



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

PROPAGACIÓN DE CAOS PARA SISTEMAS DE
PARTÍCULAS DE INTERACCIÓN DE SALTO PURO

TESIS PARA OPTAR AL GRADO DE DOCTOR EN CIENCIAS DE
LA INGENIERÍA MENCIÓN MODELACIÓN MATEMÁTICA

ROBERTO AMARU CORTEZ MILÁN

PROFESOR GUÍA:
JOAQUÍN FONTBONA TORRES

MIEMBROS DE LA COMISIÓN:
SERVET MARTÍNEZ AGUILERA
ALEJANDRO RAMÍREZ CHUAQUI
DANIEL REMENIK ZISIS
JAIME SAN MARTÍN ARISTEGUI

SANTIAGO DE CHILE
2015

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE: Doctor en Ciencias de la Ingeniería, Mención Modelación Matemática
POR: Roberto Amaru Cortez Milán
FECHA: 26 de octubre de 2015
PROFESOR GUÍA: Joaquín Fontbona

PROPAGACIÓN DE CAOS PARA SISTEMAS DE PARTÍCULAS DE INTERACCIÓN DE SALTO PURO

En la mecánica estadística, la *ecuación de Boltzmann espacialmente homogénea* (en honor de Ludwig Eduard Boltzmann, quien introdujo la primera versión en 1872) describe a nivel macroscópico la evolución temporal de la distribución de las velocidades de una enorme cantidad de moléculas de un gas en \mathbb{R}^3 , las cuales obedecen las leyes de la mecánica clásica y están sujetas a colisiones a nivel microscópico. Ecuaciones de similares características han sido introducidas recientemente en variadas situaciones; por ejemplo, para modelar la redistribución de riqueza en una población, en el contexto de la *Econofísica*.

Con el fin de validar matemáticamente la ecuación de Boltzmann y a la vez deducir propiedades de la misma, en 1956 Kac propuso estudiar un *sistema de partículas*, el cual es un proceso estocástico a valores en $(\mathbb{R}^3)^N$ que representa las velocidades de N partículas que evolucionan continuamente en el tiempo y cambian su estado mediante saltos aleatorios correspondientes a las colisiones entre ellas. Es sabido que este sistema aproxima a la ecuación, en el sentido que se cumple la propiedad de *propagación de caos*: la medida empírica del sistema converge débilmente a la solución de la ecuación en el límite cuando $N \rightarrow \infty$. En los últimos años ha habido gran interés por cuantificar esta convergencia, con dependencia explícita en N y en el tiempo t , e idealmente uniforme en t , pues esto validaría plenamente a la distribución estacionaria de la ecuación como el estado del gas en equilibrio termodinámico.

En la presente tesis se estudia la propagación de caos para algunos sistemas de partículas, incluyendo a los modelos recién descritos. En el Capítulo 2 se trabaja con un sistema a valores en un espacio general, y se obtiene un resultado de propagación de caos en convergencia débil en el espacio de trayectorias. En el Capítulo 3 se estudia una clase de sistemas de partículas en \mathbb{R} que incluye a algunos modelos de redistribución de riqueza y a una versión simplificada de la ecuación de Boltzmann, introducida por Kac. Se desarrolla una técnica de *coupling* que permite obtener resultados de propagación de caos con tasas polinomiales moderadas en N y t . Finalmente, en el Capítulo 4 se utiliza esta técnica en el contexto de la ecuación de Boltzmann y se obtiene el resultado principal de la tesis (el cual mejora significativamente la tasa uniforme obtenida por Mischler y Mouhot en 2013):

Teorema. *Para la ecuación de Boltzmann espacialmente homogénea en el caso de las moléculas de Maxwell, se tiene una tasa uniforme de propagación de caos, en distancia de Wasserstein-2 al cuadrado, de orden casi $N^{-1/6}$.*



A Javier.

Agradecimientos

Gracias a mi familia por estar presente desde siempre.

Gracias a Javier y Nancy por su apoyo durante todo este proceso.

Gracias a Joaquín por su rol como profesor y por su calidez como persona. Gracias por su preocupación por mi bienestar tanto académico como personal, y por su confianza en el trabajo que yo podía realizar. Espero podamos seguir colaborando.

Gracias a todos los amigos de la vida que estuvieron presentes de una u otra forma durante estos años: Raúl, N, V, Cris, Lisset, Nacho, Óscar, Erwin, Mauro, Héctor, Clarita y Natalia. Aunque algunos ahora estén lejos, sé que seguiremos compartiendo momentos.

Agradezco el apoyo económico de:

- Beca de Doctorado Proyecto Mecsup UCH0607,
- Beca de Movilidad Doctoral de Estadías Cortas de Investigación en Francia para Estudiantes Tesistas del Departamento de Postgrado y Postítulo de la Vicerrectoría de Asuntos Académicos y el Instituto Francés de Chile – Embajada de Francia, convocatoria 2013–2014,
- Financiamiento del Núcleo Milenio de Modelos Estocásticos de Sistemas Complejos y Desordenados NC130062,
- Beca CMM para término de tesis.

Tabla de contenido

1	Introducción	1
1.1	La ecuación de Boltzmann	2
1.2	La ecuación de Boltzmann espacialmente homogénea	4
1.3	Sistemas de partículas y propagación de caos	6
1.4	Proceso no-lineal	8
1.5	Avances en propagación de caos	10
1.6	Modelos de intercambio económico	11
1.7	Resultados y plan de la tesis	14
2	Propagación de caos trayectorial en espacios generales	17
2.1	Introducción	17
2.2	Descripción del modelo	18
2.3	Resultados previos	19
2.4	Cálculos preliminares	20
2.5	Nuestro resultado	21
2.6	Construcción del proceso	29
3	Quantitative propagation of chaos for generalized Kac particle systems	33
3.1	Introduction and main result	33
3.1.1	The kinetic equation	33
3.1.2	Particle system and propagation of chaos	34
3.1.3	Main result	35
3.1.4	Particular cases and comparison with known results	38
3.1.5	The nonlinear process and idea of the proof	39
3.1.6	Plan of the paper	41
3.2	Coupling of the particle system and the nonlinear processes	41
3.2.1	The particle system	41
3.2.2	Coupling with the nonlinear processes	42
3.3	Proof of the main result	44
3.4	Proof of intermediate lemmas	49
4	Quantitative uniform propagation of chaos for Maxwell molecules	59
4.1	Introduction and main results	59

Tabla de contenido

4.1.1	The Boltzmann equation	59
4.1.2	Particle system and propagation of chaos	60
4.1.3	Main results	62
4.1.4	Comparison to known results and approaches	65
4.1.5	Weak solutions and nonlinear processes	67
4.1.6	Idea of the proofs and plan of the paper	68
4.2	Construction	69
4.2.1	The particle system	69
4.2.2	Coupling with system of cutoff nonlinear processes	71
4.3	Estimates and technical results	75
4.4	Proof of the main results	83
5	Conclusiones	87
	Bibliografía	91

Capítulo 1

Introducción

La presente tesis doctoral trata sobre las llamadas *ecuaciones cinéticas* y los *sistemas de partículas* asociados a ellas. Estos modelos se utilizan para representar y estudiar la evolución temporal de una gran variedad de fenómenos que aparecen en las ciencias y la ingeniería, cuya característica común es que *están compuestos por una enorme cantidad de objetos similares que interactúan entre sí*. Por razones históricas, el ejemplo más importante proviene de la mecánica estadística, y específicamente de la mecánica de gases, en el cual un gas suele representarse como una colección de moléculas microscópicas confinadas en una región del espacio y que se desplazan e interactúan mediante colisiones, de acuerdo a las leyes de la mecánica clásica. La evolución temporal de la distribución de partículas queda descrita por la *ecuación de Boltzmann*, la cual es parte fundamental de lo que hoy se conoce como *teoría cinética de gases*. El estudio de esta ecuación, iniciada en 1872 por el mismo Boltzmann y que continúa en activo desarrollo hasta el día de hoy, ha generado ramificaciones más allá de la mecánica estadística, encontrando aplicaciones en áreas tales como la Biología, la Economía y las finanzas.

A continuación se revisará brevemente, en orden más o menos cronológico, cómo ha sido el desarrollo del estudio de la ecuación de Boltzmann y sus variantes y su aplicación al contexto económico, destacando los aspectos que son de interés para la presente tesis. Especial énfasis se dará a los sistemas de partículas y la propiedad de *propagación de caos*, la cual es el objeto principal de estudio de esta tesis. Esta revisión se basa principalmente en [61], ver también [14] y [50].

1.1 La ecuación de Boltzmann

La *ecuación de Boltzmann*, introducida por Ludwig Eduard Boltzmann en 1872 [7], es una ecuación integro-diferencial ampliamente utilizada en la mecánica estadística, la cual describe la evolución temporal de la distribución de las posiciones y velocidades en \mathbb{R}^3 de las moléculas de un gas, lo cual permite estudiar sus características macroscópicas, tales como su temperatura, densidad y presión, a partir de las reglas microscópicas de interacción entre las moléculas. Su deducción se basa en algunos principios básicos de modelación:

- el gas está compuesto por una enorme cantidad de moléculas o partículas idénticas,
- las partículas se rigen por las leyes de la mecánica clásica, y están sujetas a interacciones de a pares y elásticas, y
- la evolución temporal de la distribución de las partículas debe emerger como consecuencia de estas interacciones.

Para ser más precisos: consideremos un gas compuesto por N moléculas o partículas en un dominio acotado $\Omega \subseteq \mathbb{R}^3$ sujetas a la siguiente dinámica determinista:

$$\partial_t x_t^i = v_t^i, \quad \partial_t v_t^i = -\frac{1}{\epsilon} \sum_{j \neq i} \nabla J \left(\frac{x_t^i - x_t^j}{\epsilon} \right), \quad (1.1)$$

donde $(x_t^i, v_t^i) \in \Omega \times \mathbb{R}^3$ son la posición y la velocidad de la partícula i -ésima, y $J : \mathbb{R}^3 \rightarrow \mathbb{R}$ es una función de potencial. Adicionalmente, debe imponerse alguna condición de borde apropiada. Cuando J posee soporte compacto, las partículas se desplazan en línea recta con velocidad constante hasta que dos de ellas se encuentran a distancia de orden ϵ y se produce una interacción que altera sus respectivas velocidades. En el caso que las partículas se comportan como *esferas duras* (es decir, cuando $J(x) = \infty$ para $|x| < 1$ y $J(x) = 0$ si $|x| > 1$, informalmente), la interacciones corresponden a reflexiones instantáneas.

Anotemos $z_t^i = (x_t^i, v_t^i)$, y supongamos que en $t = 0$ las partículas z_t^1, \dots, z_t^N se escogen al azar con una cierta distribución simétrica dada $F_0^N \in \mathcal{P}((\Omega \times \mathbb{R}^3)^N)$, de modo que el sistema resulta ser intercambiable ($\mathcal{P}(E)$ denota la colección de medidas de probabilidad sobre el espacio métrico E). Denotemos $F_t^N \in \mathcal{P}((\Omega \times \mathbb{R}^3)^N)$ a la distribución de las partículas en un instante posterior $t > 0$, lo cual corresponde simplemente a hacer evolucionar la distribución inicial F_0^N siguiendo la dinámica descrita por (1.1); anotemos también $F_t^k \in \mathcal{P}((\Omega \times \mathbb{R}^3)^k)$ a la distribución de (z_t^1, \dots, z_t^k) . En este contexto, interesa estudiar la evolución de $F_t^1(dx, dv)$, la cual representa la probabilidad de encontrar una partícula en la posición x con velocidad v en el instante t . En el límite cuando $N \rightarrow \infty$ y manteniendo la relación $N\epsilon^2 = 1$ (lo cual corresponde a que el gas está *suficientemente diluído*, de manera que las interacciones

entre más de 2 partículas pueden despreciarse), bajo algunos supuestos, se espera que la distribución F_t^1 esté bien aproximada por la *ecuación de Boltzmann*:

$$\partial_t f + v \cdot \nabla_x f = \mathcal{Q}(f, f), \quad (1.2)$$

donde $f = f_t(x, v)$ representa la densidad de partículas en la posición x y con velocidad v en el instante t , y $\mathcal{Q}(f, f) = \mathcal{Q}(f_t, f_t)(x, v)$ es el *operador cuadrático de colisión*. Éste viene dado por

$$\mathcal{Q}(f, f) = \frac{1}{2} \int_{\mathbb{R}^3} dv_* \int_{\mathbb{S}^2} d\sigma B(|v - v_*|, \theta) [f' f'_* - f f_*], \quad (1.3)$$

donde, como es estándar en la literatura, anotamos $f_* = f_t(x, v_*)$, $f' = f_t(x, v')$ y $f'_* = f_t(x, v'_*)$, donde v' y v'_* son las *velocidades pre-colisionales*, dadas por

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma, \quad (1.4)$$

y θ es el *ángulo de desviación* entre $v - v_*$ y $v' - v'_*$, definido mediante $\cos \theta = \sigma \cdot (v - v_*) / |v - v_*|$. Notar que esta interacción preserva *momentum* y *energía*, es decir, se cumple $v + v_* = v' + v'_*$ y $|v|^2 + |v_*|^2 = |v'|^2 + |v'_*|^2$. La función $B(|v - v_*|, \theta)$ es el *kernel de colisión*, cuya forma particular depende del tipo de interacciones físicas en estudio, es decir, depende del potencial J en (1.1). Una excelente revisión de esta ecuación desde un enfoque matemático puede encontrarse en [61].

La interpretación de (1.2) es la siguiente: la variación temporal de la densidad de partículas se explica por el transporte de ellas en la variable x , correspondiente al término $v \cdot \nabla_x f$, y por los cambios de velocidades producto de las colisiones en la variable v , codificados en el operador \mathcal{Q} . Éste último se puede separar, de manera directa desde (1.3), en un término de *ganancia* \mathcal{Q}^+ y otro de *pérdida* \mathcal{Q}^- :

$$\mathcal{Q}(f, f) = \mathcal{Q}^+(f, f) - \mathcal{Q}^-(f, f).$$

Informalmente: para (x, v) fijo, el término \mathcal{Q}^- recorre todas las posibles partículas en la posición x y con alguna velocidad v_* que pueden interactuar con las partículas en (x, v) y hacer que éstas cambien su velocidad a alguna otra distinta, produciendo una pérdida de partículas con velocidad v . Similarmente, el término \mathcal{Q}^+ recorre todas las posibles partículas en la posición x con velocidades pre-colisionales v' y v'_* que producto de su interacción hagan que v' adquiera a v como nueva velocidad, produciendo una ganancia de partículas con velocidad v .

Cabe mencionar que esta ecuación fue previamente escrita en una formulación débil por Maxwell, en 1866. Sin embargo, Boltzmann realizó un trabajo considerable respecto a la interpretación y consecuencias de la ecuación, además de hacerla ampliamente conocida en su tratado [7], el cual tendría enorme influencia en la física teórica por varias décadas. En particular, Boltzmann probó el llamado *Teorema-H*, el cual corresponde a la primera demostración matemática del hecho que la entropía de un gas crece con el tiempo.

1.2 La ecuación de Boltzmann espacialmente homogénea

El problema de deducir rigurosamente la ecuación de Boltzmann (1.2) a partir de las N partículas sujetas a la dinámica determinista (1.1) en el límite cuando $N \rightarrow \infty$, es famoso y de larga data. A nivel informal, la manera correcta de realizar este procedimiento límite fue identificado por Grad [36] en 1949, y una formulación matemáticamente correcta del problema fue planteada más tarde por Cercignani [13]. El primer resultado al respecto fue dado por Lanford [44], el cual es válido solamente para tiempos cortos; avances recientes en esa dirección han sido obtenidos por Gallagher, Saint-Raymond y Texier [35]. Sin embargo, el problema original ha permanecido mayormente abierto, lo cual se explica principalmente por el carácter irreversible de la ecuación.

Debido a la dificultad de tratar directamente este problema, es usual considerar la siguiente versión simplificada: suponiendo que la distribución de partículas del gas *no depende de la posición*, es decir, $f = f_t(v)$, desaparece el término de transporte en (1.2), obteniendo la *ecuación de Boltzmann espacialmente homogénea*:

$$\partial_t f = \frac{1}{2} \int_{\mathbb{R}^3} dv_* \int_{\mathbb{S}^2} d\sigma B(|v - v_*|, \theta) [f' f'_* - f f_*]. \quad (1.5)$$

Esta ecuación es el objeto principal de estudio de esta tesis (aunque no el único). Ha sido estudiada extensamente durante las últimas décadas, ver por ejemplo las revisiones de Cercignani [14], Desvillettes [19] y Villani [61]. La naturaleza de las interacciones descritas en (1.4) implica que, al menos informalmente, la solución de (1.5) satisface *conservación de masa, momentum y energía*, es decir, las siguientes expresiones preservan su valor inicial:

$$\int f_t(v) dv, \quad \int v f_t(v) dv, \quad \int |v|^2 f_t(v) dv,$$

por lo cual uno supone naturalmente que $\int f_t(v) dv = 1$.

En términos heurísticos, la ecuación (1.5) describe la evolución de la densidad de velocidades de una cantidad infinita de partículas sujetas a las siguientes interacciones: a tasa $B(|v - v_*|, \theta)$, una partícula con velocidad v colisiona con otra con velocidad v_* con ángulo de desviación θ y, como resultado de la colisión, adquieren nuevas velocidades v' y v'_* dadas por (1.4), donde el vector $\sigma \in \mathbb{S}^2$ se escoge al azar entre todos los que satisfacen $\sigma \cdot (v - v_*) / |v - v_*| = \cos \theta$, ver figura 1.1. Notar que acá v' y v'_* toman el rol de las *velocidades post-colisionales*, lo cual puede hacerse gracias a la simetría de la interacción.

Comúnmente se trabaja con un kernel de colisión que tiene la forma

$$B(|v - v_*|, \theta) \text{ sen } \theta = |v - v_*|^\gamma \beta(\theta),$$

para un cierto $\gamma \in \mathbb{R}$ y función $\beta : [0, \pi] \rightarrow \mathbb{R}_+$. En el caso de *esferas duras*, se cumple que $\gamma = 1$ y $\beta \equiv 1$. Cuando el potencial J de (1.1) es una función radial de la forma $1/r^{s-1}$ para

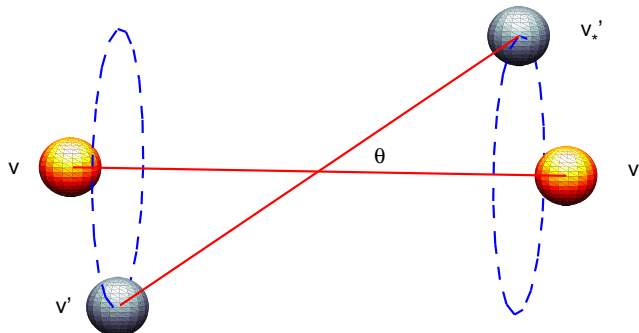


Figura 1.1: Esquema del cambio de velocidades en una interacción en el contexto de la ecuación de Boltzmann espacialmente homogénea. Acá, $v, v_* \in \mathbb{R}^3$ son las velocidades pre-colisionales, y $v', v'_* \in \mathbb{R}^3$ son las respectivas velocidades post-colisionales.

algún $s > 2$ (es decir, las partículas interactúan mediante una fuerza de repulsión de orden $1/r^s$), se tiene que

$$\gamma = \frac{s-5}{s-1} \in (-3, 1), \quad \text{y} \quad \beta(\theta) \sim \theta^{-1-\nu} \text{ cuando } \theta \rightarrow 0, \quad \nu = \frac{2}{s-1} \in (0, 2).$$

Dependiendo del valor de s , distinguimos los siguientes casos: cuando $s = 5$, es decir, fuerza de repulsión de orden $1/r^5$, se tiene que $\gamma = 0$ y $\nu = 1/2$, y hablamos de *moléculas de Maxwell*; notar que en este caso el kernel de colisión depende solamente del ángulo de desviación θ y no de la velocidad relativa de las partículas. El caso en que $s > 5$, es decir, $\gamma \in (0, 1)$ y $\nu \in (0, 1/2)$, se denomina *potenciales duros*; y cuando $2 < s < 5$, es decir $\gamma \in (-3, 0)$ y $\nu \in (1/2, 2)$, hablamos de *potenciales suaves*.

La singularidad no-integrable en θ se debe a la gran cantidad de *colisiones rasantes* (*grazing collisions* en inglés), es decir, colisiones en que las partículas interactúan a una distancia relativamente grande, de manera que apenas desvían su trayectoria. Heurísticamente, esto significa que cada partícula interactúa infinitas veces en cada intervalo de tiempo finito, pero estas interacciones son con ángulos de desviación muy pequeños. Tanto por razones técnicas como algorítmicas, es usual considerar versiones *truncadas* (*cutoff* en inglés) de estos modelos, de manera que $\sup_{\theta} \beta(\theta) < \infty$.

El problema de existencia y unicidad de soluciones para la ecuación de Boltzmann espacialmente homogénea (1.5) se considera mayormente resuelto. Ver por ejemplo [60] para las moléculas de Maxwell, [31, 21] para los potenciales duros, y [4, 48] para las esferas duras. Típicamente se trabaja con una formulación débil de la ecuación, de modo que las soluciones f_t son medidas de probabilidad en \mathbb{R}^3 con segundo momento finito. Utilizaremos la notación $f_t(dv)$ ó $f_t(v)dv$ para referirnos a la distribución f_t , aunque no necesariamente ésta posea

densidad. Condiciones que aseguran que la densidad exista pueden encontrarse por ejemplo en [26, 48, 3].

1.3 Sistemas de partículas y propagación de caos

A pesar de que la versión espacialmente homogénea de la ecuación de Boltzmann tiene una forma más simple que la original, su tratamiento aún ofrecía enormes dificultades, debido a su carácter no-lineal e irreversible. Problemas como existencia, unicidad, comportamiento en el tiempo largo, etc., no tuvieron avances significativos por muchos años.

Por este motivo, en 1956 Kac [42] propuso, como primer paso, estudiar una versión aún más simple de (1.5), donde las velocidades en \mathbb{R}^3 son reemplazadas por velocidades unidimensionales $v \in \mathbb{R}$. La unidimensionalidad de las velocidades es incompatible con la conservación de momentum y energía simultáneamente, por lo que Kac asumió solamente conservación de energía. Específicamente, la heurística en este contexto es la siguiente: las partículas interactúan a tasa constante, y cuando partículas con velocidades $v, v_* \in \mathbb{R}$ colisionan, sus nuevas velocidades vienen dadas respectivamente por

$$v' = (\cos \theta)v - (\sin \theta)v_* \quad \text{y} \quad v'_* = (\sin \theta)v + (\cos \theta)v_*, \quad (1.6)$$

donde el ángulo $\theta \in [-\pi, \pi]$ se escoge uniformemente al azar. La ecuación asociada a lo anterior es la siguiente, y la llamamos *ecuación de Boltzmann-Kac*:

$$\partial_t f = \frac{1}{2\pi} \int_{\mathbb{R}} dv_* \int_{-\pi}^{\pi} d\theta [f' f'_* - f f_*], \quad (1.7)$$

donde, como es usual, denotamos $f = f_t(v)$, $f_* = f_t(v_*)$, $f' = f_t(v')$ y $f'_* = f_t(v'_*)$.

Además de esta ecuación, Kac propuso estudiar una *sistema de N partículas* lineal, con el objetivo de deducir propiedades relevantes de la ecuación a través del sistema en lugar de en la ecuación directamente. Este sistema debía ser distinto al original, sujeto a la dinámica determinista (1.1), pues debido al supuesto de homogeneidad ya no es posible involucrar a la variable espacial en la modelación. Para $N \in \mathbb{N}$ fijo, el sistema de partículas propuesto por Kac es un proceso de Markov $\mathbf{V}_t^N = (V_t^{1,N}, \dots, V_t^{N,N})$ en \mathbb{R}^N que evoluciona como sigue: con tasa $N/2$ se seleccionan dos partículas al azar $v = V_t^{i,N}$ y $v_* = V_t^{j,N}$, se selecciona un ángulo $\theta \in [-\pi, \pi]$ con distribución uniforme, y las velocidades obtenidas producto de la interacción corresponden a v' y v'_* respectivamente, dadas por (1.6). Este procedimiento se repite indefinidamente, donde la elección de $i \neq j$ y de $\theta \in [-\pi, \pi]$ se hace independiente del resto. Alternativamente, el sistema de partículas puede definirse a través de su generador infinitesimal \mathcal{A}^N , el cual está dado por

$$\mathcal{A}^N \Phi(\mathbf{v}) = \frac{1}{4(N-1)\pi} \sum_{i \neq j} \int_{-\pi}^{\pi} d\theta [\Phi(\mathbf{a}_{ij}(\theta, \mathbf{v})) - \Phi(\mathbf{v})]$$

para todo $\mathbf{v} = (v^1, \dots, v^N)$ y toda función test $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}$, donde $\mathbf{a}_{ij}(\theta, \mathbf{v}) \in \mathbb{R}^N$ es el vector con $(\cos \theta)v^i - (\sin \theta)v^j$ en la coordenada i -ésima, $(\sin \theta)v^i + (\cos \theta)v^j$ en la coordenada j -ésima, y v^k en el resto. Además, las partículas parten en $t = 0$ de manera independiente y con ley $f_0(v)dv$. Notar que, producto de las colisiones, las partículas *no son independientes* en instantes $t > 0$.

Al igual que en el caso de la dinámica determinista, se espera que el sistema recién definido esté bien aproximado por la ecuación de Boltzmann–Kac (1.7). En efecto, Kac demostró lo siguiente: para todo $t \geq 0$ y $k \in \mathbb{N}$,

$$\lim_{N \rightarrow \infty} \mathcal{L}(V_t^{1,N}, \dots, V_t^{k,N}) = f_t^{\otimes k} \quad \text{débilmente,} \quad (1.8)$$

donde $\mathcal{L}(\mathbf{X})$ denota la ley de un vector aleatorio \mathbf{X} , $(f_t)_{t \geq 0}$ es la solución de (1.7), y $f_t^{\otimes k}$ denota la medida de probabilidad en \mathbb{R}^k con densidad $f_t(v^1)dv^1 \cdots f_t(v^k)dv^k$. Es decir: para t fijo y N grande, cualquier colección fija de partículas en el sistema se comporta aproximadamente como variables independientes con densidad f_t ; en otras palabras, la independencia en $t = 0$ *se propaga* a tiempos posteriores $t > 0$, en el límite cuando $N \rightarrow \infty$. Por lo anterior, esta propiedad se conoce hoy en día como *propagación de caos* (interpretando “caos” como “independencia”). Es la propiedad fundamental que estudiaremos en esta tesis para variados sistemas de partículas. Para un sistema intercambiable como $V_t^{1,N}, \dots, V_t^{N,N}$, esto equivale a que para todo $t \geq 0$ la medida empírica del sistema en el instante t converge a f_t , es decir,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \delta_{V_t^{i,N}} = f_t \quad (1.9)$$

en distribución, como variables aleatorias en $\mathcal{P}(\mathbb{R})$, ver detalles en [57].

Para la ecuación de Boltzmann espacialmente homogénea (1.5), también es posible definir un sistema de partículas adecuado. En este caso el sistema es un proceso de Markov en $(\mathbb{R}^3)^N$, que también denotamos $\mathbf{V}_t^N = (V_t^{1,N}, \dots, V_t^{N,N})$, con generador infinitesimal dado por

$$\mathcal{A}^N(\Phi)(\mathbf{v}) = \frac{1}{2(N-1)} \sum_{i \neq j} \int_{\mathbb{S}^2} d\sigma B(|v^i - v^j|, \theta) [\Phi(\mathbf{a}_{ij}(\mathbf{v}, \sigma)) - \Phi(\mathbf{v})],$$

donde $\mathbf{a}_{ij}(\mathbf{v}, \sigma) \in (\mathbb{R}^3)^N$ corresponde al vector \mathbf{v} con sus coordenadas i -ésima y j -ésima reemplazadas por v' y v'_* , respectivamente, dadas por (1.4) con $v = v^i$ y $v_* = v^j$. La dinámica del sistema \mathbf{V}_t^N es la siguiente: por cada par de partículas i y j con velocidades $v = V_t^{i,N}$ y $v_* = V_t^{j,N}$, ellas interactúan con ángulo de desviación θ a tasa $B(|v - v_*|, \theta)/[2(N-1)]$, y luego actualizan sus velocidades a v' y v'_* , respectivamente, donde el vector $\sigma \in \mathbb{S}^2$ se escoge al azar entre todos los que satisfacen $\sigma \cdot (v - v_*)/|v - v_*| = \cos \theta$. Como se dijo previamente, cuando $\beta(\theta)$ tiene una singularidad no-integrable en $\theta \sim 0$, cada partícula del sistema cambia su estado infinitas veces en todo intervalo temporal finito, aunque con ángulos de desviación muy pequeños (colisiones rasantes).

La propiedad de propagación de caos de este sistema de partículas con respecto a la distribución límite f_t dada por la ecuación de Boltzmann espacialmente homogénea (1.5) ha sido estudiada extensamente desde el artículo de Kac [42]. Por ejemplo, McKean [46] y Grünbaum [38] obtuvieron propagación de caos para algunos modelos con kernel de colisión acotado, mientras que Sznitman obtuvo un resultado similar para un kernel no-acotado en el caso de esferas duras.

La propagación de caos que entregan estos resultados está en términos de convergencia débil, como se describe en (1.8) o (1.9). Sin embargo, típicamente la convergencia débil no es suficientemente informativa, y uno busca resultados más cuantitativos, con tasas explícitas en N y t con respecto a alguna distancia adecuada para medidas de probabilidad. Más aún: el fin último es probar la *propagación de caos uniforme*, es decir, con una tasa de convergencia explícita en N y que *no depende del tiempo t* . Un resultado de este tipo sería de suma importancia, pues indicaría que la ecuación representa el comportamiento de una gran cantidad de partículas sujetas a colisiones incluso para tiempos muy grandes. Esto significaría, por ejemplo, que el estado estacionario de la ecuación (1.5) efectivamente corresponde a la distribución de velocidades del gas en equilibrio termodinámico.

Numerosos trabajos pueden encontrarse en la literatura respecto al problema de propagación de caos uniforme. Muchos de ellos se centran en el *problema de gap espectral de Kac*, correspondiente a estudiar la tasa a la que el sistema de partículas converge a su distribución de equilibrio en un contexto L^2 (idealmente, con tasa uniforme en la cantidad de partículas), ver por ejemplo [41, 9, 8, 10]. Sin embargo, esta estrategia no ha logrado relacionar el comportamiento asintótico de la solución de la ecuación con el del sistema de partículas.

Cabe mencionar que un sistema de partículas como los recién descritos, en el cual *ambas* partículas involucradas en una interacción cambian su estado, se conoce como sistema *Bird*. En contraste, es posible definir el llamado sistema *Nanbu*, en el cual una sola de las partículas cambia su estado al interactuar con alguna otra; ejemplos de este tipo de sistema pueden encontrarse en [27, 30]. Un sistema de tipo Nanbu es menos pertinente desde el punto de vista físico, aunque de igual manera se espera que cumpla la propiedad de propagación de caos, por lo cual puede utilizarse en simulaciones numéricas para aproximar la solución de la ecuación correspondiente.

1.4 Proceso no-lineal

En 1978, Tanaka [58, 59] introdujo un nuevo objeto de gran importancia teórica y conceptual. Trabajando en el contexto de las moléculas de Maxwell, Tanaka caracterizó el comportamiento límite, cuando $N \rightarrow \infty$, de una partícula del sistema *como proceso*, y no solamente como variable aleatoria en un instante t fijo. El proceso límite lo llamamos *proceso no-lineal*, y puede interpretarse como el comportamiento de una partícula cualquiera inmersa en la población

infinita de partículas que componen el gas.

El proceso no-lineal, que denotamos $(U_t)_{t \geq 0}$, puede definirse de dos formas: mediante un problema de martingala no-lineal (ver (2.5)), o, bajo supuestos adicionales sobre el kernel de colisión, directamente mediante una ecuación diferencial estocástica con respecto a una medida puntual de Poisson (ver (3.9) ó (4.17)). Para simplificar la exposición, por ahora definiremos el proceso de manera informal describiendo su dinámica en el caso de las moléculas de Maxwell: en $t = 0$ se escoge U_0 con ley f_0 ; y para tiempos posteriores, a tasa $\beta(\theta)/\sin \theta$, el proceso $v = U_t$ colisiona con ángulo de desviación θ con una segunda “partícula” v_* *escogida aleatoriamente con ley f_t* , y producto de la colisión el proceso queda con velocidad v' dada por (1.4), donde el vector $\sigma \in \mathbb{S}^2$ se escoge uniformemente al azar entre todos los que cumplen $\sigma \cdot (v - v_*)/|v - v_*| = \cos \theta$. Esto se repite indefinidamente, de manera que todas las elecciones aleatorias, en particular la elección de las variables f_t -distribuidas, se hacen independiente al resto.

El proceso recién descrito satisface que $\mathcal{L}(U_t) = f_t$, por lo cual $(U_t)_{t \geq 0}$ es un objeto más general que la ecuación de Boltzmann espacialmente homogénea. De hecho, usualmente se espera que el sistema de partículas satisfaga la *propagación de caos trayectorial* con respecto a la ley del proceso no-lineal: si anotamos $\mathcal{L}(U)$ a la ley de U en el espacio de trayectorias càdlàg con la topología de Skorokhod, se espera que para todo $k \in \mathbb{N}$ fijo,

$$\lim_{N \rightarrow \infty} \mathcal{L}(V^{1,N}, \dots, V^{k,N}) = \mathcal{L}(U)^{\otimes k} \quad \text{débilmente.}$$

Para una versión de la ecuación de Boltzmann con kernel acotado, Graham y Méléard [37] probaron lo anterior con tasas explícitas en distancia de variación total en el espacio de trayectorias; Desvillettes, Graham y Méléard [20] obtuvieron un resultado similar para una versión singular de la ecuación de Boltzmann-Kac (1.7).

Además de su importancia debido a que es el límite trayectorial de las partículas del sistema (lo cual se estudiará en un contexto general en el Capítulo 2), el proceso no-lineal es relevante para los objetivos de esta tesis porque U_t es la variable f_t -distribuida idónea para obtener tasas de propagación de caos explícitas en *distancia de Wasserstein*. Específicamente: dadas μ, ν medidas de probabilidad en \mathbb{R}^3 , su distancia de Wasserstein de orden $p \geq 1$ se define como

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{X, Y} \mathbb{E}|X - Y|^p \right)^{1/p},$$

donde el ínfimo anterior se toma sobre todas las variables aleatorias X e Y que conforman un *coupling* de μ y ν , es decir, están definidas sobre el mismo espacio de probabilidad y cumplen $\mathcal{L}(X) = \mu$ y $\mathcal{L}(Y) = \nu$ (ver [62] para detalles sobre la distancia de Wasserstein). Por lo tanto, si $(U_t)_{t \geq 0}$ es *cualquier* proceso no lineal definido en el mismo espacio de probabilidad que el sistema de partículas $(V_t^{1,N}, \dots, V_t^{N,N})$, debido a que U_t tiene ley f_t , es directo de la definición de \mathcal{W}_p que

$$\mathcal{W}_p(\mathcal{L}(V_t^{1,N}), f_t)^p \leq \mathbb{E}|V_t^{1,N} - U_t|^p.$$

Lo anterior provee un método para obtener estimaciones en distancia de Wasserstein: basta construir un buen *coupling* entre un proceso no-lineal U_t y un sistema de partículas, de manera que una partícula del sistema y U_t difieran lo menos posible, para obtener una cota superior de \mathcal{W}_p . Este método es utilizado por ejemplo por Graham y Méléard [37] y Fournier y Mischler [30], y será la base de la técnica que se introducirá en los capítulos 3 y 4.

1.5 Avances en propagación de caos

Desde el artículo inicial de Kac [42] ha habido innumerables trabajos respecto a la propagación de caos asociada a la ecuación de Boltzmann espacialmente homogénea y sus variantes. El siguiente es un resumen de los resultados más importantes:

- En las décadas de los '50 y '70, varios autores y para distintas variantes de la ecuación de Boltzmann espacialmente homogénea prueban propagación de caos en términos de convergencia débil, sin tasas explícitas de convergencia. Por ejemplo: Kac [42], McKean [46] y Grünbaum [38] probaron la propagación de caos para algunos modelos con kernel de colisión acotado, mientras que Sznitman [56] obtuvo un resultado similar para un kernel no-acotado en el caso de las esferas duras, todavía sin tasa.
- A fines de los '70, Tanaka [58, 59] introduce el proceso no-lineal y lo utiliza para estudiar el comportamiento en el tiempo largo de la ecuación en el caso de las moléculas de Maxwell. Específicamente, Tanaka muestra que la distancia de Wasserstein-2 entre dos soluciones de la ecuación de Boltzmann partiendo de condiciones iniciales distintas es siempre decreciente, y más aún, toda solución converge cuando $t \rightarrow \infty$ a la distribución gaussiana con el mismo momentum y energía que la condición inicial.
- En los '90, usando técnicas de *coupling* con el proceso no-lineal introducido por Tanaka [58, 59] previamente, Graham y Méléard [37] obtuvieron un resultado de propagación de caos cuantitativo para ciertas variantes truncadas del problema, con dependencia exponencial en el tiempo. Cabe notar que el método utilizado para obtener este resultado, el cual depende fuertemente del hecho que el kernel de colisión está acotado, se basa en un argumento de *coupling* que compara el sistema de partículas y un sistema de procesos no-lineales usando la distancia de *variación total* en el espacio de trayectorias. Por lo tanto, este método no puede usarse para tratar el caso de un kernel no-acotado ni para obtener una dependencia en el tiempo menos que exponencial. Un poco más tarde, Desvillettes, Graham y Méléard [20] probaron la propagación de caos cuantitativa para una versión con kernel no-acotado de la ecuación de Boltzmann-Kac (1.7), pero aún con dependencia exponencial en el tiempo.
- Más recientemente, en el año 2013, Mischler y Mouhot [47] obtuvieron el primer resultado de propagación de caos *uniforme* para la ecuación de Boltzmann espacialmente

homogénea en el caso de las moléculas de Maxwell y de las esferas duras, con una tasa de convergencia lenta en distancia de Wasserstein-1, pero explícita en la cantidad de partículas. A diferencia de los enfoques probabilistas que utilizan técnicas de *coupling* con el proceso no-lineal, su método es de naturaleza principalmente analítica, y se basa en el estudio detallado de la evolución temporal de la ley del sistema en cada instante.

- Por otro lado, en un trabajo reciente, Fournier y Mischler [30] obtuvieron un resultado de propagación de caos en distancia de Wasserstein-2 al cuadrado de orden $t^2 N^{-1/2}$ en el caso de las moléculas de Maxwell. A pesar de no ser uniforme en t , esta tasa es óptima en N , ver [28]. Sin embargo, este trabajo considera únicamente un sistema de partículas de tipo *Nambu*, en el cual solo una partícula cambia su velocidad en cada colisión, lo cual hace al modelo menos pertinente desde el punto de vista físico (aunque sí es interesante numéricamente). La técnica utilizada se basa nuevamente en un argumento de *coupling* con el proceso no-lineal, y las estimaciones de este trabajo serán útiles en el Capítulo 4 para obtener un resultado similar en el contexto del sistema de partículas con colisiones efectivas.
- Uno de los últimos trabajos en el área corresponde a lo realizado por Rousset [55], quien estudia el sistema de partículas correspondiente a las moléculas de Maxwell y prueba un resultado de convergencia al equilibrio en el tiempo largo en distancia de Wasserstein-2, en el espíritu de lo realizado por Tanaka [58, 59]. Esta convergencia es cuantitativa y uniforme en la cantidad de partículas, lo cual será fundamental para obtener, en conjunto con las técnicas desarrolladas en el Capítulo 4, el resultado principal de esta tesis, correspondiente a la propagación de caos cuantitativa y uniforme en el caso de las moléculas de Maxwell, ver Teorema 4.3.

