THE EXISTENCE OF RESPONSE TORI FOR HAMILTONIAN WITH NORMAL DEGENERACY

L. XU, W. SI, AND M. WU

Abstract. In this paper, we prove the existence of response tori for a general Hamiltonian with normal degeneracy, that is,

$$H = \langle \omega, I\rangle + \sum_{i=1}^{n} \lambda_i \frac{x_i^{l_i}}{l_i} + \sum_{i=1}^{n} \frac{y_i^2}{2} + \varepsilon P(\omega t, z),$$

where $I \in \mathbb{R}^d$, $z := (x_1, \cdots, x_n, y_1, \cdots, y_n)^\top \in \mathbb{R}^{2n}$, $\theta := \omega t \in \mathbb{T}^d$ and $\omega \in \mathbb{R}^d$ is the Diophantine frequency. The order numbers $l_i > 2$ are fixed integers, $\lambda_i \neq 0$ are fixed constants for $i = 1, \cdots, n$ and $0 < \varepsilon \ll 1$ is a sufficiently small parameter. When P is independent of y, it can be seen as the energy function of several quasi-periodically forced oscillator equations, that is,

$$\begin{cases} \ddot{x}_1 + \lambda_1 x_1^{l_1-1} + \varepsilon f_1(\omega t, x) = 0, \\ \ddot{x}_2 + \lambda_2 x_2^{l_2-1} + \varepsilon f_2(\omega t, x) = 0, \\ \qquad \vdots \\ \ddot{x}_n + \lambda_n x_n^{l_n-1} + \varepsilon f_n(\omega t, x) = 0, \end{cases} \tag{0.1}$$

where $f_i := \frac{\partial P(\omega t, x)}{\partial x_i}$ for $i = 1, 2, \cdots, n$.

Most of the previous results focus on a single oscillator equation and prove the existence of response solutions under certain non-degenerate assumptions. In the present paper, we will consider high dimensional system (0.1) coupled by oscillator equations in different degenerate types. We will prove that the response solutions still exist around perturbed equilibria, which reveals the mechanics of the existence of response solution for a system coupled by degenerate nonlinear oscillator equations.

For the sake of generality, we will actually consider a general Hamiltonian normal form and prove the persistence of invariant tori with fixed Diophantine frequency ω by the methods of finding relative equilibria, improving the order of perturbations, KAM iterations, and measure estimates. The result will then be applied to the problem of the existence of response solutions of the above system (0.1).

1. Introduction

In the present paper, we consider a general Hamiltonian normal form as follows

$$H = \langle \omega, I\rangle + \sum_{i=1}^{n} \lambda_i \frac{x_i^{l_i}}{l_i} + \sum_{j=1}^{n} \frac{y_j^2}{2} + \varepsilon P(\theta, z), \tag{1.1}$$

2000 *Mathematics Subject Classification.* Primary 37J40, 70H08.

Key words and phrases. Normal degeneracy, KAM theory, response solutions.

Corresponding author: W.Si. Email: siwenmath@sdu.edu.cn.

The first author was partially supported by National Natural Science Foundation of China (Grant No. 12271204), and the Department project of Science and Technology of Jilin Province (Grant No. 20200201265JC). The second author was partially supported by the National Natural Science Foundation of China (Grant No. 12001315,11971261,11571201,12071255), Shandong Provincial Natural Science Foundation, China (Grant No. ZR2020MA015).

where $I \in \mathbb{R}^d$, $z := (x, y)^\top \in \mathbb{R}^{2n}$, $\theta \in \mathbb{T}^d$ and $\omega \in \mathbb{R}^d$ is the Diophantine frequency. The order numbers $l_i > 2$ are fixed integers satisfying $l_i \neq l_j$ for $1 \leq i, j \leq n$. The constants $\lambda_i \neq 0$, $i = 1, 2, \cdots, n$, are fixed constants and $0 < \varepsilon < \varepsilon_* \ll 1$ is a sufficiently small parameter. The function H is real analytic with respect to (θ, I, z). Moreover, the Hamiltonian system is associated with standard symplectic form $\mathrm{d}\theta \wedge \mathrm{d}I + \mathrm{d}x \wedge \mathrm{d}y$.

When the perturbation P is independent of y, the Hamiltonian (1.1) can be seen as an energy function of a system coupled by several oscillator equations forced by small quasi-periodic functions, that is,

$$\begin{cases} \ddot{x}_1 + \lambda_1 x_1^{l_1 - 1} + \varepsilon f_1(\omega t, x) = 0, \\ \qquad \vdots \\ \ddot{x}_n + \lambda_n x_n^{l_n - 1} + \varepsilon f_n(\omega t, x) = 0, \end{cases} \tag{1.2}$$

where $f_i = \frac{\partial P}{\partial x_i}$, $i = 1, 2, \cdots, n$. We mention that a response solution of system (1.2) is a quasi-periodic solution $x(t) = (x_1(\omega t, \varepsilon), \ \cdots, x_n(\omega t, \varepsilon))^\top$ with the same frequency ω as in the forcing functions f_i, $i = 1, 2, \cdots, n$. The existence of response solutions play an important role in studying the harmonic responses and oscillatory properties. In the present paper, we will obtain the existence of the response solutions of equation (1.2) by the persistence of invariant tori with fixed Diophantine frequency ω of Hamiltonian (1.1).

Plenty results in the existence of the response solutions have been obtained with respect to a single oscillator equation with a quasi-periodic forced function, that is,

$$\ddot{x} + c\dot{x} + a^2 x + \lambda x^{l-1} = \varepsilon f(\omega t, x, \dot{x}), \tag{1.3}$$

where $a, \ c, \ \lambda$ are fixed constants, $l > 2$ is a fixed integer, f is a real analytic function with respect to $(\theta, x, \dot{x})$ with $\theta := \omega t$, ε is a small parameter. When $a \neq 0, c = 0$, the system can be seen as a harmonic oscillator with nonlinear term. We say the equation is in *non-degenerate* case since $x = 0$ is non-degenerate equilibrium for the unforced equation. As an early application of normal form reduction, Moser [19] firstly proved the existence of response solutions under the assumption that f satisfying reversible condition, i.e., $f(-\omega t, x, -\dot{x}) = f(\omega t, x, \dot{x})$. The result was generalized to the case $c \neq 0$ but sufficiently small in[12]. Recently, the existence of response solutions for (1.3) with forced function in Liouvillean type frequency has been proved in [18], [23] in the case that $d = 2$ and later generalized to the case $d > 2$ in [3], [29].

When $a = c = 0$, $x = 0$ is a degenerate equilibrium of the unforced equation, the existence of the response solutions as well as the persistence of invariant tori become challenging. When equation (1.3) is independent of $\dot{x}$, there exists a Hamiltonian function $H: \ \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^2 \to \mathbb{R}$ which is an integral of equation (1.3). In the extended phrase space $\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^2$ with standard symplectic structure, the Hamiltonian H can be written as

$$H(\theta, I, x, y, \varepsilon) = \langle \omega, I \rangle + \lambda \frac{x^l}{l} + \frac{y^2}{2} + \varepsilon P(\theta, x). \tag{1.4}$$

Hence, the existence of response solutions is equal to the persistence of the invariant tori with fixed frequency ω of Hamiltonian (1.4).

The persistence results of Hamiltonian normal form with different non-degenerate conditions were demonstrated [7], [8], [20], [31] based on modified KAM iterations. Other results related on the existence of quasi-periodical solutions were proved via variation method, see [13], [14], [15], [32] for details. For instance, You [31] firstly considered the case that l is even and $\lambda < 0$, it was proved that Hamiltonian (1.4) admits a family of d-invariant tori with a frequency ω_* which slightly shifts from ω. Note that the assumption of the perturbation in [31] is only the smallness and real analyticity, since $(x, y) = (0, 0)$ is a saddle-like critical point of the unperturbed system

(1.4) for $\lambda < 0$. Otherwise, when $(x, y) = (0, 0)$ is a center-like critical point, the persistence results only hold on certain cantor set due to the existence of small divisors.

As it was formulated in [21], the authors consider the following completely degenerate Hamiltonian

$$H(\theta, I, x, y, \varepsilon) = \langle \omega, I\rangle + \lambda\frac{x^l}{l} + \frac{y^m}{m} + \varepsilon P(\theta, x, y), \tag{1.5}$$

where $\lambda \neq 0$, m, $n \geq 2$ are positive integers, P is real analytic with respect to (θ, x, y). Under certain non-degenerate assumptions, it was proved when $\lambda < 0$, the systems (1.5) admits a family of invariant response tori as long as $\varepsilon \in (0, \varepsilon_*)$ is sufficiently small, otherwise, there exists a almost full measure Cantor set $O \subset (0, \varepsilon_*)$ such that the persistence result holds for $\varepsilon \in O$. Although adding an assumption to perturbation P, the result proved the existence of response solution for the motion equation with respect to Hamiltonian (1.5) for fixed Diophantine vector ω.

A nature question is what will happen to the existence of response tori (solutions) when several oscillator equations coupled together. A similar problem was considered by L. Corsi and G. Gentilde in [6] but for the case that $\lambda = 0$, that is,

$$\ddot{x} = \varepsilon f(\omega t, x),$$

where $x \in \mathbb{T}^d$, $d \geq 1$, f is real analytic and ε is sufficiently small. The existence of response solutions was proved for $d > 1$ in [6] under the assumption that f is even with respect to ωt, that is, $f(-\omega t, x) = f(\omega t, x)$ and for $d = 1$ in [5] without any further non-degenerate condition but only smallness on forced function f. As a consequence, we aim to prove the persistence of response tori for Hamiltonian (1.1), which leads to the existence of response solutions of equation (1.2).

Define the average of a function with respect to θ as $[f(\cdot, z)] := \int_{\mathbb{T}^d} f(\theta, z)\mathrm{d}\theta$ and denote that

$$p_i := \left[\frac{\partial P(\cdot, 0)}{\partial x_i}\right], \quad i = 1, 2, \cdots, n.$$

Then we formulate our main result under the following assumptions:

A1) Assume that ω is a Diophantine vector, that is,

$$|\langle k, \omega\rangle| > \frac{\gamma}{|k|^\tau},$$

where $\gamma > 0$, $\tau > d - 1$ are fixed constants.

A2) For $i = 1, 2, \cdots, n$, assume that $p_i \neq 0$. Moreover, $p_i/\lambda_i < 0$ when l_i is odd.

As it is classified in [22] and [31], the d-dimensional tori of unperturbed Hamiltonian (1.4) is in hyperbolic type if $\lambda < 0$. Hence, we say that the d-dimensional tori of unperturbed Hamiltonian (1.1) is in hyperbolic type if $\lambda_i < 0$ for $i = 1, 2, \cdots, n$. Otherwise, we say the d-dimensional tori of unperturbed Hamiltonian is in mixed type. Then we formulate our main result as follows.

Main Theorem. *Consider Hamiltonian systems (1.1) and assume* **A1)**, **A2)** *hold. Then the followings hold.*

(1) *If $\lambda_i < 0$ for $i = 1, 2, \cdots, n$, then there exists a sufficiently small parameter $0 < \varepsilon_* \ll 1$ such that, as $0 < \varepsilon \leq \varepsilon_*$, the Hamiltonian systems admit a C^N smooth family of real analytic, hyperbolic response tori around a family of hyperbolic type relative equilibria, where $N \geq 1$ is a fixed integer.*

(2) *If there exits at least one $\lambda_i > 0$ for certain $1 \le i \le n$, then there exists a sufficiently small parameter $0 < \varepsilon_* \ll 1$ and a Cantor set $O_\infty \subset (0, \varepsilon_*)$ with measure estimate $\frac{|\text{meas } O_\infty|}{\varepsilon_*} = 1 - O(\varepsilon_*^{1-\sigma})$ such that, as $\varepsilon \in O_\infty$, the Hamiltonian systems admit a C^N Whitney smooth family of real analytic response tori around a family of mixed type of relative equilibria, where $\sigma := \min\{\frac{1}{l_1-1}, \ \cdots, \frac{1}{l_n-1}\}$ and $N \ge 2n^2 - n$.*

As it is mentioned above, the Main Theorem can be applied to prove the existence of response solutions for a couple of nonlinear oscillator equations. Hence we consider equations (1.2) and assume the following conditions hold.

A3) There exists a real analytic function $P : \mathbb{T}^d \times \mathbb{R}^n \to \mathbb{R}$ such that

$$f_i = \frac{\partial P}{\partial x_i}, \qquad i = 1, 2, \cdots, n.$$

A4) For $i = 1, 2, \cdots, n$, denote $f_i = [f(\cdot, 0)]$ and assume that $f_i \neq 0$. Moreover, $f_i/\lambda_i < 0$ when l_i is odd.

Corollary 1. *Consider equations (1.2) and assume* **A1)**, **A3)**, **A4)** *hold. Then the followings hold.*

(1) *If $\lambda_i < 0$ for $i = 1, 2, \cdots, n$, then there exists a sufficiently small parameter $0 < \varepsilon_* \ll 1$ such that, as $0 < \varepsilon \le \varepsilon_*$, the equations (1.2) admit a C^N smooth family of real analytic response solutions around a family of relative equilibria in hyperbolic type, where $N \ge 1$ is a fixed integer.*

(2) *If there exits at least one $\lambda_i > 0$ for certain $1 \le i \le n$, then there exists a sufficiently small parameter $0 < \varepsilon_* \ll 1$ and a Cantor set $O_\infty \subset (0, \varepsilon_*)$ of with measure estimate $\frac{|\text{meas } O_\infty|}{\varepsilon_*} = 1 - O(\varepsilon_*^{1-\sigma})$ such that, as $\varepsilon \in O_\infty$, the equations (1.2) admit a C^N Whitney smooth family of real analytic responsive solutions around a family of relative equilibria in mixed type, where $\sigma := \min\{\frac{1}{l_1-1}, \ \cdots, \frac{1}{l_n-1}\}$ and $N \ge 2n^2 - n$.*

Remark 1.1. *The Main Theorem will be proved via KAM iterations since we deal with the hyperbolic type as well as the mixed type. We mention that the existence of response tori in hyperbolic type can be proved simply via the uniform contraction mapping principle, which requires no Diophantine condition on ω. See e.g.* [2], [30] *for general situations.*

Remark 1.2. *Comparing to the previous results in the persistence of lower dimensional tori for a multi-scale Hamiltonian system, for instance, we consider the following Hamiltonian normal form in* [26]*, that is*

$$H = \langle \omega, I \rangle + \frac{1}{2} \langle M(\omega, \varepsilon) z, z \rangle + \varepsilon P(\theta, I, z, \varepsilon),$$

where ω varying in a closed region in $\mathbb{R}^d$. We have proved that under certain non-degenerate assumption, most of the tori $T_\omega = \{\omega\} \times \{I = 0\} \times \{z = 0\}$ persist but the tangent frequency shifts to $\tilde{\omega}$ with the estimate that $|\tilde{\omega} - \omega| = O(\varepsilon)$. The main difference in the present paper is we prove the persistence of the tori with fixed frequency ω, consequently, we take ε as a parameter varying in a small interval.

We also mention that, the difference in measure estimate between hyperbolic type and mixed type is due to the reason that there are no small divisors during the KAM iterations in hyperbolic type. Hence, we could obtain the persistence of a C^N-smooth family of response tori for Hamiltonian (1.1), as well as a C^N-smooth family of response solutions for coupled equations (1.2) in hyperbolic type, for any integer $N \ge 1$.

The rest sections are organized as follows. In section 2, we will solve the average equation with respect to (1.1) to obtain a new Hamiltonian H_0 with non-singular normal frequency. In section 3, we will perform a finite steps of KAM iterations to Hamiltonian H_0 to obtain a new normal form H_* with sufficiently small perturbation. The smallness of the perturbation ensures the standard KAM iteration and the measure estimate can be directly applied on Hamiltonian H_*. Hence we will prove the Main Theorem by applying standard KAM method to H_* in section 4 such that we obtain the persistence of the invariant tori with fixed frequency ω. In section 5, we will prove the measure estimate. It is different from pervious ones since we take the ε as a parameter instead of the frequency ω.

2. Normalization

In this section, we will normalize the Hamiltonian normal form (1.1) based on the conditions **A1) - A2)**. The normalization procedure includes finding relative equilibria and removing Hamiltonian (1.1) into the vicinity of relative equilibria. As a result, the transformed Hamiltonian in the vicinity of relative equilibria is of multi-scale in ε, their order of perturbations also need to be improved in order to perform infinite steps of KAM iterations.

2.1. Notations and weighted norms.

We first introduce some notations and norms which will be used in the following proof.

