Calculation of nonperturbative part width of band matrices

Yuchen Ma, Yijian Zou, Peijun Zhu, Wenge Wang
Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China *
(Dated: November 15, 2015)

Studies have shown close relationship among width of nonperturbative parts(NPT), half width of local spectral density of states (LDOS), and half width of eigenfunctions (EFs). New approaches are developed to calculate NPT width of band matrices using generalized Brillouin-Wigner perturbation theory. Behaviors of NPT width in the limits of strong and weak perturbation are predicted by matrix truncation analysis. Iterative algorithms based on matrix transformation and path summation are developed to numerically calculate NPT width with less time complexity. Similarity of behaviors of NPT width, half width of LDOS, half width of eigenfunctions, and localization length is investigated for Wigner-band random matrices.

PACS numbers: ?

I. INTRODUCTION

Band matrices[8] are used to describe many quantum systems, such as atomic nuclei, cold atoms, and disordered systems. Local spectral density of states (LDOS), also known as strength function in nulear physics, is an important property that determines many behaviors of the system, especially in the study of relaxation process. The width and shape of LDOS play a key role to the behavior of fidelity[1–7]. Closed related to LDOS is the shape of eigenfunctions (EFs) and their half width. When coupling between near eigenstates is weak, we may use Schrodinger perturbation theory to calculate its LDOS shape. However, this method is not applicable to strongly-coupled Hamiltonian. Some progress have been made to generalize the perturbation theory to strong coupled Hamiltonian. By using generalized Brillouin-Wigner perturbation theory, we can divide every eigenstate of the perturbed Hamiltonian into perturbative part(PT) and nonperturbative part(NPT). The shape of PT region of an eigenstate, given by perturbative expansion using coefficients in the NPT region, is an exponential decay. Thus wave functions are dominant in the NPT region, so the width of NPT is closely related to half width of LDOS and half width of eigenstates. [8]

We first recall our model of band matrix and concept of NPT using generalized Brillouin-Wigner perturbation theory(GBWPT), and then develop a matrix truncation method to predict width of NPT of Wigner-band random matrices in the limits of strong and weak perturbation. Next, we introduce a general algorithm to calculate width of NPT in band matrix that has a time complexity $O(b^2N)$ with band-width denoted as b and dimension of matrix denoted as N. We verify this algorithm in two ways, through matrix transformation and path summation. Finally, we use the algorithm to calculate NPT width of Wigner-band random matrices and show its consistency with LDOS and half width of eigenstates.

*Electronic address: mayuchen@mail.ustc.edu.cn

II. THEORY AND MODEL

A. Generalized Brillouin-Wigner perturbation theory

In this section, we recall basic contents of the GBWPT. Consider a Hamiltonian of the form

$$H(\lambda) = H_0 + \lambda V, \tag{1}$$

where H_0 is an unperturbed Hamiltonian and λV represents a perturbation with a running parameter λ. Eigenstates of $H(\lambda)$ and H_0 are denoted by $|\alpha\rangle$ and $|k\rangle$, respectively,

$$H(\lambda)|\alpha\rangle = E_\alpha(\lambda)|\alpha\rangle, \quad H_0|k\rangle = E_k^0|k\rangle, \tag{2}$$

with α and k in energy order. Components of the EFs are denoted by $C_{\alpha k} = \langle k|\alpha\rangle$.

In the GBWPT, for each perturbed state $|\alpha\rangle$, the set of the unperturbed states $|k\rangle$ is divided into two substes, denoted by S_α and $\overline{S}_\alpha$. The related projection operators,

$$P_{S_\alpha} = \sum_{|k\rangle\in S_\alpha} |k\rangle\langle k|, Q_{\overline{S}_\alpha} = \sum_{|k\rangle\in\overline{S}_\alpha} |k\rangle\langle k| = 1 - P_{S_\alpha}, \tag{3}$$

divide the perturbed state into two parts, $|\alpha_s\rangle \equiv P_{S_\alpha}|\alpha\rangle$, $|\alpha_{\overline{s}}\rangle \equiv Q_{\overline{S}_\alpha}|\alpha\rangle$. As shown in Ref.[8], if the subset S_α (equivalently $\overline{S}_\alpha$) is chosen such that

$$\lim_{n\to\infty} \langle\phi|(T_\alpha^\dagger)^n T_\alpha^n|\phi\rangle = 0 \quad \forall\phi, \tag{4}$$

where

$$T_\alpha = \frac{1}{E_\alpha - H_0} Q_{\overline{S}_\alpha} \lambda V, \tag{5}$$

then, making use of the part $|\alpha_s\rangle$, the other part $|\alpha_{\overline{s}}\rangle$ can be expanded in a convergent perturbation expansion, i.e.,

$$|\alpha_{\overline{s}}\rangle = T_\alpha|\alpha_s\rangle + T_\alpha^2|\alpha_s\rangle + \cdots + T_\alpha^n|\alpha_s\rangle + \cdots. \tag{6}$$

Let us consider an operator W_α in the subspace spanned by unperturbed states $|k\rangle \in \overline{S}_\alpha$, namely,

$$W_\alpha \equiv Q_{\overline{S}_\alpha} V \frac{1}{E_\alpha - H^0} Q_{\overline{S}_\alpha}, \tag{7}$$

and use $|\nu_\alpha\rangle$ and $u_{\alpha\nu}$ to denote its eigenvectors and eigenvalues, $W_\alpha|\nu_\alpha\rangle = u_{\alpha\nu}|\nu_\alpha\rangle$. It is easy to verify that the condition (4) is equivalent to the requirement that $|\lambda u_{\alpha\nu}| < 1$ for all $|\nu_\alpha\rangle$. Expanding the state vector $Q_{\overline{S}_\alpha}\lambda V|\alpha_s\rangle$ in $|\nu_\alpha\rangle$, this gives

$$Q_{\overline{S}_\alpha}\lambda V|\alpha_s\rangle = \sum_\nu h_\nu|\nu_\alpha\rangle. \tag{8}$$

Substituting Eq.(5) and Eq.(8) into Eq.(6), after simple derivation, it is found that, for each unperturbed state $|j\rangle$ in the set $\overline{S}_\alpha$, the component $C_{\alpha j} = \langle j|\alpha\rangle$ has the following expression,

$$C_{\alpha j} = \frac{1}{E_\alpha - E_j^0}\sum_\nu \left[\frac{h_\nu}{1-\lambda u_{\alpha\nu}}\langle j|\nu_\alpha\rangle\right](\lambda u_{\alpha\nu})^{m-1}, \tag{9}$$

where m is the smallest positive integer for $\langle j|(Q_{\overline{S}_\alpha}V)^m|\alpha_s\rangle$ not equal to zero [8]. Note that, since $|\lambda u_{\alpha\nu}| < 1$, each term on the right hand side of Eq.(9) decreases exponentially with increasing m, except for $m = 1$.

In quantum chaotic systems $H(\lambda)$, all good quantum numbers of the unperturbed system H_0 have been destroyed, except that related to the energy. In the study of statistical properties of EFs in the unperturbed basis in these systems, it proves useful to consider only those sets S_α, each of which is composed of unperturbed states within a connected energy region, namely,

$$S_\alpha = \{|k\rangle : p_1 \le k \le p_2\}. \tag{10}$$

These are the sets S_α that have the simplest structure and are useful in the study of statistical properties of EFs of chaotic systems. Therefore, in the study to be given below in this paper, we consider only sets S_α of the form in Eq.(10). In fact, for an unperturbed system H_0 that has more than one good quantum numbers, one may consider S_α possessing more complex structures, but, as to be discussed later, we found that consideration of such more complex S_α does not bring information more than that by S_α in Eq.(10).

There usually exist many sets S_α for which Eq.(4) can be satisfied. Among them, the most important is the smallest one. We call the smallest set S_α, under the condition (4), the *non-perturbative (NPT) region* of the state $|\alpha\rangle$ and, correspondingly, the set $S_{\overline{\alpha}}$ the *perturbative (PT) region*. Due to the specific form in Eq.(10), the NPT region of $|\alpha\rangle$ has the smallest value of $(p_2 - p_1)$.

