

A GENERAL NONLINEAR VERSION OF ROTH'S THEOREM ON THE REAL LINE

XIANG LI, DUNYAN YAN, HAIXIA YU AND XINGSONG ZHANG*

(Communicated by L. Liu)

Abstract. Let $N > 1$ be a real number and $\varepsilon > 0$ be given. In this paper, we will prove that, for a measurable subset S of $[0, N]$ with positive density ε , there must be patterns of the form $(x, x+t, x+\gamma(t))$ such that

$$x, x+t, x+\gamma(t) \in S,$$

where γ is convex and has some curvature constraints, $t > \delta(\varepsilon, \gamma)\gamma^{-1}(N)$ and $\delta(\varepsilon, \gamma)$ is a positive constant depending only on ε and γ , γ^{-1} is the inverse function of γ . Our result extends Bourgain's result [2] to the general curve γ . We use Bourgain's energy pigeonholing argument and Li's σ -uniformity argument.

1. Introduction

A classical question in pure mathematics is to ask what conditions need to be imposed on a subset of the integers to guarantee that it contains an arithmetic progression. In 1953, Roth's remarkable article [34] on the existence of triples in arithmetic progression in subsets of integers with positive upper density tells us that if a subset of integers is large enough, then there probably exists certain additive structure in it. Furthermore, it implies that the following original version of Roth's theorem: Let $S \subset \mathbb{Z}$ be a subset of the positive integers of positive upper density, that is

$$\limsup_{N \rightarrow \infty} \frac{\#\{S \cap (-N, N)\}}{2N} > 0.$$

Then, the set S must contain non-trivial arithmetic progression of length 3, i.e., existing patterns of the form $(x, x+t, x+2t)$ such that

$$x, x+t, x+2t \in S,$$

where $x \in \mathbb{Z}$ and $t > 0$. Here, Roth used exponential sums and the Hardy-Littlewood circle methods from Fourier analysis. His idea is to assume for contradiction that one

Mathematics subject classification (2020): Primary 42B20; Secondary 05D10.

Keywords and phrases: Roth's theorem, convex curve, stationary phase method, σ -uniformity.

Dunyan Yan is supported by the National Natural Science Foundation of China (Grant No. 11871452), Haixia Yu is supported by the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2020A1515110241) and the Fundamental Research Funds for the Central Universities (Grant No. 20lgpy144).

* Corresponding author.

had a set of integers of positive density which contained no arithmetic progressions of length three, and then to use Fourier analysis to construct a new set of integers with even higher density which still contains no such progressions. Eventually one would be forced to construct a set of density over 1, which yields a contradiction.

REMARK 1.1. If we replace the length 3 in Roth's theorem by a positive integer k with $k \geq 3$, we then obtain Szemerédi's theorem. It was proved by Szemerédi [37]. Indeed, Szemerédi [36] first considered the case $k = 4$ and extended to the general case [37]. The theorem implies Van der Waerden's theorem and also gives an answer to the long-standing conjecture of Erdős and Turán [14]. From [34], we knew that Roth's proof based on the methods of Fourier analysis, but Szemerédi [36, 37] used a method from combinatorial mathematics. Furthermore, the theorem can also be proved by the techniques from ergodic theory (see, Furstenberg [17], Furstenberg and Katznelson [18]) and additive number theory (see, Gowers [19, 20]). On the other hand, Szemerédi's theorem has been extended to the polynomial Szemerédi's theorem [1] and multidimensional Szemerédi's theorem [18, Theorem B]. Later, the theorem was generalized to the version of Szemerédi's theorem relative to the primes [21, 22] and polynomial Szemerédi's theorem for the primes [39], multidimensional Szemerédi's theorem for the primes [38] and the references contained therein.

The study on Roth's theorem inspired the emergence of new mathematical ideas and methods. Bourgain [3] give a new proof of Roth's theorem. Later, by the spirit of [3], Bourgain [2] studied the nonlinear version of Roth's theorem in \mathbb{R} , which was closely related to the bilinear Hilbert transform along the curve t^2 . Bourgain's result can be stated as follows.

THEOREM 1.2. ([2, Theorem 1]) *For any given $\varepsilon > 0$, let S be a measurable subset of $[0, N]$ with $|S| > \varepsilon N$. Then we can find patterns of the form $(x, x+t, x+t^2)$, such that*

$$x, x+t, x+t^2 \in S,$$

where $t > \delta(\varepsilon)N^{\frac{1}{2}}$ and $\delta(\varepsilon)$ is a positive constant depending only on ε .

Theorem 1.2 is also suitable for t^d with $d \in \mathbb{N}$ and $d > 2$. Recently, Durcik, Guo and Roos [13] extend this result to the curve $(t, \mathcal{P}(t))_{t \in \mathbb{R}}$, where $\mathcal{P}(t)$ is a monic polynomial of degree $d > 1$ without constant term. As Bourgain [2], the paper [13] is also closely related to the bilinear Hilbert transform along the monic polynomial $(t, \mathcal{P}(t))_{t \in \mathbb{R}}$. There also are other versions of Roth's theorem; see, for example, Roth's theorem on \mathbb{R}^d [11], polynomial Roth's theorem on sets of fractional dimensions [16], nonlinear Roth's theorem in finite fields [4], polynomial Roth's theorem in finite fields [12] and the references contained therein.

In this paper, we further extend Theorem 1.2 to a wider class of curves. We now state our main theorem.

THEOREM 1.3. Let $\gamma(t) \in C^3((0, \infty))$ with $\gamma(0) = \gamma'(0) = \lim_{t \rightarrow 0^+} \frac{\gamma'}{\gamma''}(t) = 0$, and convex on $(0, \infty)$, satisfying

$$C_0 \leq \left(\frac{\gamma'}{\gamma''} \right)'(t) \leq C_1, \tag{1.1}$$

where $t \in (0, \infty)$ and $C_1 > C_0 > 0$ are constants. Moreover, let $N > 1$ be real number, $\varepsilon > 0$ be given and S be a measurable subset of $[0, N]$ with $|S| > \varepsilon N$. Then we can find patterns of the form $(x, x+t, x+\gamma(t))$, such that

$$x, x+t, x+\gamma(t) \in S,$$

where $t > \delta(\varepsilon, \gamma)\gamma^{-1}(N)$, $\delta(\varepsilon, \gamma)$ is a positive constant depending only on ε and γ , γ^{-1} is the inverse function of γ .

REMARK 1.4. It is easy to check that $\gamma(t) := t^d$ with $d \in \mathbb{N}$ and $d \geq 2$ satisfies the conditions of Theorem 1.3. Obviously, Theorem 1.3 covers Bourgain's Theorem 1.2. The following cases are some other curves which satisfy the conditions of Theorem 1.3:

- (i) for any $t \in [0, \infty)$, $\gamma_1(t) := t^\alpha$, $\alpha \in (1, \infty)$;
- (ii) for any $t \in [0, \infty)$, $\gamma_2(t) := t^\alpha \log(1+t)$, $\alpha \in (1, \infty)$;
- (iii) for any $t \in [0, \infty)$, $\gamma_3(t) := \int_0^t \tau^\alpha \log(1+\tau) d\tau$, $\alpha \in (0, \infty)$;
- (iv) for any $t \in [0, \infty)$, $\gamma_4(t) := \int_0^t \tau^\alpha \arctan \tau d\tau$, $\alpha \in (0, \infty)$;
- (v) for any $t \in [0, \infty)$ and $K \in \mathbb{N}$, $\gamma_5(t) := \sum_{i=1}^K t^{\alpha_i}$ under $\alpha_i \in (1, \infty)$ for all $i = 1, 2, \dots, K$.

One of the motivations of our study on Theorem 1.3 arises from the $L^p(\mathbb{R}) \times L^q(\mathbb{R}) \rightarrow L^r(\mathbb{R})$ boundedness of the bilinear Hilbert transform $H_\gamma(f, g)$ along the curve γ defined as

$$H_\gamma(f, g)(x) := \text{p.v.} \int_{-\infty}^{\infty} f(x-t)f(x-\gamma(t)) \frac{dt}{t},$$

where p, q, r satisfy $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, $p > 1$, $q > 1$. Here and hereafter, p.v. $\int_{-\infty}^{\infty}$ denotes the principal-value integral. This originated from Calderón [5] in order to study the Cauchy transform along Lipschitz curves. If $\gamma(t) := -t$, the operator is the standard bilinear Hilbert transform. Lacey and Thiele [26, 27] obtained the $L^p(\mathbb{R}) \times L^q(\mathbb{R}) \rightarrow L^r(\mathbb{R})$ boundedness with $r > \frac{2}{3}$. If $\gamma(t) := t^d$ or $\gamma(t) := \mathcal{P}(t)$, a polynomial of degree d without linear term and constant term, $d \in \mathbb{N}$ and $d > 1$, for the boundedness of $H_\gamma(f, g)$, we refer the reader to Li [29] and Li and Xiao [30]. For more general curve γ , Lie [31] introduced a class \mathcal{NF}^C of curves and obtained the $L^2(\mathbb{R}) \times L^2(\mathbb{R}) \rightarrow L^1(\mathbb{R})$ boundedness of $H_\gamma(f, g)$ for $\gamma \in \mathcal{NF}^C$. Later, it was extended to the $L^p(\mathbb{R}) \times L^q(\mathbb{R}) \rightarrow L^r(\mathbb{R})$ boundedness with $r \geq 1$ in Lie [32]. More recently, Guo and Xiao [23]

obtained the $L^2(\mathbb{R}) \times L^2(\mathbb{R}) \rightarrow L^1(\mathbb{R})$ boundedness of $H_\gamma(f, g)$, where $\gamma \in \mathbf{F}(-1, 1)$, the definition of the class $\mathbf{F}(-1, 1)$ of curves can be found in p. 970 in [23].

As what we have stated before, we know that the proofs of the Bourgain’s Theorem 1.2 and the polynomial Roth’s theorem [13] are closely related to the boundedness of $H_\gamma(f, g)$ along the homogeneous curve $\gamma(t) := t^2$ and the monic polynomial $\gamma(t) := \mathcal{P}(t)$, respectively. Based on the development of the $L^p(\mathbb{R}) \times L^q(\mathbb{R}) \rightarrow L^r(\mathbb{R})$ boundedness of $H_\gamma(f, g)$, whose boundedness has been obtained for more general curves, it is natural to consider the general nonlinear version of Roth’s Theorem, i.e. Theorem 1.3. Our conditions are easier to check than \mathcal{NF}^C in Lie [31, 32] and $\mathbf{F}(-1, 1)$ in Guo and Xiao [23].

Another motivation of our study is offered by the Hilbert transform $H_\gamma f$ along the curve γ defined as

$$H_\gamma f(x_1, x_2) := \text{p. v.} \int_{-\infty}^{\infty} f(x_1 - t, x_2 - \gamma(t)) \frac{dt}{t},$$

which was initiated by Fabes and Rivière [15] and Jones [25] in order to understand the behavior of the constant coefficient parabolic differential operators. Later, $H_\gamma f$ was extended to cover more general classes of curves; see, for example, [6, 7, 8, 10, 33]. $H_\gamma(f, g)$ is closely associated to $H_\gamma f$, since they have the same multiplier. Indeed, we can rewrite $H_\gamma(f, g)(x)$ and $H_\gamma f(x_1, x_2)$ as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\xi) \hat{g}(\eta) \left(\text{p. v.} \int_{-\infty}^{\infty} e^{-i\xi t - i\eta \gamma(t)} \frac{dt}{t} \right) e^{ix\xi} e^{ix\eta} d\xi d\eta,$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{f}(\xi, \eta) \left(\text{p. v.} \int_{-\infty}^{\infty} e^{-i\xi t - i\eta \gamma(t)} \frac{dt}{t} \right) e^{ix_1 \xi} e^{ix_2 \eta} d\xi d\eta.$$

Therefore, we may find many similarities between $H_\gamma(f, g)$ and $H_\gamma f$.

The last but not least, our motivation for generalizing of nonlinear version of Roth’s theorem comes from itself. Theorem 1.3 is a natural generalization, and it is also the inevitable development of this theorem. Compared with the usual linear setting, i.e. $\gamma(t) := 2t$, and Bourgain’s Theorem 1.2, i.e. $\gamma(t) := t^d$ with $d \in \mathbb{N}$ and $d \geq 2$, our result leads to different phenomena.

The rest of this paper is organized as follows. In Section 2, we first obtain some properties for γ and collect a lemma from [2]. Furthermore, we reduce our Theorem 1.3 to the key Lemma 2.5, this can be realized by Proposition 2.4. Section 3 is devoted to proof of Lemma 2.5. We first split our operator as the sum of $A_1(f, g)$ and $A_2(f, g)$ by the critical points of the phase function. The former part $A_1(f, g)$ far from the critical points, whose estimate can be obtained by Van der Corput’s lemma. The second part $A_2(f, g)$ closes to the critical points, which is the most difficult part. In order to obtain its estimate, we used the TT^* argument, Hörmander’s theorem [24, Theorem 1.1], the stationary phase method and σ -uniformity argument [29, Theorem 7.1].

Throughout this paper, we use C to denote a positive constant that is independent of the main parameters involved but whose value may vary from line to line. Moreover,

we use $C(\varepsilon, \gamma, \dots)$ or $\delta(\varepsilon, \gamma, \dots)$ to denote a positive constant depending on the indicated parameters $\varepsilon, \gamma, \dots$. The positive constants with subscripts, such as C_1 and C_2 , are the same in different occurrences. For two real functions f and g , we use $f \lesssim g$ or $g \gtrsim f$ to denote $f \leq Cg$ and, if $f \lesssim g \lesssim f$, we write $f \approx g$. We use $\mathcal{S}(\mathbb{R})$ to denote Schwartz class on \mathbb{R} . \mathbb{R} means the set of real numbers, \mathbb{C} means the set of complex numbers, \mathbb{Z} means the set of integers, $\mathbb{Z}_- := \mathbb{Z} \setminus \mathbb{N}$ with $\mathbb{N} := \{0, 1, 2, \dots\}$. For any $A \in \mathbb{R}$, $[A]$ is the unique integer such that $0 \leq A - [A] < 1$, and $\lceil A \rceil$ is the unique integer such that $0 \leq \lceil A \rceil - A < 1$. For any set E , we use $\mathbf{1}_E$ to denote its characteristic function, $\#E$ denotes the cardinality of it. \hat{f} denotes the Fourier transform of f , \check{f} is the inverse Fourier transform of f .

