# Eigenvalue-based method and form-factor determinant representations for integrable XXZ Richardson-Gaudin models

Pieter W. Claeys,[1,2,*] Stijn De Baerdemacker,[1,2,3] Mario Van Raemdonck,[1,2,3] and Dimitri Van Neck[1,2]

[1]*Center for Molecular Modeling, Ghent University, Technologiepark 903, 9052 Ghent, Belgium*
[2]*Department of Physics and Astronomy, Ghent University, Proeftuinstraat 86, 9000 Ghent, Belgium*
[3]*Department of Inorganic and Physical Chemistry, Ghent University, Krijgslaan 281 (S3), 9000 Ghent, Belgium*

We propose an extension of the numerical approach for integrable Richardson-Gaudin models based on a new set of eigenvalue-based variables [A. Faribault *et al.*, Phys. Rev. B **83**, 235124 (2011); O. El Araby *et al.*, *ibid.* **85**, 115130 (2012)]. Starting solely from the Gaudin algebra, the approach is generalized towards the full class of XXZ Richardson-Gaudin models. This allows for a fast and robust numerical determination of the spectral properties of these models, avoiding the singularities usually arising at the so-called singular points. We also provide different determinant expressions for the normalization of the Bethe ansatz states and form factors of local spin operators, opening up possibilities for the study of larger systems, both integrable and nonintegrable. These expressions can be written in terms of the new set of variables and generalize the results previously obtained for rational Richardson-Gaudin models [A. Faribault and D. Schuricht, J. Phys. A **45**, 485202 (2012)] and Dicke-Jaynes-Cummings-Gaudin models [H. Tschirhart and A. Faribault, J. Phys. A **47**, 405204 (2014)]. Remarkably, these results are independent of the explicit parametrization of the Gaudin algebra, exposing a universality in the properties of Richardson-Gaudin integrable systems deeply linked to the underlying algebraic structure.

## I. INTRODUCTION

Exactly solvable systems play an important role in the understanding of the physics of quantum many-body systems. They offer insight into the behavior of strongly correlated systems in ways that would otherwise be impossible. One class of these systems is the class of Richardson-Gaudin (RG) integrable systems, which can be derived from a generalized Gaudin algebra [1,2]. The pairing model in the reduced BCS approximation, used to describe superconductivity, has been shown to be RG integrable [3], as has the $p_x + i p_y$ pairing Hamiltonian [4,5], the central spin model [6], factorizable pairing models in heavy nuclei [7], an extended $d + id$ pairing Hamiltonian [8], and several atom-molecule Hamiltonians such as the inhomogeneous Dicke model [9,10]. For these models, diagonalizing the Hamiltonian in an exponentially scaling Hilbert space can be reduced to solving a set of nonlinear equations scaling linearly with system size. For most of these systems, computationally inexpensive expressions are also known for several form factors and overlaps, which can be used to investigate observables in systems where the Hilbert space is too large for traditional exact methods. Unfortunately, the Bethe ansatz or Richardson-Gaudin equations [6,11,12] that need to be solved are highly nonlinear and give rise to singularities, making a straightforward numerical solution challenging [13,14]. Several methods have been introduced as a way to resolve this difficulty, such as a change in variables [14,15], a (pseudo)deformation of the algebra [16,17], or a Heine-Stieltjes connection, reducing the problem to a differential equation [18]. The ground-state energy in the thermodynamic limit has also been obtained by treating the interaction as an effective temperature [19].

Our method consists of a generalization of the numerical method first proposed by Faribault *et al.* [20] for a set of nondegenerate rational RG models and later extended to degenerate rational models by El Araby *et al.* [21]. The equations for the Dicke model were independently presented by Babelon and Talalaev [22]. In this method, an alternative set of equations is derived in terms of variables related to the eigenvalues of the constants of motion. One of the advantages of this method is that it offers not only the solutions of the RG equations but also efficient numerical expressions for form factors and overlaps [23,24]. Such expressions for the rational model have already been used to study quantum quenches [25] and decoherence in quantum dots [26] and should facilitate the use of the XXZ model as a variational or projected approach for nonintegrable systems, similar to the use of the XXX model in quantum chemistry [27–29].

In the following, the results for the rational XXX model are generalized towards the full class of XXZ RG integrable systems, and possible applications are discussed.

## II. RICHARDSON-GAUDIN MODELS

### A. Definitions

Unlike its classical counterpart, no single unique definition is known for quantum integrability [30]. One class of systems usually denoted integrable is the set of RG integrable systems since these support as many (nontrivial) conserved operators commuting with the Hamiltonian as there are degrees of freedom in the system [1,31]. These systems can be diagonalized exactly by means of a Bethe ansatz wave function, where diagonalizing the full Hamiltonian can be reduced to solving a set of nonlinear coupled equations for the variables in the ansatz [6,11].

