Chinese Physics Letters

LETTER

The Roaming Bethe Roots: An Effective Bethe Ansatz Beyond Integrability

Cite as: Wenlong Zhao *et al*. 2026 *Chin. Phys. Lett.* **43** 100001

Recommended Articles

A Composite Ansatz for Calculation of Dynamical Structure Factor
Yupei Zhang, Chongjie Mo, Ping Zhang, and Wei Kang
Chin. Phys. Lett. **41**, (2024)

Properties of One-Dimensional Highly Polarized Fermi Gases
Ya-Dong Song, and Xiao-Ming Cai
Chin. Phys. Lett. **35**, 110301 (2018)

Exact Solution of a Non-Hermitian Generalized Rabi Model
Yusong Cao, and Junpeng Cao
Chin. Phys. Lett. **38**, 080202 (2021)

Spectrum of the Open Asymmetric Simple Exclusion Process with Arbitrary Boundary Parameters
WEN Fa-Kai, YANG Zhan-Ying, CUI Shuai, CAO Jun-Peng, and YANG Wen-Li
Chin. Phys. Lett. **32**, 050503 (2015)

A series journal of Chinese Physical Society distributed by IOP Publishing

Chinese Physics Letters

Chin. Phys. Lett. **43** 100001 (2026)

LETTER

DOI: 10.1088/0256-307X/43/10/100001
CSTR: 32039.14.0256-307X.43.10.100001

The Roaming Bethe Roots: An Effective Bethe Ansatz Beyond Integrability

Wenlong Zhao(赵文龙)[1], Yunfeng Jiang(江云峰)[2,5,*], and Rui-Dong Zhu(朱睿东)[1,3,4,*]

[1]School of Physical Science and Technology & Institute for Quantum Science, Soochow University, Suzhou 215006, China
[2]School of Physics & Shing-Tung Yau Center, Southeast University, Nanjing 211189, China
[3]Suzhou Municipal Key Laboratory of Quantum-Intelligence Fusion Computing, Suzhou 215000, China
[4]Jiangsu Key Laboratory of Frontier Material Physics and Devices, Soochow University, Suzhou 215006, China
[5]Peng Huanwu Center for Fundamental Theory, Hefei 230026, China

[*]Corresponding authors. Email: jinagyf2008@seu.edu.cn; rdzhu@suda.edu.cn

We propose an effective Bethe ansatz (EBA) for solving quantum many-body systems near an integrable point. Our approach retains the functional form of the Bethe wave function while renormalizing the Bethe roots to account for integrability-breaking interactions. These effective roots are determined by minimizing physically motivated cost functions. The resulting off-shell Bethe states serve as approximate eigenstates of the non-integrable models. We assess the quality of the approximation using various physical observables, including the energy eigenvalue, state fidelity, and bipartite entanglement entropy. Our tests show that for models with weak integrability-breaking, the EBA provides a high-quality approximation to the exact eigenstates over a wide range of deformation parameters. In contrast, for models with strong integrability-breaking interactions, the efficacy of EBA degrades relatively quickly as the deformation parameter increases. The efficacy of the method thus offers a useful probe for characterizing the strength of integrability breaking. Within its regime of accuracy, it also provides a new representation of the eigenstates of nearly integrable models, enabling one to exploit the algebraic structure inherited from integrability.

Introduction

Integrable models are special many-body systems that are exactly solvable. Their solvability originates from an extensive set of conserved quantities, which impose strong constraints on the dynamics and lead to many unique physical properties. A fundamental question in integrable models is what happens when integrability is broken. In classical systems, the Kolmogorov–Arnold–Moser theorem [1–3] shows that weak integrability-breaking perturbations preserve most phase space tori; as the perturbation increases, the tori gradually break down, leading to chaos. In the quantum regime, systems near integrable points also exhibit properties reminiscent of integrability, such as approximate conservation laws [4–9], prethermalization [10–14], and anomalous transport [15–22]. These observations may be interpreted as evidence that the influence of integrability or similar physical properties does not cease abruptly under deformation, but instead diminishes continuously with growing perturbation strength.

This picture implies that sufficiently close to an integrable point, the influence of integrability remains dominant. It provides useful intuition about the physical properties one can expect for models near integrability. At the same time, it offers hints for how to actually solve such models. Integrability is a powerful tool when applicable. Near an integrable point, one can pursue a hybrid method that exploits the integrability structure while incorporating integrability-breaking effects in a controlled manner. Indeed, one may treat the integrability-breaking part as a perturbation and apply standard perturbation theory [23,24], or adopt the Hamiltonian truncation approach [25–27].

In this work, we pursue a different route to studying models near integrability, based on the scattering picture. A key idea behind integrable models is factorized scattering, where the system can be viewed as a collection of stable particles interacting in a factorizable way. What happens to factorized scattering when integrability is broken? It is reasonable to assume that sufficiently close to the integrable point, the factorized scattering picture remains valid, but certain parameters must be renormalized. This is analogous to Fermi liquid theory, where interactions renormalize parameters such as the effective mass. When and how does the factorized scattering picture break down as integrability-breaking interactions are turned on? To investigate this question concretely, in this work, we focus on models solvable by the Bethe ansatz.

In the Bethe ansatz, the wave function is built upon the assumption of factorized scattering [28–31]. In finite volume, the rapidities of pseudoparticles are quantized by the Bethe ansatz equations (BAE). When integrability is broken, the Bethe ansatz no longer yields exact solutions. However, we could assume that sufficiently close to an integrable point, the functional form of the Bethe wave function remains valid while the Bethe roots are deformed. To determine these deformed roots, we introduce a numerical optimization approach: we define suitable cost functions that depend on the deformed roots and employ modern optimization algorithms to find the minima. This approach, using the Bethe ansatz as a variational ansatz in non-integrable models, restores the original spirit of the ansatz as a trial wave function and shall be called an effective Bethe ansatz (EBA). A similar idea has been explored to solve the (integrable) XXZ [32,33] and Lieb–Liniger models [34]; reference [35] employed the Richardson–Gaudin state as a variational ansatz for approximating near-integrable models.

The resulting off-shell Bethe states with the deformed Bethe roots serve as an approximation of the exact eigenstates. To assess the validity of the EBA, we compare energies, fidelities, and bipartite entanglement entropies of the effective Bethe states with exact eigenstates. A clear distinction emerges between two types of model: under the deformation of Eq. (7) (details are provided in the following sections), the approximation degrades much more slowly than under Eq. (8), remaining accurate over a broad parameter range. The rate of breakdown thus serves as a diagnostic for the strength of integrability breaking. Moreover, we found that near specific points of the parameter space, such as a level crossing point or a phase transition point, the

Received 4 July 2026; accepted 26 August 2026; published 28 October 2026
© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Check for updates

approximation undergoes a sharp decline, suggesting that the EBA can also be a useful probe for quantum criticality.