2. Preliminaries

2.1. Curve γ

We start by introducing some simple properties of curve γ which we need in the course of proof. We conclude these properties as the following lemma

LEMMA 2.1. *Let γ as defined in Theorem 1.2, then for any $t \in (0, \infty)$ we have*

- (i) *there exist positive constants C_2 and C_3 , such that $C_2 \leq \frac{t\gamma''}{\gamma'}(t) \leq C_3$;*
- (ii) *there exist positive constants C_4 and C_5 , such that $C_4 \leq \frac{t\gamma'(t)}{\gamma(t)} \leq C_5$*
- (iii) *there exist positive constants $C_7 > C_6 > 1$, such that $C_6 \leq \frac{\gamma'(2t)}{\gamma(t)} \leq C_7$;*
- (iv) *there exist the same constants C_6 and C_7 as above, such that $2C_6 \leq \frac{\gamma(2t)}{\gamma(t)} \leq 2C_7$;*
- (v) *there exist positive constants C_8 and C_9 , such that $-\frac{C_8}{t^2} \leq (\frac{\gamma''}{\gamma'})'(t) \leq -\frac{C_9}{t^2}$.*

Proof. For (i), we denote

$$F(t) = C'_1 t - \frac{\gamma'}{\gamma''}(t).$$

After deriving the variable t , we have $F'(t) = C'_1 - (\frac{\gamma'}{\gamma''})'(t)$, if we set $C'_1 \geq C_1$, then $F'(t) \geq 0$ for any $t \in (0, \infty)$. Since $F(0) = -\frac{\gamma'}{\gamma''}(0) = 0$, we obtain that $F(t) \geq 0$ for any $t \in (0, \infty)$, which means

$$\frac{t\gamma''(t)}{\gamma'(t)} \geq \frac{1}{C_1}.$$

Then we denote

$$G(t) = \frac{\gamma'}{\gamma''}(t) - C'_0 t,$$

if we set $C'_0 \leq C_0$, in the same way, we get

$$\frac{t\gamma''(t)}{\gamma'(t)} \leq \frac{1}{C_0}.$$

For (ii), Since γ' is increasing on $(0, \infty)$ and $\gamma(0) = \gamma'(0) = 0$, by the mean value theorem again, for any $t \in (0, \infty)$, there exists $\tau_2 \in (0, t)$ such that

$$\frac{t\gamma'(t)}{\gamma(t)} = \frac{t\gamma'(t) - 0\gamma'(0)}{\gamma(t) - \gamma(0)} = \frac{\gamma'(\tau_2) + \tau_2\gamma''(\tau_2)}{\gamma'(\tau_2)}.$$

Thus, by (i), we can see that

$$1 + \frac{1}{C_1} \leq \frac{t\gamma'(t)}{\gamma(t)} \leq 1 + \frac{1}{C_0}. \tag{2.1}$$

For (iii), we denote

$$\tau(t) = \ln \gamma'(t).$$

Then we can obtain that

$$\tau(2t) - \tau(t) = \tau'(\theta t) \cdot t,$$

where $\theta \in (1, 2)$ and $\tau'(t) = \frac{\gamma''(t)}{\gamma'(t)}$, combining with (i) we have

$$e^{\frac{1}{2C_1}} \leq \frac{\gamma'(2t)}{\gamma'(t)} \leq e^{\frac{1}{C_0}}.$$

For (iv), repeating the process in (ii), we know for any $t \in (0, \infty)$, there exists $\tau_1 \in (0, t)$ such that

$$\frac{\gamma(2t)}{\gamma(t)} = \frac{\gamma(2t) - \gamma(0)}{\gamma(t) - \gamma(0)} = \frac{2\gamma'(2\tau_1)}{\gamma'(\tau_1)}.$$

This, combined with (iii), we have

$$2e^{\frac{1}{2C_1}} \leq \frac{\gamma(2t)}{\gamma(t)} \leq 2e^{\frac{1}{C_0}}. \tag{2.2}$$

For (v), (1.1) implies

$$C_0 \leq \frac{(\gamma'')^2 - \gamma'\gamma'''}{(\gamma'')^2}(t) \leq C_1,$$

from (i) we can replace γ'' in the denominator by γ' , then we have

$$-\frac{1}{C_0} \frac{1}{t^2} \leq \left(\frac{\gamma''}{\gamma'}\right)'(t) \leq -\frac{1}{C_1} \frac{1}{t^2}. \quad \square$$

2.2. Reduction of Theorem 1.3 to Lemma 2.5

In this Section, we will reduce the proof of Theorem 1.3 to the key Lemma 2.5. We first collect the following Lemma 2.2 from [2].

LEMMA 2.2. ([2, Lemma 6]) *For a nonnegative function f supported on $[0, 1]$ and $t_1, t_2 > 0$, we have*

$$\int_0^1 f(x) (P_{t_1} * f)(x) (P_{t_2} * f)(x) dx \geq \tilde{C} \left(\int_0^1 f(x) dx \right)^3,$$

for some positive constant \tilde{C} depending only on P , where $\{P_t\}_{t>0}$ is the standard Poisson semi-group and $P_t(\cdot) := \frac{1}{t}P(\frac{\cdot}{t})$.

It is easy to see that Theorem 1.3 is a consequence of the following Proposition 2.3.

PROPOSITION 2.3. *Let $N > 1$ be a real number, $\varepsilon > 0$ be given, and γ satisfy all of conditions in Theorem 1.3. Suppose that f is a function on \mathbb{R} with $0 \leq f \leq 1$ and $\int_0^N f(x) dx \geq \varepsilon N$. Then*

$$\int_0^N \int_0^{\gamma^{-1}(N)} f(x)f(x+t)f(x+\gamma(t)) dt dx > \delta(\varepsilon, \gamma)N\gamma^{-1}(N), \tag{2.3}$$

where $\delta(\varepsilon, \gamma)$ is a positive constant depending only on ε and γ .

By changing of variable $x \rightarrow Nx$, $t \rightarrow \gamma^{-1}(N)t$ and letting $\phi(x) := f(Nx)$, (2.3) is equivalent to

$$\int_0^1 \int_0^1 \phi(x)\phi\left(x + \frac{\gamma^{-1}(N)t}{N}\right)\phi\left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N}\right) dt dx > \delta(\varepsilon, \gamma). \tag{2.4}$$

Therefore, it suffices to prove the following Proposition 2.4.

PROPOSITION 2.4. *Let $N > 1$ be a real number, $\varepsilon > 0$ be given and γ satisfy all of conditions in Theorem 1.3. Suppose that f be a function supported on $[0, 1]$ with $0 \leq f \leq 1$ and $\int_0^1 f(x) dx \geq \varepsilon$. Then there exists a positive constant $\delta(\varepsilon, \gamma)$ depending only on ε and γ such that*

$$\int_0^1 \int_0^1 f(x)f\left(x + \frac{\gamma^{-1}(N)t}{N}\right)f\left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N}\right) dt dx > \delta(\varepsilon, \gamma). \tag{2.5}$$

To prove Proposition 2.4, we use the forthcoming Lemma 2.5, which will be proved in Section 3. Let τ be a nonnegative smooth bump function supported on $\{t \in \mathbb{R} : \frac{1}{2} \leq t \leq 2\}$ with $\hat{\tau}(0) = 1$. We denote $\tau_t(\cdot) := \frac{1}{t}\tau(\frac{\cdot}{t})$ for $t > 0$.

LEMMA 2.5. *Let $N > 1$ be a real number, $l \in \mathbb{N}$, γ satisfies all of conditions in Theorem 1.3 and f be the same as in Proposition 2.4. Then there exists a positive constant β such that, for all $g \in \mathcal{S}(\mathbb{R})$ with $\text{supp } \hat{g} \subset [2^m, 2^{m+1}]$ and $m \geq 0$, the following inequality*

$$\begin{aligned} & \left\| \int_{\mathbb{R}} f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} \\ & \leq C(2C_7)^l 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \end{aligned} \tag{2.6}$$

holds for some positive constant C depending only on γ .

Proof of Proposition 2.4. Let

$$I := \int_0^1 \int_0^1 f(x)f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) f \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) dt dx.$$

For $1 \leq l' \leq l \leq l''$, it is easy to see that

$$\begin{aligned} 2^l I & \geq \int_0^1 \int_0^1 f(x)f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) f \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt dx \\ & = I_1 + I_2 + I_3, \end{aligned} \tag{2.7}$$

where

$$\begin{aligned} I_1 & := \int_0^1 \int_0^1 f(x)f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) (P_{2^{-l'}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt dx, \\ I_2 & := \int_0^1 \int_0^1 f(x)f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) (P_{2^{-l''}} * f - P_{2^{-l'}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt dx, \\ I_3 & := \int_0^1 \int_0^1 f(x)f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) (f - P_{2^{-l''}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt dx. \end{aligned}$$

We first estimate I_3 . Let $g := f - P_{2^{-l''}} * f$. We take inhomogeneous Littlewood-Paley decomposition on \hat{g} and obtain that

$$\hat{g}(\xi) = \sum_{k \in \mathbb{N}} (\hat{g}\chi_k)(\xi) + \sum_{k \in \mathbb{Z}_-} (\hat{g}\chi_k)(\xi),$$

where $\chi_k(\xi) := \mathbf{1}_{\{\xi \in \mathbb{R}: 2^k \leq \xi < 2^{k+1}\}}(\xi)$. Letting $\widehat{g}_k := \hat{g}\chi_k$, from the definition of Poisson kernel, we have¹

$$\widehat{g}_k(\xi) = (f - P_{2^{-l''}} * f)\widehat{\gamma}(\xi)\chi_k(\xi) = \widehat{f}(\xi) \left(1 - e^{-2\pi 2^{-l''}|\xi|} \right) \chi_k(\xi). \tag{2.8}$$

¹Here and hereafter, we also denote $(f)\widehat{\gamma}(\cdot)$ means the Fourier transform of f .

Furthermore, as $0 \leq f \leq 1$ and f is supported on $[0, 1]$, by the triangle inequality, we have that

$$|I_3| \leq \left\| \sum_{k \in \mathbb{Z}} \int_0^1 f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g_k \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} \\ \leq I_{31} + I_{32} + I_{33},$$

where

$$I_{31} := \left\| \int_0^1 f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) \left(\sum_{k \in \mathbb{Z}_-} g_k \right) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} ; \\ I_{32} := \sum_{0 \leq k \leq l''-1} \left\| \int_0^1 f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g_k \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} ; \\ I_{33} := \sum_{k \geq l''} \left\| \int_0^1 f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g_k \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} .$$

For I_{33} , by Lemma 2.5 and the fact that $\|g_k\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}$, we have

$$I_{33} \lesssim (2C_7)^l \sum_{k \geq l''} 2^{-\beta k} \|f\|_{L^2(\mathbb{R})} \|g_k\|_{L^2(\mathbb{R})} \tag{2.9} \\ \lesssim (2C_7)^l \sum_{k \geq l''} 2^{-\beta k} \|f\|_{L^2(\mathbb{R})}^2 \\ \lesssim 2^{(\log_2 2C_7)l - \beta l''} \|f\|_{L^2(\mathbb{R})}^2 .$$

For I_{32} , by Lemma 2.5, it can be bounded by

$$(2C_7)^l \sum_{0 \leq k \leq l''-1} 2^{-\beta k} \|f\|_{L^2(\mathbb{R})} \|g_k\|_{L^2(\mathbb{R})} .$$

Using Plancherel's theorem and (2.8), we have that

$$\|g_k\|_{L^2(\mathbb{R})} = \|\widehat{g}_k\|_{L^2(\mathbb{R})} = \left\| \widehat{f}(\cdot) \left(1 - e^{-2\pi 2^{-l''}|\cdot|} \right) \chi_k(\cdot) \right\|_{L^2(\mathbb{R})} \\ \lesssim 2^{-l''+k} \|\widehat{f}\chi_k\|_{L^2(\mathbb{R})} \lesssim 2^{-l''+k} \|f\|_{L^2(\mathbb{R})} .$$

We may assume that $0 < \beta < 1$. Hence

$$I_{32} \lesssim (2C_7)^l \sum_{0 \leq k \leq l''-1} 2^{-\beta k} 2^{-l''+k} \|f\|_{L^2(\mathbb{R})}^2 \tag{2.10} \\ \lesssim (2C_7)^l 2^{-l''+(1-\beta)l''} \|f\|_{L^2(\mathbb{R})}^2 \\ \lesssim 2^{(\log_2 2C_7)l - \beta l''} \|f\|_{L^2(\mathbb{R})}^2 .$$

For I_{31} , applying Hölder’s inequality and Plancherel’s theorem, we have

$$\begin{aligned}
 I_{31} &\leq \|f\|_{L^2(\mathbb{R})} \left\| \sum_{k \in \mathbb{Z}_-} g_k \right\|_{L^2(\mathbb{R})} \\
 &= \|f\|_{L^2(\mathbb{R})} \left\| \hat{f}(\cdot) \left(1 - e^{-2\pi 2^{-l''}|\cdot|} \right) \mathbf{1}_{\{\xi \in \mathbb{R}: 0 < \xi < 1\}}(\cdot) \right\|_{L^2(\mathbb{R})} \\
 &\lesssim 2^{-l''} \|f\|_{L^2(\mathbb{R})}^2 \\
 &\lesssim 2^{(\log_2 2C_7)l - \beta l''} \|f\|_{L^2(\mathbb{R})}^2.
 \end{aligned} \tag{2.11}$$

The last inequality follows from the fact that $1 \leq l \leq l''$ and $0 < \beta < 1$.

Combining (2.9), (2.10) with (2.11), for any $\varepsilon > 0$, yields that

$$|I_3| \lesssim 2^{(\log_2 2C_7)l - \beta l''} \|f\|_{L^2(\mathbb{R})}^2 \leq 2^{-100} \tilde{C} \varepsilon^3 \tag{2.12}$$

holds if we take l'' large enough with respect to l , where \tilde{C} can be found in Lemma 2.2.

