-\chi)\mu_\varepsilon$. In particular, (i) there exists $\mathsf{L}_2 > 0$ such that $\tilde{r}_2$ is L_2-Lipschitz-continuous on $\bar{\mathsf{M}}$ with respect to $\|\cdot\|_2$, since $\tilde{r}_2 \in \mathrm{C}^1(\mathbb{R}^d, [0,1])$, and (ii) for any $x \in \mathsf{M}_\varepsilon$, $\tilde{r}_2(x) > \mu_\varepsilon$.

Smoothness of r_1. We now prove that r_1 is 1-Lipschitz continuous on M with respect to $\|\cdot\|_2$. Let $x \in \mathsf{M}$. Note that $r_1(x) = (\|\mathfrak{g}(x)\|_2^2)^{-1/4}$ and we thus have

$$\nabla r_1(x) = (-1/4)\|\mathfrak{g}(x)\|_2^{-5/2} h(x) : D\mathfrak{g}(x) ,$$

where $h(x) = \partial_{\mathfrak{g}(x)}\|\mathfrak{g}(x)\|_2^2 = 2\|\mathfrak{g}(x)\|_2 u(x)u(x)^\top$, $u(x)$ being a normal eigenvector of $\mathfrak{g}(x)$ corresponding to the eigenvalue $\|\mathfrak{g}(x)\|_2$. Hence,

$$\begin{aligned}
\|\nabla r_1(x)\|_2 &\leq (1/2)\|\mathfrak{g}(x)\|_2^{-3/2}\|u(x)u(x)^\top : D\mathfrak{g}(x)\|_2 \\
&\leq (1/2)\|\mathfrak{g}(x)\|_2^{-3/2}\|u(x)u(x)^\top\|_2\|D\mathfrak{g}(x)\|_2 \\
&\leq (1/2)\|\mathfrak{g}(x)\|_2^{-3/2} \times 2\|\mathfrak{g}(x)\|_2^{3/2} \leq 1 \qquad \text{(Definition 1-(c))}
\end{aligned}$$

which proves the result on r_1.

Smoothness of r. Let $x \in \mathsf{M}$. We can face two cases: either, $x \in \mathsf{M}_{-\varepsilon}$, then $r_2(x) = \tilde{r}_2(x)$; or, $x \in \mathsf{M}_\varepsilon$, then $\tilde{r}_2(x) > \mu_\varepsilon \geq r_1(x)$ and $r_2(x) \geq r_1(x)$ by definition of M_ε. Thus, we have for any $x \in \mathsf{M}$, $r^\star(x) = \min(r_1(x), \tilde{r}_2(x))$ where r_1 and $\tilde{r}_2$ are respectively 1 and L_2 Lipschitz-continuous on M with respect to $\|\cdot\|_2$. By observing that $2\min(r_1, \tilde{r}_2) = r_1 + \tilde{r}_2 - |r_1 - \tilde{r}_2|$, we set $\mathsf{L} = 1 + \mathsf{L}_2$ and thus obtain item (b) of Lemma 27. Finally, item (c) of Lemma 27 directly derives from item (b), since M is bounded. □

Note that the extra-assumption on $\mathfrak{g}^{-1}$ used in Lemma 27 is not directly ensured by self-concordance but can be proved when M is a polytope, as shown below.

Lemma 28. *Consider a polytope M defined by m constraints with $m > d$ such that $\mathsf{M} = \{x : \mathrm{A}x < b\}$, where $\mathrm{A} \in \mathbb{R}^{m\times d}$ is a full-rank matrix and $b \in \mathbb{R}^m$. We endow M with the Riemannian metric $\mathfrak{g}(x) = D^2\phi(x)$ where $\phi : \mathsf{M} \to \mathbb{R}$ is the logarithmic barrier given for any $x \in \mathsf{M}$ by $\phi(x) = -\sum_{i=1}^m \ln\left(b_i - \mathrm{A}_i^\top x\right)$. In particular, M and $\mathfrak{g}$ verify **A**1 and **A**2. Then, the function $r : \mathsf{M} \to (0, +\infty)$ defined by $r(x) = \|\mathfrak{g}^{-1}(x)\|_2$, for any $x \in \mathsf{M}$, is bounded from above.*

Proof. Consider such manifold M and metric $\mathfrak{g}$. We aim to show that the smallest eigenvalue of $\mathfrak{g}(x)$ is bounded from below for any $x \in \mathsf{M}$ by a constant $c > 0$, which does not depend on x, *i.e.*, for any $h \in \mathbb{R}^d$ and any $x \in \mathsf{M}$, $\mathfrak{g}(x)[h,h] \geq c\|h\|_2^2$.

Since A is full-ranked, $\mathrm{A}^\top\mathrm{A}$ is positive-definite. In particular, for any $h \in \mathbb{R}^d$, $(\mathrm{A}^\top\mathrm{A})[h,h] \geq \lambda_{\min}(\mathrm{A}^\top\mathrm{A})\|h\|_2^2$, where $\lambda_{\min}(\mathrm{A}^\top\mathrm{A}) > 0$ is the smallest eigenvalue of $\mathrm{A}^\top\mathrm{A}$. We recall that we have for any $x \in \mathsf{M}$, $\mathfrak{g}(x) = \mathrm{A}^\top S(x)^{-2}\mathrm{A}$, where $S(x) = \mathrm{Diag}(b_i - \mathrm{A}_i^\top x)_{i\in[m]}$. Let $i \in [m]$. The function $r_i : x \in \mathsf{M} \mapsto S(x)_{i,i}^{-2}$ has the following properties: (i) r_i is continuous on M, (ii) $r_i(x) > 0$ for any $x \in \mathsf{M}$ and (iii) $r(x) \to +\infty$ as $x \to \partial\mathsf{M}$. Thus, there exists $c_i > 0$ such that for any $x \in \mathsf{M}$, $r_i(x) \geq c_i$. We define $\tilde{c} = \min_{i\in[m]} c_i$ and we have for any $x \in \mathsf{M}$, $S(x)^{-2} \succeq \tilde{c}\mathrm{I}_d$. We now define $c = \tilde{c}\lambda_{\min}(\mathrm{A}^\top\mathrm{A})$ and we have for any $x \in \mathsf{M}$

$$\mathfrak{g}(x)[h,h] \geq \tilde{c}\cdot(\mathrm{A}^\top\mathrm{A})[h,h] \geq c\|h\|_2^2 .$$