1.6 Modelos de intercambio económico

Durante los últimos 20 años, de manera paralela a (y a la vez inspiradas por) los modelos de la mecánica estadística recién descritos, ha aparecido una (relativamente) nueva gama de ecuaciones con similares características, pero en el contexto de la Economía, como parte de la disciplina hoy conocida como *Econofísica*. Típicamente, estas ecuaciones modelan la evolución temporal de la distribución de riqueza en una población de, digamos, una ciudad o un país. La idea básica de estos modelos económicos es aplicar los mismos principios que en el contexto de la mecánica estadística:

- la población está compuesta por una gran cantidad (infinita) de individuos o *agentes económicos* idénticos,
- se realizan *intercambios económicos* (comercio) entre pares de agentes, y

- la evolución temporal de la distribución de riqueza en la población debe emerger como consecuencia de estas interacciones económicas.

Cada agente de la población posee una cierta cantidad de riqueza $v \in \mathbb{R}$ (en algunos casos se impone $v \geq 0$; en otros se permite que $v < 0$ y se interpreta como deuda). Pares de agentes realizan intercambios económicos utilizando (parte de) sus respectivas riquezas, de acuerdo a reglas aleatorias bien definidas. Ver [50] para una revisión extensa sobre este tema.

Uno de los objetivos principales al estudiar estos modelos es caracterizar la distribución de equilibrio de la ecuación, y específicamente, determinar bajo qué condiciones (a nivel de interacciones entre los individuos) esta distribución posee una *cola pesada* tipo Pareto, como se observa empíricamente en las economías reales. Específicamente, decimos que una distribución posee este tipo de cola cuando la proporción de la población que posee riqueza mayor a v decae como $v^{-\alpha}$, donde $\alpha \geq 1$ es el llamado *índice de Pareto*. Este índice puede interpretarse como una medida de la (in)equidad de la distribución de la riqueza: cuando α es cercano a 1, significa que una minoría de la población acumula la mayor parte de la riqueza y se asocia con economías con desigualdad alta, mientras que un $\alpha \gg 1$ corresponde a economías con desigualdad baja. Por ejemplo, estudios empíricos muestran que la economía de Estados Unidos posee α de orden 1,6 [23], mientras que en Japón se estima que α varía entre 1,8 y 2,2 [32]. Cuando la distribución posee cola que decae exponencialmente (informalmente, el caso $\alpha = \infty$), hablamos de *cola liviana*, y se interpreta como una economía con muy baja desigualdad, lo cual no se condice con los datos de economías reales.

Los primeros estudios en este contexto comenzaron en el año 2000 y se basaron principalmente en simulaciones numéricas, ver por ejemplo los trabajos de Drăgulescu y Yakovenko [22], Chakraborti y Chakraborti [15], y Chatterjee, Chakraborti y Manna [16]. Por ejemplo, en [15] se considera el siguiente modelo de intercambio económico: cuando agentes con riquezas $v, v_* \in \mathbb{R}$ interactúan, sus riquezas producto del intercambio son respectivamente

$$v' = \lambda v + \epsilon(1 - \lambda)(v + v_*), \quad v'_* = \lambda v_* + (1 - \epsilon)(1 - \lambda)(v + v_*), \quad (1.10)$$

donde $\lambda \in [0, 1]$ es una constante llamada *propensión al ahorro*, y ϵ es una variable aleatoria uniforme en $[0, 1]$. En otras palabras: en cada intercambio, los agentes ahorran una proporción λ de su riqueza, y el $1 - \lambda$ restante va a un pozo común que se reparte de manera aleatoria entre ambos. Notar que este modelo es *conservativo*, en el sentido que la riqueza total se conserva en cada intercambio, es decir, $v' + v'_* = v + v_*$.

Posteriormente, cerca del año 2005, se realizaron estudios más rigurosos mediante la introducción de ecuaciones cinéticas correspondientes a distintos tipos de interacciones, incluyendo a las recién descritas, ver por ejemplo Cordier, Pareschi y Toscani [17], Matthes y Toscani [45] y Bassetti, Ladelli y Toscani [5]. El tipo de interacción considerada en estos trabajos generaliza (1.10):

$$v' = Lv + Rv_*, \quad v'_* = \tilde{L}v_* + \tilde{R}v, \quad (1.11)$$

donde L , R , \tilde{L} y \tilde{R} son variables aleatorias en \mathbb{R} con distribución conocida. Lo anterior también puede verse como una generalización de la versión simplificada de la ecuación de Boltzmann introducida por Kac [42]: basta tomar $L = \cos \theta = \tilde{L}$ y $R = -\sin \theta = -\tilde{R}$, con $\theta \in [-\pi, \pi]$ escogido uniformemente, para recuperar (1.6). Consecuentemente, la ecuación cinética que describe la evolución de la distribución de riqueza en la población correspondiente a la interacción (1.11), anotada $f = f_t(v)$, es similar a (1.7):

$$\partial_t f = -f + \mathcal{Q}^+(f, f), \quad (1.12)$$

donde el operador \mathcal{Q}^+ asocia a una medida $\mu \in \mathcal{P}(\mathbb{R})$ una nueva medida $\mathcal{Q}^+(\mu, \mu)$ mediante

$$\int \phi(v) \mathcal{Q}^+(\mu, \mu)(dv) = \iint \frac{1}{2} \mathbf{E}[\phi(Lv + Rv_*) + \phi(\tilde{L}v_* + \tilde{R}v)] \mu(dv_*) \mu(dv),$$

para toda función test ϕ , donde \mathbf{E} denota la esperanza con respecto al vector $(L, R, \tilde{L}, \tilde{R})$.

Este modelo es suficientemente flexible como para admitir distribuciones de equilibrio con variados comportamientos, dependiendo de las características específicas de L , R , \tilde{L} y \tilde{R} . Por ejemplo, las primeras versiones de este modelo trabajan bajo la condición

$$|L|^p + |\tilde{R}|^p = 1 \quad \text{c.s.} \quad \text{y} \quad |R|^p + |\tilde{L}|^p = 1 \quad \text{c.s.}, \quad (1.13)$$

para un cierto exponente $p > 0$ (observar que en el modelo de Kac dado por (1.6) se cumple lo anterior con $p = 2$). Interacciones conservativas, es decir, que satisfacen $v' + v'_* = v + v_*$ (como en (1.10)), corresponden a $p = 1$, con L , R , \tilde{L} y \tilde{R} no-negativos; en tal caso se prueba que la distribución estacionaria de la ecuación necesariamente posee una cola liviana, por lo cual el modelo no se considera realista desde el punto de vista económico. Versiones posteriores trabajan bajo la siguiente condición más débil:

$$\mathbb{E}(|L|^p + |\tilde{R}|^p) = 1 \quad \text{y} \quad \mathbb{E}(|R|^p + |\tilde{L}|^p) = 1, \quad (1.14)$$

la cual, en el caso $p = 1$ con L , R , \tilde{L} y \tilde{R} no-negativos, se interpreta como intercambios *conservativos en media*, en cuyo caso se prueba que la distribución estacionaria puede admitir colas pesadas con un índice de pareto explícito en términos de los momentos de L , R , \tilde{L} y \tilde{R} , ver detalles en [45]. Esta condición corresponde a *intercambios riesgosos*, en que la riqueza total de los agentes involucrados puede no conservarse, produciendo eventuales pérdidas o ganancias netas en cada interacción. Variantes de este y otros modelos han sido introducidas con el fin de obtener una distribución de equilibrio con colas pesadas, ver por ejemplo [17, 49, 51].

Cabe mencionar que hasta la fecha no existen trabajos que estudien de manera rigurosa los sistemas de partículas y la propiedad de propagación de caos asociados a estos modelos cinéticos de redistribución de riqueza; lo único que existe al respecto corresponde a enfoques numéricos como los llevados a cabo en [22, 15, 16]. En el Capítulo 3 de esta tesis se lleva a cabo el primer estudio de propagación de caos cuantitativo (con tasas explícitas) específico

para la ecuación (1.12), el cual ha dado a lugar al trabajo [18]. Sin embargo, los sistemas de partículas sí han sido utilizados en la modelación en otros contextos económicos y/o financieros, ver por ejemplo el trabajo de Föllmer [34], uno de los primeros al respecto. En particular, Remenik [54] estudia un sistema de partículas bajo condiciones muy generales, proveyendo un marco teórico genérico que incluye tanto a los modelos de redistribución de riqueza mencionados más arriba como a otros tipo financiero. Allí se demuestra un resultado de propagación de caos en convergencia débil (sin tasa) para la ley empírica del sistema vista como proceso a valores en el espacio de medidas de probabilidad, lo cual se conoce como *ley funcional de los grandes números*. En el Capítulo 2 de esta tesis se prueba la propagación de caos del sistema con un enfoque *trayectorial*, extendiendo el resultado de Remenik.

1.7 Resultados y plan de la tesis

En la presente tesis se estudia la propiedad de propagación de caos para algunos sistemas de partículas como los descritos en las secciones previas y con distintos grados de generalidad, incluyendo tanto a modelos con aplicaciones económicas y/o financieras como a otros de tipo puramente cinético, en particular, y muy especialmente, la ecuación de Boltzmann espacialmente homogénea (1.5).

En líneas generales, los resultados de esta tesis son principalmente 3:

- (I) *propagación de caos trayectorial con convergencia débil (sin tasa) para un sistema de partículas en un espacio polaco general*, ver Teorema 2.1;
- (II) *propagación de caos cuantitativa para la generalización de la ecuación de Boltzmann–Kac (1.7) dada por (1.12)*, ver Teorema 3.1; y
- (III) *propagación de caos cuantitativa para la ecuación de Boltzmann espacialmente homogénea en el caso de las moléculas de Maxwell*, ver los teoremas 4.2 y 4.3.

Estos puntos serán desarrollados en extenso en los capítulos 2, 3 y 4, respectivamente. A continuación se presenta una descripción general de lo realizado en cada uno de ellos:

Capítulo 2. Se trabaja en el contexto de [54], el cual, debido a su generalidad, incluye a una gran variedad de modelos (económicos, financieros, físicos, etc.). Se considera un sistema de N partículas que toman valores en un espacio polaco general W , las cuales cambian su estado mediante saltos de acuerdo a 2 mecanismos aleatorios: saltos espontáneos de cada partícula e interacciones de a pares de partículas. Las tasas de ambos tipos de salto se asumen uniformemente acotadas, y pueden depender de la medida

empírica del sistema completo, además del estado de la(s) partícula(s) involucrada(s). En [54], bajo hipótesis de continuidad de las tasas de salto, se prueba el siguiente resultado de propagación de caos: la medida empírica del sistema converge en distribución, como elemento (aleatorio) en el espacio de trayectorias a valores en las medidas de probabilidad de W , a un proceso determinista caracterizado por ser la solución a una ecuación integro-diferencial escrita en forma débil. Esto se conoce como *ley funcional de los grandes números*.

En esta tesis se estudia la propagación de caos *trayectorial* de este sistema, por lo cual la medida empírica del mismo es vista como un elemento (aleatorio) en el espacio de medidas de probabilidad sobre el espacio de trayectorias a valores en W , lo cual es estrictamente más general que lo anterior. El objeto límite en este caso debe ser un proceso estocástico (no determinista) a valores en W , y corresponde, naturalmente, al proceso no-lineal. En este contexto general, primero es necesario definir dicho proceso, lo cual se hace mediante un problema de martingala no-lineal, ver (2.5). Luego, para demostrar la propagación de caos deseada, se sigue el argumento clásico utilizado en [37], es decir, se prueba que: (i) hay unicidad de la solución del problema de martingala no-lineal, (ii) la sucesión (en N) de medidas empíricas es tensa, y (iii) todo punto acumulación de dicha sucesión resuelve el problema de martingala no-lineal. El resultado obtenido, es decir, la propagación de caos trayectorial del sistema, corresponde al Teorema 2.1, el cual extiende al Teorema 1 de [54].

Capítulo 3. Se trabaja con la ecuación (1.12), la cual incluye a la ecuación de Boltzmann-Kac (1.7) y a ciertos modelos de redistribución de riqueza como los estudiados en [45]. Se define el correspondiente sistema de partículas y el proceso no-lineal asociado, ver (3.6) y (3.9). Asumiendo una hipótesis similar a (1.14), para $p = 1$ ó $p = 2$, (ver (3.7)), se estudia la propagación de caos del sistema utilizando la distancia de Wasserstein \mathcal{W}_p . Para tal efecto, en esta tesis se sigue una estrategia similar a la utilizada en [30] para un sistema de tipo Nanbu: utilizando una medida puntual de Poisson y mediante herramientas de *coupling* óptimo, se asocia a cada partícula $V^{i,N}$ un proceso no-lineal $U^{i,N}$, para todo $i = 1, \dots, N$, de manera de asegurar que $\mathbb{E}|V_t^{i,N} - U_t^{i,N}|^p$ es pequeño. Sin embargo, para poder usar esto en el contexto de un sistema de tipo Bird como el que se trabaja en el Capítulo 3, se hace necesario introducir una técnica especial de *coupling* que permita tratar con este tipo de colisiones efectivas, ver la construcción de la Sección 3.2. Como consecuencia del tipo de colisiones en consideración, a diferencia de lo que ocurre en [30], los procesos no lineales $U^{1,N}, \dots, U^{N,N}$ que se obtienen *no son independientes*. Luego, para poder obtener estimaciones útiles utilizando esta construcción, es necesario probar que estos procesos no-lineales se vuelven asintóticamente independientes, lo cual se logra con estimaciones uniformes en el tiempo, ver el Lema 3.6. Todo lo anterior permite obtener una tasa de propagación de caos en distancia de Wasserstein- p que depende linealmente en t y de manera polinomial moderada en N , ver Teorema 3.1. Este corresponde al primer resultado de propagación de caos cuantitativa específico para ecuaciones que modelan la redistribución de riqueza. Lo desarrollado en el Capítulo 3 ha dado lugar al trabajo [18], que prontamente aparecerá

publicado en *Annals of Applied Probability*.

Capítulo 4. Se trabaja con la ecuación de Boltzmann espacialmente homogénea (1.5) en el caso de las moléculas de Maxwell. Se estudia el sistema de partículas correspondiente, incluyendo al caso con *cutoff*, y se estudia la propagación de caos en distancia de Wasserstein-2, siguiendo la misma estrategia y estimaciones utilizadas en [30] para el correspondiente sistema de tipo Nanbu. Sin embargo, para tratar con las interacciones de tipo Bird, se hace necesario utilizar la técnica de *coupling* desarrollada en el Capítulo 3, construyendo el sistema de partículas de manera acoplada al correspondiente sistema de procesos no-lineales dependientes, ver Sección 4.2. Esto permite obtener un resultado en distancia de Wasserstein-2 al cuadrado, similar al Teorema 1.4 de [30], con la misma dependencia de orden t^2 en el tiempo, aunque con dependencia en el número de partículas de orden $N^{-1/3}$ en lugar de $N^{-1/2}$, ver Teorema 4.2. Cuando se utilizan las estimaciones del Capítulo 4 en conjunto con los resultados de convergencia al equilibrio de [55], se obtiene el resultado más importante de esta tesis: propagación uniforme de orden casi $N^{-1/6}$, ver Teorema 4.3. Este resultado mejora significativamente la tasa (en N) obtenida por Mischler y Mouhot [47]. Lo desarrollado en el Capítulo 4 será prontamente sometido para revisión y publicación.

Capítulo 2

Propagación de caos trayectorial en espacios generales

2.1 Introducción

Estudiaremos un sistema de N partículas basados en el trabajo de Daniel Remenik [54]. Las partículas pueden cambiar de estado sólo mediante saltos de acuerdo a dos mecanismos: cada partícula por separado decide saltar a otro estado a una cierta tasa (salto espontáneo), o bien la partícula decide interactuar con otra a cierta tasa y conjuntamente escogen dos nuevos estados a los cuales saltar (salto conjunto). Todo esto ocurre de manera aleatoria y continua en el tiempo de acuerdo a tasas y distribuciones de salto dadas, las cuales permitiremos que dependan del estado actual de la o las partículas involucradas en el salto, y también de la ley empírica de todas las partículas en el sistema.

Históricamente, sistemas de este tipo han sido utilizados para aproximar ciertas ecuaciones diferenciales de la mecánica estadística que modelan la dinámica de gases, en particular la ecuación de Boltzman, ver por ejemplo [37]. La interacción entre dos partículas descrita previamente representa una colisión física, y los nuevos estados de ellas corresponden a las nuevas velocidades adquiridas por las partículas producto de la colisión. Sin embargo, hay un reciente interés por aplicar los modelos y ecuaciones de la mecánica estadística en el contexto económico: la idea básica es que las partículas representan agentes económicos y cada colisión corresponde a un intercambio de bienes, e interesa obtener propiedades globales del sistema, como la convergencia de la ley empírica en tiempo largo o cuando la cantidad de agentes es grande. Ver [30, 42, 47] para ejemplos del uso de de los sistemas de partículas en la mecánica estadística, y [15, 5, 49] para su usos en un contexto económico. También pueden encontrarse aplicaciones en la biología, ver por ejemplo [29].

2.2 Descripción del modelo

Dado un espacio métrico E , anotamos $\mathcal{B}(E)$ a la colección de conjuntos medibles de E , o bien la colección de funciones medibles y acotadas de E en \mathbb{R} , según el contexto. $\mathcal{P}(E)$ denota la colección de medidas de probabilidad en $(E, \mathcal{B}(E))$, la cual usualmente dotamos de la topología de la convergencia débil. $D(\mathbb{R}_+, E)$ denota el espacio de trayectorias càdlàg en E dotado de la topología de Skorohod, y análogamente para $D([0, T], E)$. Para $\mu \in \mathcal{P}(E)$ y $f \in \mathcal{B}(E)$, anotamos $\langle \mu, f \rangle = \int_E f d\mu$. $C_b(E)$ denota el espacio de funciones continuas y acotadas de E en \mathbb{R} .

Trabajaremos en un espacio W polaco localmente compacto, que representa el conjunto de los posibles estados de una partícula, de modo que el sistema completo toma valores en W^N . Anotamos $x^N = (x^{1N}, \dots, x^{NN})$ a un elemento genérico de W^N , y $\bar{x}^N = N^{-1} \sum_{i=1}^N \delta_{x^{iN}}$ al elemento de $\mathcal{P}(W)$ correspondiente a la medida empírica del vector x^N . $X = (X_t)_{t \geq 0}$ y $X^N = (X_t^{iN})_{t \geq 0}^{i=1, \dots, N}$ denotan los procesos canónicos en $D(\mathbb{R}_+, W)$ y en $D(\mathbb{R}_+, W^N)$, respectivamente. Por último, denotemos $\bar{X}^N = N^{-1} \sum_{i=1}^N \delta_{X^{iN}}$ a la medida empírica del proceso X^N y $\bar{X}_t^N = N^{-1} \sum_{i=1}^N \delta_{X_t^{iN}}$ a su ley marginal en el tiempo t .

Describamos en mayor detalle la forma en que evoluciona el sistema de N partículas. Si $x^N \in W^N$ es el estado actual del sistema, los mecanismos de cambio de estado son:

Salto espontáneo: una partícula x^{iN} en el estado w decide saltar con tasa $\gamma(w, \bar{x}^N, W)$, y el nuevo estado w' se escoge con ley proporcional a $\gamma(w, \bar{x}^N, dw')$.

Salto conjunto: una partícula x^{iN} en el estado w decide interactuar con una partícula x^{jN} en el estado v con tasa $N^{-1} \lambda(w, v, \bar{x}^N, W \times W)$, y luego escogen conjuntamente nuevos estados w' y v' con ley proporcional a $\lambda(w, v, \bar{x}^N, dw' \otimes dv')$.

Las funciones de tasa $\gamma(w, \mu, dw')$ y $\lambda(w, v, \mu, dw' \otimes dv')$ son no-negativas y están acotadas uniformemente por constantes $\bar{\gamma}$ y $\bar{\lambda}$, respectivamente. Cumplirán condiciones básicas de medibilidad en $w, v \in W$ y de continuidad con respecto a la ley $\mu \in \mathcal{P}(W)$, que especificaremos más adelante. No supondremos simetría de λ .

Concretamente, el sistema de N partículas corresponde a un proceso de Markov en W^N , el cual queda completamente determinado por su generador infinitesimal $\mathcal{A}^N f(x^N)$: para toda función $f \in \mathcal{B}(W^N)$, éste corresponde a

$$\begin{aligned} & \sum_{i=1}^N \int_W [f(x_i^N(w')) - f(x^N)] \gamma(x^{iN}, \bar{x}^N, dw') \\ & + \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \int_{W \times W} [f(x_{i,j}^N(w', v')) - f(x^N)] \lambda(x^{iN}, x^{jN}, \bar{x}^N, dw' \otimes dv'), \end{aligned} \tag{2.1}$$

donde $x_i^N(w') \in W^N$ denota al vector obtenido al reemplazar la i -ésima coordenada de x^N por $w' \in W$. Similarmente, para $i \neq j$, $x_{i,j}^N(w', v') \in W^N$ corresponde a reemplazar las coordenadas i y j de x^N por w' y v' , respectivamente; si $i = j$ ponemos $x_{i,i}^N(w', v') = x_i^N(w')$, lo cual corresponde a permitir auto-interacciones con ley del salto proporcional a $\lambda(x^{iN}, x^{iN}, \bar{x}^N, dw' \otimes W)$. La distribución inicial de cada partícula se escoge de acuerdo a una ley $P_0 \in \mathcal{P}(W)$ dada, de manera independiente. Todo esto define una probabilidad P^N en el espacio $D(\mathbb{R}_+, W^N)$ tal que el proceso $(X_t^{iN})_{t \geq 0}^{i=1, \dots, N}$ es de Markov con respecto a P^N y evoluciona de acuerdo a las reglas de salto descritas previamente; anotamos \mathcal{L}^N y E^N a la ley y esperanza bajo la medida P^N . Con esos supuestos, las partículas son *intercambiables*, es decir, la ley del sistema es la misma bajo cualquier permutación de los índices $\{1, \dots, N\}$.

2.3 Resultados previos

Nos interesa estudiar la convergencia del sistema cuando $N \rightarrow \infty$. En [54] se prueba, bajo hipótesis bastante generales sobre las tasas de salto, que $(\bar{X}_t^N)_{t \geq 0}$, visto como un proceso a valores en $\mathcal{P}(W)$, converge en distribución en $D([0, T], \mathcal{P}(W))$ a un proceso determinista y continuo $(P_t)_{t \geq 0}$ caracterizado por satisfacer una ecuación integro-diferencial escrita en forma débil. Esto se conoce como *ley funcional de los grandes números*.

Una propiedad comúnmente más fuerte que lo anterior sería probar la convergencia en distribución de \bar{X}^N en el espacio $\mathcal{P}(D([0, T], W))$ a una cierta ley determinista que también llamamos P . Dado que las partículas son intercambiables, la Proposición 2.2 de [57] nos dice que esto equivale a que el sistema cumpla la propiedad de *propagación de caos* con respecto a P : para todo $k \in \mathbb{N}^*$ fijo, se cumple

$$\mathcal{L}^N(X^{1N}, \dots, X^{kN}) \rightarrow P^{\otimes k} \quad (2.2)$$

débilmente cuando $N \rightarrow \infty$. Es decir, cualquier colección finita fija de partículas tienden a hacerse independientes e idénticamente distribuidas con ley P a medida que N crece. Un resultado de este tipo entrega más información que la convergencia de $(\bar{X}_t^N)_{t \geq 0}$ como proceso a valores en $\mathcal{P}(W)$, ya que permite aproximar funcionales que dependen de la trayectoria completa de una o varias partículas y no sólo de sus leyes marginales.

En [37], Graham y Méléard trabajan en $W = \mathbb{R}^d$ y agregan un operador diferencial de segundo orden al generador infinitesimal del proceso, de modo de considerar el caso con difusión. Bajo hipótesis de existencia de segundo momento de P_0 y de regularidad del operador diferencial y de las tasas de salto, ellos demuestran la propiedad de propagación de caos utilizando métodos de problemas de martingala. La técnica utilizada sigue la trilogía clásica:

- Unicidad de la solución del problema de martingala límite, el cual caracteriza la ley P .
- Tensión de $\mathcal{L}^N(\bar{X}^N)$ en $\mathcal{P}(\mathcal{P}(D([0, T], \mathbb{R}^d)))$.

- Todo punto de acumulación de $\mathcal{L}^N(\bar{X}^N)$ coincide con δ_P .

Como primera tarea, queremos probar la propagación de caos en nuestro contexto. Utilizaremos la técnica recién descrita basándonos en la demostración del Teorema 4.5 de [37], con adaptaciones en los pasos que haga falta.

2.4 Cálculos preliminares

Para probar la convergencia deseada, necesitaremos introducir notación y objetos adecuados, en particular ciertas martingalas asociadas a cada partícula. Por la teoría general de procesos de Markov, sabemos que para toda función $f \in \mathcal{B}(W^N)$, el proceso $f(X_t^N) - \int_0^t \mathcal{A}^N f(X_s^N) ds$ es una P^N -martingala (ver p.162 de [24]). Dado $\phi \in \mathcal{B}(W)$, aplicando lo anterior a la función $f^{\phi i}(x^N) = \phi(x^{iN})$, desde (2.1) obtenemos la siguiente P^N -martingala partiendo de 0 asociada a ϕ , para todo $i = 1, \dots, N$:

$$\begin{aligned} M_t^{\phi iN} &= \phi(X_t^{iN}) - \phi(X_0^{iN}) - \int_0^t K\phi(X_s^{iN}, \bar{X}_s^N) ds \\ &\quad + \frac{1}{N} \int_0^t \int_W [\phi(w') - \phi(X_s^{iN})] \lambda(X_s^{iN}, X_s^{iN}, \bar{X}_s^N, W \otimes dw') ds, \end{aligned} \tag{2.3}$$

donde hemos introducido el operador $K : \mathcal{B}(W) \times \mathcal{P}(W) \rightarrow \mathcal{B}(W)$ mediante

$$\begin{aligned} K\phi(w, \mu) &= \int_W [\phi(w') - \phi(w)] \beta(w, \mu, dw'), \quad \text{donde} \\ \beta(w, \mu, dw') &= \gamma(w, \mu, dw') + \int_W [\lambda(w, v, \mu, dw' \otimes W) \\ &\quad + \lambda(v, w, \mu, W \otimes dw')] \mu(dv). \end{aligned}$$

Informalmente, cuando $N \rightarrow \infty$ uno espera que el último término en (2.3) desaparezca y que las medidas empíricas \bar{X}^N estén convergiendo, de modo que el “proceso límite” satisfaga una ecuación similar sin dicho término; esto nos lleva naturalmente a considerar un problema de martingala, que introduciremos más abajo. Para estudiar la relación entre las martingalas $M^{\phi iN}$ y $M^{\phi jN}$ (con $i \neq j$) utilizamos su *covariación cuadrática predecible*, que en este caso corresponde a

$$\begin{aligned} &\langle M^{\phi iN}, M^{\phi jN} \rangle_t \\ &= \frac{1}{N} \int_0^t \int_{W \times W} [\phi(w') - \phi(X_s^{iN})][\phi(v') - \phi(X_s^{jN})] \\ &\quad \times \{ \lambda(X_s^{iN}, X_s^{jN}, \bar{X}_s^N, dw' \otimes dv') + \lambda(X_s^{jN}, X_s^{iN}, \bar{X}_s^N, dv' \otimes dw') \} ds. \end{aligned} \tag{2.4}$$

El $1/N$ en la ecuación anterior es un indicio de la propiedad de propagación de caos del sistema, y es lo que nos permitirá desacoplar las partículas cuando $N \rightarrow \infty$.

Por otro lado, consideremos el siguiente *problema de martingala no lineal*: encontrar una ley $P \in \mathcal{P}(D([0, T], W))$, con proyección en $t = 0$ igual al P_0 prescrito, tal que para toda función $\phi \in \mathcal{B}(W)$,

$$M_t^\phi = \phi(X_t) - \phi(X_0) - \int_0^t K\phi(X_s, P_s) ds \quad (2.5)$$

es una P -martingala. La no-linealidad proviene del hecho que P debe hacer martingala a M_t^ϕ , el cual a su vez está definido en términos de P . Por la forma de (2.3), la solución de este problema (suponiendo que existe) será nuestro candidato natural para la ley límite en la propiedad de propagación de caos (2.2).

Si P es solución del problema de martingala no lineal, tomando $E^P(\cdot)$ en (2.5) y notando que $\mathcal{L}(X_t) = P_t$, obtenemos

$$\langle P_t, \phi \rangle = \langle P_0, \phi \rangle + \int_0^t \langle P_s, K\phi(\cdot, P_s) \rangle ds, \quad (2.6)$$

la cual es una ecuación integro-diferencial en forma débil para el flujo de leyes marginales de P , que coincide con la ecuación (S1) de [54]. Usualmente esta ecuación entrega menos información que (2.5), pero es más fácil de trabajar directamente.

2.5 Nuestro resultado

Consideremos las siguientes hipótesis:

(H1) Las medidas $\gamma(w, \mu, \cdot)$ y $\lambda(w, v, \mu, \cdot)$ (en W y $W \times W$ respectivamente) están definidas para todo $\mu \in \mathcal{P}(W)$ y pueden descomponerse como

$$\begin{aligned} \gamma(w, \mu, \cdot) &= \gamma_0(w, \mu) \gamma_1(w, \mu, \cdot) \\ \lambda(w, v, \mu, \cdot) &= \lambda_0(w, v, \mu) \lambda_1(w, v, \mu, \cdot), \end{aligned}$$

donde $\gamma_0(w, \mu) = \gamma(w, \mu, W)$ y $\lambda_0(w, v, \mu) = \lambda(w, v, \mu, W \times W)$ son medibles en w y v y están acotadas uniformemente por $\bar{\gamma} < \infty$ y $\bar{\lambda} < \infty$, respectivamente. $\gamma_1(w, \mu, \cdot)$ y $\lambda_1(w, v, \mu, \cdot)$ son núcleos de transición de $W \times \mathcal{P}(W)$ en W y de $W \times W \times \mathcal{P}(W)$ en $W \times W$, respectivamente.

(H2) Las aplicaciones

$$\begin{aligned} \mathcal{P}(W) \ni \mu &\mapsto \int_W \beta(w, \mu, \cdot) \mu(dw) \in \mathcal{M}(W) \\ \mathcal{P}(W) \ni \mu &\mapsto \beta(\cdot, \mu, W) \mu(\cdot) \in \mathcal{M}(W) \end{aligned}$$

son Lipschitz con respecto a la norma de variación total. Recordemos que ésta se define como $\|\mu\|_{\text{TV}} = \sup_{\|\phi\|_\infty \leq 1} |\langle \mu, \phi \rangle|$. Acá $\mathcal{M}(W)$ denota la colección de medidas finitas y acotadas en W .

(H3) La aplicación

$$W \times \mathcal{P}(W) \ni (w, \mu) \mapsto \beta(w, \mu, \cdot) \in \mathcal{M}(W)$$

es continua cuando dotamos a $\mathcal{P}(W)$ y $\mathcal{M}(W)$ de la topología de la convergencia débil.

Notemos que (H3) implica que las aplicaciones mencionadas en (H2) son además continuas con respecto a la topología de la convergencia débil. Esto significa que nuestras hipótesis implican las del Teorema 1 en [54].

Estamos en condiciones de enunciar nuestro resultado.

Teorema 2.1. *Supongamos las hipótesis (H1), (H2) y (H3). Entonces el sistema posee la propiedad de propagación de caos con respecto a la medida $P \in \mathcal{P}(D([0, T], W))$ que es la única solución del problema de martingala (2.5).*

La demostración pasa por 3 lemas que cubren los pasos de la trilogía clásica, siguiendo la demostración del Teorema 4.5 de [37]. En lo que sigue, anotamos simplemente $D = D([0, T], W)$.

Lema 2.2. *Bajo (H1) y (H2), se cumple unicidad de la solución del problema de martingala (2.5).*

Demostración. Primero probemos la unicidad de las soluciones de (2.6), repitiendo el argumento en la demostración del Teorema 1 de [54]. Para $\mu, \nu \in \mathcal{P}(W)$ y $\phi \in \mathcal{B}(W)$ cualquiera, utilizando la hipótesis (H2) tenemos:

$$\begin{aligned} & |\langle \mu, K\phi(\cdot, \mu) \rangle - \langle \nu, K\phi(\cdot, \nu) \rangle| \\ & \leq \left| \int_W \int_W \phi(w') [\beta(w, \mu, dw')\mu(dw) - \beta(w, \nu, dw')\nu(dw)] \right| \\ & \quad + \left| \int_W \phi(w) [\beta(w, \mu, W)\mu(dw) - \beta(w, \nu, W)\nu(dw)] \right| \\ & \leq \|\phi\|_\infty \left\| \int_W \beta(w, \mu, \cdot)\mu(dw) - \beta(w, \nu, \cdot)\nu(dw) \right\|_{\text{TV}} \\ & \quad + \|\phi\|_\infty \|\beta(\cdot, \mu, W)\mu(\cdot) - \beta(\cdot, \nu, W)\nu(\cdot)\|_{\text{TV}} \\ & \leq C\|\phi\|_\infty \|\mu - \nu\|_{\text{TV}}, \end{aligned}$$

donde C es una constante. Dadas $P, Q \in D([0, T], \mathcal{P}(W))$ dos soluciones de (2.6), aplicando lo anterior a $\mu = P_s$ y $\nu = Q_s$, obtenemos

$$\begin{aligned} |\langle P_t - Q_t, \phi \rangle| & \leq \int_0^t |\langle P_s, K\phi(\cdot, P_s) \rangle - \langle Q_s, K\phi(\cdot, Q_s) \rangle| ds \\ & \leq C\|\phi\|_\infty \int_0^t \|P_s - Q_s\|_{\text{TV}} ds, \end{aligned}$$

y tomando supremo sobre $\|\phi\|_\infty \leq 1$ se obtiene que $u_t = \|P_t - Q_t\|_{TV}$ satisface $u_t \leq C \int_0^t u_s ds$. Aplicando el lema de Gronwall, se concluye que $u_t = 0$, luego $P_t = Q_t$ para todo t . Esto prueba la unicidad de las soluciones de (2.6) en $D([0, T], \mathcal{P}(W))$. Llamemos $(P_t^*)_{t \in [0, T]}$ a esta única solución, en caso que exista.

Luego, cualquier solución P del problema de martingala no lineal (2.5) debe satisfacer la ecuación integro-diferencial (2.6), y entonces sus leyes marginales P_t coinciden con P_t^* . Por lo tanto, P satisface que para toda $\phi \in \mathcal{B}(W)$,

$$\phi(X_t) - \phi(X_0) - \int_0^t K\phi(X_s, P_s^*) ds$$

es una P -martingala. Pero este es un problema de martingala *lineal* y *no homogéneo*, el cual posee solución única gracias al Teorema 7.3 de [24], ya que $K\phi(\cdot, P_t^*)$ es un operador acotado. \square

Lema 2.3. *Bajo (H1), (H2) y (H3), $(\mathcal{L}^N(\bar{X}^N))_{N \in \mathbb{N}}$ es tensa en $\mathcal{P}(\mathcal{P}(D))$.*

Demostración. Dado que las partículas son intercambiables, por la Proposición 2.2 de [57] basta probar que $(\mathcal{L}^N(X^{1N}))_{N \in \mathbb{N}}$ es tensa en $\mathcal{P}(D)$. Utilizaremos el siguiente criterio: de acuerdo al Teorema 13.2 de [6], dado un espacio métrico E y una sucesión $(Y^N)_{N \in \mathbb{N}}$ de elementos aleatorios de $D([0, T], E)$, la colección $(\mathcal{L}(Y^N))_{N \in \mathbb{N}}$ es tensa en $\mathcal{P}(D([0, T], E))$ si y sólo si se satisfacen las siguientes dos condiciones:

(I) Para todo $\varepsilon > 0$ existe $C \subseteq E$ compacto tal que

$$\limsup_{N \rightarrow \infty} \mathbb{P}(\exists t \in [0, T], Y_t^N \notin C) < \varepsilon;$$

(II) Para todo $\varepsilon > 0$

$$\lim_{\delta \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{P}(\omega'(Y^N, \delta) \geq \varepsilon) = 0.$$

Acá $\omega'(z, \delta)$ denota el “módulo de continuidad cádlág” de $z \in D([0, T], E)$, definido como

$$\omega'(z, \delta) = \inf_{\{t_i\}} \sup_i \omega(z, [t_{i-1}, t_i]),$$

donde el ínfimo se toma sobre todas las particiones $0 = t_0 < t_1 < \dots < t_k = T$ con paso mayor que δ , y $\omega(z, I) = \sup_{t, s \in I} \rho(z_t, z_s)$ para $I \subseteq [0, T]$, donde $\rho(\cdot, \cdot)$ es la métrica de E .

Para demostrar las condiciones (I) y (II) en el contexto de la sucesión $(X^{1N})_{N \in \mathbb{N}}$, consideremos el evento $A^{iN\delta}$ en que X^{iN} posee dos o más saltos en $[0, T]$ separados por un tiempo menor a δ . Probaremos que

$$P^N(A^{iN\delta}) = O(\delta) \text{ uniformemente en } N. \tag{2.7}$$

Dejemos pendiente la demostración y utilicemos esto para probar (I) y (II).

Observemos que $\omega'(X^{iN}, \delta) = 0$ fuera del evento $A^{iN\delta}$, pues la partición $\{t_0, \dots, t_k\}$ en la definición de ω' puede escogerse exactamente como los saltos de X^{iN} . Luego, dado $\varepsilon > 0$ tenemos

$$P^N(\omega'(X^{1N}, \delta) \geq \varepsilon) \leq P^N(A^{1N\delta}) + P^N((A^{1N\delta})^c \cap \{\omega'(X^{1N}, \delta) \geq \varepsilon\}) = O(\delta),$$

lo cual prueba (II).

Ahora probemos (I) para la sucesión $(X^{1N})_{N \in \mathbb{N}}$, con $E = W$. Sea $\varepsilon > 0$, y escojamos $k \in \mathbb{N}$ suficientemente grande tal que para $\delta = T/k$ se tiene $P^N(A^{1N\delta}) < \varepsilon/3$ para todo N , lo cual puede hacerse gracias a (2.7).