We then estimate I_2 . By Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}
 |I_2| &\leq \int_0^1 \left\| f(x) f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) \right\|_{L^2(\mathbb{R}_x)} \\
 &\quad \times \left\| (P_{2^{-l''}} * f - P_{2^{-l'}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \right\|_{L^2(\mathbb{R}_x)} \tau_{2^{-l}}(t) dt \\
 &\leq \|P_{2^{-l''}} * f - P_{2^{-l'}} * f\|_{L^2(\mathbb{R})}.
 \end{aligned} \tag{2.13}$$

For I_1 , we construct a new term \tilde{I}_1 . Our aim is to replace the estimate of I_1 by \tilde{I}_1 .

Let

$$\tilde{I}_1 := \int_0^1 \int_0^1 f(x) f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) (P_{2^{-l'}} * f)(x) \tau_{2^{-l}}(t) dt dx.$$

We claim that there is only a tiny difference between \tilde{I}_1 and I_1 . The difference between \tilde{I}_1 and I_1 can be written as

$$\begin{aligned}
 &\int_0^1 \int_0^1 f(x) f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) \\
 &\quad \left[(P_{2^{-l'}} * f)(x) - (P_{2^{-l'}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \right] \tau_{2^{-l}}(t) dt dx.
 \end{aligned}$$

By the mean value theorem, it implies that

$$\begin{aligned}
 \left| (P_{2^{-l'}} * f)(x) - (P_{2^{-l'}} * f) \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \right| &\leq 2^{l'} \|(P^{l'})_{2^{-l'}} * f\|_{L^\infty(\mathbb{R})} \frac{\gamma(\gamma^{-1}(N)t)}{N} \\
 &\lesssim 2^{l'} \frac{\gamma(\gamma^{-1}(N)2^{1-l})}{N}.
 \end{aligned}$$

The last inequality follows from Young's inequality and the monotonicity of γ on $(0, \infty)$. For fixed $N > 1$, from (iv) in Lemma 2.1, it implies that $\gamma(\gamma^{-1}(N)2^{1-l}) \leq \frac{N}{2^{l-1}}$. Thus for any $\varepsilon > 0$, we can choose l large enough with respect to l' such that

$$2^{l'} \frac{\gamma(\gamma^{-1}(N)2^{1-l})}{N} \lesssim 2^{l'-l} \leq 2^{-100} \tilde{C}\varepsilon^3.$$

This easily leads to

$$|\tilde{I}_1 - I_1| \leq 2^{-100} \tilde{C}\varepsilon^3. \tag{2.14}$$

We now turn to the estimate of \tilde{I}_1 . Let $\tilde{\tau}(\cdot) := \tau(\cdot)$, then \tilde{I}_1 can be written as

$$\tilde{I}_1 = - \int_0^1 f(x) (P_{2^{-l'}} * f)(x) \left(\tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right)(x) dx = \tilde{I}_{11} + \tilde{I}_{12},$$

where

$$\begin{aligned} \tilde{I}_{11} &:= - \int_0^1 f(x) (P_{2^{-l'}} * f)(x) \left[\left(\tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right)(x) - \left(P_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right)(x) \right] dx; \\ \tilde{I}_{12} &:= - \int_0^1 f(x) (P_{2^{-l'}} * f)(x) \left(P_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right)(x) dx. \end{aligned}$$

It follows from Lemma 2.2 that

$$|\tilde{I}_{12}| \geq \tilde{C}\varepsilon^3. \tag{2.15}$$

For \tilde{I}_{11} , by Hölder's inequality, it can be bounded by

$$\left\| \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right\|_{L^2(\mathbb{R})}.$$

Furthermore, by the triangle inequality and Young's convolution inequality, we have that

$$\left\| \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{lN}}} * f \right\|_{L^2(\mathbb{R})} \leq I_a + I_b + I_c,$$

where

$$\begin{aligned} I_a &:= \left\| \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * P_{\frac{\gamma^{-1}(N)}{2^{l''N}}} * f - \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f \right\|_{L^2(\mathbb{R})}; \\ I_b &:= \left\| \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} - \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{lN}}} * P_{\frac{\gamma^{-1}(N)}{2^{l''N}}} \right\|_{L^1(\mathbb{R})}; \end{aligned}$$

$$I_c := \left\| \frac{P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} - \tilde{\tau}_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * P_{\frac{\gamma^{-1}(N)}{2^{l'N}}}}{L^1(\mathbb{R})} \right\|.$$

We apply Young’s convolution inequality again for I_a to obtain that

$$I_a \leq \left\| \frac{P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f}{L^2(\mathbb{R})} \right\|.$$

By rescaling in I_b and I_c , we get

$$I_b = \left\| \tilde{\tau}_{2^{-l}} - \tilde{\tau}_{2^{-l}} * P_{2^{-l'}} \right\|_{L^1(\mathbb{R})}$$

and

$$I_c = \left\| P_{2^{-l'}} - \tilde{\tau}_{2^{-l}} * P_{2^{-l'}} \right\|_{L^1(\mathbb{R})}.$$

By the mean value theorem, if l'' is chosen large enough with respect to l , and l large large enough with respect to l' , then I_b and I_c are bounded from above by $2^{-100}\tilde{C}\varepsilon^3$. Therefore,

$$|\tilde{I}_{11}| \leq \left\| \frac{P_{\frac{\gamma^{-1}(N)}{2^{l''N}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f}{L^2(\mathbb{R})} \right\| + 2^{-99}\tilde{C}\varepsilon^3. \tag{2.16}$$

Putting these estimates $|I_3|$, $|I_2|$, $|\tilde{I}_1 - I_1|$, $|\tilde{I}_{12}|$ and $|\tilde{I}_{11}|$ from (2.12), (2.13), (2.14), (2.15) and (2.16) together, and noticing $\tilde{I}_1 = \tilde{I}_{11} + \tilde{I}_{12}$, from (2.7), we obtain

$$\tilde{C}\varepsilon^3 \leq 2^l I + \left\| P_{2^{-l'}} * f - P_{2^{-l'}} * f \right\|_{L^2(\mathbb{R})} + \left\| \frac{P_{\frac{\gamma^{-1}(N)}{2^{l''N}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f}{L^2(\mathbb{R})} \right\| + 2^{-90}\tilde{C}\varepsilon^3.$$

Therefore, by the pigeonhole argument, we see that either

$$I > 2^{-l-10}\tilde{C}\varepsilon^3$$

or

$$\left\| P_{2^{-l'}} * f - P_{2^{-l'}} * f \right\|_{L^2(\mathbb{R})} + \left\| \frac{P_{\frac{\gamma^{-1}(N)}{2^{l''N}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{l'N}}} * f}{L^2(\mathbb{R})} \right\| > 2^{-10}\tilde{C}\varepsilon^3.$$

Starting with $l_0 = 1$, the previous considerations enable us to construct a sequence $l_0 < l_1 < \dots < l_k < \dots$ satisfying $l_{k+1} \leq C(\varepsilon, \gamma)l_k$ from some positive constant $C(\varepsilon, \gamma)$ depending only on ε and γ such that for each k either

$$I > 2^{-l_{k+1}-10}\tilde{C}\varepsilon^3, \tag{2.17}$$

or

$$\|P_{2^{-l_k}} * f - P_{2^{-l_{k+1}}} * f\|_{L^2(\mathbb{R})} + \left\| P_{\frac{\gamma^{-1}(N)}{2^{kN}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{k+1}N}} * f \right\|_{L^2(\mathbb{R})} > 2^{-10} \tilde{C} \varepsilon^3. \tag{2.18}$$

Notice that, for any positive constant K , we have the following estimate:

$$\sum_{k=0}^K \left(\|P_{2^{-l_k}} * f - P_{2^{-l_{k+1}}} * f\|_{L^2(\mathbb{R})}^2 + \left\| P_{\frac{\gamma^{-1}(N)}{2^{kN}}} * f - P_{\frac{\gamma^{-1}(N)}{2^{k+1}N}} * f \right\|_{L^2(\mathbb{R})}^2 \right) \leq \tilde{C} \|f\|_{L^2(\mathbb{R})}^2 \leq c_0, \tag{2.19}$$

where c_0 is a positive constant independent of K and f . Taking K large enough to satisfy $K \frac{(2^{-10} \tilde{C} \varepsilon^3)^2}{2} > c_0$, if (2.18) holds for all k satisfying $0 < k \leq K$, then the sum in (2.19) leads to $K \frac{(2^{-10} \tilde{C} \varepsilon^3)^2}{2} \leq c_0$, which yields a contradiction. Thus, there exists k satisfying $0 < k \leq K$ such that (2.17) established. Note that the sequence $1 = l_0 < l_1 < \dots < l_k < \dots$ satisfying $l_{k+1} \leq C(\varepsilon, \gamma) l_k$, then l_{k+1} in (2.17) can be bounded from above by $C(\varepsilon, \gamma)^K$, where K satisfies $K \frac{(2^{-10} \tilde{C} \varepsilon^3)^2}{2} > c_0$. Therefore, we obtain a uniform lower estimate on I by letting

$$\delta(\varepsilon, \gamma) := 2^{-C(\varepsilon, \gamma) K_0 - 10} \tilde{C} \varepsilon^3,$$

with $K_0 := \frac{4c_0}{(2^{-10} \tilde{C} \varepsilon^3)^2}$. \square

3. Proof of the key Lemma 2.5

We first denote the bilinear operator $A(f, g)$ as

$$A(f, g)(x) := \int_{\mathbb{R}} f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt.$$

In this Section, we want to obtain that there exists a positive constant β such that

$$\|A(f, g)\|_{L^1([0,1])} \lesssim (2C_7)^l 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}. \tag{3.1}$$

Since the support of \hat{g} is on a given dyadic interval, we may also take a dyadic decomposition on the frequency support of f . Therefore, we define the frequency projection operator \mathbb{P}_k as

$$\widehat{\mathbb{P}_k f}(\xi) := \hat{f}(\xi) \chi_k(\xi),$$

where $\chi_k(\xi) := \mathbf{1}_{\{\xi \in \mathbb{R}: 2^k \leq |\xi| < 2^{k+1}\}}$ and $k \in \mathbb{Z}$. Thus, $A(f, g)(x)$ can be written as

$$\sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} (\mathbb{P}_k f) \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt.$$

Furthermore, by the Fourier inversion formula, the expression above can be written as

$$\sum_{k \in \mathbb{Z}} \iint_{\mathbb{R}^2} \widehat{\mathbb{P}_k f}(\xi) \hat{g}(\eta) e^{i\xi x} e^{i\eta x} m_l(\xi, \eta) d\xi d\eta,$$

where

$$m_l(\xi, \eta) := \int_{\mathbb{R}} e^{i2^{-l} \frac{\gamma^{-1}(N)}{N} t \xi} e^{i \frac{\gamma(\gamma^{-1}(N)2^{-l}t)}{N} \eta} \tau(t) dt.$$

Since the support of τ is $[\frac{1}{2}, 2]$ and the support of \hat{g} is $[2^m, 2^{m+1}]$ with $m \geq 0$, the support of $\widehat{\mathbb{P}_k f}$ varies with k , we can take $|\xi| \approx 2^k$, $\eta \approx 2^m$ and $t \approx 1$. This, combining with (iv) in Lemma 2.1, we may further expect that the main contribution to $A(f, g)$ comes from such k 's which satisfy

$$2^{-l} \frac{\gamma^{-1}(N)}{N} 2^k = \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N} 2^m,$$

which implies that

$$k = m + M_0$$

with

$$M_0 := \log_2 \left(\frac{\gamma(\gamma^{-1}(N)2^{-l})}{\gamma^{-1}(N)2^{-l}} \right). \tag{3.2}$$

Here we will always regard M_0 as an integer, if not so, we can take $M'_0 := \lfloor M_0 \rfloor$ to replace M_0 by M'_0 . Then, we can rewrite $A(f, g)$ as the following form.

$$\begin{aligned} A(f, g)(x) &= \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}} (\mathbb{P}_{k+m+M_0} f) \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \\ &= \sum_{k \in \mathbb{Z}} \iint_{\mathbb{R}^2} (\mathbb{P}_{k+m+M_0} f)(\xi) \hat{g}(\eta) e^{i\xi x} e^{i\eta x} m_l(\xi, \eta) d\xi d\eta. \end{aligned}$$

According to the stationary phase method, we can divide $A(f, g)$ into the following two parts by the range of $|k|$. Write

$$A(f, g)(x) = A_1(f, g)(x) + A_2(f, g)(x),$$

where

$$A_1(f, g)(x) := \sum_{k \in \mathbb{Z}, |k| \geq D} \iint_{\mathbb{R}^2} (\mathbb{P}_{k+m+M_0} f)(\xi) \hat{g}(\eta) e^{i\xi x} e^{i\eta x} m_l(\xi, \eta) d\xi d\eta,$$

and

$$A_2(f, g)(x) := \sum_{k \in \mathbb{Z}, |k| < D} \iint_{\mathbb{R}^2} (\mathbb{P}_{k+m+M_0} f)(\xi) \hat{g}(\eta) e^{i\xi x} e^{i\eta x} m_l(\xi, \eta) d\xi d\eta,$$

where D is a large enough positive constant depending on curve γ such that $|2^k - 1| \gtrsim 1$ holds for all $k \in \mathbb{Z}$ and $|k| \geq D$.

