

THE RANDOM BATCH METHOD FOR N -BODY QUANTUM DYNAMICS

FRANÇOIS GOLSE, SHI JIN, AND THIERRY PAUL

ABSTRACT. This paper discusses a numerical method for computing the evolution of large interacting system of quantum particles. The idea of the random batch method is to replace the total interaction of each particle with the $N - 1$ other particles by the interaction with $p \ll N$ particles chosen at random at each time step, multiplied by $(N - 1)/p$. This reduces the computational cost of computing the interaction potential per time step from $O(N^2)$ to $O(N)$. For simplicity, we consider only in this work the case $p = 1$ — in other words, we assume that N is even, and that at each time step, the N particles are organized in $N/2$ pairs, with a random reshuffling of the pairs at the beginning of each time step. We obtain a convergence estimate for the Wigner transform of the single-particle reduced density matrix of the particle system at time t that is both uniform in $N > 1$ and independent of the Planck constant \hbar . The key idea is to use a new type of distance on the set of quantum states that is reminiscent of the Wasserstein distance of exponent 1 (or Monge-Kantorovich-Rubinstein distance) on the set of Borel probability measures on \mathbf{R}^d used in the context of optimal transport.

1. INTRODUCTION

Consider the quantum Hamiltonian for N identical particles with unit mass located at positions $x_1, \dots, x_N \in \mathbf{R}^d$:

$$(1) \quad \mathcal{H}_N := \sum_{m=1}^N -\frac{1}{2}\hbar^2 \Delta_{x_m} + \frac{1}{N-1} \sum_{1 \leq l < n \leq N} V(x_l - x_n),$$

where \hbar is the reduced Planck constant. The N - particles in this system interact via a binary (real-valued) potential V assumed to be even, bounded and sufficiently regular (at least of class $C^{1,1}$ on \mathbf{R}^d). The coupling constant $\frac{1}{N-1}$ is chosen in order to balance the summations in the kinetic energy (involving N terms) and in the potential energy (involving $\frac{1}{2}N(N-1)$ terms).

We seek to compute the solution $\Psi_N \equiv \Psi_N(t, x_1, \dots, x_N) \in \mathbf{C}$ of the Schrödinger equation

$$(2) \quad i\hbar \partial_t \Psi_N(t, x_1, \dots, x_N) = \mathcal{H}_N \Psi_N(t, x_1, \dots, x_N), \quad \Psi_N|_{t=0} = \Psi_N^{in}$$

where $t \geq 0$ is the time while $x_m \in \mathbf{R}^d$ is the position of the m th particle. When solving (2), the computation is exceedingly expensive due to the smallness of \hbar which demands small time steps Δt and small mesh sizes of order \hbar for the convergence of the numerical scheme, due to the oscillation in the wave function Ψ_N with frequency of order $1/\hbar$ (see [2, 18]). On top of this, any numerical scheme

1991 *Mathematics Subject Classification.* 82C10, 82C22 (65M75).

Key words and phrases. Time-dependent Schrödinger equations, Random batch method, Mean-field limit, Wasserstein distance.

for (2) requires computing, at each time step, the sum of the interaction potential for each particle pair in the N -particle system, i.e. the sum of $\frac{1}{2}N(N-1)$ terms. For large values of N , the cost of this computation, which is of order $O(N^2)$, may become significant at each time step. The purpose of the Random Batch Method (RBM) described below is precisely to reduce significantly the computational cost of computing the interacting potential from $O(N^2)$ to $O(N)$.

Throughout this paper, we assume for simplicity that $N \geq 2$ is an even integer. Let $\sigma_1, \sigma_2, \dots, \sigma_j, \dots$ be a random sequence of mutually independent permutations distributed uniformly in \mathfrak{S}_N . Each permutation $\sigma \in \mathfrak{S}_N$ defines a partition of $\{1, \dots, N\}$ into $N/2$ batches of two indices (pairs) as follows:

$$\{1, \dots, N\} = \bigsqcup_{k=1}^{N/2} \{\sigma(2k-1), \sigma(2k)\}.$$

Pick a time step $\Delta t > 0$, set

$$\mathbf{T}_t(l, n) := \begin{cases} 1 & \text{if } \{l, n\} = \{\sigma_{[\frac{t}{\Delta t}]+1}(2k-1), \sigma_{[\frac{t}{\Delta t}]+1}(2k)\} \text{ for some } k = 1, \dots, \frac{N}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

and consider the time-dependent Hamiltonian

$$(3) \quad \mathcal{H}_N(t) := \sum_{m=1}^N -\frac{1}{2}\hbar^2 \Delta_{x_m} + \sum_{1 \leq l < n \leq N} \mathbf{T}_t(l, n) V(x_l - x_n).$$

In other words, at each time step, the particle labels $m = 1, \dots, N$ are reshuffled randomly, then grouped pairwise, and the potential applied to the m th particle by the system of $N-1$ other particles is replaced with the interaction potential of that particle with the other — *only one in this case* — particle in the same group (batch).

The motivation of the RBM is that the computation of the solution $\tilde{\Psi}_N \in \mathbf{C}$ of the time-dependent, random batch Schrödinger equation

$$(4) \quad i\hbar \partial_t \tilde{\Psi}_N(t, x_1, \dots, x_N) = \mathcal{H}_N(t) \tilde{\Psi}_N(t, x_1, \dots, x_N), \quad \tilde{\Psi}_N|_{t=0} = \tilde{\Psi}_N^{in}$$

is much less costly than computing the solution Ψ_N of the N -body Schrödinger (2) for large values of N . Clearly, for each time step the cost of computing the interaction potential is reduced from $O(N^2)$ to $O(N)$. We remark that the computational cost of reshuffling the N labels is $O(N)$ by Durstenfeld's algorithm [11]. Of course, one needs to prove that (4) is a “good approximation of Ψ_N ” for a sufficiently small time-step Δt .

Our goal in the present paper is to show that the RBM converges in some sense as $\Delta t \rightarrow 0$, with an error estimate that is both

- (a) independent of N , and
- (b) uniform in $\hbar \in (0, 1)$.

Obviously, one wishes to use the RBM for finite, albeit possibly large, values of N . It is therefore an obvious advantage to have an error estimate for the RBM that is *independent of N* , instead of an asymptotic rate of convergence that would be valid only in the limit as $N \rightarrow +\infty$. Specifically, the error bound obtained in this way is small *even for moderate values of N* . This explains the need for condition (a). Moreover, the RBM is known to converge in the case of classical dynamics (see [17]). It is therefore natural to seek an error estimate for the quantum RBM

method which does not deteriorate in the semiclassical regime, and this accounts for condition (b).

Our main results on this problem are gathered in the next section, especially in Theorem 2.1.

There are many variants of the RBM presented above. For instance, one could divide the N particles in batches of p (instead of only 2, but with $p \ll N$) particles to enhance the accuracy, or reduce the variance of the method (assuming of course that N is a multiple of p for simplicity). Likewise, one could replace the PDE (4) with some numerical approximation thereof — for instance one could approximate the solution of (4) by alternating direction method, where, at each time-step, one replaces the resolution of (4) by that of $N/2$ Schrödinger 2-body equations for each particle pair belonging to the same batch.

The RBM for the classical dynamics of large particle systems has been proposed and analyzed in [17]. However, obtaining error bounds which satisfy the conditions (a)-(b) listed above on random batch algorithms for quantum particle systems requires completely new ideas, especially on the problem of metrizing the state space as the number N of interacting particles tends to infinity. For that reason, the present paper discusses only the simplest possible formulation of the RBM, specifically the approximation of the solution of (2) by that of (4) with batches of only 2 particles, in order to focus our attention on the essential features of this problem.

The origins of the random batch method introduced in [17] for classical interacting particle systems can be found in stochastic programming (see for instance the discussion of the stochastic gradient method in [22]), and more specifically in the applications of that method in the context of machine learning (see [1, 24] and the references therein). A detailed presentation of stochastic approximation methods can be found in [19], while [3] provides a nice introduction to the dynamical aspects of these methods. In [17], and also in the problem under study in this paper, an error estimate of the RBM is established for *unsteady, time-dependent* problems, while in stochastic optimization methods such as the stochastic Gradient descent methods, one uses pseudo-time and the goal is to prove convergence toward the steady state.

2. MATHEMATICAL SETTING AND MAIN RESULT

We have introduced the N -body quantum dynamics and its random batch approximation via the Schrödinger equations (2) and (4). However, it will be more convenient to couch the analysis leading to our error estimates in terms of the corresponding von Neumann equations, which we recall below.

Henceforth we denote $\mathfrak{H} := L^2(\mathbf{R}^d; \mathbf{C})$ and $\mathfrak{H}_N = \mathfrak{H}^{\otimes N} \simeq L^2((\mathbf{R}^d)^N; \mathbf{C})$ for each $N \geq 2$. The algebra of bounded operators on \mathfrak{H} is denoted by $\mathcal{L}(\mathfrak{H})$, while $\mathcal{L}^1(\mathfrak{H}) \subset \mathcal{L}(\mathfrak{H})$ and $\mathcal{L}^2(\mathfrak{H})$ are respectively the two-sided ideals of trace-class and Hilbert-Schmidt operators on \mathfrak{H} . The operator norm of $A \in \mathcal{L}(\mathfrak{H})$ is denoted $\|A\|$. A density operator on \mathfrak{H} is a trace-class operator R on \mathfrak{H} such that

$$R = R^* \geq 0 \quad \text{and} \quad \text{trace}_{\mathfrak{H}}(R) = 1.$$

An example¹ of density operator on \mathfrak{H} is the \mathfrak{H} -orthogonal projection on $\mathbf{C}\psi$ for $\psi \in \mathfrak{H}$ satisfying $\|\psi\|_{\mathfrak{H}} = 1$, henceforth denoted $|\psi\rangle\langle\psi|$. The set of density operators on a separable Hilbert space H is henceforth denoted $\mathcal{D}(H)$.

The N -body von Neumann equation is the following differential equation with unknown $t \mapsto R(t)$, an operator-valued function of t :

$$(5) \quad i\hbar\partial_t R_N(t) = \mathcal{H}_N R_N(t) - R_N(t)\mathcal{H}_N =: [\mathcal{H}_N, R_N(t)], \quad R_N(0) = R_N^{in}.$$

Since $V \in C(\mathbf{R}^d)$ is bounded real-valued, the N -body quantum Hamiltonian \mathcal{H}_N has a self-adjoint extension to \mathfrak{H}_N , so that the solution of (2) is $\Psi_N(t, \cdot) = e^{-it\mathcal{H}_N}\Psi_N^{in}$, while the solution of (5) is given by

$$(6) \quad R_N(t) = e^{-it\mathcal{H}_N/\hbar} R_N^{in} e^{it\mathcal{H}_N/\hbar}.$$

In particular, if $R_N^{in} \in \mathcal{D}(\mathfrak{H}_N)$, then $R_N(t) \in \mathcal{D}(\mathfrak{H}_N)$ for each $t \geq 0$.

Likewise, if $R_N^{in} = |\Psi_N^{in}\rangle\langle\Psi_N^{in}|$, then $R_N(t) = |\Psi_N(t)\rangle\langle\Psi_N(t)|$ for each $t \geq 0$. Conversely, if $R_N(t)$ is a rank-one density operator, its range is of the form $\mathbf{C}\Psi_N(t)$ with $\|\Psi_N(t)\|_{\mathfrak{H}_N} = 1$, and this defines a unique $\Psi_N(t)$ so that $R_N(t) = |\Psi_N(t)\rangle\langle\Psi_N(t)|$ up to multiplication by a complex number of modulus one. In other words, $R_N(t)$ is in one-to-one correspondence with the quantum state associated to $\Psi_N(t)$, that is to say, in accordance with the Born interpretation, with the complex line in \mathfrak{H}_N spanned by $\Psi_N(t)$. This explains the connection between (2) and (5).