Bethe Ansatz as an Ansatz

We consider a Hamiltonian of the form

$$H(\lambda) = H_0 + \lambda H_1, \tag{1}$$

where H_0 is an integrable spin chain solvable by the Bethe ansatz. We denote the eigenstates by $|\{u_j\}\rangle$, which depends on a set of rapidities $\{u_j\}$ that satisfy the BAE. The Hamiltonian is deformed by λH_1, where λ controls the strength of the deformation. The deformation may preserve integrability or, more generally, break it. Our central assumption is that, for sufficiently small λ, the Bethe state remains a good approximate eigenstate of $H(\lambda)$, as long as $\{u_j\}$ are chosen properly. The range of validity for λ depends on the nature of the deformation operator H_1.

We distinguish two classes of deformations[*]: *weak* and *strong* integrability breaking. The two types exhibit different physical properties. Weakly integrability-breaking models exhibit approximate conservation laws [4,5,7,36–39] and display behavior close to that of integrable systems, such as anomalously large heat conductivity [40] and prethermalization [13,41,42]. In contrast, strongly integrability-breaking perturbations tend to produce generic non-integrable behavior, including chaotic level statistics [43–45], particle decay and production [24,46], soliton confinement [47–53], and false vacuum decay [54–57]. It is therefore natural to expect that the EBA will remain accurate for weakly integrability-broken models over a relatively wide parameter window, whereas for strongly integrability-broken models the approximation will deteriorate rapidly, remaining useful only for very small λ.

Assuming the Bethe state provides a good approximation in the appropriate regime, the deformed Bethe roots can be obtained by treating the off-shell Bethe state as a variational ansatz with parameters $\{u_j\}$. We optimize these parameters by minimizing a suitable functional. Specifically, for the ground state, we minimize

$$\mathcal{H}_0(\lambda;\{u_j\}) \equiv \frac{\langle\{u_j\}|H(\lambda)|\{u_j\}\rangle}{\langle\{u_j\}|\{u_j\}\rangle}. \tag{2}$$

The resulting rapidities are denoted by $\{\bar{u}_j\}$. A convenient starting point for the optimization is provided by the original Bethe roots of the undeformed integrable spin chain. The resulting normalized ground state is then $|\psi_0\rangle = |\{\bar{u}_j\}\rangle/\sqrt{\langle\{\bar{u}_j\}|\{\bar{u}_j\}\rangle}$.

For excited states, we minimize a modified functional,

$$\mathcal{H}_1(\lambda;\{u_j\}) = \mathcal{H}_0(\lambda;\{u_j\}) + K_1\frac{|\langle\{u_j\}|\psi_0\rangle|^2}{\langle\{u_j\}|\{u_j\}\rangle}, \tag{3}$$

where $K_1 \gg 1$ is a large constant that effectively enforces orthogonality to $|\psi_0\rangle$. In general, this condition alone does not uniquely select the first excited state; the minimization may converge to other excited states. Lower-lying excitations can be singled out by initializing the search near the corresponding undeformed Bethe roots. The resulting rapidities are denoted by $\{\bar{u}_j^{(1)}\}$, and the normalized excited state is $|\psi_1\rangle = |\{\bar{u}_j^{(1)}\}\rangle/\sqrt{\langle\{\bar{u}_j^{(1)}\}|\{\bar{u}_j^{(1)}\}\rangle}$.

This procedure can be iterated by defining

$$\mathcal{H}_n(\lambda;\{u_j\}) = \mathcal{H}_{n-1}(\lambda;\{u_j\}) + K_n\frac{|\langle\{u_j\}|\psi_{n-1}\rangle|^2}{\langle\{u_j\}|\{u_j\}\rangle} \tag{4}$$

with $K_n \gg 1$. Minimizing $\mathcal{H}_n$ yields rapidities $\{\bar{u}_j^{(n)}\}$ for the n-th excited state, giving the approximate eigenstate $|\psi_n\rangle = |\{\bar{u}_j^{(n)}\}\rangle/\sqrt{\langle\{\bar{u}_j^{(n)}\}|\{\bar{u}_j^{(n)}\}\rangle}$.

We note that when the target Hamiltonian $H(\lambda)$ possesses degenerate energy levels, the cost function (2) or (4) will have multiple distinct optima. In such cases, the optimized Bethe roots are not unique: different initial guesses for the effective Bethe roots lead to different local minima, each corresponding to a particular linear combination within the degenerate subspace.

Models and Results

We now test our proposal on concrete models. We take H_0 to be the periodic Heisenberg XXX spin chain,

$$H_0 = H_{\rm XXX} = \sum_{n=1}^{L}\boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_{n+1}, \tag{5}$$

where $\boldsymbol{\sigma}_n = (\sigma_n^x,\sigma_n^y,\sigma_n^z)$ are the Pauli matrices. The model is solved exactly by the algebraic Bethe ansatz (ABA), which we briefly review in Appendix. Within the ABA framework, a Bethe state takes the form

$$|\{u_j\}\rangle = B(u_1)\ldots B(u_M)|\Omega\rangle, \tag{6}$$

where the operator $B(u)$ is defined in Appendix and $|\Omega\rangle = |\uparrow^L\rangle$ denotes the ferromagnetic reference state. For the deformation term H_1, we consider two representative cases:

- A weak integrability-breaking deformation

$$H_1 = \sum_{n=1}^{L}\boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_{n+2}, \tag{7}$$

  corresponding to the J_1-J_2 model, and charges conserved up to $O(\lambda^2)$ corrections were constructed in Ref. [36], demonstrating integrability at $O(\lambda)$ (see also Refs. [58,59]).
- A strong integrability-breaking deformation

$$H_1 = \sum_{n=1}^{L}(-1)^n\sigma_n^z, \tag{8}$$

  which introduces a staggered magnetic field.

Notably, the weak integrability-breaking term (7) preserves $SU(2)$ symmetry and is an (renormalization group) RG-irrelevant perturbation [60] that leaves the spectrum gapless, whereas the strong breaking term (8) breaks $SU(2)$ and also translation symmetries, and it gives an RG-relevant perturbation (positive RG eigenvalue) that opens a gap.

We analyze several quantities to assess the quality of the approximation. As a basic test, we first compare the energy expectation value $E^{(\rm EBA)} = \langle\psi_{\rm EBA}|H(\lambda)|\psi_{\rm EBA}\rangle/\langle\psi_{\rm EBA}|\psi_{\rm EBA}\rangle$ obtained from the EBA with the exact energy from exact diagonalization (ED). For $L = 4$, we use the EBA method to obtain the full spectrum of both models[†], and the largest relative error occurs for the antiferromagnetic ground state. To evaluate how close the EBA state is to the exact eigenstate, we compute the fidelity $F = |\langle\psi_{\rm ex}|\psi_{\rm EBA}\rangle|^2/(\langle\psi_{\rm EBA}|\psi_{\rm EBA}\rangle\langle\psi_{\rm ex}|\psi_{\rm ex}\rangle)$, where $|\psi_{\rm EBA}\rangle$ is the EBA state and $|\psi_{\rm ex}\rangle$ is the exact eigenstate. A fidelity close to 1 indicates that the state serves as a good approximation to the exact state. Finally, since entanglement structure plays an important role in characterizing many-body wave functions, we also compute the bipartite entanglement entropy for different subsystem sizes and compare the results for the EBA state with those for the exact states. All these quantities evolve in a synchronized manner. In addition to physical quantities, we examine how the effective Bethe roots are modified under continuous deformation. The results for the two types of deformations are presented below.