In particular, $\mathfrak{g}(x)^{-1} \preceq (1/c)\mathrm{I}_d$, *i.e.*, $\|\mathfrak{g}(x)^{-1}\|_2 \leq (1/c)$, which concludes the proof. □

I.2 Markov kernels of Algorithm 1

Based on the model of $\bar{\mathsf{E}}_h$ defined in (24), we define the set $\bar{\mathsf{E}}_h^\Phi \subset \mathrm{T}^\star\mathsf{M}$, which will ensure that the maps derived from implicit integrators of n-BHMC are properly expressed

$$\bar{\mathsf{E}}_h^\Phi = (s \circ \mathrm{S}_{h/2})^{-1}(\mathrm{dom}_{\Phi_h} \cap \Phi_h^{-1}(\mathrm{dom}_{\Phi_h})) = (s \circ \mathrm{S}_{h/2})(\mathrm{dom}_{\Phi_h} \cap \Phi_h^{-1}(\mathrm{dom}_{\Phi_h})) . \qquad (53)$$

For any $(z, z') \in \bar{\mathsf{E}}_h^{\Phi} \times \mathrm{T}^\star\mathsf{M}$, we define the acceptance probability $\bar{a}(z'|z)$ to move from z to z' by

$$\bar{a}(z'|z) = a(z' \mid z)\mathbb{1}_{(s\circ \mathrm{S}_{h/2})(z)=(\Phi_h\circ\Phi_h\circ s\circ \mathrm{S}_{h/2})(z)} = a(z' \mid z)\mathbb{1}_{z=(\mathrm{R}_h^{\Phi}\circ \mathrm{R}_h^{\Phi})(z)} \ ,$$

where $a(z' \mid z)$ is the acceptance probability defined in (9). We denote by $\mathrm{Q}_n : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel of the (homogeneous) Markov chain $(x_n, p_n)_{n\in[N]}$ generated by Algorithm 1. We also denote by:

(a) $\mathrm{Q}_0 : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel referring to Step 1 (also Step 5) in Algorithm 1, defined in (25).

(b) $\mathrm{Q}_{n,1} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel referring to Step 2-3-4 in Algorithm 1.

We provide below details on Markov kernels Q_n and $\mathrm{Q}_{n,1}$.

Kernel $\mathrm{Q}_{n,1}$. This kernel is deterministic and corresponds to the *numerical* integration of the Hamiltonian up until time h. For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\mathrm{Q}_{n,1}(z, \mathrm{d}z') = \mathbb{1}_{\bar{\mathsf{E}}_h^{\Phi}}(z)(s_\#\mathrm{Q}_{n,2})(z, \mathrm{d}z') + \mathbb{1}_{(\bar{\mathsf{E}}_h^{\Phi})^c}(z)\delta_{s(z)}(\mathrm{d}z') \ , \tag{54}$$

where

$$\mathrm{Q}_{n,2}(z, \mathrm{d}z') = \bar{a}(\mathrm{R}_h^{\Phi}(z) \mid z)\delta_{\mathrm{R}_h^{\Phi}(z)}(\mathrm{d}z') + [1 - \bar{a}(\mathrm{R}_h^{\Phi}(z) \mid z)]\delta_z(\mathrm{d}z') \ . \tag{55}$$

Kernel Q_n. This kernel corresponds to one step of Algorithm 1 (*i.e.*, comprising Steps 1 to 5). For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\mathrm{Q}_n(z, \mathrm{d}z') = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} \mathrm{Q}_0(z, \mathrm{d}z_1)\mathrm{Q}_{n,1}(z_1, \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z') \ .$$

I.3 Proof of reversibility in Algorithm 3

Let $h > 0$. Using notation from **A**3, we recall definition of the set A_h introduced in Section 4.2

$$\mathsf{A}_h = \{z \in \mathrm{T}^\star\mathsf{M} \ : \ h < \min_{\tilde{z}\in \mathsf{B}_{\|\cdot\|_z}(z,1)} h_\star(\tilde{z})\} \subset \mathrm{dom}_{\Phi_h} \ . \tag{56}$$

We also recall that $\bar{\pi}$, as defined in (2), admits a density with respect to the product Lebesgue measure given for any $z = (x, p) \in \mathrm{T}^\star\mathsf{M}$ by

$$(\mathrm{d}\bar{\pi}/(\mathrm{d}x\mathrm{d}p))(x, p) = (1/Z)\exp[-(1/2)\|p\|^2_{\mathfrak{g}(x)^{-1}}]\det(\mathfrak{g}(x))^{-1/2}\exp[-V(x)] \ .$$

Since Algorithm 3 can be thought as a *restrictive* version of Algorithm 1, the Markov kernels from Tn-BHMC are similar to the kernels from n-BHMC, defined in Appendix I.2. In particular, the transition kernel corresponding to the Gaussian momentum update (Step 1 and 5 in Algorithm 3, Step 1 and 5 in Algorithm 1) is the same and is reversible (up to momentum reversal) with respect to $\bar{\pi}$ (see Lemma 17). Namely, we replace

(a) the set $\bar{\mathsf{E}}_h^{\Phi}$, defined in (53), by $\tilde{\mathsf{E}}_h^{\Phi}$,

(b) the kernels $\mathrm{Q}_{n,1}$ and $\mathrm{Q}_{n,2}$, respectively defined in (54) and (55), by Q_1 and Q_2,

where

$$\tilde{\mathsf{E}}_h^{\Phi} = (s\circ\mathrm{S}_{h/2})^{-1}(\mathsf{A}_h \cap \Phi_h^{-1}(\mathsf{A}_h)) = (s\circ\mathrm{S}_{h/2})(\mathsf{A}_h \cap \Phi_h^{-1}(\mathsf{A}_h)) \ , \tag{57}$$

$$\mathrm{Q}_1(z, \mathrm{d}z') = \mathbb{1}_{\tilde{\mathsf{E}}_h^{\Phi}}(z)(s_\#\mathrm{Q}_2)(z, \mathrm{d}z') + \mathbb{1}_{(\tilde{\mathsf{E}}_h^{\Phi})^c}(z)\delta_{s(z)}(\mathrm{d}z') \ , \tag{58}$$

$$\mathrm{Q}_2(z, \mathrm{d}z') = a(\mathrm{R}_h^{\Phi}(z) \mid z)\delta_{\mathrm{R}_h^{\Phi}}(\mathrm{d}z') + [1 - a(\mathrm{R}_h^{\Phi}(z) \mid z)]\delta_z(\mathrm{d}z') \ . \tag{59}$$