Likewise, the random batch von Neumann equation is the differential equation with unknown $t \mapsto \tilde{R}_N(t)$, an operator-valued function of t :

$$(7) \quad i\hbar\partial_t \tilde{R}_N(t) = [\mathcal{H}_N(t), \tilde{R}_N(t)], \quad \tilde{R}_N(0) = R_N^{in}.$$

The formula giving $\tilde{R}_N(t)$ is

$$\tilde{R}_N(t) = U(t, 0)\tilde{R}_N(0)U(0, t)$$

where, for each $0 \leq s \leq t$,

$$U(s, t) := e^{-\frac{i(s - \lfloor s/\Delta t \rfloor \Delta t)}{\hbar} \mathcal{H}_N(\lfloor \frac{s}{\Delta t} \rfloor \Delta t)} \prod_{j=\lfloor s/\Delta t \rfloor}^{\lfloor t/\Delta t \rfloor - 1} e^{\frac{i\Delta t}{\hbar} \mathcal{H}_N(j\Delta t)} e^{\frac{i(t - \lfloor t/\Delta t \rfloor \Delta t)}{\hbar} \mathcal{H}_N(\lfloor \frac{t}{\Delta t} \rfloor \Delta t)},$$

$$U(t, s) := U(s, t)^*.$$

Henceforth we denote for simplicity

$$(8) \quad \mathcal{U}(t, s)A := U(t, s)AU(s, t)$$

for each $A \in \mathcal{L}(\mathfrak{H})$; hence

$$(9) \quad \tilde{R}_N(t) = \mathcal{U}(t, 0)R_N^{in}, \quad t \geq 0.$$

Since our purpose is to find an error estimate for the RBM that is independent of the particle number N , we first need to define in terms of $R_N(t)$ and $\tilde{R}_N(t)$ *quantities of interest* to be compared that are *independent of N* . For instance one cannot hope to use the trace-norm of $\tilde{R}_N(t) - R_N(t)$ since both $\tilde{R}_N(t)$ and $R_N(t)$, and the trace-norm itself for elements of $\mathcal{L}^1(\mathfrak{H}_N)$ significantly depend on N . (There are other reasons for not using the trace-norm in this context, which

¹Throughout this paper, we use Dirac's bra-ket notation. A square integrable function $\psi \equiv \psi(x) \in \mathbf{C}$ viewed as a vector in \mathfrak{H} is denoted $|\psi\rangle$, while the notation $\langle\psi|$ designates the linear functional

$$\mathfrak{H} \ni \phi \mapsto \int_{\mathbf{R}^d} \overline{\psi(x)}\phi(x)dx =: \langle\psi|\phi\rangle \in \mathbf{C}.$$

will be explained later.) A common practice when considering large systems of identical particles is to study the reduced density operators. Assume that R_N^{in} has an integral kernel $r^{in} \equiv r^{in}(x_1, \dots, x_N; y_1, \dots, y_N)$ satisfying the symmetry

$$(10) \quad r^{in}(x_1, \dots, x_N; y_1, \dots, y_N) = r^{in}(x_{\sigma(1)}, \dots, x_{\sigma(N)}; y_{\sigma(1)}, \dots, y_{\sigma(N)})$$

for a.e. $(x_1, \dots, x_N; y_1, \dots, y_N) \in \mathbf{R}^{2dN}$ and each permutation $\sigma \in \mathfrak{S}_N$. Then, for each $t \geq 0$, the N -body density operator $R_N(t)$ solution of (5) satisfies the same symmetry, i.e. it has an integral kernel of the form $r(t; x_1, \dots, x_N; y_1, \dots, y_N)$ such that

$$(11) \quad r(t; x_1, \dots, x_N; y_1, \dots, y_N) = r(t; x_{\sigma(1)}, \dots, x_{\sigma(N)}; y_{\sigma(1)}, \dots, y_{\sigma(N)})$$

for a.e. $(x_1, \dots, x_N; y_1, \dots, y_N) \in \mathbf{R}^{2dN}$, all $t \geq 0$ and each permutation $\sigma \in \mathfrak{S}_N$. The 1-particle reduced density operator of $R_N(t) \in \mathcal{D}(\mathfrak{H}_N)$ is $R_{N,1}(t) \in \mathcal{D}(\mathfrak{H})$ defined by the integral kernel

$$(12) \quad r_{\mathbf{1}}(t, x, y) := \int_{(\mathbf{R}^d)^{N-1}} r(t; x, z_2, \dots, z_N; y, z_2, \dots, z_N) dz_2 \dots dz_N.$$

(This operation is legitimate for a trace-class operator R on \mathfrak{H}_N : indeed, R has an integral kernel $r(x_1, \dots, x_N; y_1, \dots, y_N)$ such that

$$(z_1, \dots, z_N) \mapsto r(x_1 + z_1, \dots, x_N + z_N; x_1, \dots, x_N)$$

belongs to $C(\mathbf{R}_{z_1, \dots, z_N}^{dN}; L^1(\mathbf{R}_{x_1, \dots, x_N}^{dN}))$ according to Footnote 1 on p. 61 in [14].)

Even if R_N^{in} satisfies the symmetry (10), in general $\tilde{R}_N(t)$ does not satisfy the symmetry analogous to (11) for $t > 0$ (with r replaced with \tilde{r} , an integral kernel for $\tilde{R}_N(t)$) because the random batch potential

$$\sum_{1 \leq l < n \leq N} \mathbf{T}_t(l, n) V(x_l - x_n)$$

is not invariant under permutations of the particle labels, at variance with the N -body potential

$$\frac{1}{N-1} \sum_{1 \leq l < n \leq N} V(x_l - x_n).$$

For that reason, the 1-particle reduced density operator of $\tilde{R}_N(t)$ is $\tilde{R}_{N,1}(t) \in \mathcal{D}(\mathfrak{H})$ defined for all $t > 0$ by the integral kernel

$$(13) \quad \tilde{r}_{\mathbf{1}}(t, x, y) := \frac{1}{N} \sum_{j=1}^N \int_{(\mathbf{R}^d)^{N-1}} \tilde{r}(t; Z_{j,N}[x], Z_{j,N}[y]) d\hat{Z}_{j,N},$$

with the notation

$$Z_{j,N}[x] := z_1, \dots, z_{j-1}, x, z_{j+1}, \dots, z_N, \quad d\hat{Z}_{j,N} = dz_1 \dots dz_{j-1} dz_{j+1} \dots dz_N.$$

(Obviously (13) holds with $r_{\mathbf{1}}$ and r in the place of $\tilde{r}_{\mathbf{1}}$ and \tilde{r} respectively because of the symmetry (11).)

Our main result on the convergence of the RBM for the N -body von Neumann equation (5) (or for the N -body Schrödinger equation (2)) is stated in terms of the Wigner functions of the density operators $R_N(t)$ and $\tilde{R}_N(t)$. We first recall the definition of the Wigner function of an operator $S \in \mathcal{L}^2(\mathfrak{H})$: let $s \equiv s(x, y)$ be an integral kernel for S . Then $s \in L^2(\mathbf{R}^d \times \mathbf{R}^d)$ and the Wigner function of S is the element of $L^2(\mathbf{R}^d \times \mathbf{R}^d)$ defined by the formula

$$W_h[S](x, \cdot) := \frac{1}{(2\pi)^d} \mathcal{F}(y \mapsto s(x + \frac{1}{2}\hbar y, x - \frac{1}{2}\hbar y)) \quad \text{for a.e. } x \in \mathbf{R}^d,$$

where \mathcal{F} designates the Fourier transform on $L^2(\mathbf{R}^d)$. If the argument of \mathcal{F} is integrable in y , then

$$W_h[S](x, \xi) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} s(x + \frac{1}{2}\hbar y, x - \frac{1}{2}\hbar y) e^{-i\xi \cdot y} dy.$$

When $S = |\psi\rangle\langle\psi|$ with $\psi \in \mathfrak{H}$, the Wigner function of S is often denoted $W_h[\psi]$. The reader is referred to [20] for more details on the Wigner function.

For each integer $M \geq 1$, we also introduce the dual norm

$$\|f\|_{-M} := \sup \left\{ \left| \iint_{\mathbf{R}^d \times \mathbf{R}^d} f(x, \xi) \overline{a(x, \xi)} dx d\xi \right| \left| \begin{array}{l} a \in C_c(\mathbf{R}^d \times \mathbf{R}^d), \text{ and} \\ \max_{\substack{|\alpha|, |\beta| \leq M \\ |\alpha| + |\beta| > 0}} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)} \leq 1 \end{array} \right. \right\}.$$

With the material introduced above, we can state our main results in the following theorem.

Theorem 2.1. *Assume that $N \geq 2$ and that $V \in C(\mathbf{R}^d)$ is a real-valued function such that*

$$V(z) = V(-z) \text{ for all } z \in \mathbf{R}^d, \quad \lim_{|z| \rightarrow +\infty} V(z) = 0, \quad \text{and} \quad \int_{\mathbf{R}^d} (1 + |\omega|^2) |\hat{V}(\omega)| d\omega < \infty.$$

Let $R_N^{in} \in \mathcal{D}(\mathfrak{H}_N)$, and let $R_N(t)$ and $\tilde{R}_N(t)$ be defined respectively by (6) and (9). Let $R_{N,1}(t)$ and $\tilde{R}_{N,1}(t)$ be the single-particle reduced density operators defined in terms of $R_N(t)$ and $\tilde{R}_N(t)$ by (12).

Then there exists a constant $\gamma_d > 0$ depending only on the dimension d of the configuration space such that, for each $t > 0$, one has

$$(14) \quad \begin{aligned} & \| |W_h[\mathbf{E}\tilde{R}_{N,1}(t)] - W_h[R_{N,1}(t)]| \|_{-[d/2]-3} \\ & \leq 2\gamma_d \Delta t e^{6t \max(1, \sqrt{d}L(V))} \Lambda(V) (2 + 3t\Lambda(V) \max(1, \Delta t) + 4\sqrt{d}L(V)t\Delta t). \end{aligned}$$

where \mathbf{E} is the mathematical expectation and Δt the reshuffling time-step in the definition of the random batch Hamiltonian (3), while

$$L(V) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\omega|^2 |\hat{V}(\omega)| d\omega, \quad \Lambda(V) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \sum_{\mu=1}^d |\omega^\mu| |\hat{V}(\omega)| d\omega,$$

where ω^ν is the ν -th component of ω .

Notice that the above result holds for the most general N -particle initial density operator R_N^{in} .