[*]So far, a rigorous definition of weak and strong integrability-breaking perturbations is still lacking. In fact, for a sufficiently large deformation parameter, all perturbations lead to strong integrability breaking. In this context, the distinction between weak and strong should be understood qualitatively. For the same deformation parameter, some perturbations break integrability more rapidly than others.

[†]If one uses random initial values for the optimization, it will be very hard to find the state that corresponds to the singular physical Bethe state in the XXX model (in our convention, $u_1 = -i$, $u_2 = 0$). If one starts from the known solution to the BAE, then it will not be hard to identify it.

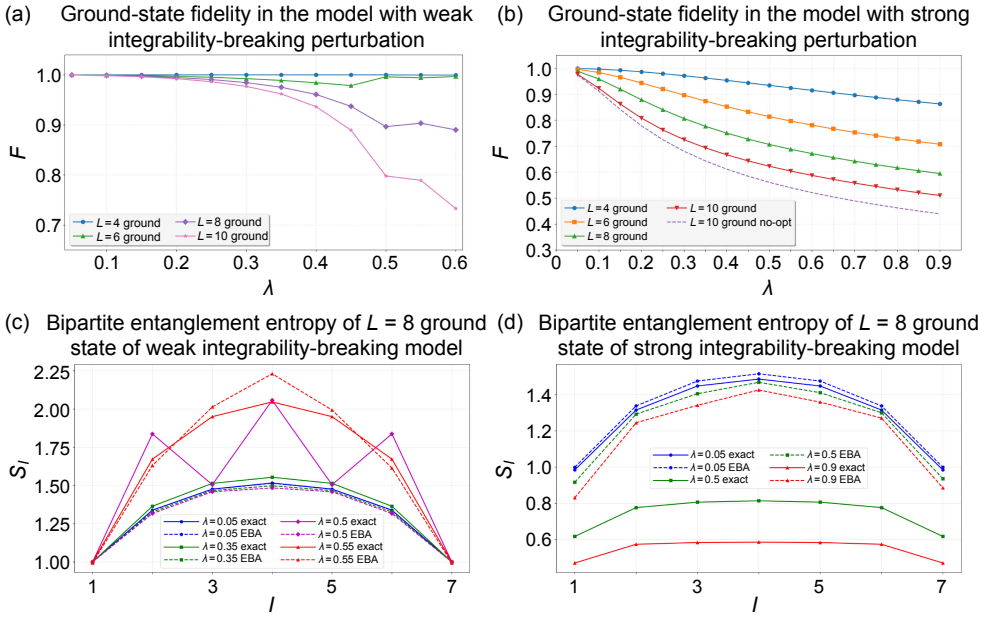

Fig. 1. Fidelity and Entanglement Entropy computed in two models.

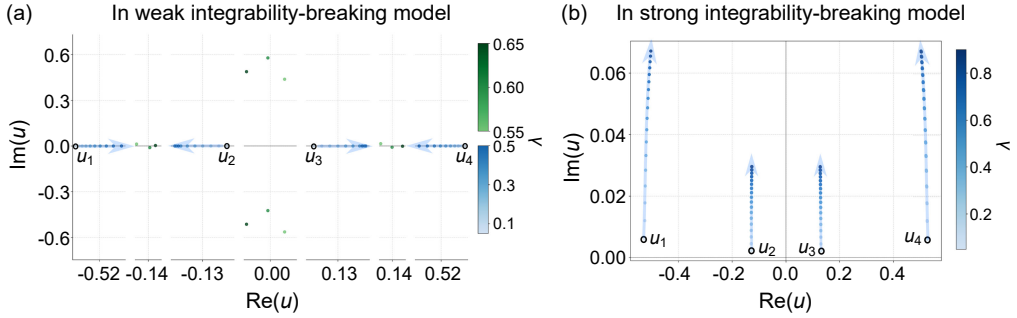

Fig. 2. Evolution of the effective Bethe roots of the ground state of two models with $L = 8$. The black circles mark the positions of the Bethe roots corresponding to the undeformed ground state. The arrows indicate the directions that the corresponding Bethe roots move as the deformation parameter increases. For the weak integrability-breaking model, the effective Bethe roots changed abruptly at the MG point $\lambda = 0.5$, indicated by the green dots in the figure.

Weak Integrability-Breaking

The energy for the ground state and the first excited state of different lengths are given in Table 1. We compared the result for $\lambda = 0.1$ with lengths from $L = 4$ to $L = 10$, the error of the energy is within 0.1% for the ground state and 0.3% for the 1st excited state. We have also computed at other deformation parameters ranging from 0.1 to 0.9 and higher lengths (for ground state at $\lambda = 0.1$, 0.3 up to $L = 20$, with errors bounded by 0.11% and 1.2%, respectively).

The fidelity of the ground state is shown in Fig. 1(a), which remains close to 1 in the region $\lambda < 0.5$. At $\lambda = 0.5$, the fidelity exhibits a sudden drop for all states. The $\lambda = 0.5$ is known as the Majumdar–Ghosh (MG) point [61]. In a finite chain, a level-crossing occurs at the MG point between the ground state and the 1st excited state. The entanglement entropy shown in Fig. 1(c), where an abrupt change is also observed at $\lambda = 0.5$[‡]. The distribution of the effective Bethe roots also undergoes a sudden change in pattern when crossing this point, as illustrated in Fig. 2(a). In the thermodynamic limit, the ground state and the 1st excited state in the finite chain merge into a twofold-degenerate ground state, and the system is known to lie in a dimerized phase [62]. The MG point is simply a special solvable point inside the dimerized phase, and the level-crossing observed does not lead to a quantum phase transition. There is indeed a phase transition point from a gapless spin-fluid phase to the gapped dimerized phase at $\lambda \approx 0.2411$ [63–65]. It is indicated by a similar abrupt slope change in the fidelity curve and a similar sudden jump of effective Bethe roots of the 1st excited state around $\lambda \sim 0.25$ in our approach [see Fig. 3(a); there is again a jump at the MG point]. These results suggest that the effective Bethe ansatz may serve as a probe for potential quantum phase transitions, even in relatively short spin chains.

Strong Integrability-Breaking

For comparison, in the model with a strong integrability-breaking deformation, the effectiveness of EBA is visibly reduced. The energy spectrum is given in Table 2, where we take the same value of deformation parameter $\lambda = 0.1$. A mismatch already occurs for the first decimal digit for $L = 10$ with ~1% error. The rate of breakdown of the EBA approximation thus serves as a diagnostic for the type of perturbations.