For each $r, s > 0$, we denote

$$D(r,s) = \mathbb{T}^d_r \times \mathbf{B}_s,$$

where

$$\mathbf{B}_s := \{z = (x_1, \cdots, x_n, y_1, \cdots y_n) \in \mathbb{C}^{2n} : \ |z| \le s\}$$

is the ball of radius s in $\mathbb{C}^{2n}$ and

$$\mathbb{T}^d_r := \{\theta = (\theta_1, \dots, \theta_d) \in \mathbb{C}^d/(2\pi\mathbb{Z})^d : \ |\mathrm{Im}\ \theta_j| \le r, \quad j = 1, 2, \dots, d\}$$

is the strip neighborhood of size s of the d-torus $\mathbb{T}^d = \mathbb{R}^d/(2\pi\mathbb{Z})^d$ in $\mathbb{C}^d$. For given $\varepsilon_* > 0$, let $\mathcal{D} := (0, \varepsilon_*)$. We say the function

$$f(\theta, z, \varepsilon) := \sum_{\imath \in \mathbb{Z}^{2n}_+,\ k \in \mathbb{Z}^d} f_k(\varepsilon) z^\imath \mathrm{e}^{\sqrt{-1}\langle k, \theta \rangle},$$

is real analytic in $(\theta, z) \in D(r, s)$ and C^N-(Whitney) smooth in $\varepsilon \in \mathcal{D}$ for certain fixed integer $N \ge 1$, if the norm $\|\cdot\|_{D(r,s)\times\mathcal{D}}$ defined as follows is finite, that is

$$\|\partial^i_\varepsilon f\|_{D(r,s)\times\mathcal{D}} = \sum_{k\in\mathbb{Z}^d, \imath \in \mathbb{Z}^2_+} \sup_{\varepsilon\in\mathcal{D}} |\partial^i_\varepsilon f_k(\varepsilon)| s^{|\imath|} \mathrm{e}^{r|k|} < +\infty, \quad \forall\ i = 0, 1, \cdots, N,$$

where $\partial^i_\varepsilon f_k(\varepsilon) = |f_k(\varepsilon)| + \cdots + \varepsilon^i |\frac{\mathrm{d}^i f_k(\varepsilon)}{\mathrm{d}\varepsilon^i}|$ and $|k| = \sum_{l=1}^d |k_l|$ for $k = (k_1, \cdots, k_d) \in \mathbb{Z}^d$. Taking $s = 0$ in the above, we can define the $\|\cdot\|_{r,\mathcal{D}}$ norm for any function $f : \mathbb{T}^d_r \to \mathbb{C}$,

$$f(\theta, \varepsilon) = \sum_{k\in\mathbb{Z}^d} f_k(\varepsilon) \mathrm{e}^{\sqrt{-1}\langle k, \theta\rangle},$$

which is analytic in θ and C^N-(Whitney) smooth in $\varepsilon \in \mathcal{D}$. The Banach algebra of all such functions under the $\|\cdot\|_{r,B}$ norm is denoted by

$$C^N(\mathbb{T}^d_r \times B) = \{f(\theta, \varepsilon) : \|\partial^i_\varepsilon f(\theta, \varepsilon)\|_{r,B} < +\infty,\ i = 0, 1, \cdots, N\}.$$

As mentioned above, for any function $f(\theta, z)$ in $D(r, s)$, we denote its average with respect to θ by

$$[f(\cdot, z, \varepsilon)] = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} f(\theta, z, \varepsilon)\ \mathrm{d}\theta.$$

Moreover, the notation $\mathrm{D}^j f(\theta, z, \varepsilon)$ denotes the partial derivatives of function f with respect to z in the j-th order, that is,

$$\mathrm{D}^j f(\theta, z, \varepsilon) = \sum_{\jmath \in \mathbb{Z}_+^{2n}, j=|\jmath|} \frac{\partial^\jmath f(\theta, z, \varepsilon)}{\partial z^\jmath}.$$

Without loose of generality, we will frequently use c or c_i, $i = 0, 1, \cdots, 6$ to denote the intermediate constants depending on domain constants $r,\ s,\ \eta > 0$, Diophantine constants γ, τ and the norms of known functions. We also use $\|\cdot\|$ to denote the weighted norms of (vector-valued) functions, as well as the norms of matrix operators in the following proof.

2.2. **Average equations and relative equilibria.** The average equations are referred to the averaged part of the Hamiltonian vector fields in the normal direction. We will find relative equilibria by solving such average equations corresponding to (1.1). The result is formulated as follows.

Lemma 2.1. *Consider the average equations corresponding to Hamiltonian (1.1) and assume* **A2)** *holds. Then, there exits a family of nonzero solutions in form of* $z_\varepsilon = (x_\varepsilon, y_\varepsilon)^\top$*, where*

$$\begin{aligned} x_\varepsilon &= (\varepsilon^{\frac{1}{l_1-1}} x_1^* + O(\varepsilon^{\frac{1}{l_1-1}+\sigma}),\ \cdots,\ \varepsilon^{\frac{1}{l_n-1}} x_n^* + O(\varepsilon^{\frac{1}{l_n-1}+\sigma}))^\top, \\ y_\varepsilon &= (\varepsilon y_1^* + O(\varepsilon^{1+\sigma}), \cdots,\ \varepsilon y_n^* + O(\varepsilon^{1+\sigma}))^\top, \end{aligned}$$

where $0 < \varepsilon \ll 1$*,* $\sigma := \min\{\frac{1}{l_1-1}, \cdots, \frac{1}{l_n-1}\}$ *and* $x_i^* \neq 0$ *for* $i = 1, 2, \cdots, n$*.*

Proof. The corresponding average equations in normal direction with respect to Hamiltonian (1.1) are as follows,

$$\left\{ \begin{array}{ll} \frac{\partial H}{\partial x_i} := \lambda_i x_i^{l_i-1} + \varepsilon p_i + O(\varepsilon|z|) + O(\varepsilon^2) = 0, & \forall\ i = 1, 2, \cdots, n, \\ \frac{\partial H}{\partial y_i} := y_i + \varepsilon q_i + O(\varepsilon|z|) + O(\varepsilon^2) = 0, & \forall\ i = 1, 2, \cdots, n, \end{array} \right. \tag{2.1}$$

where,

$$p_i := \left[\frac{\partial P(\cdot, 0)}{\partial x_i}\right], \quad q_i := \left[\frac{\partial P(\cdot, 0)}{\partial y_i}\right], \quad i = 1, 2, \cdots, n.$$

Introduce the re-scale transformations

$$x_i \to \varepsilon^{\frac{1}{l_i-1}} x_i, \qquad y_j \to \varepsilon y_i, \quad i = 1, 2, \cdots, n. \tag{2.2}$$

By substituting the transformations into (2.1) and dividing the equations by ε, we obtain that

$$\left\{ \begin{array}{ll} H_i(x, y, \varepsilon) := \lambda_i x_i^{l_i-1} + p_i + O(\varepsilon^a|z|) + O(\varepsilon) = 0, & i = 1, 2, \cdots, n, \\ H_{n+i}(x, y, \varepsilon) := y_j + q_j + O(\varepsilon^a|z|) + O(\varepsilon) = 0, & i = 1, \cdots, n, \end{array} \right. \tag{2.3}$$

where $\sigma := \min\{\frac{1}{l_1-1}, \cdots, \frac{1}{l_n-1}\}$. Define that

$$x_i^* = (-a_i/\lambda_i)^{\frac{1}{l_i-1}}, \qquad y_i^* = -b_i, \quad i = 1, 2, \cdots, n. \tag{2.4}$$

Based on **A2)**, x_i^*, $i = 1, 2, \cdots, n$, are well defined and $x_i^* \neq 0$. Denote $x_* = (x_1^*,\ \cdots, x_n^*)^\top$, $y_* = (y_1^*, \cdots, y_n^*)$, it yields that $H_i(x_*, y_*, 0) = 0$ for $i = 1, 2, \cdots, 2n$ and

$$\mathrm{D}H(x_*, y_*, 0) = \det \begin{pmatrix} \frac{\partial H_1(x_*, y_*, 0)}{\partial x_1} & \cdots & \frac{\partial H_1(x_*, y_*, 0)}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial H_{2n}(x_*, y_*, 0)}{\partial x_1} & \cdots & \frac{\partial H_{2n}(x_*, y_*, 0)}{\partial y_n} \end{pmatrix} = \prod_{i=1}^n \lambda_i (l_i - 1)(x_i^*)^{l_i-2} \neq 0.$$

By the Implicit Function Theorem, we obtain a family of nonzero solutions for equations (2.3) in form of

$$x_{i,\varepsilon} = x_i^* + O(\varepsilon^\sigma), \qquad y_{i,\varepsilon} = y_i^* + O(\varepsilon^\sigma) \quad i = 1, \cdots, n,$$

where $(x_*, y_*)^\top$ are defined as in (2.4). By tracing back to the re-scaling transformation, the average equation (2.1) admits a family of solutions in form of

$$z_\varepsilon = (\varepsilon^{\frac{1}{l_1-1}} x_{1,\varepsilon}, \cdots, \varepsilon^{\frac{1}{l_n-1}} x_{n,\varepsilon}, \ \varepsilon y_{1,\varepsilon}, \cdots, \varepsilon y_{n,\varepsilon})^\top.$$

Since that H_i are C^N-smoothly depending on ε, it follows from the Implicit Function Theorem that the relative equilibria z_ε forms a C^N-smooth family with respect to $\varepsilon \in \mathcal{D}$ for any fixed positive integer N. □

Now we remove Hamiltonian (1.1) in the vicinity of the relative equilibria obtained in the Lemma 2.1, we obtain the new Hamiltonian normal form as follows.

Lemma 2.2. *Consider Hamiltonian (1.1) and assume* **A2)** *holds. Then, introducing the linear transformation* $L: \ z \to z + z_\varepsilon$, *such that the Hamiltonian (1.1) can be reduced into the following form*

$$\tilde{H} = H \circ L = \tilde{e} + \langle \omega, I \rangle + \langle \tilde{M} z, z \rangle + \tilde{h}(z, \varepsilon) + \varepsilon \tilde{G}(\theta, z, \varepsilon) + \varepsilon \tilde{E}(\theta, z, \varepsilon), \tag{2.5}$$

where, $\tilde{e}$ *is a constant term depending on* ε, *the normal frequency* M *is a* $2n \times 2n$ *non-singular symmetric matrix in form of*

$$\tilde{M} = A(\varepsilon) + \varepsilon \tilde{A}(\varepsilon),$$

$$A(\varepsilon) = \text{diag}\{\varepsilon^{a_1} m_1(\varepsilon), \cdots, \ \varepsilon^{a_n} m_n(\varepsilon), \ 1, \cdots, \ 1\}. \tag{2.6}$$

The order numbers $a_i = \frac{l_i - 2}{l_i - 1}$, $m_i = \lambda_i (l_i - 1)(x_i^*)^{l_i - 2} + O(\varepsilon^\sigma)$, $i = 1, 2, \cdots, n$. *The functions* $\tilde{h} := O(|z|^3)$, $\tilde{G} := O(|z|^3)$ *and the perturbation is in the following form*

$$\tilde{E} \ := \sum_{|\imath| \le 2} \tilde{E}_\imath(\theta, \varepsilon) z^\imath, \qquad [\tilde{E}_\imath(\cdot, \varepsilon)] = 0, \ |\imath| = 0, 1, 2. \tag{2.7}$$

Moreover, the Hamiltonian $\tilde{H}$ *is real analytic with respect to* $(I, \theta, z) \in D(r - \eta, s - \eta)$ *and* C^N *smoothly depending on* $\varepsilon \in \mathcal{D} := (0, \varepsilon_*)$, *where* $0 < \eta < \min\{r, s\}/8$ *and* ε_* *is sufficiently small.*

Proof. Replacing z by $z + z_\varepsilon$, we obtain that the following calculation results:

$$\begin{aligned}
\tilde{H} \ &= \ H \circ L = \langle \omega, y \rangle + \sum_{i=1}^{n-1} \lambda_i (l_i - 1) x_{i,\varepsilon}^{l_i - 2} \frac{x_i^2}{2} + \lambda_n \frac{x^2}{2} + \sum_{j=1}^{n} \frac{y_j^2}{2} + \varepsilon \langle \left[\frac{\partial^2 P(\cdot, z_\varepsilon, \varepsilon)}{\partial z^2} \right] z, z \rangle \\
&+ \ \sum_{i=1}^{n-1} \sum_{k=3}^{l_i} \frac{\lambda_i}{l_i} C_{l_i}^k x_{i,\varepsilon}^{l_i - k} x^k + \varepsilon (P - P(\theta, z_\varepsilon, \varepsilon) - \langle \frac{\partial P(\theta, z_\varepsilon, \varepsilon)}{\partial z}, z \rangle - \langle \frac{\partial^2 P(\theta, z_\varepsilon, \varepsilon)}{\partial z^2} z, z \rangle) \\
&\quad + \sum_{i=1}^{n} \lambda_i x_{i,\varepsilon}^{l-1} x_i + \sum_{j=1}^{n} y_{j,\varepsilon} y_j + \varepsilon \langle \left[\frac{\partial P(\cdot, z_\varepsilon, \varepsilon)}{\partial z} \right], z \rangle + \sum_{i=1}^{n} \frac{\lambda_i}{l_i} x_{i,\varepsilon}^l + \sum_{j=1}^{n} \frac{y_{j,\varepsilon}^2}{2} + \varepsilon P(\theta, z_\varepsilon, \varepsilon) \\
&\quad + \varepsilon \langle \frac{\partial P(\theta, z_\varepsilon, \varepsilon)}{\partial z}, z \rangle - \langle \left[\frac{\partial P(\cdot, z_\varepsilon, \varepsilon)}{\partial z} \right], z \rangle + \varepsilon \langle \frac{\partial^2 P(\theta, z_\varepsilon, \varepsilon)}{\partial z^2} z, z \rangle - \langle \left[\frac{\partial^2 P(\cdot, z_\varepsilon, \varepsilon)}{\partial z^2} \right] z, z \rangle.
\end{aligned}$$

Since z_ε solves average equations (2.1), we have

$$\sum_{i=1}^{n} \lambda_i x_{i,\varepsilon}^{l-1} x_i + \sum_{j=1}^{n} y_{j,\varepsilon} y_j + \varepsilon \langle \left[\frac{\partial P(\cdot, z_\varepsilon, \varepsilon)}{\partial z} \right], z \rangle = 0.$$

The lemma is proved by denoting that $a_i,\ m_i,\ i=1,2,\cdots,n$ as in above, and

$$\begin{aligned}
\tilde{A}(\varepsilon) &= \left[\frac{\partial^2 P(\cdot,z_\varepsilon,\varepsilon)}{\partial z^2}\right], \quad \tilde{h}(z,\varepsilon)=\sum_{i=1}^{n-1}\sum_{k=3}^{l_i}\frac{\lambda_i}{l_i}C_{l_i}^k x_{i,\varepsilon}^{l_i-k}x_i^k,\\
\tilde{G}(\theta,z,\varepsilon) &= P(\theta,z+z_\varepsilon,\varepsilon)-P(\theta,z_\varepsilon,\varepsilon)-\langle\frac{\partial P(\theta,z_\varepsilon,\varepsilon)}{\partial z},z\rangle-\langle\frac{\partial^2 P(\theta,z_\varepsilon,\varepsilon)}{\partial z^2}z,z\rangle,\\
\tilde{E}_0(\theta,z,\varepsilon) &= P(\theta,z_\varepsilon,\varepsilon)-[P(\cdot,z_\varepsilon,\varepsilon)],\\
\tilde{E}_1(\theta,z,\varepsilon) &= \langle\frac{\partial P(\theta,z_\varepsilon,\varepsilon)}{\partial z},z\rangle-\langle\left[\frac{\partial P(\cdot,z_\varepsilon,\varepsilon)}{\partial z}\right],z\rangle,\\
\tilde{E}_2(\theta,z,\varepsilon) &= \langle\frac{\partial^2 P(\theta,z_\varepsilon,\varepsilon)}{\partial z^2}z,z\rangle-\langle\left[\frac{\partial^2 P(\cdot,z_\varepsilon,\varepsilon)}{\partial z^2}\right]z,z\rangle,\\
\tilde{e} &= \sum_{i=1}^{n}\frac{\lambda_i}{l_i}x_{i,\varepsilon}^l+\sum_{j=1}^{n}\frac{y_{j,\varepsilon}^2}{2}+\varepsilon[P(\cdot,z_\varepsilon,\varepsilon)].
\end{aligned}$$

□

Since the family of relative equilibria z_ε solves the average equations, it yields that the average of the perturbation E equals zero. By performing one step of average process, we can improve the order of the perturbation in (2.5) into at least the order of $O(\varepsilon^{2a})$, where $a:=\max\{a_1,\cdots,a_n\}$.