In the case that λ is sufficiently small and no unperturbed energy is equal to E_α, the condition (4) can be satisfied for a set S_α that includes only one unperturbed state $|k_0\rangle$ whose energy $E_{k_0}^0$ is the closest to E_α. In this case, the NPT region of $|\alpha\rangle$ contains one unperturbed state only and the expansion in Eq.(6) reduces to that of the ordinary Brillouin-Wigner perturbation theory. When the perturbation strength is increased beyond some value of λ, the condition (4) will no longer be satisfied for $S_\alpha = \{|k_0\rangle\}$ and the NPT region will contain more than one unperturbed states, usually including $|k_0\rangle$. It is easy to see that at least the average value of $(p_2 - p_1)$ should increase with increasing λ.

As an application of the GBWPT, let us consider a Hamiltonian that has a band structure, with a band width b, that is, $\langle k|V|k'\rangle = 0$ for all k and k' satisfying $|k - k'| > b$. For a NPT region S_α, from Eq.(9), it is seen that the EF decays exponentially or the like, for k increasing from $p_2 + b$ and for k decreasing from $p_1 - b$. We call the two regions $[p_1 - b, p_1]$ and $[p_2, p_2 + b]$ two *shoulders* of the NPT region. Then, the main body of the EF should lie within the region $[p_1 - b, p_2 + b]$, namely, the NPT region plus its shoulders.

Physical meanings of this decomposition are in two aspects. First, Eq.(6) shows that we can obtain all coefficients $C_{\alpha i} = \langle i|\alpha\rangle$ by GBWPT expansion if all coefficients $C_{\alpha j} = \langle j|\alpha\rangle, p_1 \le j \le p_2$ in NPT part are known. The width of NPT, defined as the dimension of Q_{min} subspace, i.e., $p_2 - p_1$,

$$N_p = p_2 - p_1, \tag{11}$$

is the number of coefficients we must know at first. Second, we can predict that $C_{\alpha i} = \langle i|\alpha\rangle$ drops exponentially as i departs from α in PT region, which implies localization of eigenstate $|\alpha\rangle$. Thus, generally, the NPT width is comparable to half width of eigenstates and half width of LDOS.

B. Band Matrix

Band matrices have nonvanishing element only in and near the diagonal line. Specifically, a matrix U is called a band matrix with band width b if all elements U_{ij} vanish when $|i-j| > b$. We study the Hamiltonian $H = H_0 + \lambda V$ whose matrix is a band matrix with band width b under representation of H_0, with known nondegenerate eigenvalues E_k^0 of H_0 arranged ascending and known corresponding eigenstates $|k\rangle$.

Of great importance among band matrices are Wigner-band random matrices. In this particular case, we have $E_i^0 = i, i = 1, 2, \cdots, N$ and $V_{ij} (i \ne j)$ as standard normal random numbers, with mean value 0 and variance 1. We can work out the expectation value of NPT width analytically to predict some properties of NPT width for Wigner-band matrices. In the following chapter, we introduce a matrix

$$U = Q\frac{1}{E_\alpha - H_0}\lambda V Q, \tag{12}$$

where Q is defined in the previous chapter as projection operator on the subspace of NPT. (We will omit the subscript α_s from now) It can be shown that Eq.(5) is equivalent to $s(U) < 1$. Here, $s(U)$ denotes the maximum of $|u_m|$, where u_m are eigenvalues of U., i.e, the maximal modulus of eigenvalues of T denoted as $s(T)$.

The elements of U can be written as

$$U_{ij} = \begin{cases} 0, & p_1 \le i,j \le p_2, \\ \dfrac{\lambda V_{ij}}{E_\alpha - E_i^0}, & \text{other.} \end{cases} \tag{13}$$

C. Half width of LDOS, Half width of eigenfunctions and localization length

Halfwidth of eigenstates $|\alpha\rangle$, defined as the halfwidth of the curve $\psi_\alpha(E) = \sum_j |C_{\alpha j}|^2 \delta(E - E_j)$ is known to determine a lot of properties of quantum systems.[8] To calculate average half width of eigenstates for random matrices, we apply an energy-shift method to sum over different eigenstates. We set a fixed base energy E_0, and calculate the energy shift $\Delta E_\alpha = E_0 - E_\alpha$, then we translate the curve $\psi_\alpha(E)$ by ΔE_α to make E_0 and E_α coincide. Then we sum over many eigenstates to obtain a bell-shape curve, whose halfwidth is the average halfwidth of eigenstates. In this way, what we obtain is the half width of

$$\psi(E) = \sum_{\alpha,j} |C_{\alpha j}|^2 \delta(E - E_j + E_\alpha). \tag{14}$$

Half width of local density of spectrum(LDOS) is a quantity closely related to halfwidth of eigenstates.[8] LDOS is defined as the curve $\phi_j(E) = \sum_\alpha |C_{\alpha j}|^2 \delta(E - E_\alpha)$. To calculate average half width of LDOS, we again apply an energy-shift method analogous to what we use to calculate average half width of eigenstates. In this way, what we obtain is the half width of

$$\phi(E) = \sum_{j,\alpha} |C_{\alpha j}|^2 \delta(E - E_\alpha + E_j). \tag{15}$$

Comparing Eq.(14) and Eq.(15), we see that $\psi(E) = \phi(-E)$, thus half widths of them are the same. In this way of averaging over eigenstates, we must obtain the same result for half widths of eigenstates and LDOS.

Another quantity to describe the property of eigenstates is localiztion length[8], defined as

$$L = 1/\sum_j |C_{\alpha j}|^4. \tag{16}$$

This quantity is closely related to half width of eigenstates.

We will investigate the relationship between these three quantities and NPT width in later chapters.

III. ANALYTICAL STUDY

A. NPT width for $b = 1$:small λ

For $b = 1$, the elements of matrix U in Eq.(13) are written as $U_{ij} = \lambda V_{ij}\delta_{i,j\pm1}/(E_\alpha - E_i^0)$ for i, j outside of the interval $[p_1, p_2]$.

Let us first discuss the case of small λ. In this case, for most random matrices, $p_1 = [E_\alpha]$ and $p_2 = [E_\alpha] + 1$, and $N_p = 1$. If $N_p \neq 1$, we approximately take $N_p = 2$, neglecting the probability for $N_p \ge 3$.

In this case, $s(U)$ can be approximated by the maximal modulus of eigenvalues of two submatrices truncated from U, $s(U) = max\{s(U_{\text{up}}), s(U_{\text{down}})\}$, where U_{up} and U_{down} are the following five-dimensional matrices

$$U_{\text{up}} = \lambda \begin{pmatrix} 0 & \dfrac{V_{p_1-4,p_1-5}}{E_\alpha - p_1 + 5} & 0 & 0 & 0 \\ \dfrac{V_{p_1-4,p_1-5}}{E_\alpha - p_1 + 4} & 0 & \dfrac{V_{p_1-3,p_1-4}}{E_\alpha - p_1 + 4} & 0 & 0 \\ 0 & \dfrac{V_{p_1-3,p_1-4}}{E_\alpha - p_1 + 3} & 0 & \dfrac{V_{p_1-2,p_1-3}}{E_\alpha - p_1 + 3} & 0 \\ 0 & 0 & \dfrac{V_{p_1-2,p_1-3}}{E_\alpha - p_1 + 2} & 0 & \dfrac{V_{p_1-1,p_1-2}}{E_\alpha - p_1 + 2} \\ 0 & 0 & 0 & \dfrac{V_{p_1-1,p_1-2}}{E_\alpha - p_1 + 1} & 0 \end{pmatrix}, \tag{17}$$

$$U_{\text{down}} = -\lambda \begin{pmatrix} 0 & \dfrac{V_{p_2+2,p_2+1}}{p_2 - E_\alpha + 1} & 0 & 0 & 0 \\ \dfrac{V_{p_2+2,p_2+1}}{p_2 - E_\alpha + 2} & 0 & \dfrac{V_{p_2+3,p_2+2}}{p_2 - E_\alpha + 2} & 0 & 0 \\ 0 & \dfrac{V_{p_2+3,p_2+2}}{p_2 - E_\alpha + 3} & 0 & \dfrac{V_{p_2+4,p_2+3}}{p_2 - E_\alpha + 3} & 0 \\ 0 & 0 & \dfrac{V_{p_2+4,p_2+3}}{p_2 - E_\alpha + 4} & 0 & \dfrac{V_{p_2+5,p_2+4}}{p_2 - E_\alpha + 4} \\ 0 & 0 & 0 & \dfrac{V_{p_2+5,p_2+4}}{p_2 - E_\alpha + 5} & 0 \end{pmatrix}. \tag{18}$$

We can see these two matrices are just the nearest five-dimensional matrix to E_α. The reason for this estimation is that the elements outside the above two matrices are generally much smaller than the elements inside them.