This error estimate satisfies both conditions (a) and (b). That it satisfies (a) is obvious, since N appears on neither side of (14). That it satisfies (b) is seen with the help of Theorem III.1 in [20]. Indeed, for each $t \geq 0$, the operators $R_N(t)$ and $\tilde{R}_N(t)$ define two bounded families of elements of $\mathcal{D}(\mathfrak{H})$ indexed by $N \geq 2$ and $\hbar \in (0, 1)$. Thus, $W_h[\mathbf{E}\tilde{R}_{N,1}(t)]$ and $W_h[R_{N,1}(t)]$ are relatively compact in the dual space \mathcal{A}' defined in Proposition III.1 of [20], and the limit points of these families as $\hbar \rightarrow 0$ are positive measures $\tilde{\mu}(t)$ and $\mu(t)$ on the phase space $\mathbf{R}^d \times \mathbf{R}^d$. With relatively mild tightness assumptions on the behavior of $W_h[\mathbf{E}\tilde{R}_{N,1}(t)]$ and $W_h[R_{N,1}(t)]$ in the limit as $|x| + |\xi| \rightarrow \infty$, these ‘‘Wigner measures’’ $\tilde{\mu}(t)$ and $\mu(t)$ encode the behavior of the reduced density operators $\tilde{R}_{N,1}(t)$ and $R_{N,1}(t)$ in the semiclassical regime. After checking how the dual norm $\| \cdot \|_{-[d/2]-3}$ behaves on weakly-* converging sequences in \mathcal{A}' , one can therefore hope that (14) implies an

estimate for the difference of the N -body reduced Wigner measure $\mu(t)$, and the expected value of its random batch analogue $\tilde{\mu}(t)$, since the right hand side of (14) does not involve \hbar .

The error estimate (14) can also be stated directly in terms of the density operators $R_{N,1}(t)$ and $\mathbf{E}\tilde{R}_{N,1}(t)$: see formula (32) in section 4.1 below. This new formulation of the error bound involves a new metric d_\hbar on the set of density operators on \mathfrak{H} , introduced in Definition 4.1. Using this new metric d_\hbar instead of the (more traditional) trace norm is the key to satisfying the uniformity conditions (a) and (b) formulated in the introduction — i.e. the uniformity in the particle number N and in the Planck constant \hbar respectively. The novelty of this approach is that using the new metric d_\hbar solves *both* the stiffness issue pertaining to the fast oscillations at frequency of order $1/\hbar$, and the “curse of dimensionality”, i.e. the fact that any numerical error in the evaluation of the interaction potential is amplified N times in the N -particle quantum Hamiltonian \mathcal{H}_N . To clarify completely this point, we shall compare in detail the key estimates obtained in the proof of Theorem 2.1 by using the new metric d_\hbar with the corresponding bounds that one would get with the trace norm (see section 4.2 below).

3. PROOF OF THEOREM 2.1

The proof of Theorem 2.1 makes critical use of rather different key ingredients (such as the mutual independence of the reshuffling permutations σ_j , semiclassical estimates on the interaction terms, together with a careful choice of test operators in the weak formulations of the N -body and random batch dynamics, and a quantitative version of the Calderon-Vaillancourt theorem), and as a result, is rather involved. We shall therefore decompose our argument in seven steps. Each step addresses one of the key issues in the error estimate obtained in Theorem 2.1.

3.1. Using the weak formulations of (2) and (4). For each $A \in \mathcal{L}(\mathfrak{H})$ and all $k = 2, \dots, N-1$, we set

$$J_1 A := A \otimes I_{\mathfrak{H}}^{\otimes(N-1)}, \quad J_k A := I_{\mathfrak{H}}^{\otimes(k-1)} \otimes A \otimes I_{\mathfrak{H}}^{\otimes(N-k)}, \quad J_N A = I_{\mathfrak{H}}^{\otimes(N-1)} \otimes A,$$

where $I_{\mathfrak{H}}$ designates the identity operator on the Hilbert space \mathfrak{H} .

With $A \in \mathcal{L}(\mathfrak{H})$ to be specified later, let

$$B_N(s) := \mathcal{U}(s, t) \frac{1}{N} \sum_{k=1}^N J_k A.$$

By the Duhamel formula,

$$B_N(t) = e^{-it\mathcal{H}_N/\hbar} B_N(0) e^{+it\mathcal{H}_N/\hbar} + \frac{1}{i\hbar} \int_0^t e^{-i(t-s)\mathcal{H}_N/\hbar} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] e^{+i(t-s)\mathcal{H}_N/\hbar} ds,$$

with the notation

$$V_{mn} := \text{multiplication by } V(x_m - x_n).$$

Because of (6) and (9), one has

$$\begin{aligned} \text{trace}_{\mathfrak{H}_N}(\tilde{R}_N(t) B_N(t)) &= \text{trace}(U(t, 0) \tilde{R}_N(0) U(t, 0)^* U(t, 0) B_N(0) U(t, 0)^*) \\ &= \text{trace}_{\mathfrak{H}_N}(\tilde{R}_N(0) B_N(0)) \end{aligned}$$

by cyclicity of the trace, and because $U(t, s)$ is unitary on \mathfrak{H}_N . On the other hand

$$\begin{aligned} \text{trace}_{\mathfrak{H}_N}(R_N(t)B_N(t)) &= \text{trace}(e^{-it\mathcal{H}_N/\hbar}R_N^{in}e^{+it\mathcal{H}_N/\hbar}B_N(t)) \\ &= \text{trace}_{\mathfrak{H}_N}(R_N^{in}e^{+it\mathcal{H}_N/\hbar}B_N(t)e^{-it\mathcal{H}_N/\hbar}) \end{aligned}$$

so that

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N}(R_N(t)B_N(t)) - \text{trace}(R_N^{in}B_N(0)) \\ &= \frac{1}{i\hbar} \int_0^t \text{trace} \left(\Psi_N e^{\frac{is}{\hbar}\mathcal{H}_N} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] e^{-\frac{is}{\hbar}\mathcal{H}_N} \right) ds \\ &= \frac{1}{i\hbar} \int_0^t \text{trace} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] \right) ds \end{aligned}$$

again by cyclicity of the trace. Therefore

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N}((\tilde{R}_N(t) - R_N(t))B_N(t)) \\ &= -\frac{1}{i\hbar} \int_0^t \text{trace} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] \right) ds. \end{aligned}$$

With our choice of $B_N(t)$, this last identity is recast as

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N} \left((\tilde{R}_N(t) - R_N(t)) \frac{1}{N} \sum_{k=1}^N J_k A \right) \\ &= \frac{i}{\hbar} \sum_{j=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \int_{(j-1)\Delta t}^{j\Delta t} \text{trace} \left(\left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] R_N(s) \right) ds \\ & \quad + \frac{i}{\hbar} \int_{\lfloor \frac{t}{\Delta t} \rfloor \Delta t}^t \text{trace} \left(\left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] R_N(s) \right) ds. \end{aligned}$$

Hence, taking the expectation over random reshufflings, and using the definitions (12) and (13) of reduced density operators, we arrive at the identity

$$\begin{aligned} & \text{trace}_{\mathfrak{H}}((\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t))A) = \text{trace}_{\mathfrak{H}_N} \left((\mathbf{E}\tilde{R}_N(t) - R_N(t)) \frac{1}{N} \sum_{k=1}^N J_k A \right) \\ &= \frac{i}{\hbar} \sum_{j=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \int_{(j-1)\Delta t}^{j\Delta t} \text{trace} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] \right) ds \\ & \quad + \frac{i}{\hbar} \int_{\lfloor \frac{t}{\Delta t} \rfloor \Delta t}^t \text{trace} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] \right) ds. \end{aligned}$$

3.2. Using the independence of $\sigma_1, \sigma_2, \dots$ For all $S \in \mathcal{L}(\mathfrak{H}_N)$, denote

$$\mathcal{U}_0(t)S := (e^{ith\Delta/2})^{\otimes N} S (e^{-ith\Delta/2})^{\otimes N}.$$

Observe that

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, B_N(s) \right] \right) \\ &= \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} (\mathbf{T}_s(m, n) - \frac{1}{N-1}) V_{mn}, \Delta B_N(s, j) \right] \right), \end{aligned}$$

with

$$\Delta B_N(s, j) := B_N(s) - \mathcal{U}_0(s - j\Delta t)B_N(j\Delta t),$$

since

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \mathcal{U}_0(s - j\Delta t) B_N(j\Delta t) \right] \right) \\ &= \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} \mathbf{E} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \mathcal{U}_0(s - j\Delta t) \mathbf{E} B_N(j\Delta t) \right] \right) = 0. \end{aligned}$$

The penultimate equality follows from the independence of the σ_j 's, since

$$B_N(j\Delta t) = B_N(j\Delta t + 0)$$

involves only $\sigma_{j+1}, \dots, \sigma_{[t/\Delta t]+1}$, while $\mathcal{U}_0(s - j\Delta t)$ is deterministic and $\mathbf{T}_s(m, n)$ only depends on σ_j . As for the last equality, it comes from the identity

$$\mathbf{E} \mathbf{T}_s(m, n) = \frac{1}{N-1}, \quad \text{for all } 1 \leq m < n \leq N \text{ and } s \geq 0.$$

For this last identity, see the proof of Lemma 3.1 in [17], and especially the second formula after (3.13) on p. 8 in [17].

Therefore

$$\begin{aligned} (15) \quad & \text{trace}_{\mathfrak{H}} \left((\mathbf{E} \tilde{R}_{N,1}(t) - R_{N,1}(t)) A \right) = \text{trace}_{\mathfrak{H}_N} \left((\mathbf{E} \tilde{R}_N(t) - R_N(t)) \frac{1}{N} \sum_{k=1}^N J_k A \right) \\ &= \frac{i}{\hbar} \sum_{j=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \int_{(j-1)\Delta t}^{j\Delta t} \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right) ds \\ & \quad + \frac{i}{\hbar} \int_{\lfloor \frac{t}{\Delta t} \rfloor \Delta t}^t \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, B_N(s) \right] \right) ds. \end{aligned}$$

3.3. Semiclassical potential estimate. The previous formula makes it obvious that our error analysis requires estimating commutators of various operators with the interaction potential. Besides, all these terms involve a $1/\hbar$ prefactor. With a view towards obtaining uniform as $\hbar \rightarrow 0$ error estimates, one should avoid by all means using bounds of the type

$$\left\| \frac{1}{\hbar} [V_{mn}, S] \right\| \leq \frac{2}{\hbar} \|V\|_{L^\infty(\mathbf{R}^d)} \|S\|.$$

We shall use instead the following lemma (see [16] on p. 1048).