As shown in Fig. 1(b), the fidelity rapidly drops below 0.9 once the perturbation is turned on to $\lambda \sim 0.1$, and the entanglement entropy similarly deviates from its exact value around the same point [see Fig. 1(d)]. This indicates that the integrability structure is more

[‡]The ground states at the MG point $\lambda = 0.5$ are twofold-degenerate and dimerized. The following states, $|D_A\rangle = \otimes_{j=1}^{L/2}|\text{singlet}\rangle_{2j-1,2j}$ and $|D_B\rangle = \otimes_{j=1}^{L/2}|\text{singlet}\rangle_{2j,2j+1}$, are eigenstates at the MG point. A clear even-odd effect is evident in the bipartite entanglement entropy of these states. Taking one end of the region between the 1st and the L-th site, the entanglement entropy alternates between $\log_2 2 = 1$ (when cutting through a singlet pair) and 0 for $|D_A\rangle$, and between $\log_2 2 = 1$ and $2\log_2 2 = 2$ for $|D_B\rangle$. Figure 1(c) gives the bipartite entanglement entropy of the state that has large overlap with the EBA state. It is a linear combination of $|D_A\rangle$ and $|D_B\rangle$ (which are not orthogonal to each other), and its entanglement entropy varies between 1 and about 2 with an even-odd effect. Unfortunately, the EBA does not capture such an effect at this special point.

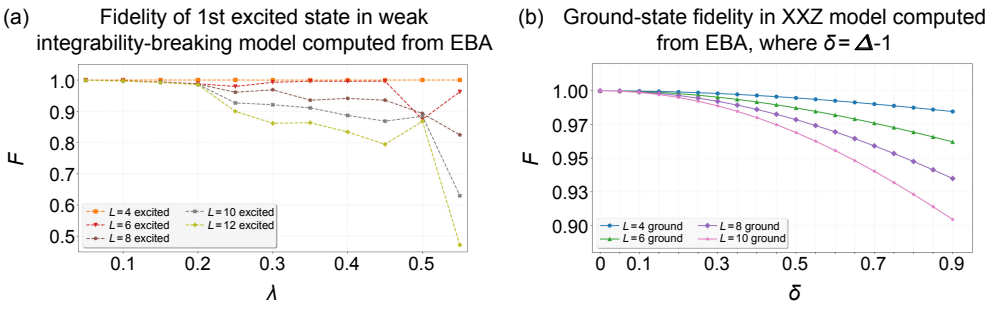

Fig. 3. Fidelity computed in the weak integrability-breaking model and the XXZ model.

Table 1. The ground state energy ($M = L/2$) and the first-excited state energy ($M = L/2-1$) computed from ED, our EBA method, and the 1st-order perturbation theory for the weakly integrability-broken model at $\lambda = 0.1$. For comparison, the exact energy levels of the antiferromagnetic ground state $E_0^{H_0}$ and the first excited state $E_1^{H_0}$ of the Heisenberg model H_0 are also listed. $E^{(\text{no-opt})}$ stands for the energy expectation value of the XXX Bethe states (without optimization of Bethe roots). $\text{Var}(H)/\Delta E^2$ measures eigenstate quality with ΔE denoting the energy difference between the lowest and next-to-lowest levels in the same sector as the target state.

L	4	6	8	10
$E_0^{H_0}$	−8	−11.2111	−14.6044	−18.0618
$E_0^{(\text{ED})}$	−7.6	−10.7201	−13.9896	−17.3147
$E_0^{(\text{EBA})}$	−7.6	−10.7119	−13.9779	−17.2996
$\text{Var}(H)/\Delta E^2$	0	1.5×10^{-2}	3.1×10^{-2}	5.8×10^{-2}
$E_0^{(\text{no-opt})}$	−7.6	−10.7119	−13.9779	−17.2996
$E_0^{(\text{1st-pt})}$	−7.6	−10.7119	−13.9779	−17.2996
$E_1^{H_0}$	−4	−8.4721	−12.5137	−16.3688
$E_1^{(\text{ED})}$	−3.6	−8.0448	−11.9742	−15.6963
$E_1^{(\text{EBA})}$	−3.6	−8.0253	−11.9531	−15.6740
$\text{Var}(H)/\Delta E^2$	0	2.1×10^{-2}	3.8×10^{-2}	6.4×10^{-2}
$E_1^{(\text{no-opt})}$	−3.6	−8.0249	−11.9529	−15.6739
$E_1^{(\text{1st-pt})}$	−3.6	−8.0249	−11.9529	−15.6739

strongly disturbed under such a deformation. For $M = 1$, one can easily understand it: the Bethe ansatz gives a plane wave, while the translational invariance is broken in this model.

We also note that the effect of optimization of Bethe roots is more evident in the strong integrability-breaking model [see Tables 1 and 2, $E^{(\text{EBA})}$ vs $E^{(\text{no-opt})}$]. As expected, optimization becomes increasingly beneficial for larger deformation λ and size L [see Figs. 1(b) and B1, no-opt curves]. In the weak integrability-breaking model, the optimization improves the fidelity of the ground state only for $\lambda \gtrsim 0.4$, and becomes indispensable beyond the MG point $\lambda = 0.5$, where the Bethe roots dive into the complex plane.

In perturbation theory, when the unperturbed spectrum contains degenerate levels, the correct approximation eigenstates of the perturbed system are generally given by linear combinations of the degenerate unperturbed states at the leading order. To handle such cases, the effective Bethe ansatz must be upgraded: instead of a single off-shell Bethe state $|\{u_j\}\rangle$, a superposition of multiple off-shell Bethe states $\sum_n c_n|\{u_j^{(n)}\}\rangle$ in general needs to be employed as the variational ansatz. Through the optimization process, the EBA algorithm then naturally constructs the correct linear combinations that lift the degeneracy.

For instance, in the simple case of $L = 4$ chain with a staggered magnetic field $\lambda = 0.1$, our method identifies the perturbed eigenstates as a superposition of two degenerate states from the integrable XXX point ($|\bar{u}_1 = \pm 0.5\rangle$ with $E = 0$), yielding

Table 2. The ground state energy ($M = L/2$) and the first-excited state energy ($M = L/2-1$) computed from ED, our EBA method, and the 1st-order perturbation theory for the strongly integrability-breaking model at $\lambda = 0.1$. For comparison, the exact energy levels of the antiferromagnetic ground state $E_0^{H_0}$ and the first excited state $E_1^{H_0}$ of the Heisenberg model H_0 are also listed. The no-opt rows give the energies from the XXX Bethe states without optimization of Bethe roots.