Due to the similarity between $U_{\rm up}$ and $U_{\rm down}$, we can discuss $s(U_{up})$ only. Below, we give a condition under which $s(U_{up}) > 1$. To be precise, we introduce some notations that will be frequently used. Denote $f = E_\alpha - p_1 + 1$ and $g = p_2 - E_\alpha + 1$. Denote by r_α the decimal part of E_α, then $f = 1 + r_\alpha, g = 2 - r_\alpha$. Denote by v_i $(i = 1, 2, 3, 4)$ the random matrix element $V_{p,p-1}$ used in $U_{\rm up}$ from upper to lower lines.

Now, We note that the eigenvalue μ of $U_{\rm up}$ satisfies

$$h(\mu) = \mu^4 - A\mu^2 + B = 0, \tag{19}$$

in which A and B are determined by v_is and f. Given that the discriminant of Eq.(19) is positive(which we will verify later), we may solve

$$\mu_\pm^2 = \frac{A \pm \sqrt{A^2 - 4B}}{2}. \tag{20}$$

Then $s(U_{\rm up}) > 1$ is equivalent to $\mu_+^2 > 1$, which is then $A > 2$ or $A > B + 1$.

Now we write down A and B.

$$\begin{aligned} A &= \lambda^2[\frac{v_1^2}{(f+3)(f+4)} + \frac{v_2^2}{(f+2)(f+3)} \\ &+ \frac{v_3^2}{(f+1)(f+2)} + \frac{v_4^2}{f(f+1)}], \end{aligned} \tag{21}$$

$$\begin{aligned} B &= \lambda^4[\frac{v_1^2 v_3^2}{(f+1)(f+2)(f+3)(f+4)} \\ &+ \frac{v_2^2 v_4^2}{f(f+1)(f+2)(f+3)} \\ &+ \frac{v_1^2 v_4^2}{f(f+1)(f+3)(f+4)}]. \end{aligned} \tag{22}$$

Denote the four additive terms in medium bracket in Eq.(21) as a_1, a_2, a_3, a_4. We may directly examine the discriminant of Eq.(19) that

$$\begin{aligned} A^2 - 4B &> A^2 - 4B - 4a_2a_3 \\ &= (a_2 - a_3)^2 + (a_1 - a_4)^2 - 2(a_2 - a_3)(a_1 - a_4) \\ &> 0. \end{aligned} \tag{23}$$

We may estimate the value for A and B by imposing $f = 1$ and all v_is $= 1$, then $A = 4\lambda/5$ and $B = 3\lambda^4/40$, then for at least $\lambda < 1$, the probability of $A < 2$ and $B \ll 1$ is very high. Then, when λ is relatively small, for most cases, $B \ll 1$ and $A < 2$. So the condition for $s(U_{up}) > 1$ is simplified to $A > 1$, i.e.,

$$\begin{aligned} X_{up} = &\frac{v_1^2}{(f+3)(f+4)} + \frac{v_2^2}{(f+2)(f+3)} \\ &+ \frac{v_3^2}{(f+1)(f+2)} + \frac{v_4^2}{f(f+1)} > \frac{1}{\lambda^2}. \end{aligned} \tag{24}$$

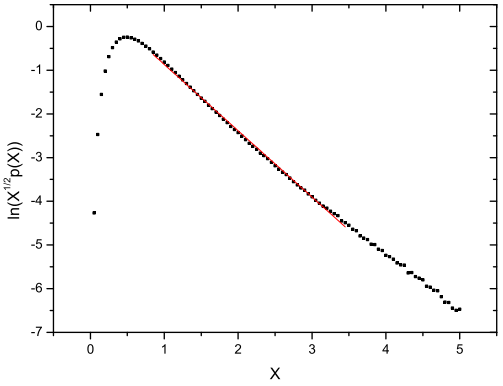

FIG. 1: Fitting the distribution function of X with $p(X) = CX^{-1/2}e^{-\beta X}(X \geq 1)$, where $X > 1/\lambda^2$ implies NPT width is larger than 1.

Substituting f by $g = 3 - f$ we get $X_{\rm down} > 1/\lambda^2$, the condition for $s(U_{\rm down}) > 1$. Denote $X = max\{X_{\rm up}, X_{\rm down}\}$, then $s(U) > 1$ is equivalent to $X > 1/\lambda^2$. We make a Monte Carlo simulation and fit for $p(X) = CX^{-1/2}e^{-\beta X}(X \geq 1)$, we omit the points in the region of large X because the probability for X in that region is very low($\sim 10^{-3}$). Our fitting result is shown in Fig.(1).

Now we calculate the average width of NPT: The probability for $N_p = 1$ is

$$P(N_p = 1) = 1 - \int_{1/\lambda^2}^{+\infty} p(X)dX, \tag{25}$$

then we set $P(N_p = 2) = 1 - P(N_p = 1)$ if we neglect the probability that $N_p > 2$, thus we obtain

$$\langle N_p \rangle = 1 + \int_{1/\lambda^2}^{+\infty} p(X)dX. \tag{26}$$

Completing the integral, we obtain

$$\langle N_p \rangle = 1 + C\sqrt{\frac{\pi}{\beta}}{\rm erfc}(\sqrt{\frac{\beta}{\lambda^2}})(\lambda \lesssim 1). \tag{27}$$

B. NPT width for $b = 1$:large λ

Next, we move to the case of large λ, in which N_p is large. We estimate the $s(U)$ by the maximum value of $s(M_i)$, in which M_is are a series of m-dimensional submatrices truncated from U. Generally, we require $b < m \ll \lambda < N$. For the case where $b = 1$, we simply set $m = 5$.

To be precise, we introduce a number sequence $l_i = p_1 - (m-1)i - 1$ and a m-dimensional matrix sequence $M_0, M_1, \cdots, M_n$. Each element M_i is obtained by truncating U by its l_ith,l_{i+1}th lines, and l_ith,l_{i+1}th columns.

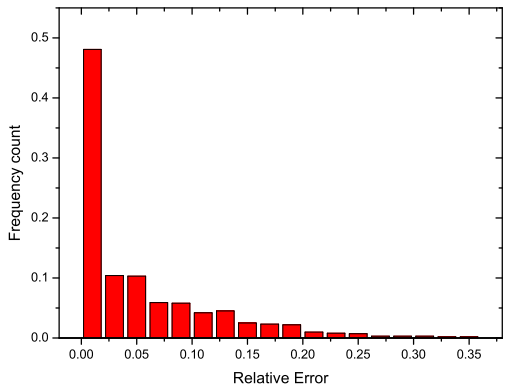

FIG. 2: Error distribution of estimation in Eq.(29)

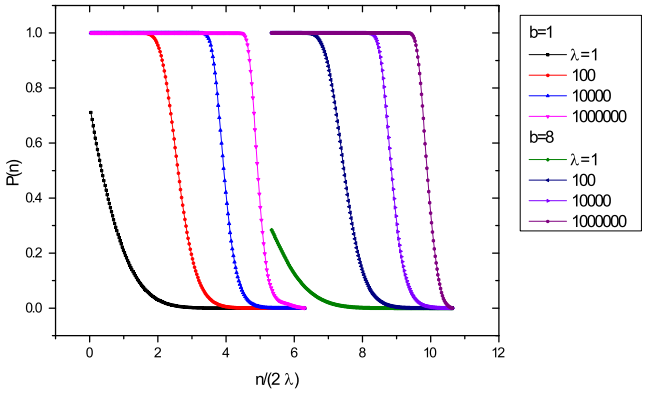

FIG. 3: Curves of $P(n)$ as b and λ vary. We see that when λ us sufficiently large, the curve is like heaviside for both b.