Lemma 3.1. *Let $f \equiv f(x)$ be an element of $C_0^1(\mathbf{R}^d; \mathbf{C})$ such that \hat{f} and $\widehat{\nabla} f$ belong to $L^1(\mathbf{R}^d)$. Then, for each $T \in \mathcal{L}(\mathfrak{H})$, one has*

$$\|[f, T]\| \leq \Lambda(f) \max_{1 \leq \nu \leq d} \|[x^\nu, T]\|,$$

with

$$\Lambda(f) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \sum_{\mu=1}^d |\omega^\mu| |\hat{f}(\omega)| d\omega.$$

Proof. Let $E_\omega \in \mathcal{L}(\mathfrak{H})$ be the operator defined by $E_\omega \psi(x) = e^{i\omega \cdot x} \psi(x)$ for all $x \in \mathbf{R}^d$. Then

$$[f, T] = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \hat{f}(\omega) [E_\omega, T] d\omega$$

and

$$[E_\omega, T] E_\omega^* = \int_0^1 \frac{d}{dt} (E_{t\omega} T E_{t\omega}^*) dt = \int_0^1 E_{t\omega} [i\omega \cdot x, T] E_{t\omega}^* dt$$

so that

$$\| [E_\omega, T] \| \leq \max_{1 \leq \nu \leq d} \| [x^\nu, T] \| \sum_{\mu=1}^d |\omega^\mu|.$$

Hence

$$\| [f, T] \| \leq \max_{1 \leq \nu \leq d} \| [x^\nu, T] \| \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\hat{f}(\omega)| \sum_{\mu=1}^d |\omega^\mu| d\omega,$$

which implies the desired bound. \square

We use this lemma to control the terms $[V_{mn}, B_N(s)]$ and $[V_{mn}, \Delta B_N(s, j)]$ for each $m, n = 1, \dots, N$. First, one has

$$(16) \quad \| [V_{mn}, B_N(s)] \| \leq \Lambda(V) \left(\max_{1 \leq \mu \leq d} \| [x_m^\mu, B_N(s)] \| + \max_{1 \leq \nu \leq d} \| [x_n^\nu, B_N(s)] \| \right),$$

and

$$(17) \quad \| [V_{mn}, \Delta B_N(s, j)] \| \leq \Lambda(V) \left(\max_{1 \leq \mu \leq d} \| [x_m^\mu, \Delta B_N(s, j)] \| + \max_{1 \leq \nu \leq d} \| [x_n^\nu, \Delta B_N(s, j)] \| \right).$$

Hence

$$(18) \quad \left\| \sum_{1 \leq m < n \leq N} \left[\left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, B_N(s) \right] \right\| \leq 2\Lambda(V) \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, B_N(s)] \|,$$

and

$$(19) \quad \left\| \sum_{1 \leq m < n \leq N} \left[\left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right\| \leq 2\Lambda(V) \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, \Delta B_N(s, j)] \|.$$

This estimate is one of the key points in our analysis. It is based on the fact that the bounds (16) and (17) involve commutators with *only* the variables x_m and x_n which appear in V_{mn} . This observation is at the origin of the *independence of the right hand side of (14) in N* .

For each $m = 1, \dots, N$, we henceforth denote by $\tilde{m}(t)$ the unique index in $\{1, \dots, N\}$ different from m and in the same batch as m at time t . In other words, $\tilde{m}(t)$ is defined by the following two conditions:

$$(20) \quad \tilde{m}(t) \neq m \quad \text{and} \quad \mathbf{T}_t(m, \tilde{m}(t)) = 1.$$

By the Duhamel formula

$$(21) \quad \Delta B_N(s, j) = \frac{1}{i\hbar} \int_{j\Delta t}^s \frac{1}{2} \sum_{l=1}^N \mathcal{U}_0(s-\tau) [V_{l, \tilde{l}(s)}, B_N(\tau)] d\tau,$$

so that, for each $m = 1, \dots, N$ and each $\mu = 1, \dots, d$, one has

$$\| [x_m^\mu, \Delta B_N(s, j)] \| \leq \frac{1}{\hbar} \int_s^{j\Delta t} \frac{1}{2} \sum_{l=1}^N \| [x_m^\mu, \mathcal{U}_0(s-\tau) [V_{l, \tilde{l}(s)}, B_N(\tau)]] \| d\tau.$$

An elementary computation shows that

$$[x_m^\mu, \mathcal{U}_0(\theta)S] = \mathcal{U}_0(\theta) ([x_m^\mu, S] + \theta [-i\hbar \partial_{x_m^\mu}, S]);$$

hence

$$\| [x_m^\mu, \mathcal{U}_0(\theta)S] \| \leq \| [x_m^\mu, S] \| + |\theta| \| [-i\hbar \partial_{x_m^\mu}, S] \|.$$

Therefore

$$\begin{aligned}
& \| [x_m^\mu, \Delta B_N(s, j)] \| \\
& \leq \frac{1}{2\hbar} \int_s^{j\Delta t} \sum_{l=1}^N (\| [x_m^\mu, [V_{l, \tilde{l}(s)}, B_N(\tau)]] \| + (\tau - s) \| [-i\hbar \partial_{x_m^\mu}, [V_{l, \tilde{l}(s)}, B_N(\tau)]] \|) d\tau \\
& \leq \frac{1}{2\hbar} \int_s^{j\Delta t} \sum_{l=1}^N (\| [V_{l, \tilde{l}(s)}, [x_m^\mu, B_N(\tau)]] \| + (\tau - s) \| [V_{l, \tilde{l}(s)}, [-i\hbar \partial_{x_m^\mu}, B_N(\tau)]] \|) d\tau \\
& \quad + \frac{1}{\hbar} \int_s^{j\Delta t} (\tau - s) \| [-i\hbar \partial_\mu V_{m, \tilde{m}(s)}, B_N(\tau)] \| d\tau.
\end{aligned}$$

(In the formula above $\partial_\mu V$ denotes the partial derivative of V with respect to its μ -th variable (i.e. the μ -th coordinate in the position variable). In other words, $\partial_\mu V_{m, \tilde{m}(s)}$ is the multiplication operator defined on \mathfrak{H}_N by the formula $(\partial_\mu V_{m, \tilde{m}(s)} \Psi_N)(x_1, \dots, x_N) = (\partial_\mu V)(x_m - x_{\tilde{m}(s)}) \Psi_N(x_1, \dots, x_N)$ for all $\Psi_N \in \mathfrak{H}_N$.)

Thus

$$\begin{aligned}
& \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, \Delta B_N(s, j)] \| \\
& \leq \frac{\Lambda(V)}{\hbar} \int_s^{j\Delta t} \sum_{1 \leq l, m \leq N} \max_{1 \leq \lambda, \mu \leq d} \| [x_l^\lambda, [x_m^\mu, B_N(\tau)]] \| d\tau \\
(22) \quad & + \frac{\Lambda(V)\Delta t}{\hbar} \int_s^{j\Delta t} \sum_{1 \leq l, m \leq N} \max_{1 \leq \lambda, \mu \leq d} \| [x_l^\lambda, [-i\hbar \partial_{x_m^\mu}, B_N(\tau)]] \| d\tau \\
& + 2\sqrt{d}L(V)\Delta t \int_s^{j\Delta t} \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, B_N(\tau)] \| d\tau.
\end{aligned}$$

3.4. First order semiclassical estimates. In view of the inequality above, a key issue is therefore to bound operators of the form

$$[x_m^\mu, B_N(\tau)], \quad [-i\hbar \partial_{x_m^\mu}, B_N(\tau)],$$

and

$$[x_l^\lambda, [x_m^\mu, B_N(\tau)]], \quad [x_l^\lambda, [-i\hbar \partial_{x_m^\mu}, B_N(\tau)]], \quad [-i\hbar \partial_{x_l^\lambda}, [x_m^\mu, B_N(\tau)]].$$

This is done in the present section and the next. The bounds on the first two quantities above follow the proof of Lemma 4.1 in [16], which is itself based on the earlier analysis in Appendix B of [4], or Appendix C of [5].

We recall that

$$i\hbar \partial_s B_N(s) = [\mathcal{H}_N(s), B_N(s)], \quad B_N(t) = \frac{1}{N} \sum_{m=1}^N J_m A.$$

Hence

$$i\hbar \partial_s [x_m^\mu, B_N(s)] = [\mathcal{H}_N(s), [x_m^\mu, B_N(s)]] + [[x_m^\mu, \mathcal{H}_N(s)], B_N(s)],$$

and

$$[x_m^\mu, \mathcal{H}_N(s)] = [x_m^\mu, -\frac{1}{2}\hbar^2 \Delta_{x_m}] = i\hbar (-i\hbar \partial_{x_m^\mu}),$$

so that

$$i\hbar \partial_s [x_m^\mu, B_N(s)] = [\mathcal{H}_N(s), [x_m^\mu, B_N(s)]] + i\hbar [-i\hbar \partial_{x_m^\mu}, B_N(s)].$$

Thus

$$\| [x_m^\mu, B_N(s)] \| \leq \| [x_m^\mu, B_N(t)] \| + \int_s^t \| [-i\hbar \partial_{x_m^\mu}, B_N(\tau)] \| d\tau.$$

Likewise

$$i\hbar\partial_s[-i\hbar\partial_{x_m^\mu}, B_N(s)] = [\mathcal{H}_N(s), [-i\hbar\partial_{x_m^\mu}, B_N(s)]] + [[-i\hbar\partial_{x_m^\mu}, \mathcal{H}_N(s)], B_N(s)],$$

and

$$[-i\hbar\partial_{x_m^\mu}, \mathcal{H}_N(s)] = \frac{1}{2} \sum_{l=1}^N [-i\hbar\partial_{x_m^\mu}, V(x_l - x_{\tilde{l}(s)})] = -i\hbar\partial_\mu V(x_m - x_{\tilde{m}(s)}),$$

so that

$$i\hbar\partial_s[-i\hbar\partial_{x_m^\mu}, B_N(s)] = [\mathcal{H}_N(s), [-i\hbar\partial_{x_m^\mu}, B_N(s)]] - i\hbar[\partial_\mu V(x_m - x_{\tilde{m}(s)}), B_N(s)],$$

Therefore

$$\begin{aligned} \|[-i\hbar\partial_{x_m^\mu}, B_N(s)]\| &\leq \|[-i\hbar\partial_{x_m^\mu}, B_N(t)]\| \\ &\quad + \sqrt{d}L(V) \int_s^t (\max_{1 \leq \kappa \leq d} \|x_m^\kappa, B_N(\tau)\| + \max_{1 \leq \kappa \leq d} \|x_{\tilde{m}(\tau)}^\kappa, B_N(\tau)\|) d\tau \end{aligned}$$

Set

$$(23) \quad M_1(s) := \sum_{m=1}^N \max_{1 \leq \mu \leq d} \|x_m^\mu, B_N(s)\| + \sum_{m=1}^N \max_{1 \leq \mu \leq d} \|[-i\hbar\partial_{x_m^\mu}, B_N(s)]\|.$$

Then, one has

$$M_1(s) \leq M_1(t) + \max(1, 2\sqrt{d}L(V)) \int_s^t M_1(\tau) d\tau$$

so that, by Gronwall's lemma

$$(24) \quad M_1(s) \leq M_1(t) e^{(t-s)\max(1, 2\sqrt{d}L(V))}.$$

Finally, since

$$(25) \quad B_N(t) = \frac{1}{N} \sum_{k=1}^N J_k A,$$

one has

$$[x_m^\mu, B_N(t)] = \frac{1}{N} J_m [x^\mu, A], \quad [-i\hbar\partial_{x_m^\mu}, B_N(t)] = \frac{1}{N} J_m [-i\hbar\partial_\mu, A].$$

Hence

$$\| [x_m^\mu, B_N(t)] \| \leq \frac{1}{N} \| [x^\mu, A] \|, \quad \| [-i\hbar\partial_{x_m^\mu}, B_N(t)] \| \leq \frac{1}{N} \| [-i\hbar\partial_\mu, A] \|,$$

so that

$$(26) \quad M_1(t) \leq \max_{1 \leq \mu \leq d} \| [x^\mu, A] \| + \max_{1 \leq \mu \leq d} \| [-i\hbar\partial_\mu, A] \|$$

and therefore

$$M_1(s) \leq (\max_{1 \leq \mu \leq d} \| [x^\mu, A] \| + \max_{1 \leq \mu \leq d} \| [-i\hbar\partial_\mu, A] \|) e^{(t-s)\max(1, 2\sqrt{d}L(V))}.$$

The bound (26) is the second key point in obtaining a convergence rate in (14) that is independent of N . Indeed, one might suspect that the summation over N in the right hand sides of (18) and (19) leads to a bound of order $O(N)$, but the semiclassical propagation bound (26) and the careful choice of the initial condition $B_N(t)$ in (25) show that this is not the case. To summarize, the duality argument used in (16)-(19) and the semiclassical bound (23)-(26) allow treating separately the action of each position variable x_1, \dots, x_N corresponding to each particle. In section 4.2 below, we compare the estimates above with those obtained by using the

operator norm instead of $M_1(t)$ to control $B_N(t)$. We shall see in particular that using the operator norm to estimate $B_N(t)$ would mix all the variables x_1, \dots, x_N , thereby losing the benefits of dealing a pairwise interaction potential as observed in (16).