We denote by $\mathrm{Q} : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$, the transition kernel of the (homogeneous) Markov chain $(x_n, p_n)_{n\in[N]}$ generated by Algorithm 3. For any $(z, z') \in \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}$, we have

$$\mathrm{Q}(z, \mathrm{d}z') = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} \mathrm{Q}_0(z, \mathrm{d}z_1)\mathrm{Q}_1(z_1, \mathrm{d}z_2)\mathrm{Q}_0(z_2, \mathrm{d}z') \ . \tag{60}$$

We first turn to the reversibility up to momentum reversal of Q_1 with respect to $\bar{\pi}$. We start with the following lemma which is key to establish this result.

Lemma 29. *Assume* **A**1*,* **A**2*,* **A**3*. Then, for any compact set* $\mathsf{G} \subset \mathrm{T}^\star\mathsf{M}$*, for any* $g \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ *such that* $\mathrm{supp}(g) \subset \mathsf{G} \times \mathsf{G}$*, we have*

$$\int_{\mathsf{G}_h} g(z, \Phi_h(z))\mathrm{d}z = \int_{\mathsf{G}_h} g(\Phi_h(z), z)\mathrm{d}z \ ,$$

where $\mathsf{G}_h = \mathsf{F}_h \cap \mathsf{G} \cap \Phi_h^{-1}(\mathsf{G})$*,* $\mathsf{F}_h = \mathsf{A}_h \cap \Phi_h^{-1}(\mathsf{A}_h)$*,* A_h *being defined in* (56)*.*

Proof. Assume **A**1, **A**2, **A**3. Let $\mathsf{G} \subset \mathrm{T}^\star\mathsf{M}$ be a compact set, $g \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ such that $\mathrm{supp}(g) \subset \mathsf{G} \times \mathsf{G}$. We first define the sets $\mathsf{F}_h = \mathsf{A}_h \cap \Phi_h^{-1}(\mathsf{A}_h)$, $\mathsf{G}_h = \mathsf{F}_h \cap \mathsf{G} \cap \Phi_h^{-1}(\mathsf{G})$ and the integrals I' and J'

$$I' = \int_{\mathsf{G}_h} g(z, \Phi_h(z))\mathrm{d}z \ , \quad J' = \int_{\mathsf{G}_h} g(\Phi_h(z), z)\mathrm{d}z \ . \tag{61}$$

We recall the existence under **A**1 and **A**2 of $\mathtt{L} > 0$ and $\mathtt{M} > 0$ such that $r^\star$, given by **A**3 and defined in (52), is $\mathtt{L}$-Lipschitz-continuous on M with respect to $\|\cdot\|_2$ and bounded from above by $\mathtt{M}$ (see Lemma 27).

We define $r : \mathrm{T}^\star\mathsf{M} \to (0, +\infty)$ by $r(x,p) = r^\star(x)/\mathtt{m}$ for any $(x,p) \in \mathrm{T}^\star\mathsf{M}$, where $\mathtt{m} = \max\{1/\lambda, 4\mathtt{M}, 4\mathtt{L}\}$ and λ is given by **A**3. Using the properties of $r^\star$, it comes that

(a) r is $\mathtt{L}$-Lipschitz on $\mathrm{T}^\star\mathsf{M}$ with respect to $\|\cdot\|_2$,

(b) $r(x,p) \leq 1/(4\mathtt{L}C_x)$ for any $(x,p) \in \mathrm{T}^\star\mathsf{M}$, where C_x is defined in Lemma 12,

(c) $r \leq 1/4$, and

(d) $r \leq \lambda r^\star$.

Note that $\mathsf{G} \subset \cup_{z \in \mathsf{G}} \mathsf{B}_{\|\cdot\|_z}(z, r(z))$. Since G is a compact set, there exist $K \in \mathbb{N}$ and $(z_i)_{i \in [K]} \in \mathrm{T}^\star\mathsf{M}^K$ such that $\mathsf{G} \subset \bigcup_{i=1}^K \mathsf{B}_i$, where $\mathsf{B}_i = \mathsf{B}_{\|\cdot\|_{z_i}}(z_i, r(z_i))$.

We consider the sequence $\{\mathsf{V}_i\}_{i=1}^K$ constructed as follows: $\mathsf{V}_1 = \mathsf{G} \cap \mathsf{B}_1$ and for any $i \in \{2, \ldots, K\}$, $\mathsf{V}_i = (\mathsf{G} \cap \mathsf{B}_i) \cap (\cup_{j=1}^{i-1}\mathsf{V}_j)^{\mathrm{c}}$. Then, we have that, for any $i \in \{1, \ldots, N\}$, $\cup_{j=1}^i \mathsf{V}_j = \mathsf{G} \cap (\cup_{j=1}^i \mathsf{B}_j)$, and that for any $i_1, i_2 \in \{1, \ldots, K\}$, $\mathsf{V}_{i_1} \cap \mathsf{V}_{i_2} = \emptyset$ if $i_1 \neq i_2$. Therefore, we get that $\mathsf{G} = \sqcup_{i=1}^K \mathsf{V}_i$ and $\Phi_h^{-1}(\mathsf{G}) = \sqcup_{i=1}^K \Phi_h^{-1}(\mathsf{V}_i)$. In particular, for any $\mathsf{A} \in \mathcal{B}(\mathrm{T}^\star\mathsf{M})$ and any $\zeta \in \mathrm{C}_c(\mathrm{T}^\star\mathsf{M}, \mathbb{R})$

$$\int_{\mathsf{G} \cap \mathsf{A}} \zeta(z)\mathrm{d}z = \sum_{i=1}^K \int_{\mathsf{V}_i \cap \mathsf{A}} \zeta(z)\mathrm{d}z \ , \tag{62}$$

and for any $\tilde{\mathsf{A}} \in \mathcal{B}(\mathrm{T}^\star\mathsf{M})$ and any $\tilde{\zeta} \in \mathrm{C}_c(\mathrm{T}^\star\mathsf{M}, \mathbb{R})$

$$\int_{\Phi_h^{-1}(\mathsf{G}) \cap \tilde{\mathsf{A}}} \tilde{\zeta}(z)\mathrm{d}z = \sum_{i=1}^K \int_{\Phi_h^{-1}(\mathsf{V}_i) \cap \tilde{\mathsf{A}}} \tilde{\zeta}(z)\mathrm{d}z \ . \tag{63}$$