Therefore, the elements of M_i take the form

$$M_{i,pq} = U_{l_{i+1}+p-1,l_{i+1}+q-1} \tag{28}$$

By definition, the array M_i contains all nonvanishing element of U, and each M_i is independent to another. We estimate that

$$s(U) = \max_{0\le i\le n} s(M_i) \tag{29}$$

The difference of the left side and right side is shown in Fig.(2). We see that for most cases the difference is relatively small.

The multiplied factor $1/(E_\alpha - E_i^0)$ can be seen as a constant for each M_i, because When $i < p_1$ and $|i-j| < m \ll N_p$, we can estimate

$$\left|\frac{1}{E_\alpha - E_i^0} - \frac{1}{E_\alpha - E_j}\right| \approx \frac{m}{(E_\alpha - E_i^0)^2} \approx \frac{m}{N_p^2} \approx 0, \tag{30}$$

Denote $a_i = E_\alpha - E_{p_1} + (m-1)i + 1$, which represents the l_ith diagonal element of $1/(E_\alpha - H_0)$, then we can define $M{=}a_i M_i$, nearly independent of i due to Eq.(30). M is called a standard m-dimensional matrix whose elements take the form

$$M_{ij} = v_i \delta_{i,j\pm 1}, v_i \sim N(0,1). \tag{31}$$

We denote the density distribution function of $s(M)$ as $h(x)$, and corresponding cumulative distribution function $H(x)$. Then making use of Eq.(30), Eq.(12) and Eq.(29), we obtain

$$s(U) = \max_{0\le i\le n} (\lambda s_i/a_i), s_i \sim h(x). \tag{32}$$

Denote the density distribution function of $s(U)$ by $w(s)$ and corresponding cumulative distribution function $W(s)$. Using Eq.(32), we obtain

$$W(s) = \prod_{i=0}^{n} H(\frac{a_i s}{\lambda}). \tag{33}$$

As $m \ll \lambda$, we can regard a_i/λ as a quasi-continuous variable, so we can first take logarithm in both sides of Eq.(33), then transform the summation into integration,

$$\ln W(s) \approx \frac{\lambda}{m-1} \int_{a_0/\lambda}^{+\infty} \ln H(ts) dt, \tag{34}$$

where the upper limit is set $+\infty$ because we can regard the matrix as sufficient large. Note that $a_0 \approx N_p/2$, thus we obtain from Eq.(34) that

$$P(s(U) < 1) = \exp\left(\frac{\lambda}{m-1} \int_{N_p/2\lambda}^{+\infty} \ln H(t) dt\right). \tag{35}$$

Denote the right side of Eq.(35) by $1 - P(N_p)$, then the probability that NPT width is N_p is $P(n-1) - P(n)$, then

$$\begin{aligned} \langle N_p \rangle &= \sum_{n=1}^{\infty} n(P(n-1) - P(n)) \\ &= \sum_{n=1}^{\infty} P(n) - \lim_{n\to\infty} nP(n), \end{aligned} \tag{36}$$

where

$$P(n) = 1 - \exp\left(\frac{\lambda}{m-1} \int_{n/2\lambda}^{+\infty} \ln H(t) dt\right). \tag{37}$$

Now we discuss the properties of $P(n)$. When $n \ll \lambda$, the integral is almost $-\infty$, then $P(n) \approx 1$. As n exceeds λ, the integral swiftly increases to 0, thus $P(n) \approx 0$.

Next we evaluate the expression in the right hand of Eq.(37) by making use of our numerical results of $h(t)$(see Appendix A) and method of complex trapezoidal integration. We show the figure of $P(n)$ versus $n/2\lambda$ in Fig.(3) for different λ. The curve becomes "ladder" like when λ is very large, and $nP(n) \to 0 (n \to \infty)$, then we estimate from Eq.(36) and Eq.(37) that

$$\langle N_p \rangle = n_c, \tag{38}$$

where n_c is the point at which $P(n)$ decreases to its half height, i.e. $P(n_c) = 1/2$. Using Eq.(37), we obtain

$$\int_{n_c/2\lambda}^{+\infty} \ln H(t)dt = -\frac{(m-1)\ln 2}{\lambda}. \tag{39}$$

As λ is large, the absolute value of the right side of Eq.(39) is small, then n_c will also be large, which enables us to use the asymptotic property of $H(t)$ as $t \to +\infty$. Discussions in Appendix A show that

$$\int_{x}^{+\infty} \ln H(t)dt \sim -\frac{e^{-ax^2}}{x^2}(x \to +\infty). \tag{40}$$

The exponential decay is dominant in the expression, so we can simply omit the $1/x^2$ term, then using Eqs.(38),(39) and (40), we obtain

$$\langle N_p \rangle = C'\lambda\sqrt{\ln \lambda}(\lambda \gg 1). \tag{41}$$

If we are dealing with a perturbation whose intensity λ is restricted to a relatively small range, we can simply take $\langle N_p \rangle$ proportional to λ.

C. NPT width of Wigner-band random matrix:band width $b \geq 2$

We argue that Eq.(27) and Eq.(41) are still applicable when $b \geq 2$, despite that coefficients in the equation need to be fitted.

First, in the limit of weak perturbation, we substitute U_{up} and U_{down}(Eqs.(13)(14)) by the nearest $5b$-dimensional matrix to E_α. If we omit higher order small value in their eigenvalue equation than $O(\lambda^2)$, we will obtain a condition similar to Eq.(18), despite that the left side is composed of sum of $5b-1$ square of Gaussian random numbers. Thus, we can use the same kind of distribution function $p(X) = CX^{-1/2}e^{-\beta X}$ to fit the distribution of greatest eigenvalue of U_{up} and U_{down}. In the end, we also come to Eq.(27), but it is only applicable to relatively smaller range of λ.(to ensure higher orders of v_is can be neglected)

Second, in developing Eq.(41), we only make use of the properties of $H(t)$, the cumulative distribution function of the maximal eigenvalue of a standard m-dimensional random matrix. We can choose m to be sufficiently larger than b but also sufficiently smaller than λ. The asymptotic behavior of $H(t)$, is also Eq.(40), as is shown in Appendix A. Then our result, Eq.(41) remains unchanged, despite that C' needs to be redetermined.

The predictions of Eq.(27) and Eq.(41) are compared with numeral results later in this paper, see Figs.(4)-(6).

IV. NUMERAL RESULT

A. Iterative Algorithm

We introduce our algorithm in this section, but leave its verification in Appendix. For a given $N \times N$ Wigner band random matrix $H = H_0 + \lambda V$ and a given energy E_α of an eigenstate $|\alpha\rangle$, we may calculate the NPT width by the following algorithm.

Step 1: Separate the matrix H_0 and V into blocks as

$$H_0 = \begin{pmatrix} H_{0p} & 0 \\ 0 & H_{0n} \end{pmatrix}, V = \begin{pmatrix} V_p & * \\ * & V_n \end{pmatrix}, \tag{42}$$

in which p and n stand for up and down. H_{0p} and V_p are both $[E_\alpha] \times [E_\alpha]$ square matrices, and H_{0n} and V_n are $N - [E_\alpha]$ dimensional square matrices. This option ensures that $E_\alpha - H_{0p}$ is positive definite and $E_\alpha - H_{0n}$ is negative definite.

Step 2: Compute S_p and S_n by as following,

$$S_p = \frac{1}{\sqrt{E_\alpha - H_{0p}}}\lambda V_p \frac{1}{\sqrt{E_\alpha - H_{0p}}}, \tag{43}$$

$$S_n = \frac{1}{\sqrt{-(E_\alpha - H_{0n})}}\lambda V_n \frac{1}{\sqrt{-(E_\alpha - H_{0n})}}. \tag{44}$$

Step 3: Compute $I+S_p$, where I is the identity matrix, and use Guassian elimination procedure and gradually eliminate the elements of $I + S_p$ in the lower triangle. In the procedure, whenever we eliminate all lower-triangle elements in the ith column, we will obtain a diagonal element y_{i+1} in the $(i+1)$th row and column. We terminate the procedure when we first find a diagonal element $y_{i_{c1}} < 0$.