Notemos que $(\mathcal{L}((\bar{X}_t^N)_{t \in [0, T]}))_{N \in \mathbb{N}}$ es tensa en $\mathcal{P}(D([0, T], \mathcal{P}(W)))$, pues de hecho está convergiendo, de acuerdo al Teorema 1 de [54]. Esto implica que la sucesión de procesos $((\bar{X}_t^N)_{t \in [0, T]})_{N \in \mathbb{N}}$ cumple la condición (I) con $E = \mathcal{P}(W)$, es decir, para todo $\eta > 0$ existe $\mathcal{C} \subseteq \mathcal{P}(W)$ compacto tal que $\limsup_N P^N(\exists t \in [0, T], \bar{X}_t^N \notin \mathcal{C}) < \eta$. Pero los compactos de $\mathcal{P}(W)$ son conjuntos tensos (y cerrados), luego existe $C \subseteq W$ compacto tal que $\mu(C^c) \leq \eta$ para todo $\mu \in \mathcal{C}$. En particular, para $\eta = \varepsilon/[3(k+1)]$ obtenemos un compacto $C = C_\eta$ tal que

$$\limsup_{N \rightarrow \infty} P^N(\exists t \in [0, T], \bar{X}_t^N(C^c) > \eta) < \eta. \quad (2.8)$$

Sea $0 = t_0 < t_1 < \dots < t_k = T$ la malla equiespaciada de $[0, T]$ de paso δ . Notemos que fuera del evento $A^{1N\delta}$, que exista $t \in [0, T]$ tal que $X_t^{1N} \notin C$ equivale a que exista $j \in \{0, \dots, k\}$ tal que $X_{t_j}^{1N} \notin C$. Usando esto y la intercambiabilidad de las partículas, tenemos:

$$\begin{aligned} & P^N(\exists t \in [0, T], X_t^{1N} \notin C) \\ & \leq P^N(A^{1N\delta}) + P^N((A^{1N\delta})^c \cap \{\exists t \in [0, T], X_t^{1N} \notin C\}) \\ & \leq \frac{\varepsilon}{3} + \sum_{j=0}^k P^N(X_{t_j}^{1N} \notin C) \\ & = \frac{\varepsilon}{3} + \sum_{j=0}^k E^N(\bar{X}_{t_j}^N(C^c)) \\ & = \frac{\varepsilon}{3} + \sum_{j=0}^k [E^N(\bar{X}_{t_j}^N(C^c) \mathbf{1}_{\{\bar{X}_{t_j}^N(C^c) \leq \eta\}}) + E^N(\bar{X}_{t_j}^N(C^c) \mathbf{1}_{\{\bar{X}_{t_j}^N(C^c) > \eta\}})] \\ & \leq \frac{\varepsilon}{3} + \sum_{j=0}^k [\eta + P^N(\bar{X}_{t_j}^N(C^c) > \eta)] \\ & \leq \frac{\varepsilon}{3} + (k+1)\eta + (k+1)P^N(\exists t \in [0, T], \bar{X}_t^N(C^c) > \eta). \end{aligned}$$

Tomando \limsup_N y utilizando (2.8) y la elección de $\eta = \varepsilon/[3(k+1)]$, obtenemos que $(X^{1N})_{N \in \mathbb{N}}$ satisface (I), como queríamos.

Por último, probemos (2.7). Utilizando la construcción explícita de X^N presentada en la Sección 2.6, notemos que

$$\begin{aligned} \mathcal{N}^{iN}(ds) &= \mathcal{N}^N(ds, i, [0, 1], [0, 1]) + \mathcal{M}^N(ds, \{i\}, \{1, \dots, N\}, [0, 1], [0, 1]) \\ &\quad + \mathcal{M}^N(ds, \{1, \dots, N\} \setminus \{i\}, \{i\}, [0, 1], [0, 1]) \end{aligned}$$

es una medida puntual de Poisson en $[0, \infty)$ cuyos átomos representan los posibles instantes de salto de la partícula i , tanto espontáneos como conjuntos. La intensidad de \mathcal{N}^{iN} es

$$\left(\bar{\gamma} + \bar{\lambda} + (N-1) \frac{\bar{\lambda}}{N} \right) ds,$$

lo cual está acotado uniformemente en N por $(\bar{\gamma} + 2\bar{\lambda})ds$. Luego, se satisface que la probabilidad de que haya 2 o más átomos de \mathcal{N}^{iN} en $[0, T]$ separados por menos que δ es un $O(\delta)$. Como los saltos de X^{iN} están contenidos en los átomos de \mathcal{N}^{iN} , obtenemos lo deseado. \square

Lema 2.4. *Bajo (H1), (H2) y (H3), existe P solución del problema de martingala no lineal, y todo punto de acumulación débil de $(\mathcal{L}^N(\bar{X}^N))_{N \in \mathbb{N}}$ coincide con δ_P .*

Demostración. El argumento es idéntico al de la demostración del Teorma 4.5 de [37]. Llamemos $\Pi^N = \mathcal{L}^N(\bar{X}^N) \in \mathcal{P}(\mathcal{P}(D))$, las cuales forman una sucesión tensa por el Lema 2.3. Sea entonces Π un punto de acumulación de una subsucesión de $(\Pi^N)_{N \in \mathbb{N}}$ que llamamos igual; queremos probar que Π -casi toda ley $Q \in \mathcal{P}(D)$ resuelve el problema de martingala no lineal. Para esto, consideremos la función $F : \mathcal{P}(D) \rightarrow \mathbb{R}$ dada por

$$F(Q) = \left\langle Q, \left(\phi(X_t) - \phi(X_s) - \int_s^t K \phi(X_r, Q_r) dr \right) g(X) \right\rangle,$$

donde $g : D \rightarrow \mathbb{R}$ corresponde a $g(X) = g_1(X_{s_1}) \cdots g_q(X_{s_q})$, para ciertos $0 \leq s_1, \dots, s_q < s < t$, $\phi, g_1, \dots, g_q \in C_b(W)$ fijos. Queremos probar que F es nula Π -casi seguramente, y la idea es para pasar las propiedades de las martingalas $M^{\phi iN}$ al límite. Sin embargo, como las proyecciones en s_1, \dots, s_q, s, t no son continuas con respecto a la topología de Skorohod, la función F no es continua. Para salvar este inconveniente, consideremos el conjunto

$$J = \{r \in [0, T] : \Pi(Q : Q(|\Delta X_r| > 0) > 0) > 0\},$$

donde $|\Delta X_r| = \rho(X_r, X_{r-})$, siendo ρ la métrica de W . Veremos que J es a lo más numerable, y F será continua Π -casi seguramente si los instantes s_1, \dots, s_q, s, t se escogen fuera de él. Para mayor claridad, separamos el argumento en pasos.

1. J es numerable. Consideremos

$$J_k = \{r \in [0, T] : \Pi(Q : Q(|\Delta X_r| > 1/k) > 1/k) > 1/k\},$$

y veamos que J_k es finito: si no lo fuera, existe una sucesión $r_n \in J_k$ convergiendo de manera propia y monótona a un cierto $r \in [0, T]$. Llamemos

$$A_n = \{Q : Q(|\Delta X_{r_n}| > 1/k) > 1/k\}$$

y sea $A = \limsup_n A_n$, luego $\Pi(A) \geq 1/k$. Sea $B_n = \{|\Delta X_{r_n}| > 1/k\}$ y $B = \limsup_n B_n$. Dado que todo $Q \in A$ está en infinitos conjuntos A_n , es claro que $Q(B) \geq 1/k$. Pero notemos que toda trayectoria en B no posee límite por la izquierda o por la derecha en r (dependiendo del crecimiento o decrecimiento de r_n). Esto significa que Π está asignando masa estrictamente positiva sobre un conjunto de medidas Q que a su vez asignan masa a trayectorias fuera de D , lo cual contradice el hecho que $\Pi \in \mathcal{P}(\mathcal{P}(D))$.

Además, se tiene que $J = \bigcup_k J_k$. En efecto, dado $r \in J$ se tiene lo siguiente para cierto k grande:

$$\begin{aligned} \frac{1}{k} &< \Pi(Q : Q(|\Delta X_r| > 0) > 0) \\ &= \lim_{l \rightarrow \infty} \Pi(Q : Q(|\Delta X_r| > 0) > 1/l), \end{aligned}$$

donde hemos utilizado la continuidad de la probabilidad Π . Luego, para l suficientemente grande podemos eliminar el límite de la desigualdad anterior, obteniendo

$$\frac{1}{k} < \Pi(Q : Q(|\Delta X_r| > 0) > 1/l),$$

posiblemente redefiniendo k . Notemos que los eventos $\{|\Delta X_r| > 1/l\}$ convergen crecientemente a $\{|\Delta X_r| > 0\}$ cuando $l \rightarrow \infty$, y por continuidad de las medidas Q se tiene que $\{Q : Q(|\Delta X_r| > 1/l) > 1/k\}$ crecen a $\{Q : Q(|\Delta X_r| > 0) > 1/k\}$. Aplicando Π en la convergencia anterior, nuevamente por continuidad de la probabilidad se tiene que

$$\Pi(Q : Q(|\Delta X_r| > 1/l) > 1/k) \rightarrow \Pi(Q : Q(|\Delta X_r| > 0) > 1/k).$$

Pero como el lado derecho es mayor que $1/k$, para l grande el lado izquierdo también lo será. Nuevamente redefiniendo k si hace falta, obtenemos

$$\Pi(Q : Q(|\Delta X_r| > 1/k) > 1/k) > \frac{1}{k},$$

lo cual muestra que $r \in J_k$ y entonces $J \subseteq \bigcup_k J_k$. La otra inclusión es directa, con lo cual se tiene la igualdad. Como J_k es finito, hemos probado que J es a lo más numerable, como deseábamos.

2. Π -continuidad de F . Específicamente, afirmamos que si los instantes $I = \{s_1, \dots, s_q, s, t\}$ se escogen fuera de J , entonces F es Π -casi seguramente continua. Para verlo, consideremos el conjunto $A_I = \{Q : \forall r \in I, Q(|\Delta X_r| > 0) = 0\}$, el cual tiene Π -medida 1, pues

$$\Pi(A_I^c) \leq \sum_{r \in I} \Pi(Q : Q(|\Delta X_r| > 0) > 0),$$

lo cual vale 0 por definición de J . Por otro lado, la aplicación $\int_s^t K\phi(z_r, Q_r)dr$ es continua como función del par $(z, Q) \in D \times \mathcal{P}(D)$. En efecto: si $z^n \rightarrow z$ en D y $Q^n \rightarrow Q$ en $\mathcal{P}(D)$, se tendrá que $z_r^n \rightarrow z_r$ y $Q_r^n \rightarrow Q_r$ para todo r salvo una cantidad a lo más numerable (ver p.134 y 138 en [6]). Para todo tal r , usando la hipótesis (H3) y la continuidad de ϕ , tenemos que

$$\begin{aligned} K\phi(z_r^n, Q_r^n) &= \int_W [\phi(w') - \phi(z_r^n)]\beta(z_r^n, Q_r^n, dw') \\ &= \langle \beta(z_r^n, Q_r^n, \cdot), \phi \rangle - \phi(z_r^n)\langle \beta(z_r^n, Q_r^n, \cdot), 1 \rangle \\ &\rightarrow \langle \beta(z_r, Q_r, \cdot), \phi \rangle - \phi(z_r)\langle \beta(z_r, Q_r, \cdot), 1 \rangle \\ &= K\phi(z_r, Q_r). \end{aligned}$$

Es decir, $K\phi(z_r^n, Q_r^n) \rightarrow K\phi(z_r, Q_r)$ para casi todo $r \in [s, t]$. Como ϕ y β están acotados uniformemente, por convergencia dominada obtenemos la continuidad deseada de $\int_s^t K\phi(z_r, Q_r)dr$. Si definimos la aplicación

$$(z, R) \mapsto h(z, R) = \left(\phi(z_t) - \phi(z_s) - \int_s^t K\phi(z_r, R_r)dr \right) g(z),$$

la continuidad de ϕ, g_1, \dots, g_q y la elección de los instantes s_1, \dots, s_q, s, t fuera de J implican que h es $Q \otimes \delta_Q$ -casi seguramente continua, para cualquier $Q \in A_I$ fijo. Luego, si Q^n converge débilmente a Q , se tendrá que $Q^n \otimes \delta_{Q^n}$ converge a $Q \otimes \delta_Q$ (Teorema 2.8 [6]), y entonces $\langle Q^n \otimes \delta_{Q^n}, h \rangle$ converge a $\langle Q \otimes \delta_Q, h \rangle$. Pero para todo $R \in \mathcal{P}(D)$, se tiene que $\langle R \otimes \delta_R, h \rangle$ es exactamente $F(R)$, lo cual prueba la continuidad de F en todo $Q \in A_I$, es decir, F es continua Π -casi seguramente.

3. Caracterización de Π . Finalmente, probemos que Π -casi todo $Q \in \mathcal{P}(D)$ resuelve el problema de martingala no lineal. Como F es Π -continua, también lo es F^2 , y entonces $\Pi^N(F^2)^{-1}$ converge débilmente a $\Pi(F^2)^{-1}$ como distribuciones en \mathbb{R} . Esto implica que $\langle \Pi, F^2 \rangle \leq \liminf_N \langle \Pi^N, F^2 \rangle$, lo cual puede verse fácilmente con el Teorema de Representación de Skorohod (Teorema 6.7 de [6]): existen variables aleatorias $Z, (Z_N)_{N \in \mathbb{N}}$ definidas sobre un espacio de probabilidad común $(\Omega, \mathcal{F}, \mathbb{P})$ tales que $\Pi(F^2)^{-1} = \mathcal{L}(Z)$, $\Pi^N(F^2)^{-1} = \mathcal{L}(Z_N)$, y $Z_N(\omega) \rightarrow Z(\omega)$ para todo $\omega \in \Omega$. Aplicando el Lema de Fatou, se obtiene la desigualdad:

$$\langle \Pi, F^2 \rangle = \int_{\Omega} Z(\omega)\mathbb{P}(d\omega) \leq \liminf_{N \rightarrow \infty} \int_{\Omega} Z_N(\omega)\mathbb{P}(d\omega) = \liminf_{N \rightarrow \infty} \langle \Pi^N, F^2 \rangle.$$

Recordemos que $\Pi^N = \mathcal{L}^N(\bar{X}^N)$. Luego:

$$\begin{aligned}
 \langle \Pi^N, F^2 \rangle &= E^N \langle \bar{X}^N, h(X, \bar{X}^N) \rangle^2 \\
 &= E^N \left[\frac{1}{N} \sum_{i=1}^N h(X^{iN}, \bar{X}^N) \right]^2 \\
 &= \frac{1}{N^2} \sum_{i=1}^N E^N h(X^{iN}, \bar{X}^N)^2 + \frac{1}{N^2} \sum_{i \neq j} E^N h(X^{iN}, \bar{X}^N) h(X^{jN}, \bar{X}^N) \\
 &= \frac{1}{N} E^N h(X^{1N}, \bar{X}^N)^2 + \frac{N-1}{N} E^N h(X^{1N}, \bar{X}^N) h(X^{2N}, \bar{X}^N),
 \end{aligned}$$

donde en el último paso hemos utilizado la intercambiabilidad de las partículas. La primera esperanza está acotada gracias a que ϕ , g , λ y γ lo están, con lo cual el primer término converge a 0 cuando $N \rightarrow \infty$. Sin embargo, en el segundo término necesitaremos utilizar (2.4) para hacer aparecer un $1/N$. Recordemos que la martingala $M^{\phi iN}$ tiene la forma

$$M_t^{\phi iN} = \phi(X_t^{iN}) - \phi(X_0^{iN}) - \int_0^t K \phi(X_r^{iN}, \bar{X}_r^N) dr + \frac{1}{N} \int_0^t \tilde{K} \phi(X_r^{iN}, \bar{X}_r^N) dr,$$

donde \tilde{K} es el operador que aparece en el último término de (2.3). Con esto, la segunda esperanza queda:

$$\begin{aligned}
 &E^N h(X^{1N}, \bar{X}^N) h(X^{2N}, \bar{X}^N) \\
 &= E^N \left(M_t^{\phi 1N} - M_s^{\phi 1N} - \frac{1}{N} \int_s^t \tilde{K} \phi(X_r^{1N}, \bar{X}_r^N) dr \right) g(X^{1N}) \\
 &\quad \times \left(M_t^{\phi 2N} - M_s^{\phi 2N} - \frac{1}{N} \int_s^t \tilde{K} \phi(X_r^{2N}, \bar{X}_r^N) dr \right) g(X^{2N}) \\
 &= E^N (M_t^{\phi 1N} - M_s^{\phi 1N}) (M_t^{\phi 2N} - M_s^{\phi 2N}) g(X^{1N}) g(X^{2N}) \\
 &\quad - \frac{2}{N} E^N (M_t^{\phi 1N} - M_s^{\phi 1N}) g(X^{1N}) g(X^{2N}) \int_s^t \tilde{K} \phi(X_r^{2N}, \bar{X}_r^N) dr \\
 &\quad + \frac{1}{N^2} E^N g(X^{1N}) g(X^{2N}) \int_s^t \tilde{K} \phi(X_r^{1N}, \bar{X}_r^N) dr \int_s^t \tilde{K} \phi(X_r^{2N}, \bar{X}_r^N) dr.
 \end{aligned}$$

Nuevamente, gracias al $2/N$ y $1/N^2$ y a las cotas uniformes en ϕ , g , γ y λ , el segundo y tercer término convergen a 0. El primer término es más delicado: condicionando en \mathcal{F}_s^N (siendo $(\mathcal{F}_r^N)_{r \geq 0}$ la filtración natural de X^N) y sabiendo que g depende sólo de instantes anteriores a s , obtenemos

$$\begin{aligned}
 &E^N (M_t^{\phi 1N} - M_s^{\phi 1N}) (M_t^{\phi 2N} - M_s^{\phi 2N}) g(X^{1N}) g(X^{2N}) \\
 &= E^N \left[g(X^{1N}) g(X^{2N}) E^N [(M_t^{\phi 1N} - M_s^{\phi 1N}) (M_t^{\phi 2N} - M_s^{\phi 2N}) \mid \mathcal{F}_s^N] \right] \\
 &= E^N \left[g(X^{1N}) g(X^{2N}) E^N [\langle M^{\phi 1N}, M^{\phi 2N} \rangle_t - \langle M^{\phi 1N}, M^{\phi 2N} \rangle_s \mid \mathcal{F}_s^N] \right] \\
 &= E^N \left[g(X^{1N}) g(X^{2N}) E^N \left[\frac{1}{N} \int_s^t \hat{K} \phi(X_r^{1N}, X_r^{2N}, \bar{X}_r^N) dr \mid \mathcal{F}_s^N \right] \right],
 \end{aligned}$$

donde \hat{K} es el operador que aparece a la derecha en (2.4). Usando el $1/N$ y las cotas uniformes, este término se va a 0 cuando $N \rightarrow \infty$.

Con todo lo anterior, tenemos que $\langle \Pi^N, F^2 \rangle$ converge a 0, y entonces

$$\langle \Pi, F^2 \rangle \leq \liminf_{N \rightarrow \infty} \langle \Pi^N, F^2 \rangle = 0,$$

es decir F es nula Π -casi seguramente. Esto quiere decir que Π -casi todo $Q \in \mathcal{P}(D)$ satisface

$$E^Q \left[\left(\phi(X_t) - \phi(X_s) - \int_s^t K \phi(X_r, Q_r) dr \right) g_1(X_{s_1}) \cdots g_q(X_{s_q}) \right] = 0,$$

y por la arbitrariedad de $g_1, \dots, g_q \in C_b(W)$ y de los instantes $s_1, \dots, s_q < s < t$ en el conjunto denso $[0, T] \setminus J$, esto significa que Q resuelve el problema de martingala no lineal (que $Q_0 = P_0$ es directo). Esto prueba la existencia de soluciones, y como ya sabemos que se cumple unicidad, necesariamente $\Pi = \delta_P$, siendo P la única solución. \square

Demostración (Teorema 2.1). Gracias a los lemas previos, $(\mathcal{L}^N(\bar{X}^N))_{N \in \mathbb{N}}$ converge débilmente a δ_P , donde P es la única solución del problema de martingala no lineal (2.5). Es decir, $(\bar{X}^N)_{N \in \mathbb{N}}$ converge en distribución en $\mathcal{P}(D)$ a la medida determinista P . Como mencionamos previamente, esto equivale a que el sistema cumpla la propiedad de propagación de caos con respecto a P , gracias a la Proposición 2.2 de [57] y a la intercambiabilidad de las partículas. \square

2.6 Construcción del proceso

A continuación daremos una construcción explícita del proceso X^N definido mediante su generador infinitesimal (2.1). La idea es definir el proceso a partir de los átomos de ciertas medidas puntuales de Poisson adecuadas, que contendrán toda la aleatoriedad necesaria para describir el sistema X^N .

Específicamente, consideremos los siguientes tres objetos construidos de manera independiente en algún espacio de probabilidad suficientemente grande:

- Variables aleatorias $X_0^{1N}, \dots, X_0^{NN}$ a valores en W con distribución P_0 e independientes, correspondientes a la condición inicial del sistema.
- Una medida puntual de Poisson $\mathcal{N}^N(ds, di, d\theta, d\rho)$ en $[0, \infty) \times \{1, \dots, N\} \times [0, 1]^2$ con medida de intensidad $\bar{\gamma} ds (\sum_{k=1}^N \delta_k(di)) d\theta d\rho$. \mathcal{N}^N definirá los saltos espontáneos de las partículas.

- Una medida puntual de Poisson $\mathcal{M}^N(ds, di, dj, d\theta, d\rho)$ en $[0, \infty) \times \{1, \dots, N\}^2 \times [0, 1]^2$ con medida de intensidad $N^{-1}\bar{\lambda}ds(\sum_{k,l=1}^N \delta_{(k,l)}(di, dj))d\theta d\rho$. \mathcal{M}^N definirá los saltos conjuntos de las partículas.

Definimos también la filtración $(\mathcal{F}_t^N)_{t \geq 0}$ generada por estos objetos, y anotamos \mathbb{P} y \mathbb{E} la probabilidad y esperanza en el correspondiente espacio de probabilidad. En palabras, \mathcal{N}^N escoge átomos al azar en $[0, \infty)$ con tasa $N\bar{\gamma}$, y por cada uno de ellos, también escoge un índice $i \in \{1, \dots, N\}$ al azar y dos variables aleatorias θ y ρ independientes y uniformes en el intervalo $[0, 1]$. La componente i corresponderá al índice de la partícula que saltará espontáneamente; θ entregará la aleatoriedad necesaria para escoger el nuevo estado de la partícula, y ρ se utilizará para decidir si es que corresponde o no realizar dicho salto espontáneo. Similarmente para \mathcal{M}^N .

Las medidas puntuales \mathcal{N}^N y \mathcal{M}^N pueden consolidarse en una sola medida puntual \mathcal{O}^N mediante el siguiente procedimiento:

- Fijemos un $j_0 \in \{1, \dots, N\}$ arbitrario. A cada átomo (s, i, θ, ρ) de \mathcal{N}^N le agregamos la marca determinista $(j_0, 0)$, obteniendo el átomo $(s, i, j_0, \theta, \rho, 0)$. Esto genera una medida puntual de Poisson $\tilde{\mathcal{N}}^N(ds, di, dj, d\theta, d\rho, db)$ en $[0, \infty) \times \{1, \dots, N\}^2 \times [0, 1]^2 \times \{0, 1\}$ con intensidad

$$\bar{\gamma}ds \left(\sum_{k=1}^N \delta_{(k,j_0)}(di, dj) \right) d\theta d\rho \delta_0(db).$$

- A cada átomo (s, i, j, θ, ρ) de \mathcal{M}^N le agregamos la marca determinista 1, obteniendo el átomo $(s, i, j, \theta, \rho, 1)$. Esto genera una medida puntual de Poisson $\tilde{\mathcal{M}}^N(ds, di, dj, d\theta, d\rho, db)$ en el mismo conjunto $[0, \infty) \times \{1, \dots, N\}^2 \times [0, 1]^2 \times \{0, 1\}$ con intensidad

$$N^{-1}\bar{\lambda}ds \left(\sum_{k,l=1}^N \delta_{(k,l)}(di, dj) \right) d\theta d\rho \delta_1(db).$$

- Definimos¹ $\mathcal{O}^N = \tilde{\mathcal{N}}^N \cup \tilde{\mathcal{M}}^N$. Por independencia, se obtiene que $\mathcal{O}^N(ds, di, dj, d\theta, d\rho, db)$ es una medida puntual de Poisson en $[0, \infty) \times \{1, \dots, N\}^2 \times [0, 1]^2 \times \{0, 1\}$, cuya intensidad es la suma de las respectivas intensidades de $\tilde{\mathcal{N}}^N$ y $\tilde{\mathcal{M}}^N$, es decir

$$dsd\theta d\rho \left[\bar{\gamma} \left(\sum_{k=1}^N \delta_{(k,j_0)}(di, dj) \right) \delta_0(db) + N^{-1}\bar{\lambda} \left(\sum_{k,l=1}^N \delta_{(k,l)}(di, dj) \right) \delta_1(db) \right].$$

La coordenada $b \in \{0, 1\}$ indica la procedencia de cada átomo (\mathcal{N}^N cuando $b = 0$, ó \mathcal{M}^N cuando $b = 1$).

¹Anotamos "U" para referirnos a la suma de medidas, lo cual, en el contexto de medidas puntuales, equivale a unir sus soportes.

Por otro lado, notemos que gracias al Lema 3.22 de [43], existen aplicaciones medibles

$$\begin{aligned}\Gamma &: W \times \mathcal{P}(W) \times [0, 1] \rightarrow W, & \text{y} \\ \Lambda &: W^2 \times \mathcal{P}(W) \times [0, 1] \rightarrow W^2\end{aligned}$$

tales que si θ es una variable uniforme en $[0, 1]$, entonces $\Gamma(w, \mu, \theta)$ tiene distribución $\gamma_1(w, \mu, \cdot)$ y $\Lambda(w, v, \mu, \theta)$ tiene distribución $\lambda_1(w, v, \mu, \cdot)$, para todo $w, v \in W$, $\mu \in \mathcal{P}(W)$. Con estas aplicaciones podremos utilizar la aleatoriedad de la componente θ de \mathcal{O}^N para definir el o los nuevos estados de la o las partículas involucradas en cada salto.

Ahora, sean $(s_k, i_k, j_k, \theta_k, \rho_k, b_k)_{k=1}^\infty$ los átomos de \mathcal{O}^N ordenados temporalmente, es decir, $0 := s_0 < s_1 < s_2 < s_3 < \dots$. Con estos átomos, podemos construir el sistema de partículas $X_t^N = (X_t^{1N}, \dots, X_t^{NN})$ partiendo de la condición inicial $X_0^N = (X_0^{1N}, \dots, X_0^{NN})$ previamente escogida, y definiendo $(X_{s_k}^N)_{k=1}^\infty$ recursivamente mediante el siguiente procedimiento:

- Caso $b_k = 0$: si $\rho_k > \gamma_0(X_{s_{k-1}}^{i_k N}, \bar{X}_{s_{k-1}}^N) / \bar{\gamma}$, definimos $X_{s_k}^N = X_{s_{k-1}}^N$ (no se ejecuta el salto espontáneo). En otro caso, definimos

$$X_{s_k}^{\ell N} = \begin{cases} \Gamma(X_{s_{k-1}}^{i_k N}, \bar{X}_{s_{k-1}}^N, \theta_k) & \text{si } \ell = i_k, \\ X_{s_{k-1}}^{\ell N} & \text{en otro caso.} \end{cases}$$

- Caso $b_k = 1$: si $\rho_k > \lambda_0(X_{s_{k-1}}^{i_k N}, X_{s_{k-1}}^{j_k N}, \bar{X}_{s_{k-1}}^N) / \bar{\lambda}$, definimos $X_{s_k}^N = X_{s_{k-1}}^N$ (no se ejecuta el salto conjunto). En otro caso, anotando (Λ_1, Λ_2) las funciones componentes de Λ , definimos:

$$X_{s_k}^{\ell N} = \begin{cases} \Lambda_1(X_{s_{k-1}}^{i_k N}, X_{s_{k-1}}^{j_k N}, \bar{X}_{s_{k-1}}^N, \theta_k) & \text{si } \ell = i_k, \\ \Lambda_2(X_{s_{k-1}}^{i_k N}, X_{s_{k-1}}^{j_k N}, \bar{X}_{s_{k-1}}^N, \theta_k) & \text{si } \ell = j_k, \\ X_{s_{k-1}}^{\ell N} & \text{en otro caso.} \end{cases}$$

- Finalmente, se define $X_s^N = X_{s_{k-1}}^N$ para todo $s \in (s_{k-1}, s_k)$.

A pesar de que este procedimiento parece complicado, el principio básico es simple: por cada átomo de \mathcal{N}^N se realiza un salto espontáneo con probabilidad proporcional al valor de γ_0 en el estado actual del sistema, escogiendo la partícula que salta al azar, y cuyo nuevo estado se escoge con la ley deseada utilizando la función Γ . Análogamente para los átomos de \mathcal{M}^N , correspondientes a los saltos conjuntos del sistema.

Mediante un cálculo de rutina, se verifica que $(X_t^N)_{t \geq 0}$ es un proceso de Markov con respecto a la filtración $(\mathcal{F}_t^N)_{t \geq 0}$ y que posee generador infinitesimal dado por (2.1), como deseábamos.

Chapter 3

Quantitative propagation of chaos for generalized Kac particle systems

This chapter is largely based on the article [18], which has been accepted for publication on Annals of Applied Probability.

3.1 Introduction and main result

3.1.1 The kinetic equation

We consider the collection $(P_t)_{t \geq 0}$ of probability measures on \mathbb{R} , solution of the following nonlinear kinetic-type equation:

$$\partial_t P_t = -P_t + \mathcal{Q}^+(P_t). \quad (3.1)$$

Here \mathcal{Q}^+ is a generalized Wild convolution, which associates with every measure μ on \mathbb{R} a new measure $\mathcal{Q}^+(\mu)$ given by

$$\int \phi(u) \mathcal{Q}^+(\mu)(du) = \iint \frac{1}{2} \mathbf{E}(\phi(Lu + Rv) + \phi(\tilde{L}v + \tilde{R}u)) \mu(dv) \mu(du), \quad (3.2)$$

for all bounded measurable functions ϕ , where $(L, R, \tilde{L}, \tilde{R})$ is a given random vector in \mathbb{R}^4 (with known distribution) and \mathbf{E} denotes the expectation with respect to it.

Equations (3.1)-(3.2) describe the behaviour of an infinite number of objects or “particles” subjected to binary interactions. The state of each particle is characterized by a scalar $u \in \mathbb{R}$, and $P_t(du)$ represents the proportion of particles in state u at time $t \geq 0$. The microscopic binary interactions, which occur randomly at constant rate, are heuristically described as

follows: when a particle at state u interacts with a particle at state v , their states change according to the rule

$$(u, v) \mapsto (Lu + Rv, \tilde{L}v + \tilde{R}u). \quad (3.3)$$

This model is a generalization of Kac's one dimensional simplification of the (more realistic) Boltzmann equation for a spatially homogeneous dilute gas in \mathbb{R}^3 , in which the interacting objects represent actual physical particles. Specifically, in Kac's model introduced in [42], the state of a particle is its one dimensional velocity, and the interactions correspond to random exchanges of velocities that occur at binary collisions that preserve kinetic energy, so that $L = \cos \theta = \tilde{L}$, $R = -\sin \theta = -\tilde{R}$, with θ randomly chosen in $[0, 2\pi)$. We refer the reader to [47] and the references therein for historical background on Kac and Boltzmann's equations.

A further source of models of the type described by equations (3.1)-(3.2) is the kinetic description of the evolution of the wealth distribution in a simplified economy, studied for instance in [45] (see also the references therein). In that setting the state of a particle represents the wealth of an economic agent, and the binary interactions correspond to trades or economic exchanges between them. Early versions of that model assumed

$$|L|^p + |\tilde{R}|^p = 1 \quad \text{a.s.}, \quad |\tilde{L}|^p + |R|^p = 1 \quad \text{a.s.}, \quad (3.4)$$

for some $p \geq 1$ (notice that in Kac's model (3.4) is satisfied with $p = 2$). In the case $p = 1$, for non-negative L, R, \tilde{L} and \tilde{R} , condition (3.4) can be seen as *exact* conservation of total wealth in each interaction. The weaker condition

$$\mathbf{E}(|L|^p + |\tilde{R}|^p) = 1, \quad \mathbf{E}(|\tilde{L}|^p + |R|^p) = 1, \quad (3.5)$$

interpreted as conservation of wealth only *in the mean* (so that risky trades with possible gain or loss of total wealth in each interaction are allowed), has also been considered in order to obtain wider classes of equilibrium distributions for the nonlinear dynamics (see [45], [5]).

3.1.2 Particle system and propagation of chaos

In order to rigorously justify the interpretation of the model (3.1)-(3.2) as representing the evolution of an infinite number of interacting particles or agents, one considers a finite system of N of such particles, which we denote $\mathbf{X}_t = (X_t^1, \dots, X_t^N)$, starting independently with common law P_0 and such that, at each binary interaction, the states of both involved particles are modified according to the rule (3.3). In the terminology of particle approximations of the Boltzmann equation, a particle system with such (true) binary interactions is called of *Bird type*, as opposed to particle systems of *Nanbu type*, in which only one particle changes its state after interaction with some other.

Specifically, the particle system \mathbf{X} has infinitesimal generator

$$\mathcal{A}^N \phi(\mathbf{x}) = \frac{1}{2(N-1)} \sum_{i \neq j} \int_{\mathbb{R}^4} [\phi(\mathbf{x} + a_{ij}(\eta, x^i, x^j)) - \phi(\mathbf{x})] \Lambda(d\eta) \quad (3.6)$$

for all $\mathbf{x} = (x^1, \dots, x^N) \in \mathbb{R}^N$ and every test function ϕ on \mathbb{R}^N , where $\eta = (\xi, \zeta, \tilde{\xi}, \tilde{\zeta})$ denotes a generic point in \mathbb{R}^4 , Λ is the joint law of $(L, R, \tilde{L}, \tilde{R})$ and $a_{ij}(\eta, u, v)$ is the vector of \mathbb{R}^N whose i -th and j -th components are $(\xi - 1)u + \zeta v$ and $(\tilde{\xi} - 1)v + \tilde{\zeta}u$, respectively, and which is equal to 0 in the other components.

Convergence of such a particle system, more precisely of its empirical measures $\frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$ towards the unique solution $(P_t)_{t \geq 0}$ of the nonlinear evolution (3.1) as N goes to infinity, has been studied in more general frameworks and from several points of view, see for instance [37], [47] and the references therein (in particular, well posedness of (3.1) is by now standard). Since the particles are exchangeable, the convergence of the empirical measure to P_t for large N , as a random variable in the space of probability measures in \mathbb{R} endowed with the weak topology, is equivalent to the property of *propagation of chaos* of \mathbf{X}_t with respect to P_t (see [57] for background): for every fixed $k \in \mathbb{N}$, the joint law of X_t^1, \dots, X_t^k converges weakly to $P_t^{\otimes k}$ as N goes to ∞ . That is, when N is large, any fixed number of particles of the system behaves at time t approximately like independent random variables of law P_t . This property was introduced and first established by Kac himself in [42] for the particle system bearing his name, and is nowadays known to hold for a large class of particle models, under general mild assumptions.

3.1.3 Main result

Typically, weak convergence results are not sufficiently informative, and one looks for more quantitative statements. In this article we will study the Bird-type N -particle system $\mathbf{X} = (X^1, \dots, X^N)$ and its propagation of chaos property, in the cases $p = 1$ and $p = 2$. Our main goal is to obtain rates of convergence, as $N \rightarrow \infty$, for the Wasserstein distance between the empirical measure of the particle system at time t and its limiting law P_t , with explicit estimates on N and t that grow reasonably fast as functions of t .

Let $p \in \{1, 2\}$ be fixed. In the case $p = 2$, we will assume the additional condition $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$, which is certainly satisfied in Kac's model. As a generalization of (3.5), we will work under the assumption

$$\frac{1}{2} \mathbf{E}(|L|^p + |R|^p + |\tilde{L}|^p + |\tilde{R}|^p) \leq 1. \quad (3.7)$$

With some abuse of language, for each value of $p \in \{1, 2\}$ we will say that the model is *inelastic* if the latter inequality is strict. In that case, the interaction between particles produce an average loss of energy when $p = 2$ (see for instance the inelastic Kac model in [52]) or of “wealth” (in the context of [45]) when $p = 1$. Also, to avoid trivial situations, in all what follows we will assume that the model is non-degenerate, that is, $\mathbf{E}(|R| + |\tilde{R}|) > 0$; this means that the system produces at least some effective interactions.

Let us fix some notation. $\mathcal{P}(E)$ denotes the space of probability measures on the metric

space E . For $\mathbf{x} \in \mathbb{R}^N$ and any $i = 1, \dots, N$ we define the empirical measures $\bar{\mathbf{x}} = \frac{1}{N} \sum_j \delta_{x^j}$ and $\bar{x}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}$, both being elements of $\mathcal{P}(\mathbb{R})$. Define $M_q(\mu) = \int |u|^q \mu(du)$ the absolute q -moment of $\mu \in \mathcal{P}(\mathbb{R})$. Given a random vector \mathbf{Z} on \mathbb{R}^N we denote its law by $\mathcal{L}(\mathbf{Z}) \in \mathcal{P}(\mathbb{R}^N)$, and the joint law of its first k components by $\mathcal{L}^k(\mathbf{Z}) \in \mathcal{P}(\mathbb{R}^k)$.

Recall that for $\mu, \nu \in \mathcal{P}(\mathbb{R}^k)$ their p -Wasserstein distance $\mathcal{W}_p(\mu, \nu)$ is defined to be the cost of the optimal transfer plan between μ and ν , that is,

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\pi} \int_{\mathbb{R}^k \times \mathbb{R}^k} d_{k,p}(\mathbf{x}, \mathbf{y})^p \pi(d\mathbf{x}, d\mathbf{y}) \right)^{1/p} = \left(\inf_{\theta, \vartheta} \mathbb{E} d_{k,p}(\theta, \vartheta)^p \right)^{1/p},$$

where the first infimum is taken over all measures π on $\mathbb{R}^k \times \mathbb{R}^k$ with marginals μ and ν , and the second infimum is taken over all pairs of random vectors θ and ϑ such that $\mathcal{L}(\theta) = \mu$ and $\mathcal{L}(\vartheta) = \nu$ (see for instance [62] for background on Wasserstein distances). We will use the *normalized* distance $d_{k,p}$ on \mathbb{R}^k given by

$$d_{k,p}(\mathbf{x}, \mathbf{y}) = \left(\frac{1}{k} \sum_{i=1}^k |x^i - y^i|^p \right)^{1/p}, \quad (3.8)$$

which is natural when one cares about the dependence on the dimension.

In order to obtain good rates of convergence in N which moreover are well behaved with respect to t , it is convenient to introduce the concave function

$$\alpha_q = 1 - \frac{1}{2} \mathbf{E}(|L|^q + |R|^q + |\tilde{L}|^q + |\tilde{R}|^q), \quad \forall q \geq 0.$$

We also define

$$q^* = \sup\{q : M_q(P_0) < \infty, \alpha_q > 0\}.$$

These objects play an important role in [45], since, when $p = 1$ and q^* is non-trivial (that is, $1 < q^* < \infty$), q^* corresponds to the Pareto index of the stationary distribution of P_t . More importantly in the present context, the moments of order $q < q^*$ of P_t can be controlled uniformly in time (see Lemma 5 below). Assuming (3.7) and $M_p(P_0) < \infty$, the concavity of α_q implies that either $q^* \in [p, \infty]$ or $q^* = -\infty$. Also, define for all $q \in \{p\} \cup (p, q^*)$

$$\bar{\alpha}_{p,q} = \inf_{p \leq r \leq q} \alpha_r = \min(\alpha_p, \alpha_q).$$

Note that if $\alpha_p = 0$, then $\bar{\alpha}_{p,q} = 0$ for all such q , so this function is meaningful only in the case $\alpha_p > 0$, in which case it will be useful to obtain uniform (in time) estimates.