Using (62) and (61), we obtain $I' = \sum_{i=1}^K I_i'$, where $I_i' = \int_{\mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h} g(z, \Phi_h(z))\mathrm{d}z$ for any $i \in [K]$. We are now going to show that, for any $i \in [K]$

$$I_i' = \int_{\Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h} g(\Phi_h(z), z)\mathrm{d}z \ . \tag{64}$$

Let $i \in [K]$. We proceed by making the following case disjunction:

(a) Either $\mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h = \emptyset$, and then $I_i' = 0$. We prove by contradiction in this case that $\Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h = \emptyset$. Assume that there exists $z \in \Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h$. By definition of F_h, $z \in \mathsf{A}_h$, $\Phi_h(z) \in \mathsf{A}_h$, and we get that $\Phi_h(\Phi_h(z)) = z$ using **A**3. Hence, we have:

- $\Phi_h(z) \in \mathsf{V}_i$,
- $\Phi_h(\Phi_h(z)) = z \in \mathsf{G}$,
- $\Phi_h(z) \in \mathsf{A}_h$,
- $\Phi_h(\Phi_h(z)) = z \in \mathsf{A}_h$.

Therefore, we get $\Phi_h(z) \in \mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h = \emptyset$, which is absurd. Finally, (64) holds since

$$I_i' = 0 = \int_{\Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h} g(\Phi_h(z), z)\mathrm{d}z \ .$$

(b) Either, there exists some $\tilde{z} \in \mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G})$ such that $\tilde{z} \in \mathsf{F}_h$. In particular, $\tilde{z} \in \mathsf{B}_i$, and thus $\|\tilde{z} - z_i\|_{z_i} < r(z_i) < 1$. In this case, by combining **A3** with Lemma 9-(c), we have for any $z' \in \mathrm{T}^\star\mathsf{M}$

$$\|z'\|_{\tilde{z}} \leq (1 - \|\tilde{z} - z_i\|_{z_i})^{-1}\|z'\|_{z_i} < (1 - r(z_i))^{-1}\|z'\|_{z_i} \; .$$

By considering $z' = z_i - \tilde{z}$, it comes that

$$\|z_i - \tilde{z}\|_{\tilde{z}} < (1 - r(z_i))^{-1}\|\tilde{z} - z_i\|_{z_i} < (1 - r(z_i))^{-1}r(z_i) \; .$$

By combining properties (a) and (b) of r with Lemma 12, we have the following upper bound of $r(z_i)$

$$r(z_i) \leq r(\tilde{z}) + \mathsf{L}\|\tilde{z} - z_i\|_2 \leq r(\tilde{z}) + \mathsf{L}C_{x_i}\|\tilde{z} - z_i\|_{z_i} < r(\tilde{z}) + \mathsf{L}C_{x_i}r(z_i) < r(\tilde{z}) + 1/4 \; ,$$

and thus

$$\|z_i - \tilde{z}\|_{\tilde{z}} < (1 - r(\tilde{z}) - 1/4)^{-1}(r(\tilde{z}) + 1/4) < 1 \; ,$$

using property (c) of r. Hence, $z_i \in \mathsf{B}_{\|\cdot\|_{\tilde{z}}}(\tilde{z}, 1)$. Moreover, $\tilde{z} \in \mathsf{F}_h \subset \mathsf{A}_h$, and then it comes that $h < h_\star(z_i)$. Therefore, by combining property (d) of r with **A3**, we have

(i) $\mathsf{V}_i \subset \mathsf{B}_i \subset \mathsf{B}_{\|\cdot\|_{z_i}}(z_i, \lambda r^\star(z_i)) \subset \mathrm{dom}_{\Phi_h}$,
(ii) the restriction of Φ_h to V_i is a C^1-diffeomorphism such that $\Phi_h \circ \Phi_h = \mathrm{Id}$.

This last result provides proper assumptions on Φ_h to operate a change of variable in I'_i. We now define $\mathsf{G}_{1,h} = \Phi_h(\mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h)$ and $\mathsf{G}_{2,h} = \Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h$, and prove that $\mathsf{G}_{1,h} = \mathsf{G}_{2,h}$ in two steps.

(i) We first prove that $\mathsf{G}_{1,h} \subset \mathsf{G}_{2,h}$. Let $z \in \mathsf{G}_{1,h}$. Then, there exists $z' \in \mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h$ such that $z = \Phi_h(z')$. Since Φ_h is an involution on V_i, we have $\Phi_h(z) = z'$. Since $z' \in \mathsf{V}_i \cap \mathsf{A}_h$, it comes that $z \in \Phi_h^{-1}(\mathsf{V}_i) \cap \Phi_h^{-1}(\mathsf{A}_h)$. Moreover, $z' \in \Phi_h^{-1}(\mathsf{G}) \cap \Phi_h^{-1}(\mathsf{A}_h)$, and thus, $z \in \mathsf{G} \cap \mathsf{A}_h$. Then, we have $z \in \mathsf{G}_{2,h}$, which proves this first result.
(ii) We now prove that $\mathsf{G}_{2,h} \subset \mathsf{G}_{1,h}$. Let $z \in \mathsf{G}_{2,h}$. In particular, $z \in \mathsf{G}$. Then, there exists $j \in [K]$ such that $z \in \mathsf{V}_j$. Therefore, $z \in \mathsf{B}_j$ and thus $\|z - z_j\|_{z_j} < r(z_j) < 1$. Since $z \in \mathsf{A}_h$, we obtain that $h < h_\star(z_j)$ with the same computations as those written above. In particular, this proves with **A3** that Φ_h is an involution on V_j. Then, $z = \Phi_h(\Phi_h(z))$, with $\Phi_h(z) \in \mathsf{V}_i \cap \mathsf{A}_h$ and $\Phi_h(z) \in \Phi_h^{-1}(\mathsf{G}) \cap \Phi_h^{-1}(\mathsf{A}_h)$, since $z \in \mathsf{G} \cap \mathsf{A}_h$. Therefore, we have $z \in \mathsf{G}_{1,h}$, which proves this last result.