3.5. Second Order Semiclassical Estimates. One has

$$\begin{aligned} i\hbar\partial_s[x_n^\nu, [x_m^\mu, B_N(s)]] &= [\mathcal{H}_N(s), [x_n^\nu, [x_m^\mu, B_N(s)]]] \\ &\quad + [[x_n^\nu, \mathcal{H}_N(s)], [x_m^\mu, B_N(s)]] + i\hbar[x_n^\nu, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &= [\mathcal{H}_N(s), [x_n^\nu, [x_m^\mu, B_N(s)]]] \\ &\quad + i\hbar[-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]] + i\hbar[x_n^\nu, [-i\hbar\partial_{x_m^\mu}, B_N(s)]]. \end{aligned}$$

Notice that

$$[-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]] = [x_m^\mu, [-i\hbar\partial_{x_n^\nu}, B_N(s)]] .$$

Likewise

$$\begin{aligned} i\hbar\partial_s[-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]] &= [\mathcal{H}_N(s), [-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]]] \\ &\quad + [[-i\hbar\partial_{x_n^\nu}, \mathcal{H}_N(s)], [x_m^\mu, B_N(s)]] \\ &\quad + i\hbar[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &= [\mathcal{H}_N(s), [-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]]] \\ &\quad - i\hbar[\partial_\nu V(x_n - x_{\tilde{n}(s)}), [x_m^\mu, B_N(s)]] \\ &\quad + i\hbar[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] , \end{aligned}$$

and

$$\begin{aligned} i\hbar\partial_s[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] &= [\mathcal{H}_N(s), [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]]] \\ &\quad + [[-i\hbar\partial_{x_n^\nu}, \mathcal{H}_N(s)], [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &\quad + i\hbar[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &= [\mathcal{H}_N(s), [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]]] \\ &\quad - i\hbar[\partial_\nu V(x_n - x_{\tilde{n}(s)}), [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &\quad + i\hbar[-i\hbar\partial_{x_n^\nu}, [\partial_\mu V(x_m - x_{\tilde{m}(s)}), B_N(s)]] . \end{aligned}$$

Since

$$\begin{aligned} [-i\hbar\partial_{x_n^\nu}, [\partial_\mu V(x_m - x_{\tilde{m}(s)}), B_N(s)]] &= [\partial_\mu V(x_m - x_{\tilde{m}(s)}), [-i\hbar\partial_{x_n^\nu}, B_N(s)]] \\ &\quad + [[-i\hbar\partial_{x_n^\nu}, \partial_\mu V(x_m - x_{\tilde{m}(s)})], B_N(s)] \end{aligned}$$

and

$$\begin{aligned} [-i\hbar\partial_{x_n^\nu}, \partial_\mu V(x_m - x_{\tilde{m}(s)})] &= -i\hbar\partial_\nu\partial_\mu V(x_m - x_{\tilde{m}(s)})(\delta_{n,m} - \delta_{n,\tilde{m}(s)}) \\ &= -i\hbar\partial_\nu\partial_\mu V(x_n - x_{\tilde{n}(s)}) + i\hbar\partial_\nu\partial_\mu V(x_{\tilde{n}(s)} - x_n) = 0 . \end{aligned}$$

Indeed, V is even, so that $\partial_\nu\partial_\mu V$ is also even. Hence

$$\begin{aligned} i\hbar\partial_s[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] &= [\mathcal{H}_N(s), [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]]] \\ &\quad - i\hbar[\partial_\nu V(x_n - x_{\tilde{n}(s)}), [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \\ &\quad + i\hbar[\partial_\mu V(x_m - x_{\tilde{m}(s)}), [-i\hbar\partial_{x_n^\nu}, B_N(s)]] . \end{aligned}$$

Using the Duhamel formula, one finds that

$$\begin{aligned} \|[x_n^\nu, [x_m^\mu, B_N(s)]]\| &\leq \|[x_n^\nu, [x_m^\mu, B_N(s)]]\| + \int_s^t \|[x_m^\mu, [-i\hbar\partial_{x_n^\nu}, B_N(\tau)]]\| d\tau \\ &\quad + \int_s^t \|[x_n^\nu, [-i\hbar\partial_{x_m^\mu}, B_N(\tau)]]\| d\tau \end{aligned}$$

and

$$\begin{aligned} \|[-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]]\| &\leq \|[-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(t)]]\| \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_n^\kappa, [x_m^\mu, B_N(\tau)]] \| d\tau \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_{\tilde{n}(\tau)}^\kappa, [x_m^\mu, B_N(\tau)]] \| d\tau \\ &\quad + \int_s^t \| [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(\tau)]] \| d\tau, \end{aligned}$$

while

$$\begin{aligned} \|[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]]\| &\leq \|[-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(t)]]\| \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_n^\kappa, [-i\hbar\partial_{x_m^\mu}, B_N(\tau)]] \| d\tau \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_{\tilde{n}(\tau)}^\kappa, [-i\hbar\partial_{x_m^\mu}, B_N(\tau)]] \| d\tau \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_m^\kappa, [-i\hbar\partial_{x_n^\nu}, B_N(\tau)]] \| d\tau \\ &\quad + \sqrt{d}L(V) \int_s^t \max_{1 \leq \kappa \leq d} \| [x_{\tilde{m}(\tau)}^\kappa, [-i\hbar\partial_{x_n^\nu}, B_N(\tau)]] \| d\tau. \end{aligned}$$

Set

$$\begin{aligned} (27) \quad M_2(s) &:= \sum_{1 \leq m, n \leq N} \max_{1 \leq \mu, \nu \leq d} \| [x_n^\nu, [x_m^\mu, B_N(s)]] \| \\ &\quad + \sum_{1 \leq m, n \leq N} \max_{1 \leq \mu, \nu \leq d} \| [-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(s)]] \| \\ &\quad + \sum_{1 \leq m, n \leq N} \max_{1 \leq \mu, \nu \leq d} \| [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(s)]] \|. \end{aligned}$$

Therefore

$$M_2(s) \leq M_2(t) + 6 \max(1, \sqrt{d}L(V)) \int_s^t M_2(\tau) d\tau.$$

Since

$$B_N(t) = \frac{1}{N} \sum_{k=1}^N J_k A$$

one has

$$\begin{aligned} (28) \quad [x_n^\nu, [x_m^\mu, B_N(t)]] &= \frac{\delta_{mn}}{N} J_m [x^\nu, [x^\mu, A]] \\ [-i\hbar\partial_{x_n^\nu}, [x_m^\mu, B_N(t)]] &= \frac{\delta_{mn}}{N} J_m [-i\hbar\partial_\nu, [x^\mu, A]] \\ [-i\hbar\partial_{x_n^\nu}, [-i\hbar\partial_{x_m^\mu}, B_N(t)]] &= \frac{\delta_{mn}}{N} J_m [-i\hbar\partial_\nu, [-i\hbar\partial_\mu, A]]. \end{aligned}$$

Therefore

$$\begin{aligned} (29) \quad M_2(t) &\leq \max_{1 \leq \mu, \nu \leq d} \| [x^\nu, [x^\mu, A]] \| + \max_{1 \leq \mu, \nu \leq d} \| [-i\hbar\partial_\nu, [x^\mu, A]] \| \\ &\quad + \max_{1 \leq \mu, \nu \leq d} \| [-i\hbar\partial_\nu, [-i\hbar\partial_\mu, A]] \|, \end{aligned}$$

and

$$\begin{aligned}
(30) \quad M_2(s) &\leq \max_{1 \leq \mu, \nu \leq d} \|[x^\nu, [x^\mu, A]]\| e^{6(t-s) \max(1, \sqrt{d}L(V))} \\
&+ \max_{1 \leq \mu, \nu \leq d} \|[-i\hbar \partial_\nu, [x^\mu, A]]\| e^{6(t-s) \max(1, \sqrt{d}L(V))} \\
&+ \max_{1 \leq \mu, \nu \leq d} \|[-i\hbar \partial_\nu, [-i\hbar \partial_\mu, A]]\| e^{6(t-s) \max(1, \sqrt{d}L(V))}.
\end{aligned}$$

3.6. Implications of the semiclassical potential estimates. At this point, we use (19), (22) and (18) together with the bounds obtained in the last two steps.

First, we find that

$$\begin{aligned}
&\frac{1}{\hbar} \left| \int_{[\frac{t}{\Delta t}]\Delta t}^t \text{trace} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, B_N(s) \right] \right) ds \right| \\
&\leq \frac{1}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \mathbf{E} \left\| \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, B_N(s) \right] \right\| ds \\
&\leq \frac{1}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \frac{1}{N-1} \sum_{1 \leq m < n \leq N} \|[V_{mn}, B_N(s)]\| ds \\
&\quad + \frac{1}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \frac{1}{2} \sum_{m=1}^N \|[V_{m, \tilde{m}(s)}, B_N(s)]\| ds,
\end{aligned}$$

so that

$$\begin{aligned}
&\frac{1}{\hbar} \left| \int_{[\frac{t}{\Delta t}]\Delta t}^t \text{trace} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, B_N(s) \right] \right) ds \right| \\
&\leq \frac{\Lambda(V)}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \frac{1}{N-1} \sum_{1 \leq m < n \leq N} \left(\max_{1 \leq \kappa \leq d} \|[x_m^\kappa, B_N(s)]\| + \max_{1 \leq \kappa \leq d} \|[x_n^\kappa, B_N(s)]\| \right) ds \\
&\quad + \frac{\Lambda(V)}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \frac{1}{2} \sum_{m=1}^N \left(\max_{1 \leq \kappa \leq d} \|[x_m^\kappa, B_N(s)]\| + \max_{1 \leq \kappa \leq d} \|[x_{\tilde{m}(s)}^\kappa, B_N(s)]\| \right) ds \\
&\leq \frac{2\Lambda(V)}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \sum_{m=1}^N \max_{1 \leq \kappa \leq d} \|[x_m^\kappa, B_N(s)]\| ds \leq \frac{2\Lambda(V)}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t M_1(s) ds \\
&\leq 2\Lambda(V) \Delta t e^{t \max(1, 2\sqrt{d}L(V))} \frac{M_1(t)}{\hbar}.
\end{aligned}$$

Next

$$\begin{aligned}
&\frac{1}{\hbar} \left| \int_{(j-1)\Delta t}^{j\Delta t} \text{trace} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right) ds \right| \\
&\leq \frac{1}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \mathbf{E} \left\| \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right\| ds \\
&\leq \frac{1}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \frac{1}{N-1} \sum_{1 \leq m < n \leq N} \|[V_{mn}, \Delta B_N(s, j)]\| ds \\
&\quad + \frac{1}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \frac{1}{2} \sum_{m=1}^N \|[V_{m, \tilde{m}(s)}, \Delta B_N(s, j)]\| ds,
\end{aligned}$$

so that

$$\begin{aligned}
& \left| \frac{1}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \text{trace} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right) ds \right| \\
& \leq \frac{\Lambda(V)}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \frac{1}{N-1} \sum_{1 \leq m < n \leq N} \left(\max_{1 \leq \kappa \leq d} \| [x_m^\kappa, \Delta B_N(s, j)] \| + \max_{1 \leq \kappa \leq d} \| [x_n^\kappa, \Delta B_N(s, j)] \| \right) ds \\
& \quad + \frac{\Lambda(V)}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \frac{1}{2} \sum_{m=1}^N \left(\max_{1 \leq \kappa \leq d} \| [x_m^\kappa, \Delta B_N(s, j)] \| + \max_{1 \leq \kappa \leq d} \| [x_{\bar{m}(s)}^\kappa, \Delta B_N(s, j)] \| \right) ds \\
& \leq \frac{2\Lambda(V)}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, \Delta B_N(s, j)] \|.
\end{aligned}$$