L	4	6	8	10
$E_0^{H_0}$	−8	−11.2111	−14.6044	−18.0618
$E_0^{(\text{ED})}$	−8.0265	−11.2753	−14.7211	−18.2433
$E_0^{(\text{EBA})}$	−8.0089	−11.2226	−14.6181	−18.0774
$\text{Var}(H)/\Delta E^2$	5.7×10^{-3}	2.2×10^{-2}	5.3×10^{-2}	9.5×10^{-2}
$E_0^{(\text{no-opt})}$	−8	−11.2111	−14.6044	−18.0618
$E_0^{(\text{1st-pt})}$	−8	−11.2111	−14.6044	−18.0618
$E_1^{H_0}$	−4	−8.4721	−12.5137	−16.3688
$E_1^{(\text{ED})}$	−4.0050	−8.4851	−12.5380	−16.4080
$E_1^{(\text{EBA})}$	−4.0017	−8.4747	−12.5171	−16.3729
$\text{Var}(H)/\Delta E^2$	1.3×10^{-3}	7.9×10^{-3}	2.1×10^{-2}	4.1×10^{-2}
$E_1^{(\text{no-opt})}$	−4	−8.4721	−12.5137	−16.3688
$E_1^{(\text{1st-pt})}$	−4	−8.4721	−12.5137	−16.3688

$$\frac{|\bar{u}_1 = 0.5\rangle \pm \mathrm{i}|\bar{u}_1 = -0.5\rangle}{\sqrt{2}} \tag{9}$$

with energies $E = \mp 0.2$. This behavior is reminiscent of the Stark effect in conventional perturbation theory. The single Bethe states $|\bar{u}_1 = \pm 0.5\rangle$ are no longer eigenstates in the perturbed system.

Improvements of the Ansatz

So far we have used the off-shell Bethe states of the XXX spin chain as an ansatz, where the tunable parameters are the effective Bethe roots. It is possible to improve the ansatz by introducing more tunable parameters while preserving integrability. In this section, we discuss a few natural generalizations of the ansatz.

More General R-Matrix

The off-shell Bethe ansatz of XXX model is constructed by using the R-matrix of the isotropic 6-vertex model. It is straightforward to generalize this construction by employing the anisotropic or q-deformed R-matrix of XXZ type. In this way, we can introduce q as another tunable parameter.

Interestingly, one may use the XXX off-shell Bethe state to approximate an on-shell Bethe state of the XXZ spin chain. As the anisotropic parameter $\Delta = (q + q^{-1})/2$ deviates from 1, the fidelity decreases slowly and smoothly, as shown in Fig. 3(b). Naturally, if we include q as an optimization parameter, we achieve a perfect match with the exact XXZ eigenstate.

In this example, the imperfection of the EBA originates from an

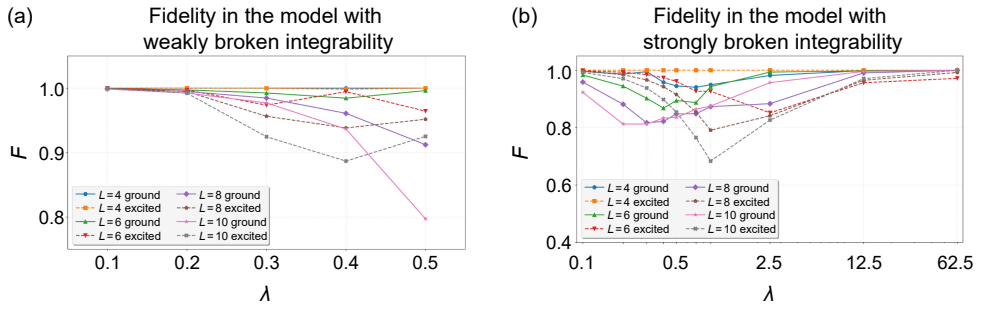

Fig. 4. Fidelity computed in two models based on an EBA generated from the 8-vertex R-matrix.

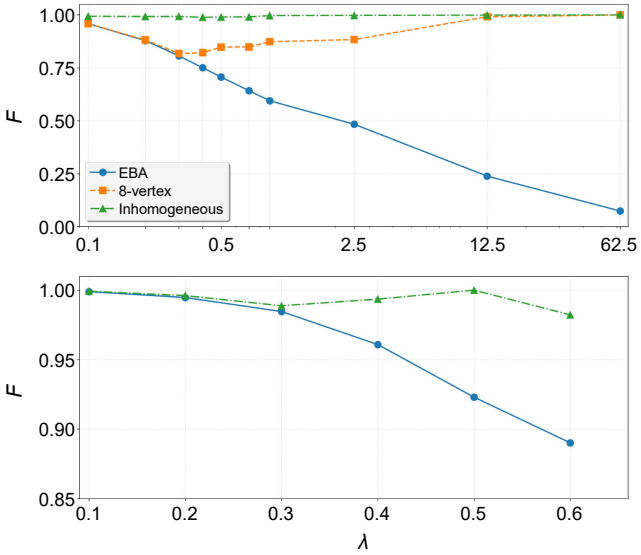

Fig. 5. Comparison of ground-state fidelity F for three ansätze in the $L = 8$ chain. Upper panel: strong-breaking model; lower panel: weak-breaking model. In the weak-breaking case, the 8-vertex extension offers little improvement over the standard EBA for the ground state (which is difficult to discern in the plot), whereas the inhomogeneous ansatz maintains much better fidelity $F \geq 0.98$ across the full parameter regime in the lower figure.

inappropriate ansatz, instead of integrability breaking. This motivates us to adopt the following more general 8-vertex ansatz, which introduces two additional parameters:

$$R(u;\beta,\gamma) = \begin{pmatrix} u+\frac{\mathrm{i}(1+\beta)}{2} & & & \frac{\mathrm{i}(1-\gamma)}{2} \\ & u+\frac{\mathrm{i}(1-\beta)}{2} & \mathrm{i} & \\ & \mathrm{i} & u+\frac{\mathrm{i}(1-\beta)}{2} & \\ \frac{\mathrm{i}(1-\gamma)}{2} & & & u+\frac{\mathrm{i}(1+\beta)}{2} \end{pmatrix}.$$

By optimizing both the Bethe roots and the parameters β and γ, we obtained a significantly improved approximation. For the XXZ model, the fidelity can be enhanced to unity, with an error of order of 10^{-13}. The optimized value of γ is very close to 1, consistent with the fact that XXZ spin chains correspond to the 6-vertex model. The improvement can also be seen in models where integrability is broken. In the weak integrability-breaking case, the fidelity is slightly improved towards 1 at a generic point, and a mysterious improvement for the 1st excited state is observed at the MG point $\lambda = 0.5$ [see Fig. 4(a)]. An unexpected phenomenon is observed in the strong integrability-breaking model: the fidelity revives to 1 when the magnetic field becomes large enough [see Fig. 4(b)]. We remark that although the 8-vertex "R-matrix" employed here is generally not a solution to the Yang–Baxter equation [§], an interesting observation is that for the 1st excited state in the strong integrability-breaking model with $L = 10$ and $\lambda \sim 1500$, most components of the Yang–Baxter equation evaluated with the optimized R-matrix are found to be very close to zero. This suggests that some sector of the free Hamiltonian $H = \sum_{n=1}^{L}(-1)^n\sigma_n^z$ might also be described by the 8-vertex R-matrix.