Given the fact that Φ_h is an involution on V_i, we operate the change of variable $z \mapsto \Phi_h(z)$ in I'_i and obtain

$$I'_i = \int_{\Phi_h(\mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h)} g(\Phi_h(z), z)\mathrm{d}z = \int_{\Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h} g(\Phi_h(z), z)\mathrm{d}z \; ,$$

which gives (64).

Therefore, combining (62), (63) and (64), we get

$$I' = \sum_{i=1}^K \int_{\mathsf{V}_i \cap \Phi_h^{-1}(\mathsf{G}) \cap \mathsf{F}_h} g(z, \Phi_h(z))\mathrm{d}z = \sum_{i=1}^K \int_{\Phi_h^{-1}(\mathsf{V}_i) \cap \mathsf{G} \cap \mathsf{F}_h} g(\Phi_h(z), z)\mathrm{d}z = J' \; ,$$

which concludes the proof. □

We are now ready to establish the reversibility up to momentum reversal of Q_1, defined in (58), with respect to $\bar{\pi}$.

Lemma 30. *Assume* **A1***,* **A2***,* **A3***. Then, for any $h > 0$, the Markov kernel Q_1 with step-size h, defined in* (58)*, is reversible up to momentum reversal with respect to $\bar{\pi}$.*

Proof. Assume **A1**, **A2**, **A3**. Let $h > 0$. We recall that the kernels Q_1 and Q_2 are respectively defined in (58) and (59). We define the transition kernel $\mathrm{Q}_3 : \mathrm{T}^\star\mathsf{M} \times \mathcal{B}(\mathrm{T}^\star\mathsf{M}) \to [0, 1]$ by

$$\mathrm{Q}_3(z, \mathrm{d}z') = \mathbb{1}_{\bar{\mathsf{E}}_h^\Phi}(z)\mathrm{Q}_2(z, \mathrm{d}z') + \mathbb{1}_{(\bar{\mathsf{E}}_h^\Phi)^c}(z)\delta_z(\mathrm{d}z') \; , \tag{65}$$

such that $\mathrm{Q}_1 = s_\#\mathrm{Q}_3$. The rest of the proof is divided into two parts. First, we prove that Q_3 is reversible with respect to $\bar{\pi}$. Then, we prove that Q_1 is reversible up to momentum reversal with respect to $\bar{\pi}$.

(a) Let $f \in \mathrm{C}(\mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M}, \mathbb{R})$ with compact support. We consider a compact set K with respect to the topology induced by the set $\{\mathsf{B}_{\|\cdot\|_z}(z,r), z \in \mathrm{T}^\star\mathsf{M}, r \in (0,1)\}$ such that $\mathrm{supp}(f) \subset \mathsf{K} \times \mathsf{K}$. According to Definition 4, we aim to show that

$$\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,z')\mathrm{Q}_3(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) = \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z',z)\mathrm{Q}_3(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) . \tag{66}$$

We denote by I the left integral of (66). By combining (59) and (65), we have $I = I_1 + I_2 + I_3$ where

$$I_1 = \int_{\tilde{\mathsf{E}}_h^\Phi} \bar{\pi}(z)a(\mathrm{R}_h^\Phi(z) \mid z)f(z,\mathrm{R}_h^\Phi(z))\mathrm{d}z ,$$
$$I_2 = \int_{\tilde{\mathsf{E}}_h^\Phi} \bar{\pi}(z)[1 - a(\mathrm{R}_h^\Phi(z) \mid z)]f(z,z)\mathrm{d}z ,$$
$$I_3 = \int_{(\tilde{\mathsf{E}}_h^\Phi)^{\mathrm{c}}} \bar{\pi}(z)f(z,z)\mathrm{d}z .$$

Since $\mathrm{supp}(f) \subset \mathsf{K} \times \mathsf{K}$, we have for any $z \in \mathsf{K}^{\mathrm{c}}$, $f(z,\cdot) = 0$ and $f(\cdot,z) = 0$. Note also that for any $z \in ((\mathrm{R}_h^\Phi)^{-1}(\mathsf{K}))^{\mathrm{c}}$, $\mathrm{R}_h^\Phi(z) \notin \mathsf{K}$, and thus $f(\mathrm{R}_h^\Phi(z),\cdot) = 0$ and $f(\cdot,\mathrm{R}_h^\Phi(z)) = 0$. By combining these preliminary results with (9), we get

$$\begin{aligned} I_1 &= \int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}\cap(\mathrm{R}_h^\Phi)^{-1}(\mathsf{K})} \bar{\pi}(z)a(\mathrm{R}_h^\Phi(z) \mid z)f(z,\mathrm{R}_h^\Phi(z))\mathrm{d}z \\ &= (1/Z)\int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}\cap(\mathrm{R}_h^\Phi)^{-1}(\mathsf{K})} \min\{e^{-H(z)}, e^{-H(\mathrm{R}_h^\Phi(z))}\}f(z,\mathrm{R}_h^\Phi(z))\mathrm{d}z \\ I_2 &= \int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}} \bar{\pi}(z)[1 - a(\mathrm{R}_h^\Phi(z) \mid z)]f(z,z)\mathrm{d}z , \\ I_3 &= \int_{(\tilde{\mathsf{E}}_h^\Phi)^{\mathrm{c}}} \bar{\pi}(z)f(z,z)\mathrm{d}z . \end{aligned}$$