Recall that

$$\begin{aligned}
& \sum_{m=1}^N \max_{1 \leq \mu \leq d} \| [x_m^\mu, \Delta B_N(s, j)] \| \\
& \leq \frac{\Lambda(V)}{\hbar} \max(1, \Delta t) \int_s^{j\Delta t} M_2(\tau) d\tau + 2\sqrt{d}L(V)\Delta t \int_s^{j\Delta t} M_1(\tau) d\tau \\
& \leq \Lambda(V)\Delta t \max(1, \Delta t) e^{6t \max(1, \sqrt{d}L(V))} \frac{M_2(t)}{\hbar} \\
& \quad + 2\sqrt{d}L(V)\Delta t^2 e^{t \max(1, 2\sqrt{d}L(V))} M_1(t).
\end{aligned}$$

Therefore

$$\begin{aligned}
& \left| \frac{1}{\hbar} \int_{(j-1)\Delta t}^{j\Delta t} \text{trace} \left(R_N(s) \left[\sum_{1 \leq m < n \leq N} \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}, \Delta B_N(s, j) \right] \right) ds \right| \\
& \leq 2\Lambda(V)^2 \Delta t^2 \max(1, \Delta t) e^{6t \max(1, \sqrt{d}L(V))} \frac{M_2(t)}{\hbar^2} \\
& \quad + 4\sqrt{d}\Lambda(V)L(V)\Delta t^3 e^{t \max(1, 2\sqrt{d}L(V))} \frac{M_1(t)}{\hbar}.
\end{aligned}$$

Putting together all these elements of information leads to the bound (31)

$$\begin{aligned}
& \left| \text{trace}_{\mathfrak{H}} \left((\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t))A \right) \right| = \left| \text{trace}_{\mathfrak{H}_N} \left((\mathbf{E}\tilde{R}_N(t) - R_N(t)) \frac{1}{N} \sum_{k=1}^N J_k A \right) \right| \\
& \leq 2\Lambda(V)(1 + 2\sqrt{d}L(V)t\Delta t)\Delta t e^{t \max(1, 2\sqrt{d}L(V))} \frac{M_1(t)}{\hbar} \\
& \quad + 2\Lambda(V)^2 t \Delta t \max(1, \Delta t) e^{6t \max(1, \sqrt{d}L(V))} \frac{M_2(t)}{\hbar^2}.
\end{aligned}$$

Observe that the right hand side of (31) is independent of N . This explains why the error estimate in (14) is independent of N .

3.7. Specializing to the case where A is a Weyl operator. In order to finish the proof of Theorem 2.1, we restrict our attention to a convenient class of test operators A , for which the quantities $M_1(t)/\hbar$ and $M_2(t)/\hbar^2$ are bounded as $\hbar \rightarrow 0$.

We first recall the definition of a Weyl operator on \mathfrak{H} . For each $a \in \mathcal{S}'(\mathbf{R}^d \times \mathbf{R}^d)$, one defines a linear map $\text{OP}_\hbar^W[a] : \mathcal{S}(\mathbf{R}^d) \mapsto \mathcal{S}'(\mathbf{R}^d)$ by the following duality formula:

$$\langle \text{OP}_\hbar^W[a] \psi, \bar{\phi} \rangle_{\mathcal{S}'(\mathbf{R}^d), \mathcal{S}(\mathbf{R}^d)} := \langle a, W_\hbar[[\phi]\langle \psi |] \rangle_{\mathcal{S}'(\mathbf{R}^d), \mathcal{S}(\mathbf{R}^d)}.$$

We recall that $\mathcal{S}(\mathbf{R}^d)$ is the Schwartz class of infinitely differentiable functions defined on \mathbf{R}^d which are rapidly decaying at infinity (i.e. $O(|x|^{-n})$ as $|x| \rightarrow +\infty$) and whose partial derivatives of all orders are also rapidly decaying at infinity. Then $\mathcal{S}'(\mathbf{R}^d)$ is the set of tempered distributions (i.e. the topological dual of $\mathcal{S}(\mathbf{R}^d)$), while $\langle T, \phi \rangle_{\mathcal{S}'(\mathbf{R}^d), \mathcal{S}(\mathbf{R}^d)}$ designates the evaluation of the tempered distribution T on the test function ϕ .

Elementary computations show that

$$\frac{i}{\hbar}[x^\nu, \text{OP}_\hbar^W[a]] = -\text{OP}_\hbar^W[\partial_{\xi^\nu} a], \quad \frac{i}{\hbar}[-i\hbar\partial_{x^\nu}, \text{OP}_\hbar^W[a]] = \text{OP}_\hbar^W[\partial_{x^\nu} a].$$

Boulkhemair's improvement [7] of the Calderon-Vaillancourt theorem states that, for each integer $d \geq 1$, there exists $\gamma_d > 0$ such that, for each $a \in \mathcal{S}'(\mathbf{R}^d \times \mathbf{R}^d)$ satisfying the condition $\partial_x^\alpha \partial_\xi^\beta a \in L^\infty(\mathbf{R}^d \times \mathbf{R}^d)$ whenever $|\alpha|$ and $|\beta| \leq [d/2] + 1$,

$$\max_{|\alpha|, |\beta| \leq [d/2] + 1} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)} \leq 1 \implies \|\text{OP}_\hbar^W[a]\| \leq \gamma_d.$$

Therefore, choosing $A = \text{OP}_\hbar^W[a]$ implies that

$$\frac{M_1(t)}{\hbar} \leq 2\gamma_d \max_{|\alpha|, |\beta| \leq [d/2] + 2} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)},$$

and

$$\frac{M_2(t)}{\hbar^2} \leq 3\gamma_d \max_{|\alpha|, |\beta| \leq [d/2] + 3} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)},$$

so that

$$\begin{aligned} & \left| \text{trace}_{\mathfrak{H}} \left((\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)) \text{OP}_\hbar^W[a] \right) \right| \\ & \leq 2\Lambda(V)(1 + 2\sqrt{d}L(V)t\Delta t)\Delta t e^{t \max(1, 2\sqrt{d}L(V))} \frac{M_1(t)}{\hbar} \\ & \quad + 2\Lambda(V)^2 t \Delta t \max(1, \Delta t) e^{6t \max(1, \sqrt{d}L(V))} \frac{M_2(t)}{\hbar^2} \\ & \leq 2\Delta t e^{6t \max(1, \sqrt{d}L(V))} \Lambda(V)(2 + 3t\Lambda(V) \max(1, \Delta t) + 4\sqrt{d}L(V)t\Delta t) \\ & \quad \times \gamma_d \max_{|\alpha|, |\beta| \leq [d/2] + 3} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)}. \end{aligned}$$

Since $\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)$ is a self-adjoint element of $\mathcal{L}^1(\mathfrak{H})$, there exists a complete orthonormal sequence ϕ_k in \mathfrak{H} such that

$$\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t) = \sum_{k \geq 1} \rho_k |\phi_k\rangle \langle \phi_k|$$

with

$$\rho_k \in \mathbf{R} \quad \text{and} \quad \sum_{k \geq 1} |\rho_k| < \infty.$$

Therefore, if $a \in \mathcal{S}'(\mathbf{R}^d \times \mathbf{R}^d)$ satisfies

$$\max_{|\alpha|, |\beta| \leq [d/2] + 3} \|\partial_x^\alpha \partial_\xi^\beta a\|_{L^\infty(\mathbf{R}^d \times \mathbf{R}^d)} < \infty,$$

then

$$\begin{aligned} \left\langle W_\hbar[\mathbf{E}\tilde{R}_{N,1}(t)] - W_\hbar[R_{N,1}(t)], a \right\rangle_{\mathcal{S}'(\mathbf{R}^d), \mathcal{S}(\mathbf{R}^d)} &= \sum_{k \geq 1} \rho_k \langle W_\hbar[|\phi_k\rangle \langle \phi_k|], a \rangle_{\mathcal{S}'(\mathbf{R}^d), \mathcal{S}(\mathbf{R}^d)} \\ &= \text{trace}_{\mathfrak{H}} \left((\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)) \text{OP}_\hbar^W[a] \right). \end{aligned}$$

In other words

$$\begin{aligned} & \left\| \|W_h[\mathbf{E}\tilde{R}_{N,1}(t)] - W_h[R_{N,1}(t)]\| \right\|_{-[d/2]-3} \\ & \leq 2\gamma_d \Delta t e^{6t \max(1, \sqrt{d}L(V))} \Lambda(V) (2 + 3t\Lambda(V) \max(1, \Delta t) + 4\sqrt{d}L(V)t\Delta t). \end{aligned}$$

4. REMARKS ON THEOREM 2.1 AND ITS PROOF

4.1. Metrizing the set of density operators. Our analysis in the present paper can be expressed in terms of the following metric on the set of density operators on the Hilbert space \mathfrak{H} .

Definition 4.1. For $R, S \in \mathcal{D}(\mathfrak{H})$, set $D := -i\partial$ and

$$d_h(R, S) := \sup \left\{ \left| \text{trace}_{\mathfrak{H}}((R-S)A) \right| \left| \begin{array}{l} A \in \mathcal{L}(\mathfrak{H}) \text{ and for all } 1 \leq \mu, \nu \leq d \\ \hbar \| [x^\mu, A] \| + \hbar \| [\hbar D_\mu, A] \| + \| [x^\nu, [x^\mu, A]] \| \\ + \| [\hbar D_\nu, [x^\mu, A]] \| + \| [\hbar D_\nu, [\hbar D_\mu, A]] \| \leq 5\hbar^2 \end{array} \right. \right\}.$$

The distance is analogous to several distances introduced earlier in the literature. The first is obviously the Monge-Kantorovich(-Rubinstein) distance, also referred to as the Wasserstein distance of exponent 1: see formula (7.1) in chapter 7 of [23]. However, the Monge-Kantorovich distance is defined on the set of Borel probability measures on the Euclidean space \mathbf{R}^d , and not on density operators on \mathfrak{H} .

An analogue of the Monge-Kantorovich distance has been proposed by Connes on the set of states on a C^* -algebra endowed with an unbounded Fredholm module: see Proposition 4 in [8], or §5 in the Introduction and §1 in chapter 6 of [9]. See also the review paper [21] for a more thorough discussion of this distance. However, the Connes distance is the noncommutative analogue of a Riemannian metric on compact spin manifold, as explained in Proposition 1 of [8] — see also formula (2.9) in [21] which does not involve a spin structure — and not on a phase space, i.e. not on a cotangent bundle. The analogue of the Monge-Kantorovich or Wasserstein distance of exponent 1, or of Connes' distance in our setting would be

$$MK_1^h(R, S) := \sup_{\substack{A \in \mathcal{L}(\mathfrak{H}) \\ \max_{1 \leq \mu \leq d} (\| [x^\mu, A] \|, \| [\hbar D_\mu, A] \|) \leq \hbar}} \left| \text{trace}_{\mathfrak{H}}((R-S)A) \right|.$$

Since the correspondence principle associates the Poisson bracket $\{\cdot, \cdot\}$ to $\frac{i}{\hbar}[\cdot, \cdot]$, and since $\{x^\mu, \cdot\} = -\partial_{\xi^\mu}$ while $\{\xi^\mu, \cdot\} = -\partial_{x^\mu}$, the constraint $\| [\hbar D_\mu, A] \| \leq \hbar$ corresponds to Lipschitz continuity in the position variable x^μ , while the constraint $\| [x^\mu, A] \| \leq \hbar$ corresponds to Lipschitz continuity in the momentum variable ξ^μ . The distance d_N^T used in [16] to prove the uniformity in the Planck constant \hbar of the mean-field limit in quantum mechanics (see formula (43) and Theorem 1.1 in [16]) is essentially based on the same idea as MK_1^h .