Step 4: Use the same procedure in step 3 for $I - S_p$ to obtain a $y_{i_{c2}} < 0$, then $p_1 = max\{i_{c1}, i_{c2}\}$.

Step 5: For $I+S_n$ and $I-S_n$, we eliminate the elements in the upper triangle by Guassian elimination from the last line. Just like step 3 and 4, we choose p_2 as the smaller i_c we obtain for two matrices, where i_c is the row and column index for the diagonal element y_{i_c} that we find less than zero at the first time in the elimination procedure.

Step 6: Calculate NPT width $N_p = p_2 - p_1$.

The algorithm above is applicable only to the case where $p_2 - p_1 > b$, in which the elements in $*$ in V do not take part in the elimination. Therefore, for a given band matrix, we first apply our iterative algorithm above to obtain p_2 and p_1, if $N_p = p_2 - p_1 \leq b$, then we use the ordinary method to solve N_p.

Similar to Guassian elimination, the algorithm has a time complexity $O(Nb^2)$. When λ is relatively large, the elimination will be much less than N times, reducing the time consumption in practice. When λ is small such that $N_p < b$, although our algorithm does not apply, ordinary method to determine p_1 and p_2 is quite quick, because we only need to try p_1 and p_2 from near E_α to far from E_α to find the minimum $p_2 - p_1$ such that $s(U) < 1$.

Detailed verification of the algorithm is in Appendix B and C.

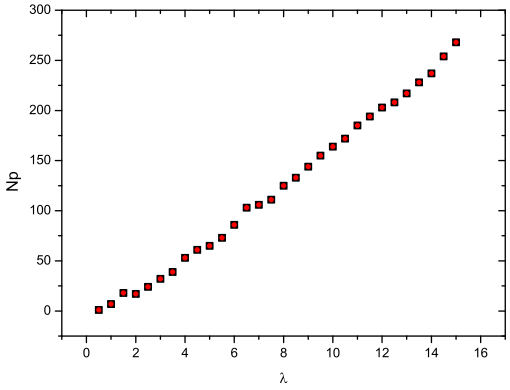

FIG. 4: Confirmation of the new algorithm. For each perturbation strength λ, we randomly choose one Wigner band random matrix with $b = 15$. Black squares represent NPT width calculated with ordinary method, and red circles represent NPT width calculated with our new iterative algorithm. Their consistency indicates that our algorithm is correct.

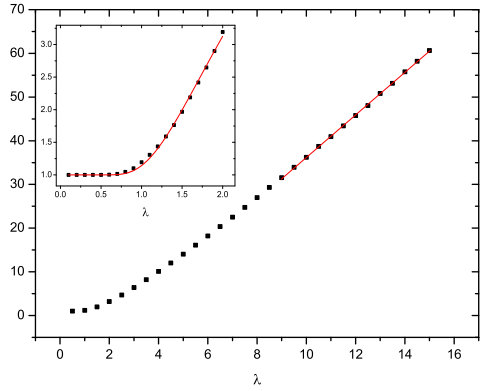

FIG. 5: Average NPT width towards perturbation strength λ for Wigner band random matrices with band width $b = 1$. When λ is small, we fit the scatters with Eq.(27). When λ is relatively large, we fit linear.

B. Confirmation of our algorithm

We randomly choose several random matrices H with different λ, and calculate their NPT width with our algorithm and ordinary algorithm to show their consistency, in Fig.(4). Fig.(4) indicates that our algorithm is correct.

C. NPT width

We use our algorithm to calculate the average NPT width in respect to λ of 10^4 Wigner-band random matrices as b varies. We fit the regions where λ is small, with Eq.(27) and Eq.(41), in Fig.(5)-(6). The figures also indicate that average NPT width increases linearly with λ for large but not very large λ. For very large λ, we plot $\langle N_p \rangle / \lambda$ towards $\sqrt{\ln \lambda}$ to examine Eq.(41) for $b = 1$, as shown in Fig.(7).

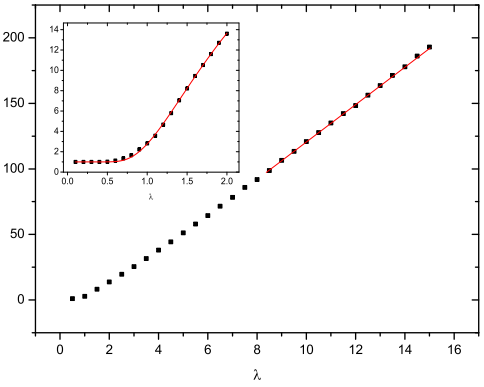

FIG. 6: Average NPT width towards perturbation strength λ for Wigner band random matrices with band width $b = 8$. When λ is small, we fit the scatters with Eq.(27). When λ is relatively large, we fit linear.

D. Comparison of NPT width, LDOS halfwidth and halfwidth of eigenstates

We plot NPT width of Wigner band random matrices and above three quantities towards λ to show their relationship in Fig.(8)-(10). We note that behaviors of four quantities do not change much when b varies.. As λ increases from 0, all quantities increase in a concave curve. When λ becomes larger, the increasing speed of localization length gradually drops and finally the localization length remains nearly constant as λ increase. This is so-called dynamical localization,[8] before which four quantities are comparable. After dynamical localization, the other three quantities increase nearly linearly, with nearly the same slope. Thus, the increase of NPT width can partly characterize the increasing of LDOS halfwidth and halfwidth of eigenstates. As b rises up, dynamical localization happens for larger λ, so four quantities are comparable in a larger range.

Now we give an explanation for the numerical results that NPT width nearly equals to half width of LDOS and eigenstates. Wave functions of eigenstates, shown in Fig.(11), generally consist of several major peaks together with some minor peaks. Due to randomness of the matrices, we anticipate that distribution of major peaks is random. On the other hand, wave functions out of NPT region decays exponentially, thus major peaks may only occur inside the NPT region. Fig.(12) shows that for typical Wigner-band random matrices, major peaks distribute almost evenly inside NPT region. Averaging

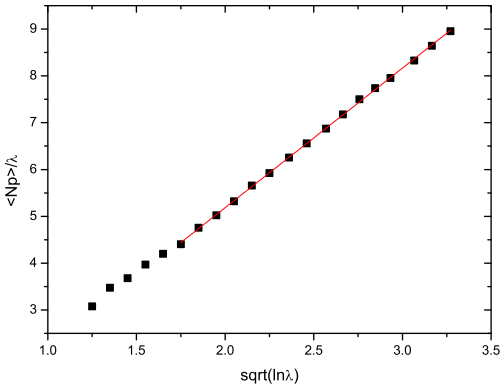

FIG. 7: Average NPT width towards perturbation strength λ for Wigner band random matrices with band width $b = 1$ when λ is very large. We change the axis to make a linear fit with Eq.(41).

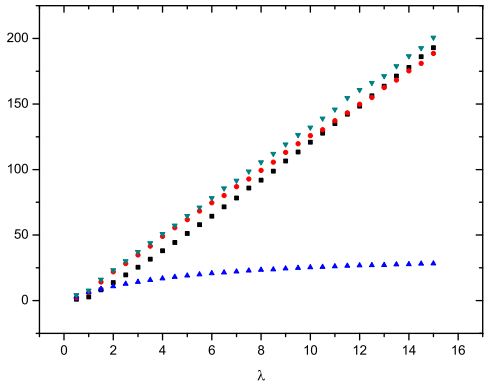

FIG. 9: Average NPT width(black squares), LDOS halfwidth(dark green triangles), half width of eigenstates(red circles) and localization length(blue triangles) for Wigner band random matrices with band width $b = 8$.