Lemma 2.3. *Consider the Hamiltonian (2.5) on domain $D(r-\eta,s-\eta)\times\mathcal{D}$. As ε_* is sufficiently small, then for any $\varepsilon\in(0,\varepsilon_*)$, there exists a C^N-smooth family of transformations $\Phi_{0,\varepsilon}: D(r-2\eta,\ s-2\eta)\to D(r-\eta,\ s-\eta)$, under which the Hamiltonian (2.5) can be transformed into the following form:*

$$H_0=H\circ\Phi_{0,\varepsilon}=\langle\omega,I\rangle+\langle z,M_0z\rangle+h_0(z,\varepsilon)+\varepsilon^a G_0(\theta,z,\varepsilon)+\varepsilon^{2a}P_0(\theta,z,\varepsilon)+e_0(\varepsilon).$$

where M_0 is a nonsingular matrix with $|M_0^{-1}|=O(\varepsilon^{-a})$, $h_0,\ G_0:=O(|z|^3)$, and

$$\|\partial_\varepsilon^i \mathrm{D}^j P_0\|_{D(r-2\eta,s-2\eta)\times\mathcal{D}}\le c_1\varepsilon^{2-2a},\qquad i=0,1,\cdots,N,\ j=0,1,2,$$

where c_1 is a positive constant depending on n,d,s,r,η and independent of ε.

Proof. For fixed $\varepsilon\in\mathcal{D}$, define that

$$K=([\log\frac{1}{\varepsilon}]+1)^2, \tag{2.8}$$

where for fixed constant a, $[a]>0$ denotes the maximum integer less as a. We will truncate the Fourier series of $\tilde{E}$ up to order K-th term, i.e., we write the perturbation into its Fourier series and the truncated form $\bar{E}_i,\ i=0,1,2$ are of the form

$$\bar{E}_0=\sum_{0<|k|\le K}E_{k0}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad \bar{E}_1=\sum_{0<|k|\le K}\langle E_{k1},z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad \bar{E}_2=\sum_{0<k\le K}\langle E_{k2}z,z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},$$

$$\hat{E}_0=\sum_{|k|>K}E_{k0}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad \hat{E}_1=\sum_{|k|>K}\langle E_{k1},z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad \hat{E}_2=\sum_{k>K}\langle E_{k2}z,z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle}.$$

It follows from the definition of K that for $i=0,1,\cdots,N$ and $j=0,1,2$, we have

$$\begin{aligned}
&\|\partial_\varepsilon^i \mathrm{D}^j\hat{E}\|_{D(r-5\eta/4,s-5\eta/4)\times\mathcal{D}}\le c\sum_{|k|>K}\mathrm{e}^{-\frac{\eta|k|}{4}}\le c\int_K^\infty t^{d+1}\mathrm{e}^{-\frac{\eta}{4}}\mathrm{d}t\le c\varepsilon, \qquad (2.9)\\
&\|\partial_\varepsilon^i \mathrm{D}^j\bar{E}\|_{D(r-\eta,s-\eta)\times\mathcal{D}}\le c,
\end{aligned}$$

where $\hat{E} := \hat{E}_0 + \hat{E}_1 + \hat{E}_2,\ \bar{E} := \bar{E}_0 + \bar{E}_1 + \bar{E}_2$. Now we seek for a canonical transformation as the time-1 map ϕ_F^1 of the flow ϕ_F^t which is generated by the following function,

$$\begin{aligned} F &= F_0(\theta,\varepsilon) + F_1(\theta,z,\varepsilon) + F_2(\theta,z,\varepsilon) \\ &= \sum_{0<|k|\le K} f_{k0}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle} + \sum_{0<|k|\le K} \langle f_{k1}, z\rangle \mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle} + \sum_{0<|k|\le K} \langle f_{k2}z, z\rangle \mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle}, \end{aligned}$$

where $f_{kj},\ j=0,1,2,\ 0<|k|\le K$, are (vector-valued or matrix-valued) coefficients which will be determined later. Since that

$$\begin{aligned} H\circ\phi_F^1 &= N\circ\phi_F^1 + h\circ\phi_F^1 + \varepsilon(G + \sum_{j=0}^{2}\bar{E}_j)\circ\phi_F^1 + (\varepsilon\sum_{j=0}^{2}\hat{E}_j)\circ\phi_F^1 \\ &= N + \tilde{h} + \varepsilon\tilde{G} + \{\tilde{h}_{\ge 4}, F_1\} + \{\tilde{h}_, F_2\} + (\{N,F\} + \varepsilon\sum_{j=0}^{2}\bar{E}_j + \{\tilde{h}_{=3}, F_1\}) + \tilde{e} \\ &\quad +\varepsilon\int_0^1 \{\tilde{G} + \sum_{j=0}^{2}\bar{E}_j, F\}\circ\phi_F^t \mathrm{d}t + (\varepsilon\hat{E}_2)\circ\phi_F^1 \\ &\quad + \int_0^1 \{\{(1-t)(N+\tilde{h}), F\}, F\}\circ\phi_F^t \mathrm{d}t. \end{aligned}$$

where $N := \langle\omega, I\rangle + \langle\tilde{M}z, z\rangle$, $\{\cdot,\cdot\}$ denotes the Poisson bracket, $\tilde{h}_{=3} := \sum_{i=1}^n h_{i,3}x_i^3$ denotes the third order terms in $\tilde{h}$ and $\tilde{h}_{\ge 4}$ denotes the terms in the fourth order or high than the fourth order. Firstly, we solve the following quasi-homological equation

$$\{N, F_0\} + \varepsilon\bar{E}_0 = 0, \tag{2.10}$$

$$\{N, F_1\} + \varepsilon\bar{E}_1 = 0, \tag{2.11}$$

$$\{N, F_2\} + \varepsilon\bar{E}_2 + \{\tilde{h}_{=3}, F_1\} = 0. \tag{2.12}$$

Substitute $N,\ F,\ \bar{E}_j,\ j=0,1,2$, into equations (2.10)-(2.12), we obtain that

$$\sqrt{-1}\langle k,\omega\rangle f_{k0} - \varepsilon E_{k0} = 0, \tag{2.13}$$

$$\sqrt{-1}\langle k,\omega\rangle I_{2n} - \tilde{M}(\varepsilon)Jf_{k1} - \varepsilon E_{k1} = 0, \tag{2.14}$$

$$\sqrt{-1}\langle k,\omega\rangle f_{k2} + \tilde{M}(\varepsilon)Jf_{k2} - f_{k2}J\tilde{M}(\varepsilon) - \varepsilon E_{k2} - E_{k3} = 0, \tag{2.15}$$

where

$$E_{k3} := \mathrm{diag}\{3h_{1,3}f_{k1,n+1}, \cdots, 3h_{n,3}f_{k1,2n}, 0, \cdots, 0\},$$

$f_{k1,j},\ j=1,\cdots,2n$, denotes the j-th components of f_{k1}, J is the $2n\times 2n$ standard symplectic matrix and $\tilde{M}$ is defined as in (2.6). Denote that

$$\begin{aligned} L_{k0} &:= \sqrt{-1}\langle k,\omega\rangle, \\ L_{k1} &:= \sqrt{-1}\langle k,\omega\rangle I_{2n} - \tilde{M}(\varepsilon)J, \\ L_{k2} &:= \sqrt{-1}\langle k,\omega\rangle I_{4n^2} - \tilde{M}(\varepsilon)J\otimes I_{2n} - I_{2n}\otimes\tilde{M}(\varepsilon)J, \end{aligned}$$

where $\otimes$ denotes the Tensor product.

Define a positive constant to simplify the notations in the following estimates, that is

$$C_\eta = \sum_{0<|k|\le K} \mathrm{e}^{-\frac{\eta}{4}|k|}|k|^{(N+1)(4n^2\tau+4n-1)} < +\infty.$$

Consider homological equation (2.13), we have for any $0<|k|\le K,\ i=0,1,\cdots,N$ that

$$f_{k0} = \varepsilon L_{k0}^{-1}E_{k0}, \qquad |\partial_\varepsilon^i f_{k0}| \le \varepsilon|\partial_\varepsilon^i E_{k0}|\frac{|k|^\tau}{\gamma} \le c\varepsilon|k|^\tau \mathrm{e}^{-|k|(r-\eta)}. \tag{2.16}$$

It follows that

$$\begin{aligned}(2.17)\qquad \|\partial_\varepsilon^i F_0(\theta,\varepsilon)\|_{D(r-5\eta/4)\times\mathcal{D}} &\leq \varepsilon \sum_{0<|k|\leq K} |L_{k0}^{-1}||E_{k0}|\mathrm{e}^{|k|(r-5\eta/4)} \\ &\leq c\varepsilon \sum_{0<|k|\leq K} |k|^\tau \mathrm{e}^{-\frac{\eta|k|}{4}} \leq \varepsilon c C_\eta.\end{aligned}$$

Let

$$\varepsilon^{a_m/2}K^\tau = o(\gamma), \tag{2.18}$$

where $a_m = \min\{a_1, \cdots, a_n\}$. Now we prove operators $L_{k1},\ L_{k2}$ are invertible. Denote that

$$L_{k1}^0 := \sqrt{-1}\langle k,\omega\rangle I_{2n} - A(\varepsilon)J,$$
$$L_{k2}^0 := \sqrt{-1}\langle k,\omega\rangle I_{4n^2} - A(\varepsilon)J\otimes I_{2n} - I_{2n}\otimes A(\varepsilon)J,$$

where $A(\varepsilon)$ is defined as in (2.6). It is easy to calculate

$$|\mathrm{det} L_{k1}^0| = \prod_i^n |\mathcal{K}^2 - \varepsilon^{a_i} m_i(\varepsilon)|,$$

where $\mathcal{K} = \sqrt{-1}\langle k,\omega\rangle$. Based on (2.18), we have that for any $0<|k|\leq K$, $i=1,\cdots,n$

$$|\mathcal{K}^2 - \varepsilon_{m_i}^{a_i}| \geq |\mathcal{K}^2||1-\varepsilon^{a_i} m_i |k|^{2\tau}/\gamma^2| \geq \frac{\gamma^2}{2|k|^{2\tau}}.$$

It follows that

$$|\mathrm{det} L_{k1}^0| \geq c\frac{\gamma^{2n}}{|k|^{2\tau}}.$$

Since that $L_{1k} = L_{1k}^0 + \varepsilon\tilde{A}(\varepsilon)$, it follows from (2.18) that

$$|\mathrm{det} L_{k1}|_{\mathcal{D}} \geq |\mathrm{det} L_{k1}^0|(1 - c(\varepsilon^{\alpha_m/2}|k|^\tau/\gamma) - \cdots - c(\varepsilon^{\alpha_m/2}|k|^\tau/\gamma)^{2n}) \geq c\frac{\gamma^{2n}}{|k|^{2n\tau}},$$

Similarly, we have that

$$\det L_{k2}^0 = \prod_{i,j=1}^n |\mathcal{K}^2 \pm \sqrt{\varepsilon^{a_i} m_i(\varepsilon)} \mp \sqrt{\varepsilon^{a_j} m_j(\varepsilon)}| \geq c\frac{\gamma^{4n^2}}{|k|^{4n^2\tau}}.$$

It follows that

$$\begin{aligned}|\mathrm{det} L_{k2}|_{\mathcal{D}} &\geq |\mathrm{det} L_{k2}^0|(1 - c\varepsilon^{\alpha_{n-1}/2}|k|^\tau/\gamma - \cdots - c(\varepsilon^{\alpha_{n-1}/2}|k|^\tau/\gamma)^{4n^2}) \\ &\geq c\frac{\gamma^{4n^2}}{|k|^{4n^2\tau}},\end{aligned}$$

Since that

$$L_{kq}^{-1} = \frac{\mathrm{adj} L_{kq}}{\det L_{kq}},$$

where $\mathrm{adj}L_{kq}$ denotes the adjoint matrix of L_{kq}, $q=1,2$. Then we have that

$$|L_{k1}^{-1}|_{\mathcal{D}} \leq c\frac{|k|^{2n\tau+2n-1}}{\gamma^{2n}},$$

$$|L_{k2}^{-1}|_{\mathcal{D}} \leq c\frac{|k|^{4n^2\tau+4n^2-1}}{\gamma^{4n^2}}.$$

Together with the following formula,

$$\partial_\varepsilon^i L_{kq}^{-1} = -\sum_{i'=1}^i C_i^{i'}(\partial_\varepsilon^{i-i'} L_{kq}^{-1}\partial_\varepsilon^{i'} L_{kq})L_{kq}^{-1}, \quad i=0,1,\cdots,N,\ q=1,2,$$

there exists a positive constant c such that for $q=1,2,\ i=0,1,\cdots,N$, we have

$$|\partial_\varepsilon^i L_{kq}^{-1}|_{\mathcal{D}} \le c\frac{|k|^{(i+1)((2n)^q\tau+(2n)^q-1)}}{\gamma^{(i+1)(2n)^q}}. \tag{2.19}$$

The estimate (2.19) yields that equations (2.11), (2.12) are uniquely solvable for any $\varepsilon\in\mathcal{D}$, $0<|k|\le K$ and there exists a positive constant c_0 such that the following estimates hold:

$$\begin{aligned}
&\|\partial_\varepsilon^i \mathrm{D}^j F(\theta,z,\varepsilon)\|_{D(r-3\eta/2,s-3\eta/2)\times\mathcal{D}} \le \varepsilon c_0(C_\eta+C_\eta^2), \quad i=0,1,\cdots,N,\ j=0,1,2,\\
&|\partial_\varepsilon^i \mathrm{D}(\phi_F^t-id)|_{D(r-3\eta/2,s-3\eta/2)\times\mathcal{D}} \le \varepsilon c_0(C_\eta+C_\eta^2), \quad t\in[0,1],\ i=0,1,\cdots,N.
\end{aligned} \tag{2.20}$$

By taking ε_* sufficiently small such that

$$\varepsilon^{1-a}c_0(C_\eta+C_\eta^2)\le\eta,$$

it yields from standard arguments in the proof of KAM-type theorems that the transformation

$$\Phi_{0,\varepsilon}:=\phi_F^1: D(r-2\eta,s-2\eta)\times\mathcal{D}\to D(r-\eta,s-\eta)\times\mathcal{D}$$

is well defined. As a consequence, we obtain the new Hamiltonian as follows

$$\begin{aligned}
H_0 &:= H\circ\Phi_{0,\varepsilon}\\
&= \langle\omega,I\rangle+\langle M_0(\varepsilon)z,z\rangle+h_0(z,\varepsilon)+\varepsilon^a G_0(z,\varepsilon,\theta)+\varepsilon^{2a}P_0(z,\varepsilon,\theta)+e_0(\varepsilon).
\end{aligned} \tag{2.21}$$

where $M_0:=\tilde{M}$, $h_0:=\tilde{h}$, $e_0:=\varepsilon[E_0]+\tilde{e}$ and

$$G_0 = \varepsilon^{-a+1}\tilde{G}+\varepsilon^{-a}\{h,F_2\}+\varepsilon^{-a}\{h_{\ge4},F_1\}, \tag{2.22}$$

$$\begin{aligned}
P_0 &= \varepsilon^{-2a+1}\int_0^1\{\tilde{G}+\sum_{j=0}^2\bar{E}_j,F\}\circ\phi_F^t\mathrm{d}t+(\varepsilon^{-2a+1}\hat{E}_2)\circ\phi_F^1\\
&\quad+\varepsilon^{-2a}\int_0^1\{\{(1-t)(N+\tilde{h}),F\},F\}\circ\phi_F^t\mathrm{d}t.
\end{aligned} \tag{2.23}$$

Based on estimates (2.9), (2.18) and (2.20), we obtain that there exists a positive constant c_1 depending on constants $n,\ d,\ \gamma,\ s,\ r,\ \eta$ such that for $i=0,1,\cdots,N,\ j=0,1,2$, the following estimates hold:

$$\begin{aligned}
\|\partial_\varepsilon^i\mathrm{D}^jG_0\|_{D(r-2\eta,\ s-2\eta)\times\mathcal{D}} &\le c(C_\eta+C_\eta^2)\varepsilon^{1-a}\le c_1\varepsilon^{1-a},\\
\|\partial_\varepsilon^i\mathrm{D}^jP_0\|_{D(r-2\eta,\ s-2\eta)\times\mathcal{D}} &\le c(C_\eta+C_\eta^2)^2\varepsilon^{2-2a}\le c_1\varepsilon^{2-2a}.
\end{aligned}$$

Since that ε is sufficiently small, we also have

$$|M_0^{-1}| \le |(A+\varepsilon\tilde{A}(\varepsilon))^{-1}|\le\frac{|\tilde{A}^{-1}|}{1-\varepsilon|\tilde{A}^{-1}||\tilde{A}|}\le\frac{\varepsilon^{-a}}{1-O(\varepsilon_*)}\le c_1\varepsilon^{-a}.$$

□

3. Improve the order of perturbation

Consider Hamiltonian (2.21) on a new domain $(\theta,z)\in D(r_0,s_0)$, $\varepsilon\in\mathcal{D}$, where $r_0:=r-2\eta$, $s_0:=\varepsilon^{\frac{2-2a-\iota}{3}}\ll s-2\eta$ for fixed $0<\iota<2-2a$. Let $\mu_0:=\varepsilon^{\frac{2-2a}{3}}$, $\gamma_0:=\gamma^{4n^2(N+1)}$, where γ is the Diophantine constant. The estimate of the perturbation P_0 can be rewritten as

$$\|\partial_\varepsilon^i\mathrm{D}^jP_0\|_{D(r_0,s_0)\times\mathcal{D}}\le\gamma_0 s_0^2\mu_0, \qquad i=0,1,\cdots,N,\ j=0,1,2.$$

Note that the gap parameter γ_0 and iterative parameter μ_0 are much bigger than ε^a. It means that the perturbation is not small enough for the convergence of measure estimate. As a consequence, we apply a finite number of averaging process to further improve the order of perturbation till it is high enough for usual KAM iteration step can be directly conducted. Since we do not average out the first degree terms in P_0, the perturbation can not be push up to the order of $O(\varepsilon^{4a})$ directly. Instead of that, we sharply shrink the domain z to ensure the new perturbation become much smaller at each iterative step.