Next we will analyze $A_1(f, g)$ and $A_2(f, g)$, respectively. We now define the phase function in m_l as $\phi(t)$ and obtain its derivative $\phi'(t)$ and second derivative $\phi''(t)$ as follows:

$$\begin{aligned} \phi(t) &:= 2^{-l} \frac{\gamma^{-1}(N)}{N} t \xi + \frac{\gamma(\gamma^{-1}(N)2^{-l}t)}{N} \eta; \\ \phi'(t) &= 2^{-l} \frac{\gamma^{-1}(N)}{N} \xi + \gamma^{-1}(N) 2^{-l} \frac{\gamma'(\gamma^{-1}(N)2^{-l}t)}{N} \eta; \\ \phi''(t) &= \left(\gamma^{-1}(N)2^{-l}\right)^2 \frac{\gamma''(\gamma^{-1}(N)2^{-l}t)}{N} \eta. \end{aligned}$$

For $A_1(f, g)$, we use duality to analyze its $L^1(\mathbb{R})$ -norm. Taking $h \in L^\infty([0, 1])$, we consider the following expression

$$\sum_{k \in \mathbb{Z}, |k| \geq D} \int_0^1 \int_{\mathbb{R}} (\mathbb{P}_{k+m+M_0} f) \left(x + \frac{\gamma^{-1}(N)t}{N}\right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N}\right) h(x) \tau_{2^{-l}}(t) dt dx.$$

By Fourier inversion formula, the above expression can be rewritten as

$$\sum_{k \in \mathbb{Z}, |k| \geq D} \iint_{\mathbb{R}^2} \hat{f}(\xi) \hat{g}(\eta) \chi_{k+m+M_0}(\xi) \hat{h}(-\xi - \eta) m_l(\xi, \eta) d\xi d\eta. \tag{3.3}$$

We now analyze $2^{-l} \frac{\gamma^{-1}(N)}{N} \xi$ and $\gamma^{-1}(N) 2^{-l} \frac{\gamma'(\gamma^{-1}(N)2^{-l}t)}{N} \eta$ in ϕ' , respectively. Indeed, for the former part, noticing $|\xi| \approx 2^{k+m+M_0}$ and the definition of M_0 in (3.2), we have

$$\left| 2^{-l} \frac{\gamma^{-1}(N)}{N} \xi \right| \approx 2^{-l} \frac{\gamma^{-1}(N)}{N} 2^{k+m+M_0} \approx 2^{k+m} \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N}. \tag{3.4}$$

For the second part, noticing $\eta \approx 2^m$ and $t \approx 1$, from (ii) and (iv) in Lemma 2.1, we have that

$$\left| \gamma^{-1}(N) 2^{-l} \frac{\gamma'(\gamma^{-1}(N)2^{-l}t)}{N} \eta \right| \approx \gamma^{-1}(N) 2^{-l} \frac{\gamma(\gamma^{-1}(N)2^{-l}t)}{N \gamma^{-1}(N)2^{-l}t} 2^m \approx 2^m \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N}. \tag{3.5}$$

Furthermore, from (iv) in Lemma 2.1, it is easy to see that

$$(2C_7)^{-l} \leq \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N} \leq 2^{-l}. \tag{3.6}$$

Therefore, we have

$$|\phi'(t)| \gtrsim 2^m \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N} |2^k - 1|.$$

Note that $|2^k - 1| \gtrsim 1$ holds for all $k \in \mathbb{Z}$ and $|k| \geq D$. This, combining with (3.6), leads to

$$|\phi'(t)| \gtrsim 2^m (2C_7)^{-l}. \tag{3.7}$$

On the other hand, from convexity of γ , we have that $\gamma''(t) > 0$ holds for all $t \in (0, \infty)$, which further implies that $\phi''(t) > 0$ holds for all $t \in (0, \infty)$. Therefore, ϕ' is monotonic. By Van der Corput's lemma, for example, (see [35], p. 332, Proposition 2), we have that

$$|m_l(\xi, \eta)| \lesssim (2C_7)^l 2^{-m}.$$

Applying Cauchy-Schwarz inequality to (3.3) and Plancherel's theorem, we can control (3.3) by

$$\begin{aligned} & (2C_7)^l 2^{-m} \left(\iint_{\mathbb{R}^2} |\hat{g}(\eta)|^2 \sum_{k \in \mathbb{Z}, |k| \geq D} |\hat{f}(\xi) \chi_{k+m+M_0}(\xi)|^2 d\xi d\eta \right)^{\frac{1}{2}} \\ & \times \left(\iint_{\mathbb{R}^2} |\hat{h}(-\xi - \eta)|^2 \chi_m(\eta) d\xi d\eta \right)^{\frac{1}{2}} \\ & \lesssim (2C_7)^l 2^{-\frac{m}{2}} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}. \end{aligned}$$

Letting $h(x) := \mathbf{1}_{(0,1)}(x)$, we have

$$\|A_1(f, g)\|_{L^1([0,1])} \lesssim (2C_7)^l 2^{-\frac{1}{2}m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \tag{3.8}$$

with $\beta := \frac{1}{2}$ as desired.

For $A_2(f, g)$, where $|k|$ is small. Without loss of generality, we may take $k = 0$ for the purpose of simplifying notation and write

$$A_2(f, g)(x) = \iint_{\mathbb{R}^2} (\mathbb{P}_{m+M_0} \tilde{f})(\xi) \hat{g}(\eta) e^{i\xi x} e^{i\eta x} m_l(\xi, \eta) d\xi d\eta$$

with abuse of notations. Therefore, in the rest part of the section, our aim is to prove the following inequality:

$$\begin{aligned} & \left\| \int_{\mathbb{R}} f \left(x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,1])} \\ & \lesssim (2C_7)^l 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}, \end{aligned}$$

where β is a positive constant, $\text{supp } \hat{f} \subset \{\xi \in \mathbb{R} : 2^{m+M_0} \leq |\xi| \leq 2^{m+M_0+1}\}$ and $\text{supp } \hat{g} \subset [2^m, 2^{m+1}]$.

For the convenience of computation, we will change variables several times. Changing variables $x \rightarrow 2^{-M_0-m}x$, then the inequality above is changed into the following form:

$$\begin{aligned} & \left\| \int_{\mathbb{R}} f \left(2^{-M_0-m}x + \frac{\gamma^{-1}(N)t}{N} \right) g \left(2^{-M_0-m}x + \frac{\gamma(\gamma^{-1}(N)t)}{N} \right) \tau_{2^{-l}}(t) dt \right\|_{L^1([0,2^{M_0+m}])} \\ & \lesssim (2C_7)^l 2^{M_0+m} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}. \end{aligned} \tag{3.9}$$

Let

$$f(2^{-M_0-m}x) \rightarrow f(x) \quad \text{and} \quad g(2^{-m}x) \rightarrow g(x).$$

Then (3.9) becomes

$$\left\| \int_{\mathbb{R}} f(x + \lambda t) g(2^{-M_0}x + \lambda Q(t)) \tau(t) dt \right\|_{L^1([0, 2^{M_0+m}])} \lesssim (2C_7)^l 2^{\frac{M_0}{2}} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}, \tag{3.10}$$

where

$$\lambda := 2^{M_0+m-l} \frac{\gamma^{-1}(N)}{N} \quad \text{and} \quad Q(t) := 2^{-M_0} \frac{\gamma(\gamma^{-1}(N)2^{-l}t)}{\gamma^{-1}(N)2^{-l}} \tag{3.11}$$

and

$$\text{supp } \hat{f} \subset [-2, -1] \cup [1, 2] \quad \text{and} \quad \text{supp } \hat{g} \subset [1, 2].$$

Therefore, it suffices to obtain (3.10). This, combining with (3.8), further implies Lemma 2.5. Indeed, the properties of λ and Q will play a significant role in proving (3.10). Now we conclude their properties as the following Lemma 3.1.

LEMMA 3.1. *If λ and Q are defined as in (3.11), then we have*

$$(2C_7)^{-l} 2^m \lesssim \lambda \lesssim 2^{-l} 2^m \quad \text{and} \quad \|Q\|_{C^2([\frac{1}{2}, 2])} \approx 1.$$

Proof of Lemma 3.1. From the definition of M_0 in (3.2), we have

$$\lambda = 2^m \frac{\gamma(\gamma^{-1}(N)2^{-l})}{\gamma^{-1}(N)2^{-l}} 2^{-l} \frac{\gamma^{-1}(N)}{N} = 2^m \frac{\gamma(\gamma^{-1}(N)2^{-l})}{N}.$$

This, combining with (3.6), implies the first estimate about λ .

For $\|Q\|_{C^2([\frac{1}{2}, 2])}$, we need to estimate Q , Q' and Q'' , respectively, where

$$Q'(t) = 2^{-M_0} \gamma'(\gamma^{-1}(N)2^{-l}t)$$

and

$$Q''(t) = 2^{-M_0} \gamma^{-1}(N) 2^{-l} \gamma''(\gamma^{-1}(N)2^{-l}t).$$

For Q , applying (iv) in Lemma 2.1 and the definition of M_0 in (3.2) again, we have

$$Q(t) \approx Q(1) = \frac{\gamma^{-1}(N)2^{-l}}{\gamma(\gamma^{-1}(N)2^{-l})} \frac{\gamma(\gamma^{-1}(N)2^{-l})}{\gamma^{-1}(N)2^{-l}} = 1. \tag{3.12}$$

For Q' , from Theorem (iii) in Lemma 2.1 and γ' is increasing and $\gamma' \geq 0$ on $(0, \infty)$, which implies

$$1 \leq \frac{\gamma'(2t)}{\gamma'(t)} \leq C_7$$

for all $t \in (0, \infty)$. This, combined with (3.2) and (ii) in Lemma 2.1, we have

$$Q'(t) \approx Q'(1) = \frac{\gamma^{-1}(N)2^{-l}\gamma'(\gamma^{-1}(N)2^{-l})}{\gamma(\gamma^{-1}(N)2^{-l})} \approx 1. \tag{3.13}$$

For Q'' , by (i) in Lemma 2.1, we have

$$Q''(t) \approx 2^{-M_0}\gamma^{-1}(N)2^{-l}\frac{\gamma'(\gamma^{-1}(N)2^{-l}t)}{\gamma^{-1}(N)2^{-l}t}$$

for all $t \in [\frac{1}{2}, 2]$. Noticing $t \approx 1$ and (3.13), we obtain

$$Q''(t) \approx Q'(t) \approx 1. \tag{3.14}$$

From (3.12), (3.13) and (3.14), we have

$$\|Q\|_{C^2([\frac{1}{2}, 2])} \approx 1$$

which is our desired result. This completes the proof of Lemma 3.1. \square

We now turn to (3.10). By using dual function $h \in L^\infty([0, 2^{m+M_0}])$ again, it is enough to verify the following inequality:

$$\begin{aligned} & \left| \iint_{\mathbb{R}^2} f(x + \lambda t)g(2^{-M_0}x + \lambda Q(t))h(x)\tau(t) dt dx \right| \\ & \lesssim (2C_7)^l 2^{\frac{M_0}{2}} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}. \end{aligned} \tag{3.15}$$

As h is supported on $[0, 2^{m+M_0}]$, by Hölder’s inequality, it is easy to see that

$$\|h\|_{L^2(\mathbb{R})} \leq 2^{\frac{m+M_0}{2}} \|h\|_{L^\infty(\mathbb{R})}.$$

Therefore, it suffices to prove the following new inequality:

$$\begin{aligned} & \left| \iint_{\mathbb{R}^2} f(x + \lambda t)g(2^{-M_0}x + \lambda Q(t))h(x)\tau(t) dt dx \right| \\ & \lesssim (2C_7)^l 2^{-\frac{m}{2}} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^2(\mathbb{R})}. \end{aligned} \tag{3.16}$$

By applying Fourier inversion formula to f and g , the left hand side of (3.16) can be written as

$$\left| \iint_{\mathbb{R}^2} \hat{f}(\xi)\hat{g}(\eta)\hat{h}(-\xi - 2^{-M_0}\eta) \left(\int_{\mathbb{R}} e^{i\lambda(t\xi + Q(t)\eta)} \tau(t) dt \right) d\xi d\eta \right|.$$

We denote the phase function of the integral in t by

$$\Phi_{\xi, \eta}(t) := t\xi + \eta Q(t). \tag{3.17}$$

Here, we want to obtain (3.16) by using the stationary phase method, which forces us to analyze the critical points of phase function. If $\Phi_{\xi, \eta}$ has no critical points, then we can use usual skills such as integration by parts and Cauchy-Schwarz inequality

to obtain the desired bound. Otherwise, we consider the case that $\Phi_{\xi,\eta}$ have critical points. Since for fixed ξ and η ,

$$\Phi'_{\xi,\eta}(t) = \xi + \eta Q'(t).$$

From the monotonically increasing property of Q' , we have $\Phi'_{\xi,\eta}$ is monotonically increasing, so there are only one critical point in internal $[\frac{1}{2}, 2]$. Therefore, in the remainder part of this section, we will always assume the following equation

$$\Phi'_{\xi,\eta}(t) = 0$$

has a unique solution

$$t_c := t_c(\xi, \eta) \in \left[\frac{1}{2}, 2\right].$$

Then we can denote the corresponding dual phase function as

$$\Psi(\xi, \eta) := \Phi_{\xi,\eta}(t_c) = t_c \xi + \eta Q(t_c).$$

We now consider the following two cases:

$$\begin{cases} \text{Case1: } |M_0| \leq (1 - \kappa)m; \\ \text{Case2: } |M_0| > (1 - \kappa)m, \end{cases}$$

respectively. Here κ is small, positive universal constant that is to be determined later.