However, neither the variant MK_1^h of the Connes distance nor the distance d_N^T in [16] can be used to obtain the convergence rate (14) for the random batch method, and a new idea is needed at this point.

This new idea is to use iterated commutators in the constraints entering the definition of d_h in order to fit the specifics of the random batch method. More precisely, the need for iterated commutators comes from the key step using the independence of the reshuffling permutations $\sigma_1, \sigma_2, \dots$ leading to (15) and the quantity $\Delta B_N(s, j)$. Indeed, estimating $\Delta B_N(s, j)$ systematically involves iterated brackets, as shown in (21) and the subsequent formulas.

It should be noted that, because the constraints in the definition of d_{\hbar} involve iterated commutators, this metric may not have an interpretation in the context of optimal transport through the Kantorovich duality theorem (Theorem 1.3 in [23]) as the Wasserstein (or Monge-Kantorovich) metric of exponent 1, or as the Connes distance, for which an analogue of the Kantorovich duality theorem has been recently worked out [10]. For this reason, it could be that the metric d_{\hbar} considered here is of a different nature than either MK_1^{\hbar} or the Connes distance.

The main properties of d_{\hbar} are summarized in the following proposition.

Proposition 4.2. *For each $\hbar > 0$, the function d_{\hbar} is defined on $\mathcal{D}(\mathfrak{H}) \times \mathcal{D}(\mathfrak{H})$ and takes its values in $[0, +\infty]$. Moreover*

- (i) *the function d_{\hbar} is an extended metric on $\mathcal{D}(\mathfrak{H})$: it is symmetric in its two arguments, satisfies the triangle inequality, and separates points in $\mathcal{D}(\mathfrak{H})$;*
- (ii) *there exists $\gamma_d > 0$ (depending only on the space dimension d) such that, for all $R, S \in \mathcal{D}(\mathfrak{H})$*

$$\| \|W_{\hbar}[R] - W_{\hbar}[S]\| \|_{-[d/2]-3} \leq \gamma_d d_{\hbar}(R, S).$$

Proof. That d_{\hbar} is symmetric in its arguments and satisfies the triangle inequality is obvious from the definition. That d_{\hbar} separates points in $\mathcal{D}(\mathfrak{H})$ follows from (ii). Indeed

$$d_{\hbar}(R, S) = 0 \implies \| \|W_{\hbar}[R] - W_{\hbar}[S]\| \|_{-[d/2]-3} = 0,$$

so that

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} W_{\hbar}[R - S](x, \xi) \overline{a(x, \xi)} dx d\xi = 0$$

for all $a \equiv a(x, \xi) \in \mathcal{S}(\mathbf{R}^d \times \mathbf{R}^d)$ such that $\|\partial_x^{\alpha} \partial_{\xi}^{\beta} a\|_{L^{\infty}(\mathbf{R}^{2d})} \leq 1$ for all multi-indices $\alpha, \beta \in \mathbf{N}^d$ such that $|\alpha|, |\beta| \leq [d/2] + 3$, and therefore, by homogeneity and density of $\mathcal{S}(\mathbf{R}^d \times \mathbf{R}^d)$ in $L^2(\mathbf{R}^d \times \mathbf{R}^d)$, for all $a \equiv a(x, \xi) \in L^2(\mathbf{R}^d \times \mathbf{R}^d)$. Hence $W_{\hbar}[R] = W_{\hbar}[S]$ in $L^2(\mathbf{R}^d \times \mathbf{R}^d)$, and since the Fourier transform is invertible on $L^2(\mathbf{R}^d \times \mathbf{R}^d)$, this implies that R and S have integral kernels a.e. equal, so that $R = S$. This proves Property (i) taking Property (ii) for granted.

The proof of Property (ii) is essentially a repetition of the last step in the proof of Theorem 2.1 (section 3.7). Indeed

$$\begin{aligned} & \| \|W_{\hbar}[R] - W_{\hbar}[S]\| \|_{-[d/2]-3} \\ = & \sup \left\{ \left| \int_{\mathbf{R}^d \times \mathbf{R}^d} W_{\hbar}[R - S](x, \xi) a(x, \xi) dx d\xi \right| \text{ s.t. } \max_{|\alpha|, |\beta| \leq [d/2] + 3} |\partial_x^{\alpha} \partial_{\xi}^{\beta} a(x, \xi)| = 1 \right\} \\ & = \sup \left\{ \left| \text{trace}_{\mathfrak{H}} \left((R - S) \text{OP}_{\hbar}^W[a] \right) \right| \text{ s.t. } \max_{|\alpha|, |\beta| \leq [d/2] + 3} |\partial_x^{\alpha} \partial_{\xi}^{\beta} a(x, \xi)| = 1 \right\} \\ & \leq \gamma_d d_{\hbar}(R, S), \end{aligned}$$

since

$$\begin{aligned} \max_{|\alpha|, |\beta| \leq [d/2] + 3} |\partial_x^{\alpha} \partial_{\xi}^{\beta} a(x, \xi)| = 1 & \implies \hbar \| [x^{\mu}, A] \| + \hbar \| [\hbar D_{\mu}, A] \| + \| [x^{\nu}, [x^{\mu}, A]] \| \\ & + \| [\hbar D_{\nu}, [x^{\mu}, A]] \| + \| [\hbar D_{\nu}, [\hbar D_{\mu}, A]] \| \leq 5\gamma_d \hbar^2 \end{aligned}$$

by Boulkhemair's variant [7] of the Calderon-Vaillancourt theorem. \square

The error estimate in Theorem 2.1 could have been couched in terms of the distance d_{\hbar} . Indeed, the inequality (31) at the end of the penultimate step in the

proof of Theorem 2.1 can be recast as

$$(32) \quad d_h(\mathbf{E}\tilde{R}_{N,1}(t), R_{N,1}(t)) \leq 10\Lambda(V)^2 t \Delta t \max(1, \Delta t) e^{6t \max(1, \sqrt{d}L(V))} \\ + 10\Lambda(V)(1 + 2\sqrt{d}L(V)t\Delta t)\Delta t e^{t \max(1, 2\sqrt{d}L(V))}.$$

Up to unessential modifications in the constants, the error estimate in Theorem 2.1 is a consequence of this inequality and Proposition 4.2 (ii).

4.2. On the choice of $\|\cdot\|_{-[d/2]-3}$ or d_h to express the error bound. The idea of using the metric d_h presented in the previous section might seem strange. One might find it more natural to use more traditional metrics on density operators, such as the trace norm, for instance. Indeed, for all $t, s \in \mathbf{R}$, the map $\mathcal{U}(t, s)$ is an isometry for the trace norm, because the map $U(t, s)$ is a unitary operator on \mathfrak{H}_N .

Estimating the difference $\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)$ in trace norm can be done along the same line as in section 3. Although this is not the simplest route to obtaining this estimate, it will be easier to compare the inequalities at each step in this estimate with the ones using d_h . Indeed

$$\|\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)\|_1 = \sup_{\|A\| \leq 1} |\text{trace}((\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t))A)|,$$

and using (15) shows that

$$\begin{aligned} & \|\mathbf{E}\tilde{R}_{N,1}(t) - R_{N,1}(t)\|_1 \\ & \leq \frac{1}{\hbar} \int_{[\frac{t}{\Delta t}]\Delta t}^t \sup_{\|A\| \leq 1} \left| \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \mathcal{V}_{mn}, B_N(s) \right] \right) \right| ds \\ & + \frac{1}{\hbar} \sum_{j=1}^{[\frac{t}{\Delta t}]} \int_{(j-1)\Delta t}^{j\Delta t} \sup_{\|A\| \leq 1} \left| \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \mathcal{V}_{mn}, \Delta B_N(s, j) \right] \right) \right| ds, \end{aligned}$$

with the notation

$$\mathcal{V}_{mn} := \left(\mathbf{T}_s(m, n) - \frac{1}{N-1} \right) V_{mn}.$$

Then

$$\|A\| \leq 1 \implies \left\| \frac{1}{N} \sum_{k=1}^N J_k A \right\| \leq 1 \implies \|B_N(s)\| \leq 1$$

for all $s \in [0, t]$, and

$$(33) \quad \begin{aligned} & \left| \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \mathcal{V}_{mn}, B_N(s) \right] \right) \right| \\ & \leq \sum_{1 \leq m < n \leq N} |\text{trace}_{\mathfrak{H}_N} (R_N(s) [\mathbf{E} \mathcal{V}_{mn}, B_N(s)])| \\ & \leq 2N \|V\|_{L^\infty(\mathbf{R}^d)} \|R_N(s)\|_1 \|B_N(s)\| = 2N \|V\|_{L^\infty(\mathbf{R}^d)}, \end{aligned}$$

since

$$\begin{aligned} \mathbf{T}_s(m, n) = 0 & \implies \|\mathcal{V}_{mn}\| = \frac{1}{N-1} \|V\|_{L^\infty(\mathbf{R}^d)} \\ \mathbf{T}_s(m, n) = 1 & \implies \|\mathcal{V}_{mn}\| = \frac{N-2}{N-1} \|V\|_{L^\infty(\mathbf{R}^d)}. \end{aligned}$$

Likewise

$$(34) \quad \left| \text{trace}_{\mathfrak{H}_N} \left(R_N(s) \mathbf{E} \left[\sum_{1 \leq m < n \leq N} \mathcal{V}_{mn}, \Delta B_N(s, j) \right] \right) \right| \\ \leq \sum_{1 \leq m < n \leq N} |\text{trace}_{\mathfrak{H}_N} (R_N(s) \mathbf{E}[\mathcal{V}_{mn}, \Delta B_N(s, j)])| \\ \leq 2N \|V\|_{L^\infty(\mathbf{R}^d)} \|\Delta B_N(s, j)\|,$$

and (21) implies that

$$(35) \quad \|\Delta B_N(s, j)\| \leq \frac{1}{\hbar} \Delta t N \|V\|_{L^\infty(\mathbf{R}^d)}.$$

Putting all these estimates together results in the upper bound

$$\|\mathbf{E} \tilde{R}_{N,1}(t) - R_{N,1}(t)\|_1 \leq \frac{2}{\hbar} \Delta t \cdot N \|V\|_{L^\infty(\mathbf{R}^d)} \\ + \frac{2}{\hbar} \left[\frac{t}{\Delta t} \right] \Delta t N \|V\|_{L^\infty(\mathbf{R}^d)} \cdot \frac{1}{\hbar} \Delta t N \|V\|_{L^\infty(\mathbf{R}^d)} \\ \leq \frac{2N}{\hbar} \Delta t \|V\|_{L^\infty(\mathbf{R}^d)} \left(1 + \frac{Nt}{\hbar} \|V\|_{L^\infty(\mathbf{R}^d)} \right),$$

which is neither uniform as $N \rightarrow \infty$ nor as $\hbar \rightarrow 0$, and therefore satisfies neither of our requirements (a) and (b) at the end of section 1.