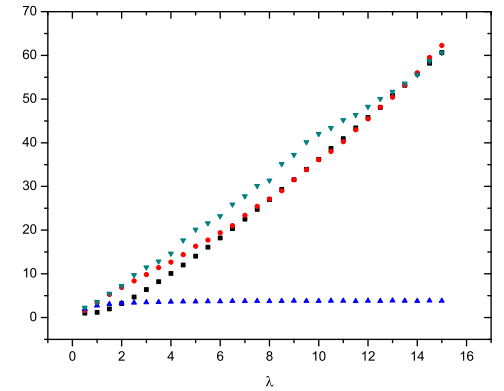

FIG. 8: Average NPT width(black squares), LDOS halfwidth(dark green triangles), half width of eigenstates(red circles) and localization length(blue triangles) for Wigner band random matrices with band width $b = 1$.

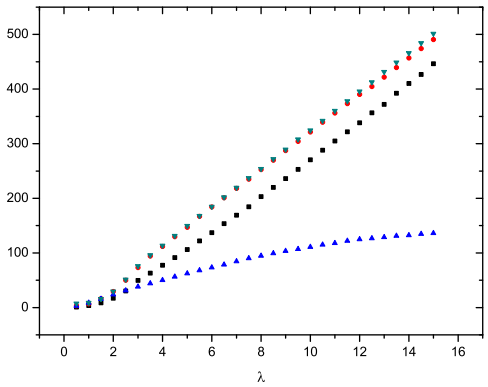

FIG. 10: Average NPT width(black squares), LDOS halfwidth(dark green triangles), half width of eigenstates(red circles) and localization length(blue triangles) for Wigner band random matrices with band width $b = 50$.

over these peaks, mean shape of eigenfunctions thus exhibit a steady distribution inside the NPT region but a sharp decay outside the NPT region and the shoulders (width b). Thus half width of eigenstates is close to NPT width. We obtain LDOS by expanding unperturbed eigenstates with perturbed ones, just inversely of how we obtain eigenfunctions, so half width of LDOS is also close to NPT width.

Therefore, given that the average eigenfunction changes steadily inside NPT region, we can predict that it decays exponentially outside the NPT region plus two shoulders with width b. See Fig.(13). Thus, when the eigenfunction $|C_{\alpha i}|^2$ decreases to half its maximum, the corresponding energy E_i is near the beginning of exponential decay (the exterior boundary of shoulders). The average shape of LDOS is more or less like that of eigenfunction. Then $w_{\text{eigenfunction}} \approx w_{\text{LDOS}} = N_p + 2b\eta$, where η is a number near to 1, and w denotes the halfwidth.

We plot η against λ in Fig.(14). We can see that $\eta \approx 1$ for middle-strength perturbation, and is small in weak or strong perturbation. For small enough perturbation, it weakly mixes adjacent eigenstates, thus the effect of band width is not obvious. For larger perturbation, difference between halfwidth of LDOS and NPT width is smaller because the $\sqrt{\ln \lambda}$ term in NPT width. (Eq.(41))

V. CONCLUSION

In this paper, we have studied the non-perturbative and perturbative parts of Wigner band random matrix. We have explicitly developed the formula for NPT width

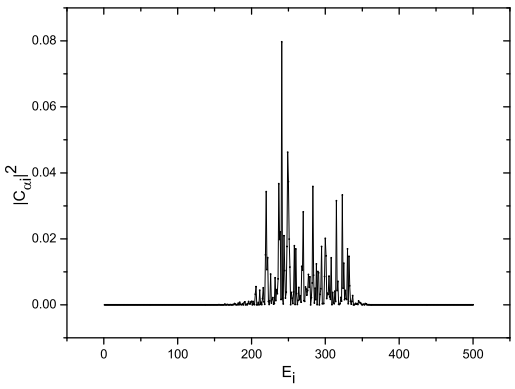

FIG. 11: Typical shape of $|C_{\alpha i}|^2 = |\langle i|\alpha\rangle|^2$ versus E_i^0 of Wigner-band random matrices. The shape consists of several major peaks($|C_{\alpha i}|^2 > 0.03$) and some minor peaks, while the major peaks are distributed randomly inside NPT region. The figure above is an eigenfunction for $N = 500$,$b = 8$,$\lambda = 15$.

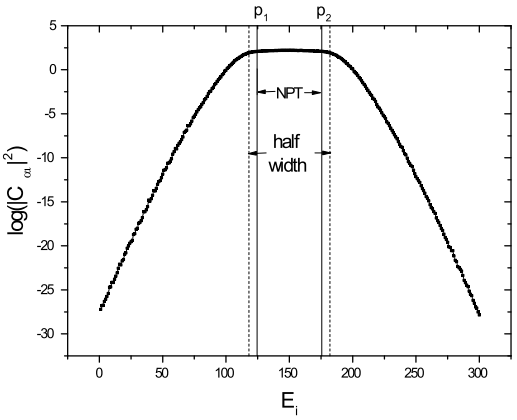

FIG. 13: Average shape of eigenfunctions of Wigner-band random matrices, $b = 8$,$\lambda = 5$,$N = 500$. The shape is steady inside NPT region but decreases exponentially outside shoulders. Half width of eigenfunctions is near NPT width plus two shoulders.

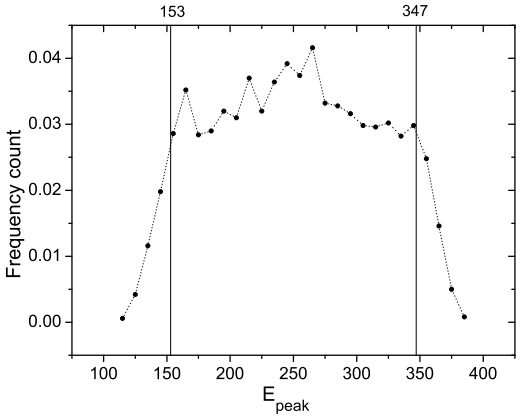

FIG. 12: For typical Wigner-band random matrices, $N = 500$,$b = 8$,$\lambda = 15$, distribution of major peaks is almost an even distribution inside NPT region between solid vertical lines.

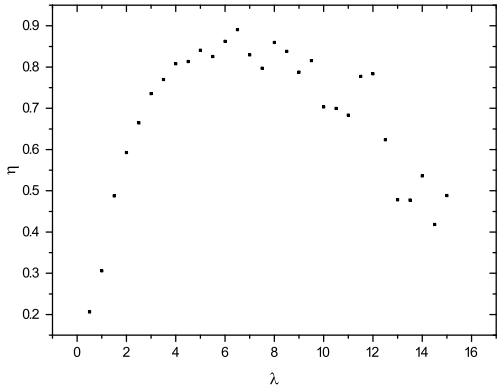

FIG. 14: Value of η against λ for $b = 8$, $N = 500$. $\eta \approx 1$ for middle-strength perturbation, and is small in weak or strong perturbation.

towards perturbation strength λ in the limits of weak and strong perturbation. We have shown that when perturbation is weak, the average NPT width $\langle N_p\rangle$ increases very slowly as shown in Eq.(27), and when perturbation is strong, $\langle N_p\rangle$ increases slightly quicker than linear to λ, as shown in Eq.(41).

Numeral calculation of NPT width is simplified by the recursive algorithm, with a time complexity less than $O(b^2N)$. We confirm this new approach by comparison to the ordinary method and then use the new method to calculate the relationship between $\langle N_p\rangle$ and λ, for different band width $b = 1, 8, 50$, respectively. We fit the numeral results with analytical formulas Eqs.(20) and (34).

We calculate LDOS width, halfwidth of eigenfunctions, and localizaton length, and compare them with NPT width. We find that LDOS width and NPT width both increase nearly linearly, with nearly the same slope when λ is large. Localization length remains nearly constant when λ is large, due to dynamical localization. Halfwidth of eigenfunctions is nearly the same with LDOS width. When λ is small, four properties differ slightly, before dynamical localization appears. The range for λ when these four properties converge increases as b increases.

Our research suggests that for some cases, NPT width for a band matrix is near to LDOS width, localiztion length and halfwidth of eigenfunctions, and NPT width is easier to calculate with new methods. Further research may reveal the relationship of the four properties more precisely.