3.1. Case 1: $|M_0| \leq (1 - \kappa)m$

We follow the approach of [29] and use TT^* method to obtain (3.16) under this Case 1. By Proposition 3 in chapter VIII in [35], we have

$$\int_{\mathbb{R}} e^{i\lambda\Phi_{\xi,\eta}(t)} \tau(t) dt = \lambda^{-\frac{1}{2}} e^{i\lambda\Psi(\xi,\eta)} a(\xi, \eta) + R_{\xi,\eta}(\lambda), \tag{3.18}$$

where $a(\xi, \eta)$ is a smooth compactly supported function and the remainder term $R_{\xi,\eta}$ satisfies

$$|R_{\xi,\eta}(\lambda)| \lesssim \lambda^{-1}.$$

Furthermore, noticing that $|a| = |Q''(t_c)|^{-\frac{1}{2}} |\tau(t_c)|$, by Lemma 3.1, we know that $|a| \lesssim 1$. From Lemma 3.1, we have $\lambda^{-1} \lesssim (2C_7)^l 2^{-m}$, which further implies

$$\begin{aligned} & \left| \iint_{\mathbb{R}^2} \hat{f}(\xi) \hat{g}(\eta) \hat{h}(-\xi - 2^{-M_0}\eta) R_{\xi,\eta}(\lambda) d\xi d\eta \right| \\ & \lesssim (2C_7)^l 2^{-m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^2(\mathbb{R})}. \end{aligned} \tag{3.19}$$

This is the desired estimate in (3.16) for the remainder term $R_{\xi, \eta}$. Therefore, our main task is to prove that

$$\left| \iint_{\mathbb{R}^2} \hat{f}(\xi) \hat{g}(\eta) e^{i\lambda \Psi(\xi, \eta)} a(\xi, \eta) \hat{h}(-\xi - 2^{-M_0} \eta) d\xi d\eta \right| \lesssim (2C_7)^{\frac{1}{2}} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^2(\mathbb{R})}. \tag{3.20}$$

By changing variable $\xi \rightarrow \xi - 2^{-M_0} \eta$ and applying Cauchy-Schwarz inequality to separate the function h , our main task is to verify the following expression

$$\left\| \int_{\mathbb{R}} \hat{f}(\xi - 2^{-M_0} \eta) \hat{g}(\eta) a(\xi - 2^{-M_0} \eta, \eta) e^{i\lambda \Psi(\xi - 2^{-M_0} \eta, \eta)} d\eta \right\|_{L^2(\mathbb{R}_\xi)} \lesssim (2C_7)^{\frac{1}{2}} 2^{-\beta m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}. \tag{3.21}$$

Expanding the square of the $L^2(\mathbb{R})$ -norm on the left side gives

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} \hat{f}(\xi - 2^{-M_0} \eta) \hat{g}(\eta) a(\xi - 2^{-M_0} \eta, \eta) e^{i\lambda \Psi(\xi - 2^{-M_0} \eta, \eta)} d\eta \right) \times \left(\int_{\mathbb{R}} \overline{\hat{f}(\xi - 2^{-M_0} \eta') \hat{g}(\eta') a(\xi - 2^{-M_0} \eta', \eta')} e^{-i\lambda \Psi(\xi - 2^{-M_0} \eta', \eta')} d\eta' \right) d\xi.$$

By changing variables

$$\eta' \rightarrow \eta - \alpha \quad \text{and} \quad \xi - 2^{-M_0} \eta \rightarrow \xi, \tag{3.22}$$

we transform the square of left hand of (3.21) into

$$\iiint_{\mathbb{R}^3} F_\alpha(\xi) G_\alpha(\eta) \Upsilon_\alpha(\xi, \eta) e^{i\lambda \Xi_\alpha(\xi, \eta)} d\eta d\xi d\alpha,$$

where

$$\begin{cases} F_\alpha(\xi) := \hat{f}(\xi) \overline{\hat{f}(\xi + 2^{-M_0} \alpha)}; \\ G_\alpha(\eta) := \hat{g}(\eta) \overline{\hat{g}(\eta - \alpha)}; \\ \Upsilon_\alpha(\xi, \eta) := a(\xi, \eta) \overline{a(\xi + 2^{-M_0} \alpha, \eta - \alpha)}; \\ \Xi_\alpha(\xi, \eta) := \Psi(\xi, \eta) - \Psi(\xi + 2^{-M_0} \alpha, \eta - \alpha). \end{cases} \tag{3.23}$$

We split the integration into two parts by considering the value of $|\alpha|$. Take $\alpha_0 > 0$ be a constant as the threshold of the range of α and its value will be determined later.

If $|\alpha| \leq \alpha_0$. Note that $|a| \lesssim 1$, which implies that $|\Upsilon_\alpha| \lesssim 1$. Furthermore, By Hölder’s inequality and Plancherel’s theorem, we can then write

$$\|F_\alpha\|_{L^1(\mathbb{R})} \leq \|f\|_{L^2(\mathbb{R})}^2 \quad \text{and} \quad \|G_\alpha\|_{L^1(\mathbb{R})} \leq \|g\|_{L^2(\mathbb{R})}^2,$$

which, together with $|\Upsilon_\alpha| \lesssim 1$, show that

$$\left| \int_{|\alpha| \leq \alpha_0} \iint_{\mathbb{R}^2} F_\alpha(\xi) G_\alpha(\eta) \Upsilon_\alpha(\xi, \eta) e^{i\lambda \Xi_\alpha(\xi, \eta)} d\eta d\xi d\alpha \right| \lesssim \alpha_0 \|f\|_{L^2(\mathbb{R})}^2 \|g\|_{L^2(\mathbb{R})}^2. \tag{3.24}$$

If $|\alpha| \geq \alpha_0$, the oscillation of the phase function Ξ_α can cause cancellation. We want to obtain our estimates by Hörmander's theorem on the oscillatory integrals with non-degenerate phase [24]. Therefore, we need to obtain the following Lemma 3.2, let us postpone the proof of Lemma 3.2 for the moment. Indeed, by (3.27) in Lemma 3.2 below and Hörmander's theorem [24, Theorem 1.1], we conclude that

$$\begin{aligned} & \left| \int_{|\alpha| \geq \alpha_0} \iint_{\mathbb{R}^2} F_\alpha(\xi) G_\alpha(\eta) \Upsilon_\alpha(\xi, \eta) e^{i\lambda \Xi_\alpha(\xi, \eta)} d\eta d\xi d\alpha \right| \tag{3.25} \\ & \lesssim (\lambda \alpha_0)^{-1/2} \int_{|\alpha| \geq \alpha_0} \|F_\alpha\|_{L^2(\mathbb{R})} \|G_\alpha\|_{L^2(\mathbb{R})} d\alpha \\ & \lesssim (2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}} (\alpha_0)^{-1/2} \int_{|\alpha| \geq \alpha_0} \|F_\alpha\|_{L^2(\mathbb{R})} \|G_\alpha\|_{L^2(\mathbb{R})} d\alpha, \end{aligned}$$

where the last inequality follows from Lemma 3.1. Note that

$$\begin{aligned} \left(\int_{|\alpha| \geq \alpha_0} \|F_\alpha\|_{L^2(\mathbb{R})}^2 d\alpha \right)^{\frac{1}{2}} &= \left(\int_{|\alpha| \geq \alpha_0} \int_{\mathbb{R}} |\hat{f}(\xi) \overline{\hat{f}(\xi + 2^{-M_0 \alpha})}|^2 d\xi d\alpha \right)^{\frac{1}{2}} \\ &\leq 2^{\frac{M_0}{2}} \|f\|_{L^2(\mathbb{R})}^2, \end{aligned}$$

and

$$\left(\int_{|\alpha| \geq \alpha_0} \|G_\alpha\|_{L^2(\mathbb{R})}^2 d\alpha \right)^{\frac{1}{2}} = \left(\int_{|\alpha| \geq \alpha_0} \int_{\mathbb{R}} |\hat{g}(\eta) \overline{\hat{g}(\eta - \alpha)}|^2 d\eta d\alpha \right)^{\frac{1}{2}} \leq \|g\|_{L^2(\mathbb{R})}^2.$$

Continuing the calculation in (3.25), by Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} & \left| \int_{|\alpha| \geq \alpha_0} \iint_{\mathbb{R}^2} F_\alpha(\xi) G_\alpha(\eta) \Upsilon_\alpha(\xi, \eta) e^{i\lambda \Xi_\alpha(\xi, \eta)} d\eta d\xi d\alpha \right| \\ & \lesssim (2C_7)^{\frac{1}{2}} 2^{\frac{M_0}{2} - \frac{m}{2}} (\alpha_0)^{-1/2} \|f\|_{L^2(\mathbb{R})}^2 \|g\|_{L^2(\mathbb{R})}^2. \tag{3.26} \end{aligned}$$

Thus, if $|M_0| \leq (1 - \kappa)m$ for some fixed small absolute constant $\kappa > 0$, then by letting $\alpha_0 := (2C_7)^{\frac{1}{2}} 2^{-\frac{\kappa}{2}m}$, and noticing (3.24) and (3.26), we see that our desired estimate (3.20) holds with $\beta := \frac{\kappa}{8}$.

LEMMA 3.2. For Ξ_α in (3.23), we have

$$|\partial_\xi \partial_\eta \Xi_\alpha(\xi, \eta)| \gtrsim |\alpha|. \tag{3.27}$$

Proof of Lemma 3.2. We first calculate $\partial_\xi \partial_\eta \Psi$. Recall that

$$\Psi(\xi, \eta) = t_c \xi + \eta Q(t_c),$$

where

$$\Phi'_{\xi, \eta}(t_c) = \xi + \eta Q'(t_c) = 0,$$

and Q can be found in (3.11). By a simple calculation, we have

$$\partial_\eta \Psi(\xi, \eta) = \partial_\eta t_c \cdot \xi + Q(t_c) + \partial_\eta t_c \cdot \eta Q'(t_c) = Q(t_c). \tag{3.28}$$

Furthermore, we also have

$$\partial_\xi t_c = -\frac{1}{\eta Q''(t_c)} \quad \text{and} \quad \partial_\eta t_c = -\frac{Q'(t_c)}{\eta Q''(t_c)}. \tag{3.29}$$

Then we obtain

$$\partial_\xi \partial_\eta \Psi(\xi, \eta) = Q'(t_c) \cdot \partial_\xi t_c = -\frac{Q'(t_c)}{\eta Q''(t_c)}.$$

Let us set $H := \partial_\xi \partial_\eta \Psi$, by the mean value theorem, it is easy to see that

$$\partial_\xi \partial_\eta \Xi_\alpha(\xi, \eta) = H(\xi, \eta) - H(\xi + 2^{-M_0} \alpha, \eta - \alpha) = (\nabla H)(\tilde{\xi}, \tilde{\eta}) \cdot (-2^{-M_0} \alpha, \alpha), \tag{3.30}$$

where $(\tilde{\xi}, \tilde{\eta}) := (\xi + \theta_1 2^{-M_0} \alpha, \eta - \theta_2 \alpha)$ with $\theta_1, \theta_2 \in (0, 1)$.

We now claim that

$$|\partial_\xi H(\xi, \eta)| \approx 1 \quad \text{and} \quad |\partial_\eta H(\xi, \eta)| \approx 1. \tag{3.31}$$

Indeed, it implies that

$$\begin{aligned} \partial_\xi H(\xi, \eta) &= \frac{1}{\eta^2 Q''(t_c)} \left(1 - \frac{Q'''(t_c) Q'(t_c)}{(Q''(t_c))^2} \right); \\ \partial_\eta H(\xi, \eta) &= \frac{Q'(t_c)}{\eta^2 Q''(t_c)} \left(1 - \frac{Q'''(t_c) Q'(t_c)}{(Q''(t_c))^2} \right) + \frac{Q'(t_c)}{\eta^2 Q''(t_c)}. \end{aligned}$$

By the definitions of Q in (3.11) and M_0 in (3.2), we have that

$$\begin{aligned} \frac{1}{Q''(t_c)} &= \frac{\gamma(\gamma^{-1}(N)2^{-l})}{(\gamma^{-1}(N)2^{-l})^2 \gamma'(\gamma^{-1}(N)2^{-l}t_c)}; \\ \frac{Q'(t_c)}{Q''(t_c)} &= \frac{\gamma'(\gamma^{-1}(N)2^{-l}t_c)}{\gamma^{-1}(N)2^{-l} \gamma'(\gamma^{-1}(N)2^{-l}t_c)}; \\ 1 - \frac{Q'''(t_c) Q'(t_c)}{(Q''(t_c))^2} &= \frac{(\gamma'(\gamma^{-1}(N)2^{-l}t_c))^2 - \gamma'''(\gamma^{-1}(N)2^{-l}t_c) \gamma'(\gamma^{-1}(N)2^{-l}t_c)}{(\gamma'(\gamma^{-1}(N)2^{-l}t_c))^2}. \end{aligned}$$

Furthermore, from (v) in Lemma 2.1, we also have

$$\frac{C_9}{t^2} \leq \frac{(\gamma'(t))^2 - \gamma'''(t) \gamma'(t)}{(\gamma'(t))^2} \leq \frac{C_8}{t^2}, \quad t \in (0, \infty).$$

Therefore, by (i) in Lemma 2.1, we get

$$1 - \frac{Q'''(t_c)Q'(t_c)}{(Q''(t_c))^2} \approx \frac{(\gamma'(\gamma^{-1}(N)2^{-l}t_c))^2}{(\gamma''(\gamma^{-1}(N)2^{-l}t_c))^2(\gamma^{-1}(N)2^{-l}t_c)^2} \approx 1. \tag{3.32}$$

From (i) (ii) (iv) in Lemma 2.1 and $t_c \in [\frac{1}{2}, 2]$, it now follows that

$$\frac{1}{Q''(t_c)} \approx \frac{Q'(t_c)}{Q''(t_c)} \approx 1. \tag{3.33}$$

Since we will only use Lemma 3.2 to prove (3.25), and $\text{supp } \hat{g} \subset [1, 2]$ in it, we have $\eta \subset [1, 2]$. This, combining with (3.32) and (3.33), leads to (3.31).

For (3.30), by (3.31), we assert that

$$|\partial_\xi \partial_\eta \Xi_\alpha(\xi, \eta)| \gtrsim |1 - 2^{-M_0}| \cdot |\alpha|. \tag{3.34}$$

On the other hand, from (3.6) and the definition of M_0 in (3.2), we have that

$$(C_7)^{-l} \frac{N}{\gamma^{-1}(N)} \leq 2^{M_0} \leq \frac{N}{\gamma^{-1}(N)}, \tag{3.35}$$

which further implies that we may take 2^{M_0} large enough for any given $l \in \mathbb{N}$, since $\frac{N}{\gamma^{-1}(N)} = \frac{\gamma(\gamma^{-1}(N))}{\gamma^{-1}(N)}$, and the fact that $\frac{\gamma(t)}{t}$ is strictly increasing on $(0, \infty)$ and we always take N large enough in Roth's theorem. Therefore, from (3.34), we have

$$|\partial_\xi \partial_\eta \Xi_\alpha(\xi, \eta)| \gtrsim |\alpha|,$$

as our desired result. This completes the proof of Lemma 3.2. \square

3.2. Case 2: $|M_0| \geq (1 - \kappa)m$

In this Subsection, we want to obtain the desired estimate (3.15), which leads to (3.16), under this Case 2 by using σ -uniformity argument [29, Theorem 7.1]. This, in combination with the estimate (3.16) under the Case 1, completes the proof of Lemma 2.5.