3.1. **One circle of KAM step.** Suppose that we have arrived at the ν-th step and obtained the following real analytic Hamiltonian,

$$H = \langle \omega, I\rangle + \langle z, Mz\rangle + h(z,\varepsilon) + \varepsilon G(\theta, z, \varepsilon) + \varepsilon^{2a} P(\theta, z, \varepsilon), \tag{3.1}$$

which is defined on a phase domain $(\theta, z) \in D(r, s)$ and depending smoothly on $\varepsilon \in \mathcal{D}$. Since that the Hamiltonian vector field X_H is corresponding to (θ, I, z), we omit the constant term during the KAM process. In addition, we have that M is nonsingular and symmetry for each $\varepsilon \in \mathcal{D}$ and satisfies

$$\|\partial_\varepsilon^i (M - M_0)\|_{\mathcal{D}} \le \varepsilon \mu^{\frac{1}{4}}, \qquad i = 0, 1, \cdots, N. \tag{3.2}$$

The functions $h(z, \varepsilon),\ G(\theta, z, \varepsilon) = O(|z|^3)$ and

$$\|\partial_\varepsilon^i \mathrm{D}^j P\|_{D(r,s)\times\mathcal{D}} \le \gamma_0 s^2 \mu, \qquad i = 0, 1, \cdots, N,\ j = 0, 1, 2$$

for some $0 < \mu \ll \mu_0$, $0 < s \ll s_0$. We try to find a canonical transformation $\Phi_+ :\ D(r_+, s_+)\times\mathcal{D} \to D(r, s) \times \mathcal{D}$, which transforms the Hamiltonian (3.1) into the following form

$$H_+ := H \circ \Phi_+ = \langle \omega, I\rangle + \langle z, M_+ z\rangle + h_+(z, \varepsilon) + \varepsilon^a G_+(\theta, z, \varepsilon) + \varepsilon^{2a} P_+(\theta, z, \varepsilon),$$

where the matrix M_+, the functions h_+ , G_+ are in the same forms as $M,\ h,\ G$, respectively. The new perturbation P_+ is much smaller than P on some smaller domains, that is,

$$\|\partial_\varepsilon^i \mathrm{D}^j P_+\|_{D(r_+,s_+)\times\mathcal{D}} \le \gamma_0 s_+^2 \mu_+, \qquad i = 0, 1, \cdots, N,\ j = 0, 1, 2,$$

for some $r_+ \le r$, $s_+ \ll s$, $\mu_+ \ll \mu$. The normal form reduction Proposition states as follows.

Proposition 3.1. *Consider the Hamiltonian (2.21) in $D(r_0, s_0) \times \mathcal{D}$ and assume ε_* is sufficiently small. Then there exists a C^N-smooth family of real analytic transformations $\Phi_* : D(r_*,\ s_*)\times\mathcal{D} \to D(r_0,\ s_0) \times \mathcal{D}$, where $r_*,\ s_*$ are positive constant depending on $r_0,\ s_0$ that will be specific later. Under this transformation, Hamiltonian (2.21) can be transformed as follows*

$$H_* = H_0 \circ \Phi_* = \langle \omega, I\rangle + \langle z, M_* z\rangle + h_*(z, \varepsilon) + G_*(\theta, z, \varepsilon) + P_*(\theta, z, \varepsilon), \tag{3.3}$$

where M_* is a nonsingular symmetric matrix with $\|M_*^{-1}\|_{\mathcal{D}} = O(\varepsilon^{-a})$, the function $h_*,\ G_* = O(|z|^3)$ and the following estimates hold

$$\|\partial_\varepsilon^i (M_* - M_0)\|_{\mathcal{D}} \le \varepsilon^a \mu_0^{\frac{3}{4}}, \quad i = 0, 1, \cdots, N,$$
$$\|\partial_\varepsilon^i \mathrm{D}^j P_*\|_{D(r_*,s_*)\times\mathcal{D}} \le \gamma_*^{3(N+1)} s_*^2 \mu_*^3,\ i = 0, 1, \cdots, N,\ j = 0, 1, 2,$$

where $\gamma_* = \varepsilon^{2n^2-n}$, $\mu_* = \varepsilon^{2a}$.

We mention that, for simplicity, we have omitted the subscript ν and use '+' to denote subscript $\nu + 1$ in (3.1) and in the following proof. We will also use c_i, c to denote any positive intermediate constants which are independent of $\varepsilon,\ \mu,\ \nu$ during the iteration process. Define

$$\begin{aligned}
r_+ &= \frac{r}{2}+\frac{r_0}{4},\\
s_+ &= \frac{1}{4}\alpha s,\ \alpha=\mu^{\frac{1}{3}},\\
\mu_+ &= \mu^{7/6},\\
K_+ &= ([\log\frac{1}{\mu}]+1)^3,\\
D_{i\alpha} &= D(r_+ + \frac{i-1}{4}(r-r_+), \frac{i}{4}\alpha s),\ i=1,2,3,4\\
D(s) &= \{z\in\mathbb{C}^{2n}: |z|<s,\ s>0\},\\
\Gamma(r-r_+) &= \sum_{0<|k|\le K_+}|k|^{(N+1)(4n^2\tau+4n-1)}\mathrm{e}^{-|k|\frac{r-r_+}{4}}.
\end{aligned}$$

Firstly, we write P in the Taylor-Fourier series and let R be the truncation, that is

$$P=\sum_{k\in\mathbb{Z}^d,\imath\in Z_+^{2n}} p_{k\imath}z^{\imath}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},$$

$$R=\sum_{|k|\le K_+}(p_{k0}+\langle p_{k1},z\rangle+\langle z,p_{k2}z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle}, \tag{3.4}$$

where K_+ is defined as above.

Lemma 3.1. *Assume that*

H1) $\displaystyle\int_{K_+}^{\infty} t^{d+1}e^{-t\frac{r-r_+}{16}}\mathrm{d}t\le\mu.$

Then, there is a constant c_1 depending on $n,\ d,\ r$ such that

$$\begin{aligned}
\|\partial_\varepsilon^i(P-R)\|_{D_{4\alpha}\times\mathcal{D}} &\le c_1\mathcal{C}\gamma_0 s^2\mu^2,\\
\|\partial_\varepsilon^i R\|_{D_{4\alpha}\times\mathcal{D}} &\le \mathcal{C}\gamma_0 s^2\mu.
\end{aligned}$$

Proof. See [16] for the proof. □

Now we rewrite $R:=R_0+R_1+R_2$, where

$$R_0=\sum_{|k|\le K_+}p_{k0}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad R_1=\sum_{|k|\le K_+}\langle p_{k1},z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\quad R_2=\sum_{|k|\le K_+}\langle z,p_{k2}z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle}.$$

We aim to eliminate R by introducing a canonical transformation ϕ_F^1 which is the time-1 map of the Hamiltonian flow generated by a function $F:=F_0+F_1+F_2$ of the following form,

$$\begin{aligned}
F_0 &= \sum_{0<|k|\le K_+}f_{k0}\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\\
F_1 &= \sum_{0\le|k|\le K_+}\langle f_{k1},z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle},\\
F_2 &= \sum_{0<|k|\le K_+}\langle z,f_{k2}z\rangle\mathrm{e}^{\sqrt{-1}\langle k,\theta\rangle}.
\end{aligned}$$

Since that

$$\begin{aligned} H\circ\phi_F^1 &= N+\varepsilon^{2a}[R_2]+h+\varepsilon^aG+q+\varepsilon^{2a}[R_0]+\{N,F\}+\varepsilon^{2a}(R-[R_0]-[R_2]) \\ &\quad +\{h,F_1+F_2\}+\int_0^1(1-t)\{\{N+h,F\},F\}\circ\phi_F^t\mathrm{d}t \\ &\quad +\int_0^1\{\varepsilon^aG+\varepsilon^{2a}R,F\}\circ\phi_F^t\mathrm{d}t+\varepsilon^{2a}(P-R)\circ\phi_F^1, \end{aligned}$$

where $N:=\langle\omega,I\rangle+\langle z,Mz\rangle$. We determine F by solving homological equation

$$\{N,F\}+\varepsilon^{2a}(R-[R_0]-[R_2])=0. \tag{3.5}$$

Substitute $N,\ F,\ R$ into equation (3.5), we obtain the following equations by comparing the coefficients

$$\sqrt{-1}\langle k,\omega\rangle f_{k0}=p_{k0},\qquad 0<|k|\le K_+, \tag{3.6}$$

$$(\sqrt{-1}\langle k,\omega\rangle I_{2n}-MJ)f_{k1}=p_{k1},\qquad 0<|k|\le K_+, \tag{3.7}$$

$$\sqrt{-1}\langle k,\omega\rangle f_{k2}+MJf_{k2}-f_{k2}JM=p_{k2},\qquad 0<|k|\le K_+, \tag{3.8}$$

$$Mf_{01}=-p_{01}. \tag{3.9}$$

Denote that

$$\begin{aligned} L_{k1} &:= \sqrt{-1}\langle k,\omega\rangle I_{2n}-MJ, \\ L_{k2} &:= \sqrt{-1}\langle k,\omega\rangle I_{4n^2}-MJ\otimes I_{2n}-I_{2n}\otimes MJ, \end{aligned}$$

we have the following lemma.

Lemma 3.2. *Assume that*

H2) $\varepsilon^{a_m/2}K_+^\tau=o(\gamma)$,

where $a_m=\min\{a_1,\cdots,a_n\}$, γ *is the Diophantine constant. Then for* $0<|k|\le K_+,\ \varepsilon\in(0,\varepsilon_*)$, *the operators* $L_{k1},\ L_{k2}$ *and matrix* M *are invertible. Moreover, there exists a positive constant* c_2 *such that following estimate holds,*

$$|\partial_\varepsilon^i L_{kq}^{-1}|_{\mathcal{D}}\le c_2\frac{|k|^{(i+1)((2n)^q+(2n)^q-1)}}{\gamma^{(i+1)(2n)^q}},\quad i=0,1,\cdots,N,\ q=1,2. \tag{3.10}$$

Proof. The proof of estimates (3.10) are the same as the proof of (2.19). Moreover, we have that that

$$\begin{aligned} \|M^{-1}\|_{\mathcal{D}} &= \|(I+(M-M_0+\varepsilon\tilde{A}))^{-1}A^{-1}\|_{\mathcal{D}} \\ &\le \frac{\|A^{-1}\|}{1-\|A_0^{-1}\|\|M-M_0+\varepsilon\tilde{A}\|}\le c\varepsilon^{-a}. \end{aligned} \tag{3.11}$$

□

It follows from Lemma 3.2 that equations (3.7)-(3.9) are uniquely solvable for $|k|<K_+$ and $\varepsilon\in\mathcal{D}$ and there exists a positive constant c_3 such that

$$\begin{aligned} &|\partial_\varepsilon^i f_{01}|_{\mathcal{D}}\le c_3\varepsilon^a\gamma_0 s\mu, \\ &|\partial_\varepsilon^i f_{kj}|_{\mathcal{D}}\le c_3\varepsilon^{2a}\gamma_0^{-(i+1)(2n)^j}s^{2-j}\mu\mathrm{e}^{-\frac{r-r_+}{4}},\quad k\ne0 \\ &\|\partial_\varepsilon^i\mathrm{D}^j[F_1]\|_{D_{3\alpha}\times\mathcal{D}}\le c_3\varepsilon^a\gamma_0 s^{2-j}\mu, \\ &\|\partial_\varepsilon^i\mathrm{D}^j(F_0+F_1-[F_1]+F_2)\|_{D_{3\alpha}\times\mathcal{D}}\le c_3\varepsilon^{2a}\gamma_0^{-(i+1)4n^2}s^{2-j}\mu\Gamma(r-r_+), \\ &\|\partial_\varepsilon^i\mathrm{D}^jF\|_{D_{3\alpha}\times\mathcal{D}}\le c_3(\varepsilon^a\gamma_0s^{2-j}\mu+\varepsilon^{2a}\gamma_0^{-(i+1)4n^2}s^{2-j}\mu\Gamma(r-r_+), \end{aligned} \tag{3.12}$$

where $i=0,\cdots,N,\ j=0,1,2$ and $0<|k|\le K_+$.

Lemma 3.3. *Suppose that the following assumptions hold,*

- **H3)** $c_3\mu\Gamma(r-r_+)+c_3\mu<\frac{1}{4}(r-r_+)$;
- **H4)** $c_3s\mu\Gamma(r-r_+)+c_3s\mu<s_+$.

Let ϕ_F^t be the flow generated by F. We have that

1) *For all $0\le t\le 1$, $\phi_F^t:D_\alpha\to D_{4\alpha}$ are well defined for $\varepsilon\in\mathcal{D}$.*
2) *Let $\Phi_+=\phi_F^1$. Then for all $\varepsilon\in\mathcal{D}$,*
$$\Phi_+:D_+\to D.$$
3) *There is a constant c_3 such that*
$$|\phi_F^1-id|_{D_\alpha\times\mathcal{D}}\le c_3(\varepsilon^a\gamma_0 s\mu+\varepsilon^{2a}s\mu\Gamma(r-r_+)),$$
$$|\mathrm{D}\Phi_+-Id|_{D_\alpha\times\mathcal{D}}\le c_3(\varepsilon^a\gamma_0\mu+\varepsilon^{2a}\mu\Gamma(r-r_+)),$$
for all $0\le t\le 1$.

Omitting the constant term, we arrived at the new Hamiltonian in the following form
$$H_+:=H\circ\phi_F^1=\langle\omega,I\rangle+\langle z,M_+z\rangle+h_++\varepsilon^aG_++\varepsilon^{2a}P_+,$$
where
$$M_+ := M+\frac{\partial^2\{h_{=3},\bar{F}_1\}}{\partial z^2}+\varepsilon^{2a}[R_2], \tag{3.13}$$
$$h_+ := h+\{h_{\ge 4},[F_1]\},$$
$$G_+ := G,$$
$$P_+ := \varepsilon^{-2a}\{h,F_1-[F_1]+F_2\}+\int_0^1(1-t)\{\{N+h,F\},F\}\circ\phi_F^t\mathrm{d}t \tag{3.14}$$
$$+\varepsilon\int_0^1\{\varepsilon^{-a}G+R,F\}\circ\phi_F^t\mathrm{d}t+\varepsilon^{2a}(P-R)\circ\phi_F^1,$$
where $h_{=3}$ is the three degree term in h and $h_{\ge4}:=h-h_3$. It is obvious that there exists c_4 depending on $c_1,\ c_3$ such that
$$\|\partial_\varepsilon^i(M_+-M)\|_{\mathcal{D}}\le c_4\varepsilon^a\gamma_0 s\mu\le\varepsilon\mu^{\frac{1}{4}},\quad i=0,1,\cdots,N, \tag{3.15}$$
$$\|\partial_\varepsilon^i(h_+-h)\|_{\mathcal{D}}\le c_4\varepsilon^a\gamma_0 s\mu\le\varepsilon\mu^{\frac{1}{4}},\quad i=0,1,\cdots,N,$$
by assuming that μ is sufficiently small. For the new perturbation P_+, we have the following estimate.