Inhomogeneities

The weak integrability-breaking model features long-range interactions beyond the nearest-neighbor level. To approximate long-range interacting spin chains, a natural extension of our ansatz is to further incorporate inhomogeneous parameters $\{\theta_j\}$ into the transfer matrix (see Appendix for the definition) since the inhomogeneous model is intrinsically long-range interacting. Using the EBA generated from this transfer matrix increases the fidelity F to $F \geq 0.99$ for all values of λ in the strong integrability-breaking case and $F \geq 0.98$ in the region $0 \leq \lambda \leq 0.6$ in the weakly breaking model (see Fig. 5). This improvement comes at the cost of optimizing both the inhomogeneities and the rapidities, which raises the computational costs. An optimal compromise between computational cost and accuracy is to make a proper choice of the inhomogeneity, which reduces the number of tunable parameters.

Discussions and Outlook

We have introduced an EBA for quantum systems in the vicinity of an integrable point. The central idea is that sufficiently close to integrability, the functional form of the Bethe wave function remains valid, while the Bethe roots undergo renormalization due to integrability-breaking perturbations. These effective roots are determined by minimizing physically motivated cost functions. The resulting effective Bethe states provide approximations to the exact eigenstates of non-integrable models. With a suitable refinement of the ansatz, the approximation can be remarkably accurate.

The idea of approximating exact eigenstates of non-integrable models by appropriate off-shell Bethe states offers numerous avenues for further development. For a given non-integrable model, what is the optimal choice of off-shell Bethe state and cost function? It is plausible that different choices of cost functions may be suited to different purposes. For example, in the XXZ model, one can also start from the XX point to perform the EBA analysis, and for $\Delta \lesssim 0.5$, it outperforms the EBA based on the XXX ansatz. Since the XX model maps exactly to free fermions, the EBA of XX version is essentially the same as the Hartree–Fock approximation, and similar ideas have been widely applied in the study of quantum many-body physics by using the Slater–Gaussian state (or its superposition) as a variational ansatz [66–69]. To gain a deeper understanding, it is necessary to explore more general classes of models and a broader range of physical

[§]From the perspective of variational method, introducing more parameters generally yields better approximations (with higher computational costs), thus the 8-vertex ansatz adopted here is a naive extension of the original EBA. While the ansatz itself can be chosen arbitrarily, when the R-matrix satisfies the Yang–Baxter equation (e.g., the R-matrix of the XXZ model), it may further offer some physical interpretation to the results.

observables. A natural extension of this work is to apply the method to spin-1 models, such as those arising from deformations of the integrable Takhtajan–Babujian model [70,71], which includes the AKLT model [72,73] as a special case. Progress in this direction has been made and is presented in Ref. [74].

A deeper physical understanding of why the off-shell Bethe ansatz or the factorized scattering picture breaks down for sufficiently strong integrability-breaking interactions would be desirable. In quantum field theory, away from integrability, stable particles may decay, and particle production and annihilation processes become allowed, rendering scattering inelastic. Can such a mechanism also be realized in the spin chain setting? Or are there other underlying mechanisms at play? Answering these questions would deepen our understanding of integrability breaking and may provide guidance for developing improved ansätze further from integrability.

The algebraic structure of off-shell Bethe states may prove useful for investigating a wider class of physical observables. One important direction concerns dynamical properties, such as quench dynamics and quantum transport. How well does the approximation hold under time evolution? Does the fidelity decay over time or remain robust? Addressing these questions could yield new insights into phenomena such as prethermalization and anomalous transport.

Acknowledgements

We are grateful for the valuable feedback received during the International Workshop on “Challenges in Integrability”, where part of this work was presented. This work was supported by the National Natural Science Foundation of China (Grant Nos. 12247103, 12575073 for Y. J., and 12105198 partially for R. Z.).