We are now ready to state our main theorem (see also Corollary 3.8 for a trajectorial result):

Theorem 3.1. *Let $(P_t)_{t \geq 0}$ be the unique solution of (3.1) and let \mathbf{X} be the particle system starting with law $P_0^{\otimes N}$ and with generator (3.6). For $p = 1$ or $p = 2$, assume $\alpha_p \geq 0$ and $M_p(P_0) < \infty$. If $p = 2$, assume also that $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$ and $q^* > 2$. Then*

- for any $q \in \{1\} \cup (1, q^*)$ and any $\gamma < (2 + 1/q)^{-1}$ in the case $p = 1$, or
- for any $q \in (2, q^*)$, $q \neq 4$ and for $\gamma = \min(1/3, \frac{q-2}{2q-2})$ in the case $p = 2$,

there exists a constant C , depending on p, q, γ and some moments of P_0 and $(L, R, \tilde{L}, \tilde{R})$ of order at most q , such that:

(i) for all $k \leq N$ and for all $t \geq 0$,

$$\mathcal{W}_p^p(\mathcal{L}^k(\mathbf{X}_t), P_t^{\otimes k}) \leq C \left(\frac{t(1+t)^{p-1} e^{-\frac{p}{q}\bar{\alpha}_{p,q}t}}{N^\gamma} + \mathbf{1}_{k \neq 1} \frac{k \min(1, t) e^{-\alpha_p t}}{N} \right),$$

(ii) for all $t \geq 0$,

$$\mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{X}}_t, P_t) \leq \frac{C(1+t)^p e^{-\frac{p}{q}\bar{\alpha}_{p,q}t}}{N^\gamma}.$$

Remark 3.2. • The power γ in Theorem 3.1 is a consequence of using recently established sharp quantitative estimates in Wasserstein distance for the empirical measures of exchangeable or i.i.d. collections of random variables (which improve or extend a classical result in [53]). More specifically, the rate $N^{-\gamma}$ with $\gamma < (2 + 1/q)^{-1}$ in the case $p = 1$ comes from Theorem 1.2 of [40], whereas the value $\gamma = \min(1/3, \frac{q-2}{2q-2})$ in the case $p = 2$ comes from Theorem 1 of [28]. On the other hand, the dependence on t results from our estimates, which rely on Gronwall's lemma.

- The restriction $q \neq 4$ in the case $p = 2$ comes from Theorem 1 of [28]. As those authors mention, the case $q = 4$ would produce additional logarithmic terms, which in our case translate into a rate of order $N^{-1/3}$ times a logarithmic function of N .
- In the elastic case (that is, $\alpha_p = 0 = \bar{\alpha}_{p,q}$), (i) and (ii) give estimates that grow linearly with time (in the case $p = 2$ both sides are squared). In the inelastic case, which corresponds to $\alpha_p, \bar{\alpha}_{p,q} > 0$, all estimates are uniform in time.
- From a physical point of view, it is interesting to consider models where infinitely many particles interact over finite time intervals, such as the Kac equation without cutoff. The techniques used in the proof of Theorem 3.1 can also be applied to cutoffed approximations of that equation and, in the case that (in the notation of [20]) the classical condition $\int_0^\pi \theta^2 \beta(\theta) d\theta < \infty$ on the cross-section function $\beta : [-\pi, \pi] \rightarrow \mathbb{R}_+$ is satisfied, they yield a constant that does not depend on the cutoff parameter, see Remark 3.9.

3.1.4 Particular cases and comparison with known results

The Kac equation

Note that if the stronger condition (3.4) is satisfied (or holds with \leq instead of equality), then $|L|$, $|R|$, $|\tilde{L}|$ and $|\tilde{R}|$ are all ≤ 1 a.s., which implies that α_q is strictly increasing with q . Thus, $\bar{\alpha}_{p,q} = \alpha_p$ for $q \geq p$ and the value of q^* will depend only on the finiteness of the moments of P_0 . In Kac's model, since (3.4) is satisfied for $p = 2$, if P_0 has finite moment of order $4 + \varepsilon$, then $q^* > 4$ and Theorem 3.1 gives

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{X}}_t, P_t) \leq \frac{C(1+t)^2}{N^{1/3}}.$$

Several similar results can be found in the literature. The closest one corresponds to quantitative rates for the Nanbu system associated with Kac's model, which are found for instance in the proof of Proposition 6.2 of [27]. The authors state there a \mathcal{W}_2^2 convergence rate that also depends quadratically on t and is optimal on N , in the sense that it is equal to the \mathcal{W}_2^2 rate of convergence of the empirical measure of an i.i.d. sample towards their common law. The latter is of order $N^{-1/2}$, according to Theorem 1 of [28]. Thus, the Bird-type particle system seems to produce a slower rate of convergence than the corresponding Nanbu-type system. An interesting question is whether this difference is a mere consequence of the techniques used in our proof (more specifically, some order is lost when one uses Lemma 3.7) or is intrinsically related to the type of binary interactions (Bird or Nanbu) in the system.

A similar result as the one of [27] can be found in [30] where, motivated by the numerical approximation of the Boltzmann equation for hard spheres, hard potentials and Maxwellian gases, a pathwise coupling argument was developed for Nanbu particle systems, which extends a coupling construction based on optimal transport developed in [25]. That pathwise approach however does not readily extend to the particle systems of Bird type we are interested in, which in turn provide a physically more transparent description of the relevant interaction phenomena.

As for the Bird particle system, in [37] the authors obtain an explicit rate in total variation distance on the path space, between the law of one particle and the law of the nonlinear process (to be introduced later). However, due to the generality of their hypotheses and the strong pathwise distance they use, the convergence rate depends exponentially on the length of the time interval that is considered. Similarly, in Theorem 4.3 of [20] the authors state a propagation of chaos result in \mathcal{W}_2 for the law at time t of one particle in the system with cutoff, towards the law P_t of the nonlinear dynamics without cutoff. Since some relations between N and the cutoff parameter must be satisfied when removing the latter, that result gives estimates that are logarithmic in N and grow exponentially with t .

On the other hand, the general theory developed in [47] provides a framework and a methodology to establish quantitative (in t and N) propagation of chaos estimates which

can be applied in the present framework. For instance, in their Theorem 5.2, a \mathcal{W}_1 estimate for the Boltzmann equation in the Maxwell molecules case is obtained, which is uniform in time and decays with N in a polynomial way (see also Step 3 of the proof of Theorem 8 in [12] for results in \mathcal{W}_2 distance); we expect that similar bounds can be obtained with their techniques for the Kac model. The actual dependence on N they give seems however hard to trace in general, and we have not been able to deduce with their techniques an estimate in Wasserstein distance as sharp as ours in terms of N . Also, their approach does not provide any information on the way in which trajectories of particles get closer to those of the limiting processes. On the other side, unfortunately our techniques (ultimately relying on Gronwall's lemma) do not seem to yield uniform in time estimates for the elastic Kac equation, even if P_0 were compactly supported.

Finally, we observe that for the inelastic Kac model, $\bar{\alpha}_{2,q} = \alpha_2 > 0$ for all q , hence Theorem 3.1 does give a uniform-in-time rate in that case:

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{X}}_t, P_t) \leq \frac{C e^{-\frac{\alpha_2}{2+\varepsilon}t}}{N^{1/3}},$$

if P_0 is again assumed to have finite moment of order $q > 4$. This exponential decay is not surprising: when P_0 has finite moment of order 2 it is known that $M_2(P_t)$ decays exponentially fast, see [52]. Nevertheless, to our knowledge, our quantitative (in N) propagation of chaos result is new for the inelastic Kac model.

Models for economic exchanges and wealth distribution

Working with $p = 1$ as in [45] and assuming only that the first moment of P_0 is finite, Theorem 3.1 gives in the elastic case a propagation of chaos result for \mathcal{W}_1 of order almost $N^{-1/3}$, with estimates growing linearly with time. Any additional finite q -moment of P_0 (with $q < q^*$) can be used to improve the rate in N , up to almost $N^{-1/(2+1/q^*)}$. In the case of exact conservation of wealth (condition (3.4)) we have $q^* = \infty$ and we obtain a rate of $N^{-(1/2-\varepsilon)}$, which is almost optimal according to Theorem 1 of [28]. To our knowledge, this is the first quantitative propagation of chaos result for kinetic equations modeling the evolution of wealth distribution.

3.1.5 The nonlinear process and idea of the proof

Following ideas pioneered by Tanaka in the case of the Boltzmann equation (see [58] and [59]), it is also possible to establish the convergence of the *pathwise* law of a particle, to the law of some process obtained by the following construction: consider a Poisson point measure \mathcal{M} on $\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}$ with intensity $dt\bar{\Lambda}(d\xi, d\zeta)P_t(dv)$, where $\bar{\Lambda} = \frac{1}{2}(\mathcal{L}(L, R) + \mathcal{L}(\tilde{L}, \tilde{R}))$, and

let $(V_t)_{t \geq 0}$ be the jump process on \mathbb{R} defined as the unique solution starting with law P_0 of the stochastic equation

$$dV_t = \int_{\mathbb{R}^2} \int_{\mathbb{R}} [(\xi - 1)V_{t-} + \zeta v] \mathcal{M}(dt, d\xi, d\zeta, dv). \quad (3.9)$$

It is not hard to see that such a jump process V exists, it is uniquely defined and it satisfies $\mathcal{L}(V_t) = P_t$ for all $t \geq 0$. We call P the pathwise law of V , and any process with law P is called a *nonlinear process*; it represents the trajectory of any fixed particle in the (infinite) population subjected to the random interactions described above in (3.3).

To prove our results, we will couple the Bird particle system \mathbf{X}_t with a system $\mathbf{U}_t = (U_t^1, \dots, U_t^N)$ where each U^i is a copy of the nonlinear process V , constructed in such a way that it remains close to X^i . To achieve this, we will use techniques of optimal coupling inspired by those used in [25] and [30], in order to carefully choose the jumps of the nonlinear process U^i as similar as possible to those of the particle X^i . However, contrary to those papers which deal with Nanbu-type particle systems (in which each randomness source acts on the trajectory of only one of the particles), ensuring closeness of X^i and U^i simultaneously for all $i = 1, \dots, N$ will imply that the processes U^1, \dots, U^N are *not* independent. Therefore, to obtain the desired estimates we will need, in a second step, to “decouple” the system \mathbf{U}_t as N goes to infinity, which we will be able to do with estimates that are uniform in time, see Lemma 3.6 below.

Let us point out that the coupling construction we will introduce can in principle be replicated in higher dimensions, and with more general interaction rules, which is why we preferred to avoid the use of specific one-dimensional features in its construction, see for instance Remark 3.10. We thus expect these techniques to be applicable in physically more relevant situations, hopefully including (at least some instances of) the Boltzmann equation. Also, we think it should be possible to adapt this coupling construction in order to quantitatively study “Bird-type” Brownian particle approximations of a certain Gaussian white-noise driven nonlinear process, associated with the Landau equation arising in the grazing collisions limit of the Boltzmann equation. Such a process was studied in [33] and [39], and a particle approximation result with a “Nanbu type” Brownian particle system was proved in [25], by means of a coupling construction based on optimal transport. The corresponding particle system of Bird-type is studied in [11] using the functional tools developed in [47], but there seems to be so far no suitable coupling argument available in order to deal with such class of particle systems.¹

¹When the present work was just finished, the authors learnt from Nicolas Fournier that the latter question was currently being studied by him, François Bolley and Arnaud Guillin.

3.1.6 Plan of the paper

In Section 3.2 we give the explicit construction of the particle system \mathbf{X}_t , and more importantly, we couple it with the system $\mathbf{U}_t = (U_t^1, \dots, U_t^N)$ of dependent nonlinear processes that we will use throughout the rest of this article. In Section 3.3 we prove Theorem 3.1. The proof of some intermediate lemmas, including statements of Section 3.2 and the “decoupling” of the process \mathbf{U}_t , is left for the final Section 3.4.

3.2 Coupling of the particle system and the nonlinear processes

3.2.1 The particle system

Let us fix the number of particles $N \in \mathbb{N}$. Although most of the subsequent objects will depend on N , for notational simplicity we will not make this dependence explicit. We will define both the particle system \mathbf{X} and the nonlinear processes \mathbf{U} by means of integral equations driven by the same Poisson point measure. To this end, let us first introduce the function $\mathbf{i} : [0, N] \rightarrow \{1, \dots, N\}$ given by $\mathbf{i}(\rho) = \lfloor \rho \rfloor + 1$, and the set $\mathcal{C} \subseteq [0, N]^2$

$$\mathcal{C} = \{(\rho, \sigma) \in [0, N]^2 : \mathbf{i}(\rho) \neq \mathbf{i}(\sigma)\}.$$

Note that $|\mathcal{C}| = N(N - 1)$, see Figure 3.1. As in (3.6), denote $\eta = (\xi, \zeta, \tilde{\xi}, \tilde{\zeta})$ a generic point in \mathbb{R}^4 and $\Lambda = \mathcal{L}(L, R, \tilde{L}, \tilde{R})$. Now, let $\mathcal{N}(dt, d\eta, d\rho, d\sigma)$ be a Poisson point measure on $[0, \infty) \times \mathbb{R}^4 \times [0, N]^2$ with intensity

$$\frac{N}{2} dt \Lambda(d\eta) d\rho d\sigma \frac{1}{|\mathcal{C}|} \mathbf{1}_{\mathcal{C}}(\rho, \sigma) = \frac{1}{2(N - 1)} dt \Lambda(d\eta) d\rho d\sigma \mathbf{1}_{\mathcal{C}}(\rho, \sigma).$$

In words, \mathcal{N} picks atoms in $[0, \infty)$ at constant rate of $N/2$, and for each such atom it also independently samples a tuple $(\xi, \zeta, \tilde{\xi}, \tilde{\zeta})$ from Λ and a pair (ρ, σ) uniformly on \mathcal{C} . We will use (ρ, σ) to choose the indices of the particles that interact at each jump. Consider also N independent random variables $(X_0^1, \dots, X_0^N) =: \mathbf{X}_0$, independent from \mathcal{N} , each having distribution P_0 . Finally, set $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ to be the complete right continuous filtration generated by \mathbf{X}_0 and \mathcal{N} . We denote \mathbb{P} and \mathbb{E} the probability and expectation in the corresponding probability space.

The particle system $\mathbf{X} = (X^1, \dots, X^N)$ is defined as the solution, starting from \mathbf{X}_0 , of the following integral equation:

$$d\mathbf{X}_t = \int_{\mathbb{R}^4} \int_{[0, N]^2} \sum_{i, j=1}^N \mathbf{1}_{\{\mathbf{i}(\rho)=i, \mathbf{i}(\sigma)=j\}} a_{ij}(\eta, X_{t-}^i, X_{t-}^j) \mathcal{N}(dt, d\eta, d\rho, d\sigma). \quad (3.10)$$

(Recall that $a_{ij}(\eta, u, v)$ is the vector of \mathbb{R}^N whose i -th and j -th components are $(\xi - 1)u + \zeta v$ and $(\tilde{\xi} - 1)v + \tilde{\zeta}u$, respectively, and is equal to 0 in the other components). Given the timely ordered atoms $(t_n, \eta_n, \rho_n, \sigma_n)_{n \geq 0}$ of \mathcal{N} (that is, $t_n \leq t_{n+1}$ for all $n \geq 0$), a solution of this equation can be constructed as follows: recursively define \mathbf{X}_{t_n} as

$$X_{t_n}^\ell = \begin{cases} \xi_n X_{t_{n-1}}^i + \zeta_n X_{t_{n-1}}^j, & \ell = i, \\ \tilde{\xi}_n X_{t_{n-1}}^j + \tilde{\zeta}_n X_{t_{n-1}}^i, & \ell = j, \\ X_{t_{n-1}}^\ell, & \ell \neq i, j, \end{cases} \quad (3.11)$$

where $(i, j) = (\mathbf{i}(\rho_n), \mathbf{i}(\sigma_n))$, and set $\mathbf{X}_t = \mathbf{X}_{t_n}$ for all $t \in (t_n, t_{n+1})$. Uniqueness for (3.10) also holds, since there is no choice to make in this construction. It is straightforward to verify that \mathbf{X} has generator (3.6).

Thus, the system \mathbf{X} is what we want it to be: at rate $N/2$ we choose two distinct indices $i = \mathbf{i}(\rho)$ and $j = \mathbf{i}(\sigma)$, and then we update the particles X^i and X^j according to the rule described in (3.3). The fact that we use continuous variables (ρ, σ) to choose the indices (i, j) (instead of a discrete pair chosen uniformly from the set $\{1, \dots, N\}^2 \setminus \{i = j\}$) will be crucial to define our system \mathbf{U} of N nonlinear processes.

3.2.2 Coupling with the nonlinear processes

From (3.10) it follows that for any $i = 1, \dots, N$, the process X^i satisfies

$$dX_t^i = \int_{\mathbb{R}^2} \int_{[0, N)} [(\xi - 1)X_{t^-}^i + \zeta X_{t^-}^{i(\tau)}] \mathcal{N}^i(dt, d\xi, d\zeta, d\tau), \quad (3.12)$$

where \mathcal{N}^i is defined as

$$\begin{aligned} \mathcal{N}^i(dt, d\xi, d\zeta, d\tau) &= \mathcal{N}(dt, (d\xi \times d\zeta \times \mathbb{R}^2), [i - 1, i), d\tau) \\ &\quad + \mathcal{N}(dt, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [i - 1, i)). \end{aligned} \quad (3.13)$$

Clearly, \mathcal{N}^i is a Poisson point measure on $[0, \infty) \times \mathbb{R}^2 \times [0, N)$ with intensity

$$dt \bar{\Lambda}(d\xi, d\zeta) \frac{d\tau}{N-1} \mathbf{1}_{A^i}(\tau),$$

where $\bar{\Lambda} = \frac{1}{2}(\mathcal{L}(L, R) + \mathcal{L}(\tilde{L}, \tilde{R}))$, and $A^i = [0, N) \setminus [i - 1, i)$. In other words, \mathcal{N}^i selects only the atoms of \mathcal{N} that produce a jump of X^i , that is, the atoms in which $\mathbf{i}(\rho) = i$ or $\mathbf{i}(\sigma) = i$, see Figure 3.1.

Let us examine the expression (3.12) in more detail. First note that since τ is chosen uniformly in A^i , the variable $X_{t^-}^{i(\tau)}$ corresponds to a sample from the (random) probability measure $\bar{\mathbf{X}}_{t^-}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{X_{t^-}^j}$. Thus, from the point of view of the process X^i , the dynamics

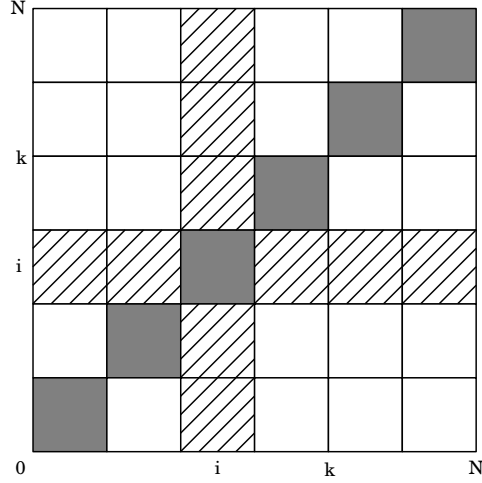


Figure 3.1: For $N = 6$, the set \mathcal{C} is the region $[0, N]^2$ without the diagonal squares. For $i = 3$, the Poisson point measure \mathcal{N}^i selects the atoms (t, η, ρ, σ) of \mathcal{N} in which (ρ, σ) falls on the hatched region.

is as follows: at rate 1, a number $v = X_{t-}^{i(\tau)}$ is sampled from the measure $\bar{\mathbf{X}}_{t-}^i$, and then the value of the process is updated according to the rule $X_{t-}^i \mapsto \xi X_{t-}^i + \zeta v$, where (ξ, ζ) is chosen with law $\bar{\Lambda}$.

Comparing (3.9) and (3.12), the key observation is the following: if for each jump time t one replaces $X_{t-}^{i(\tau)}$ in (3.12) with a realization v of the law $P_t(dv)$, the resulting process has law P . In view of this, we would like to define the system of nonlinear processes $\mathbf{U} = (U^1, \dots, U^N)$ based on this idea, but using a realization of P_t that is optimally coupled to the realization $X_{t-}^{i(\tau)}$ of the measure $\bar{\mathbf{X}}_{t-}^i$. In doing this, some measurability issues need to be taken into account.

Lemma 3.3 (coupling). *For every $p \geq 1$ and $i \in \{1, \dots, N\}$ there exist a measurable mapping $\Pi^i : \mathbb{R}_+ \times \mathbb{R}^N \times A^i \rightarrow \mathbb{R}$, $(t, \mathbf{x}, \tau) \mapsto \Pi_t^i(\mathbf{x}, \tau)$, with the following property: for every $t \geq 0$ and $\mathbf{x} \in \mathbb{R}^N$, if τ is uniformly chosen from A^i , then the pair $(\Pi_t^i(\mathbf{x}, \tau), x^{i(\tau)})$ is an optimal coupling between P_t and $\bar{\mathbf{x}}^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{x_j}$ with respect to the cost function $c(u, v) = |u - v|^p$. Moreover, if \mathbf{Y} is any exchangeable random vector in \mathbb{R}^N , then $\mathbb{E} \int_{j=1}^j \phi(\Pi_t^i(\mathbf{Y}, \tau)) d\tau = \langle P_t, \phi \rangle$ for any $j \in \{1, \dots, N\}$, $j \neq i$, and any bounded measurable function ϕ .*

For simplicity, in our notation we have not made explicit the dependence of Π_t^i on p (however, see Remark 3.10). Now, we can define U^i as the solution of

$$dU_t^i = \int_{\mathbb{R}^2} \int_{[0, N]} [(\xi - 1)U_{t-}^i + \zeta \Pi_t^i(\mathbf{X}_{t-}, \tau)] \mathcal{N}^i(dt, d\xi, d\zeta, d\tau), \quad (3.14)$$

where \mathcal{N}^i is the same Poisson point measure as in (3.12). The proof of Lemma 3.3 will imply that the mapping $((t, \omega), \xi, \zeta, \tau) \mapsto (\xi - 1)U_{t-}^i(\omega) + \zeta \Pi_t^i(\mathbf{X}_{t-}(\omega), \tau)$ above is measurable with respect to the product of the predictable sigma field (in (t, ω)) and the Borel sigma field of

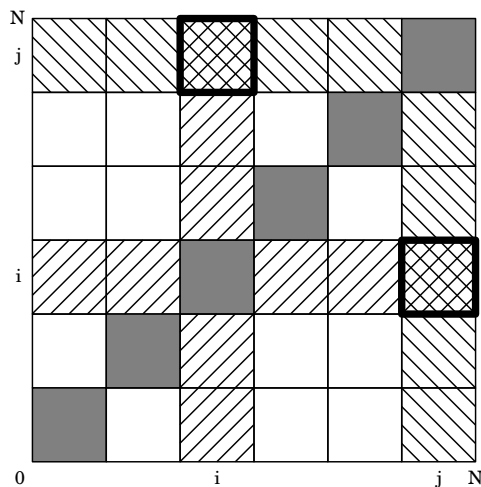


Figure 3.2: For $N = 6$, $i = 3$ and $j = 6$, the atoms of \mathcal{N} such that (ρ, σ) falls on the highlighted region will appear in both \mathcal{N}^i and \mathcal{N}^j .

$\mathbb{R}^2 \times [0, N)$. This ensures that the integral in (3.14) has the usual properties of integrals with respect to Poisson point processes.

We summarize our construction in the following

Lemma 3.4. *Let $p \geq 1$ be fixed. For each $i = 1, \dots, N$ there is a unique solution U^i of (3.14), and it is a nonlinear process. Moreover, the collection $(X^1, U^1), \dots, (X^N, U^N)$ is exchangeable.*

Thus, the system $\mathbf{U} = (U^1, \dots, U^N)$ is indeed a tuple of N nonlinear processes. However, as we already mentioned, they are *not independent*, since \mathcal{N}^i and \mathcal{N}^j share a portion of \mathcal{N} , namely, the atoms of \mathcal{N} whose coordinates (ρ, σ) lie in $[i-1, i) \times [j-1, j)$ or $[j-1, j) \times [i-1, i)$, see Figure 3.2. In particular, whenever such an atom occurs the processes U^i and U^j jump simultaneously, using a single realization of $(L, R, \tilde{L}, \tilde{R})$, and samples of P_t that also are correlated.

3.3 Proof of the main result

Before proving our results, let us first state two lemmas that constitute our basic tools; they will be proven in Section 3.4. The first one provides uniform bounds for the moments of P_t ; it can be seen as a version of Theorem 3.2 in [45]:

Lemma 3.5 (moment bounds). *For $p = 1$ or $p = 2$, assume $\alpha_p \geq 0$ and $M_p(P_0) < \infty$. If $p = 2$, assume also that $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$. Then for any $q \in \{p\} \cup (p, q^*)$ there exists a*

constant C , depending on q and some moments of P_0 and $(L, R, \tilde{L}, \tilde{R})$ of order at most q , such that

$$M_q(P_t) \leq C e^{-\bar{\alpha}_p q t} \quad \forall t \geq 0.$$

The second lemma is fundamental in our developments since it decouples the non-independent nonlinear processes uniformly in time, even in the case $\alpha_p = 0$:

Lemma 3.6 (decoupling). *For $p = 1$ or $p = 2$, assume $\alpha_p \geq 0$ and $M_p(P_0) < \infty$. If $p = 2$, assume also that $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$. Then there exists a constant C , depending only on the p -moment of P_0 and $(L, R, \tilde{L}, \tilde{R})$, such that for all $k = 2, \dots, N$ and $t \geq 0$,*

$$\mathcal{W}_p^p(\mathcal{L}^k(\mathbf{U}_t), P_t^{\otimes k}) \leq \frac{C(k-1) \min(1, t) e^{-\alpha_p t}}{N-1}.$$

Proof of Theorem 3.1. Define the constants $\alpha_p^L = \frac{1}{2} \mathbf{E}(|L|^p + |\tilde{L}|^p)$ and $\alpha_p^R = \frac{1}{2} \mathbf{E}(|R|^p + |\tilde{R}|^p)$, so $\alpha_p = 1 - \alpha_p^L - \alpha_p^R$. We first treat the case $p = 1$. Thus, we work with the processes U^i solution of (3.14) using the functions Π_t^i of Lemma 3.3 with $p = 1$. Let us prove (i) first. We estimate the quantity $f_t = \mathbb{E}|X_t^1 - U_t^1|$ which provides an upper bound for $\mathcal{W}_1(\mathcal{L}^1(\mathbf{X}_t), P_t)$. Using (3.12) and (3.14), for all $0 \leq s \leq t$ we have

$$\begin{aligned} & |X_t^1 - U_t^1| - |X_s^1 - U_s^1| \\ &= \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} \left(|\xi(X_{r-}^1 - U_{r-}^1) + \zeta(X_{r-}^{i(\tau)} - \Pi_r^1(\mathbf{X}_{r-}, \tau))| \right. \\ & \quad \left. - |X_{r-}^1 - U_{r-}^1| \right) \mathcal{N}^1(dr, d\xi, d\zeta, d\tau). \end{aligned} \quad (3.15)$$

Recall that the intensity of \mathcal{N}^1 is $(N-1)^{-1} dt \bar{\Lambda}(d\xi, d\zeta) d\tau \mathbf{1}_{A^1}(\tau)$, where $\bar{\Lambda} = (\mathcal{L}(L, R) + \mathcal{L}(\tilde{L}, \tilde{R}))/2$. By the compensation formula, $t \mapsto f_t$ is absolutely continuous and we obtain

$$\begin{aligned} f_t - f_s &\leq \mathbb{E} \int_s^t \int_{\mathbb{R}^2} \int_{A^1} \left((|\xi| - 1) |X_r^1 - U_r^1| \right. \\ & \quad \left. + |\zeta| |X_r^{i(\tau)} - \Pi_r^1(\mathbf{X}_r, \tau)| \right) \frac{d\tau}{N-1} \bar{\Lambda}(d\xi, d\zeta) dr \\ &= \mathbb{E} \int_s^t \left((\alpha_1^L - 1) |X_r^1 - U_r^1| + \alpha_1^R \mathcal{W}_1(\bar{\mathbf{X}}_r^1, P_r) \right) dr, \end{aligned} \quad (3.16)$$

where in the last step we have used the fact that when τ is uniform in A^1 , $(\Pi_s^1(\mathbf{x}, \tau), x^{i(\tau)})$ is an optimal coupling between P_s and $\bar{\mathbf{x}}^1$. We deduce that for almost all $t \geq 0$

$$\partial_t f_t \leq -(1 - \alpha_1^L) f_t + \alpha_1^R \mathbb{E} \mathcal{W}_1(\bar{\mathbf{X}}_t^1, P_t). \quad (3.17)$$

Recall that $\bar{\mathbf{U}}_t^i = \frac{1}{N-1} \sum_{j \neq i} \delta_{U_t^j}$ for $i = 1, \dots, N$. The triangle inequality for \mathcal{W}_1 gives us

$$\begin{aligned} \mathbb{E} \mathcal{W}_1(\bar{\mathbf{X}}_t^1, P_t) &\leq \mathbb{E} \mathcal{W}_1(\bar{\mathbf{X}}_t^1, \bar{\mathbf{U}}_t^1) + \mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t^1, P_t) \\ &\leq \mathbb{E} |X_t^1 - U_t^1| + \mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t^1, P_t), \end{aligned} \quad (3.18)$$

where the last inequality comes from the fact that $(X_t^{i(\tau)}, U_t^{i(\tau)})$ is a coupling between $\bar{\mathbf{X}}_t^1$ and $\bar{\mathbf{U}}_t^1$ when τ is uniformly chosen in A^1 , and from the exchangeability of $(X^i, U^i)_{i=1, \dots, N}$. Putting this together with (3.17), we obtain

$$\partial_t f_t \leq -\alpha_1 f_t + \alpha_1^R \mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t^1, P_t). \quad (3.19)$$

Next, we need an estimate for $\mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t^1, P_t)$. Since the system (U^2, \dots, U^N) is exchangeable, using a recent result (Theorem 1.2 of [40]), we obtain the following: for each $q > 0$ and each $\gamma < (2 + 1/q)^{-1}$, there exists a constant $C_{q, \gamma}$ such that

$$\mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t^1, P_t) \leq C_{q, \gamma} M_q(P_t)^{1/q} \left(\mathcal{W}_1(\mathcal{L}(U_t^2, U_t^3), P_t^{\otimes 2}) + \frac{1}{N-1} \right)^\gamma. \quad (3.20)$$

Now, Lemma 3.6 in the case $p = 1$ and $k = 2$ implies $\mathcal{W}_1(\mathcal{L}(U_t^1, U_t^2), P_t^{\otimes 2}) \leq C/N$, where C is some constant, which can change from line to line in what follows. From this, Lemma 3.5, and (3.19)-(3.20) we have $\partial_t f_t \leq -\alpha_1 f_t + CN^{-\gamma} e^{-\frac{1}{q} \bar{\alpha}_{1, q} t}$, and then Gronwall's lemma yields

$$f_t \leq \frac{C}{N^\gamma} \int_0^t e^{-\alpha_1(t-s)} e^{-\frac{1}{q} \bar{\alpha}_{1, q} s} ds,$$

since $f_0 = 0$. Bounding $e^{-\alpha_1(t-s)} \leq e^{-\frac{1}{q} \bar{\alpha}_{1, q} (t-s)}$ gives (i) in the case $p = 1$ and $k = 1$. From this and Lemma 3.6, the case $k \geq 2$ follows:

$$\begin{aligned} \mathcal{W}_1(\mathcal{L}^k(\mathbf{X}_t), P_t^{\otimes k}) &\leq \mathcal{W}_1(\mathcal{L}^k(\mathbf{X}_t), \mathcal{L}^k(\mathbf{U}_t)) + \mathcal{W}_1(\mathcal{L}^k(\mathbf{U}_t), P_t^{\otimes k}) \\ &\leq \mathbb{E}|X_t^1 - U_t^1| + \frac{Ck \min(1, t) e^{-\alpha_1 t}}{N}. \end{aligned} \quad (3.21)$$

We now prove (ii): as in (3.18) we have

$$\begin{aligned} \mathbb{E} \mathcal{W}_1(\bar{\mathbf{X}}_t, P_t) &\leq \mathbb{E}|X_t^1 - U_t^1| + \mathbb{E} \mathcal{W}_1(\bar{\mathbf{U}}_t, P_t) \\ &\leq \frac{Ct e^{-\frac{1}{q} \bar{\alpha}_{1, q} t}}{N^\gamma} + \frac{C e^{-\frac{1}{q} \bar{\alpha}_{1, q} t}}{N^\gamma}, \end{aligned}$$

where the last inequality comes from (i) in the case $k = 1$, and from (3.20) (with $\bar{\mathbf{U}}_t$ and N in place of $\bar{\mathbf{U}}_t^1$ and $N - 1$) together with Lemma 3.6 in the case $k = 2$. From the previous inequality, (ii) follows; moreover, the same estimate is also valid for $\mathbb{E} \mathcal{W}_1(\bar{\mathbf{X}}_t^1, P_t)$.

Now we treat the case $p = 2$. The proof is similar to the previous case, with adaptations where required. We work with the processes U^i solution of (3.14) using the functions Π_t^i of Lemma 3.3 with $p = 2$. As before, to prove the case $k = 1$ we want to estimate $f_t =$

$\mathbb{E}(X_t^1 - U_t^1)^2$. We proceed as in (3.15): from (3.12) and (3.14), we have for all $0 \leq s \leq t$

$$\begin{aligned}
 & (X_t^1 - U_t^1)^2 - (X_s^1 - U_s^1)^2 \\
 &= \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} \left([\xi(X_{r^-}^1 - U_{r^-}^1) + \zeta(X_{r^-}^{i(\tau)} - \Pi_r^1(\mathbf{X}_{r^-}, \tau))]^2 \right. \\
 &\quad \left. - [X_{r^-}^1 - U_{r^-}^1]^2 \right) \mathcal{N}^1(dr, d\xi, d\zeta, d\tau) \\
 &= \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} \left([\xi^2 - 1](X_{r^-}^1 - U_{r^-}^1)^2 + \zeta^2(X_{r^-}^{i(\tau)} - \Pi_r^1(\mathbf{X}_{r^-}, \tau))^2 \right. \\
 &\quad \left. + 2\xi\zeta(X_{r^-}^1 - U_{r^-}^1)(X_{r^-}^{i(\tau)} - \Pi_r^1(\mathbf{X}_{r^-}, \tau)) \right) \mathcal{N}^1(dr, d\xi, d\zeta, d\tau).
 \end{aligned} \tag{3.22}$$

Taking expectations, the last term in the integral vanishes thanks to condition $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$. As in (3.16)-(3.17), this yields

$$\partial_t f_t \leq -(1 - \alpha_2^L) f_t + \alpha_2^R \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{X}}_t^1, P_t) \tag{3.23}$$

Defining $g_t = \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^1, P_t)$ and using the triangle inequality of \mathcal{W}_2 we have

$$\begin{aligned}
 \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{X}}_t^1, P_t) &\leq \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{X}}_t^1, \bar{\mathbf{U}}_t^1) + 2\mathbb{E} \mathcal{W}_2(\bar{\mathbf{X}}_t^1, \bar{\mathbf{U}}_t^1) \mathcal{W}_2(\bar{\mathbf{U}}_t^1, P_t) + \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^1, P_t) \\
 &\leq f_t + 2f_t^{1/2} g_t^{1/2} + g_t,
 \end{aligned} \tag{3.24}$$

where in the last inequality the term f_t is obtained with the same argument as in (3.18), and the term $f_t^{1/2} g_t^{1/2}$ comes from the Cauchy-Schwarz inequality. From this and (3.23) we obtain

$$\partial_t f_t \leq -\alpha_2 f_t + 2\alpha_2^R f_t^{1/2} g_t^{1/2} + \alpha_2^R g_t.$$

Using a version of Gronwall's lemma (see e.g. Lemma 4.1.8 of [2]) together with Jensen's inequality, we obtain

$$f_t \leq \alpha_2^R e^{-\alpha_2 t} (2 + 8\alpha_2^R t) \int_0^t e^{\alpha_2 s} g_s ds. \tag{3.25}$$

Now, we need an estimate for $g_t = \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^1, P_t)$. Unfortunately, we do not have at our disposal a result similar to (3.20), which is valid only for \mathcal{W}_1 . To bypass this, we will make use of the following lemma (proved in Section 3.4); it has the spirit of (3.20) in the sense that it will allow us to work with $\mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t), P_t^{\otimes n})$ instead of $\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^1, P_t)$, but at the price of the extra term $\varepsilon_{n,2}(P_t)$.

Lemma 3.7. *Let $\mathbf{Y} = (Y^1, \dots, Y^m)$ be an exchangeable random vector, and let μ be a probability measure on \mathbb{R} . Then, for any $p \geq 1$ and $n \leq m$, $n \in \mathbb{N}$, we have*

$$\begin{aligned}
 \frac{1}{2^{p-1}} \mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Y}}, \mu) &\leq \frac{kn}{m} \left(\mathcal{W}_p^p(\mathcal{L}^n(\mathbf{Y}), \mu^{\otimes n}) + \varepsilon_{n,p}(\mu) \right) \\
 &\quad + \frac{\ell}{m} \left(\mathcal{W}_p^p(\mathcal{L}^\ell(\mathbf{Y}), \mu^{\otimes \ell}) + \varepsilon_{\ell,p}(\mu) \right)
 \end{aligned}$$

where k and ℓ are the unique non-negative integers satisfying $m = kn + \ell$, with $\ell \leq n - 1$. Here $\varepsilon_{n,p}(\mu) := \mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Z}}, \mu)$, where $\mathbf{Z} = (Z^1, \dots, Z^n)$ are i.i.d. and μ distributed.

Note that $\mathcal{W}_2^2(\mathcal{L}^\ell(\mathbf{U}_t), P_t^{\otimes \ell}) + \varepsilon_{\ell,2}(P_t) \leq 8M_2(P_t)$. Using this Lemma with $p = 2$, $m = N - 1$, $\mathbf{Y} = (U_t^2, \dots, U_t^{N-1})$ and $\mu = P_t$, we obtain that for every $n \leq N - 1$

$$\begin{aligned} \mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^1, P_t) &\leq \mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t), P_t^{\otimes n}) + \varepsilon_{n,p}(P_t) + \frac{n-1}{N-1}8M_2(P_t) \\ &\leq C \left(\frac{ne^{-\alpha_2 t}}{N} + \varepsilon_{n,2}(P_t) \right), \end{aligned}$$

where in the last inequality we have used Lemma 3.5 and Lemma 3.6 with $p = 2$ and $k = n$; again C is some constant that can change from line to line. Putting this into (3.25) gives

$$f_t \leq C(1+t) \left(\frac{nte^{-\alpha_2 t}}{N} + \int_0^t e^{-\alpha_2(t-s)} \varepsilon_{n,2}(P_s) ds \right).$$

Given $q \in (2, q^*)$, $q \neq 4$, from Theorem 1 of [28] we know that $\varepsilon_{n,2}(P_t) \leq CM_q^{2/q}(P_t)n^{-\eta}$, where $\eta = \min(1/2, \frac{q-2}{q})$. Choosing $n = \lfloor N^{1/(1+\eta)} \rfloor$ and using Lemma 3.5 with $p = 2$ yields

$$f_t \leq C(1+t) \left(\frac{te^{-\alpha_2 t}}{N^\gamma} + \frac{1}{N^\gamma} \int_0^t e^{-\alpha_2(t-s)} e^{-\frac{2}{q}\bar{\alpha}_{2,q}s} ds \right).$$

where $\gamma = \eta/(1+\eta) = \min(1/3, \frac{q-2}{2q-2})$. Bounding $e^{-\alpha_2(t-s)} \leq e^{-\frac{2}{q}\bar{\alpha}_{2,q}(t-s)}$ gives (i) in the case $p = 2$ and $k = 1$. The case $k \geq 2$ follows as in (3.21).