Lemma 3.4. *There exists a constant c_5 such that*
$$\|\partial_\varepsilon^i\mathrm{D}^jP_+\|_{D_+\times\mathcal{D}}\le c_5(s^3\mu\Gamma(r-r_+)+s^3\mu^2\Gamma^2(r-r_+)+\gamma_0s^2\mu^2\Gamma(r-r_+)+\gamma_0s^2\mu^2), \tag{3.16}$$
for $i=0,1,\cdots,N$, $j=0,1,2$. Consequently, if

- **H5)** $c_5(s^3\mu\Gamma(r-r_+)+s^3\mu^2\Gamma^2(r-r_+)+\gamma_0s^2\mu^2\Gamma(r-r_+)+\gamma_0s^2\mu^2)\le\gamma_0s_+^2\mu_+$,

then
$$\|\partial_\varepsilon^i\mathrm{D}^jP_+\|_{D_+\times\mathcal{D}}\le\gamma_0s_+^2\mu_+. \tag{3.17}$$

Proof. The proof follows easily from the expression of P_+ as (3.14) and the estimates of F as in (3.12). Moreover, Lemma 3.1-Lemma 3.4 complete one cycle of KAM iteration. □

3.2. **Proof of Theorem 3.1.** Recursively applying the definitions of quantities at the very beginning of subsection 3.1, we have the following iterative sequences

$$\begin{aligned} r_\nu &= r_0(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}), \\ s_\nu &= \frac{1}{4}\alpha_{\nu-1}s_{\nu-1},\ \alpha_\nu=\mu_\nu^{\frac{1}{3}}, \\ \mu_\nu &= \mu_{\nu-1}^{\frac{7}{6}}, \\ K_\nu &= ([\log(\frac{1}{\mu_{\nu-1}})]+1)^3 \end{aligned}$$

for $\nu=1,2,\cdots$. It is easy to deduce that

$$r_\nu - r_{\nu+1} = \frac{r_0}{2^{\nu+2}}, \quad \mu_\nu=\mu_0^{(\frac{7}{6})^\nu}=\varepsilon^{(\frac{1-a-\bar{b}}{3})(\frac{7}{6})^\nu}\le \varepsilon^{\frac{1-b}{3}(\frac{7}{6})^\nu} \quad \nu=1,2,\cdots, \tag{3.18}$$

from which the hypotheses H1), H3)-H5) can be verified for all $\nu=1,2,\cdots$ as μ is sufficiently small. However, H2) only holds for a finite number of ν's. More precisely, we define

$$\nu_* = \left[\frac{\ln(9(2n^2-n)(N+1)+18a)-\ln(2-2a)}{\ln 7/6}\right]+1, \tag{3.19}$$

where $[\cdot]$ denotes the maximum integer less than x. As long as

$$\varepsilon[(\log\frac{1}{\varepsilon})(\frac{7}{6})^{\nu_*}+1]^3 \ll \gamma,$$

the assumption H2) holds for all $\nu=1,2,\cdots,\nu_*$. By repeating the iterative process inductively, we have obtained a sequence of Hamiltonian

$$H^\nu = H^{\nu-1}\circ\Phi^\nu = \langle\omega,I\rangle+\langle z,M_\nu(\omega)z\rangle+h_\nu+\varepsilon^a G_\nu+\varepsilon^{2a}P_\nu(\theta,z,\varepsilon)$$

defined on $D(r_\nu,s_\nu)\times\mathcal{D}$ for all $\nu=1,2,\cdots,\nu_*$. Define that $\Phi_*:=\Phi_0\circ\cdots\circ\Phi_{\nu_*-1}$, we obtain the following Hamiltonian

$$H_* = H\circ\Phi_* = \langle\omega,I\rangle+\langle z,M_*z\rangle+h_*(z,\varepsilon)+G_*(\theta,z,\varepsilon)+P_*(\theta,z,\varepsilon)$$

defined on $D(r_*,s_*)\times\mathcal{D}$, where $r_*=r_{\nu_*},\ s_*=s_{\nu_*}, M_*=M_{\nu_*}, h_*=h_{\nu_*},\ G_*=\varepsilon^a G_{\nu_*}, P_*=\varepsilon^{2a}P_{\nu_*}$. Based on (3.15), we have that for $i=0,1,\cdots,N$

$$\|\partial_\varepsilon^0(M_*-M_0)\|_{\mathcal{D}}\le c_0\varepsilon^a(\mu_{\nu_*-1}+\mu_{\nu_*-2}+\cdots+\mu_0)\le\varepsilon^a\mu_0^{3/4},$$

which guarantees that $\|M_*\|^{-1}=O(\varepsilon^{-a})$. Moreover, it follows form (3.18) and (3.19) that

$$\mu_{\nu_*}=\mu_0^{(\frac{7}{6})^\nu}\le\varepsilon^{3(2n^2-n)(N+1)+6a},$$

It yields for $i=0,1,\cdots,N$, that

$$\|\partial_\varepsilon^i P^*\|_{D(r_*,s_*)\times\mathcal{D}}\le\varepsilon^{2a}\gamma_0 s_{\nu_*}^2\mu_{\nu_*}\le s_*^2\varepsilon^{3(2n^2-n)(N+1)+6a}\le\gamma_*^{3(N+1)}s_*^2\mu_*^3, \tag{3.20}$$

by denoting $\gamma_*:=\varepsilon^{2n^2-n},\ \mu_*:=\varepsilon^{2a},\ s_*:=s_{\nu_*}$.

4. Infinite steps of KAM iteration

Since we have pushed the perturbation to a sufficiently high order such that we can take $\varepsilon\in\mathcal{D}$ as a normal parameter and directly apply an infinite steps of classical KAM theorem to prove the persistence of the d-tori for most of $\varepsilon\in\mathcal{D}$. In order to make the iteration processes simpler, we consider the following re-scale transformation,

$$I\to\gamma_*^2\mu_*^2 I,\quad z\to\gamma_*\mu_* z,\quad H_*\to\frac{H_*}{\gamma_*^2\mu_*^2}$$

to the normal form (3.3). Then the re-scaled Hamiltonian reads

$$H^0 := \frac{H_*}{\gamma_*^2\mu_*^2} := \langle \omega, I\rangle + \langle z, M^0(\varepsilon)z\rangle + P^0(\theta, z, \varepsilon) \tag{4.1}$$

defined on new region $D(r_0, s_0) \times \mathcal{D}$, where $r_0 := r_*,\ s_0 := s_*,\ O_0 = \mathcal{D} = (0, \varepsilon_*),\ M^0 := M_*$ being non-singular matrix on $\mathcal{D}$ with $|(M^0)^{-1}| = O(\varepsilon^{-a})$. Moreover,

$$P^0 = \frac{P_* + h_* + G_*}{\varepsilon^2\gamma_*^2\mu_*^2}.$$

It follows from (3.20) that

$$|\partial_\varepsilon^i P^0|_{D(r_0,s_0)\times\mathcal{D}} \le \frac{\|\partial_\varepsilon^i P_*\|_{D_*\times\mathcal{D}}}{\varepsilon^4\gamma_*^2\mu_*^2} \le \gamma_0^{N+1}s_0^2\mu_0,$$

where $\gamma_0 := \gamma_* = \varepsilon^{2n^2-n},\ \mu_0 := \mu_* = \varepsilon^{2a},\ i = 0, 1, \cdots, N.$

Remark 4.1. *Without great loose of generality, we still use $r_0,\ s_0$ to denote the domain parameters, $\gamma_0,\ \mu_0$ to denote the gap parameter and iterative parameter, respectively. These four parameters and the corresponding sequences are not related to the ones in Section 3. We also mention that, after re-normalization by finite steps of averaging process, the gap parameter γ_0 becomes much smaller that the constant γ in Diophantine condition* **A1)***.*

4.1. **Iteration and convergence.**

Consider the following sequences

$$\begin{aligned}
&r_\nu = r_0(1 - \sum_{i=1}^{\nu}\frac{1}{2^{i+1}}),\\
&s_\nu = \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},\\
&\alpha_\nu = \mu_\nu^{\frac{1}{3}},\\
&\mu_\nu = c_0\mu_{\nu-1}^{\frac{6}{5}},\\
&\gamma_\nu = \gamma_0(1 - \sum_{i=1}^{\nu}\frac{1}{2^{i+1}}),\\
&K_\nu = ([\log(\frac{1}{\mu_{\nu-1}})] + 1)^{3\eta},\\
&L_{1k,\nu-1} = \sqrt{-1}\langle k, \omega\rangle I_{2n} - M_{\nu-1}J, \quad 0 < |k| \le K_\nu,\\
&L_{2k,\nu-1} = \sqrt{-1}\langle k, \omega\rangle I_{4n^2} - (M_{\nu-1}J)\otimes I_{2n} - I_{2n}\otimes(M_{\nu-1}J), \quad 0 < |k| \le K_\nu,\\
&O_\nu = \{\xi \in O_{\nu-1} :\ |\det L_{1k,\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^{2n\tau}},\ |\det L_{2k,\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^{4n^2\tau}},\ 0 < |k| \le K_\nu\},
\end{aligned}$$

$\nu = 1, 2, \cdots$, where $\eta \ge \frac{\log 2}{\log 6 - \log 5}$ is a fixed constant. The following iteration lemma and convergence result are special cases of those iteration lemma in [4], [28].

Lemma 4.1. *Let μ_0 be sufficiently small. Then the followings hold for all $\nu = 1, 2, \cdots$.*

1) *There is a sequence of Whitney smooth family of symplectic, real analytic, near identity transformations*
$$\Phi^\nu : D(r_\nu, s_\nu) \to D(r_{\nu-1}, s_{\nu-1}),\ \varepsilon \in O_\nu$$
such that
$$H^\nu \quad = \quad H^{\nu-1}\circ\Phi^\nu =: \langle \omega, I\rangle + \langle z, M^\nu z\rangle + P^\nu(\theta, z, \varepsilon),$$

where

$$\|\partial_\varepsilon^i M^\nu - \partial_\varepsilon^i M^0\|_{O_\nu} \le \gamma_0^{N+1}\mu_0^{\frac{1}{2}}, \tag{4.2}$$
$$\|\partial_\varepsilon^i \mathrm{D}^j P^\nu\|_{D_\nu\times O_\nu} \le \gamma_\nu^{N+1} s_\nu^2 \mu_\nu$$

for all $i = 0, 1, \cdots, N$.

2) $O_\nu = \{\omega \in O_{\nu-1}: \ |\det L_{1k,\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^{2n\tau}}, \ |\det L_{2k,\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^{4n^2\tau}}, \ K_{\nu-1} < |k| \le K_\nu\}$.

3) *The Whitney extensions of*

$$\Psi^\nu =: \Phi_\omega^1 \circ \Phi_\omega^2 \circ \cdots \circ \Phi_\omega^\nu$$

converge C^1 *uniformly to a smooth family of symplectic maps, that is,* Ψ^∞*, on* $D(\frac{r_0}{2}, \frac{s_0}{2}) \times O_\infty$*, where*

$$O_\infty = \bigcap_{\nu \ge 0} O_\nu,$$

such that

$$H^\nu = H^0 \circ \Psi^{\nu-1} \to H^\infty =: H^0 \circ \Psi^\infty = \langle \omega, I\rangle + \langle z, M^\infty z\rangle + P^\infty(\theta, z, \varepsilon)$$

with $M^\infty = \lim_{\nu\to\infty} M^\nu$*,* $P^\infty = \lim_{\nu\to\infty} P^\nu$*, and*

$$\|\mathrm{D}^j P^\infty\|_{D(\frac{r_0}{2}, \frac{s_0}{2})\times O_\infty} = 0, \qquad |j| \le 2.$$

Now we suppose that O_∞ is not empty. Remind the transformations $\Phi_{0,\varepsilon}$ and Φ_* in Lemma 2.3 and Proposition 3.1, respectively. Define $\bar\Phi^\infty := \Phi_{0,\varepsilon} \circ \Phi_* \circ \Phi^\infty$, it follows that

$$\phi_H^t \circ \bar\Phi^\infty|_{\mathbb{T}^d\times\mathbb{R}^{2n}} = \bar\Phi^\infty \circ \phi_{H_\infty}^t|_{\mathbb{T}^d\times\mathbb{R}^{2n}}$$

where ϕ_H^t and $\phi_{H^\infty}^t$ are the flow of H defined in (3.1) and H^∞ is defined as above. Define $T^{d,0} = \{\omega\} \times \{I = 0\} \times \{z = 0\}$, for any ε_∞, it yields that

$$\phi_H^t \circ \bar\Phi^\infty(T^{d,0}) = \bar\Phi^\infty \circ (\phi_{H^\infty}^t(T^{d,0})) = \bar\Phi^\infty(T^{d,0}),$$

which means the embedding tori $\Phi^\infty(T^{d,0})$ is invariant under the flow $\phi_H^t|_{\mathbb{T}^d\times\mathbb{R}^{2n}}$ with the fixed frequency ω, that is, for $\varepsilon \in O_\infty$, $\Phi^\infty(T^{d,0})$ forms a C^N (Whitney) smooth family of invariant tori with fixed frequency ω for Hamiltonian normal form (1.1).

Remark 4.2. *Based on assumption* **A3)***, there exists an energy function in form of Hamiltonian (1.1) such that the lower-dimensional, response invariant tori of Hamiltonian (1.1) also form the quasi-periodic response solutions of the motion equation (1.2), which prove the Main Theorem as well as Corollary 1.*

5. Measure estimate

For each $\nu = 0, 1, \cdots$ and $k \in \mathbb{Z}^n \setminus \{0\}$, denote

$$R_k^{\nu+1} = R_{k,1}^{\nu+1} \bigcup R_{k,2}^{\nu+1},$$

where

$$R_{k,1}^{\nu+1} = \{\varepsilon \in O_\nu: \ |\det L_{1k,\nu}| \le \frac{\gamma_\nu}{|k|^{2n\tau}}, \ K_\nu < |k| \le K_{\nu+1}\},$$
$$R_{k,2}^{\nu+1} = \{\varepsilon \in O_\nu: \ |\det L_{2k,\nu}| \le \frac{\gamma_\nu}{|k|^{4n^2\tau}}, \ K_\nu < |k| \le K_{\nu+1}\}.$$

By Lemma 4.1, we obtain that

$$O_0 \setminus O_\infty = \bigcup_{\nu=0}^{\infty} \bigcup_{K_\nu < |k| \le K_{\nu+1}} R_k^{\nu+1}. \tag{5.1}$$

In the following, we will prove that the O_∞ is almost full with respect to O_0 in the mixed type and it is equal to O_0 in the hyperbolic type. Before measure estimate, we introduce the following lemmas.

Lemma 5.1. ([28, Lemma 2.1]) *Suppose that $g(x)$ is a differentiable function on the closure $\bar{I} \subset I$, where I is a finite open interval. Let $I_h = \{x : \ |g(x)| \le h, \ x \in I\}, \ h > 0$. If $x \in I$, $|\frac{\mathrm{d}g(x)}{\mathrm{d}x}| \ge D > 0$, where D is a constant, then $|I_h| \le 2hD^{-1}$.*

Lemma 5.2. *Assume that M is a $2n \times 2n$ symmetric matrix, then*

$$\det(\lambda I_{2n} - MJ) = P_{2n}$$
$$\det(\lambda I_{4n^2} - I_{2n} \otimes (MJ) - (MJ) \otimes I_{2n}) = \lambda^{2n} P_{4n^2-2n},$$

where P_j is a j-degree, even polynomial function with respect to variable λ.