References

[1] Kolmogorov A N 1954 *Proc. USSR Acad. Sci.* **98** 527
[2] Arnol’d V I 1963 *Russ. Math. Surv.* **18** 9
[3] Moser J 1962 *Matematika* **6** 51
[4] Bargheer T, Beisert N, and Loebbert F 2008 *J. Stat. Mech.* **2008** L11001
[5] Bargheer T, Beisert N, and Loebbert F 2009 *J. Phys. A* **42** 285205
[6] Mierzejewski M, Prosen T, and Prelovšek P 2015 *Phys. Rev. B* **92** 195121
[7] Malikis S, Kurlov D, and Gritsev V 2020 arXiv:2007.01715 [cond-mat.stat-mech]
[8] Vanovac S, Surace F M, and Motrunich O I 2024 *Phys. Rev. B* **110** 144309
[9] Brandino G P, Caux J S, and Konik R M 2015 *Phys. Rev. X* **5** 041043
[10] Berges J, Borsanyi S, and Wetterich C 2004 *Phys. Rev. Lett.* **93** 142002
[11] Kollar M, Wolf F A, and Eckstein M 2011 *Phys. Rev. B* **84** 054304
[12] Langen T, Gasenzer T, and Schmiedmayer J 2016 *J. Stat. Mech.* **2016** 064009
[13] Mallayya K, Rigol M, and Roeck W D 2019 *Phys. Rev. X* **9** 021027
[14] Moeckel M and Kehrein S 2008 *Phys. Rev. Lett.* **100** 175702
[15] Žnidarič M 2011 *Phys. Rev. Lett.* **106** 220601
[16] Mierzejewski M, Pawłowski J, Prelovsek P, and Herbrych J 2022 *SciPost Phys.* **13** 013
[17] Jung P and Rosch A 2007 *Phys. Rev. B* **76** 245108
[18] Huang Y, Karrasch C, and Moore J E 2013 *Phys. Rev. B* **88** 115126
[19] Karrasch C, Kennes D M, and Heidrich-Meisner F 2015 *Phys. Rev. B* **91** 115130
[20] Steinigeweg R, Herbrych J, Zotos X, and Brenig W 2016 *Phys. Rev. Lett.* **116** 017202
[21] Durnin J, Bhaseen M J, and Doyon B 2021 *Phys. Rev. Lett.* **127** 130601
[22] Bastianello A, De Luca A, and Vasseur R 2021 *J. Stat. Mech.* **2021** 114003
[23] Zamolodchikov A B 1989 *Int. J. Mod. Phys. A* **4** 4235
[24] Delfino G, Mussardo G, and Simonetti P 1996 *Nucl. Phys. B* **473** 469
[25] Yurov V P and Zamolodchikov A B 1991 *Int. J. Mod. Phys. A* **6** 4557
[26] Yurov V P and Zamolodchikov A B 1990 *Int. J. Mod. Phys. A* **5** 3221
[27] Fonseca P and Zamolodchikov A 2001 arXiv:hep-th/0112167
[28] Bethe H 1931 *Z. Phys.* **71** 205
[29] Lieb E H and Liniger W 1963 *Phys. Rev.* **130** 1605
[30] Yang C N 1967 *Phys. Rev. Lett.* **19** 1312
[31] Baxter R J 1982 *Exactly solved models in statistical mechanics* (London, Academic Press)
[32] Nepomechie R I 2021 *Quantum Inf. Comput.* **21** 255
[33] Raveh D and Nepomechie R I 2024 *J. Phys. A* **57** 355303
[34] De Palo S, Citro R, and Orignac E 2020 *Phys. Rev. B* **101** 045102
[35] Claeys P W, Caux J S, Van Neck D, and De Baerdemacker S 2017 *Phys. Rev. B* **96** 155149
[36] Surace F M and Motrunich O 2023 *Phys. Rev. Res.* **5** 043019
[37] Pozsgay B, Jiang Y, and Takács G 2020 *JHEP* **03** 092
[38] Marchetto E, Sfondrini A, and Yang Z 2020 *Phys. Rev. Lett.* **124** 100601
[39] Doyon B, Durnin J, and Yoshimura T 2022 *SciPost Phys.* **13** 072
[40] Jung P, Helmes R W, and Rosch A 2006 *Phys. Rev. Lett.* **96** 067202
[41] Mallayya K and Rigol M 2021 *Phys. Rev. B* **104** 184302
[42] Mori T, Ikeda T N, Kaminishi E, and Ueda M 2018 *J. Phys. B* **51** 112001
[43] Rabson D A, Narozhny B N, and Millis A J 2004 *Phys. Rev. B* **69** 054403
[44] Santos L F and Rigol M 2010 *Phys. Rev. E* **81** 036206
[45] Brandino G P, Konik R M, and Mussardo G 2010 *J. Stat. Mech.* **2010** P07013
[46] Robinson N J, Essler F H L, Cabrera I, and Coldea R 2014 *Phys. Rev. B* **90** 174406
[47] McCoy B M and Wu T T 1978 *Phys. Rev. D* **18** 1259
[48] Fonseca P and Zamolodchikov A 2006 arXiv:hep-th/0612304
[49] Kormos M, Collura M, Takács G, and Calabrese P 2017 *Nat. Phys.* **13** 246
[50] Rutkevich S B 2024 *Phys. Rev. B* **109** 014411
[51] Lagnese G, Surace F M, Kormos M, and Calabrese P 2022 *J. Phys. A* **55** 124003
[52] Gao Y, Jiang Y, and Wu J 2025 *JHEP* **07** 072
[53] Litvinov A, Meshcheriakov P, and Shestopalov E 2025 *Phys. Rev. D* **112** 085021
[54] Coleman S R 1977 *Phys. Rev. D* **15** 2929
[55] Callan J C G and Coleman S R 1977 *Phys. Rev. D* **16** 1762
[56] Lagnese G, Surace F M, Kormos M, and Calabrese P 2021 *Phys. Rev. B* **104** L201106
[57] Szász-Schagrin D and Takács G 2022 *Phys. Rev. D* **106** 025008
[58] Serban D 2013 *JHEP* **08** 128
[59] Orlov P, Tiutiakina A, Sharipov R, Petrova E, Gritsev V, and Kurlov D V 2023 *Phys. Rev. B* **107** 184312
[60] Beria M, Brandino G P, Lepori L, Konik R M, and Sierra G 2013 *Nucl. Phys. B* **877** 457
[61] Majumdar C K and Ghosh D K 1969 *J. Math. Phys.* **10** 1388
[62] Haldane F D M 1982 *Phys. Rev. B* **25** 4925
[63] Tonegawa T and Harada I 1987 *J. Phys. Soc. Jpn.* **56** 2153
[64] Okamoto K and Nomura K 1992 *Phys. Lett. A* **169** 433
[65] Liu G H, Wang C H, and Deng X Y 2011 *Physica B* **406** 100
[66] Bravyi S and Gosset D 2017 *Commun. Math. Phys.* **356** 451
[67] Boutin S and Bauer B 2021 *Phys. Rev. Res.* **3** 033188
[68] Buijsman W and Lev Y B 2024 *SciPost Phys.* **17** 055
[69] Albert N, Zhang Y, and Tu H H 2024 arXiv:2410.04204 [cond-mat.str-el]
[70] Babujian H M 1982 *Phys. Lett. A* **90** 479
[71] Takhtajan L A 1982 *Phys. Lett. A* **87** 479
[72] Affleck I, Kennedy T, Lieb E H, and Tasaki H 1987 *Phys. Rev. Lett.* **59** 799
[73] Affleck I, Kennedy T, Lieb E H, and Tasaki H 1988 *Commun. Math. Phys.* **115** 477
[74] Wang Z and Zhu R D 2026 *J. Phys. A* **59** 305001
[75] Faddeev L D 1995 *Int. J. Mod. Phys. A* **10** 1845
[76] Slavnov N A 2022 *Algebraic Bethe ansatz and Correlation Functions* (Singapore: World Scientific)

Appendix

Appendix A: Algebraic Bethe Ansatz

This appendix provides a concise review of the algebraic Bethe ansatz (ABA) for the Heisenberg spin chain. For a more detailed introduction, we refer to Refs. [75,76].

XXX Spin Chain. We begin with the 6-vertex R-matrix, which acts on the tensor product space $\mathbb{C}_a^2 \otimes \mathbb{C}_b^2$ and is given by

$$R_{ab}(u) = \begin{pmatrix} u+\mathrm{i} & 0 & 0 & 0 \\ 0 & u & \mathrm{i} & 0 \\ 0 & \mathrm{i} & u & 0 \\ 0 & 0 & 0 & u+\mathrm{i} \end{pmatrix}. \tag{A1}$$

The central object in the ABA framework is the monodromy matrix, defined as

$$M_a(u) \equiv R_{aL}(u-\theta_L)\ldots R_{a1}(u-\theta_1), \tag{A2}$$

where L is the length of the spin chain and 'a' denotes a two-dimensional auxiliary space, the parameters $\{\theta_i\}$ are called inhomogeneities. For the special case $\theta_1 = \theta_2 = \ldots = \theta_L$, we recover the homogeneous model. In this auxiliary space, the monodromy matrix can be expressed as a 2×2 matrix

$$M_a(u) = \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}_a. \tag{A3}$$

The transfer matrix, which serves as the generating function for the conserved charges, is obtained by taking the trace over the auxiliary space,

$$T(u) = \mathrm{Tr}_a M_a(u) = A(u) + D(u), \tag{A4}$$

and has the fundamental property that $[T(u), T(v)] = 0$. The Hamiltonian of the XXX chain is then derived from the transfer matrix of the homogeneous model (with $\theta_j = \mathrm{i}/2$) as

$$H_{\mathrm{XXX}} \propto \frac{\mathrm{d}}{\mathrm{d}u} \log \boldsymbol{T}(u)\Big|_{u=\mathrm{i}/2}. \tag{A5}$$

Consequently, eigenvectors of the transfer matrix are also eigenvectors of the Hamiltonian. Such eigenstates can be constructed by repeatedly acting with the B-operator on the pseudovacuum state $|\Omega\rangle = |\uparrow^L\rangle$,

$$|\{u_j\}\rangle = B(u_1)\ldots B(u_N)|\Omega\rangle. \tag{A6}$$

For this state to be an eigenstate of the transfer matrix, the complex parameters $\{u_j\}$ must satisfy the Bethe ansatz equations (BAE)

$$\left(\frac{u_j + \frac{\mathrm{i}}{2}}{u_j - \frac{\mathrm{i}}{2}}\right)^L = \prod_{\substack{k=1 \\ k\neq j}}^{N} \frac{u_j - u_k + \mathrm{i}}{u_j - u_k - \mathrm{i}}, \quad j = 1, \cdots, N. \tag{A7}$$

A Bethe state whose rapidities satisfy the BAE is referred to as on-shell; otherwise, it is called off-shell.