Finally, (ii) in the case $p = 2$ follows from (3.24) with a similar argument as in the case $p = 1$. This concludes the proof. \square

Corollary 3.8. *Under the same hypotheses and notation of Theorem 3.1, we have for all $T \geq 0$,*

$$\mathbb{E} \sup_{t \in [0, T]} |X_t^1 - U_t^1|^p \leq \frac{C}{N^\gamma} \int_0^T (1+t)^p e^{-\frac{p}{q}\bar{\alpha}_{p,q}t} dt.$$

Proof. From (3.15), discarding the negative term in the integral, we have

$$\begin{aligned} \sup_{t \in [0, T]} |X_t^1 - U_t^1| &\leq \int_{(0, T]} \int_{\mathbb{R}^2} \int_{[0, N]} \left(|\xi| |X_{t^-}^1 - U_{t^-}^1| \right. \\ &\quad \left. + |\zeta| |X_{t^-}^{i(\tau)} - \Pi_t^1(\mathbf{X}_{t^-}, \tau)| \right) \mathcal{N}^1(dt, d\xi, d\zeta, d\tau). \end{aligned}$$

With the same argument that produced the term $\mathcal{W}_1(\bar{\mathbf{X}}_r^1, P_r)$ in (3.16), the conclusion follows taking expectations and using the previous estimates for $\mathbb{E}|X_t^1 - U_t^1|$ and $\mathbb{E}\mathcal{W}_1(\bar{\mathbf{X}}_t^1, P_t)$. This proves the case $p = 1$, and the case $p = 2$ follows from (3.22) with a similar argument. \square

Remark 3.9. To illustrate how our methods can indeed be used in non-cutoff contexts, consider Kac's model: $L = \cos \theta = \tilde{L}$ and $R = -\sin \theta = -\tilde{R}$, where θ is chosen according to an even cross-section function $\beta : [-\pi, \pi] \rightarrow \mathbb{R}_+$ that possibly is singular at 0, but satisfies

the classical condition $\int_0^\pi \theta^2 \beta(\theta) d\theta < \infty$, see [20] for details. Define $\beta_\varepsilon(\theta) = \mathbf{1}_{|\theta| > \varepsilon} \beta(\theta)$ for a given cutoff level $\varepsilon > 0$, and associate with it the collection $(P_t^\varepsilon)_{t \geq 0}$ solving $\partial_t P_t^\varepsilon = \kappa_\varepsilon(-P_t^\varepsilon + \mathcal{Q}_\varepsilon^+(P_t^\varepsilon))$, where $\kappa_\varepsilon = \int_{-\pi}^\pi \beta_\varepsilon(\theta) d\theta$ and $\mathcal{Q}_\varepsilon^+$ is defined as

$$\int \phi(u) \mathcal{Q}_\varepsilon^+(\mu)(du) = \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{-\pi}^\pi \phi(u \cos \theta - v \sin \theta) \frac{\beta_\varepsilon(\theta) d\theta}{\kappa_\varepsilon} \mu(dv) \mu(du).$$

The particle system \mathbf{X}^ε and nonlinear processes \mathbf{U}^ε are constructed in a way similar as in (3.12) and (3.14) but now using a Poisson measure $\mathcal{N}^{\varepsilon, i}(dt, d\theta, d\tau)$ with intensity $(N - 1)^{-1} dt \beta_\varepsilon(\theta) d\theta d\tau \mathbf{1}_{A^i}(\tau)$ and functions $\Pi_t^{\varepsilon, i}$ that couple optimally with P_t^ε instead of P_t . Note that:

- The even moments of P_t^ε are controlled uniformly in time and independently of ε (see for instance Lemma A.5 in [27] in the case of the non-cutoff nonlinear process; also, an induction similar to the one used in the proof of Lemma 3.5 yields the desired uniform bounds for $M_q(P_t^\varepsilon)$ when q is even).
- The decoupling property of Lemma 3.6 is also valid for the system \mathbf{U}_t^ε , with constants independent of ε : in (3.32), all the terms involve either $1 - L$ or R^2 , which correspond to $1 - \cos \theta$ and $\sin^2 \theta$, respectively, both of order θ^2 .
- In (3.25) the constant α_2^R corresponds to $\int_{-\pi}^\pi \sin^2 \theta \beta_\varepsilon(\theta)$.

Thus, the argument can be replicated and the final constant will depend on $\int_0^\pi \theta^2 \beta_\varepsilon(\theta) d\theta$, which remains bounded as we let the cutoff $\varepsilon \rightarrow 0$. Assuming for instance that $M_6(P_0) < \infty$, this yields a constant C independent of $\varepsilon > 0$ such that

$$\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{X}}_t^\varepsilon, P_t^\varepsilon) \leq \frac{C(1+t)^2}{N^{1/3}}.$$

However, we have not been able to obtain a trajectorial result in the non-cutoff case: discarding the negative term in the integral of (3.22) produces the term $\int_0^\pi \cos^2 \theta \beta_\varepsilon(\theta) d\theta$ which no longer stays bounded when $\varepsilon \rightarrow 0$.

3.4 Proof of intermediate lemmas

Proof of Lemma 3.3. For fixed $n \in \mathbb{N}$, given $\mathbf{y} = (y^1, \dots, y^n) \in \mathbb{R}^n$ recall that we write $\bar{\mathbf{y}} = \frac{1}{n} \sum_j \delta_{y^j}$. The mapping $(t, \mathbf{y}) \mapsto (P_t, \bar{\mathbf{y}})$ from $\mathbb{R}_+ \times \mathbb{R}^n$ to $\mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R})$ is continuous when $\mathcal{P}(\mathbb{R})$ is endowed with the weak topology (weak continuity of $t \mapsto P_t$ is clear from the pathwise properties of the nonlinear process). Thus, thanks to a measurable selection result

(see for instance Corollary 5.22 of [62]), there exists a measurable mapping $(t, \mathbf{y}) \mapsto \pi_{t, \bar{\mathbf{y}}}$ such that $\pi_{t, \bar{\mathbf{y}}} \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ is an optimal transference plan between P_t and $\bar{\mathbf{y}}$. We now define

$$G(t, \mathbf{y}, B) = \frac{\pi_{t, \bar{\mathbf{y}}}(B \times \{y^1\})}{\pi_{t, \bar{\mathbf{y}}}(\mathbb{R} \times \{y^1\})} = \pi_{t, \bar{\mathbf{y}}}(B \times \{y^1\} \mid \mathbb{R} \times \{y^1\}),$$

for $t \geq 0$, $\mathbf{y} \in \mathbb{R}^n$ and any Borel set $B \subseteq \mathbb{R}$. We claim that G is a probability kernel from $\mathbb{R}_+ \times \mathbb{R}^n$ into \mathbb{R} . Indeed, it suffices to show that for every such B the mapping $(t, \mathbf{y}) \mapsto \pi_{t, \bar{\mathbf{y}}}(B \times \{y^1\})$ is measurable, which in turn follows from the measurability of $(t, \mathbf{y}) \mapsto (\pi_{t, \bar{\mathbf{y}}}, \mathbf{y})$ and the identity

$$\pi_{t, \bar{\mathbf{y}}}(B \times \{y^1\}) = \lim_{\varepsilon \rightarrow 0} \sum_{\ell \in \mathbb{N}} \pi_{t, \bar{\mathbf{y}}}(B \times D_\ell^\varepsilon) \mathbf{1}_{D_\ell^\varepsilon}(y^1),$$

where $(D_\ell^\varepsilon)_{\ell \in \mathbb{N}}$ is a measurable partition of \mathbb{R} with $\text{diam}(D_\ell^\varepsilon) \leq \varepsilon$.

Now, given $N \geq 1$, with the kernel G defined above for $n = N - 1$ we can associate a measurable mapping $g : \mathbb{R}_+ \times \mathbb{R}^{N-1} \times [0, 1] \rightarrow \mathbb{R}$ or randomization of G such that $g(t, \mathbf{y}, \theta)$ has distribution $G(t, \mathbf{y}, \cdot)$ whenever θ is a uniform random variable in $[0, 1]$ (see for instance Lemma 3.22 of [43]). For $\mathbf{x} \in \mathbb{R}^N$ we now put

$$\Pi_t^i(\mathbf{x}, \tau) = \sum_{j \neq i}^N \mathbf{1}_{\{\mathbf{i}(\tau)=j\}} g(t, \mathbf{x}^{(ij)}, \tau - \lfloor \tau \rfloor), \quad \tau \in A^i, \quad (3.26)$$

where $\mathbf{x}^{(ij)} \in \mathbb{R}^{N-1}$ denotes the vector \mathbf{x} with its i coordinate removed, the j coordinate in the first position, and the remaining coordinates in positions $2, \dots, N-1$ in increasing order. We now show that when τ is uniform in A^i , $\Pi_t^i(\mathbf{x}, \tau)$ and $x^{\mathbf{i}(\tau)}$ have joint distribution $\pi_{t, \bar{\mathbf{x}}^i}$. Denoting \mathbf{P}^i the law of this random variable τ and using the fact that $g(t, \mathbf{x}^{(ij)}, \theta)$ has law $\pi_{t, \bar{\mathbf{x}}^i}(du \times \{x^j\} \mid \mathbb{R} \times \{x^j\})$ when θ is uniform in $[0, 1]$, we have for every fixed measurable set $B \subseteq \mathbb{R}$ and every $j \neq i$:

$$\begin{aligned} \mathbf{P}^i(\Pi_t^i(\mathbf{x}, \tau) \in B, x^{\mathbf{i}(\tau)} = x^j) &= \sum_{\ell: x^\ell = x^j, \ell \neq i} \int_{\ell-1}^{\ell} \mathbf{1}_B(g(t, \mathbf{x}^{(i\ell)}, \tau - \lfloor \tau \rfloor)) \frac{d\tau}{N-1} \\ &= \frac{1}{N-1} \sum_{\ell: x^\ell = x^j, \ell \neq i} \frac{\pi_{t, \bar{\mathbf{x}}^i}(B \times \{x^\ell\})}{\pi_{t, \bar{\mathbf{x}}^i}(\mathbb{R} \times \{x^\ell\})} \\ &= \frac{|\{\ell : x^\ell = x^j, \ell \neq i\}|}{(N-1)\pi_{t, \bar{\mathbf{x}}^i}(\mathbb{R} \times \{x^j\})} \pi_{t, \bar{\mathbf{x}}^i}(B \times \{x^j\}), \end{aligned}$$

where the quotient in the last line equals 1. This shows that $(\Pi_t^i(\mathbf{x}, \tau), x^{\mathbf{i}(\tau)})$ has distribution $\pi_{t, \bar{\mathbf{x}}^i}$ and concludes the proof of the existence of Π^i .

It remains to show that $\mathbb{E} \int_{j-1}^j \phi(\Pi_t^i(\mathbf{Y}, \tau)) d\tau = \langle P_t, \phi \rangle$ when \mathbf{Y} is exchangeable, $j \neq i$ and ϕ is bounded and measurable. We get from (3.26) that

$$\begin{aligned} \int_{j-1}^j \phi(\Pi_t^i(\mathbf{Y}, \tau)) d\tau &= \int_0^1 \phi(g(t, \mathbf{Y}^{(ij)}, \tau)) d\tau \\ &= \int_{\mathbb{R}} \phi(u) \pi_{t, \bar{\mathbf{Y}}^i}(du \times \{Y^j\} \mid \mathbb{R} \times \{Y^j\}), \end{aligned}$$

where we have again used that $g(t, \mathbf{Y}^{(ij)}, \theta)$ has distribution $\pi_{t, \bar{\mathbf{Y}}^i}(du \times \{Y^j\} \mid \mathbb{R} \times \{Y^j\})$ when θ is uniform in $[0, 1]$. From the exchangeability of \mathbf{Y} , it is clear that the last expression has the same distribution, for all $j \neq i$. Thus, its expected value must be the same for all $j \neq i$, and since

$$\langle P_t, \phi \rangle = \int_{A^i} \phi(\Pi_t^i(\mathbf{Y}, \tau)) \frac{d\tau}{N-1} = \sum_{j \neq i} \int_{j-1}^j \phi(\Pi_t^i(\mathbf{Y}, \tau)) \frac{d\tau}{N-1},$$

the conclusion follows. \square

Remark 3.10. Since we are working on \mathbb{R} , the increasing coupling between P_t and $\bar{\mathbf{x}}^i$ is in fact an optimal coupling (see for instance Theorem 6.0.2 in [2]), which allows for a simpler proof of Lemma 3.3. However, we opted to give a proof that remains valid on \mathbb{R}^d with the hope that this coupling can be used in a more general setting.

Proof of Lemma 3.4. Existence and uniqueness for (3.14) are obtained with a construction similar to (3.11). To show that U^i is a nonlinear process, define $\tilde{\mathcal{N}}^i(dt, d\xi, d\zeta, dv)$ to be the point measure on $\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}$ with atoms $(t, \xi, \zeta, \Pi_t^i(\mathbf{X}_{t-}, \tau))$ for every atom (t, ξ, ζ, τ) of \mathcal{N}^i ; since the dependence on \mathbf{X} is predictable, one can use the compensation formula to compute the Laplace functional of $\tilde{\mathcal{N}}^i$ and conclude that $\tilde{\mathcal{N}}^i$ is a Poisson point measure with intensity $dt \bar{\Lambda}(d\xi, d\zeta) P_t(dv)$ (see the end of the proof of Lemma 3.6 for a similar argument in a more general setting). Then, (3.9) is satisfied for $V = U^i$ with $\mathcal{M} = \tilde{\mathcal{N}}^i$, implying that $\mathcal{L}(U^i) = P$.

We now prove that the collection $(\mathbf{X}, \mathbf{U}) = ((X^1, U^1), \dots, (X^N, U^N))$ is exchangeable. Regarding (\mathbf{X}, \mathbf{U}) as an $(\mathbb{R} \times \mathbb{R})^N$ -valued process, from (3.10) and (3.14) we obtain the following equation:

$$(d\mathbf{X}_t, d\mathbf{U}_t) = \int_{\mathbb{R}^4} \sum_{i \neq j} \int_{[i-1, i)} \int_{[j-1, j)} b_{ij}(t, \eta, \rho, \sigma, \mathbf{X}_{t-}, \mathbf{U}_{t-}) \mathcal{N}(dt, d\eta, d\rho, d\sigma), \quad (3.27)$$

where, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, $b_{ij}(t, \eta, \rho, \sigma, \mathbf{x}, \mathbf{y})$ is the $(\mathbb{R} \times \mathbb{R})^N$ -vector given by

$$b_{ij}(t, \eta, \rho, \sigma, \mathbf{x}, \mathbf{y})^\ell = \begin{cases} \left(((\xi - 1)x^i + \zeta x^j, (\xi - 1)y^i + \zeta \Pi_t^i(\mathbf{x}, \sigma)) \right) & \text{if } \ell = i, \\ \left(((\tilde{\xi} - 1)x^j + \tilde{\zeta} x^i, (\tilde{\xi} - 1)y^j + \tilde{\zeta} \Pi_t^j(\mathbf{x}, \rho)) \right) & \text{if } \ell = j, \\ (0, 0) & \text{otherwise.} \end{cases}$$

Now, consider a permutation α of the indices $\{1, \dots, N\}$, and denote $\alpha(\mathbf{x})$ the vector $\mathbf{x} = (x^1, \dots, x^N)$ with its coordinates permuted according to it, that is, $\alpha(\mathbf{x})^\ell = x^{\alpha(\ell)}$ for all $1 \leq \ell \leq N$. From the definition of Π_t^i given by (3.26), it follows that for any $i \neq j$ and any $\theta \in [0, 1)$,

$$\Pi_t^i(\mathbf{x}, j + \theta) = \Pi_t^{\alpha^{-1}(i)}(\alpha(\mathbf{x}), \alpha^{-1}(j) + \theta).$$

Using this, a routine computation shows that the system $\alpha(\mathbf{X}, \mathbf{U})$ of permuted pairs also satisfies (3.27), with a permutation of the i.i.d. collection $(X_0^1, X_0^1), \dots, (X_0^N, X_0^N)$ as initial condition. We conclude that $\alpha(\mathbf{X}, \mathbf{U})$ has the same law as (\mathbf{X}, \mathbf{U}) . \square

Proof of Lemma 3.5. Call $h_t^q = \int |u|^q P_t(du)$. We first prove the statement for the case $p = 2$. Using (3.1)-(3.2) with $\phi = |\cdot|^2$ yields $\partial_t h_t^2 = -\alpha_2 h_t^2 + \mathbf{E}(LR + \tilde{L}\tilde{R})(h_t^1)^2$, and since $\mathbf{E}(LR + \tilde{L}\tilde{R}) = 0$ this implies $h_t^2 = h_0^2 e^{-\alpha_2 t}$. Assume now that $q \in (2, q^*)$ is an integer. Using (3.1)-(3.2) with $\phi = |\cdot|^q$, we have

$$\begin{aligned} \partial_t h_t^q &= -h_t^q + \frac{1}{2} \iint \mathbf{E}(|Lu + Rv|^q + |\tilde{L}v + \tilde{R}u|^q) P_t(du) P_t(dv) \\ &\leq -\alpha_q h_t^q + \frac{1}{2} \sum_{i=1}^{q-1} \binom{q}{i} h_t^i h_t^{q-i} \mathbf{E}(|L|^i |R|^{q-i} + |\tilde{L}|^i |\tilde{R}|^{q-i}). \end{aligned} \quad (3.28)$$

Using loose bounds for $\binom{q}{i}$ we obtain

$$h_t^q \leq h_0^q e^{-\alpha_q t} + C \sum_{i=1}^{q-1} \int_0^t e^{-\alpha_q(t-s)} h_s^i h_s^{q-i} ds,$$

where C is a constant that does not depend on t , and may change from line to line. We now apply induction: the case $q = 2$ was already proven, and for $q \in (2, q^*)$ integer, assuming the desired property for all integer in $\{2, \dots, q-1\}$ and using the bound $h_t^1 \leq (h_t^2)^{1/2} \leq C e^{-\alpha_2 t/2}$, we obtain

$$\begin{aligned} h_t^q &\leq h_0^q e^{-\alpha_q t} + C \int_0^t e^{-\alpha_q(t-s)} e^{-\alpha_2 s/2} e^{-\bar{\alpha}_{2,q-1} s} ds \\ &\quad + C \sum_{i=2}^{q-2} \int_0^t e^{-\alpha_q(t-s)} e^{-\bar{\alpha}_{2,i} s} e^{-\bar{\alpha}_{2,q-i} s} ds. \end{aligned}$$

Note that $\alpha_q > 0$, since $2 < q < q^*$, and recall that $\bar{\alpha}_{2,q} := \inf_{2 \leq r \leq q} \alpha_r = \min(\alpha_2, \alpha_q)$. Thus, if $\alpha_2 = 0$ then $\bar{\alpha}_{2,i} = \bar{\alpha}_{2,q-i} = \bar{\alpha}_{2,q} = 0$ and the last inequality yields $h_t^q \leq h_0^q + C \int_0^t e^{-\alpha_q(t-s)} ds \leq C$, as desired. On the other hand, if $\alpha_2 > 0$, we bound $\alpha_2, \alpha_q, \bar{\alpha}_{2,i}$ and $\bar{\alpha}_{2,q-i}$ from below by $\bar{\alpha}_{2,q} > 0$ and obtain $h_t^q \leq h_0^q e^{-\bar{\alpha}_{2,q} t} + C \int_0^t e^{-\bar{\alpha}_{2,q}(t+s/2)} ds \leq C e^{-\bar{\alpha}_{2,q} t}$, which concludes the induction and the proof in the case $p = 2$ and integer $q \in \{2\} \cup (2, q^*)$.

Assume now that $2 < q = m + \varepsilon < q^*$ with $m \in \{2, \dots\}$ and $\varepsilon \in (0, 1)$. Bounding $|x + y|^q \leq (|x| + |y|)^m (|x|^\varepsilon + |y|^\varepsilon)$ in (3.28) and using the binomial theorem as before, we obtain

$$\begin{aligned} \partial_t h_t^q &= -h_t^q + \frac{1}{2} \sum_{i=0}^m \binom{m}{i} \iint \mathbf{E} \left((|Lu|^\varepsilon + |Rv|^\varepsilon) |Lu|^i |Rv|^{m-i} \right. \\ &\quad \left. + (|\tilde{L}v|^\varepsilon + |\tilde{R}u|^\varepsilon) |\tilde{L}v|^i |\tilde{R}u|^{m-i} \right) P_t(du) P_t(dv) \\ &= -\alpha_q h_t^q + \sum_{i=0}^{m-1} \binom{m}{i} \mathbf{E} \left(|L|^{i+\varepsilon} |R|^{m-i} + |L|^{m-i} |R|^{i+\varepsilon} \right. \\ &\quad \left. + |\tilde{L}|^{i+\varepsilon} |\tilde{R}|^{m-i} + |\tilde{L}|^{m-i} |\tilde{R}|^{i+\varepsilon} \right) h_t^{i+\varepsilon} h_t^{m-i}, \end{aligned}$$

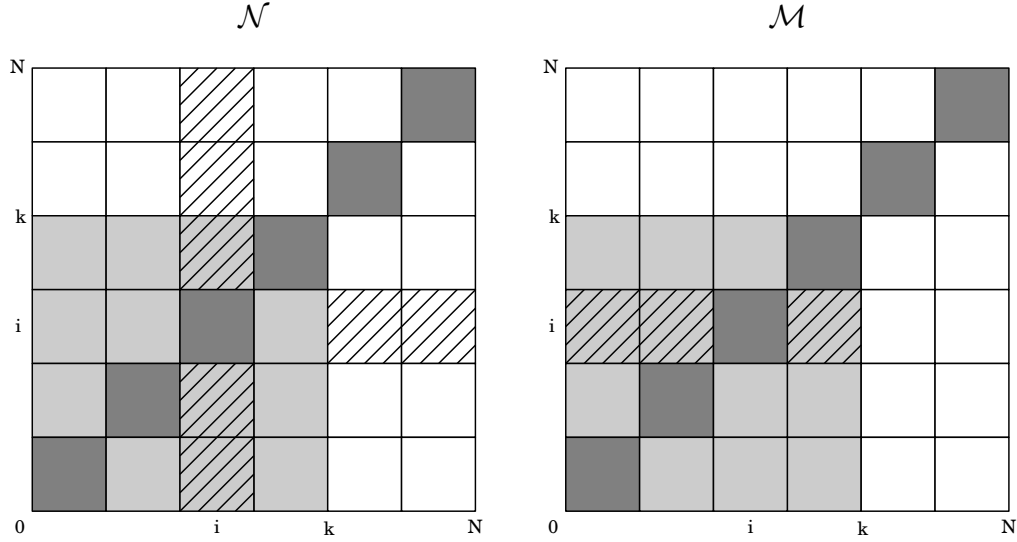


Figure 3.3: For $N = 6$, $k = 4$ and $i = 3$, the Poisson point measure \mathcal{M}^i selects the atoms of \mathcal{N} (left) and \mathcal{M} (right) in which (ρ, σ) falls on the hatched region. This ensures that no such atom appears in two \mathcal{M}^j 's.

which yields

$$h_t^q \leq h_0^q e^{-\alpha_q t} + C \sum_{i=0}^{m-1} \int_0^t e^{-\alpha_q(t-s)} h_s^{i+\varepsilon} h_s^{m-i} ds.$$

Note that $h_t^r \leq (h_t^2)^{r/2} \leq C e^{-r\alpha_2 t/2}$ for $r \in (0, 2)$. This and the fact that the property is true for the integers, allow us to use induction on m in a way similar as before, and conclude the proof in the case $p = 2$.

A similar argument, with the induction starting at $q = 1$, proves the case $p = 1$. \square

Proof of Lemma 3.6. Let us first prove the case $p = 1$. Given $k \in \{2, \dots, N\}$ fixed, we want to construct k independent nonlinear processes V^1, \dots, V^k such that $\mathbb{E}|U_t^i - V_t^i|$ is small. To achieve this, we will decouple U^1, \dots, U^k by replacing the shared atoms of $\mathcal{N}^1, \dots, \mathcal{N}^k$ with new, independent atoms. To this end, let \mathcal{M} be an independent copy of \mathcal{N} (also independent from \mathbf{X}_0), and define for each $i \in \{1, \dots, k\}$

$$\begin{aligned} \mathcal{M}^i(dt, d\xi, d\zeta, d\tau) &= \mathcal{N}(dt, (d\xi \times d\zeta \times \mathbb{R}^2), [i-1, i], d\tau) \\ &\quad + \mathcal{N}(dt, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [i-1, i]) \mathbf{1}_{[k, N)}(\tau) \\ &\quad + \mathcal{M}(dt, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [i-1, i]) \mathbf{1}_{[0, k)}(\tau). \end{aligned} \quad (3.29)$$

Note that \mathcal{M}^i is, like \mathcal{N}^i , a Poisson point measure on $\mathbb{R}_+ \times \mathbb{R}^2 \times [0, N)$ with intensity $(N-1)^{-1} dt \bar{\Lambda}(\xi, \zeta) d\tau \mathbf{1}_{A^i}(\tau)$, and that $\mathcal{M}^1, \dots, \mathcal{M}^k$ are independent, see Figure 3.3. Following (3.14), we define V^i as the solution of

$$dV_t^i = \int_{\mathbb{R}^2} \int_{[0, N)} [(\xi - 1)V_{t-}^i + \zeta \Pi_t^i(\mathbf{X}_{t-}, \tau)] \mathcal{M}^i(dt, d\xi, d\zeta, d\tau), \quad (3.30)$$

with $V_0^i = U_0^i$. Clearly, V^1, \dots, V^k are nonlinear processes, and we need to show that they are independent. This does not follow directly from the independence of the \mathcal{M}^i 's since there is a dependence on \mathbf{X} in (3.30). Thus, we will need some further analysis to conclude this independence, which we postpone until the end of this proof.

Assuming that V^1, \dots, V^k are independent, we have

$$\mathcal{W}_1(\mathcal{L}^k(\mathbf{U}_t), P_t^{\otimes k}) \leq \mathbb{E} \left(\frac{1}{k} \sum_{i=1}^k |U_t^i - V_t^i| \right) = \mathbb{E}|U_t^1 - V_t^1|,$$

where in the last step we used the fact that all the (U^i, V^i) 's have the same law. To estimate the last term $h_t = \mathbb{E}|U_t^1 - V_t^1|$, we proceed as in (3.15): from (3.13), (3.14), (3.29) and (3.30), we have for all $0 \leq s \leq t$:

$$h_t = h_s + \mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} (J_r^1 + J_r^2 + J_r^3), \quad (3.31)$$

where J_r^1 is the term associated with the simultaneous jumps of U^1 and V^1 , J_r^2 corresponds to the jumps of U^1 alone, and J_r^3 gives the jumps of V^1 alone. Specifically,

$$\begin{aligned} J_r^1 &= (|\xi| - 1) |U_{r-}^1 - V_{r-}^1| \left(\mathcal{N}(dr, (d\xi \times d\zeta \times \mathbb{R}^2), [0, 1], d\tau) \right. \\ &\quad \left. + \mathcal{N}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[k,N)}(\tau) \right), \\ J_r^2 &= \left(|\xi(U_{r-}^1 - V_{r-}^1) + \zeta \Pi_r^1(\mathbf{X}_{r-}, \tau) + (\xi - 1)V_{r-}^1| - |U_{r-}^1 - V_{r-}^1| \right) \\ &\quad \mathcal{N}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[0,k)}(\tau), \\ J_r^3 &= \left(|\xi(U_{r-}^1 - V_{r-}^1) - \zeta \Pi_r^1(\mathbf{X}_{r-}, \tau) - (\xi - 1)U_{r-}^1| - |U_{r-}^1 - V_{r-}^1| \right) \\ &\quad \mathcal{M}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[0,k)}(\tau). \end{aligned}$$

Then:

$$\mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} J_r^1 = \left(\frac{1}{2}(\mathbf{E}|L| - 1) + \frac{1}{2}(\mathbf{E}|\tilde{L}| - 1) \frac{N - k}{N - 1} \right) \int_s^t h_r dr.$$

Using the triangle inequality in the term J_r^2 :

$$\begin{aligned} \mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} J_r^2 &\leq \mathbb{E} \int_s^t \int_1^k \left((\mathbf{E}|\tilde{L}| - 1) |U_r^1 - V_r^1| \right. \\ &\quad \left. + \mathbf{E}|\tilde{R}| |\Pi_r^1(\mathbf{X}_r, \tau)| + \mathbf{E}|\tilde{L} - 1| |V_r^1| \right) \frac{dr d\tau}{2(N - 1)}. \end{aligned}$$

From Lemma 3.3 we know that $\mathbb{E} \int_{i-1}^i |\Pi_r^1(\mathbf{X}_r, \tau)| d\tau = M_1(P_r)$ for all $i = 2, \dots, k$. Using that V_r^1 has law P_r , we obtain

$$\begin{aligned} &\mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} J_r^2 \\ &\leq \frac{k - 1}{2(N - 1)} \left((\mathbf{E}|\tilde{L}| - 1) \int_s^t h_r dr + (\mathbf{E}|\tilde{L} - 1| + \mathbf{E}|\tilde{R}|) \int_s^t M_1(P_r) dr \right). \end{aligned}$$

With a similar argument, the last inequality is also valid with J_r^3 in the left hand side. Putting all this into (3.31), we have

$$\begin{aligned} h_t \leq & h_s - \left(1 - \frac{1}{2}\mathbf{E}(|L| + |\tilde{L}|) + \frac{1}{2}(1 - \mathbf{E}|\tilde{L}|)\frac{k-1}{N-1}\right) \int_s^t h_r dr \\ & + \frac{(\mathbf{E}|\tilde{L} - 1| + \mathbf{E}|\tilde{R}|)(k-1)}{N-1} \int_s^t M_1(P_r) dr. \end{aligned}$$

Recall the constants $\alpha_1^L = \frac{1}{2}\mathbf{E}(|L| + |\tilde{L}|)$, $\alpha_1^R = \frac{1}{2}\mathbf{E}(|R| + |\tilde{R}|)$ and $\alpha_1 = 1 - \alpha_1^L - \alpha_1^R$. Also, put $b = \frac{1}{2}(1 - \mathbf{E}|\tilde{L}|)$, which can be assumed non-negative without loss of generality (if not, exchange the roles of (L, R) and (\tilde{L}, \tilde{R})). From the previous inequality and from Lemma 3.5 in the case $q = 1$, it follows that for almost all $t \geq 0$,

$$\partial_t h_t \leq - \left(\alpha_1 + \alpha_1^R + b\frac{k-1}{N-1}\right) h_t + \frac{C(k-1)e^{-\alpha_1 t}}{N-1},$$

and now Gronwall's lemma gives

$$h_t \leq \frac{C(k-1)e^{-\alpha_1 t}}{(N-1)(\alpha_1^R + b\frac{k-1}{N-1})} \left[1 - e^{-(\alpha_1^R + b\frac{k-1}{N-1})t}\right].$$

Using the inequality $1 - e^{-x} \leq x$, the desired result follows for the case $p = 1$.

In the case $p = 2$, we construct the system V^1, \dots, V^k exactly as before, but using the functions Π_t^i provided by Lemma 3.3 with cost $|x - y|^2$. To obtain the desired inequality for $\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{U}_t), P_t^{\otimes k})$ it suffices to work with $h_t = \mathbb{E}(U_t^1 - V_t^1)^2$. We also have (3.31), where J_r^1 , J_r^2 and J_r^3 now are given by

$$\begin{aligned} J_r^1 &= (\xi^2 - 1)(U_{r-}^1 - V_{r-}^1)^2 \left(\mathcal{N}(dr, (d\xi \times d\zeta \times \mathbb{R}^2), [0, 1], d\tau) \right. \\ &\quad \left. + \mathcal{N}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[k, N]}(\tau) \right), \\ J_r^2 &= \left((\xi(U_{r-}^1 - V_{r-}^1) + \zeta \Pi_r^1(\mathbf{X}_{r-}, \tau) + (\xi - 1)V_{r-}^1)^2 - (U_{r-}^1 - V_{r-}^1)^2 \right) \\ &\quad \mathcal{N}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[0, k]}(\tau), \\ J_r^3 &= \left((\xi(U_{r-}^1 - V_{r-}^1) - \zeta \Pi_r^1(\mathbf{X}_{r-}, \tau) - (\xi - 1)U_{r-}^1)^2 - (U_{r-}^1 - V_{r-}^1)^2 \right) \\ &\quad \mathcal{M}(dr, (\mathbb{R}^2 \times d\xi \times d\zeta), d\tau, [0, 1]) \mathbf{1}_{[0, k]}(\tau). \end{aligned}$$

Using that $\mathbb{E} \int_{i-1}^i \Pi_t^1(\mathbf{X}_t, \tau)^2 d\tau = M_2(P_t)$ for all $i = 2, \dots, k$, we obtain:

$$\begin{aligned} \mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0, N]} J_r^1 &= \left(\frac{1}{2}(\mathbf{E}L^2 - 1) + \frac{1}{2}(\mathbf{E}\tilde{L}^2 - 1)\frac{N-k}{N-1} \right) \int_s^t h_r dr, \\ \mathbb{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0, N]} J_r^2 &= \int_s^t \left((\mathbf{E}\tilde{L}^2 - 1)h_r + \mathbf{E}\tilde{R}^2 M_2(P_r) + \mathbf{E}(\tilde{L} - 1)^2 M_2(P_r) \right) \end{aligned}$$

$$\begin{aligned}
 & + 2\mathbf{E}(\tilde{L}\tilde{R})\mathbf{E}(U_r^1 - V_r^1) \int_1^k \Pi_r^1(\mathbf{X}_r, \tau) \frac{d\tau}{k-1} \\
 & + 2\mathbf{E}(\tilde{L}(\tilde{L}-1))\mathbf{E}(U_r^1 - V_r^1)V_r^1 \\
 & + 2\mathbf{E}((\tilde{L}-1)\tilde{R})\mathbf{E}V_r^1 \int_1^k \Pi_r^1(\mathbf{X}_r, \tau) \frac{d\tau}{k-1} \Big) \frac{(k-1)dr}{2(N-1)}, \\
 \mathbf{E} \int_{(s,t]} \int_{\mathbb{R}^2} \int_{[0,N)} J_r^3 & = \int_s^t \left((\mathbf{E}\tilde{L}^2 - 1)h_r + \mathbf{E}\tilde{R}^2 M_2(P_r) + \mathbf{E}(\tilde{L}-1)^2 M_2(P_r) \right. \\
 & - 2\mathbf{E}(\tilde{L}\tilde{R})\mathbf{E}(U_r^1 - V_r^1) \int_1^k \Pi_r^1(\mathbf{X}_r, \tau) \frac{d\tau}{k-1} \\
 & - 2\mathbf{E}(\tilde{L}(\tilde{L}-1))\mathbf{E}(U_r^1 - V_r^1)U_r^1 \\
 & \left. + 2\mathbf{E}((\tilde{L}-1)\tilde{R})\mathbf{E}U_r^1 \int_1^k \Pi_r^1(\mathbf{X}_r, \tau) \frac{d\tau}{k-1} \right) \frac{(k-1)dr}{2(N-1)}.
 \end{aligned}$$

From this and (3.31), we have for almost all $t \geq 0$

$$\begin{aligned}
 \partial_t h_t & = -h_t \left(\left(1 - \frac{1}{2}\mathbf{E}(L^2 + \tilde{L}^2) \right) + \frac{k-1}{2(N-1)}\mathbf{E}(\tilde{L}-1)^2 \right) \\
 & + M_2(P_t) \frac{k-1}{N-1} (\mathbf{E}(\tilde{L}-1)^2 + \mathbf{E}\tilde{R}^2) \\
 & + \mathbf{E}((\tilde{L}-1)\tilde{R})\mathbf{E}(U_t^1 + V_t^1) \int_1^k \Pi_t^1(\mathbf{X}_t, \tau) \frac{d\tau}{N-1}.
 \end{aligned} \tag{3.32}$$

We also have $\int_{i-1}^i \mathbf{E}(U_t^1 + V_t^1) \Pi_t^1(\mathbf{X}_t, \tau) d\tau \leq 2M_2(P_t)$ for all $i = 2, \dots, k$, thanks to the Cauchy-Schwarz and Jensen's inequalities. Recall the constants $\alpha_2^L = \frac{1}{2}\mathbf{E}(L^2 + \tilde{L}^2)$, $\alpha_2^R = \frac{1}{2}\mathbf{E}(R^2 + \tilde{R}^2)$, $\alpha_2 = 1 - \alpha_2^L - \alpha_2^R$, and put $b = \frac{1}{2}\mathbf{E}(\tilde{L}-1)^2$. Using Lemma 3.5 in the case $p = q = 2$, we thus obtain

$$\partial_t h_t \leq - \left(\alpha_2 + \alpha_2^R + b \frac{k-1}{N-1} \right) h_t + \frac{C(k-1)M_2(P_0)e^{-\alpha_2 t}}{N-1},$$

and the conclusion follows from Gronwall's lemma as before.