We denote by J the right integral of (66). By symmetry, we have $J = J_1 + J_2 + J_3$ where

$$\begin{aligned} J_1 &= \int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}\cap(\mathrm{R}_h^\Phi)^{-1}(\mathsf{K})} \bar{\pi}(z)a(\mathrm{R}_h^\Phi(z) \mid z)f(\mathrm{R}_h^\Phi(z),z)\mathrm{d}z \\ &= (1/Z)\int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}\cap(\mathrm{R}_h^\Phi)^{-1}(\mathsf{K})} \min\{e^{-H(z)}, e^{-H(\mathrm{R}_h^\Phi(z))}\}f(\mathrm{R}_h^\Phi(z),z)\mathrm{d}z , \\ J_2 &= \int_{\tilde{\mathsf{E}}_h^\Phi\cap\mathsf{K}} \bar{\pi}(z)[1 - a(\mathrm{R}_h^\Phi(z) \mid z)]f(z,z)\mathrm{d}z , \\ J_3 &= \int_{(\tilde{\mathsf{E}}_h^\Phi)^{\mathrm{c}}} \bar{\pi}(z)f(z,z)\mathrm{d}z . \end{aligned}$$

We directly have $I_2 = J_2$ and $I_3 = J_3$. Let us now prove that $I_1 = J_1$.

We recall that $s \circ \mathrm{S}_{h/2}$ is a symplectic C^1-diffeomorphism (see Section 3.1) and $\tilde{\mathsf{E}}_h^\Phi = (s \circ \mathrm{S}_{h/2})(\mathsf{F}_h)$ where $\mathsf{F}_h = \mathsf{A}_h \cap \Phi_h^{-1}(\mathsf{A}_h)$, using (57). We define $\mathsf{K}_h = \mathsf{F}_h \cap (s\circ\mathrm{S}_{h/2})(\mathsf{K}) \cap \Phi_h^{-1}((s\circ\mathrm{S}_{h/2})(\mathsf{K}))$, such that $\tilde{\mathsf{E}}_h^\Phi \cap \mathsf{K} \cap (\mathrm{R}_h^\Phi)^{-1}(\mathsf{K}) = (s \circ \mathrm{S}_{h/2})(\mathsf{K}_h)$, and we operate the change of variable $z \mapsto (s \circ \mathrm{S}_{h/2})(z)$ in I_1 and J_1

$$I_1 = (1/Z)\int_{\mathsf{K}_h} \min\{e^{-H((s\circ\mathrm{S}_{h/2})(z))}, e^{-H((s\circ\mathrm{S}_{h/2}\circ\Phi_h)(z))}\}f((s \circ \mathrm{S}_{h/2})(z), (s \circ \mathrm{S}_{h/2} \circ \Phi_h)(z))\mathrm{d}z ,$$
$$J_1 = (1/Z)\int_{\mathsf{K}_h} \min\{e^{-H((s\circ\mathrm{S}_{h/2})(z))}, e^{-H((s\circ\mathrm{S}_{h/2}\circ\Phi_h)(z))}\}f((s \circ \mathrm{S}_{h/2} \circ \Phi_h)(z), (s \circ \mathrm{S}_{h/2})(z))\mathrm{d}z .$$

We now define the map $g : \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M} \to \mathbb{R}$ and the set $\mathsf{G} \subset \mathrm{T}^\star\mathsf{M}$ by

$$g(z,z') = \min\{e^{-H((s\circ\mathrm{S}_{h/2})(z))}, e^{-H((s\circ\mathrm{S}_{h/2})(z'))}\}f((s \circ \mathrm{S}_{h/2})(z), (s \circ \mathrm{S}_{h/2})(z')) ,$$
$$\mathsf{G} = (s \circ \mathrm{S}_{h/2})(\mathsf{K}) .$$

Note that G is a compact set of $\mathrm{T}^\star\mathsf{M}$ by continuity of $s \circ \mathrm{S}_{h/2}$ and g is a continuous function by continuity of H and $s \circ \mathrm{S}_{h/2}$. Then, we obtain $I_1 = J_1$, by applying Lemma 29 with g and G. Finally, we obtain $I = J$ and thus prove (66) for any continuous function f with compact support.

(b) Let $f : \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M} \to \mathbb{R}$ be a continuous function with compact support. We have

$$\begin{aligned} &\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,z')\mathrm{Q}_1(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\ &= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,z')(s_{\#}\mathrm{Q}_3)(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\ &= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,s(z'))\mathrm{Q}_3(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z'\text{)} \\ &= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z',s(z))\mathrm{Q}_3(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(using (66))} \\ &= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'),s(z))(s_{\#}\mathrm{Q}_3)(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z'\text{)} \\ &= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'),s(z))\mathrm{Q}_1(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) , \end{aligned}$$

which concludes the proof.

□

We are now ready to prove Theorem 3, which states that Q is reversible up to momentum reversal with respect to $\bar{\pi}$.

Proof of Theorem 3. Assume **A**1, **A**2, **A**3. This proof is very similar to the proof of Theorem 15 (see Appendix E.2). Let $f : \mathrm{T}^\star\mathsf{M} \times \mathrm{T}^\star\mathsf{M} \to \mathbb{R}$ be a continuous function with compact support. We have

$$\begin{aligned}
&\int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(z,z')\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(z,z')\mathrm{Q}_0(z,\mathrm{d}z_1)\mathrm{Q}_1(z_1,\mathrm{d}z_2)\mathrm{Q}_0(z_2,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(see (60))} \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(s(z_1),z')\mathrm{Q}_0(z,\mathrm{d}z_1)\mathrm{Q}_1(s(z),\mathrm{d}z_2)\mathrm{Q}_0(z_2,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 17)} \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(s(z_1),z')\mathrm{Q}_0(s(z),\mathrm{d}z_1)\mathrm{Q}_1(z,\mathrm{d}z_2)\mathrm{Q}_0(z_2,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z) \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(s(z_1),z')\mathrm{Q}_0(z_2,\mathrm{d}z_1)\mathrm{Q}_1(z,\mathrm{d}z_2)\mathrm{Q}_0(s(z),\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 30)} \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(s(z_1),z')\mathrm{Q}_0(z_2,\mathrm{d}z_1)\mathrm{Q}_1(s(z),\mathrm{d}z_2)\mathrm{Q}_0(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(momentum reversal on } z) \\
&= \int_{(\mathrm{T}^\star\mathsf{M})^4} f(s(z_1),s(z))\mathrm{Q}_0(z_2,\mathrm{d}z_1)\mathrm{Q}_1(z',\mathrm{d}z_2)\mathrm{Q}_0(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z) && \text{(Lemma 17)} \\
&= \int_{\mathrm{T}^\star\mathsf{M}\times\mathrm{T}^\star\mathsf{M}} f(s(z'),s(z))\mathrm{Q}(z,\mathrm{d}z')\bar{\pi}(\mathrm{d}z)\ .
\end{aligned}$$

Moreover, $s_\#\bar{\pi} = \bar{\pi}$. Hence, by combining Definition 4 and Lemma 5, we obtain the result of Theorem 3. □

J Comparison with Kook et al. (2022a).