As we known, σ -uniformity argument [29, Theorem 7.1] is a useful tool to obtain the $L^2(\mathbb{R})$ boundedness for some operators, which allows us to restrict our discussion on a subspace of $L^2(\mathbb{R})$. Indeed, this argument can be traced back to Christ et al. [9] and Gowers [20]. Here, we quoting a lemma stated in [13, Lemma 4.4] or [23, Lemma 3.3], which is a slight variant of this σ -uniformity argument. We first state the definition of σ -uniformity.

DEFINITION 3.3. Let $\sigma \in (0, 1)$, $\mathcal{I} \subset \mathbb{R}$ be a bounded interval, and $\mathcal{U}(\mathcal{I})$ be a nontrivial subset of $L^2(\mathcal{I})$ such that $\sup_{u \in \mathcal{U}(\mathcal{I})} \|u\|_2 < \infty$. A function $f \in L^2(\mathcal{I})$ is called σ -uniform in $\mathcal{U}(\mathcal{I})$ if

$$\left| \int_{\mathcal{I}} f(x) \overline{u(x)} dx \right| \leq \sigma \|f\|_{L^2(\mathcal{I})}$$

for all $u \in \mathcal{U}(\mathcal{I})$.

We now state the σ -uniformity argument from [29].

LEMMA 3.4. ([29, Theorem 7.1]) *Let \mathcal{L} be a bounded sublinear functional from $L^2(\mathcal{I})$ to \mathbb{C} , and S_σ be the set of all functions that are σ -uniform in $\mathcal{U}(\mathcal{I})$. Denote*

$$A_\sigma := \sup \left\{ \frac{|\mathcal{L}(f)|}{\|f\|_{L^2(\mathcal{I})}} : f \in S_\sigma, f \neq 0 \right\}$$

and

$$K := \sup_{u \in \mathcal{U}(\mathcal{I})} |\mathcal{L}(u)|.$$

Then for all $f \in L^2(\mathcal{I})$, we have

$$|\mathcal{L}f| \leq \max \{A_\sigma, 2\sigma^{-1}K\} \|f\|_{L^2(\mathcal{I})}.$$

In the following, we will apply Lemma 3.4 to

$$\mathcal{L}(g) := \iint_{\mathbb{R}^2} f(x + \lambda t) g(2^{-M_0}x + \lambda Q(t)) h(x) \tau(t) dt dx, \tag{3.36}$$

and σ is a constant, whose exact value will be determined later. We also denote the interval \mathcal{I} means either $[1, 2]$ or $[-2, -1]$, and define

$$\mathcal{U}(\mathcal{I}) := \left\{ \eta \rightarrow A(\xi, \eta) e^{i\alpha\eta + i\lambda\Psi(\xi, \eta)} : \alpha \in \mathbb{R}, 2^{-100} \leq |\xi| \leq 2^{100} \right\},$$

where $A(\xi, \eta)$ is a compactly supported smooth function that is to be determined later. Thus, by Lemma 3.4, we can divide the analysis into the following three parts.

3.2.1. Part 1: Estimates for A_σ

We assume that $\hat{g}|_{\mathcal{I}}$ is σ -uniform in $\mathcal{U}(\mathcal{I})$. From the support of h , we know

$$x + \lambda t \in \left[\frac{1}{2}\lambda, 2^{m+M_0} + 2\lambda \right]$$

for $t \in [\frac{1}{2}, 2]$, then localizing in the spatial variable x , we can rewrite $\mathcal{L}(g)$ as

$$\sum_{l \in \mathbb{N}, 0 \leq l < 2^m} \iint_{\mathbb{R}^2} (\mathbf{1}_{\mathcal{I}_l} f)(x + \lambda t) g(2^{-M_0}x + \lambda Q(t)) (\mathbf{1}_{\mathcal{I}_l} h)(x) \tau(t) dt dx, \tag{3.37}$$

where

$$\mathcal{I}_l := 2^{M_0} [l, l + 1] =: [\alpha_l, \alpha_{l+1}],$$

and

$$\mathcal{J}_t := \left[t2^{M_0} + \frac{1}{2}\lambda, (t+1)2^{M_0} + 2\lambda \right].$$

Note that $x \in \mathcal{J}_t$ implies $x + \lambda t \in \mathcal{J}_t$. Let

$$f_t := \mathbf{1}_{\mathcal{J}_t} f \quad \text{and} \quad h_t := \mathbf{1}_{\mathcal{J}_t} h,$$

by the Fourier transform, we obtain that

$$\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \iiint_{\mathbb{R}^3} \widehat{f}_t(\xi) e^{ix\xi} \widehat{g}(\eta) e^{i2^{-M_0}x\eta} \left(\int_{\mathbb{R}} e^{i\lambda\Phi_{\xi,\eta}(t)} \tau(t) dt \right) h_t(x) dx d\xi d\eta, \quad (3.38)$$

where $\Phi_{\xi,\eta}$ can be found in (3.17). Due to the fact that $|2^{-M_0}x\eta - 2^{-M_0}\alpha_t\eta| \leq 1$ for all $x \in \mathcal{J}_t$, where α_t is an arbitrary point chosen from \mathcal{J}_t , it is natural to replace $e^{i2^{-M_0}x\eta}$ by $e^{i2^{-M_0}\alpha_t\eta}$, we can then write

$$e^{i2^{-M_0}(x-\alpha_t)\eta} = \sum_{s=0}^{\infty} \frac{i^s}{s!} (2^{-M_0}(x-\alpha_t))^s \eta^s$$

by the Taylor series expression. Applying this into (3.38), then we need to calculate every term of the Taylor series expansion, separately. Similar to [12, 23, 30], we will here only consider the term of $s = 0$ for simplicity of notation, because the process is the same for all $s \in \mathbb{N}$. Thus, we should only to bound

$$\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \iint_{\mathbb{R}^2} \widehat{g}(\eta) \widehat{f}_t(\xi) \widehat{h}_t(-\xi) e^{i2^{-M_0}\alpha_t\eta} \left(\int_{\mathbb{R}} e^{i\lambda\Phi_{\xi,\eta}(t)} \tau(t) dt \right) d\xi d\eta.$$

Applying the stationary phase method again in (3.18), and noticing $\lambda^{-\frac{1}{2}} \leq (2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}}$ from Lemma 3.1, forces us to control the following expression

$$(2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}} \sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \iint_{\mathbb{R}^2} \widehat{g}(\eta) A(\xi, \eta) e^{i2^{-M_0}\alpha_t\eta + i\lambda\Psi(\xi,\eta)} \widehat{f}_t(\xi) \widehat{h}_t(-\xi) d\eta d\xi, \quad (3.39)$$

where $A(\xi, \eta)$ is a compactly supported smooth function. We should to explain that $A(\xi, \eta)$ is equal to $a(\xi, \eta)$ in (3.18) if we restrict $s = 0$, but for general $s \in \mathbb{N}$, $A(\xi, \eta)$ is equal to $a(\xi, \eta)$ multiple of some other functions; see, for example, η^s . Therefore, we replace $a(\xi, \eta)$ by $A(\xi, \eta)$ in this time, because it will not cause any problems. On the other hand, we also have omitted the remainder term as in (3.18), which can be treated as (3.19).

By using the definition of σ -uniformity, which with respect to g , we have

$$\left| \int_{\mathbb{R}} \widehat{g}(\eta) A(\xi, \eta) e^{i2^{-M_0}\alpha_t\eta + i\lambda\Psi(\xi,\eta)} d\eta \right| \leq \sigma \|g\|_{L^2(\mathbb{R})}.$$

Furthermore, by Hölder’s inequality and Plancherel’s theorem, we can bound (3.39) by

$$(2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}} \sigma \sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|f_t\|_{L^2(\mathbb{R})} \|h_t\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})}. \tag{3.40}$$

By Cauchy-Schwarz inequality, we further bound (3.40) by

$$(2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}} \sigma \|g\|_{L^2(\mathbb{R})} \left(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|f_t\|_{L^2(\mathbb{R})}^2 \right)^{\frac{1}{2}} \left(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|h_t\|_{L^2(\mathbb{R})}^2 \right)^{\frac{1}{2}}. \tag{3.41}$$

We now want to bound $(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|f_t\|_{L^2(\mathbb{R})}^2)^{\frac{1}{2}}$ by the spacial localization of f_t on the interval \mathcal{J}_t . We can write

$$\begin{aligned} \left(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|f_t\|_{L^2(\mathbb{R})}^2 \right)^{\frac{1}{2}} &= \left[\int_{\mathbb{R}} \left(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \mathbf{1}_{\mathcal{J}_t}(x) \right) |f(x)|^2 dx \right]^{\frac{1}{2}} \\ &\lesssim (\max\{1, 2^{-M_0} \lambda\})^{\frac{1}{2}} \|f\|_{L^2(\mathbb{R})}, \end{aligned} \tag{3.42}$$

where the number $\max\{1, 2^{-M_0} \lambda\}$ comes from the extent of the overlap property of \mathcal{J}_t ’s. Indeed, let $\Omega := \frac{3}{2} \frac{\lambda}{2^{M_0}}$, then $|\mathcal{J}_t| = (1 + \Omega) 2^{M_0}$ for all $t \in \mathbb{N}$ and $0 \leq t < 2^m$. Furthermore,

- ★ When $\Omega < 1$, the step length 2^{M_0} makes that \mathcal{J}_t ’s overlap at most twice, which implies that $\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \mathbf{1}_{\mathcal{J}_t} \lesssim 1$;
- ★ When $\Omega \geq 1$, the step length 2^{M_0} makes that \mathcal{J}_t ’s overlap at most $1 + \Omega$ times, which implies that $\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \mathbf{1}_{\mathcal{J}_t} \lesssim 1 + \Omega \lesssim \Omega \lesssim 2^{-M_0} \lambda$.

Thus, (3.42), combined with the following trivial estimate

$$\left(\sum_{t \in \mathbb{N}, 0 \leq t < 2^m} \|h_t\|_{L^2(\mathbb{R})}^2 \right)^{\frac{1}{2}} \lesssim \|h\|_{L^2(\mathbb{R})},$$

implies that (3.41) can be dominated by

$$\begin{aligned} &(2C_7)^{\frac{1}{2}} 2^{-\frac{m}{2}} (\max\{1, 2^{-M_0} \lambda\})^{\frac{1}{2}} \sigma \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^2(\mathbb{R})} \\ &\lesssim (2C_7)^{\frac{1}{2}} 2^{\frac{M_0}{2}} (\max\{1, 2^{-M_0} \lambda\})^{\frac{1}{2}} \sigma \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}, \end{aligned} \tag{3.43}$$

where the last estimate from the fact that

$$\|h\|_{L^2(\mathbb{R})} \leq 2^{\frac{m+M_0}{2}} \|h\|_{L^\infty(\mathbb{R})}$$

since $h \in L^\infty([0, 2^{m+M_0}])$.

Here, we need to a further estimate to the bound $(2C_7)^{\frac{1}{2}} 2^{\frac{M_0}{2}} (\max\{1, 2^{-M_0} \lambda\})^{\frac{1}{2}} \sigma$ in (3.43) under the condition that $|M_0| \geq (1 - \kappa)m$, where κ is a fixed small positive constant. Therefore,

- When $2^{-M_0}\lambda \lesssim 1$, it is easy to see that $(2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}(\max\{1, 2^{-M_0}\lambda\})^{\frac{1}{2}}\sigma \lesssim (2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}\sigma \lesssim (2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}2^{\frac{\kappa m}{2}}\sigma$;
- When $2^{-M_0}\lambda \gtrsim 1$, by Lemma 3.1, we have $(2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}(\max\{1, 2^{-M_0}\lambda\})^{\frac{1}{2}}\sigma \lesssim (2C_7)^{\frac{1}{2}}\lambda^{\frac{1}{2}}\sigma \lesssim (C_7)^{\frac{1}{2}}2^{\frac{m}{2}}\sigma$. We write m as the sum of $(1 - \kappa)m$ and κm , by $|M_0| \geq (1 - \kappa)m$, it implies $(C_7)^{\frac{1}{2}}2^{\frac{m}{2}}\sigma$ can be bounded by $(C_7)^{\frac{1}{2}}2^{\frac{|M_0|}{2}}2^{\frac{\kappa m}{2}}\sigma$. Furthermore, by (3.35), we may assume that $2^{M_0} \geq 1$ if N large enough for given $l \in \mathbb{N}$. Therefore, $(2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}(\max\{1, 2^{-M_0}\lambda\})^{\frac{1}{2}}\sigma \lesssim (C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}2^{\frac{\kappa m}{2}}\sigma$.

Putting all of things together, we further bound (3.43) by

$$(2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}2^{\frac{\kappa m}{2}}\sigma \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}.$$

Therefore, we obtain

$$A_\sigma \lesssim (2C_7)^{\frac{1}{2}}2^{\frac{M_0}{2}}2^{\frac{\kappa m}{2}}\sigma \|f\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}. \tag{3.44}$$

This finishes the estimate for A_σ for $\hat{g}|_{\mathcal{S}}$ is σ -uniform in $\mathcal{U}(\mathcal{S})$.