Proof. Since that M is a symmetric matrix and J is a standard symplectic matrix, it yields that

$$\det(\lambda I_{2n} - MJ) = \det[J(\lambda I_{2n} - (MJ))J^{-1}] = \det(\lambda I_{2n} - JM)$$

and

$$\begin{aligned} & \det(-\lambda I_{2n} - MJ) = (-1)^{2n} \det(\lambda I_{2n} + MJ) = \det(\lambda I_{2n} - MJ^\top) \\ = \ & \det[(\lambda I_{2n} - MJ^\top)^\top] = \det(\lambda I_{2n} - JM^\top) = \det(\lambda I_{2n} - JM^\top). \end{aligned}$$

It shows that $\det(\lambda I_{2n} - MJ) = \det(-\lambda I_{2n} - MJ)$, that is, $\det(\lambda I_{2n} - MJ)$ is a $2n$-degree even polynomial function with respect to λ. By the following properties of Kronecker product of matrices $A, \ B, \ C, \ D$ in the same size and constant c,

$$(cA) \otimes B = c(A \otimes B), \quad (A \otimes B) = A^\top \otimes B^\top, \quad (AB) \otimes (CD) = (A \otimes C)(B \otimes D),$$

it is easy to prove $\det(\lambda I_{4n^2} - I_{2n} \otimes (MJ) - (MJ) \otimes I_{2n})$ is an even polynomial function with respect to variable λ.

Moreover, the eigenvalues of matrix $A \otimes I + I \otimes A$ can be formulated as $\mu_{ij} = \lambda_i + \lambda_j$, where λ_i are eigenvalues of A. Since $\det(\lambda I_{2n} - MJ)$ is an even function with respect to λ, the eigenvalues of MJ can be expressed as $\pm\lambda_1, \ \cdots, \pm\lambda_n$. It follows that matrix $I_{2n} \otimes (MJ) + (MJ) \otimes I_{2n}$ has at least $2n$ zero eigenvalues, hence we have that

$$\det(\lambda I_{4n^2} - I_{2n} \otimes (MJ) - (MJ) \otimes I_{2n}) = \lambda^{2n} P_{4n^2-2n}.$$

□

Remark 5.1. *Denote $\mathcal{K} = \sqrt{-1}\langle k, \omega\rangle$. Since that M^ν is a symmetric matrix for $\nu = 1, 2, \cdots$, it directly implies that*

$$\det L_{1k,\nu} := P_{2n}, \qquad \det L_{2k,\nu} = \mathcal{K}^{2n} P_{4n^2-2n}, \tag{5.2}$$

where P_j denotes a j-degree polynomial function with respect to $\mathcal{K}$.

5.1. **Measure estimate for mixed type.**

Lemma 5.3. *In the mixed type, the remaining set O_∞ is almost full Lebesgue measure satisfying that*

$$\frac{|\text{meas } O_\infty|}{\varepsilon_*} = 1 - O(\varepsilon_*^{1-\sigma}),$$

where $0 < \varepsilon_* \ll 1$ is defined as in Lemma 2.3 and Proposition 3.1 and $\sigma := \min\{\frac{1}{l_1-1}, \ \cdots, \frac{1}{l_n-1}\}$.

Proof. Remind the estimates (3.15) and (4.2), it follows that

$$\|M^\nu - A\|_{O_\nu} \le \|M_0 - A\|_{O_\nu} + \|M_* - M_0\|_{O_\nu} + \|M^\nu - M^0\|_{O_\nu} \le \varepsilon,$$

where

$$A(\varepsilon) = \mathrm{diag}\{\varepsilon^{a_1} m_1(\varepsilon),\ \cdots,\ \varepsilon^{a_n} m_n,\ 1, \cdots,\ 1\},$$

and $a_i,\ m_i,\ i = 1, 2, \cdots, n$, are defined as in (2.6).

For fixed ν and $K_\nu < |k| \le K_{\nu+1}$, we obtain that

$$\det L_{1k,\nu} := P_{2n} = \mathcal{K}^{2n} + \alpha_1^0 \mathcal{K}^{2n-2} + \alpha_2^0 \mathcal{K}^{2n-4} + \cdots + \alpha_{n-1}^0 \mathcal{K}^2 + \alpha_n^0,$$

where

$$\tag{5.3} \begin{aligned} \alpha_1^0 &:= \sum_i \varepsilon^{a_i} m_i + O(\varepsilon), \\ \alpha_2^0 &:= \sum_{i,j} \varepsilon^{a_i + a_j} m_i m_j + O(\varepsilon^{1+a_n}), \\ \alpha_3^0 &:= \sum_{i,j,k} \varepsilon^{a_i + a_j + a_k} m_i m_j m_k + O(\varepsilon^{1+a_{n-1}+a_n}), \\ &\ \ \vdots \\ \alpha_n^0 &:= \varepsilon^{a_1 + \cdots + a_n} m_1 \cdots m_n + O(\varepsilon^{1+a_2+\cdots+a_n}). \end{aligned}$$

Hereafter, we use $\mathrm{d}_\varepsilon f(\varepsilon)$ to denote $\frac{\mathrm{d}f(\varepsilon)}{\mathrm{d}\varepsilon}$ for simplicity, where $f(\varepsilon)$ is a function only depending on ε. Then we define the polynomial functions with respect to $\mathcal{K}$ as follows:

$$\begin{aligned} P_{2n-2} &:= \frac{\mathrm{d}_\varepsilon P_{2n}}{\mathrm{d}_\varepsilon a_1^0} := \mathcal{K}^{2n-2} + \alpha_1^1 \mathcal{K}^{2n-4} + \alpha_2^1 \mathcal{K}^{2n-6} + \cdots + \alpha_{n-1}^1, \\ P_{2n-4} &:= \frac{\mathrm{d}_\varepsilon P_{2n-2}}{\mathrm{d}_\varepsilon a_1^1} := \mathcal{K}^{2n-4} + \alpha_1^2 \mathcal{K}^{2n-6} + \alpha_2^2 \mathcal{K}^{2n-8} + \cdots \alpha_{n-3}^2 \mathcal{K}^2 + \alpha_{n-2}^2, \\ &\ \ \vdots \\ P_{2n-2j} &:= \frac{\mathrm{d}_\varepsilon P_{2n-2(j-1)}}{\mathrm{d}_\varepsilon \alpha_1^{j-1}} := \mathcal{K}^{2n-2j} + \alpha_1^j \mathcal{K}^{2n-2j-2} + \cdots + \alpha_{n-j}^j, \\ &\ \ \vdots \\ P_2 &:= \frac{\mathrm{d}_\varepsilon P_{2n-2(n-2)}}{\mathrm{d}_\varepsilon \alpha_1^{n-2}} := \mathcal{K}^2 + \alpha_1^{n-1}, \end{aligned}$$

where, for fixed $j = 1, 2, \cdots, n-1$, the coefficients of polynomial function P_{2n-2j} satisfy the following inductive formula

$$\alpha_i^j := \frac{\mathrm{d}_\varepsilon \alpha_{i+1}^{j-1}}{\mathrm{d}_\varepsilon \alpha_1^{j-1}}, \quad i = 1, 2, \cdots, n-j.$$

Based on the discussion in Appendix, for any $\varepsilon \in (0, \varepsilon_*]$, there exists a positive constant c_* depending on a_i and the norm of $|m_i|$, $i = 1, 2, \cdots, n$, such that

$$\tag{5.4} |\mathrm{d}_\varepsilon \alpha_1^j| \ge c\varepsilon^{a_{n-j}-1} |m_{n-j}| \ge c_* \varepsilon^{\sigma - 1}, \quad j = 0, 1, \cdots, n-1.$$

Define that

$$R_2 := \{\varepsilon \in O_\nu :\ |P_2| \le \frac{\varepsilon}{|k|^{2\tau}}, \quad K_\nu < |k| \le K_{\nu+1}\},$$

based on Lemma 5.1, we have that

$$|\mathrm{meas}\ R_2| \le \frac{\varepsilon^{2-\sigma}}{c_*} \sum_{K_\nu < |k| \le K_{\nu+1}} \frac{1}{|k|^{2\tau}},$$

where $|\cdot|$ denote the measure of the set. Now we define the following sets for $j = 1, 2, \cdots, n-1$

$$R_{2j} = \{\varepsilon \in O_\nu : \ |P_{2j}| \leq \frac{\varepsilon^j}{|k|^{2\tau j}}, \quad K_\nu < |k| \leq K_{\nu+1}\}.$$

Assume that for fixed $1 \leq j_0 < n-1$, we have obtained the measure estimate of R_{j_0}, that is

$$|\text{meas } R_{j_0}| \leq \frac{\varepsilon^{2-a_1} + \cdots + \varepsilon^{2-a_{j_0}}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}} \leq \frac{j_0 \varepsilon^{2-\sigma}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}}$$

Then we define a new set

$$\tilde{R}_{2j_0+2} = \{\varepsilon \in O_\nu \setminus R_{j_0} : \ |P_{2j_0+2}| \leq \frac{\varepsilon^{j_0+1}}{|k|^{(2j_0+2)\tau}} \quad K_\nu < |k| \leq K_{\nu+1}\}.$$

Since that for $\varepsilon \in O_\nu \setminus R_{j_0}$, we have

$$|\mathrm{d}_\varepsilon P_{2j_0+2}| = |\mathrm{d}_\varepsilon \alpha_1^{n-(j_0+1)}||P_{2j_0}| \geq c_* \varepsilon^{a_{j_0+1}-1+j_0} \frac{1}{|k|^{2j_0\tau}},$$

it follows from Lemma 5.1 that

$$\begin{aligned}
|\text{meas } R_{2j_0+2}| &\leq |\text{meas } \tilde{R}_{2j_0+2}| + |\text{meas } R_{2j_0}| \\
&\leq \frac{\varepsilon^{2-a_{j_0+1}}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}} + \frac{\varepsilon^{2-a_1} + \cdots + \varepsilon^{2-a_{j_0}}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}} \\
&\leq \frac{(j_0+1)\varepsilon^{2-\sigma}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}}.
\end{aligned}$$

By the Mathematical inductive method, we have that

$$|R_{2n-2}| \leq \frac{(n-1)\varepsilon^{2-\sigma}}{c_*} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{2\tau}},$$

where

$$R_{2n-2} = \{\varepsilon \in O_\nu : \ |P_{2n-2}| \leq \frac{\varepsilon^{n-1}}{|k|^{(2n-2)\tau}}, \quad K_\nu < |k| \leq K_{\nu+1}\}.$$

Since that for $\varepsilon \in O_\nu \setminus R_{2n-2}$, we have

$$|\mathrm{d}_\varepsilon P_{2n}| = |\mathrm{d}_\varepsilon \alpha_1^0||P_{2n-2}| \geq c_* \varepsilon^{a_n - 1 + n - 1} \frac{1}{|k|^{(2n-2)\tau}}.$$

Remind

$$R_{k,1}^{\nu+1} = \{\varepsilon \in O_\nu : \ |P_{2n}| \leq \frac{\gamma_\nu}{|k|^{2n\tau}}, \ K_\nu < |k| \leq K_{\nu+1}\},$$

where $\gamma_\nu < \gamma_0 < \varepsilon^{2n^2-n}\gamma^{2n}$, γ is the Diophantine constant. It follows that

$$|\text{meas } \cup_{\nu=0}^\infty \cup_{K_\nu < |k| \leq K_{\nu+1}} R_{k,1}^{\nu+1}| \leq \frac{n\varepsilon^{2-a_1}}{c_*} \sum_{k \in \mathbb{Z}^d} \frac{1}{|k|^{2\tau}} \leq c_{*1} \varepsilon_*^{2-\sigma}, \tag{5.5}$$

where $c_{*1} := \frac{n}{c_*} \sum_{k \in \mathbb{Z}^d} \frac{1}{|k|^{2\tau}} > 0$.

Based on the same discussion for P_{4n^2-2n}, we obtain that

$$|\text{meas } \cup_{\nu=0}^\infty \cup_{K_\nu < |k| \leq K_{\nu+1}} R_{k,2}^{\nu+1}| \leq c_{*2} \varepsilon_*^{2-\sigma},$$

where c_{*2} depending on a_i, n . As all above, we prove that

$$\frac{|\text{meas } O_\infty|}{\varepsilon^*} = 1 - \frac{|\text{meas } \bigcup_{\nu=0}^\infty \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}|}{\varepsilon} = 1 - O(\varepsilon_*^{1-\sigma}).$$

□

5.2. Measure estimate for hyperbolic type.

Lemma 5.4. *In the hyperbolic type, the remaining set* $O_\infty = (0, \varepsilon_*)$*, where* $0 < \varepsilon_* \ll 1$ *is defined as in Lemma 2.3 and Proposition 3.1.*

Proof. Remind that

$$\det L_{1k,\nu} := \mathcal{K}^{2n} + \alpha_1^0 \mathcal{K}^{2n-2} + \alpha_2^0 \mathcal{K}^{2n-4} + \cdots + \alpha_{n-1}^0 \mathcal{K}^2 + \alpha_n^0,$$

where α_i^0 are defined as in (5.3) and

$$\begin{aligned} A(\varepsilon) &= \operatorname{diag}\{\varepsilon^{a_1} m_1, \ \cdots, \ \varepsilon^{a_n} m_n, \ 1, \cdots, \ 1\}, \\ m_i &= \lambda_i (l_i - 1)(x_i^*)^{l_i - 2} + O(\varepsilon^\sigma), \qquad i = 1, 2, \cdots, n. \end{aligned}$$

Firstly, $\lambda_i < 0$ guarantees $m_i < 0$. Actually, when l_i is even, $l_i - 2$ is even so that m_i and λ_i are in the same sign. When l_i is odd, $l_i - 1$ is even, which implies $(-a_i/\lambda_i)^{\frac{1}{l_i - 1}} > 0$. Since that ε is sufficiently small so that $x_{i,\varepsilon} > 0$ and $m_i < 0$. Now we prove that $R_{k,1}^{\nu+1} = R_{k,2}^{\nu+1} = \emptyset$ for fixed $\nu = 1, 2, \cdots$ and $\lambda_i < 0$, $i = 1, 2, \cdots, n$. For the case that n is even, it is easy to see that $\mathcal{K}^{2n} > 0$ and $\alpha_k^0 \mathcal{K}^{2n-2k} > 0$, $k = 1, 2, \cdots, n$. More specifically, $\alpha_k^0 > 0$, $\mathcal{K}^{2n-2k} > 0$ when k is even and $\alpha_k^0 < 0$, $\mathcal{K}^{2n-2k} < 0$ when k is odd. Otherwise, n is odd, it follow that $\mathcal{K}^{2n} < 0$ and $\alpha_k^0 \mathcal{K}^{2n-2k} < 0$, $k = 1, 2, \cdots, n$. As all above, we obtain that for all $\varepsilon \in O_\nu$,

$$|\det L_{1k,\nu}| > |\mathcal{K}^{2n}| > \frac{\gamma^{2n}}{|k|^{2n\tau}} > \frac{\gamma_\nu}{|k|^{2n\tau}}.$$

since $\gamma_\nu < \gamma_0 \ll \gamma^{2n}$. Based on the same discussion, we also have that for all $\varepsilon \in O_\nu$,

$$|\det L_{2k,\nu}| > |\mathcal{K}^{4n^2}| > \frac{\gamma^{4n^2}}{|k|^{4n^2\tau}} > \frac{\gamma_\nu}{|k|^{4n^2\tau}}.$$

It follows that for all $\nu = 1, 2, \cdots$, $O_\nu = O_\infty = (0, \varepsilon_*)$ holds. We mention that, for hyperbolic case, one can directly apply classical KAM iterations to Hamiltonian (2.21) to prove the Main theorem. Furthermore, it is obvious that measure estimate for hyperbolic type does not involve any derivatives of $\det L_{ik,\nu}$, $i = 1, 2$, with respect to ε, hence one can choose any integer $N \geq 1$ in all of the KAM iterations mentioned above which leads that the persisted tori form a C^N-smoothly family for any integer $N \geq 1$. □

6. Appendix

In this subsection, we prove estimate (5.4), which is the key point for measure estimate. Hereafter, we also use c to denote the constant independent of parameter ε. Based on Lemma 5.2. we obtain that $\det L_{1k,\nu}$ is a $2n$-th degree polynomial function with respect to $\mathcal{K}$ in the following form

$$\det L_{1k,\nu} := P_{2n} = \mathcal{K}^{2n} + \alpha_1^0 \mathcal{K}^{2n-2} + \alpha_2^0 \mathcal{K}^{2n-4} + \cdots + \alpha_{n-1}^0 \mathcal{K}^2 + \alpha_n^0, \tag{6.1}$$

where

$$\begin{aligned} \alpha_1^0 &:= \sum_i \varepsilon^{a_i} m_i + O(\varepsilon), \\ {\alpha^0}_2 &:= \sum_{i,j} \varepsilon^{a_i + a_j} m_i m_j + O(\varepsilon^{1 + a_n}), \\ {\alpha^0}_3 &:= \sum_{i,j,k} \varepsilon^{a_i + a_j + a_k} m_i m_j m_k + O(\varepsilon^{1 + a_{n-1} + a_n}), \\ &\ \vdots \\ \alpha_n^0 &:= \varepsilon^{a_1 + \cdots + a_n} m_1 \cdots m_n + O(\varepsilon^{1 + a_2 + \cdots + a_n}). \end{aligned}$$