It is instructive to compare the rather naive estimates above with the more subtle corresponding estimate in the proof of Theorem 2.1.

For instance, comparing (33) with (16), or (34) with (17) shows clearly that (16) or (17) involve *only* the commutators with the variables x_k^μ that are present in the potential $V(x_m - x_n)$, i.e. only the two values $k = m$ or $k = n$. This key observation is at the core of section 3.3.

When summing over all possible pairs m, n either with m and n in the same batch, or over all m, n with the coupling constant $1/N$, one arrives at the bound (18), which does not involve the N factor that is present in (33). One might suspect that this N factor is hidden in the summation over $m = 1, \dots, N$ in the right hand side of (18), but in fact this summation is included in the definition (23), and the bounds (24) and (26) make it clear that no N factor can arise in this way. The key observation is obviously the bound (26) which does not involve N , since

$$B_N(t) = \frac{1}{N} \sum_{k=1}^N J_k A$$

and $[x_m^\mu, J_k A] = [-i\hbar \partial_{x_m^\mu}, J_k A] = 0$ for all $\mu = 1, \dots, d$ and all $m = 1, \dots, N$ unless $m = k$. Finally, the definition of d_\hbar implies that the test operator A satisfies both $\|[x_m^\mu, A]\| = O(\hbar)$ and $\|[-i\hbar \partial_{x_m^\mu}, J_k A]\| = O(\hbar)$ for all $m = 1, \dots, N$ and $\mu = 1, \dots, d$, so that $M_1(t) = O(\hbar)$. This nice bound (small as $\hbar \rightarrow 0$, independent of N) is propagated by the random batch dynamics as explained in section 3.4. As a result, the bound (18) does not involve the unpleasant N factor in (33), and the fact that $M_1(t) = O(\hbar)$ offsets the $1/\hbar$ factor multiplying the last time integral on the right hand side of (15), at variance with the naive estimate above.

The same advantages of using the d_\hbar metric instead of the trace norm are observed in the treatment of the “generic” term, i.e. the integral over the time interval $((j-1)\Delta t, j\Delta t)$ on the right hand side of (15). The naive estimate above, i.e. (34)

and (35), lead to an even more disastrous bound of order N^2/\hbar (there is one factor N that comes for the same reason as in (33), and an additional factor N/\hbar which comes from the estimate (35) in operator norm based on Duhamel formula for $\Delta B_N(s, j)$). Instead, one repeats with $\Delta B_N(s, j)$ the same argument as in the treatment of (18). Since the term $\Delta B_N(s, j)$ is itself the time integral of a commutator involving the random batch potential, the same rarefaction in the relevant commutators $\| [x_k^c, B] \|$ used to control $\| [V_{mn}, B_N(\tau)] \|$ is observed, except that one needs bounds for commutators of the form $\| [x_k^c, \Delta B_N(s, j)] \|$ and not $\| \Delta B_N(s, j) \|$ itself. This is the reason why we need to control iterated brackets of order 2, which is done in section 3.5. The bound (19) and the inequality (22) show that everything can be controlled in terms of the quantity $M_2(\tau)$ defined in (27). Here again, one might suspect that the summation over m, n in (27) would produce the same unpleasant N^2 factor that appears when using (34) and (35), but the bounds (30) and (29) clearly show that this is not the case.

Eventually $M_2(\tau) = O(\hbar^2)$ (uniformly in N) because of the choice of the test operator A in the definition of d_\hbar : here the key estimate is (28), which explains why the sum of N^2 terms in (27) produces a quantity that is bounded uniformly in N . That $M_2(t) = O(\hbar^2)$ follows from the condition on A in the definition of the metric d_\hbar , and this offsets the $1/\hbar^2$ due to the integral over the time interval $((j-1)\Delta t, j\Delta t)$ in (15), and to the additional time integral in the Duhamel formula (21) for $\Delta B_n(j, s)$.

Summarizing the discussion above, the metric d_\hbar is especially designed in order to handle both the large N and the small \hbar issues, i.e. requirements (a) and (b) in the introduction.

5. CONCLUSION AND PERSPECTIVES

The random batch method applied to the N -particle Schrödinger equation reduces the computational cost of the N -particle interaction potential per time-step from $O(N^2)$ to $O(N)$ operations.

We have given an error estimate (Theorem 2.1) for the simplest imaginable random batch method applied to the quantum dynamics of N identical particles. This error estimate has the advantage of being independent of the particle number N and of the Planck constant \hbar (or more precisely of the ratio of the Planck constant to the typical action of one particle in the system). On the other hand, the final estimate is stated in terms of some dual (negative) Sobolev type norm on the difference between the expected single body reduced density operators propagated from the same initial state by the random batch dynamics and by the N -particle dynamics. For the sake of simplicity, we have restricted our attention to the simplest case of batches of two particles only.

The main new mathematical ingredient in the proof is the use of the metric d_\hbar (see Definition 4.1 for the definition of this object, and Proposition 4.2 for its basic properties), which is especially tailored to handle at the same time the difficulties pertaining to the small \hbar regime (classical limit), and those pertaining to the large N regime (mean-field limit). The final statement (Theorem 2.1) of the error estimate does not involve the metric d_\hbar , but is couched in terms of the Wigner transforms [20] of the N -body and random batch density operators, a mathematical object which is familiar in the context of quantum dynamics.

Several extensions of this result should be easily obtained with the mathematical tools used in the present paper. First one can obviously consider batches of $p > 2$ particles; the error analysis is expected to be similar. Also, in practice, the random batch dynamics (4) or (7) is further approximated by some convenient numerical scheme. Of course, the numerical schemes used on (4) or (7) should satisfy the same requirements (a) and (b) (uniform convergence in N and in \hbar) listed in the introduction. For instance, time-splitting strategies for quantum dynamics converge uniformly in \hbar (see [2] and [12]), and could be used together with random batch strategies. The numerical treatment of the space variable x , however, seems much more challenging.

Acknowledgements. The work of Shi Jin was partly supported by NSFC grants No. 11871297 and No. 31571071. We thank E. Moulines for kindly indicating several references on stochastic approximation.

REFERENCES

- [1] F. Bach, E. Moulines: *Non-Asymptotic Analysis of Stochastic Approximation Algorithms for Machine Learning*, in “Proc. Advances in Neural Information Processing Systems (NIPS)”, Granada, Spain, 2011, pp. 451–459.
- [2] W. Bao, S. Jin, P.A. Markowich: *On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime*, J. Comp. Phys. **175** (2002), 487–524.
- [3] M. Benaïm: *Dynamics of stochastic approximation algorithms* in “Séminaire de Probabilités”, Strasbourg, tome 33, (1999), pp. 1–68, Springer-Verlag, Berlin Heidelberg New York, 1999.
- [4] N. Benedikter, V. Jaksic, M. Porta, C. Saffirio, B. Schlein: *Mean-field evolution of fermionic mixed states*, Commun. Pure Appl. Math. **69** (2016), 2250–2303.
- [5] N. Benedikter, M. Porta, C. Saffirio, B. Schlein: *From the Hartree dynamics to the Vlasov equation*, Arch. Ration. Mech. Anal. **221** (2016), 273–334.
- [6] N. Benedikter, M. Porta, B. Schlein: *Mean-field evolution of Fermionic systems*, Commun. Math. Phys. **331** (2014), 1087–1131.
- [7] A. Boulkhémair: *L^2 estimates for Weyl quantization*, J. Functional Anal. **165** (1999), 173–204.
- [8] A. Connes: *Compact metric spaces, Fredholm modules, and hyperfiniteness*, Ergod. Th. and Dynam. Sys. **9** (1989), 207–220.
- [9] A. Connes: “Noncommutative Geometry”, Academic Press, Inc., San Diego, CA, 1994.
- [10] F. D’Andrea, P. Martinetti: *A dual formula for the spectral distance in noncommutative geometry*, preprint [arXiv:1807.06935v1](https://arxiv.org/abs/1807.06935v1) [math.OA].
- [11] R. Durstenfeld: *Algorithm 235: random permutation*, Commun. of the ACM, **7** (1964), 420.
- [12] F. Golse, S. Jin, T. Paul: *On the Convergence of Time Splitting Methods for Quantum Dynamics in the Semiclassical Regime*, preprint [arXiv:1906.03546](https://arxiv.org/abs/1906.03546) [math.NA].
- [13] F. Golse, C. Mouhot, T. Paul: *On the Mean Field and Classical Limits of Quantum Mechanics*, Commun. Math. Phys. **343** (2016), 165–205.
- [14] F. Golse, T. Paul: *The Schrödinger Equation in the Mean-Field and Semiclassical Regime*, Arch. Rational Mech. Anal. **223** (2017), 57–94.
- [15] F. Golse, T. Paul: *Wave Packets and the Quadratic Monge-Kantorovich Distance in Quantum Mechanics*, C. R. Acad. Sci. Paris, Sér. I **356** (2018), 177–197.
- [16] F. Golse, T. Paul: *Empirical Measures and Quantum Mechanics: Applications to the Mean-Field Limit*, Commun. Math. Phys. **369** (2019), 1021–1053.
- [17] S. Jin, L. Li, J.-G. Liu: *Random Batch Method for Interacting Particle Systems*, J. Comput. Phys., **400** (2020), 108877.
- [18] S. Jin, P.A. Markowich, C. Sparber, *Mathematical and computational methods for semiclassical Schrödinger equations*, Acta Numerica, **20** (2011), 121–209.
- [19] H.J. Kushner, G.G. Yin: “Stochastic Approximation and Recursive Algorithms and Applications”, 2nd edition, Springer Verlag, New York, 2003.
- [20] P.-L. Lions, T. Paul: *Sur les mesures de Wigner*, Rev. Math. Iberoam. **9** (1993), 553–618.

- [21] P. Martinetti: *From Monge to Higgs: a survey of distance computations in noncommutative geometry*. In “Noncommutative geometry and optimal transport”, 1–46, Contemp. Math., 676, Amer. Math. Soc., Providence, RI, 2016.
- [22] A. Nemirovski, A. Juditsky, G. Lan, A. Shapiro: *Robust stochastic approximation approach to stochastic programming*, SIAM J. Optimization **19** (2009), 1574–1609.
- [23] C. Villani: “Topics in Optimal Transportation”, American Mathematical Soc, Providence (RI) (2003)
- [24] B. Ying, K. Yuan, S. Vlaski, A. H. Sayed: *Stochastic Learning under Random Reshuffling with Constant Step-sizes*, preprint [arXiv 1803.07964](https://arxiv.org/abs/1803.07964) [cs.LG].

(F.G.) ÉCOLE POLYTECHNIQUE, CMLS, 91128 PALAISEAU CEDEX, FRANCE
Email address: francois.golse@polytechnique.edu

(S.J.) SCHOOL OF MATHEMATICAL SCIENCES, INSTITUTE OF NATURAL SCIENCES, MOE-LSC, SHANGHAI JIAO TONG UNIVERSITY, SHANGHAI 200240, CHINA
Email address: shijin-m@sjtu.edu.cn

(T.P.) SORBONNE UNIVERSITÉ, CNRS, UNIVERSITÉ DE PARIS, INRIA, LABORATOIRE JACQUES-LOUIS LIONS, 75005 PARIS, FRANCE
Email address: thierry.paul@upmc.fr