It remains to show that V^1, \dots, V^k are independent. For $i = 1, \dots, k$, let $\tilde{\mathcal{M}}^i$ be the point process in $\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}$ with atoms $(t, \xi, \zeta, \Pi_t^i(X_{t-}, \tau))$ for every atom (t, ξ, ζ, τ) of \mathcal{M}^i . From (3.30) it is clear that V^i is defined only in terms of X_0^i and $\tilde{\mathcal{M}}^i$, so it suffices to show that the $\tilde{\mathcal{M}}^i$'s are independent point processes. Let f^1, \dots, f^k be non-negative compactly supported measurable functions of $(\xi, \zeta, v) \in \mathbb{R}^2 \times \mathbb{R}$, and for $t \geq 0$ define

$$Y_t = \sum_{i=1}^k \langle \tilde{\mathcal{M}}^i, f^i \mathbf{1}_{[0,t]}(s) \rangle = \sum_{i=1}^k \int_0^t \int_{\mathbb{R}^2} \int_{\mathbb{R}} f^i(\xi, \zeta, v) d\tilde{\mathcal{M}}^i(ds, d\xi, d\zeta, dv).$$

Since the Poisson point measures \mathcal{M}^i are independent, almost surely they never jump simultaneously. Thus, almost surely we have

$$e^{-Y_t} = 1 + \sum_{i=1}^k \int_{(0,t]} \int_{\mathbb{R}^2} \int_{[0,N)} e^{-Y_{s-}} [e^{-f^i(\xi, \zeta, \Pi_s^i(X_{s-}, \tau))} - 1] \mathcal{M}^i(ds, d\xi, d\zeta, d\tau).$$

The last integrand is predictable as a function of (s, ω) and of the atoms of \mathcal{M}^i . Recall that the intensity of \mathcal{M}^i is $(N-1)^{-1} ds \bar{\Lambda}(d\xi, d\zeta) d\tau \mathbf{1}_{A^i}(\tau)$ and that, for any $x \in \mathbb{R}^N$, $\Pi_s^i(x, \tau)$ has law P_s when τ is uniformly distributed in A^i . Taking expectations in the last equation, we then get

$$\mathbb{E}e^{-Y_t} = 1 + \mathbb{E} \int_0^t ds \mathbb{E}e^{-Y_s} \sum_{i=1}^k \int_{\mathbb{R}} \int_{\mathbb{R}^2} [e^{-f^i(\xi, \zeta, v)} - 1] \bar{\Lambda}(d\xi, d\zeta) P_s(dv).$$

Thus, $y_t = \mathbb{E}e^{-Y_t}$ satisfies $\partial_t y_t = y_t R_t$ with $y_0 = 1$, where R_s is the summation in the last expression. This means that $y_t = \exp(\int_0^t R_s ds)$. A monotone class argument yields

$$\mathbb{E} \exp \left(- \sum_{i=1}^k \langle \tilde{\mathcal{M}}^i, h^i \rangle \right) = \prod_{i=1}^k \exp \left(- \int_{\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}} [1 - e^{-h^i}] \bar{\Lambda}(d\xi, d\zeta) P_s(dv) ds \right)$$

for any non-negative measurable functions h^1, \dots, h^k on $\mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}$. This shows that $\tilde{\mathcal{M}}^1, \dots, \tilde{\mathcal{M}}^k$ are independent Poisson point measures with intensity $ds \bar{\Lambda}(d\xi, d\zeta) P_s(dv)$ and completes the proof. \square

Proof of Lemma 3.7. For simplicity, we will prove only the case $\ell = 0$, that is, when n divides m . Let us arrange a vector $\mathbf{y} \in \mathbb{R}^m$ as a matrix with k rows and n columns, that is, $\mathbf{y} = (y^{ij})$, with $i = 1, \dots, k$, $j = 1, \dots, n$, and write $\mathbf{y}_i = (y^{i1}, \dots, y^{in})$ and $\bar{\mathbf{y}}_i = \frac{1}{n} \sum_{j=1}^n \delta_{y^{ij}}$. Let us couple \mathbf{Y} with a random vector $\mathbf{Z} \in (\mathbb{R}^n)^k$ in such a way that each $(\mathbf{Y}_i, \mathbf{Z}_i)$ is an optimal coupling between $\mathcal{L}^n(\mathbf{Y}_i)$ and $\mu^{\otimes n}$ (with respect to the cost function $d_{n,p}^p(\cdot, \cdot)$ of (3.8), as usual). Using the latter, we have:

$$\mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Y}}, \bar{\mathbf{Z}}) \leq \frac{1}{k} \sum_{i=1}^k \mathbb{E} \frac{1}{n} \sum_{j=1}^n |Y^{ij} - Z^{ij}|^p = \mathcal{W}_p^p(\nu^n, \mu^{\otimes n}). \quad (3.33)$$

On the other hand, for each $i = 1, \dots, k$ there is a function $q^i : \mathbb{R}^m \times [0, 1] \rightarrow \mathbb{R}$ such that for all $\mathbf{z} \in \mathbb{R}^m$, the pair $(z^{i n(\theta)}, q^i(\mathbf{z}, \theta))$ with $\mathbf{i}^n(\theta) = \mathbf{i}(n\theta) = \lfloor n\theta \rfloor + 1$, is an optimal coupling between $\bar{\mathbf{z}}_i$ and μ when θ is uniformly chosen in $[0, 1]$. Now we randomize the choice of i with a uniform variable $\vartheta \in [0, 1]$ independent of θ , so $z^{\mathbf{i}^k(\vartheta) n(\theta)}$ and $q^{\mathbf{i}^k(\vartheta)}(\mathbf{z}, \theta)$ are (θ, ϑ) -realizations of $\bar{\mathbf{z}}$ and μ , respectively. Putting \mathbf{Z} in place of \mathbf{z} , this construction gives:

$$\begin{aligned} \mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Z}}, \mu) &\leq \mathbb{E} \iint_{[0,1]^2} |Z^{\mathbf{i}^k(\vartheta) n(\theta)} - q^{\mathbf{i}^k(\vartheta)}(\mathbf{Z}, \theta)|^p d\vartheta d\theta \\ &= \mathbb{E} \frac{1}{k} \sum_{i=1}^k \mathcal{W}_p^p(\bar{\mathbf{Z}}_i, \mu) = \varepsilon_{n,p}(\mu). \end{aligned}$$

(Recall that $\varepsilon_{n,p}(\mu) = \mathbb{E} \mathcal{W}_p^p(\frac{1}{n} \sum_i \delta_{\zeta^i}, \mu)$, with ζ^1, \dots, ζ^n independent and μ -distributed). With this and (3.33) we conclude in the case $\ell = 0$:

$$\begin{aligned} \mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Y}}, \mu) &\leq \mathbb{E} \left(\mathcal{W}_p(\bar{\mathbf{Y}}, \bar{\mathbf{Z}}) + \mathcal{W}_p(\bar{\mathbf{Z}}, \mu) \right)^p \\ &\leq 2^{p-1} \left(\mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Y}}, \bar{\mathbf{Z}}) + \mathbb{E} \mathcal{W}_p^p(\bar{\mathbf{Z}}, \mu) \right). \end{aligned}$$

In the case $\ell > 0$ the construction is similar, but now (\mathbf{Y}, \mathbf{Z}) must include an additional optimal coupling between $\mathcal{L}^\ell(\mathbf{Y})$ and $\mu^{\otimes \ell}$, which gives the extra term. \square

Chapter 4

Quantitative uniform propagation of chaos for Maxwell molecules

4.1 Introduction and main results

4.1.1 The Boltzmann equation

The *spatially homogeneous Boltzmann equation* predicts that the density $f_t(v)$ of particles with velocity $v \in \mathbb{R}^3$ at time $t \geq 0$ in a spatially homogeneous dilute gas subjected to binary collisions, satisfies

$$\partial_t f_t(v) = \frac{1}{2} \int_{\mathbb{R}^3} dv_* \int_{\mathbb{S}^2} d\sigma [f_t(v')f_t(v'_*) - f_t(v)f_t(v_*)]B(|v - v_*|, \theta), \quad (4.1)$$

where v' and v'_* are the *pre-collisional velocities*, given by

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2}\sigma, \quad (4.2)$$

and θ is the *deviation angle*, defined by $\cos \theta = \sigma \cdot (v - v_*)/|v - v_*|$. The *collision kernel* $B(|v - v_*|, \theta) \geq 0$ describes the rate at which collisions between pairs of particles occur, and depends on the type of physical interactions among them. Solutions of (4.1) preserve mass $\int_{\mathbb{R}^3} f_t(v)dv$, momentum $\int_{\mathbb{R}^3} v f_t(v)dv$ and kinetic energy $\int_{\mathbb{R}^3} |v|^2 f_t(v)dv$, so we may assume that $\int_{\mathbb{R}^3} f_t(v)dv = 1$ for all $t \geq 0$.

Equation (4.1) has been extensively studied for several decades. We refer the reader to Cercignani [14] for physical background on the Boltzmann equation and to Villani [61], Alexandre [1] and Mischler and Mouhot [47] for historical accounts on aspects of its mathematical theory.

Typically, one assumes that $B : \mathbb{R}_+ \times (0, \pi] \rightarrow \mathbb{R}_+$ satisfies

$$B(z, \theta) \sin \theta = z^\gamma \beta(\theta), \quad (4.3)$$

for some $\gamma \in [0, 1]$, and $\beta : (0, \pi] \rightarrow \mathbb{R}_+$ some function which, for symmetry reasons, can be assumed to be equal to 0 on $(\pi/2, \pi]$. The following three physical situations are usually considered:

- $\gamma = 1$ and $\beta \equiv 1$, known as *hard spheres* case,
- $\gamma \in (0, 1)$ and $\beta(\theta) \sim \theta^{-1-\nu}$ near 0 for $\nu = \frac{1-\gamma}{2} \in (0, 1/2)$, referred to as *hard potentials* case, or
- $\gamma = 0$ and $\beta(\theta) \sim \theta^{-1-\nu}$ with $\nu = 1/2$, called *Maxwell molecules* case.

Although the techniques of the present paper can probably be applied to all of these three cases, to simplify the exposition we decided to restrict ourselves to the Maxwell molecules case, which also yields better estimates.

4.1.2 Particle system and propagation of chaos

As a step to rigorously justify the interpretation of the Boltzmann equation as a representation of the evolution of a very large number of interacting particles, Kac [42] suggested to study the limit as N goes to ∞ , of some exchangeable stochastic system of N of such interacting particles in the sphere of radius \sqrt{N} in $(\mathbb{R}^3)^N$, defined as a continuous-time pure-jump Markov process. For a simplified one dimensional model of the nonlinear equation (4.1), he in fact proved that if the joint law of k particles at time 0 weakly converge as $N \rightarrow \infty$ to the k -fold product of some density f_0 , then the same holds true at times $t > 0$ with f_t as the limit density.

This property, termed *propagation of chaos* is equivalent to the *convergence of the empirical measure of the system at each time t , to the solution of the nonlinear equation* and has been established, under different convergence criteria, for a wide class of models including the true Boltzmann equation (4.1). For general background on propagation of chaos we refer the reader to Sznitman [57] and Mischler and Mouhot [47] (see also Section 4.1.4 below for historical and recent references).

In the Maxwell molecules case considered here, the function $\beta(\theta)$ can have a non-integrable singularity at $\theta = 0$, and thus particles may jump infinitely many times on every finite time interval (all but finitely many jumps corresponding to very small deviation angles). For both technical and numerical simulation purposes, it is thus useful to consider also truncated versions β_K , with $K \in [0, \infty)$, such that $\beta_K(\theta) \nearrow \beta_\infty(\theta) := \beta(\theta)$ as $K \nearrow \infty$. In the case that $K < \infty$, the particle system is said to be *cutoff*.

The particle system we will study is an $(\mathbb{R}^3)^N$ -valued Markov process specified for each $K \in [0, \infty]$ through its infinitesimal generator \mathcal{A}_K^N defined as follows: given a Lipschitz bounded function $\Phi : (\mathbb{R}^3)^N \rightarrow \mathbb{R}$ and $\mathbf{v} = (v^1, \dots, v^N) \in (\mathbb{R}^3)^N$, \mathcal{A}_K^N is defined by

$$\mathcal{A}_K^N \Phi(\mathbf{v}) = \frac{1}{2(N-1)} \sum_{i \neq j} \int_{\mathbb{S}^2} d\sigma [\Phi(\mathbf{a}_{ij}(\mathbf{v}, \sigma)) - \Phi(\mathbf{v})] \frac{\beta_K(\theta)}{\sin \theta}, \quad (4.4)$$

where $\mathbf{a}_{ij}(\mathbf{v}, \sigma) \in (\mathbb{R}^3)^N$ corresponds to the vector \mathbf{v} with its i -th and j -th \mathbb{R}^3 -valued components respectively replaced by v' and v'_* given by (4.2) with $v = v^i$ and $v_* = v^j$. The particles start with a given symmetric law G_0^N on $(\mathbb{R}^3)^N$. We denote¹ this stochastic interacting particle system by $\mathbf{V}_t^K = (V_t^{1,K}, \dots, V_t^{N,K})$. Hence, any pair of particles i and j with velocities $v = V_t^{i,K}$ and $v_* = V_t^{j,K}$ interact with deviation angle θ at rate $\beta_K(\theta)/[2(N-1)\sin\theta]$, and then update their velocities to v' and v'_* given by (4.2), with $\sigma \in \mathbb{S}^2$ uniformly chosen at random among unitary vectors such that $\sigma \cdot (v - v_*)/|v - v_*| = \cos\theta$; notice that v' and v'_* as defined in (4.2) now appear in the role of *post-collisional velocities*, consistently with the weak form (4.16) of equation (4.1).

An explicit construction of the system \mathbf{V}^K for the specific cutoff we shall consider (including the case $K = \infty$) will be given in Lemma 4.11, Section 4.2. Let us mention for the moment that, under the assumptions we will make, a unique (in law) Markov process with càdlàg trajectories and generator (4.4) will exist, for each value of $N \in \mathbb{N}$ and $K \in [0, \infty]$. One can check that the quantities $\sum_{i=1}^N v^i$ and $\sum_{i=1}^N |v^i|^2$, corresponding to momentum and kinetic energy, are a.s. exactly preserved by the corresponding random dynamics.

To specify the cutoff we use –which is not a usual one–, we define the functions $H : (0, \pi/2] \rightarrow \mathbb{R}_+$ and $G : \mathbb{R}_+ \rightarrow (0, \pi/2]$ by

$$H(\theta) := \int_{\theta}^{\pi/2} \beta(x) dx \quad \text{and} \quad G(z) = H^{-1}(z),$$

and, for a given cutoff level $K \in [1, \infty]$, we set

$$\beta_K(\theta) = \beta(\theta) \mathbf{1}_{\{\theta \geq G(K)\}}.$$

This cutoff was used by Fournier and Mischler [30] in the study of the *Nanbu type* particle approximations of the Boltzmann equation. We notice however that, contrary to the Markov dynamics (4.4) of binary (also called *Bird type*) interactions, only one particle changes its state upon collisions in the Nanbu system, which therefore preserves momentum and energy only in mean, and thus is less meaningful from the physical point view (but still relevant for numerical simulation purposes).

The goal of the present paper is to establish explicit propagation of chaos rates for the Kac N -particle systems described by (4.4). Our techniques rely on the pathwise approach

¹For notational simplicity, we do not make explicit the dependence of the system on N .

pioneered by Tanaka [58, 59] and Sznitman [57], and on refinements of it based on optimal transport theory, which lead to quantitative estimates in Wasserstein distance (recalled next). These ideas were developed in [25] in the context of Nanbu type Brownian particle approximations of the diffusive Landau equation, and in the aforementioned work [30] concerning Nanbu approximations of the true Boltzmann equation. The main goal of our work is to establish such quantitative results for true binary interactions of Kac particles systems in the context of the Boltzmann equation for Maxwell molecules. To do so, we will extend new pathwise coupling techniques introduced in the study of Kac's one dimensional model in [18], to the context of the present paper. Moreover, we will combine these techniques with a uniform-in- N equilibration result for Kac particles, recently established by Rousset [55], which will allow us to deduce the sharpest uniform in time propagation of chaos estimates in Wasserstein distance so far available in that setting. In the next section we state our main results. See Section 4.1.4 for a more detailed discussion on related works and results in the literature.

4.1.3 Main results

Let us first fix some notation and specify our hypotheses. Given metric space E , $p \geq 1$ and $k \in \mathbb{N}$, let $\mathcal{P}(E)$, $\mathcal{P}_p(E)$ and $\mathcal{P}_p^{\text{sym}}(E^k)$ denote the space of probability measures on E , the subspace of measure on $\mathcal{P}(E)$ with finite p -moment, and the space of symmetric probability measures on $\mathcal{P}_p(E^k)$, respectively. Given a vector $\mathbf{x} = (x^1, \dots, x^N) \in (\mathbb{R}^3)^N$, we define the empirical measures $\bar{\mathbf{x}} \in \mathcal{P}(\mathbb{R}^3)$ and $\bar{\mathbf{x}}^i \in \mathcal{P}(\mathbb{R}^3)$ for any $i \in \{1, \dots, N\}$, by

$$\bar{\mathbf{x}} := \frac{1}{N} \sum_{j=1}^N \delta_{x^j} \quad \text{and} \quad \bar{\mathbf{x}}^i := \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}. \quad (4.5)$$

The empirical measure of the particle system at time $t \geq 0$ is thus denoted by $\bar{\mathbf{V}}_t^K$. Also, given a random vector \mathbf{X} on $(\mathbb{R}^3)^N$, we will denote its law by $\mathcal{L}(\mathbf{X}) \in \mathcal{P}((\mathbb{R}^3)^N)$, and the law of its k first components by $\mathcal{L}^k(\mathbf{X}) \in \mathcal{P}((\mathbb{R}^3)^k)$, for any $k \leq N$.

For $\mu, \nu \in \mathcal{P}_2((\mathbb{R}^3)^k)$, their *quadratic Wasserstein distance* is defined as

$$\mathcal{W}_2(\mu, \nu) = \inf_{\pi} \left(\int_{(\mathbb{R}^3)^k \times (\mathbb{R}^3)^k} |\mathbf{x} - \mathbf{y}|_k^2 \pi(d\mathbf{x}, d\mathbf{y}) \right)^{1/2} = \inf_{\mathbf{X}, \mathbf{Y}} \left(\mathbb{E} |\mathbf{X} - \mathbf{Y}|_k^2 \right)^{1/2},$$

where the first infimum is taken over all $\pi \in \mathcal{P}_2((\mathbb{R}^3)^k \times (\mathbb{R}^3)^k)$ having marginals μ and ν , and the second infimum is taken over all random vectors \mathbf{X} and \mathbf{Y} on $(\mathbb{R}^3)^k$ such that $\mathcal{L}(\mathbf{X}) = \mu$ and $\mathcal{L}(\mathbf{Y}) = \nu$. Here we use the *normalized distance* $|\cdot|_k$ on $(\mathbb{R}^3)^k$ defined by

$$|\mathbf{x}|_k = \left(\frac{1}{k} \sum_{i=1}^k |x^i|^2 \right)^{1/2}.$$

It is known that the infimum is always reached, and such a π achieving the first infimum or a pair (\mathbf{X}, \mathbf{Y}) achieving the second, is known as an *optimal coupling*.

The angular cross section function β will be assumed to satisfy

$$\exists \nu \in (0, 1), \quad c_0 \theta^{-1-\nu} \leq \beta(\theta) \leq c_1 \theta^{-1-\nu} \quad \forall \theta \in (0, \pi/2), \quad (4.6)$$

for some constants $0 < c_0 < c_1$. In this case, it can be easily seen that for some constants $0 < c_2 < c_3$ we have

$$c_2(1+z)^{-1/\nu} \leq G(z) \leq c_3(1+z)^{-1/\nu} \quad \forall z > 0. \quad (4.7)$$

The initial distribution f_0 will be assumed to satisfy

$$\exists p_0 > 4, \quad \int_{\mathbb{R}^3} |v|^{p_0} f_0(dv) < \infty, \quad (4.8)$$

Consider then the hypothesis:

$$\left\{ \begin{array}{l} \mathbf{Maxwell\ molecules:} \\ \gamma = 0, \text{ (4.6) and (4.8) hold,} \end{array} \right. \quad (\text{M})$$

Note that in the usual Maxwellian case, the quantities ν and γ are linked by the relation $\nu = \frac{1-\gamma}{2}$, but in our (slightly more general) context, $\nu \in (0, 1)$ will be viewed as an independent parameter.

The next statement, whose proof can be found for instance in [60] and [61], provides the main analytical properties of equation (4.1) that we shall need. See the comments after Theorem 1.1 in [30] for a detailed discussion of these and other references, and Definition 4.5 below for the notion of weak solutions of (4.1) we will use.

Theorem 4.1. *Assume (M). Then, there exists a unique weak solution $(f_t)_{t \geq 0} \in C([0, \infty), \mathcal{P}_2(\mathbb{R}^3))$ of (4.1), which satisfies $\sup_t \int_{\mathbb{R}^3} |v|^{p_0} f_t(dv) < \infty$.*

We now state our main results.

Theorem 4.2. *Assume (M), and let $(f_t)_{t \geq 0}$ be the unique weak solution of (4.1). Let $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$ be given, and let $\mathbf{V}_t^K = (V_t^{1,K}, \dots, V_t^{N,K})$ be the particle system with generator (4.4) and starting with law G_0^N . Then, there exists a constant $C > 0$ such that for all $t \geq 0$ and all $K \in [1, \infty]$,*

$$\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{V}}_t^K, f_t) \leq C \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C(1+t)^2 N^{-1/3} + CtK^{1-2/\nu}.$$

Note that one can simply choose $G_0^N = f_0^{\otimes N}$ and then $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) = 0$, or assume that the term $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$ goes to 0 at least as fast as $N^{-1/3}$. In either case, the previous theorem gives a chaos result in squared 2–Wasserstein distance for the Bird particle system associated to the Boltzmann equation, with an explicit rate of order $N^{-1/3}$. The time dependence is quadratic.

Our second main result provides a uniform propagation of chaos estimate. In this context, the initial distribution G_0^N will typically be assumed to be a symmetric probability measure concentrated on the *Boltzmann sphere*

$$\mathcal{S}^N = \left\{ \mathbf{x} \in (\mathbb{R}^3)^N : \sum_{i=1}^N x^i = 0, \sum_{i=1}^N |x^i|^2 = N \right\}, \quad (4.9)$$

which implies that $\mathbf{V}_t^K \in \mathcal{S}^N$ for all $t \geq 0$, since the collisions preserve momentum and kinetic energy. Naturally, some f_0 –chaocity of G_0^N has to be imposed, for instance,

$$\lim_{N \rightarrow \infty} \mathcal{W}_2(G_0^N, f_0^{\otimes N}) = 0. \quad (4.10)$$

We will also need the following technical condition: denoting $G_0^k \in \mathcal{P}((\mathbb{R}^3)^k)$ the first k –marginals of G_0^N , we will assume that

$$\int_{\mathbb{R}^3} |v|^4 G_0^1(dv) < \infty, \quad \text{uniformly on } N. \quad (4.11)$$

The tensorized initial condition $G_0^N = f_0^{\otimes N}$ can also be treated by performing a suitable “standarization” procedure that forces the system to lie on \mathcal{S}^N and using the previous case. This procedure requires that f_0 satisfies the condition

$$\int_{\mathbb{R}^3} v f_0(dv) = 0, \quad \text{and} \quad \int_{\mathbb{R}^3} |v|^2 f_0(dv) = 1. \quad (4.12)$$

Additionally, when performing the standarization procedure to $f_0^{\otimes N}$, one needs to check that the analogous of condition (4.11) is satisfied, for which we will assume that

$$f_0 \text{ has a density and } \exists C_0 > 0, f_0(v) < C_0 f_\infty(v), \quad (4.13)$$

where $f_\infty(v)$ is the Gaussian density on \mathbb{R}^3 with mean 0 and variance 1, that is, $f_\infty(v) = (2\pi\sigma^2)^{-3/2} e^{-|v|^2/(2\sigma^2)}$ for $\sigma^2 = 1/3$. See Lemma 4.20. We can now state our second main result:

Theorem 4.3. *Assume (M) and let $(f_t)_{t \geq 0}$ be the unique weak solution of (4.1). Assume either that*

- (i) $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$ is concentrated on \mathcal{S}^N and satisfies (4.10) and (4.11), or

(ii) $G_0^N = f_0^{\otimes N}$ and f_0 satisfies (4.12) and (4.13).

For $K = \infty$, let $\mathbf{V}_t = (V_t^1, \dots, V_t^N)$ be the particle system with generator (4.4) and starting with law G_0^N . Then, there exists $C > 0$ and for all $\epsilon > 0$ there exists $C_\epsilon > 0$ such that

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C_\epsilon N^{-1/6+\epsilon}, \quad \forall t \geq 0.$$

Thus, we obtain a uniform propagation of chaos estimate for Maxwell molecules in squared 2–Wasserstein distance with an explicit rate of order almost $N^{-1/6}$, provided that $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$ converges to 0 with the same rate or faster (for instance, when performing the aforementioned standardization procedure to $f_0^{\otimes N}$, see Lemma 4.18). We mention at once that, together with the estimates of the present paper, the proof relies on a recent equilibration result by Rousset [55], which is why in (i) we assume that the initial distribution is concentrated on \mathcal{S}^N .

Remark 4.4. Theorems 4.2 and 4.3 can also be stated in terms of convergence of the marginal laws of the particle system. For instance, under the hypotheses of Theorem 4.3, we can also obtain for all $k \leq N$,

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{V}_t), f_t^{\otimes k}) \leq C\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C_\epsilon N^{-1/6+\epsilon} + CkN^{-1}, \quad \forall t \geq 0,$$

and similarly for Theorem 4.2.

4.1.4 Comparison to known results and approaches

The study of propagation of chaos for the Boltzmann equation was initiated in the paper [42] by Kac. Propagation of chaos for different instances of the Boltzmann equation has been proved by several authors in terms of weak convergence, without explicit rate. For instance: McKean [46] and Grünbaum [38] proved the chaos property for some models with bounded collision kernel, whereas Sznitman [56] obtained a similar result for unbounded kernel in the hard spheres case, still without rate. Tanaka (see [58] and [59]) introduced the *nonlinear process*, a stochastic jump process whose marginal laws solve the Boltzmann equation. It represents the trajectory of a “typical particle” in the infinite population, and it has been a crucial tool in the probabilistic study of the propagation of chaos property. For instance, by coupling the particle system with independent nonlinear processes, Graham and Méléard [37] obtained one of the first quantitative propagation of chaos results for some cutoff variants of the problem. Their approach however, relying on total variation distance on path space, can not be extended to non–cutoff contexts, and provides bounds which increase exponentially in time.

More recently, Mischler and Mouhot [47] obtained uniform–in–time propagation of chaos results in \mathcal{W}_1 distance for Maxwellian molecules and hard spheres, with a slow (and hard

to track) rate in the number of particles. Their method, of analytic nature, focused on the study of the evolution of the time–marginal laws of the particle system and relied on the comparison between Wasserstein and weaker probability distances.

By far the work closest to the present one is the paper by Fournier and Mischler [30], where, again using coupling with nonlinear processes, the authors obtained a propagation of chaos result with an optimal rate of order $N^{-1/2}$ in squared Wasserstein distance \mathcal{W}_2 (see [28] for the optimality of this rate), with the same dependences on t as the ones we obtain in Theorem 4.2. As concerns the analytical estimates required in our work, we will rely to a large extent on the ones established in [30]. However, the coupling techniques in that work strongly relied on the fact that Nanbu particles are driven by independent sources of randomness, which allows one to couple the particles directly with independent nonlinear processes, as in the classic probabilistic approach. Since in the case of “Bird type” or truly binary interacting particle system (including Kac’s particles systems), all particles are driven by a common source of randomness, such a coupling argument does not work anymore.

In [18], we addressed that problem in the case of Kac’s one dimensional model. More precisely, we introduced a new coupling between the interacting particle system with effective binary interactions and a system of nonlinear processes driven by the same source of randomness, which consequently turn out to be *not independent*. As part of that coupling argument, in a second step, one had to show that these nonlinear processes become, on their turn, independent as N goes to ∞ . The extension of those techniques to the Boltzmann equation is the main methodological contribution of the present work. Note that the rate of order $N^{-1/3}$ that we obtain here and in [18] for Bird particle system is slower than the optimal one for the Nanbu type systems. As was noted in [18], an interesting question is whether this sub–optimality is explained by an intrinsic difference between the Bird and Nanbu type of binary interaction, or if it is only a consequence of our techniques.

As for Theorem 4.3, the best uniform chaos result for Maxwell molecules so far can be found in [47]. In their Theorem 5.1, they obtain an estimate in \mathcal{W}_1 –distance with a slow polynomial rate on N . The rate of order almost $N^{-1/6}$ in squared \mathcal{W}_2 –distance provided by Theorem 1 is therefore a remarkable improvement. We believe however that it is still fairly sub–optimal and that it should be possible to increase the rate up to $N^{-1/3}$, by refining the techniques used in the present paper. Also, the hypotheses on the initial distribution G_0^N considered here are less restrictive: in the case (i) it can be a general symmetric law concentrated on the Boltzmann sphere satisfying the f_0 –chaocity condition (4.10) and the fourth–moment condition (4.11) (and not only those obtained by conditioning to \mathcal{S}^N), and in the case (ii) the distribution f_0 is not restricted to have bounded support.

4.1.5 Weak solutions and nonlinear processes

Let us recall the notion of weak solutions for (4.1), for which we will need the following definitions. We follow [30]: consider measurable functions $\hat{i}, \hat{j} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ such that for every $x \neq 0$,

$$\left(\frac{x}{|x|}, \frac{\hat{i}(x)}{|x|}, \frac{\hat{j}(x)}{|x|} \right)$$

is an orthonormal basis of \mathbb{R}^3 . They can also be assumed to be homogeneous functions, that is, they satisfy $\hat{i}(\lambda x) = \lambda \hat{i}(x)$ and $\hat{j}(\lambda x) = \lambda \hat{j}(x)$ for all $x \in \mathbb{R}^3$ and all $\lambda \in \mathbb{R}$. For $v, v_* \in \mathbb{R}^3$, $\theta \in (0, \pi/2)$, $\phi \in [0, 2\pi)$, and $z \geq 0$, define

$$\begin{aligned} \Gamma(x, \phi) &:= (\cos \phi) \hat{i}(x) + (\sin \phi) \hat{j}(x), \\ a(v, v_*, \theta, \phi) &:= -\frac{1 - \cos \theta}{2} (v - v_*) + \frac{\sin \theta}{2} \Gamma(v - v_*, \phi), \\ c(v, v_*, z, \phi) &:= a(v, v_*, G(z), \phi). \end{aligned} \quad (4.14)$$

Note that when ϕ ranges in $[0, 2\pi)$, the vector $v + a(v, v_*, \theta, \phi)$ goes all over the circle centered at $b = \frac{v+v_*}{2} + \cos \theta \frac{v-v_*}{2}$, with radius $r = \sin \theta \frac{|v-v_*|}{2}$, and orthogonal to $d = \frac{v-v_*}{|v-v_*|}$. Denote this circle by $\mathcal{C}(v, v_*, \theta)$, or alternatively by $\hat{\mathcal{C}}(b, r, d)$.

These definitions provide a suitable parametrization of the post-collisional velocities: it is straightforward to verify that for all $v, v_* \in \mathbb{R}^3$ and for any Lipschitz bounded measurable function Φ on \mathbb{R}^3 ,

$$\begin{aligned} &\int_{\mathbb{S}^2} d\sigma [\Phi(v') - \Phi(v)] B(|v - v_*|, \theta) \\ &= \int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi [\Phi(v + a(v, v_*, \theta, \phi)) - \Phi(v)] \beta(\theta) \\ &= \int_0^\infty dz \int_0^{2\pi} d\phi [\Phi(v + c(v, v_*, z, \phi)) - \Phi(v)], \end{aligned} \quad (4.15)$$

which is well defined under (M). Indeed, if (4.6) holds, then the fact that $|a(v, v_*, \theta, \phi)| \leq C\theta|v - v_*|$ and that $\int_0^{\pi/2} \theta \beta(\theta) d\theta < \infty$ imply that the integral is well defined. By a slight abuse of notation, we still call $v' = v + a(v, v_*, \theta, \phi)$, $v'_* = v_* - a(v, v_*, \theta, \phi)$ and also $v' = v + c(v, v_*, z, \phi)$, $v'_* = v_* - c(v, v_*, z, \phi)$.

Definition 4.5. *We say that a collection $(f_t)_{t \geq 0} \in C([0, \infty), \mathcal{P}_2(\mathbb{R}^3))$ is a weak solution for (4.1) if it preserves momentum and energy (that is, $\int v f_t(dv) = \int v f_0(dv)$ and $\int |v|^2 f_t(dv) = \int |v|^2 f_0(dv)$ for all $t \geq 0$), and if for all bounded Lipschitz function $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ and for all $t \geq 0$,*

$$\begin{aligned} \int_{\mathbb{R}^3} \Phi(v) f_t(dv) &= \int_{\mathbb{R}^3} \Phi(v) f_0(dv) + \int_0^t ds \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_s(dv) f_s(dv_*) |v - v_*|^\gamma \int_0^{\pi/2} d\theta \beta(\theta) \\ &\quad \times \int_0^{2\pi} d\phi [\Phi(v + a(v, v_*, \theta, \phi)) - \Phi(v)]. \end{aligned} \quad (4.16)$$

The nonlinear process introduced by Tanaka in [58] and [59] to provide a probabilistic interpretation of the Boltzmann equation, can be defined in the present cases through a stochastic integral equation with respect to a suitable Poisson point measure. Consider indeed a generic Poisson point measure $\mathcal{M}(dt, dz, d\phi, d\xi)$ on $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times A$ with intensity $dt dz d\phi d\xi / (2\pi|A|)$, where $A \subseteq \mathbb{R}$ is any interval (for instance, $A = [0, 1]$). Consider also a function $F_t(\xi)$, measurable in $(t, \xi) \in [0, \infty) \times A$, such that, whenever ξ is uniformly distributed on A , for each $t \geq 0$ the random variable $F_t(\xi)$ has the distribution f_t given by Theorem 4.1, which solves (4.16).² Given an f_0 -distributed random variable W_0 , independent from \mathcal{M} , consider then a process W_t solution of the following stochastic integral equation:

$$dW_t = \int_0^\infty \int_0^{2\pi} \int_A c(W_{t-}, F_t(\xi), z, \phi) \mathcal{M}(dt, dz, d\phi, d\xi). \quad (4.17)$$

Under (M), it can be shown (e.g. by adapting the arguments of Proposition 4.1 of [30]) that there exists a unique strong solution of this equation, and that it satisfies $\mathcal{L}(W_t) = f_t$, for all $t \geq 0$. A process with the same law as W is called a *nonlinear process*.

Unfortunately, some technical problems oblige us to work with the *cutoff* nonlinear process, see Remark 4.10. Given a cutoff level $L > 0$, we denote this process W^L , and it is defined as the solution to the stochastic equation

$$dW_t^L = \int_0^L \int_0^{2\pi} \int_A c(W_{t-}^L, F_t^L(\xi), z, \phi) \mathcal{M}(dt, dz, d\phi, d\xi), \quad (4.18)$$

where this time $F_t^L(\xi)$ is a ξ -realization of the solution f_t^L of (4.16) when one replaces the function β with β_L (for which the well-posedness Theorem 4.1 applies). Since in (4.18) the total rate is bounded, strong existence and uniqueness of W^L is obvious, and it holds $\mathcal{L}(W_t^L) = f_t^L$ for all t . Naturally, we will have $f_t^L \rightarrow f_t$ when $L \rightarrow \infty$, see Lemma 4.17.

4.1.6 Idea of the proofs and plan of the paper

As mentioned, probabilistic proofs of quantitative propagation of chaos results for the Boltzmann equation have mainly relied on coupling the particle systems with independent copies of nonlinear processes.

To prove our results we will use a new a coupling argument introduced in [18], which consists in coupling *in some optimal way*, for each N and cutoff parameters K and L , a suitable realization of the the particle system \mathbf{V}_t^K with generator (4.4) (which is given in (4.19)) with a system $\mathbf{U}_t^L = (U_t^{1,L}, \dots, U_t^{N,L})$ of copies of the cutoff nonlinear process W^L , and then letting $L \rightarrow \infty$. To achieve this, we will make use of optimal transport theory, in order to carefully construct the jumps of the system \mathbf{U}^L , in such a way that they mimic as

² Notice that the existence of such a function $F_t(\xi)$ is standardly granted by the measurability of the flow of probability laws $t \mapsto f_t(dx)$, see e.g. Lemma 3.22 of [43].

closely as possible the jumps of the particle system \mathbf{V}^K . As in [18], the fact that we deal with a particle system with effective binary collisions will imply that the cutoff nonlinear processes $U^{1,L}, \dots, U^{N,L}$ obtained with this coupling are *not independent*. Therefore, in a similar way as in [18], in order to obtain our estimates we will need, in a second step, to “decouple” these processes (see Lemma 4.14).

In Section 4.2 we give an explicit construction of the particle system \mathbf{V}_t^K and, more importantly, we construct the coupling with the corresponding system \mathbf{U}_t^L of non-independent cutoff nonlinear processes, which will be used throughout this paper. In Section 4.3 we state and prove several technical results. Finally, in Section 4.4 we prove Theorems 4.2 and 4.3.

4.2 Construction

In this section we provide the explicit construction of $(\mathbf{V}^K, \mathbf{U}^L)$ that we introduce in order to prove our results. They will be defined as solutions of stochastic integral equations driven by the same Poisson point measure. We follow [30] and [18].

4.2.1 The particle system

Fix the number of particles $N \in \mathbb{N}$. We introduce the function $\mathbf{i} : [0, N) \rightarrow \{1, \dots, N\}$ by $\mathbf{i}(\xi) = \lfloor \xi \rfloor + 1$, so that $\mathbf{i}(\xi)$ is a discrete index associated to the continuous variable ξ . Let $\mathcal{G} \subseteq \mathbb{R}^2$ be the set

$$\mathcal{G} = \{(\xi, \zeta) \in [0, N)^2 : \mathbf{i}(\xi) \neq \mathbf{i}(\zeta)\}.$$

Note that its area is $|\mathcal{G}| = N(N-1)$. Consider now a Poisson point measure $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$ on $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times [0, N) \times [0, N)$ with intensity

$$\frac{N}{2} dt dz \frac{d\phi d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta)}{2\pi |\mathcal{G}|} = \frac{dt dz d\phi d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta)}{4(N-1)\pi}.$$

In words, the measure \mathcal{N} picks atoms $(t, z) \in [0, \infty)^2$ with intensity $\frac{N}{2} dt dz$ and for each such atom it also independently samples an angle ϕ uniformly from $[0, 2\pi)$ and a pair (ξ, ζ) uniformly from the set \mathcal{G} . We will use the variables ξ and ζ to choose indexes $i = \mathbf{i}(\xi)$ and $j = \mathbf{i}(\zeta)$ of the particles that interact at each jump. Additionally, given $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$ and $f_0 \in \mathcal{P}_2(\mathbb{R}^3)$, let $(\mathbf{V}_0, \mathbf{U}_0)$ be a realization, independent of \mathcal{N} , of the optimal coupling between G_0^N and $f_0^{\otimes N}$. Call $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ the complete, right-continuous filtration generated by $(\mathbf{V}_0, \mathbf{U}_0)$ and \mathcal{N} . We call \mathbb{P} and \mathbb{E} the probability and expectation in the corresponding probability space.

Fix now a cutoff level $K \in [1, \infty]$. For $v, v_* \in \mathbb{R}^3$, $z \geq 0$ and $\phi \in [0, 2\pi)$, set $c_K(v, v_*, z, \phi) = c(v, v_*, z, \phi) \mathbf{1}_{\{z \leq K\}}$. We can now introduce the particle system $\mathbf{V}^K = (V^{1,K}, \dots, V^{N,K})$ as

the solution, starting from the initial condition \mathbf{V}_0 , of the stochastic equation

$$d\mathbf{V}_t^K = \int_0^\infty \int_0^{2\pi} \int_{[0,N]^2} \sum_{i \neq j} \mathbf{1}_{\{\mathbf{i}(\xi)=i, \mathbf{i}(\zeta)=j\}} \mathbf{c}_{K,ij}(\mathbf{V}_{t^-}^K, z, \phi) \mathcal{N}(dt, dz, d\phi, d\xi, d\zeta) \quad (4.19)$$

where $\mathbf{c}_{K,ij}(\mathbf{x}, z, \phi) \in (\mathbb{R}^3)^N$ is the vector with coordinates given by

$$(\mathbf{c}_{K,ij}(\mathbf{x}, z, \phi))^l = \begin{cases} c_K(x^i, x^j, z, \phi) & \text{if } l = i, \\ -c_K(x^j, x^i, z, \phi) & \text{if } l = j, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 4.6. For finite values of the cutoff parameter K (hence finite total jump rate), there is strong existence and pathwise uniqueness for the system of SDE (4.19), as a solution can be constructed in a unique way (jump by jump) from the atoms of \mathcal{N} and the initial conditions. It is clear then (by Itô's formula) that the solution is a Markov process with generator (4.4) and that the law of this processes is uniquely determined. See Lemma 4.11 below for a statement in the case $K = \infty$.

Since we required \hat{i} and \hat{j} to be homogeneous functions, in particular they are odd functions and it can be easily seen that $c(v, v_*, z, \phi) = -c(v_*, v, z, \phi)$. If a solution to (4.19) exists, it follows that for each $i = 1, \dots, N$, the particle $V^{i,K}$ satisfies the stochastic equation

$$dV_t^{i,K} = \int_0^\infty \int_0^{2\pi} \int_0^N c_K(V_{t^-}^{i,K}, V_{t^-}^{\mathbf{i}(\xi),K}, z, \phi) \mathcal{N}^i(dt, dz, d\phi, d\xi), \quad (4.20)$$

where \mathcal{N}^i is given by

$$\mathcal{N}^i(dt, dz, d\phi, d\xi) := \mathcal{N}(dt, dz, d\phi, d\xi, [i-1, i)) + \mathcal{N}(dt, dz, d\phi, [i-1, i), d\xi).$$

That is, \mathcal{N}^i selects only the atoms of \mathcal{N} such that either $\mathbf{i}(\xi) = i$ or $\mathbf{i}(\zeta) = i$. Clearly, \mathcal{N}^i is a Poisson point measure on $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times [0, N)$ with intensity

$$\frac{dtdz d\phi d\xi \mathbf{1}_{A^i}(\xi)}{2(N-1)\pi},$$

where $A^i := [0, N) \setminus [i-1, i)$. Thus, the term $V_{t^-}^{\mathbf{i}(\xi),K}$ appearing in (4.20) is a ξ -realization of the (random) probability measure $\bar{\mathbf{V}}_{t^-}^{i,K}$ (defined as in (4.5)). Therefore, from the point of view of the particle $V^{i,K}$, the dynamics is as follows: (t, z) -atoms are sampled with intensity 1 and for each such atom (not excluded by the cutoff) an angle ϕ is chosen and a particle $v_* = V_{t^-}^{\mathbf{i}(\xi),K}$ is selected at random among all the others; $v = V_{t^-}^{i,K}$ (and $v_* = V_{t^-}^{\mathbf{i}(\xi),K}$) then updates its state to $v' = V_t^{i,K}$ (and $v'_* = V_t^{\mathbf{i}(\xi),K}$) as given in (4.2).