3.2.2. Part 2: Estimates for K

We now consider the case in which $\hat{g}|_{\mathcal{S}} \in \mathcal{U}(\mathcal{S})$. By changing variable

$$x \rightarrow 2^{M_0+m}x - \lambda t,$$

we rewrite $\mathcal{L}(g)$ in (3.36) as

$$2^{M_0+m} \iint_{\mathbb{R}^2} f(2^{M_0+m}x)g(2^m x - 2^{-M_0}\lambda t + \lambda Q(t))h(2^{M_0+m}x - \lambda t)\tau(t) dt dx.$$

Using Hölder's inequality, it further bound by $\|f\|_{L^2(\mathbb{R})} \|T(g, h)\|_{L^2(\mathbb{R})}$, where

$$T(g, h)(x) := 2^{\frac{M_0+m}{2}} \int_{\mathbb{R}} g(2^m x - 2^{-M_0}\lambda t + \lambda Q(t))h(2^{M_0+m}x - \lambda t)\tau(t) dt.$$

Similar to the previous discussion, by the Fourier inversion transform for g , we rewrite $T(g, h)(x)$ as

$$2^{\frac{M_0+m}{2}} \iint_{\mathbb{R}^2} \hat{g}(\eta)e^{i\eta\lambda(2^m\lambda^{-1}x - 2^{-M_0}t + Q(t))}h(2^{M_0+m}x - \lambda t)\tau(t)d\eta dt. \tag{3.45}$$

Using our assumption that $\hat{g}|_{\mathcal{S}} \in \mathcal{U}(\mathcal{S})$, and substituting $\hat{g}(\eta)$ by $A(\xi, \eta)e^{i\alpha\eta + i\lambda\Psi(\xi, \eta)}$, $\alpha \in \mathbb{R}$ is arbitrary and ξ is a parameter, we obtain that

$$2^{\frac{M_0+m}{2}} \iint_{\mathbb{R}^2} A(\xi, \eta)e^{i\lambda(\eta z_{x,t} + \Psi(\xi, \eta))}h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda t)\tau(t)d\eta dt,$$

where

$$z_{x,t} := 2^m\lambda^{-1}x - 2^{-M_0}t + Q(t). \tag{3.46}$$

Here, we have used $x \rightarrow x - 2^{-m}\alpha$, which would not to change $\|T(g, h)\|_{L^2(\mathbb{R})}$.

We will apply the stationary phase method again, and here the phase function is

$$\tilde{\Phi}_{x,t,\xi}(\eta) := \eta z_{x,t} + \Psi(\xi, \eta). \tag{3.47}$$

As in the previous discussion, it is safe to assume that the equation

$$\partial_\eta \tilde{\Phi}_{x,t,\xi}(\eta) = 0$$

has a unique solution $\eta_c := \eta_c(x, t, \xi) \in \mathcal{S}$, where

$$\partial_\eta \tilde{\Phi}_{x,t,\xi}(\eta) = z_{x,t} + \partial_\eta \Psi(\xi, \eta). \tag{3.48}$$

Otherwise, our estimate will be trivial by Van der Corput's lemma as (3.19).

Recall that $\partial_\eta \Psi(\xi, \eta) = Q(t_c(\xi, \eta))$, see (3.28). Let us set

$$T_c := t_c(\xi, \eta_c(x, t, \xi))$$

and

$$\tilde{\Psi}_\xi(x, t) := \tilde{\Phi}_{x,t,\xi}(\eta_c) = \eta_c z_{x,t} + \Psi(\xi, \eta_c),$$

which, together with the facts that

$$\Psi(\xi, \eta_c) = T_c \xi + \eta_c Q(T_c) \quad \text{and} \quad z_{x,t} + Q(T_c) = 0,$$

show that

$$\tilde{\Psi}_\xi(x, t) = T_c \xi = Q^{-1}(-z_{x,t})\xi, \tag{3.49}$$

where Q^{-1} is the inverse function of Q .

By (3.48), (3.28), (3.29) and (3.33), respectively, then the following identity is valid.

$$\left| \partial_\eta^2 \tilde{\Phi}_{x,t,\xi}(\eta) \right| = \left| \partial_\eta^2 \Psi(\xi, \eta) \right| = \left| Q'(t_c) \partial_\eta t_c \right| = \left| \frac{(Q'(t_c))^2}{\eta Q''(t_c)} \right| \approx 1. \tag{3.50}$$

Thus, by stationary phase method, we see that

$$\begin{aligned} T(g, h)(x) &= \lambda^{-\frac{1}{2}} 2^{\frac{m+M_0}{2}} \int_{\mathbb{R}} \tilde{a}_\xi(x, t) e^{i\lambda \tilde{\Psi}_\xi(x,t)} h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda t) \tau(t) dt \\ &\quad + 2^{\frac{m+M_0}{2}} \int_{\mathbb{R}} R_{\xi,x,t}(\lambda) h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda t) \tau(t) dt, \end{aligned}$$

where $\tilde{a}_\xi(x, t)$ is a smooth compactly supported function and the remainder term $R_{\xi,x,t}(\lambda)$ can be bounded by λ^{-1} . Furthermore, by (3.50), we also have

$$|\tilde{a}_\xi| \lesssim 1. \tag{3.51}$$

As above, we may ignore the remainder term $R_{\xi,x,t}(\lambda)$. Therefore, with abuse of nota-

tions, by Lemma 3.1, it suffices to consider the major term

$$T(g, h)(x) := (2C_7)^{\frac{1}{2}} 2^{\frac{M_0}{2}} \int_{\mathbb{R}} \tilde{a}_\xi(x, t) e^{i\lambda \tilde{\Psi}_\xi(x, t)} h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda t) \tau(t) dt. \quad (3.52)$$

Expanding the square of the $L^2(\mathbb{R})$ -norm of $T(g, h)$ and changing variables gives

$$(2C_7)^l 2^{M_0} \iiint_{\mathbb{R}^3} e^{i\lambda(\tilde{\Psi}_\xi(x, t) - \tilde{\Psi}_\xi(x, t+s))} \tilde{a}_\xi(x, t) \overline{\tilde{a}_\xi(x, t+s)} \times h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda t) \overline{h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda(t+s))} \tau(t) \overline{\tau(t+s)} dt dx ds. \quad (3.53)$$

Let

$$\begin{cases} \chi_\xi(x, t) := \tilde{a}_\xi(x, t) \overline{\tilde{a}_\xi(x, t+s)}; \\ \mathbb{H}(x) := h(2^{M_0+m}x - 2^{M_0}\alpha) \overline{h(2^{M_0+m}x - 2^{M_0}\alpha - \lambda s)}; \\ \Theta(t) := \tau(t) \overline{\tau(t+s)}. \end{cases} \quad (3.54)$$

Then (3.53) equals

$$(2C_7)^l 2^{M_0} \iiint_{\mathbb{R}^3} e^{i\lambda(\tilde{\Psi}_\xi(x, t) - \tilde{\Psi}_\xi(x, t+s))} \chi_\xi(x, t) \mathbb{H}(x - 2^{-m-M_0}\lambda t) \Theta(t) dt dx ds. \quad (3.55)$$

Furthermore, by changing variable

$$x \rightarrow x + 2^{-m-M_0}\lambda t,$$

(3.55) becomes

$$(2C_7)^l 2^{M_0} \iint_{\mathbb{R}^2} \left(\int_{\mathbb{R}} e^{i\lambda(\tilde{\Psi}_\xi(x+2^{-m-M_0}\lambda t, t) - \tilde{\Psi}_\xi(x+2^{-m-M_0}\lambda t, t+s))} \times \chi_\xi(x + 2^{-m-M_0}\lambda t, t) \Theta(t) dt \right) \mathbb{H}(x) dx ds. \quad (3.56)$$

We now further analyse the phase function

$$P_{x,s}(t) := \tilde{\Psi}_\xi(x + 2^{-m-M_0}\lambda t, t) - \tilde{\Psi}_\xi(x + 2^{-m-M_0}\lambda t, t+s).$$

Indeed, from (3.46) and (3.49), the expression above is equal to

$$[Q^{-1}(-2^m\lambda^{-1}x - Q(t)) - Q^{-1}(-2^m\lambda^{-1}x + 2^{-M_0}s - Q(t+s))] \xi. \quad (3.57)$$

Let

$$\vartheta(x, t) := Q^{-1}(-2^m\lambda^{-1}x - Q(t)). \quad (3.58)$$

We can rewrite the phase function as

$$P_{x,s}(t) = \vartheta(x, t) - \vartheta(x - 2^{-m-M_0}\lambda s, t+s). \quad (3.59)$$

Now, we turn to (3.56), i.e., $\|T(g, h)\|_{L^2(\mathbb{R})}^2$. We first establish a trivial estimate for the

oscillatory integral about t . It implies from the definition of $\Theta(t)$ that $t \in [\frac{1}{2}, 2]$ and $t + s \in [\frac{1}{2}, 2]$, which further implies $|s| \leq 4$. So (3.51) leads to

$$\left| \int_{\mathbb{R}} e^{i\lambda(\tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t) - \tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t+s))} \chi_{\xi}(x+2^{-m-M_0}\lambda t, t) \Theta(t) dt \right| \lesssim 1. \tag{3.60}$$

We now consider the following two cases:

$$\begin{cases} |x| < 2^{1-l-M_0}, \\ |x| \geq 2^{1-l-M_0}, \end{cases}$$

respectively. For the former case, i.e., $|x| < 2^{1-l-M_0}$, it is easy to see that

$$\|T(g, h)\|_{L^2(\mathbb{R})}^2 \lesssim (C_7)^l \|h\|_{L^\infty(\mathbb{R})}^2. \tag{3.61}$$

For the second case, i.e., $|x| \geq 2^{1-l-M_0}$, we need to use the following Lemma 3.5 and Lemma 3.6. By Lemma 3.5, similarly to the Corollary in p. 334 in Stein’s book [35], we have

$$\begin{aligned} & \left| \int_{\mathbb{R}} e^{i\lambda(\tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t) - \tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t+s))} \chi_{\xi}(x+2^{-m-M_0}\lambda t, t) \Theta(t) dt \right| \\ & \lesssim \frac{1}{\lambda 2^l |x| |s|} + \frac{(C_7)^l}{\lambda 2^l |x| |s|^2}. \end{aligned} \tag{3.62}$$

This, in combination with (3.60), shows that

$$\begin{aligned} & \left| \int_{\mathbb{R}} e^{i\lambda(\tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t) - \tilde{\Psi}_{\xi}(x+2^{-m-M_0}\lambda t, t+s))} \chi_{\xi}(x+2^{-m-M_0}\lambda t, t) \Theta(t) dt \right| \\ & \lesssim \left(\frac{1}{\lambda 2^l |x| |s|} \right)^{\frac{1}{4}} + \left(\frac{(C_7)^l}{\lambda 2^l |x| |s|^2} \right)^{\frac{1}{4}}. \end{aligned} \tag{3.63}$$

On the other hand, from the definition of χ_{ξ} , it implies that

$$|x + 2^{-m-M_0}\lambda t| \lesssim 1.$$

By Lemma 3.1 and (3.35), it implies that

$$|2^{-m-M_0}\lambda t| \lesssim (2^{-1}C_7)^l \frac{\gamma^{-1}(N)}{N} \lesssim 1,$$

provided that we take N large enough for given $l \in \mathbb{N}$. Then, we obtain $|x| \lesssim 1$. Therefore, in this case, (3.56) is dominated by

$$\begin{aligned} & (2C_7)^l 2^{M_0} \|h\|_{L^\infty(\mathbb{R})}^2 \int_{|x| \lesssim 1} \int_{|s| \leq 4} \left(\frac{1}{\lambda 2^l |x| |s|} \right)^{\frac{1}{4}} + \left(\frac{(C_7)^l}{\lambda 2^l |x| |s|^2} \right)^{\frac{1}{4}} dx ds \\ & \lesssim \lambda^{-\frac{1}{4}} (2C_7)^{\frac{5}{4}l} 2^{M_0} \|h\|_{L^\infty(\mathbb{R})}^2. \end{aligned} \tag{3.64}$$

Putting these two cases together, by Lemma 3.1,

$$\|T(g, h)\|_{L^2(\mathbb{R})}^2 \lesssim (2C_7)^{\frac{3}{2}l} 2^{-\frac{m}{4}} 2^{M_0} \|h\|_{L^\infty(\mathbb{R})}^2. \tag{3.65}$$

Here, by (3.35), we have assumed that $2^{M_0} \geq 1$.

Furthermore, we obtain that

$$|\mathcal{L}(g)| \lesssim (2C_7)^{\frac{3}{4}l} 2^{-\frac{m}{8}} 2^{\frac{M_0}{2}} \|f\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}, \tag{3.66}$$

for $\hat{g}|_{\mathcal{J}} \in \mathcal{U}(\mathcal{J})$. Therefore, we obtain

$$K \lesssim (2C_7)^{\frac{3}{4}l} 2^{-\frac{m}{8}} 2^{\frac{M_0}{2}} \|f\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}. \tag{3.67}$$

This finishes the estimate for K for $\hat{g}|_{\mathcal{J}} \in \mathcal{U}(\mathcal{J})$.

LEMMA 3.5. [28, Lemma 2.1] *Suppose ϕ is real-valued and smooth in (a, b) , and that both*

$$|\phi'(x)| \geq \sigma_1 \quad \text{and} \quad |\phi''(x)| \leq \sigma_2$$

for any $x \in (a, b)$. Then we have

$$\left| \int_a^b e^{i\phi(t)} dt \right| \leq \frac{2}{\sigma_1} + (b-a) \frac{\sigma_2}{\sigma_1^2}.$$

LEMMA 3.6. For $P_{x,s}$ in (3.59), if $|x| \geq 2^{1-l-M_0}$, then we have

$$|\partial_t P_{x,s}(t)| \gtrsim 2^l |x| |s| \quad \text{and} \quad |\partial_t^2 P_{x,s}(t)| \lesssim (2C_7)^l |x|.$$

Proof of Lemma 3.6. By a simple calculation to (3.58), we obtain

$$\partial_t \vartheta(x, t) = -(Q^{-1})'(-\lambda^{-1}\alpha - 2^m \lambda^{-1}x - Q(t)) Q'(t) = -\frac{Q'(t)}{Q'(\vartheta(x, t))}. \tag{3.68}$$

Here, we used the fact that $(Q^{-1})'(t)Q'(Q^{-1}(t)) = 1$ since $Q(Q^{-1}(t)) = t$. Furthermore, we also have

$$\partial_t^2 \vartheta(x, t) = -\frac{Q'(t)^2}{Q'(\vartheta(x, t))} \left(\frac{Q''(t)}{Q'(t)^2} + \frac{Q''(\vartheta(x, t))}{Q'(\vartheta(x, t))^2} \right). \tag{3.69}$$

From previous argument, we have $t \in [\frac{1}{2}, 2]$. On the other hand, we also have that $\vartheta(x, t) \approx T_c$, which further implies that we may think $\vartheta(x, t) \in [\frac{1}{2}, 2]$. Indeed, from (3.46) and (3.58), which implies

$$Q(\vartheta(x, t)) = -z_{x,t} - 2^{-M_0}t.$$

This, combined with that facts that

$$Q(T_c) = -z_{x,t}$$

and (iv) in Lemma 2.1, from the definition of Q , it suffices to show

$$|-z_{x,t} - 2^{-M_0 t}| \approx |-z_{x,t}|.$$

- ◇ Estimates of $|-z_{x,t} - 2^{-M_0 t}| \lesssim |-z_{x,t}|$. From the definition of Q , we have

$$|z_{x,t}| \approx 2^{-M_0} \frac{\gamma(\gamma^{-1}(N)2^{-l})}{\gamma^{-1}(N)2^{-l}}.$$

This, combined with (3.6) and (3.35), leads to

$$(C_7)^{-l} \lesssim |z_{x,t}| \lesssim (C_7)^l.$$

Furthermore, from (3.35), if N large enough for given $l \in \mathbb{N}$, we have

$$|2^{-M_0 t}| \lesssim \frac{\gamma^{-1}(N)}{N} (C_7)^l \lesssim (C_7)^{-l} \lesssim |z_{x,t}|,$$

which further implies that $|-z_{x,t} - 2^{-M_0 t}| \lesssim |-z_{x,t}|$.