XXZ Spin Chain. The XXZ spin chain is defined by the Hamiltonian

$$H_{\mathrm{XXZ}} = \sum_{n=1}^{L} (\sigma_n^x \sigma_{n+1}^x + \sigma_n^y \sigma_{n+1}^y + \Delta \sigma_n^z \sigma_{n+1}^z), \tag{A8}$$

where Δ is the anisotropic parameter and $\Delta = 1$ corresponds to the XXX spin chain. The XXZ spin chain can also be solved by ABA, with the following R-matrix

$$R_{ab}(u) = \begin{pmatrix} \sinh(u+\eta) & 0 & 0 & 0 \\ 0 & \sinh u & \sinh\eta & 0 \\ 0 & \sinh\eta & \sinh u & 0 \\ 0 & 0 & 0 & \sinh(u+\eta) \end{pmatrix}, \tag{A9}$$

with an additional parameter η. This parameter is related to the q-deformation parameter and anisotropy by

$$q = \mathrm{e}^{\eta}, \ \Delta = \frac{1}{2}(q+q^{-1}) = \cosh\eta. \tag{A10}$$

The rest of the construction follows the same way as the XXX spin chain and will not be repeated here.

Appendix B: Numerical Optimization of Effective Bethe Roots

This appendix briefly outlines the numerical algorithm employed to optimize the parameters of the effective Bethe ansatz. The goal is to find, through variational optimization of $\mathcal{H}_n$, a set of complex parameters, including the effective Bethe roots $\{u_j\}$ and other shared algebraic parameters such as β, γ, and inhomogeneous parameters $\{\theta_j\}$ potentially existing when the ansatz is further improved.

The algorithm we use is implemented in Python. The core computations rely on the following libraries:

- **SciPy**: Used for sparse matrix representation of the Hamiltonian and for exact diagonalization to obtain benchmark energies and states to be compared with. Its module, scipy.optimize.minimize, provides an efficient optimizer, L-BFGS-B optimizer, a quasi-Newton method suitable for bound-constrained, medium-dimensional problems.
- **PyTorch**: Used to implement the construction of the effective Bethe state and the computation of the energy expectation value. Its automatic differentiation capability is leveraged to compute the exact gradient of the loss function with respect to all $2M$ $(+\alpha)$ real parameters, which is crucial for efficient optimization convergence. For relatively large systems (e.g., $L \sim 20$), where the Hilbert space dimension grows exponentially, the vectorized operations and reverse-mode automatic differentiation in PyTorch provide a significant speed advantage over manual gradient implementation or finite-difference methods.

To ensure robust convergence and avoid suboptimal local minima, the procedure employs the following initialization and search strategies:

- Random Initialization: The real and imaginary parts of effective

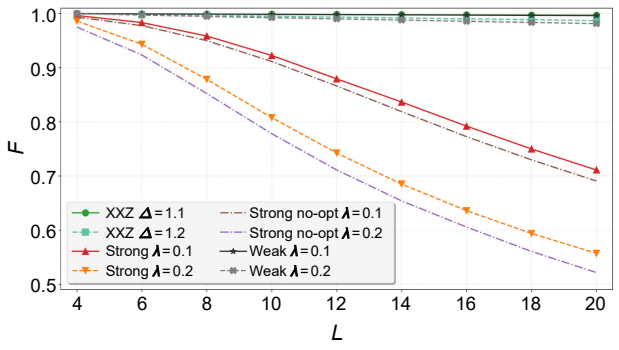

Fig. B1. Ground-state fidelity F versus system size L for the XXZ chain ($\Delta = 1.1, 1.2$) and the integrability-breaking models (strong and weak, $\lambda = 0.1$, 0.2). The curves are fitted with $F = a + bL^1 + cL^2$. The coefficients (a,b,c) are found to be: $\mathrm{XXZ}_{\Delta=1.1}$: $(1.000389, -1.47\times10^{-4}, -2\times10^{-6})$, $\mathrm{XXZ}_{\Delta=1.2}$: $(1.001526, -5.66\times10^{-4}, -8\times10^{-6})$, $\mathrm{strong}_{\lambda=0.1}$: $(1.041342, -7.968\times10^{-3}, -4.47\times10^{-4})$, $\mathrm{strong}_{\lambda=0.2}$: $(1.136773, -3.7269\times10^{-2}, 3.97\times10^{-4})$, $\mathrm{weak}_{\lambda=0.1}$: $(1.001017, -2.76\times10^{-4}, 2\times10^{-6})$, $\mathrm{weak}_{\lambda=0.2}$: $(1.005148, -1.364\times10^{-3}, 9\times10^{-6})$. All fittings achieve $R^2 > 0.996$. The coefficients of L^2-term are almost negligible except for the strong breaking model. The no-opt curves give the fidelity from the XXX Bethe states without optimization of Bethe roots.

Bethe roots $\{u_j\}$ are independently sampled from a normal distribution with mean 0 and standard deviation 10 as the initial guess.

- Parallel Attempts: The program performs multiple independent optimization runs (e.g., 30), each with a different random initial point. Each optimization is subject to a maximum number of iterations (e.g., 6,000) and convergence tolerances based on function value and gradient (ftol=1e-12, gtol=1e-10).
- Selection of Best Result: After each run, the final energy expectation and the fidelity (overlap) with the exact ground and first excited states (obtained via exact diagonalization) are recorded. The solution with the lowest energy among all attempts is selected as the final effective Bethe state.
- High-Quality Initialization via BAE: For systems near the integrable point ($\lambda \ll 1$), a highly effective initialization is obtained by first numerically solving the BAE for the underlying integrable XXX model ($\lambda = 0$) for the target sector. The resulting set of Bethe roots $\{u_j^{(0)}\}$ provides an excellent initial guess for the variational parameters $\{u_j\}$, and this strategy places the optimizer in the vicinity of the global minimum, dramatically accelerating convergence.

This pipeline successfully produced high-fidelity approximations for systems near the integrable point as discussed in the main text, and in practice, we have checked its effectiveness up to $L = 24$ (we managed to run our program for $L = 26$, but exact diagonalization was not available on the same computer for comparison). Figure B1 shows the fidelity as a function of system size L.