4.2.2 Coupling with system of cutoff nonlinear processes

The key observation is the following: in (4.20) with $K = L$, if one replaces $V_t^{-\mathbf{i}(\xi),K}$ by some realization of the probability measure f_t^L , then the resulting equation defines a cutoff nonlinear process as in (4.18). Moreover: we want to choose this f_t^L -distributed random variable in an optimal way (in the suitable sense), in order that the resulting process remains close to $V^{i,K}$. This needs to be carried out in a measurable way, which motivates the following lemma. In the sequel, all optimal couplings and optimal costs considered use the cost function $C(v, u) = |v - u|^2$.

Lemma 4.7 (coupling). *Fix $L > 0$ and $i \in \{1, \dots, N\}$. Then, there exists an \mathbb{R}^3 -valued function $\Pi_t^{i,L}(\mathbf{x}, \xi)$, measurable in $(t, \mathbf{x}, \xi) \in [0, \infty) \times (\mathbb{R}^3)^N \times A^i$, with the following property: for any $(t, \mathbf{x}) \in [0, \infty) \times (\mathbb{R}^3)^N$ and any random variable ξ uniformly chosen in A^i , the pair $(x^{\mathbf{i}(\xi)}, \Pi_t^{i,L}(\mathbf{x}, \xi))$ is an optimal coupling between $\bar{\mathbf{x}}^i$ and f_t^L . Moreover, for any exchangeable random vector $\mathbf{X} \in (\mathbb{R}^3)^N$ and any bounded measurable function h , we have $\mathbb{E} \int_{j-1}^j h(\Pi_t^{i,L}(\mathbf{X}, \xi)) d\xi = \int_{\mathbb{R}^3} h(u) f_t^L(du)$ for any $j \in \{1, \dots, N\}$, $j \neq i$.*

Proof. See Lemma 3 in [18]. □

To ensure that the post-collisional velocities of $V_t^{i,K}$ and $U_t^{i,L}$ do not differ much, we will use the functions $\Pi^{i,L}$ of Lemma 4.7 to define our system $\mathbf{U}^L = (U^{1,L}, \dots, U^{N,L})$ of cutoff nonlinear processes. This will mean that at each jump of $v = V_t^{i,K}$ together with some other particle v_* , the corresponding process $u = U_t^{i,L}$ will sample some f_t^L -distributed variable u_* to interact with, in such a way that the interactions of the system \mathbf{U}^L mimic those of the particle system \mathbf{V}^K .

However, post-collisional velocities will also depend on the angles ϕ chosen in circles of the form $\mathcal{C}(v, v_*, \theta)$ and $\mathcal{C}(u, u_*, \theta)$ associated with each collision. As remarked by Tanaka [58, 59], no continuity assumption can be made about the functions \hat{i} and \hat{j} , and, in order to control the distance between $V_t^{i,K}$ and $U_t^{i,L}$ after a collision, one has to make specific (uniformly random) choices for those angles as well. In the present paper, we will choose the angles ϕ uniformly in the circles $\mathcal{C}(v, v_*, \theta)$ and $\mathcal{C}(u, u_*, \theta)$ in such a way that their joint distribution is an optimal coupling of the uniform laws on these circles, with respect to the quadratic cost. The optimal transport cost happens to have a nice explicit formula, and the optimal transport map turns out to depend only on $v - v_*$ and $u - u_*$. This is stated in the following:

Lemma 4.8 (optimal coupling of circles). *Recall that $\hat{\mathcal{C}}(b, r, d)$ denotes the circle centered at $b \in \mathbb{R}^3$, with radius $r > 0$ and orthogonal direction $d \in \mathbb{S}^2$; alternatively, for $v, v_* \in \mathbb{R}^3$ and $\theta \in [0, \pi/2]$, $\mathcal{C}(v, v_*, \theta)$ denotes the circle centered at $\frac{v+v_*}{2} + \cos \theta \frac{v-v_*}{2}$, with radius $\sin \theta \frac{|v-v_*|}{2}$ and orthogonal to $\frac{v-v_*}{|v-v_*|}$. Then:*

(i) For any $b, \tilde{b} \in \mathbb{R}^3$, $r, \tilde{r} \geq 0$ and $d, \tilde{d} \in \mathbb{S}^2$, the optimal transport cost between the uniform distributions on the circles $\hat{\mathcal{C}}(b, r, d)$ and $\hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$ is given by

$$\mathcal{W}_2^2(\text{unif}_{\hat{\mathcal{C}}(b,r,d)}, \text{unif}_{\hat{\mathcal{C}}(\tilde{b},\tilde{r},\tilde{d})}) = |b - \tilde{b}|^2 + (r - \tilde{r})^2 + r\tilde{r}(1 - |d \cdot \tilde{d}|). \quad (4.21)$$

(ii) There exists a measurable function $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \times [0, 2\pi) \rightarrow [0, 2\pi)$ with the following property: for every $v, v_*, u, u_* \in \mathbb{R}^3$, $\theta, \vartheta \in [0, \pi/2]$ and for any random variable ϕ uniformly chosen in $[0, 2\pi)$, the pair

$$(v + a(v, v_*, \theta, \phi), u + a(u, u_*, \vartheta, \phi))$$

where $\varphi = \varphi(v - v_*, u - u_*, \phi)$, is an optimal coupling of the uniform distributions on the circles $\mathcal{C}(v, v_*, \theta)$ and $\mathcal{C}(u, u_*, \vartheta)$.

Proof. We first prove (i). Without loss of generality, assume $\tilde{b} = 0$ and $d \cdot \tilde{d} \geq 0$. Let $h = h(d, \tilde{d}) \in \mathbb{S}^2$ be a fixed measurable choice of a unitary vector orthogonal to both d and \tilde{d} (if they are parallel then there are infinitely many such h 's; if not, there are only 2; we can take for instance $h = \hat{i}(d)/|d|$ in the first case and $h = d \times \tilde{d}/|d \times \tilde{d}|$ in the second). Let also $k, \tilde{k} \in \mathbb{S}^2$ be such that (h, k, d) and $(h, \tilde{k}, \tilde{d})$ are orthonormal bases of \mathbb{R}^3 with the same orientation $s = d \cdot (h \times k) = \tilde{d} \cdot (h \times \tilde{k}) \in \{-1, 1\}$, so that $k \cdot \tilde{k} = sk \cdot (\tilde{d} \times h) = s\tilde{d} \cdot (h \times k) = d \cdot \tilde{d}$. With these bases, we can now parametrize the circles $\hat{\mathcal{C}}(b, r, d)$ and $\hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$ using angles ϕ and $\tilde{\phi} \in [0, 2\pi)$. Namely, a point $x \in \hat{\mathcal{C}}(b, r, d)$ is written as $x = b + r(\cos \phi)k + r(\sin \phi)h$, while a point $y \in \hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$ is written as $y = \tilde{r}(\cos \tilde{\phi})\tilde{k} + \tilde{r}(\sin \tilde{\phi})h$. Then, the associated cost is

$$\begin{aligned} C(\phi, \tilde{\phi}) &= |b + r(\cos \phi)k - \tilde{r}(\cos \tilde{\phi})\tilde{k} + (r \sin \phi - \tilde{r} \sin \tilde{\phi})h|^2 \\ &= |b|^2 + r^2 + \tilde{r}^2 - 2r\tilde{r}[(\sin \phi \sin \tilde{\phi}) + (\cos \phi \cos \tilde{\phi})d \cdot \tilde{d}] \\ &\quad + 2b \cdot [r(\cos \phi)k - \tilde{r}(\cos \tilde{\phi})\tilde{k} + (r \sin \phi - \tilde{r} \sin \tilde{\phi})h]. \end{aligned}$$

Using the inequality $2\alpha\beta \leq \alpha^2 + \beta^2$ in the cross-terms, we obtain $C(\phi, \tilde{\phi}) \geq \Phi(\phi) - \Psi(\tilde{\phi})$ for all $\phi, \tilde{\phi} \in [0, 2\pi)$, where

$$\begin{aligned} \Phi(\phi) &= |b|^2 + r^2 - r\tilde{r}[(\sin \phi)^2 + (\cos \phi)^2 d \cdot \tilde{d}] + 2rb \cdot [(\cos \phi)k + (\sin \phi)h] \\ \Psi(\tilde{\phi}) &= -\{\tilde{r}^2 - r\tilde{r}[(\sin \tilde{\phi})^2 + (\cos \tilde{\phi})^2 d \cdot \tilde{d}] - 2\tilde{r}b \cdot [(\cos \tilde{\phi})\tilde{k} + (\sin \tilde{\phi})h]\}. \end{aligned}$$

Moreover, the equality $C(\phi, \tilde{\phi}) = \Phi(\phi) - \Psi(\tilde{\phi})$ is reached when $\phi = \tilde{\phi}$. Using for instance Remark 5.13 of [62], this shows that taking $\phi = \tilde{\phi}$ uniformly distributed on $[0, 2\pi)$ in fact generates an optimal coupling of the uniform distributions on $\mathcal{C}(b, r, d)$ and $\mathcal{C}(\tilde{b}, \tilde{r}, \tilde{d})$. This proves point (i), since the cost of this coupling is

$$\begin{aligned} \int_0^{2\pi} C(\phi, \phi) \frac{d\phi}{2\pi} &= \frac{1}{2\pi} \int_0^{2\pi} \{ |b|^2 + r^2 + \tilde{r}^2 - 2r\tilde{r}[(\sin \phi)^2 + (\cos \phi)^2 d \cdot \tilde{d}] \\ &\quad + 2b \cdot [(\cos \phi)(rk - \tilde{r}\tilde{k}) + (\sin \phi)(r - \tilde{r})h] \} d\phi \\ &= |b|^2 + r^2 + \tilde{r}^2 - r\tilde{r}(1 + d \cdot \tilde{d}). \end{aligned}$$

We now prove (ii). Put $d = \frac{v-v_*}{|v-v_*|}$ and $\tilde{d} = \frac{u-u_*}{|u-u_*|}$. For some fixed measurable choice $(d, \tilde{d}) \mapsto h = h(d, \tilde{d})$ of a vector h orthogonal to both d and \tilde{d} , let $\phi_i = \phi_i(d, \tilde{d}) \in [0, 2\pi)$, $i = 1, 2$ be the unique angles such that

$$\frac{\Gamma(v - v_*, \phi_1)}{|v - v_*|} = \frac{\Gamma(u - u_*, \phi_2)}{|u - u_*|} =: h,$$

Note that (ϕ_1, ϕ_2) depend only on $v - v_*$ and $u - u_*$ through $d, \tilde{d}, \Gamma(v - v_*, \cdot)$ and $\Gamma(u - u_*, \cdot)$, in a measurable way. Now put

$$k := \frac{\Gamma(v - v_*, \phi_1 + \pi/2)}{|v - v_*|}, \quad \tilde{k} := \frac{\Gamma(u - u_*, \phi_2 + s\pi/2)}{|u - u_*|}.$$

Here $s = \pm 1$ is chosen such that the rotation in $\pi/2$ is performed with the same orientation. More specifically, if $d \cdot \tilde{d} \geq 0$, then $s = 1$ when the bases $(\frac{v-v_*}{|v-v_*|}, \frac{\hat{i}(v-v_*)}{|v-v_*|}, \frac{\hat{j}(v-v_*)}{|v-v_*|})$ and $(\frac{u-u_*}{|u-u_*|}, \frac{\hat{i}(u-u_*)}{|u-u_*|}, \frac{\hat{j}(u-u_*)}{|u-u_*|})$ have the same orientation, and $s = -1$ otherwise; but when $d \cdot \tilde{d} < 0$, we make the opposite choice. Now, the same argument of part (i) shows that if ϕ is a uniform random variable on $[0, 2\pi)$ then $v + a(v, v_*, \theta, \phi)$ and $u + a(u, u_*, \vartheta, s(\phi - \phi_1) + \phi_2)$ constitute an optimal coupling. Put $\varphi = s(\phi - \phi_1) + \phi_2$ and the conclusion follows. \square

Remark 4.9. • The expression on the right of (4.21) is nice: the term $|b - \tilde{b}|^2$ is the cost associated to traslation of the circles, the term $(r - \tilde{r})^2$ is the expansion or contraction cost, and $r\tilde{r}(1 - |d \cdot \tilde{d}|)$ corresponds to inclination.

- When $(v - v_*) \cdot (u - u_*) \geq 0$ and $\theta = \vartheta$, the coupling given in Lemma 4.8–(ii) reduces to the parallel spherical coupling of [55].

With the functions $\Pi^{i,L}$ and φ of Lemmas 4.7 and 4.8–(ii) in hand, we can finally introduce the system of cutoff nonlinear processes $\mathbf{U}^L = (U^{1,L}, \dots, U^{N,L})$. Mimicking (4.20), for each $K \in [0, \infty]$, $L \in [0, \infty)$, $N \in \mathbb{N}$ and $i = 1, \dots, N$, the processes $U^{i,L}$ is defined as the solution, starting from U_0^i , of the stochastic equation

$$dU_t^{i,L} = \int_0^\infty \int_0^{2\pi} \int_0^N c_L(U_t^{i,L}, \Pi_t^{i,L}(\mathbf{U}_t^L, \xi), z, \varphi_t^i) \mathcal{N}^i(dt, dz, d\phi, d\xi), \quad (4.22)$$

where we have used the shorthand

$$\varphi_t^i = \varphi(V_{t^-}^{i,K} - V_{t^-}^{i(\xi),K}, U_{t^-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_t^L, \xi), \phi). \quad (4.23)$$

In words, at jump instants of $U_t^{i,L}$, this process collides with an f_t^L -distributed random variable, which is optimally coupled to the realization $U_{t^-}^{i(\xi),L}$ of the (random) measure $\bar{\mathbf{U}}_{t^-}^{i,L}$. Since the Poisson measures \mathcal{N}^i and \mathcal{N}^j share some of its atoms, the processes $U^{i,L}$ and $U^{j,L}$ have simultaneous jumps, which implies that they are *not independent*. Thus, to make this construction useful, we will need to prove that this dependence is negligible when N is large,

which we will be able to do with estimates that uniform in t and L , see Lema 4.14. Note also that, although $U^{i,L}$ depends implicitly on K through \mathbf{V}^K (but its law does not), we will not make this dependence explicit in our notation.

We can also write the following joint SDE for the pair $(\mathbf{V}^K, \mathbf{U}^L)$, arranged as the collection of pairs $((V^{1,K}, U^{1,L}), \dots, (V^{N,K}, U^{N,L})) \in (\mathbb{R}^3 \times \mathbb{R}^3)^N$:

$$d(\mathbf{V}^K, \mathbf{U}^L)_t = \int_0^\infty \int_0^{2\pi} \int_{[0,N]^2} \sum_{i \neq j} \mathbf{1}_{\{\mathbf{i}(\xi)=i, \mathbf{i}(\zeta)=j\}} \times \mathbf{b}_{KL,ij}(\mathbf{V}_{t^-}^K, \mathbf{U}_{t^-}^L, t, z, \phi, \xi, \zeta) \mathcal{N}(dt, dz, d\phi, d\xi, d\zeta), \quad (4.24)$$

where $\mathbf{b}_{KL,ij}(\mathbf{x}, \mathbf{y}, t, z, \phi, \xi, \zeta) \in (\mathbb{R}^3 \times \mathbb{R}^3)^N$ is given by

$$\begin{aligned} & (\mathbf{b}_{KL,ij}(\mathbf{x}, \mathbf{y}, t, z, \phi, \xi, \zeta))^\ell \\ &= \begin{cases} \left(c_K(x^i, x^j, z, \phi), c_L(y^i, \Pi_t^{i,L}(\mathbf{y}, \zeta), z, \varphi(x^i - x^j, y^i - \Pi_t^{i,L}(\mathbf{y}, \zeta), \phi)) \right) & \text{if } \ell = i, \\ \left(c_K(x^j, x^i, z, \phi), c_L(y^j, \Pi_t^{j,L}(\mathbf{y}, \xi), z, \varphi(x^j - x^i, y^j - \Pi_t^{j,L}(\mathbf{y}, \xi), \phi)) \right) & \text{if } \ell = j, \\ (0, 0) & \text{otherwise.} \end{cases} \end{aligned}$$

Remark 4.10. Unfortunately, due to the lack of smoothness of the functions φ and Π_t^i , we were not able to obtain a well-posedness result for (4.24) in the case $L = \infty$. Such a well-posedness would greatly simplify the construction and proofs. We are thus restricted to work with a system of cutoff nonlinear processes.

The existence of the pair $(\mathbf{V}^K, \mathbf{U}^L)$ solving (4.24) along with their relevant properties, is stated in the next result. To a large extent the arguments of the proof are extensions of the ones in Proposition 1.2. of [30], so we will provide details only when needed.

Lemma 4.11. *Assume (M). For each $K \in [0, \infty]$, $L \in [0, \infty)$ and $N \in \mathbb{N}$, we have:*

- (i) *There is strong existence and uniqueness of a solution $(\mathbf{V}^K, \mathbf{U}^L)$ to the system of SDEs (4.24) in the case $K < \infty$, and weak existence and uniqueness if $K = \infty$.*
- (ii) *\mathbf{V}^K has the law of the unique Markov process with values in $(\mathbb{R}^3)^N$ and generator given by (4.4). In particular, we almost surely have $\sum_{i=1}^N V_t^{i,K} = \sum_{i=1}^N V_0^i$ and $\sum_{i=1}^N |V_t^{i,K}|^2 = \sum_{i=1}^N |V_0^i|^2$ for all $t \geq 0$.*
- (iii) *For each $i = 1, \dots, N$, the process $U^{i,L}$ is a cutoff nonlinear process, and in particular we have $\mathcal{L}(U_t^{i,L}) = f_t^L$.*
- (iv) *Last, the collection of pairs of processes $(V^{1,K}, U^{1,L}), \dots, (V^{N,K}, U^{N,L})$ is exchangeable.*

Proof. In the case $K < \infty$, it is clear that strong well-posedness holds for solutions \mathbf{V}^K of (4.19). Let us prove (i) in the case that $K = \infty$. Given a sequence of finite cutoff levels

$K \rightarrow \infty$, we can follow exactly the same arguments as in Step 2 of the proof of Proposition 1.2–(ii) of [30] to show that the laws of \mathbf{V}^K are tight (the required estimates are indeed simpler to get here because of the exact preservation of second moments) and, by martingale methods (and classic probability space enlargement arguments) that the accumulation points are weak solutions of equation (4.19) with $K = \infty$.

In order to prove uniqueness in law of weak solutions \mathbf{V}^∞ of equation (4.19) it is enough to show that a weak solution in a given probability space can be approximated, in a pathwise way, by (strong) solutions $\hat{\mathbf{V}}^K$ to (4.19) driven by the same Poisson measure, as $K \rightarrow \infty$. Given $\hat{\mathbf{V}}^\infty$ a weak solution to (4.19) driven by $\hat{\mathcal{N}}(dt, dz, d\phi, d\xi, d\zeta)$, define $\hat{\mathbf{V}}^K$ by

$$\begin{aligned} d\hat{\mathbf{V}}_t^K &= \int_0^\infty \int_0^{2\pi} \int_{[0, N]^2} \sum_{i \neq j} \mathbf{1}_{\{i(\xi)=i, i(\zeta)=j\}} \\ &\quad \times \mathbf{c}_{K, ij}(\hat{\mathbf{V}}_{t^-}^K, z, \phi_{t^-}^K) \hat{\mathcal{N}}(dt, dz, d\phi, d\xi, d\zeta) \end{aligned} \quad (4.25)$$

with $\hat{\mathbf{V}}_0^K = \hat{\mathbf{V}}_0$ and where $\phi_t^K = \phi_t^K(\phi, \xi, \zeta) = \varphi(\hat{V}_t^{\mathbf{i}(\xi), \infty} - \hat{V}_t^{\mathbf{i}(\zeta), \infty}, \hat{V}_t^{\mathbf{i}(\xi), K} - \hat{V}_t^{\mathbf{i}(\zeta), K}, \phi)$. Notice that $\hat{\mathbf{V}}^K$ has the same law as \mathbf{V}^K since the point measure $\hat{\mathcal{N}}^K(dt, dz, d\phi, d\xi, d\zeta) := \mathcal{N}(dt, dz, d\phi_t^K(\phi, \xi, \zeta), d\xi, d\zeta)$ has the same law as $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$. Since the coupling between circles we are using is optimal (Lemma 4.8), the bounds obtained in Lemma 5.1 of [30] are also valid in the present context.

If $\mathbf{V}_0 \in L^2$, using Itô calculus and Gronwall's lemma we are then able to prove in a similar way that, for each $T > 0$ and some function $R_T(K)$ going to 0 as $K \rightarrow \infty$, one gets

$$\sup_{t \in [0, T]} \mathbb{E}(|\hat{\mathbf{V}}_t^K - \hat{\mathbf{V}}_t^\infty|^2) \leq R_T(K)$$

and then

$$\mathbb{E}(\sup_{t \in [0, T]} |\hat{\mathbf{V}}_t^K - \hat{\mathbf{V}}_t^\infty|^2) \leq R_T(K),$$

as desired.

Given a solution \mathbf{V}^K to (4.19) (strong if $K < \infty$ and weak if $K = \infty$), since $L < \infty$, we can easily obtain \mathbf{U}^L in a unique way by means of an explicit, jump-by-jump construction. Part (i) follows. Part (ii) is easily seen by Itô calculus. Finally, parts (iii) and (iv) are obvious. \square

4.3 Estimates and technical results

The following lemmas provide useful estimates for our purposes. Typically, one wants to use these lemmas with v and v_* taken from the particle system, and u and u_* taken from the system of cutoff nonlinear processes.

Lemma 4.12. Write $R(v, u) := |v||u| + |v \cdot u| - 2v \cdot u \geq 0$. For any $v, v_*, u, u_* \in \mathbb{R}^3$, $\theta, \vartheta \in [0, \pi/2]$, write $\varphi = \varphi(v - v_*, u - u_*, \phi)$. Then:

$$\begin{aligned} & \int_0^{2\pi} \left(|v + a(v, v_*, \theta, \phi) - u - a(u, u_*, \vartheta, \varphi)|^2 - |v - u|^2 \right) \frac{d\phi}{2\pi} \\ &= - \left[(v - u) + (v_* - u_*) \right] \cdot \left[\frac{1 - \cos \theta}{2} (v - v_*) - \frac{1 - \cos \vartheta}{2} (u - u_*) \right] \\ & \quad - \frac{\sin \theta \sin \vartheta}{4} R(v - v_*, u - u_*) + \frac{1 - \cos(\theta - \vartheta)}{2} (v - v_*) \cdot (u - u_*). \end{aligned} \quad (4.26)$$

Proof. Setting

$$\begin{aligned} b &= \frac{v + v_*}{2} + \cos \theta \frac{v - v_*}{2}, & r &= \sin \theta \frac{|v - v_*|}{2}, & d &= \frac{v - v_*}{|v - v_*|}, \\ \tilde{b} &= \frac{u + u_*}{2} + \cos \vartheta \frac{u - u_*}{2}, & \tilde{r} &= \sin \vartheta \frac{|u - u_*|}{2}, & \tilde{d} &= \frac{u - u_*}{|u - u_*|}, \end{aligned}$$

we have

$$\begin{aligned} & 4(b - \tilde{b})^2 \\ &= |(v - u) + (v_* - u_*) + \cos \theta (v - v_*) - \cos \vartheta (u - u_*)|^2 \\ &= |v - u|^2 + |v_* - u_*|^2 + 2(v - u) \cdot (v_* - u_*) \\ & \quad + \cos^2 \theta |v - v_*|^2 + \cos^2 \vartheta |u - u_*|^2 - 2 \cos \theta \cos \vartheta (v - v_*) \cdot (u - u_*) \\ & \quad + 2[(v - u) + (v_* - u_*)] \cdot [\cos \theta (v - v_*) - \cos \vartheta (u - u_*)] \\ &= 3|v - u|^2 - |v_* - u_*|^2 + 2(v - u) \cdot (v_* - u_*) \\ & \quad + \cos^2 \theta |v - v_*|^2 + \cos^2 \vartheta |u - u_*|^2 - 2 \cos \theta \cos \vartheta (v - v_*) \cdot (u - u_*) \\ & \quad - 2[(v - u) + (v_* - u_*)] \cdot [(1 - \cos \theta)(v - v_*) - (1 - \cos \vartheta)(u - u_*)] \end{aligned} \quad (4.27)$$

and

$$\begin{aligned} & 4(r - \tilde{r})^2 + 4r\tilde{r}(1 - |d \cdot \tilde{d}|) \\ &= \sin^2 \theta |v - v_*|^2 + \sin^2 \vartheta |u - u_*|^2 \\ & \quad - \sin \theta \sin \vartheta (|v - v_*||u - u_*| + |(v - v_*) \cdot (u - u_*)|) \\ &= \sin^2 \theta |v - v_*|^2 + \sin^2 \vartheta |u - u_*|^2 - \sin \theta \sin \vartheta R(v - v_*, u - u_*) \\ & \quad - 2 \sin \theta \sin \vartheta (v - v_*) \cdot (u - u_*). \end{aligned} \quad (4.28)$$

Adding (4.27) and (4.28), using that $|v - v_*|^2 + |u - u_*|^2 + 2(v - u) \cdot (v_* - u_*) = |v - u|^2 + |v_* - u_*|^2 + 2(v - v_*) \cdot (u - u_*)$ and the identity $\cos \theta \cos \vartheta + \sin \theta \sin \vartheta = \cos(\theta - \vartheta)$, yields

$$\begin{aligned} & 4(b - \tilde{b})^2 + 4(r - \tilde{r})^2 + 4r\tilde{r}(1 - |d \cdot \tilde{d}|) \\ &= 4|v - u|^2 - \sin \theta \sin \vartheta R(v - v_*, u - u_*) \\ & \quad + 2(1 - \cos(\theta - \vartheta))(v - v_*) \cdot (u - u_*) \\ & \quad - 2[(v - u) + (v_* - u_*)] \cdot [(1 - \cos \theta)(v - v_*) - (1 - \cos \vartheta)(u - u_*)]. \end{aligned}$$

Thanks to Lemma 4.8–(ii), φ is an optimal transport map, and then the integral on the left side of (4.26) without the term $-|v - u|^2$ is actually the cost given by Lemma 4.8–(i), that is, $(b - \tilde{b})^2 + (r - \tilde{r})^2 + r\tilde{r}(1 - |d \cdot \tilde{d}|)$. Dividing by 4 and subtracting $|v - u|^2$ in the above identity, the result follows. \square

Corollary 4.13. *Fix any $K, L \in [0, \infty]$ with $K \leq L$, and define $\Phi_K^L := \int_K^L \frac{1 - \cos G(z)}{2} dz \geq 0$. For any $v, v_*, u, u_* \in \mathbb{R}^3$, write $\varphi = \varphi(v - v_*, u - u_*, \phi)$. Then we have*

$$\begin{aligned} & \int_0^\infty \int_0^{2\pi} (|v + c_K(v, v_*, z, \phi) - u - c_L(u, u_*, z, \varphi)|^2 - |v - u|^2) \frac{d\phi}{2\pi} dz \\ &= \Phi_0^K [-|v - u|^2 + |v_* - u_*|^2] + \Phi_K^L (u - u_*) \cdot (2v - u - u_*) \\ &\quad - R(v - v_*, u - u_*) \int_0^K \frac{\sin^2 G(z)}{4} dz. \\ &\leq \Phi_0^K [-|v - u|^2 + |v_* - u_*|^2] + C(|v| + |u| + |u_*|)^2 (1 + K)^{1-2/\nu}. \end{aligned}$$

Proof. Split the integral with respect to z into \int_0^K and \int_K^L . For the first integral we have $c_K(v, v_*, z, \phi) = a(v, v_*, \theta, \phi)$ and $c_L(u, u_*, z, \varphi) = a(u, u_*, \theta, \varphi)$ for $\theta = G(z)$; using Lemma 4.12 yields the first and third terms in the equality. For the second integral we have $c_K(v, v_*, z, \phi) = a(v, v_*, 0, \phi)$, so this time we use Lemma 4.12 with $\theta = 0$ and $\vartheta = G(z)$, which gives the second term. The inequality is obtained discarding the negative third term, noting that $\Phi_K^L \leq C \int_K^\infty G^2(z) dz$, and using (4.7). \square

The next lemma is of key importance, since it gives a decoupling estimate for the system of non-independent cutoff nonlinear processes \mathbf{U}^L . The proof, also relying on a coupling argument, follows Lemma 6 of [18].

Lemma 4.14 (decoupling). *Assume (M) and take $L \geq 1$. Then, there exists a constant C independent of L such that for all $k \in \{1, \dots, N\}$ and all $t \geq 0$,*

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{U}_t^L), (f_t^L)^{\otimes k}) \leq C \frac{k}{N}.$$

Proof. Given $k \in \{1, \dots, N\}$ fixed, we will construct k independent cutoff nonlinear processes $\tilde{U}^{1,L}, \dots, \tilde{U}^{k,L}$ such that $\mathbb{E}|U_t^{i,L} - \tilde{U}_t^{i,L}|^2$ is small, for all $i = 1, \dots, k$. To achieve this, the idea is the following: when $U_t^{i,L}$ has a simultaneous jump with some $U_t^{j,L}$ with $j \notin \{1, \dots, k\}$, then the process $\tilde{U}_t^{i,L}$ will use the same sample of f_t^L used by $U_t^{i,L}$ to define its own jump; but when $j \in \{1, \dots, k\}$, then one of the processes, $\tilde{U}_t^{i,L}$ or $\tilde{U}_t^{j,L}$, will not jump at that instant. We will use an additional, independent source of randomness to define new jumps that compensate for the missing ones. Since, when $k \ll N$, the second kind of jump occurs much less frequently, this construction will give the desired estimate.

Consider a Poisson point measure \mathcal{M} that is an independent copy of \mathcal{N} , also independent from $(\mathbf{V}_0, \mathbf{U}_0)$, and define for each $i \in \{1, \dots, k\}$

$$\begin{aligned} \mathcal{M}^i(dt, dz, d\phi, d\xi) &= \mathcal{N}(dt, dz, d\phi, [i-1, i], d\xi) \\ &\quad + \mathcal{N}(dt, dz, d\phi, d\xi, [i-1, i]) \mathbf{1}_{[k, N)}(\xi) \\ &\quad + \mathcal{M}(dt, dz, d\phi, d\xi, [i-1, i]) \mathbf{1}_{[0, k)}(\xi). \end{aligned}$$

That is, \mathcal{M}^i selects the atoms of $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$ where either $(\mathbf{i}(\xi) = i)$, or $(\mathbf{i}(\zeta) = i$ and $\mathbf{i}(\xi) \notin \{1, \dots, k\})$, and to make up for the dropped atoms it also selects new ones from $\mathcal{M}(dt, dz, d\phi, d\xi, d\zeta)$, where $\mathbf{i}(\zeta) = i$ and $\mathbf{i}(\xi) \in \{1, \dots, k\}$. This ensures that no such atom appears in two \mathcal{M}^i 's, implying that they are independent Poisson point measures, all with intensity $dtdzd\phi d\xi \mathbf{1}_{A^i}(\xi)/[2(N-1)\pi]$, just like \mathcal{N}^i .

Mimicking (4.22), we define $\tilde{U}^{i,L}$ as the solution, starting from $\tilde{U}_0^{i,L} = U_0^i$, of the stochastic equation

$$d\tilde{U}_t^{i,L} = \int_0^\infty \int_0^{2\pi} \int_0^N c_L(\tilde{U}_{t^-}^{i,L}, \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi), z, \tilde{\varphi}_{t^-}^i) \mathcal{M}^i(dt, dz, d\phi, d\xi). \quad (4.29)$$

Here we write

$$\tilde{\varphi}_{t^-}^i = \varphi(U_{t^-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi), \tilde{U}_{t^-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi), \varphi_{t^-}^i),$$

where $\varphi_{t^-}^i$ was defined in (4.23). In other words: $\tilde{\varphi}_{t^-}^i$ takes the angle $\varphi_{t^-}^i$ and maps it to $[0, 2\pi)$ in such a way that the resulting pair $(\varphi_{t^-}^i, \tilde{\varphi}_{t^-}^i)$ parametrizes (as a function of ϕ) an optimal coupling of the uniform distributions on the circles with orthogonal directions $U_{t^-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi)$ and $\tilde{U}_{t^-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi)$, whenever ϕ is uniformly chosen on $[0, 2\pi)$. The latter ensures closeness of the states of $U_t^{i,L}$ and $\tilde{U}_t^{i,L}$ after the joint jump.

If we define $\tilde{\mathcal{M}}^{i,L}(dt, dz, d\phi, dv)$ to be the point measure on $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times \mathbb{R}^3$ with atoms $(t, z, \tilde{\varphi}_{t^-}^i, \Pi_t^{i,L}(\mathbf{U}_{t^-}^L, \xi))$ for every atom (t, z, ϕ, ξ) of \mathcal{M}^i , it is clear that $\tilde{U}^{i,L}$ depends only on $\tilde{\mathcal{M}}^{i,L}$ and U_0^i . Since: (i) the dependence on \mathbf{V}^K and \mathbf{U}^L in (4.29) is predictable with respect to \mathcal{N} , \mathcal{M} and the initial data, (ii) the Poisson measures $\mathcal{M}^1, \dots, \mathcal{M}^k$ are independent and (iii) the ξ -law of $\Pi_t^{i,L}(\mathbf{x}, \xi)$ is f_t^L for every $\mathbf{x} \in \mathbb{R}^N$, one can use the compensation formula to compute the joint Laplace functional of $\tilde{\mathcal{M}}^{1,L}, \dots, \tilde{\mathcal{M}}^{k,L}$ and deduce that they are independent Poisson point measures, all with intensity $dtdzd\phi f_t^L(dv)/(2\pi)$. This shows that each $\tilde{U}^{i,L}$ is a cutoff nonlinear process and that they are independent.

Consequently, we have

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{U}_t^L), (f_t^L)^{\otimes k}) \leq \mathbb{E} \frac{1}{k} \sum_{j=1}^k |U_t^{j,L} - \tilde{U}_t^{j,L}|^2 = \mathbb{E} |U_t^{i,L} - \tilde{U}_t^{i,L}|^2.$$

Thus, it suffices to estimate the quantity $h_t := \mathbb{E} |U_t^{i,L} - \tilde{U}_t^{i,L}|^2$, for any $i \in \{1, \dots, k\}$ fixed. We can write

$$h_t - h_s = J_{s,t}^1 + J_{s,t}^2 + J_{s,t}^3 \quad (4.30)$$

where $J_{s,t}^1$ is the term associated to simultaneous jumps of $U^{i,L}$ and $\tilde{U}^{i,L}$, $J_{s,t}^2$ corresponds to jumps of $U^{i,L}$ alone, and $J_{s,t}^3$ corresponds to jumps of $\tilde{U}^{i,L}$ alone. To write this terms explicitly, let us first shorten notation: write $U_r := U_r^{i,L}$, $\tilde{U}_r := \tilde{U}_r^{i,L}$, $\Pi_r := \Pi_r^{i,L}(\mathbf{U}_r^L, \xi)$, $c_r := c(U_r, \Pi_r, z, \varphi_r^i)$ and $\tilde{c}_r := c(\tilde{U}_r, \Pi_r, z, \tilde{\varphi}_r^i)$. With these: from (4.22) and (4.29), $J_{s,t}^1$, $J_{s,t}^2$ and $J_{s,t}^3$ are given by

$$\begin{aligned} J_{s,t}^1 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N (|U_{r^-} + c_{r^-} - \tilde{U}_{r^-} - \tilde{c}_{r^-}|^2 - |U_{r^-} - \tilde{U}_{r^-}|^2) \\ &\quad \times [\mathcal{N}(dr, dz, d\phi, [i-1, i], d\xi) + \mathcal{N}(dr, dz, d\phi, d\xi, [i-1, i]) \mathbf{1}_{[k,N)}(\xi)], \\ J_{s,t}^2 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N (|U_{r^-} + c_{r^-} - \tilde{U}_{r^-}|^2 - |U_{r^-} - \tilde{U}_{r^-}|^2) \\ &\quad \times \mathcal{N}(dr, dz, d\phi, d\xi, [i-1, i]) \mathbf{1}_{[0,k)}(\xi), \\ J_{s,t}^3 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N (|U_{r^-} - \tilde{U}_{r^-} - \tilde{c}_{r^-}|^2 - |U_{r^-} - \tilde{U}_{r^-}|^2) \\ &\quad \times \mathcal{M}(dr, dz, d\phi, d\xi, [i-1, i]) \mathbf{1}_{[0,k)}(\xi). \end{aligned}$$

Recall that \mathcal{N} and \mathcal{M} have intensity $dt dz d\phi d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta) / [4(N-1)\pi]$. Note that $\int_{i-1}^i \mathbf{1}_{\mathcal{G}}(\xi, \zeta) d\zeta = \mathbf{1}_{A^i}(\xi)$, where $A^i = [0, N) \setminus [i-1, i)$. Using the compensation formula, the Poisson point measures in the integrals can be replaced by their intensities, and we obtain for $J_{s,t}^1$:

$$\begin{aligned} J_{s,t}^1 &= \mathbb{E} \int_s^t \int_0^L \int_0^{2\pi} \int_0^N (|U_r + c_r - \tilde{U}_r - \tilde{c}_r|^2 - |U_r - \tilde{U}_r|^2) [\mathbf{1}_{A^i}(\xi) + \mathbf{1}_{[k,N)}(\xi)] \frac{dr dz d\phi d\xi}{4(N-1)\pi} \\ &\leq -\Phi_0^L \mathbb{E} \int_s^t \int_0^N |U_r - \tilde{U}_r|^2 [\mathbf{1}_{A^i}(\xi) + \mathbf{1}_{[k,N)}(\xi)] \frac{dr d\xi}{2(N-1)} \\ &\leq -\frac{\Phi_0^L}{2} \int_s^t h_r dr, \end{aligned} \tag{4.31}$$

where we have used Corollary 4.13 with $v = U_r$, $u = \tilde{U}_r$, $v_* = u_* = \Pi_r$, $K = L$ and the change of variable $(\phi, \varphi_r^i) \mapsto (\tilde{\varphi}_r^i, \tilde{\varphi}_r^i)$. For $J_{s,t}^2$ we get:

$$\begin{aligned} J_{s,t}^2 &= \mathbb{E} \int_s^t \int_0^L \int_0^{2\pi} \int_0^N (|U_r + c_r - \tilde{U}_r|^2 - |U_r - \tilde{U}_r|^2) \mathbf{1}_{A^i \cap [0,k)}(\xi) \frac{dr dz d\phi d\xi}{4(N-1)\pi} \\ &\leq C \mathbb{E} \int_s^t \int_0^N (|\tilde{U}_r| + |U_r| + |\Pi_r|)^2 \mathbf{1}_{A^i \cap [0,k)}(\xi) \frac{dr d\xi}{2(N-1)} \\ &\leq C(t-s) \frac{k-1}{N-1}, \end{aligned} \tag{4.32}$$

where in the second step we have used Corollary 4.13 again, this time with $v = \tilde{U}_r$, $u = U_r$, $v_* = u_* = \Pi_r$ and $K = 0$. In the last step we have used the fact that f_r^L has uniformly bounded moments of order 2, that $\mathcal{L}(U_r) = \mathcal{L}(\tilde{U}_r) = f_r^L$, and that $\mathbb{E} \int_{j-1}^j |\Pi_r|^2 d\xi = \int_{\mathbb{R}^3} |u|^2 f_r^L(du)$ for all $j \neq i$, thanks to Lemma 4.7. Similarly for $J_{s,t}^3$: using Corollary 4.13 with $v = U_r$, $u = \tilde{U}_r$, $v_* = u_* = \Pi_r$ and $K = 0$, the bound (4.32) is also valid for $J_{s,t}^3$.