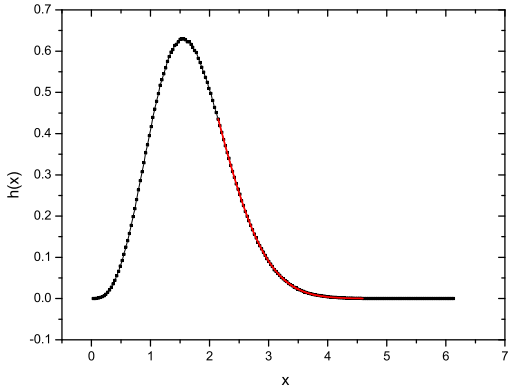

FIG. 15: Density function $h(x)$ of the maximum absolute value of eigenvalues of a standard random matrix with $b = 1$. Guassian fit of the large-λ region is shown in red curve.

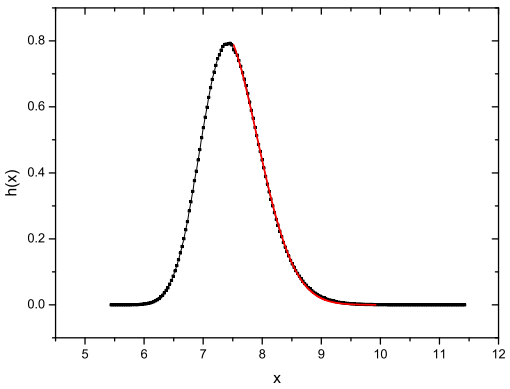

FIG. 16: Density function $h(x)$ of the maximum absolute value of eigenvalues of a standard random matrix with $b = 8$. Guassian fit of the large-λ region is shown in red curve.

Acknowledgments

We acknowledge full financial support from National Training Programs of Innovation and Entrepreneurship for Undergraduates.

APPENDIX A: ASYMPTOTIC PROPERTY OF $H(t)$ AND PROPERTIES OF $P(n)$

In previous chapters, we define $h(t)$ as the density distribution function of the maximal modulus of eigenvalues of a standard m-dimensional random matrix M_s, whose elements take the form

$$M_{s,ij} = \begin{cases} v_{ij} \sim N(0,1), & 1 \le |i-j| \le b \\ 0, & \text{others.} \end{cases} \tag{A1}$$

and a corresponding cumulative distribution function $H(t)$. We now explicitly show the curves of $h(t)$ when $b = 1$ and $b = 8$ obtained by a Monte Carlo simulation in Fig.(15) and (16). We fit the region of large t by Guassian formula and show the asymptotic property of $h(t)$ that

$$h(t) \sim e^{-t^2} (t \to \infty). \tag{A2}$$

Making use of

$$H(x) = 1 - \int_x^{+\infty} h(t)dt \tag{A3}$$

and

$$\int_x^{+\infty} t^{-n} e^{-t^2} dt \sim x^{-n-1} e^{-x^2}, \tag{A4}$$

we obtain

$$1 - H(t) \sim \frac{e^{-t^2}}{t} (t \to \infty). \tag{A5}$$

We take first order approximation that $\ln(1 - H(t)) \approx -H(t)$ and use Eq.(A4) to obtain

$$\int_x^{+\infty} \ln H(t)dt \sim \frac{e^{-x^2}}{x^2}, \tag{A6}$$

which is just Eq.(40).

APPENDIX B: VERIFICATION OF THE ITERATIVE ALGORITHM

To develop our algorithm, we first make a similarity transformation to U to create a symmetrical matrix. We first consider the upper part of U called U_p ,for Q naturally split U into two independent parts. We will denote the correspondingly upper part of V and H_0 as V_p and H_{0p}. Note that Q and $1/(E_\alpha - H_0)$ commute, so we rearrange U_p^n as

$$U_p^n = Q \frac{1}{\sqrt{E_\alpha - H_{0p}}} S_p^{n-1} \frac{1}{\sqrt{E_\alpha - H_{0p}}} \lambda Q, \tag{B1}$$

where

$$S_p = \frac{1}{\sqrt{E_\alpha - H_{0p}}} \lambda V_p \frac{1}{\sqrt{E_\alpha - H_{0p}}}. \tag{B2}$$

Then the condition $s(U_p) < 1$ is equivalent to $s(S_p) < 1$, with S_p obviously symmetrical and with the same band width.

Now we recall some results in linear algebra.[8] A real symmetrical matrix has real eigenvalues, and condition $s(S_p) < 1$ is equivalent to $I \pm S_p$ are both positive definite, where I is the identity matrix. The sufficient and necessary condition for a symmetrical matrix to be positive definite is that its ordered main subdeterminants(denoted by d_i) are all positive. We will calculate

the ordered main subdeterminants of $I \pm S_p$ by elementary row transformations. To be clear, we first deal with the case where $b = 1$.

In this simpler case, one can prove that if S_p has a positive eigenvalue x_0, it must have a corresponding eigenvalue $-x_0$, so we only need to verify the condition under which $I + S_p$ is positive definite. Recall the elementary row transformation matrices $T_{ij}(\mu)$, with all diagonal elements 1 and only nonvanishing element μ in the ith column and tth row, and multiplying $T_{ij}(\mu)$ from the left gives an elementary row transformation that does not change any its subdeterminant. Suppose the elements of S_p take the form

$$S_p(i, i+1) = S_p(i+1, i) = \xi_i (i = 1, 2, \cdots, p_1 - 2), \quad \text{(B3)}$$

then $d_1 = 1$, and we can continuously multiply $T_{i,i+1}(\mu_i)$ from the left to eliminate the element of $I + S_p$ in lower triangle to obtain d_{i+1}. For instance, our first step is that

$$T_{12}(-\xi_1)(I + S_p) = \begin{pmatrix} 1 & & & & \\ -\xi_1 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & \xi_1 & & & \\ \xi_1 & 1 & \xi_2 & & \\ & \xi_2 & 1 & & \\ & & & \ddots & \xi_{p_1-2} \\ & & & \xi_{p_1-2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \xi_1 & & & \\ 0 & 1-\xi_1^2 & \xi_2 & & \\ & \xi_2 & 1 & & \\ & & & \ddots & \xi_{p_1-2} \\ & & & \xi_{p_1-2} & 1 \end{pmatrix}, \quad \text{(B4)}$$

so $\mu_1 = -\xi_1$, and $d_2 = 1 - \xi_1^2$. If $d_2 > 0$, then we multiply $T_{23}(-\xi_2/(1-\xi_1^2))$ from the left to obtain d_3. Now we introduce the sequence $\{y_i\}$ to denote the new diagonal element before ith elimination, then $y_1 = 1, y_2 = d_2/d_1 = 1 - \xi_1^2, \cdots, y_i = d_i/d_{i-1}, \cdots$. We can construct a recursive formula for y_{i+1} by the ith elimination

$$T_{i,i+1}(-\frac{\xi_{i+1}}{y_i}) \begin{pmatrix} y_1 & \xi_1 & & & & & & & \\ & y_2 & \xi_2 & & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & y_i & \xi_i & & & & \\ & & & \xi_i & 1 & \xi_{i+1} & & & \\ & & & & \xi_{i+1} & 1 & & & \\ & & & & & & \ddots & \xi_{p_1-2} & \\ & & & & & & \xi_{p_1-2} & 1 & \end{pmatrix} = \begin{pmatrix} y_1 & \xi_1 & & & & & & & \\ & y_2 & \xi_2 & & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & y_i & \xi_i & & & & \\ & & & 0 & y_{i+1} & \xi_{i+1} & & & \\ & & & & \xi_{i+1} & 1 & & & \\ & & & & & & \ddots & \xi_{p_1-2} & \\ & & & & & & \xi_{p_1-2} & 1 & \end{pmatrix}, \quad \text{(B5)}$$

from which we obtain

$$y_{i+1} = 1 - \frac{\xi_i^2}{y_i} (i \geq 1); y_1 = 1. \quad \text{(B6)}$$

Finally we obtain our algorithm to calculate p_1 for the case $b = 1$. As p_1 is the maximal row number that ensures all ordered main subdeterminants d_i of $I + S_p$ positive, we require that $\forall i \leq p_1 - 1, y_i > 0$. Given a matrix S, we continuously apply Eq.(B6), until we obtain a $y_{i_c} < 0$, then $p_1 = i_c$. Similarly, if we set $y_N = 1$, we can obtain p_2 by adopting the same recursive formula for y_{i-1}, then obtain $N_p = p_2 - p_1$. We can also verify this algorithm by path summation in Appendix C. Now we expand our algorithm to the cases where $b \geq 2$. In this case, our way above to calculate ordered subdeterminants d_i is still applicable. Every time we eliminate elements in a column in lower triangle of $I \pm S_p$ by elementary row transformations b times, we obtain a higher ordered subdeterminant. We end the iteration until we find a negative i_c-order subdeterminant. Only difference lies in that we need to both compute i_c for $I \pm S_p$, and choose p_1 as the smaller one.