In this section, we provide a precise comparison between our work and the results stated in Kook et al. (2022a). We first dwell on the differences related to the general setting and the algorithms, and then explain how our methodology may solve the limitations of Kook et al. (2022a).

J.1 General framework

Given a convex body $\mathsf{K} \subset \mathbb{R}^d$ and a function $c : \mathbb{R}^d \to \mathbb{R}^m$, consider

$$\mathsf{M} = \{x \in \mathsf{K} : c(x) = 0\}\ . \tag{67}$$

In their paper, Kook et al. (2022a) aim at sampling from a target distribution π, with density given for $x \in \mathsf{M}$ by

$$\mathrm{d}\pi(x)/\mathrm{d}x = \exp[-V(x)]/Z\ ,$$

where $V \in \mathrm{C}^2(\mathsf{M},\mathbb{R})$, $Z = \int_{\mathsf{M}} \exp[-V(x)]\mathrm{d}x$. They make the following assumptions:

(a) K is provided with a self-concordant barrier ϕ.

(b) The differential of c at x, $Dc(x)$, is full-ranked for any $x \in \mathsf{M}$.

Although this setting might be more general than ours (consider for instance the case where K is a polytope and c is a non-trivial function), Kook et al. (2022a) focus on the case

$$\mathsf{K} = \{x \in \mathbb{R}^d : \ell \le x \le u\}, \quad c(x) = Ax - b \tag{68}$$

where $\ell, u \in \mathbb{R}^d$ with $\ell < u$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{d\times m}$. Combined with assumptions (a) and (b), A must have independent rows and the Hessian of ϕ at x given by $\mathfrak{g}(x) = D^2\phi(x)$ is a diagonal positive matrix for any x. Moreover, M defined by (67)-(68) is a polytope. However, not any polytope can be rewritten in a manner akin to (68), and therefore, the setting chosen by Kook et al. (2022a) can actually be considered as *a special case of our setting*, see Section 2.

Then, they consider the extended state-space $\bar{\mathsf{M}} = \{(x,p) \in \mathbb{R}^d \times \mathbb{R}^d : x \in \mathsf{M}, p \in \mathrm{Ker}(Dc(x))\}$ and define on $\bar{\mathsf{M}}$ a *constrained* Hamiltonian H, given by

$$\begin{aligned}
H(x,p) &= \bar{H}(x,p) + \lambda(x,p)^\top c(x)\ , \\
\text{with } \bar{H}(x,p) &= V(x) + \frac{1}{2}\log \operatorname{pdet} M(x) + \frac{1}{2}p^\top M(x)^\dagger p\ , \\
\text{and } M(x) &= Q(x)^\top \mathfrak{g}(x) Q(x)\ ,
\end{aligned}$$

where Q is the orthogonal projection into $\mathrm{Ker}(Dc(x))$, implying that M is semi-definite positive, and $\lambda(x,p)$ is a Lagrangian term given by

$$\lambda(x,p) = (Dc(x)Dc(x)^\top)^{-1}\{D^2c(x)[p, \mathrm{d}x/\mathrm{d}t] - Dc(x)\partial\bar{H}(x,p)/\partial x\} \ .$$

Remark that λ simplifies when c is chosen as in (68), but may not be well suited for higher order constraints. Then, Kook et al. (2022a) provide simplifications of H for the setting (68) based on formulas for $\mathrm{pdet}\, M$ and $M^\dagger$.

Finally, they consider the joint distribution $\bar{\pi}$ given for any $(x,p) \in \bar{\mathsf{M}}$ by

$$\mathrm{d}\bar{\pi} = (1/\bar{Z})\exp[-H(x,p)]\mathrm{d}x\mathrm{d}p,$$

where $\bar{Z} = \int_{\bar{\mathsf{M}}} \exp[-H(x,p)]\mathrm{d}x\mathrm{d}p$, for which the first marginal is π. They naturally propose to sample from $\bar{\pi}$ by implementing a version of RMHMC relying on the *constrained* Hamiltonian H, which they call CRHMC.

Similarly to our approach, they consider two cases.

(a) First, Kook et al. (2022a) assume that the Hamiltonian dynamics given by H can be explicitly computed in continuous time (which is not the case in practice). They present the continuous version of RMHMC which incorporates the simplified *constrained* Hamiltonian H, see Algorithm 1. We refer to this algorithm as c-CRHMC. This algorithm is identical to our algorithm c-BHMC provided in Appendix D, apart from the involution checking step (see Line 4 in Algorithm 2). They state in Theorem 6 that the Markov kernel corresponding to one iteration of c-CRHMC satisfies detailed balance with respect to $\bar{\pi}$, thus ensuring that it preserves $\bar{\pi}$.