Denote $\delta := \min\{|a_i - a_j|,\ |a_i| :\ 1 \le i, j \le n\}$ and rewrite the coefficients of P_{2n} as follows:

$$
\begin{aligned}
\alpha_1^0 &:= \varepsilon^{a_n} m_n + \varepsilon^{a_n+\delta}\tilde m_1^0(\varepsilon),\\
{\alpha^0}_2 &:= \varepsilon^{a_{n-1}+a_n} m_{n-1}m_n + \varepsilon^{a_{n-1}+a_n+\delta}\tilde m_2^0(\varepsilon),\\
{\alpha^0}_3 &:= \varepsilon^{a_{n-2}+a_{n-1}+a_n} m_{n-2}m_{n-1}m_n + \varepsilon^{a_{n-2}+a_{n-1}+a_n+\delta}\tilde m_3^0(\varepsilon),\\
&\ \vdots\\
{\alpha^0}_k &:= \varepsilon^{a_{n-k+1}+\cdots+a_n} m_{n-k+1}\cdots m_n + \varepsilon^{a_{n-k+1}+\cdots+a_n+\delta}\tilde m_k^0(\varepsilon),\\
&\ \vdots\\
\alpha_n^0 &:= \varepsilon^{a_1+\cdots+a_n} m_1\cdots m_n + \varepsilon^{a_1+a_2+\cdots+a_n+\delta}\tilde m_n^0(\varepsilon),
\end{aligned}
\tag{6.2}
$$

where

$$\tilde m_k^0 := \sum_{(i_1,\cdots i_k)\in I} \varepsilon^{a_{i_1}+\cdots+a_{i_k}-a_{n-k+1}-\cdots-a_n} m_{i_i}\cdots m_{i_k} + O(\varepsilon^{1+a_{n-k+2}-a_{n-k+1}}),$$

and $I := \{(i_1,\cdots,i_k) :\ 1 \le i_k \le n,\ (i_1,\cdots,i_k) \ne (n-k+1,\cdots,n)\}$. Based on Lemma 2.2, we obtain the following estimates for $p = 1,2,\cdots,N$, $k = 1,2,\cdots,n$, that is

$$
\begin{aligned}
&\|(\Pi_{p=0}^{k-1} m_{n-p})\| \le c, &&\|\tilde m_i^0(\varepsilon)\| \le c,\\
&\|\varepsilon^p \mathrm{d}_\varepsilon^p(\Pi_{p=0}^{k-1} m_{n-p})\| \le c\varepsilon^\delta, &&\|\varepsilon^p \mathrm{d}_\varepsilon^p \tilde m_i^0(\varepsilon)\| \le c.
\end{aligned}
$$

As above, we define the polynomial functions with respect to $\mathcal{K}$ as follows:

$$
\begin{aligned}
P_{2n-2} &:= \frac{\mathrm{d}_\varepsilon P_{2n}}{\mathrm{d}_\varepsilon a_1^0} := \mathcal{K}^{2n-2} + \alpha_1^1\mathcal{K}^{2n-4} + \alpha_2^1\mathcal{K}^{2n-6} + \cdots + \alpha_{n-1}^1,\\
P_{2n-4} &:= \frac{\mathrm{d}_\varepsilon P_{2n-2}}{\mathrm{d}_\varepsilon a_1^1} := \mathcal{K}^{2n-4} + \alpha_1^2\mathcal{K}^{2n-6} + \alpha_2^2\mathcal{K}^{2n-8} + \cdots \alpha_{n-3}^2\mathcal{K}^2 + \alpha_{n-2}^2,\\
&\cdots\\
P_{2n-2j} &:= \frac{\mathrm{d}_\varepsilon P_{2n-2(j-1)}}{\mathrm{d}_\varepsilon \alpha_1^j} := \mathcal{K}^{2n-2j} + \alpha_1^j\mathcal{K}^{2n-2j-2} + \cdots + \alpha_{n-j}^j,\\
&\cdots\\
P_2 &:= \frac{\mathrm{d}_\varepsilon P_{2n-2(n-2)}}{\mathrm{d}_\varepsilon \alpha_1^{n-2}} := \mathcal{K}^2 + \alpha_1^{n-1},
\end{aligned}
$$

where, for fixed $j = 1,2,\cdots,n-1$, the coefficients

$$\alpha_i^j := \frac{\mathrm{d}_\varepsilon \alpha_{i+1}^{j-1}}{\mathrm{d}_\varepsilon \alpha_1^{j-1}},\quad i = 1,2,\cdots,n-j. \tag{6.3}$$

Now we calculate the coefficients. Firstly, we have

$$
\begin{aligned}
|\mathrm{d}_\varepsilon \alpha_1^0| &= |a_n\varepsilon^{a_n-1}m_n + \varepsilon^{a_n}\mathrm{d}_\varepsilon m_n + \varepsilon^{a_n+\delta-1}\tilde m_1^0(\varepsilon) + \varepsilon^{a_n+\delta}\mathrm{d}_\varepsilon \tilde m_1^0(\varepsilon)|\\
&\ge a_n\varepsilon^{a_n-1}|m_n||1 - \frac{|\varepsilon \mathrm{d}_\varepsilon m_n|}{a_n|m_n|} - \varepsilon^\delta\frac{|\tilde m_1^0(\varepsilon)|}{a_n|m_n|} - \varepsilon^\delta\frac{|\varepsilon \mathrm{d}_\varepsilon \tilde m_1^0(\varepsilon)|}{a_n|m_n|}|,\\
&> \frac{a_n\varepsilon^{a_n-1}|m_n|}{2}.
\end{aligned}
\tag{6.4}
$$

Denote $\hat m^0 := (\varepsilon \mathrm{d}_\varepsilon m_n)/a_n$ and $\tilde m^0 := (\tilde m_1^0 + \varepsilon \mathrm{d}l_\varepsilon \tilde m_1^0)/a_n$, we simply rewrite $\mathrm{d}_\varepsilon \alpha_1^0$ as

$$\mathrm{d}_\varepsilon \alpha_1^0 := a_n\varepsilon^{a_n-1}(m_n + \hat m^0(\varepsilon) + \varepsilon^\delta \tilde m^0(\varepsilon)).$$

It follows that for $p = 0,1\cdots,N-1$,

$$\|\varepsilon^p \mathrm{d}_p \hat m^0\| \le c\varepsilon^\delta,\qquad \|\varepsilon^p \mathrm{d}_p \tilde m^0\| \le c. \tag{6.5}$$

By the inductive formula (6.3), we obtain the coefficients of P_{2n-2} as follows:

$$
\begin{aligned}
\alpha_1^1 &= \frac{\mathrm{d}_\varepsilon \alpha_2^0}{\mathrm{d}_\varepsilon \alpha_1^0} = \frac{c_1^1 \varepsilon^{a_{n-1}} m_{n-1} m_n + \varepsilon^{a_{n-1}} \hat{m}_1^1 + \varepsilon^{a_{n-1}+\delta} \tilde{m}_1^1}{m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0},\\
\alpha_2^1 &= \frac{\mathrm{d}_\varepsilon \alpha_3^0}{\mathrm{d}_\varepsilon \alpha_1^0} = \frac{c_2^1 \varepsilon^{a_{n-2}+a_{n-1}} m_{n-2} m_{n-1} m_n + \varepsilon^{a_{n-2}+a_{n-1}} \hat{m}_2^1 + \varepsilon^{a_{n-2}+a_{n-1}+\delta} \tilde{m}_2^1}{m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0},\\
&\vdots\\
\alpha_k^1 &= \frac{\mathrm{d}_\varepsilon \alpha_{k+1}^0}{\mathrm{d}_\varepsilon \alpha_1^0}\\
&= \frac{c_k^1 \varepsilon^{\sum_{p=0}^{k-1} a_{n-1-p}} ((\Pi_{p=0}^{k-1} m_{n-1-p}) m_n + \varepsilon^{\sum_{p=0}^{k-1} a_{n-1-p}} \hat{m}_k^1 + \varepsilon^{\sum_{p=0}^{k-1} a_{n-1-p}+\delta} \tilde{m}_k^1)}{m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0},\\
&\vdots\\
\alpha_{n-1}^1 &= \frac{\mathrm{d}_\varepsilon \alpha_n^0}{\mathrm{d}_\varepsilon \alpha_1^0}\\
&= \frac{c_{n-1}^1 \varepsilon^{\sum_{p=0}^{n-2} a_{n-1-p}} ((\Pi_{p=0}^{n-2} m_{n-1-p}) m_n + \hat{m}_{n-1}^1 + \varepsilon^\delta \tilde{m}_{n-1}^1)}{m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0}.
\end{aligned}
$$

where

$$
c_k^1 := \frac{a_{n-k} + \cdots + a_n}{a_n}, \qquad \hat{m}_k^1 := \varepsilon \mathrm{d}_\varepsilon (m_{n-k} \cdots m_n). \tag{6.6}
$$

Since that $|m_n + \varepsilon \mathrm{d}_\varepsilon m_n + \varepsilon^\delta \tilde{m}_1^0(\varepsilon)| > 0$, the coefficients α_k^1 are well defined. Moreover, we have that for $p = 0, 1, \cdots, N-1$,

$$
\|\varepsilon^p \mathrm{d}_\varepsilon^p \hat{m}_k^1\| \le c\varepsilon^\delta, \qquad \|\varepsilon^p \mathrm{d}_\varepsilon^p \tilde{m}_k^1(\varepsilon)\| \le c.
$$

Then, we calculate the derivative of α_1^1, that is

$$
\mathrm{d}_\varepsilon \alpha_1^1 = \frac{c_1^1 a_{n-1} \varepsilon^{a_{n-1}-1} (m_{n-1} m_n^2 + \hat{m}^1(\varepsilon) + \varepsilon^\delta \tilde{m}^1(\varepsilon))}{(m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0)^2}, \tag{6.7}
$$

where

$$
\begin{aligned}
\hat{m}^1 \quad := \quad & m_{n-1} m_n \hat{m}^0 + (\varepsilon \mathrm{d}_\varepsilon m_{n-1} m_n) m_n / a_{n-1} + (\varepsilon \mathrm{d} m_{n-1} m_n) \hat{m}^0 / a_{n-1}\\
& + \hat{m}_1^1 \hat{m}^0 / c_1^1 + (\varepsilon(\mathrm{d}_\varepsilon m_{n-1} m_n) + \varepsilon^2 (\mathrm{d}_\varepsilon^2 m_{n-1} m_n))(m_n + \hat{m}^0)/c_1^1\\
& - m_{n-1} m_n (\varepsilon \mathrm{d}_\varepsilon m_n)/a_{n-1} + \hat{m}_1^1 m_n / c_1^1 - m_{n-1} m_n (\varepsilon \mathrm{d}_\varepsilon m_n)/(a_{n-1} a_n)\\
& - m_{n-1} m_n (\varepsilon^2 \mathrm{d}_\varepsilon^2 m_n)/(a_{n-1} a_n) - \hat{m}_1^1 \varepsilon \mathrm{d}_\varepsilon m_n - \hat{m}_1^1 (\varepsilon \mathrm{d}_\varepsilon \hat{m}^0),\\
\tilde{m}^1 \quad := \quad & m_{n-1} m_n \tilde{m}^0 + \varepsilon \mathrm{d}_\varepsilon m_{n-1} m_n \tilde{m}^0 / a_{n-1} + \hat{m}_1^1 \hat{m}^0 / c_1^1\\
& - \hat{m}_1^1 (\hat{m}^0 + \varepsilon \mathrm{d}_\varepsilon \tilde{m}^0) + (\varepsilon \mathrm{d}_\varepsilon \tilde{m}_1^1 + (a_{n-1} + \delta))(m_n + \hat{m}^0 + \varepsilon^\delta \tilde{m}^0)/c_1^1 a_{n-1}\\
& - m_{n-1} m_n (\tilde{m}^0 + (\varepsilon \mathrm{d}_\varepsilon \tilde{m}^0))/a_{n-1} - \tilde{m}_1^1 (\varepsilon \partial_\varepsilon m_n + \varepsilon \mathrm{d}_\varepsilon \hat{m}^0 + \varepsilon^\delta \tilde{m}^0 + \varepsilon^{\delta+1} \mathrm{d}_\varepsilon \tilde{m}^0).
\end{aligned}
$$

It is obvious that for $p = 0, 1, \cdots, N-2$, we have

$$
\|\varepsilon^p \mathrm{d}_\varepsilon^p \hat{m}^1\| \le c\varepsilon^\delta, \qquad \|\varepsilon^p \mathrm{d}_\varepsilon^p \tilde{m}^1(\varepsilon)\| \le c.
$$

It follows that

$$
|\mathrm{d}_\varepsilon \alpha_1^1| \ge c_1^1 a_{n-1} \varepsilon^{a_{n-1}-1} |m_{n-1}|/2 > 0. \tag{6.8}
$$

Based on (6.7) and inductive formula (6.3), we obtain the following calculation results:

$$\alpha_1^2 = \frac{\mathrm{d}_\varepsilon \alpha_2^1}{\mathrm{d}_\varepsilon \alpha_1^1} = \frac{c_1^2 \varepsilon^{a_{n-2}} m_{n-2} m_{n-1} m_n^2 + \varepsilon^{a_{n-2}} \hat{m}_1^2(\varepsilon) + \varepsilon^{a_{n-2}+\delta} \tilde{m}_1^2(\varepsilon)}{m_{n-1} m_n^2 + \hat{m}^1 + \varepsilon^\delta \tilde{m}^1},$$
$$\alpha_2^2 = \frac{\mathrm{d}_\varepsilon \alpha_3^1}{\mathrm{d}_\varepsilon \alpha_1^1} = \frac{c_2^2 \varepsilon^{a_{n-3}+a_{n-2}} (m_{n-3} m_{n-2} m_{n-1} m_n^2 + \hat{m}_2^2 + \varepsilon^\delta \tilde{m}_2^2)}{m_{n-1} m_n^2 + \hat{m}^1 + \varepsilon^\delta \tilde{m}^1},$$
$$\vdots$$
$$\alpha_k^2 = \frac{\mathrm{d}_\varepsilon \alpha_{k+1}^1}{\mathrm{d}_\varepsilon \alpha_1^1} = \frac{c_k^2 \varepsilon^{\sum_{p=0}^{k-1} a_{n-2-p}} (\Pi_{p=0}^{k-1} m_{n-2-p} m_{n-1} m_n^2 + \hat{m}_k^2 + \varepsilon^\delta \tilde{m}_k^2)}{m_{n-1} m_n^2 + \hat{m}^1 + \varepsilon^\delta \tilde{m}^1},$$
$$\vdots$$
$$\alpha_{n-2}^2 = \frac{\mathrm{d}_\varepsilon \alpha_{n-1}^1}{\mathrm{d}_\varepsilon \alpha_1^1} = \frac{c_{n-2}^2 \varepsilon^{\sum_{p=0}^{n-3} a_{n-2-p}} (\Pi_{p=0}^{k-1} m_{n-2-p} m_{n-1} m_n^2 + \hat{m}_{n-1}^2 + \varepsilon^\delta \tilde{m}_{n-1}^1)}{m_{n-1} m_n^2 + \hat{m}^1 + \varepsilon^\delta \tilde{m}^1},$$

where, for $k = 1, 2, \cdots, n-2$, we have

$$|\varepsilon^p \mathrm{d}_\varepsilon^p \hat{m}_k^2| \le c\varepsilon^\delta, \ |\varepsilon^p \mathrm{d}_\varepsilon^p \tilde{m}_k^2| \le c, \quad p = 0, 1, \cdots, N-2,$$

and

$$c_k^2 := \frac{c_{k+1}^1 \sum_{p=0}^k a_{n-1-p}}{c_1^1 a_{n-1}} > 0.$$

Now assume that we have calculated out the coefficients of P_{2n-2j} for $j = 1, 2, \cdots, \nu$ and obtain the estimate

$$|\mathrm{d}_\varepsilon \alpha_1^j| \ge c_1^j a_{n-j} \varepsilon^{a_{n-j}-1} |m_{n-j}|/2 > c\varepsilon^{\sigma-1}, \tag{6.9}$$

where $c_1^j := \frac{c_2^{j-1} \sum_{p=0}^1 a_{n-j+1-p}}{c_2^{j-1} a_{n-j+1}}$ for $j = 1, 2, \cdots, \nu-1$, c_2^1 is defined in (6.6). Write the coefficients of polynomial function $P_{2n-2(\nu-1)}$ as follows