Similarly, we can use the algorithm calculate p_2, then we obtain the NPT width $N_p = p_2 - p_1$.

APPENDIX C: PATH SUMMATION DERIVATION FOR THE ITERATIVE ALGORITHM

We provide a new picture for our recursive formula Eq.(B6). Recall that NPT width is the minimal dimension of subspace P such that the generalized Brillouin-Wigner perturbation expansion Eq.(6) converges. De-

compose $|\alpha_s\rangle$ as summation of NPT eigenstates,

$$|\alpha_s\rangle = \sum_{i=p_1}^{p_2} t_i|i\rangle, \tag{C1}$$

then convergence of Eq.(6) is equivalent to convergence of

$$C_{ij} = \langle j|i\rangle + \sum_{m=1}^{n-1}\langle j|T^{m-1}|i\rangle + \langle j|T^n|i\rangle + \cdots \tag{C2}$$

for any $i \in [p_1, p_2]$, and any j. Note that the definition of T (Eq.(5)) consists of a projection operator Q, so we only need to check the convergence of Eq.(C2) for the case where $j \notin [p_1, p_2]$. Using Eq.(5) and Eq.(C2), we can write the explicit formula for C_{ij},

$$C_{ij} = \frac{\lambda V_{ji}}{E_\alpha - E_j} + \sum_{k_1\in Q}\frac{\lambda V_{jk_1}}{E_\alpha - E_j}\frac{\lambda V_{k_1 i}}{E_\alpha - E_{k_1}} + \sum_{k_1,k_2\in Q}\frac{\lambda V_{jk_1}}{E_\alpha - E_j}\frac{\lambda V_{k_1k_2}}{E_\alpha - E_{k_1}}\frac{\lambda V_{k_2 i}}{E_\alpha - E_{k_2}} + \cdots . \tag{C3}$$

By definition of NPT width, we need to find the maximal dimension of Q such that Eq.(C3) converges. For simplicity, we only consider the case where $b = 1$. Let we call the chain $k_0 = j \to k_1 \to \cdots \to k_n \to i = k_{n+1}$ a $(n+1)$-order path, then the nth term in Eq.(C3) is a summation over all n-order paths. We denote $\lambda V_{k_i k_{i+1}}/(E_\alpha - E_{k_i})$ as $f(k_i \to k_{i+1})$, then, as $b = 1$, it vanishes unless $|k_{i+1} - k_i| = 1$.

For $b = 1$, Q subspace is naturally split into to two separate parts,$[1, p_1-1]$ and $[p_2+1, N]$. Then the summation Eq.(C3) vanishes unless $i = p_1$ or $i = p_2$, where the summation only covers paths in one side of Q subspace. Two sides of Q subspaces are similar, so let us consider the case where $i = p_1$. We claim that convergence of Eq.(C3) for $j = p_1 - 1$ implies convergence for any $j \in [1, p_1 - 2]$. This is because if for one of $j \in [1, p_1 - 2]$ the path summation Eq.(C3) diverges, then for any path $j \to \cdots \to i$, we can construct a path $j + 1 \to j \to \cdots \to i$, and summation over all these paths also diverges because it is just $f(j + 1 \to j)$ times the former summation, meaning that $C_{p_1 j+1}$ also diverges. Iterate the reasoning above leads to contradiction to our condition that $C_{p_1 p_1-1}$ converges. Therefore later we only consider the case where $i = p_1$ and $j = p_1 - 1$.

Since the only way to p_1 is through $p_1 - 1$, We rewrite Eq.(C3) as

$$C_{p_1 p_1-1} = f(p_1 - 1 \to p_1)A(p_1 - 1 \to p_1 - 1), \tag{C4}$$

where $A(j \to j)$ represents the path summation over paths starting at j and ending at j, without passing through $j + 1$. Now we classify the paths from $p_1 - 1$ to $p_1 - 1$ by the number of times the path includes $p_1 - 1$ in the middle. Suppose all paths(excluded the beginning) that include $p_1 - 1$ one time contribute $g(p_1 - 1)$ to the path summation, then all paths that include $p_1 - 1$ n times contribute $g^n(p_1 - 1)$. Then

$$\begin{aligned} A(p_1 - 1 \to p_1 - 1) &= \sum_{n=1}^{\infty} g^n(p_1 - 1) \\ &= \frac{1}{1 - g(p_1 - 1)} - 1, \end{aligned} \tag{C5}$$

which converges only when $|g(p_1 - 1)| < 1$. Next we consider the paths that contribute to $g(p_1 - 1)$. We may first go to $p_1 - 2$, then return to $p_1 - 1$, then the path contributes $f(p_1 - 1 \to p_1 - 2)f(p_1 - 2 \to p_1 - 1)$ to the summation. By definition, all paths that goes to $p_1 - 2$ then returns to $p_1 - 2$ contributes to a factor $A(p_1 - 2 \to p_1 - 2)$, then

$$\begin{aligned} g(p_1 - 1) &= f(p_1 - 1 \to p_1 - 2)f(p_1 - 2 \to p_1 - 1) \\ &\times (1 + A(p_1 - 2 \to p_1 - 2)). \end{aligned} \tag{C6}$$

Using Eqs.(C5), (C6), we obtain

$$\begin{aligned} g(p_1 - 1) &= f(p_1 - 1 \to p_1 - 2)f(p_1 - 2 \to p_1 - 1) \\ &\times \frac{1}{1 - g(p_1 - 2)}. \end{aligned} \tag{C7}$$

The reasoning above applies to any $j \in [1, p_1 - 1]$. Note that $g(1) = 0$ by definition, then we obtain the recursive formula

$$g(j + 1) = f(j + 1 \to j)f(j \to j + 1)\frac{1}{1 - g(j)}; g(1) = 0. \tag{C8}$$

Direct calculation of matrix elements shows that

$$f(j + 1 \to j)f(j \to j + 1) = \xi_j^2, \tag{C9}$$

in which ξ_j^2 shares the same meaning with that in Eq.(40). Now let $y_i = 1 - g(i)$, then Eq.(B8) becomes

$$y_{i+1} = 1 - \frac{\xi_i^2}{y_i}; y_1 = 1, \tag{C10}$$

which is exactly Eq.(B6).

In this approach, path summation converges if and only if $g(p_1 - 1) < 1$, i.e, $y_{p_1-1} > 0$. Therefore, as we proceed our iteration, if we find a $y_{i_c} < 0$, then $i_c = p_1$, which is consistent with our previous derivation by elementary row transformation. p_2 can be derived by similar approach.

[1] B.Lauritzen, *et al.*, Phys. Rev. Lett. **74**, 5190 (1995).

[2] Ph. Jacquod, D.L. Shepelyansky, O. P. Sushkov, Phys.

Rev. Lett. **78**, 923 (1997)
[3] Carlos Mejia-Monasterio, *et al.*, Phys. Rev. Letts. **81**, 5189 (1998)
[4] V.V.Flambaum and F. M. Izrailev, Phys. Rev. E **61**, 2539(2000); *ibid.*, **64**, 026124 (2001)
[5] P. R. Zangara, *et al.*, Phys. Rev. A **86**, 012322 (2012).
[6] Lea F. Santos, F.Borgonovi and F.M. Izrailev, Phys. Rev. Lett. **108**, 094102 (2012).
[7] E. J. Torres-Herrera and Lea F. Santos, Phys. Rev. A **89**, 043620 (2014); *ibid.*, **90**, 033623 (2014).
[8] Wen-ge Wang, Phys. Rev. E **61**, 952 (2000); *ibid.* **65**, 036219 (2002).