(b) Then, Kook et al. (2022a) provide a practical implementation of CRHMC, which is similar to n-BHMC, based on a time-discretization of the Hamiltonian dynamics for a given time-step $h > 0$. They first split the simplified *constrained* Hamiltonian H into H_1 and H_2, by leveraging the non-separable aspect of H in H_2, as we do in Section 3.1. Then, they propose to use the first-order Euler method to solve the discretized ODE associated to H_1, as we also suggest, and the *Implicit Midpoint Integrator* (IMI) to solve the discretized ODE associated with H_2, while we implement the *Generalized Leapfrog Integrator* (GLI). We insist on the fact that these two second-order methods share the same theoretical properties (Hairer et al., 2006). Since IMI is implicit, they propose to approximate it with a numerical integrator Φ_h, which is computed with a fixed-point method (as we do) detailed in Algorithm 3. Finally, they implement CRHMC which contains the same steps as n-BHMC *apart from the involution checking step* (see Line 8 in Algorithm 1): (i) refreshing the momentum with a symplectic scheme (Step 1 in both algorithms), (ii) solving the discretized ODE associated to H in three steps according to the splitting given by H_1 and H_2 (Step 2 in both algorithms), (iii) applying a Metropolis-Hastings filter (Step 3 in both algorithms), and (iv) operating a final momentum reversal (Step 4 in n-BHMC). They state in Theorem 8 that the Markov kernel corresponding to one iteration of CRHMC satisfies detailed balance with respect to $\bar{\pi}$, thus ensuring that it preserves $\bar{\pi}$.

J.2 Theoretical gaps in the reversibility of CRHMC Kook et al. (2022a)

First, we call into question the proof of the reversibility of c-CRHMC stated in (Kook et al., 2022a, Theorem 6). Indeed, Kook et al. (2022a) implicitly assume that the time-continuous Hamiltonian dynamics have a solution at any time, but do not provide any proof of this result. We emphasize here that this result is not trivial and possibly may not hold due to the pathological behaviour of the barrier at the border of M, see Proposition 14. In contrast, the involution checking step implemented in c-BHMC verifies this condition for the *forward* and the *reversed* dynamics after momentum reversal, which then allows us to derive the reversibility of c-BHMC, see Theorem 15.

Finally, we report some gaps in the proof of reversibility of CRHMC provided in (Kook et al., 2022a, Theorem 8). Indeed, while Kook et al. (2022a) claim that (Hairer et al., 2006, Theorem VI.3.5.) applies to CRHMC which contains the *numerical* integrator detailed in (Kook et al., 2022a, Algorithm 3), in fact (Hairer et al., 2006, Theorem VI.3.5.) only applies to the *implicit* integrator (IMI). This is a fundamental limitation of their theory. This theoretical confusion thus leads to numerical errors, as we detail in Section 6. In contrast, we present in Section 4 theoretical results on the implicit integrator (GLI), which allows us to derive reasonable assumptions on the numerical integrator Φ_h to obtain reversibility of n-BHMC in Theorem 3. Contrary to Kook et al. (2022a), our proof relies on the properties of Φ_h and highlights the critical role of the involution checking step.

K Experimental details

The numerical experiments presented in Section 6 are based on the MATLAB implementation of CRHMC provided by Kook et al. (2022a). We adapt this code for n-BHMC for sake of fairness. In particular, our implementation differs from CRHMC by the use of the Generalized Leapfrog integrator and the “involution checking” mechanism.

Details on experiments with synthetic data. In these experiments, the hypercube refers to the set $[-1/2, 1/2]^d$ where $d \in \{5, 10\}$. We recall that the ground truth on the quantities we estimate is given by the Metropolis Adjusted Langevin Algorithm (MALA) (Roberts & Stramer, 2002) for the hypercube and the Independent Metropolis Hastings (IMH) sampler (Liu, 1996) for the simplex. We now give details on how the parameters of these two algorithms are chosen. For MALA, we use a constant step-size $h = 0.05$ for both dimensions, which results in an average acceptance probability equal to 0.55 for $d = 5$ and 0.44 for $d = 10$. For IMH, we use as proposal the uniform distribution, which is simply a Dirichlet distribution, resulting in an average acceptance probability of order 0.36. We recall that we use an adaptive step-size h in CRHMC and n-BHMC such that we obtain an average acceptance probability of order 0.5 in the MH filter. We now discuss the setting of the tolerance parameter η in n-BHMC. We choose η so that (i) the step-size h is at least greater than 10^{-2} over the iterations of the algorithm and (ii) the average acceptance probability is roundly equal to 0.5. This heuristic results in defining (a) for the hypercube, $\eta = 5$ if $d = 5$ and $\eta = 10$ if $d = 10$ and (b) for the simplex, $\eta = 10$ if $d = 5$ and $\eta = 200$ if $d = 10$.

Details on experiments with real-world data. For these experiments, we consider the exact same setting as in Kook et al. (2022a). In particular, we use the same adaptation strategies to update the step-size and the barrier functions.

Influence of the norm in (8). The norm chosen for the “involution checking step” is arbitrary and many options are available. In particular, one could design $\|z^{(0)} - \Phi_h(z^{(1)})\|_2 > \eta$, and thus pick the Euclidean norm in (8). However, while this approach should theoretically also reduce the bias in the method, we remark that the obtained Markov chains have very poor mixing time. This is due to the fact that in practice the update on the momentum is of order $\|\mathfrak{g}(x)\|_2^{1/2}$ while the update on the position is of order $\|\mathfrak{g}(x)^{-1}\|_2^{1/2}$. Hence the Euclidean norm is ill-suited for the “involution checking step” and the proposal states are rejected a lot more near the boundaries, see Figure 3.

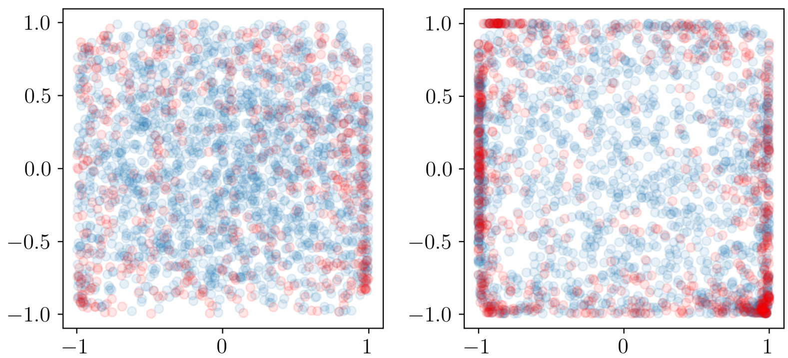

Figure 3: Outputs of n-BHMC after 25k iterations to sample from the uniform distribution over $[-1, 1]^2$ with $\eta = 10^{-3}$, $h = 0.8$, using in (8) the “self-concordant” norm (left) or the Euclidean norm (right). Red samples are rejected and blue samples are accepted.