Thus, from (4.30), (4.31) and (4.32), we obtain that $\partial_t h_t \leq -(\Phi_0^L/2)h_t + Ck/N$ for almost all $t \geq 0$, and, since $h_0 = 0$, the conclusion follows from Gronwall's lemma (the dependence on L can be dropped since Φ_0^L is bounded away from 0 when $L \geq 1$ thanks to the lower bound in (4.7)). \square

For a probability measure μ on \mathbb{R}^d , call $\varepsilon_n(\mu) := \mathbb{E}\mathcal{W}_2^2(\mu, \bar{\mathbf{Z}})$, where $\mathbf{Z} = (Z^1, \dots, Z^n) \in (\mathbb{R}^d)^n$ is a vector of n independent and μ -distributed random variables on \mathbb{R}^d . The best general estimate available for $\varepsilon_n(\mu)$ is the following, whose proof can be found in [28]: for any $p > 4$, there exists a constant $C_p < \infty$ such that for every $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$\varepsilon_n(\mu) \leq \frac{C_p(\int |v|^p \mu(dv))^{2/p}}{n^{1/2}}. \quad (4.33)$$

The following lemma will allow us to work with $\mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t^L), (f_t^L)^{\otimes n})$ instead of $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L)$, but at the price of the extra term $\varepsilon_n(f_t^L)$:

Lemma 4.15. *Let $\mathbf{X} = (X^1, \dots, X^m) \in (\mathbb{R}^d)^m$ be an exchangeable random vector, and let $\mu \in \mathcal{P}(\mathbb{R}^d)$. Then, for any $n \leq m$,*

$$\begin{aligned} \frac{1}{2}\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{X}}, \mu) &\leq \frac{kn}{m} \left(\mathcal{W}_2^2(\mathcal{L}^n(\mathbf{X}), \mu^{\otimes n}) + \varepsilon_n(\mu) \right) \\ &\quad + \frac{\ell}{m} \left(\mathcal{W}_2^2(\mathcal{L}^\ell(\mathbf{X}), \mu^{\otimes \ell}) + \varepsilon_\ell(\mu) \right), \end{aligned}$$

where k and ℓ are the unique non-negative integers satisfying $m = kn + \ell$ and $\ell \leq n - 1$.

Proof. See the proof of Lemma 7 of [18]. \square

Corollary 4.16. *Assume (M) and take $L \geq 1$. Then, there exists a constant $C > 0$ such that for any $i \in \{1, \dots, N\}$ and for all $t \geq 0$,*

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L) \leq CN^{-1/3}.$$

Moreover, the same bound is valid with $\bar{\mathbf{U}}_t^L$ in place of $\bar{\mathbf{U}}_t^{i,L}$.

Proof. In the notation of Lemma 4.15, set $m = N - 1$, $\mathbf{X} = (U_t^{j,L})_{j \neq i}$ and $\mu = f_t^L$ and given $n \leq m$, simply bound $\frac{kn}{m} \leq 1$, $\frac{\ell}{m} \leq \frac{n}{N}$, $\mathcal{W}_2^2(\mathcal{L}^\ell(\mathbf{X}), \mu^{\otimes \ell}) \leq 4\mathcal{E}$ and $\varepsilon_\ell(\mu) \leq 4\mathcal{E}$, where $\mathcal{E} = \int |v|^2 f_t^L(dv) = \int |v|^2 f_0(dv)$. Using that result we get

$$\begin{aligned} \frac{1}{2}\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L) &\leq \mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t^L), (f_t^L)^{\otimes n}) + \varepsilon_n(f_t^L) + 8\mathcal{E}\frac{n}{N} \\ &\leq C\frac{n}{N} + C\frac{1}{n^{1/2}}, \end{aligned}$$

where we have used Lemma 4.14 together with (4.33) with $p = p_0 > 4$ and the uniformity of the p_0 -moments of f_t^L . Choosing $n = \lfloor N^{2/3} \rfloor$ yields the desired result. To obtain the same estimate with $\bar{\mathbf{U}}_t^L$ on the left hand side, use $m = N$ and $\mathbf{X} = \mathbf{U}_t^L$. \square

We need to make sure that the cutoff L can be removed in a satisfactory manner:

Lemma 4.17. *Assume (M). Then there exists a constant C such that for all $t \geq 0$ and all $L > 0$,*

$$\mathcal{W}_2^2(f_t^L, f_t) \leq CtL^{1-2/\nu}.$$

Proof. The obvious coupling between W and W^L (defined in (4.17) and (4.18)) works fine: for a fixed interval $A \subseteq \mathbb{R}$, consider functions $F_t(\xi)$ and $F_t^L(\xi)$, measurable in $(t, \xi) \in [0, \infty) \times A$, such that whenever ξ is a uniform random variable on A , the pair $(F_t(\xi), F_t^L(\xi))$ is a realization of the optimal coupling between f_t and f_t^L (this can be done thanks the measurability of the flows $t \mapsto (f_t, f_t^L)$: use a measurable selection theorem of optimal couplings, see for instance Corollary 5.22 of [62], together with a result of measurable realization of probability kernels, see for instance Lemma 3.22 of [43]). Using these functions F_t and F_t^L , now define W using a Poisson point measure \mathcal{M} as in (4.17) (well-posedness follows as in Proposition 4.1 of [30]), and couple it with W^L defined as the solution to

$$dW_t^L = \int_0^L \int_0^{2\pi} \int_A c(W_{t^-}^L, F_t^L(\xi), z, \varphi_{t^-}) \mathcal{M}(dt, dz, d\phi, d\xi),$$

where $\varphi_{t^-} = \varphi(W_{t^-} - F_t(\xi), W_{t^-}^L - F_t^L(\xi), \phi)$. For $h_t = \mathbb{E}|W_t^L - W_t|^2$, we obtain from this and (4.17):

$$\begin{aligned} \partial_t h_t &= \mathbb{E} \int_0^\infty \int_0^{2\pi} \int_A (|W_t^L + c_L(W_t^L, F_t^L(\xi), z, \varphi_t) - W_t - c(W_t, F_t(\xi), z, \phi)|^2 \\ &\quad - |W_t^L - W_t|^2) \frac{dt d\phi d\xi}{2|A|\pi} \\ &\leq \mathbb{E} \int_A \left(-\Phi_0^L |W_t^L - W_t|^2 + \Phi_0^L |F_t^L(\xi) - F_t(\xi)|^2 \right. \\ &\quad \left. + C(|W_t^L| + |W_t| + |F_t(\xi)|)^2 L^{1-2/\nu} \right) \frac{d\xi}{|A|}, \end{aligned}$$

where we have used Corollary 4.13 with $K = \infty$ and the roles of K and L exchanged. By construction, we have $\int_A |F_t^L(\xi) - F_t(\xi)|^2 d\xi / |A| = \mathcal{W}_2^2(f_t^L, f_t) \leq h_t$, and using the preservation of the second moment for f_t^L and f_t , the last inequality yields $\partial_t h_t \leq CL^{1-2/\nu}$. Since $h_0 = 0$, the result follows. \square

To prove Theorem 4.3, we will need to perform the following standarization procedure to force a vector to lie on the Boltzmann sphere \mathcal{S}^N (defined in (4.9)): given any $F^N \in \mathcal{P}^{\text{sym}}((\mathbb{R}^3)^N)$, set

$$[F^N] = \mathcal{L}([\mathbf{X}]), \tag{4.34}$$

where $\mathcal{L}(\mathbf{X}) = F^N$ and $[\mathbf{X}] = \mathbf{Y} = (Y^1, \dots, Y^N)$ is given by

$$Y^i = \frac{X^i - M}{S}, \quad \text{with} \quad M = \frac{1}{N} \sum_j X^j, \quad S^2 = \frac{1}{N} \sum_j |X^j - M|^2. \tag{4.35}$$

This, of course, ensures that $[F^N]$ is concentrated on \mathcal{S}^N . Note that $[F^N]$ is well defined when $S > 0$ a.s., but it is not hard to generalize the definition and the next result to the case in which $S = 0$ with positive probability. When the X^i 's have mean 0, variance 1, and non-positive correlation (for instance, when $F^N = f_0^{\otimes N}$ for some $f_0 \in \mathcal{P}(\mathbb{R}^3)$ satisfying (4.12)), we can provide an explicit rate of convergence for $\mathcal{W}_2^2([F^N], F^N)$:

Lemma 4.18. *Let $F^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$ be given. With the notation of (4.34)-(4.35), assume that $\mathbb{E}(X^1) = 0$, $\mathbb{E}|X^1|^2 = 1$, and $\mathbb{E}(X^1 \cdot X^2) \leq 0$. Then, $\mathbb{E}\frac{1}{N} \sum_i |Y^i - X^i|^2 \leq 2N^{-1}$, and consequently,*

$$\mathcal{W}_2^2([F^N], F^N) \leq 2N^{-1}.$$

Proof. Define $M_2 = \frac{1}{N} \sum_i |X^i|^2$, so $S^2 = M_2 - |M|^2$, and now a straightforward computation shows that $\frac{1}{N} \sum_i |Y^i - X^i|^2 = 1 + M_2 - 2S$, which gives

$$\mathcal{W}_2^2([F^N], F^N) \leq \mathbb{E}(1 + M_2 - 2S) = 2\mathbb{E}(1 - S) \leq 2\mathbb{E}(1 - S^2) = 2\mathbb{E}|M|^2.$$

Since $\mathbb{E}(X^1 \cdot X^2) \leq 0$, the latter is easily seen to be $\leq 2N^{-1}$. \square

Remark 4.19. If f_∞ denotes the Gaussian density with mean 0 and variance 1, that is, $f_\infty(v) = (2\pi\sigma^2)^{-3/2} e^{-|v|^2/(2\sigma^2)}$ for $\sigma^2 = 1/3$, then the measure $[f_\infty^{\otimes N}]$ is the uniform distribution on \mathcal{S}^N .

The following lemma gives a condition on f_0 ensuring that $[f_0^{\otimes N}]$ has finite fourth moment uniformly on N (condition (4.11)), which will be required in order to reduce the proof of the case (ii) of Theorem 4.3 to the case (i).

Lemma 4.20. *Assume f_0 satisfies (4.13). Then, for $F^N = f_0^{\otimes N}$ and with the notation of (4.34)-(4.35), we have $\mathbb{E}|Y^1|^4 < \infty$ uniformly on N .*

Proof. We have $\mathbb{E}|Y^1|^4 \leq (\mathbb{E}|X^1 - M|^{4a})^{1/a} (\mathbb{E}S^{-4b})^{1/b}$ for any $a^{-1} + b^{-1} = 1$, so it suffices to bound $\mathbb{E}S^{-2q}$ uniformly on N for some $q > 2$. Using the bound provided by (4.13), we have for any $r > 0$,

$$\mathbb{P}(S^2 < r) \leq C_0^N \mathbb{P}(\tilde{S}^2 < r),$$

where \tilde{S}^2 corresponds to S^2 defined in (4.35) when the X^i 's are i.i.d. with density f_∞ . Since $3N\tilde{S}^2$ has a chi-squared distribution with $3N - 3$ degrees of freedom, using Stirling's approximation, we can easily deduce that $\mathbb{P}(\tilde{S}^2 < r) \leq C(er)^{3N/2}$ for all $r > 0$, where $C > 0$ is some constant. Consequently, for $R > 0$ to be chosen, we have

$$\mathbb{E}S^{-2q} = \int_0^\infty \mathbb{P}(S^{-2q} > t) dt \leq R + \int_R^\infty \mathbb{P}(S^2 < t^{-1/q}) dt \leq R + CC_0^N e^{3N/2} \int_R^\infty t^{-3N/(2q)} dt.$$

The last integral equals $R^{-3N/(2q)+1}/(3N/(2q) - 1)$, and taking $R = (eC_0^{2/3})^q$ gives a bound that does not depend on N , as desired. \square

4.4 Proof of the main results

Proof of Theorem 4.2. We first treat the case $K < \infty$, and take $L \in [K, \infty)$. For some $i \in \{1, \dots, N\}$ fixed, we will estimate the quantity $h_t := \mathbb{E}|V_t^{i,K} - U_t^{i,L}|^2$. To shorten notation, call $V_r := V_r^{i,K}$, $V_r^i := V_r^{i(\xi),K}$, $U_r := U_r^{i,L}$ and $\Pi_r := \Pi_r^{i,L}(U_r^L, \xi)$. From (4.20) and (4.22) it follows that for every $0 \leq s < t$

$$\begin{aligned} h_t - h_s &= \mathbb{E} \int_{(s,t]} \int_0^\infty \int_0^{2\pi} \int_0^N \left(|V_{r^-} + c_K(V_{r^-}, V_{r^-}^i, z, \phi) - U_{r^-} \right. \\ &\quad \left. - c_L(U_{r^-}, \Pi_{r^-}, z, \varphi_{r^-}^i) \right|^2 - |V_{r^-} - U_{r^-}|^2 \Big) \mathcal{N}^i(dr, dz, d\phi, d\xi). \end{aligned}$$

Using the compensation formula, \mathcal{N}^i can be replaced by its intensity $drdzd\phi d\xi \mathbf{1}_{A^i}(\xi)/[2(N-1)\pi]$, where $A^i = [0, N) \setminus [i-1, i)$. Corollary 4.13 with $v = V_r$, $v_* = V_r^i$, $u = U_r$ and $u_* = \Pi_r$, yields

$$\begin{aligned} h_t - h_s &\leq \mathbb{E} \int_s^t \int_{A^i} \left[-\Phi_0^K |V_r - U_r|^2 + \Phi_0^K |V_r^i - \Pi_r|^2 \right. \\ &\quad \left. + C(|V_r| + |U_r| + |\Pi_r|)^2 K^{1-2/\nu} \right] \frac{drd\xi}{N-1}. \end{aligned}$$

Note that $|V_r^i - \Pi_r|^2 \leq |V_r^i - U_r^i|^2 + 2|V_r^i - U_r^i||U_r^i - \Pi_r| + |U_r^i - \Pi_r|^2$, where $U_r^i := U_r^{i(\xi),L}$, and that $\mathbb{E} \int_{A^i} |V_r^i - U_r^i|^2 d\xi / (N-1) = h_r$ by exchangeability. Also, thanks to Lemma 4.7, we know that $\int_{A^i} |U_r^i - \Pi_r|^2 d\xi / (N-1) = \mathcal{W}_2^2(\bar{U}_r^{i,L}, f_r^L)$. Calling $g_t := \mathbb{E} \mathcal{W}_2^2(\bar{U}_t^{i,L}, f_t^L)$, using the Cauchy–Schwarz inequality and the fact that both the particles and the cutoff nonlinear processes have uniformly bounded second moment, we obtain $\partial_t h_t \leq C(h_t^{1/2} g_t^{1/2} + g_t + K^{1-2/\nu})$ for almost every $t \geq 0$. Using a version of Gronwall’s lemma (see for instance Lemma 4.1.8 of [2]) yields

$$h_t \leq 2h_0 + Ct(1+t)N^{-1/3} + CtK^{1-2/\nu},$$

where we have used Corollary 4.16 to bound $g_t \leq CN^{-1/3}$ uniformly on t and L (the same holds for $\mathbb{E} \mathcal{W}_2^2(\bar{U}_t^L, f_t)$). From this, we obtain

$$\mathbb{E} \mathcal{W}_2^2(\bar{V}_t^K, f_t^L) \leq 2\mathbb{E} \mathcal{W}_2^2(\bar{V}_t^K, \bar{U}_t^L) + 2\mathbb{E} \mathcal{W}_2^2(\bar{U}_t^L, f_t^L) \leq 2h_t + 2CN^{-1/3},$$

and, since $\mathcal{W}_2(f_t^L, f_t) \rightarrow 0$ when $L \rightarrow \infty$ thanks to Lemma 4.17 (recall that \mathbf{V}^K does not depend on L), we deduce that

$$\mathbb{E} \mathcal{W}_2^2(\bar{V}_t^K, f_t) \leq \liminf_{L \rightarrow \infty} \mathbb{E} \mathcal{W}_2^2(\bar{V}_t^K, f_t^L) \leq Ch_0 + Ct(1+t)N^{-1/3} + CtK^{1-2/\nu}.$$

as desired (recall that $h_0 = \mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$). For the case $K = \infty$, the argument is similar: take $L < \infty$ and now use Corollary 4.13 with the roles of L and K exchanged, which yields

$$h_t \leq 2h_0 + Ct(1+t)N^{-1/3} + CtL^{1-2/\nu}, \quad (4.36)$$

after applying Gronwall’s lemma. Letting $L \rightarrow \infty$ as before, the result follows. \square

We conclude with Theorem 4.3:

Proof of Theorem 4.3. We first treat the case (i). From (4.36), letting $L \rightarrow \infty$ gives

$$\mathcal{W}_2^2(\mathcal{L}(V_t^i), f_t) \leq 2h_0 + C(1+t)^2 N^{-1/3}. \quad (4.37)$$

On the other hand, we have

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C[\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_\infty) + \mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_\infty, \bar{\mathbf{W}}_\infty) + \mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{W}}_\infty, f_\infty) + \mathcal{W}_2^2(f_t, f_\infty)]. \quad (4.38)$$

Here \mathbf{V}_∞ is a random vector uniformly distributed on the Boltzmann sphere \mathcal{S}^N (given by (4.9)) and optimally coupled to \mathbf{V}_t ; \mathbf{W}_∞ is a random vector with law $f_\infty^{\otimes N}$ and optimally coupled to \mathbf{V}_∞ ; and f_∞ is the stationary distribution for the nonlinear equation, namely the Gaussian density $f_\infty(v) = (2\pi\sigma^2)^{-3/2}e^{-|v|^2/(2\sigma^2)}$, with $\sigma^2 = 1/3$. Since G_0^N is concentrated on the set \mathcal{S}^N , we can use Proposition 1.7 of [55] to obtain for any $k \leq N$

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{V}_t), \mathcal{L}^k(\mathbf{V}_\infty)) \leq C_\delta[1 + (t - t_*)_+]^{-2\delta}, \quad (4.39)$$

for any $\delta \in (0, 1)$, where t_* can be assumed constant (independent of δ , N and t) since it depends only on the fourth moment of V_0^i . Note that

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{V}}_\infty) \leq \mathbb{E}\frac{1}{N}\sum_i |V_t^i - V_\infty^i|^2 = \mathcal{W}_2^2(\mathcal{L}^N(\mathbf{V}_t), \mathcal{L}^N(\mathbf{V}_\infty)),$$

so the first term of (4.38) is controlled by the right hand side of (4.39). Similarly, the second term of (4.38) satisfies $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_\infty, \bar{\mathbf{W}}_\infty) \leq \mathcal{W}_2^2(\mathcal{L}^N(\mathbf{V}_\infty), f_\infty^{\otimes N})$, which is bounded by $2N^{-1}$ thanks to Lemma 4.18 and Remark 4.19. The third term satisfies $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{W}}_\infty, f_\infty) \leq CN^{-1/2}$ thanks to (4.33). For the fourth term, we can pass (4.39) in the case $k = 1$ to the limit when $N \rightarrow \infty$, since, for $t \geq 0$ fixed, equation (4.37) yields $\mathcal{L}^1(\mathbf{V}_t) \rightarrow f_t$ (using that $h_0 \rightarrow 0$ with N thanks to hypothesis (4.10)), and Remark 4.19 yields $\mathcal{L}^1(\mathbf{V}_\infty) \rightarrow f_\infty$, both with respect to \mathcal{W}_2 . Putting all this into (4.38) yields

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C_\delta[1 + (t - t_*)_+]^{-2\delta} + CN^{-1/3}.$$

For $\eta \in (0, 1/6)$ to be chosen, set $\bar{t} = t_* + N^{\eta/\delta} - 1$, so the last inequality implies that $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C_\delta N^{-2\eta}$ for all $t > \bar{t}$, whereas Theorem 4.2 in the case $K = \infty$ gives $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq Ch_0 + CN^{-1/3+2\eta/\delta}$ for all $t \leq \bar{t}$. Setting $\eta = [6(1 + 1/\delta)]^{-1} < 1/12$ gives $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq Ch_0 + C_\delta N^{-2\eta}$ for all $t \geq 0$. Since δ can be chosen arbitrarily close to 1, the result follows in the case (i).

Now we prove the case (ii), for which we reduce to the previous case. Along with the particle system \mathbf{V} that starts with initial distribution $G_0^N = f_0^{\otimes N}$, consider the particle system $\tilde{\mathbf{V}}$ satisfying the same equation (4.19) that defines \mathbf{V} , but starting with initial condition $\tilde{\mathbf{V}}_0 = [\mathbf{V}_0]$ given by the standarization procedure (4.35). We have:

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq 2\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \tilde{\bar{\mathbf{V}}}_t) + 2\mathbb{E}\mathcal{W}_2^2(\tilde{\bar{\mathbf{V}}}_t, f_t),$$

where the second term can be controlled because the system $\tilde{\mathbf{V}}$ satisfies the hypotheses of the case (i): its initial distribution $\tilde{G}_0^N = [f_0^{\otimes N}]$ is concentrated on \mathcal{S}^N by construction, satisfies (4.10) thanks to Lemma 4.18, and also satisfies (4.11) thanks to Lemma (4.20). For the first term, the idea is to verify that: (a) $\tilde{\mathbf{V}}_t = [\mathbf{V}_t]$ for all $t \geq 0$, and (b) \mathbf{V}_t satisfies the hypotheses of Lemma 4.18; this will then yield $\mathbb{E}\mathcal{W}_2^2(\tilde{\mathbf{V}}_t, \tilde{\tilde{\mathbf{V}}}_t) \leq \mathbb{E}\frac{1}{N} \sum_i |V_t^i - \tilde{V}_t^i|^2 \leq 2N^{-1}$, obtaining the desired result.

We now verify (a), that is, we will show that

$$\tilde{V}_t^i = \frac{V_t^i - M_t}{S_t}, \text{ where } M_t = \frac{1}{N} \sum_j V_t^j \text{ and } S_t^2 = \frac{1}{N} \sum_j |V_t^j - M_t|^2. \quad (4.40)$$

But this is obvious from the fact that $M_t = M_0$, $S_t = S_0 > 0$ a.s. (since the system preserves momentum and energy), and that in the Maxwellian case the dynamics preserve the standardization. Specifically: given $v, v_*, m \in \mathbb{R}^3$ and $s > 0$, the homogeneity of \hat{i} and \hat{j} and the definitions of (4.14) imply that for $\tilde{v} = (v - m)/s$ and $\tilde{v}_* = (v_* - m)/s$ we have $c(\tilde{v}, \tilde{v}_*, z, \phi) = c(v, v_*, z, \phi)/s$ (note that $\gamma = 0$, so $G(z/|v - v_*|^\gamma) = G(z)$), and then

$$\tilde{v}' = \tilde{v} + c(\tilde{v}, \tilde{v}_*, z, \phi) = \frac{v + c(v, v_*, z, \phi) - m}{s} = \frac{v' - m}{s},$$

and the same for \tilde{v}'_* . Since the function c is the one involved in the definition of the particle system $\tilde{\mathbf{V}}$ (equation (4.19)), this proves (4.40), as desired.

Finally, we now verify (b): thanks to exchangeability and condition (4.12), we obtain that the particles satisfy $\mathbb{E}V_t^i = 0$, $\mathbb{E}|V_t^i|^2 = 1$. They are also uncorrelated:

$$\mathbb{E} \left| \sum_i V_t^i \right|^2 = N + N(N-1)\mathbb{E}(V_t^1 \cdot V_t^2),$$

and since the system preserves the quantity $\sum_i V_t^i$, the latter is also equal to $N + N(N-1)\mathbb{E}(V_0^1 \cdot V_0^2)$, which then yields $\mathbb{E}(V_t^1 \cdot V_t^2) = \mathbb{E}(V_0^1 \cdot V_0^2) = 0$. Thus, we can apply Lemma 4.18 to \mathbf{V}_t . The case (ii) follows. \square

Chapter 5

Conclusiones

Respecto al trabajo realizado en la presente tesis, se puede concluir:

Objetos de estudio. Se estudiaron algunas ecuaciones cinéticas (en particular la ecuación de Boltzmann espacialmente homogénea en el caso de las moléculas de Maxwell) y la respectiva propiedad de propagación de caos de los sistemas de partículas de salto asociados. Los sistemas estudiados son de variada generalidad, e incluyen a modelos de la mecánica estadística y de la Econofísica. Se realizó una extensa revisión bibliográfica sobre el tema, presentada en el Capítulo 1. Especial énfasis se le da al problema de propagación de caos cuantitativa, con tasas explícitas en la cantidad de partículas N y en el tiempo t .

Técnicas matemáticas. Para llevar a cabo tal estudio se utilizaron técnicas matemáticas conocidas en la literatura probabilista del tema, y al mismo tiempo se desarrollaron otras nuevas. En especial, se desarrolló una técnica de *coupling* que permite obtener tasas explícitas de propagación de caos en distancia de Wasserstein para una amplia gama de sistemas de partículas de tipo Bird (es decir, con interacciones efectivas entre las partículas), como los estudiados en los capítulos 3 y 4. Dicha técnica se explica en detalle en la Sección 3.2, y se aplica nuevamente en la Sección 4.2.

Resultados. Los principales resultados obtenidos en esta tesis son los siguientes:

El Teorema 2.1 entrega un resultado de propagación de caos trayectorial con convergencia débil para un sistema de partículas que toma valores en un espacio polaco. El interés de este resultado radica en su generalidad: las tasas de salto pueden depender de la medida empírica del sistema (además del estado de las partículas involucradas en la interacción), y se trabaja bajo hipótesis muy poco restrictivas.

El Teorema 3.1 provee un resultado de propagación de caos en distancia de Wasserstein para un sistema de partículas en \mathbb{R} que incluye a la ecuación de Boltzmann–

Kac (1.7) y a modelos de redistribución de riqueza, con tasas explícitas en N y t . Es de interés pues: (i) presenta e ilustra la forma de aplicar la nueva técnica de *coupling* explicada en la Sección 3.2, (ii) provee un resultado de caos cuantitativo para la ecuación Boltzmann–Kac extensible al caso sin *cutoff*, y (iii) provee el primer resultado de propagación de caos cuantitativo específico para ecuaciones de redistribución de riqueza.

El Teorema 4.2 entrega un resultado de propagación de caos cuantitativo en distancia de Wasserstein–2 para la ecuación de Boltzmann espacialmente homogénea en el caso de las moléculas de Maxwell (incluyendo al caso con *cutoff*), el cual utiliza en su demostración la técnica de *coupling* explicada en la Sección 3.2. Su interés radica en que extiende el resultado de Fournier y Mischler [30] al caso de un sistema de tipo Bird, el cual posee mayor relevancia desde el punto de vista físico que el sistema Nanbu.

El Teorema 4.3 provee un resultado de propagación de caos cuantitativo y uniforme en distancia de Wasserstein–2 al cuadrado de orden casi $N^{-1/6}$ para la ecuación de Boltzmann espacialmente homogénea en el caso de las moléculas de Maxwell. Este resultado es de suma importancia pues mejora sustancialmente la tasa uniforme obtenida por Mischler y Mouhot [47].

Publicaciones. Producto del trabajo realizado en esta tesis se han generado 2 artículos científicos: los resultados del Capítulo 3 han dado a lugar al trabajo [18], el cual ya ha sido aceptado para ser publicado en la revista *Annals of Applied Probability*; y los resultados del Capítulo 4 están prontos a ser sometidos para revisión y publicación.

Trabajo futuro. Los resultados obtenidos plantean varias interrogantes y potenciales extensiones que den lugar a posibles trabajos a realizar en el futuro. Algunas de ellas:

Obtener un resultado como el Teorema 3.1 en el caso de la ecuación de Boltzmann–Kac sin *cutoff* con condiciones débiles de integrabilidad. Específicamente: debería ser posible utilizar la técnica de la demostración del Teorema 3.1 en conjunto con un enfoque mediante problemas de martingala como el utilizado por Desvillettes, Graham y Méléard [20], para replicar el resultado de propagación de caos al caso en que, con la notación de dicho artículo, la función $\beta(\theta)$ satisface solamente $\int \theta^2 \beta(\theta) d\theta < \infty$.

Investigar la optimalidad de la tasa $N^{-1/3}$ que aparece en los teoremas 3.1 y 4.2. Esta tasa es más lenta que la óptima, de orden $N^{-1/2}$ (ver [28]), obtenida para los sistemas de tipo Nanbu (ver [27, 30]). Sin embargo, las razones de esta diferencia no están del todo claras, y sería interesante investigar si ella se debe al tipo de interacción en el sistema (Bird o Nanbu) o es una mera consecuencia de las técnicas utilizadas en la demostración.

Mejorar la tasa uniforme del Teorema 4.3, aumentando la tasa de orden casi $N^{-1/6}$ hasta (posiblemente) $N^{-1/3}$. Específicamente: utilizando una desigualdad desarrollada en [55] u otra especialmente diseñada para este propósito, debería ser posible

aprovechar el término negativo (que en esta tesis simplemente se descarta) que aparece debido al *coupling* óptimo de círculos llevado a cabo en el Capítulo 4, para así generar un término de contracción que permita obtener la propagación uniforme directamente y sin utilizar el resultado de [55].

Incluir a los potenciales duros y esferas duras. Usando las estimaciones de [30] en conjunto con lo desarrollado en el Capítulo 4, es posible extender el Teorema 4.2 a dichos casos, al menos a nivel informal. Para hacer riguroso el argumento, bastaría, en principio, con obtener un control uniforme de ciertos momentos exponenciales del proceso no-lineal truncado (necesario para usar las estimaciones de [30]), o bien con deducir un resultado de existencia para el sistema de procesos no-lineales sin truncar.

Aplicar las técnicas desarrolladas a ecuaciones similares en las cuales haya interacciones de tipo Bird. Por ejemplo, debería ser posible adaptar este *coupling* para estudiar algunos sistemas de partículas difusivas de tipo Bird que aproximan ciertos procesos no-lineales conducidos por ruido blanco, asociados con la ecuación de Landau, ver [25].

Bibliografia

- [1] ALEXANDRE, R. A review of Boltzmann equation with singular kernels. *Kinet. Relat. Models* 2, 4 (2009), 551–646.
- [2] AMBROSIO, L., GIGLI, N., AND SAVARÉ, G. *Gradient flows in metric spaces and in the space of probability measures*, second ed. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.
- [3] ARKERYD, L. On the Boltzmann equation. I. Existence. *Arch. Rational Mech. Anal.* 45 (1972), 1–16.
- [4] ARKERYD, L. On the Boltzmann equation. II. The full initial value problem. *Arch. Rational Mech. Anal.* 45 (1972), 17–34.
- [5] BASSETTI, F., LADELLI, L., AND MATTHES, D. Central limit theorem for a class of one-dimensional kinetic equations. *Probab. Theory Related Fields* 150, 1-2 (2011), 77–109.
- [6] BILLINGSLEY, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1999. A Wiley-Interscience Publication.
- [7] BOLTZMANN, L. *Lectures on Gas Theory (Reprint of the 1896–1898 Edition)*. Dover Publications, 2011.
- [8] CARLEN, E., CARVALHO, M., AND LOSS, M. Determination of the spectral gap for kac’s master equation and related stochastic evolution. *Acta Mathematica* 191, 1 (2003), 1–54.
- [9] CARLEN, E., CARVALHO, M. C., AND LOSS, M. Many-body aspects of approach to equilibrium. In *in Journées Équations aux dérivées partielles* (2001).
- [10] CARLEN, E. A., CARVALHO, M. C., AND LOSS, M. Spectral gap for the kac model with hard sphere collisions. *Journal of Functional Analysis* 266, 3 (2014), 1787 – 1832.

- [11] CARRAPATOSO, K. Propagation of chaos for the spatially homogeneous Landau equation for Maxwellian molecules. Preprint, arXiv:1212.3724, 2014.
- [12] CARRAPATOSO, K. Quantitative and qualitative Kac's chaos on the Boltzmann's sphere. Preprint, arXiv:1205.1241, 2014.
- [13] CERCIGNANI, C. On the Boltzmann equation for rigid spheres. *Transport Theory Statist. Phys.* 2, 3 (1972), 211–225.
- [14] CERCIGNANI, C. *The Boltzmann equation and its applications*, vol. 67 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1988.
- [15] CHAKRABORTI, A., AND CHAKRABARTI, B. Statistical mechanics of money: how saving propensity affects its distribution. *The European Physical Journal B - Condensed Matter and Complex Systems* 17, 1 (2000), 167–170.
- [16] CHATTERJEE, A., CHAKRABARTI, B. K., AND MANNA, S. S. Pareto law in a kinetic model of market with random saving propensity. *Phys. A* 335, 1-2 (2004), 155–163.
- [17] CORDIER, S., PARESCI, L., AND TOSCANI, G. On a kinetic model for a simple market economy. *J. Stat. Phys.* 120, 1-2 (2005), 253–277.
- [18] CORTEZ, R., AND FONTBONA, J. Quantitative propagation of chaos for generalized Kac particle systems. Preprint, arXiv:1406.2115v3. To appear in *Ann. Appl. Probab.*, 2015.
- [19] DESVILLETES, L. Boltzmann's kernel and the spatially homogeneous Boltzmann equation. *Riv. Mat. Univ. Parma (6)* 4* (2001), 1–22. Fluid dynamic processes with inelastic interactions at the molecular scale (Torino, 2000).
- [20] DESVILLETES, L., GRAHAM, C., AND MÉLÉARD, S. Probabilistic interpretation and numerical approximation of a Kac equation without cutoff. *Stochastic Process. Appl.* 84, 1 (1999), 115–135.
- [21] DESVILLETES, L., AND MOUHOT, C. Stability and uniqueness for the spatially homogeneous Boltzmann equation with long-range interactions. *Archive for Rational Mechanics and Analysis* 193, 2 (2009), 227–253.
- [22] DRAGULESCU, A., AND YAKOVENKO, V. Statistical mechanics of money. *The European Physical Journal B - Condensed Matter and Complex Systems* 17, 4 (2000), 723–729.
- [23] DRAGULESCU, A. A., AND YAKOVENKO, V. M. Statistical mechanics of money, income, and wealth: A short survey. *AIP Conference Proceedings* 661, 1 (2003), 180–183.
- [24] ETHIER, S. N., AND KURTZ, T. G. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

- [25] FONTBONA, J., GUÉRIN, H., AND MÉLÉARD, S. Measurability of optimal transportation and convergence rate for Landau type interacting particle systems. *Probab. Theory Related Fields* 143, 3-4 (2009), 329–351.
- [26] FOURNIER, N. Finiteness of entropy for the homogeneous boltzmann equation with measure initial condition. *Ann. Appl. Probab.* 25, 2 (04 2015), 860–897.
- [27] FOURNIER, N., AND GODINHO, D. Asymptotic of grazing collisions and particle approximation for the kac equation without cutoff. *Communications in Mathematical Physics* 316, 2 (2012), 307–344.
- [28] FOURNIER, N., AND GUILLIN, A. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields* 162, 3-4 (2015), 707–738.
- [29] FOURNIER, N., AND MÉLÉARD, S. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.* 14, 4 (2004), 1880–1919.
- [30] FOURNIER, N., AND MISCHLER, S. Rate of convergence of the Nanbu particle system for hard potentials. Preprint, arXiv:1302.5810. To appear in *Ann. Probab.*, 2015.
- [31] FOURNIER, N., AND MOUHOT, C. On the well-posedness of the spatially homogeneous Boltzmann equation with a moderate angular singularity. *Comm. Math. Phys.* 289, 3 (2009), 803–824.
- [32] FUJIWARA, Y., SOUMA, W., AOYAMA, H., KAIZOJI, T., AND AOKI, M. Growth and fluctuations of personal income. *Phys. A* 321, 3-4 (2003), 598–604.
- [33] FUNAKI, T. A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrsch. Verw. Gebiete* 67, 3 (1984), 331–348.
- [34] FÖLLMER, H. Random economies with many interacting agents. *Journal of Mathematical Economics* 1, 1 (1974), 51 – 62.
- [35] GALLAGHER, I., SAINT-RAYMOND, L., AND TEXIER, B. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.
- [36] GRAD, H. On the kinetic theory of rarefied gases. *Comm. Pure Appl. Math.* 2 (1949), 331–407.
- [37] GRAHAM, C., AND MÉLÉARD, S. Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Ann. Probab.* 25, 1 (1997), 115–132.
- [38] GRÜNBAUM, F. A. Propagation of chaos for the Boltzmann equation. *Arch. Rational Mech. Anal.* 42 (1971), 323–345.

- [39] GUÉRIN, H. Solving Landau equation for some soft potentials through a probabilistic approach. *Ann. Appl. Probab.* 13, 2 (2003), 515–539.
- [40] HAURAY, M., AND MISCHLER, S. On Kac’s chaos and related problems. *J. Funct. Anal.* 266, 10 (2014), 6055–6157.
- [41] JANVRESSE, E. Spectral gap for kac’s model of boltzmann equation. *Ann. Probab.* 29, 1 (02 2001), 288–304.
- [42] KAC, M. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III* (Berkeley and Los Angeles, 1956), University of California Press, pp. 171–197.
- [43] KALLENBERG, O. *Foundations of modern probability*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2002.
- [44] LANFORD, III, O. E. Time evolution of large classical systems. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*. Springer, Berlin, 1975, pp. 1–111. Lecture Notes in Phys., Vol. 38.
- [45] MATTHES, D., AND TOSCANI, G. On steady distributions of kinetic models of conservative economies. *J. Stat. Phys.* 130, 6 (2008), 1087–1117.
- [46] MCKEAN, JR., H. P. An exponential formula for solving Boltzmann’s equation for a Maxwellian gas. *J. Combinatorial Theory* 2 (1967), 358–382.
- [47] MISCHLER, S., AND MOUHOT, C. Kac’s program in kinetic theory. *Invent. Math.* 193, 1 (2013), 1–147.
- [48] MISCHLER, S., AND WENBERG, B. On the spatially homogeneous Boltzmann equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 16, 4 (1999), 467–501.
- [49] PARESCHI, L., AND TOSCANI, G. Self-similarity and power-like tails in nonconservative kinetic models. *J. Stat. Phys.* 124, 2-4 (2006), 747–779.
- [50] PARESCHI, L., AND TOSCANI, G. *Interacting multiagent systems : kinetic equations and Monte Carlo methods*. Oxford University Press, Oxford New York, 2014.
- [51] PARESCHI, L., AND TOSCANI, G. Wealth distribution and collective knowledge: a Boltzmann approach. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 372, 2028 (2014), 20130396, 15.
- [52] PULVIRENTI, A., AND TOSCANI, G. Asymptotic properties of the inelastic Kac model. *J. Statist. Phys.* 114, 5-6 (2004), 1453–1480.
- [53] RACHEV, S. T., AND RÜSCHENDORF, L. *Mass transportation problems. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

- [54] REMENIK, D. Limit theorems for individual-based models in economics and finance. *Stochastic Process. Appl.* 119, 8 (2009), 2401–2435.
- [55] ROUSSET, M. A N -uniform quantitative Tanaka’s theorem for the conservative Kac’s N -particle system with Maxwell molecules. Preprint, arXiv:1407.1965, 2014.
- [56] SZNITMAN, A.-S. Équations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete* 66, 4 (1984), 559–592.
- [57] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.* Springer, Berlin, 1991, pp. 165–251.
- [58] TANAKA, H. On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)* (1978), Wiley, New York-Chichester-Brisbane, pp. 409–425.
- [59] TANAKA, H. Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete* 46, 1 (1978/79), 67–105.
- [60] TOSCANI, G., AND VILLANI, C. Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *J. Statist. Phys.* 94, 3-4 (1999), 619–637.
- [61] VILLANI, C. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I.* North-Holland, Amsterdam, 2002, pp. 71–305.
- [62] VILLANI, C. *Optimal transport, old and new*, vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.