$$\alpha_1^\nu = \frac{c_1^\nu \varepsilon^{a_{n-\nu}} m_{n-\nu} \Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \varepsilon^{a_{n-\nu}} \hat{m}_1^\nu(\varepsilon) + \varepsilon^{a_{n-\nu}+\delta} \tilde{m}_1^\nu}{\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^\delta \tilde{m}^{\nu-1}},$$
$$\alpha_2^\nu = \frac{c_2^\nu \varepsilon^{a_{n-\nu-1}+a_{n-\nu}} (m_{n-\nu-1} m_{n-\nu} \Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}_2^\nu + \varepsilon^\delta \tilde{m}_2^\nu)}{\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^\delta \tilde{m}^{\nu-1}},$$
$$\vdots$$
$$\alpha_k^\nu = \frac{c_k^\nu \varepsilon^{\sum_{p=0}^{k-1} a_{n-\nu-p}} (\Pi_{p=0}^{k-1} m_{n-\nu-p} \Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}_k^\nu + \varepsilon^\delta \tilde{m}_k^\nu)}{\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^\delta \tilde{m}^{\nu-1}},$$
$$\vdots$$
$$\alpha_{n-\nu}^\nu = \frac{c_{n-\nu}^\nu \varepsilon^{\sum_{p=0}^{n-\nu-1} a_{n-\nu-p}} (\Pi_{p=0}^{n-\nu-1} m_{n-\nu-p} \Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}_{n-\nu}^\nu + \varepsilon^\delta \tilde{m}_{n-\nu}^\nu)}{\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^\delta \tilde{m}^{\nu-1}},$$

where, the terms $\hat{m}^{\nu-1}$, $\tilde{m}^{\nu-1}$, $\hat{m}_k^\nu$, $\tilde{m}_k^\nu$ satisfy that

$$\|\varepsilon^p \mathrm{d}_\varepsilon \hat{m}^{\nu-1}\| \le c\varepsilon^\delta, \quad \|\varepsilon^p \mathrm{d}_\varepsilon \tilde{m}^{\nu-1}\| \le c, \quad p = 0, 1, 2, \cdots, N-\nu,$$
$$\|\varepsilon^p \mathrm{d}_\varepsilon \hat{m}_k^\nu\| \le c\varepsilon^\delta, \quad \|\varepsilon^p \mathrm{d}_\varepsilon \tilde{m}_k^\nu\| \le c, \quad p = 0, 1, 2, \cdots, N-\nu, \quad k = 1, 2, \cdots, n-\nu,$$

and

$$c_k^\nu := \frac{c_{k+1}^{\nu-1} \sum_{p=0}^k a_{n-\nu+1-p}}{c_1^{\nu-1} a_{n-\nu+1}}, \quad k = 1, 2, \cdots, n-\nu.$$

Since that $|\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^{\delta}\tilde{m}^{\nu-1}| > 0$, the coefficients α_k^{ν}, $k = 1, 2, \cdots, n-\nu$, are well defined. For the next step, we calculate the derivative of α_1^{ν} as follows:

$$\mathrm{d}_{\varepsilon}\alpha_1^{\nu} = \frac{c_1^{\nu} a_{n-\nu}\varepsilon^{a_{n-\nu}-1}(m_{n-\nu}\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-p}} + \hat{m}^{\nu}(\varepsilon) + \varepsilon^{\delta}\tilde{m}^{\nu}(\varepsilon))}{(\Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}} + \hat{m}^{\nu-1} + \varepsilon^{\delta}\tilde{m}^{\nu-1})^2}, \tag{6.10}$$

where, $\Lambda := \Pi_{p=0}^{\nu-1} m_{n-p}^{2^{\nu-1-p}}$ and

$$\begin{aligned}
\hat{m}^{\nu} \;&:=\; \varepsilon \mathrm{d}_{\varepsilon} m_{n-\nu}\Lambda\hat{m}^{\nu-1} + \hat{m}_1^{\nu}(\Lambda + \hat{m}^{\nu-1})/c_1^{\nu} + \varepsilon\mathrm{d}_{\varepsilon}\hat{m}_1^{\nu}(\Lambda + \hat{m}^{\nu-1})/c_1^{\nu} \\
&\quad - m_{n-\nu}\Lambda(\varepsilon\mathrm{d}_{\varepsilon}\Lambda + \varepsilon\mathrm{d}_{\varepsilon}\hat{m}^{\nu-1}) - \hat{m}_1^{\nu}(\varepsilon\mathrm{d}_{\varepsilon}\Lambda + \varepsilon\mathrm{d}_{\varepsilon}\hat{m}^{\nu-1})/c_1^{\nu}, \\
\tilde{m}^{\nu} \;&:=\; (\tilde{m}_1^{\nu} + \varepsilon\mathrm{d}_{\varepsilon}\tilde{m}_1^{\nu})(\Lambda + \hat{m}^{\nu-1} + \varepsilon^{\delta}\tilde{m}^{\nu-1})/c_1^{\nu} \\
&\quad - \tilde{m}_1^{\nu}(\varepsilon\mathrm{d}_{\varepsilon}\Lambda + \varepsilon\mathrm{d}_{\varepsilon}\hat{m}^{\nu-1} + \varepsilon^{\delta}\tilde{m}^{\nu-1} + \varepsilon^{1+\delta}\mathrm{d}_{\varepsilon}\tilde{m}^{\nu-1})/c_1^{\nu},
\end{aligned}$$

It is obvious that for $p = 0, 1, \cdots, n-\nu-1$

$$\|\varepsilon^p \mathrm{d}_{\varepsilon}^p \hat{m}^{\nu}\| \le c\varepsilon^{\delta}, \qquad \|\varepsilon^p \mathrm{d}_{\varepsilon}^p \tilde{m}^{\nu}\| \le c.$$

It follows that

$$|\mathrm{d}_{\varepsilon}\alpha_1^{\nu}| \ge c_1^{\nu} a_{n-\nu}\varepsilon^{a_{n-\nu}-1}|m_{n-\nu}|/2 > c_*\varepsilon^{\sigma-1}, \tag{6.11}$$

where c_* depends on a_i and the norm of m_i, $i = 1, 2, \cdots, n$. When $\nu = n-1$, the process ends; when $\nu < n-1$, by the inductive formula (6.3), the coefficients for polynomial function $P_{2n-2(\nu+1)}$ are as follows

$$\begin{aligned}
\alpha_1^{\nu+1} \;&=\; \frac{c_1^{\nu+1}\varepsilon^{a_{n-\nu-1}} m_{n-\nu-1}(\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}_1^{\nu+1}(\varepsilon) + \varepsilon^{\delta}\tilde{m}_1^{\nu+1})}{\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}^{\nu} + \varepsilon^{\delta}\tilde{m}^{\nu}}, \\
\alpha_2^{\nu+1} \;&=\; \frac{c_2^{\nu+1}\varepsilon^{a_{n-\nu}+a_{n-\nu-1}}(m_{n-\nu}m_{n-\nu-1}\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}_2^{\nu+1} + \varepsilon^{\delta}\tilde{m}_2^{\nu+1})}{\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}^{\nu} + \varepsilon^{\delta}\tilde{m}^{\nu}}, \\
&\;\;\vdots \\
\alpha_k^{\nu+1} \;&=\; \frac{c_k^{\nu+1}\varepsilon^{\sum_{p=0}^{k-1} a_{n-\nu-1-p}}(\Pi_{p=0}^{k-1} m_{n-\nu-1-p}\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}_k^{\nu+1} + \varepsilon^{\delta}\tilde{m}_k^{\nu+1})}{\Pi_{p=0}^{\nu} m_{n-p}^{2^{\nu-p}} + \hat{m}^{\nu} + \varepsilon^{\delta}\tilde{m}^{\nu}},
\end{aligned}$$

where

$$c_k^{\nu+1} := \frac{c_{k+1}^{\nu}\sum_{p=0}^{k} a_{n-\nu-p}}{c_1^{\nu} a_{n-\nu}} > 0$$

for $k = 1, 2, \cdots, n-\nu-1$ and $\nu = 1, 2, \cdots, n-2$. Together with (6.4) and (6.8), by the Mathematical Inductive method, we obtain that estimate (6.11) holds for $\nu = 0, 1, 2, \cdots, n-1$.

Remind that

$$\det L_{2k,\nu} := \mathcal{K}^{2n} P_{4n^2-2n},$$

where

$$P_{4n^2-2n} = \mathcal{K}^{4n^2-2n} + \alpha_1^0\mathcal{K}^{4n^2-2n-2} + \alpha_2^0\mathcal{K}^{4n^2-2n-4} + \cdots + \alpha_{2n^2-n-1}^0\mathcal{K}^2 + \alpha_{2n^2-n}^0.$$

By simple calculation, we obtain the coefficients of P_{4n^2-2n} as follows:

$$\begin{aligned}
\alpha_1^0 &:= \varepsilon^{a_n} m_n + \varepsilon^{a_n+\delta}\tilde{m}_1^0,\\
\alpha_2^0 &:= \varepsilon^{2a_n} m_n^2 + \varepsilon^{2a_n+\delta}\tilde{m}_2^0,\\
&\vdots\\
\alpha_{4n-3}^0 &:= \varepsilon^{(4n-3)a_n} m_n^{4n-3} + \varepsilon^{(4n-3)a_n+\delta}\tilde{m}_{4n-3}^0,\\
&\vdots\\
\alpha_{8n-10}^0 &:= \varepsilon^{(4n-3)a_n+(4n-7)a_{n-1}} m_n^{4n-3} m_{n-1}^{4n-7} + \varepsilon^{(4n-3)a_n+(4n-7)a_{n-1}+\delta}\tilde{m}_{8n-10}^0,\\
&\vdots\\
\alpha_{4n^2-2n}^0 &:= \varepsilon^{\sum_{p=0}^{n-1}(4n-3-4p)a_{n-p}} \prod_{p=0}^{n-1} m_{n-p}^{4n-3-4p} + \varepsilon^{\sum_{p=0}^{n-1}(4n-3-4p)a_{n-p}}\tilde{m}_{2n^2-n}^0,
\end{aligned}$$

where, for fixed $k = 1, 2, \cdots, 2n^2 - n$, the reminder terms satisfy that$\|\varepsilon^p \partial_\varepsilon \tilde{m}_k^0\| \leq c$. Observing the main terms in the coefficients, they are nonzero terms and the order of ε is increasing. By the same discussion as above, we prove that there exists a positive constant c depending on a_i and the norm of m_i, $i = 1, 2, \cdots, n$, that is

$$|\alpha_1^j| \geq c\varepsilon^{\sigma-1}, \qquad j = 0, 1, \cdots, 2n^2 - n - 1.$$

Acknowledgement The authors would like to thank the editors and referees for their valuable suggestions and comments which lead to a significant improvement of the paper.

References

[1] V. I. Arnold, Small denominators and problems of stability of motion in classical mechanics, *Usp. Math. Nauk.* **18** (6) (1963), 91-192.

[2] H. Cheng, R. de la Llave and F. Wang, Response solutions to the quasi-periodically forced systems with degenerate equilibrium: a simple proof of a result of W.Si and J.Si and extensions, *Nonlinearity,* **34** (2021), 372-389

[3] H. Cheng, W. Si and J. Si, Whiskered tori for forced beam equations with multi-dimensional Liouvillean, *Journal of Dynamics and Differential Equations* **32** (2020), 705-739

[4] L. Chierchia, G.Gallavotti, Drift and diffusion phase space, *Ann. Inst. H. Poincaré Phy. Th.* **69** 1994, 1-144.

[5] L. Corsi and G.Gentile G, Resonant motions in the presence of degeneracies for quasi-periodically perturbed systems *Ergod. Theor. Dyn. Syst.* **35** 2015, 1079C140.

[6] L.Corsi and G.Gentile, Resonant tori of arbitrary codimension for quasi-periodically forced systems, *Nonlinear Differ. Equ. Appl.* **24** 2017 Artical 3

[7] J.Du, Arnold-type therom about lower-dimensional invariant tori in generalized Hamiltonian systems, *J. Appl. Anal. Comput.,* **12(6)** 2022, 2621-2639.

[8] J. Du, L. Xu and Y. Li, An infinite dimensional KAM theorem with normal degeneracy, *Nonlinearity* **37(6)** (2024), 065021.

[9] G. Gentile, Degenearte lower-dimensional tori under the Bryuno condition, *Ergod. Th.Dynam. Sys.* **27** 2007, 427-457.

[10] G. Gentile, Quasi-periodic motions in strongly dissipative forced systems,*Ergod. Th.Dynam. Sys.* **30(5)** (2010), 1457-1469.

[11] G. Gentile, Construction of quasi-periodic responsive solutions in forced strongly dissipative systems, *Forum Math.* **24(4)** (2012), 791-808.

[12] M. Friedman, Quasi-periodic solutions of nonlinear ordinary differential equations with small damping, *Bull. Amer. Math. Soc.,* **73** (1967), 460-464.

[13] M. Gao, Quasi-periodic solusitons for 1D Nonlinear wave equation, *J. Appl. Anal. Comput.,***13(3),** (2023),1505-1534.

[14] N.S.Gopal, J. M. Jonnalagadda, Existence and Non-Existence of positive solutions for a discrete fractional boundary value roblem, *Journal of Nonlinear Modeling and Analysis,* **5(3),** (2023),432C445.

[15] X. Guan, W. Si. Almost-periodic bifucations for 2-dimensinonal degenerate Hamiltonian vector fields, *J. Appl. Anal. Comput.,* (2023), 3054-3073.

[16] Y. C. Han, Y. Li, and Y. Yi, Degenerate lower dimensional tori in Hamiltonian systems, *J. Differential Equations* **227** (2006), 670-691.
[17] Y. Li and Y. Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, *Trans. Amer. Math. Soc.* **357** (2005), 1565-1600.
[18] Z. Lou and J. Geng, Quasi-periodic response solutions in forced reversible systems with Liouvillean frequencies, *J. Differential Equations,* **263** (2017), 3894-3927.
[19] J. Moser, Combination tones for Duffings equation,*Comm. Pure Appl. Math.,* **18** (1965), 167-181.
[20] W. Qian, KAM Theorem and iso-energetic KAM theorem on Poission manifold, *J. Appl. Anal. Comput.,***13(2)** (2023), 1088-1107.
[21] W.Si and Y.Yi, Completely degenerate responsive tori in Hamiltonian systems, *nonlinerity*, **33** 2020, 6072-6098.
[22] W.Si and Y.Yi, Responsive sulotuions in degenerate oscillators under degenerate perturbation, *to appear*.
[23] J. Wang, J. You, and Q. Zhou, Response solutions for quasi-periodically forced harmonic oscillators, *Trans. Amer. Math. Soc.* **369(6)** (2017), 4251-4274.
[24] L.Xu, Y.Li and Y.Yi, Lower Dimensional Tori in Multi-scale, Nearly Integrabel Hamiltonian Systems,(2017) , **18-1**, 53-83.
[25] L. Xu, Y. Li, Y. Yi, Poincaré-Treshchev Mechanism in Multi-scale, Nearly Integrable Hamiltonian Systems *Journal of Nonlinear Science,* (2018), **28-1**, 337-369.
[26] L. Xu Y. Yi, Lower dimension tori of general types in multi-scale Hamitonian systems, *Nonlinearity,* **32,** 2019, 2226-2245.
[27] J. Xu and J. You, Corrigendum for the paper Invariant tori for nearly integrable Hamiltonian systems with degeneracy,
[28] J. Xu, J. You, and Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.,* **226** (1997), 375-387.
[29] X. Xu, W. Si and J. Si; Stokers Problem for Quasi-periodically Forced Reversible Systems with Multidimensional Liouvillean Frequency. *SIAM J. Appl. Dyn. Syst.* **19** (2020), 2286-2321. *Math. Z.* **257** (2007), 939.
[30] Y. Yi, A generalized integral manifold theorem, *J. Differential Equations* **102** (1993), 153-187.
[31] J.You, A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems Commun. Math. Phys. **192,** 1998, 145C68.
[32] Z. Yuan, S. Liu, Existence and multiplicity of solutions for a biharmonic kirchhoff equation in $\mathbb{R}^{5*}$, *Journal of Nonlinear Modeling and Analysis,* **6(1)**, (2024), 71-87.

L. Xu: School of Mathematics, Jilin University, Changchun, 130012

E-mail address: xulu@jlu.edu.cn

W. Si: School of mathematics, Shandong University, Jinan, 250100

E-mail address: siwenmath@sdu.edu.cn

M. Wu: School of Mathematics, Jilin University, Changchun, 130012

E-mail address: a2833481657@163.com