- ◇ Estimates of $|-z_{x,t} - 2^{-M_0 t}| \gtrsim |-z_{x,t}|$. It is enough to show that there exists $v \in (0, 1)$ such that

$$v|-z_{x,t}| \geq |2^{-M_0 t}|,$$

which equals to

$$v \frac{\gamma(\gamma^{-1}(N)2^{-l})}{\gamma^{-1}(N)2^{-l}} \gtrsim 1,$$

the later can be obtained if N large enough for given $l \in \mathbb{N}$.

We extend γ to an odd function defined on \mathbb{R} , from $t \in [\frac{1}{2}, 2]$ and $\vartheta(x, t) \in [\frac{1}{2}, 2]$, by Lemma 3.1 and (3.58), we have

$$|\partial_t^2 \vartheta(x, t)| \approx \left| \frac{Q''(t)}{Q'(t)^2} - \frac{Q''(Q^{-1}(Q(t) + 2^m \lambda^{-1}x))}{Q'(Q^{-1}(Q(t) + 2^m \lambda^{-1}x))^2} \right|. \tag{3.70}$$

By the mean value theorem

$$\begin{aligned} & \left| \frac{Q''(t)}{Q'(t)^2} - \frac{Q''(Q^{-1}(Q(t) + 2^m \lambda^{-1}x))}{Q'(Q^{-1}(Q(t) + 2^m \lambda^{-1}x))^2} \right| \\ &= \left| \left(\frac{Q''(Q^{-1}(\cdot))}{Q'(Q^{-1}(\cdot))^2} \right)' (Q(t) + \tilde{\theta} 2^m \lambda^{-1}x) \right| |2^m \lambda^{-1}x|, \end{aligned}$$

where $\tilde{\theta} \in (0, 1)$. Furthermore,

$$\begin{aligned} & \left(\frac{Q''(Q^{-1}(\cdot))}{Q'(Q^{-1}(\cdot))^2} \right)' (t) \\ &= \frac{1}{Q'(Q^{-1}(t))^2} \left(\frac{Q'''(Q^{-1}(t))Q'(Q^{-1}(t)) - Q''(Q^{-1}(t))^2}{Q'(Q^{-1}(t))^2} - \frac{Q''(Q^{-1}(t))^2}{Q'(Q^{-1}(t))^2} \right), \end{aligned}$$

Noticing that $|Q^{-1}(Q(t) + \tilde{\theta}2^m\lambda^{-1}x)| \in [\frac{1}{2}, 2]$, from the definition of Q , by (i) (v) in Lemma 2.1 and Lemma 3.1, we have

$$\left| \left(\frac{Q''(Q^{-1}(\cdot))}{Q'(Q^{-1}(\cdot))^2} \right)' (Q(t) + \tilde{\theta}2^m\lambda^{-1}x) \right| \approx 1.$$

Therefore, we obtain

$$|\partial_t^2 \vartheta(x, t)| \approx 2^m \lambda^{-1} |x|. \tag{3.71}$$

Simple calculation also leads to

$$\partial_x \partial_t \vartheta(x, t) = -2^m \lambda^{-1} \frac{Q''(\vartheta(x, t))}{Q'(\vartheta(x, t))^3} Q'(t).$$

Furthermore, from $t \in [\frac{1}{2}, 2]$ and $\vartheta(x, t) \in [\frac{1}{2}, 2]$, by Lemma 3.1

$$|\partial_x \partial_t \vartheta(x, t)| \approx 2^m \lambda^{-1}. \tag{3.72}$$

Using the mean value theorem again, then we have

$$\begin{aligned} \partial_t P_{x,s}(t) &= \partial_t (\vartheta(x, t) - \vartheta(x - 2^{-m-M_0}\lambda s, t) + \vartheta(x + 2^{-m-M_0}\lambda s, t) \\ &\quad - \vartheta(x - 2^{-m-M_0}\lambda s, t + s)) \\ &= \partial_t \partial_x \vartheta(x - \tilde{\theta}_1 2^{-m-M_0}\lambda s, t) 2^{-m-M_0}\lambda s - \partial_t^2 \vartheta(x - 2^{-m-M_0}\lambda s, t + \tilde{\theta}_2 s) s, \end{aligned}$$

where $\tilde{\theta}_1, \tilde{\theta}_2 \in (0, 1)$. From (3.71) and (3.72), which further implies that

$$|\partial_t P_{x,s}(t)| \gtrsim 2^m \lambda^{-1} |x| |s| - 2^{-M_0} |s| \gtrsim 2^l |x| |s| \tag{3.73}$$

as desired. Here, we have used Lemma 3.1 and the condition that $|x| \geq 2^{1-l-M_0}$.

From Lemma 3.1 and (3.71), which trivially leads to

$$|\partial_t P_{x,s}(t)| \lesssim (2C_7)^l |x|. \tag{3.74}$$

This completes the proof of Lemma 3.6. \square

3.2.3. Part 3: Completes the estimate (3.15)

Combining estimates (3.44) and (3.67), we obtain from Lemma 3.4 that

$$|\mathcal{L}(g)| \lesssim (2C_7)^l 2^{\frac{M_0}{2}} \max\{\sigma 2^{\frac{\kappa m}{2}}, \sigma^{-1} 2^{-\frac{m}{8}}\} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}. \tag{3.75}$$

Choosing κ small enough, and letting $\sigma := 2^{-\kappa m}$, we can bound this by

$$(2C_7)^l 2^{\frac{M_0}{2} - \frac{\kappa}{2} m} \|f\|_{L^2(\mathbb{R})} \|g\|_{L^2(\mathbb{R})} \|h\|_{L^\infty(\mathbb{R})}.$$

This completes the proof of (3.15) with $\beta := \frac{\kappa}{2}$.

Acknowledgements. X. Li and H. Yu thank Prof. Junfeng Li for the many helpful

discussions.

REFERENCES

- [1] V. BERGELSON AND A. LEIBMAN, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. 9 (1996), no. 3, 725–753.
- [2] J. BOURGAIN, *A nonlinear version of Roth's theorem for sets of positive density in the real line*, J. Analyse Math. 50 (1988), 169–181.
- [3] J. BOURGAIN, *A Szemerédi type theorem for sets of positive density in \mathbb{R}^k* , Israel J. Math. 54 (1986), no. 3, 307–316.
- [4] J. BOURGAIN, M.-C. CHANG, *Nonlinear Roth type theorems in finite fields*, Israel J. Math. 221 (2017), no. 2, 853–867.
- [5] A.-P. CALDERÓN, *Cauchy integrals on Lipschitz curves and related operators*, Proc. Nat. Acad. Sci. U.S.A. 74 (1977), no. 4, 1324–1327.
- [6] A. CARBERY, M. CHRIST, J. VANCE, S. WAINGER AND D. WATSON, *Operators associated to flat plane curves: L^p estimates via dilation methods*, Duke Math. J. 59 (1989), no. 3, 675–700.
- [7] A. CARBERY, J. VANCE, S. WAINGER AND D. WATSON, *The Hilbert transform and maximal function along flat curves, dilations, and differential equations*, Amer. J. Math. 116 (1994), no. 5, 1203–1239.
- [8] H. CARLSSON, M. CHRIST, A. CORDOBA, J. DUOANDIKOETXEA, J. L. RUBIO DE FRANCIA, J. VANCE, S. WAINGER AND D. WEINBERG, *L^p estimates for maximal functions and Hilbert transforms along flat convex curves in \mathbb{R}^2* , Bull. Amer. Math. Soc. (N.S.) 14 (1986), no. 2, 263–267.
- [9] M. CHRIST, X. LI, T. TAO AND C. THIELE, *On multilinear oscillatory integrals, nonsingular and singular*, Duke Math. J. 130 (2005), no. 2, 321–351.
- [10] M. CHRIST, A. NAGEL, E. M. STEIN AND S. WAINGER, *Singular and maximal Radon transforms: analysis and geometry*, Ann. of Math. (2) 150 (1999), no. 2, 489–577.
- [11] B. COOK, Á. MAGYAR AND M. PRAMANIK, *A Roth-type theorem for dense subsets of \mathbb{R}^d* , Bull. Lond. Math. Soc. 49 (2017), no. 4, 676–689.
- [12] D. DONG, X. LI AND W. SAWIN, *Improved estimates for polynomial Roth type theorems in finite fields*, arXiv:1709.00080, 2017.
- [13] P. DURCIK, S. GUO AND J. ROOS, *A polynomial Roth theorem on the real line*, Trans. Amer. Math. Soc. 371 (2019), no. 10, 6973–6993.
- [14] P. ERDŐS AND P. TURÁN, *On Some Sequences of Integers*, J. London Math. Soc. 11 (1936), no. 4, 261–264.
- [15] E. B. FABES AND N. M. RIVIÈRE, *Singular integrals with mixed homogeneity*, Studia Math. 27 (1966), 19–38.
- [16] R. FRASER, S. GUO AND M. PRAMANIK, *Polynomial Roth theorems on sets of fractional dimensions*, arXiv:1904.11123, 2019.
- [17] H. FURSTENBERG, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. 31 (1977), 204–256.
- [18] H. FURSTENBERG AND Y. KATZNELSON, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. 34 (1978), 275–291 (1979).
- [19] W. T. GOWERS, *A new proof of Szemerédi theorem*, Geom. Funct. Anal. 11 (2001), no. 3, 465–588.
- [20] W. T. GOWERS, *A new proof of Szemerédi theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), no. 3, 529–551.
- [21] B. GREEN, *On arithmetic structures in dense sets of integers*, Duke Math. J. 114 (2002), no. 2, 215–238.
- [22] B. GREEN AND T. TAO, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) 167 (2008), no. 2, 481–547.
- [23] J. GUO AND L. XIAO, *Bilinear Hilbert transforms associated with plane curves*, J. Geom. Anal. 26 (2016), no. 2, 967–995.
- [24] L. HÖRMANDER, *Oscillatory integrals and multipliers on FL^p* , Ark. Mat. 11 (1973), 1–11.
- [25] B. F. JONES, *A class of singular integrals*, Amer. J. Math. 86 (1964), 441–462.
- [26] M. T. LACEY AND C. THIELE, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$* , Ann. of Math. (2) 146 (1997), no. 3, 693–724.

- [27] M. T. LACEY AND C. THIELE, *On Calderón's conjecture*, Ann. of Math. (2) 149 (1999), no. 2, 475–496.
- [28] J. LI AND H. YU, *L^p boundedness of Hilbert transforms associated with variable plane curves*, arXiv: 1806.08589, 2018.
- [29] X. LI, *Bilinear Hilbert transforms along curves I: The monomial case*, Anal. PDE 6 (2013), no. 1, 197–220.
- [30] X. LI AND L. XIAO, *Uniform estimates for bilinear Hilbert transforms and bilinear maximal functions associated to polynomials*, Amer. J. Math. 138 (2016), no. 4, 907–962.
- [31] V. LIE, *On the boundedness of the bilinear Hilbert transform along “non-flat” smooth curves*, Amer. J. Math. 137 (2015), no. 2, 313–363.
- [32] V. LIE, *On the boundedness of the bilinear Hilbert transform along “non-flat” smooth curves. The Banach triangle case (L^r , $1 \leq r < \infty$)*, Rev. Mat. Iberoam. 34 (2018), no. 1, 331–353.
- [33] A. NAGEL, J. VANCE, S. WAINGER AND D. WEINBERG, *Hilbert transforms for convex curves*, Duke Math. J. 50 (1983), no. 3, 735–744.
- [34] K. F. ROTH, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104–109.
- [35] E. M. STEIN, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Mathematical Series 43, Monographs in Harmonic Analysis III, Princeton University Press, Princeton, NJ, 1993.
- [36] E. SZEMERÉDI, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. 20 (1969), 89–104.
- [37] E. SZEMERÉDI, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.
- [38] T. TAO AND T. ZIEGLER, *A multi-dimensional Szemerédi theorem for the primes via a correspondence principle*, Israel J. Math. 207 (2015), no. 1, 203–228.
- [39] T. TAO AND T. ZIEGLER, *The primes contain arbitrarily long polynomial progressions*, Acta Math. 201 (2008), no. 2, 213–305.

(Received September 20, 2020)

Xiang Li

Laboratory of Mathematics and Complex Systems
(Ministry of Education of China)

School of Mathematical Sciences, Beijing Normal University
Beijing 100875, People's Republic of China
e-mail: 201731130007@mail.bnu.edu.cn

Dunyan Yan

School of Mathematics
University of Chinese Academy of Sciences
Beijing 100049, People's Republic of China
e-mail: ydunyan@ucas.ac.cn

Haixia Yu

Department of Mathematics
Sun Yat-sen University
Guangzhou 510275, People's Republic of China
e-mail: yuhx26@mail.sysu.edu.cn

Xingsong Zhang

RDFZ CHAOYANG SCHOOL
Beijing 100028, People's Republic of China
e-mail: zhangxingsong17@mails.ucas.ac.cn