*pieterw.claeys@ugent.be

The families of RG integrable systems have their roots in a generalized Gaudin algebra [2,6] based on the $su(2)$ algebra of (quasi)spin operators [32]. Depending on the physical systems, they can either represent spin degrees of freedom, fermion pairs, or bosonic/fermionic Schwinger representations. The relevant (quasi)spin operators can be defined as satisfying the following commutation relations:

$$
\begin{aligned}
& [S_i^0, S_j^\dagger] = \delta_{ij} S_i^\dagger, \quad [S_i^0, S_j] = -\delta_{ij} S_i, \\
& [S_i^\dagger, S_j] = 2\delta_{ij} S_i^0,
\end{aligned}
\tag{1}
$$

where each separate algebra spans an $su(2)_i$ algebra associated with a level $i$ and irreducible representations (irrep) $|d_i, \mu_i\rangle$. For a set of $n$ levels, the RG integrable models are then defined by a set of $n$ mutually commuting operators,

$$
R_i = S_i^0 + g \sum_{k \neq i}^n \left[ \frac{1}{2} X_{ik} (S_i^\dagger S_k + S_i S_k^\dagger) + Z_{ik} S_i^0 S_k^0 \right].
\tag{2}
$$

Following Gaudin [6] and Dukelsky *et al.* [12], it is possible to obtain a set of conditions for which the set of operators $R_i$ commute mutually. These are given by

$$
X_{ij} = -X_{ji}, \quad Z_{ij} = -Z_{ji},
\tag{3}
$$

$$
X_{ij} X_{jk} - X_{ik}(Z_{ij} + Z_{jk}) = 0,
\tag{4}
$$

the so-called Gaudin equations. They were originally discovered by Gaudin as defining a general class of quadratic Hamiltonians in the spin variables, among which is the Gaudin magnet. The derivation by Dukelsky *et al.* differs from the derivation by Gaudin in the presence of the generator of the Cartan subalgebra $S_i^0$, which, however, does not influence these conditions. In fact, the Gaudin models are given by Dukelsky's constants of motion in the limit $g \to \infty$.

Gaudin mentioned three classes of solutions of these equations, where all classes consider $X_{ij}$ and $Z_{ij}$ to be odd functions of some real arbitrary parameters $X_{ij} = X(\epsilon_i, \epsilon_j)$. The physical interpretation of these parameters follows from the expression for the Hamiltonian constructed with these parameters.

(1) The rational model

$$
X_{ij} = Z_{ij} = \frac{1}{\epsilon_i - \epsilon_j}.
\tag{5}
$$

(2) The trigonometric model

$$
X_{ij} = \frac{1}{\sin(\epsilon_i - \epsilon_j)}, \quad Z_{ij} = \cot(\epsilon_i - \epsilon_j).
\tag{6}
$$

(3) The hyperbolic model

$$
X_{ij} = \frac{1}{\sinh(\epsilon_i - \epsilon_j)}, \quad Z_{ij} = \coth(\epsilon_i - \epsilon_j).
\tag{7}
$$

Here the rational model is also referred to as the XXX model [12], indicating that the coefficients in the expression for the constants of motion are identical for all three components of the spin. In the same vein, the trigonometric and hyperbolic models are called XXZ models. Alternative solutions are given

by

$$
\begin{aligned}
X_{ij} &= \frac{\sqrt{1 + 2\alpha\epsilon_i + \beta\epsilon_i^2}\sqrt{1 + 2\alpha\epsilon_j + \beta\epsilon_j^2}}{\epsilon_i - \epsilon_j}, \\
Z_{ij} &= \frac{1 + \alpha(\epsilon_i + \epsilon_j) + \beta\epsilon_i\epsilon_j}{\epsilon_i - \epsilon_j},
\end{aligned}
\tag{8}
$$

which was proposed by Richardson [33], and

$$
X_{ij} = \frac{2\sqrt{\epsilon_i \epsilon_j}}{\epsilon_i - \epsilon_j}, \quad Z_{ij} = \frac{\epsilon_i + \epsilon_j}{\epsilon_i - \epsilon_j},
\tag{9}
$$

which is a reparametrization of the hyperbolic model [2].

### B. Diagonalizing integrable Hamiltonians

Any general RG integrable Hamiltonian can be written as a linear combination of the constants of motion

$$
\hat{H} = \sum_{i=1}^n \eta_i R_i,
\tag{10}
$$

resulting in a wide variety of systems. The reduced BCS Hamiltonian can be found from the rational model [3], the $p_x + i p_y$ pairing Hamiltonian can be derived from the hyperbolic model [4,5], and the central spin model Hamiltonian is identical to one of the constants of motion of the rational model [6]. By introducing a bosonic degree of freedom, the inhomogeneous Dicke model can be found as a limiting case of the trigonometric model [9].

Since all constants of motion commute mutually, they also commute with any Hamiltonian that can be written as Eq. (10), so this reduces the diagonalization of the Hamiltonian to the diagonalization of one of the constants of motion. It can be shown that the eigenstates are given by a Bethe ansatz,

$$
|\psi_N\rangle = \prod_{\alpha=1}^N \left( \sum_{i=1}^n X_{i\alpha} S_i^\dagger \right) |\theta\rangle,
\tag{11}
$$

with $N$ being the number of excitations [34] and the vacuum state $|\theta\rangle = \otimes_{i=1}^n |d_i, -d_i\rangle$ being the tensor product of the lowest-weight irreps of each $su(2)_i$ copy, provided the RG equations

$$
1 + g \sum_{i=1}^n Z_{i\alpha} d_i - g \sum_{\beta \neq \alpha}^N Z_{\beta\alpha} = 0, \ \forall \alpha = 1, \ldots, N,
\tag{12}
$$

are satisfied, where the algebra has been extended to $X_{i\alpha} = X(\epsilon_i, x_\alpha)$ and $Z_{i\alpha} = Z(\epsilon_i, x_\alpha)$ by introducing a set of $N$ parameters $\{x_\alpha\}$. These parameters, also known as the RG variables or rapidities, fix the wave function and need to be determined from the RG equations (12). There are multiple strategies to derive this result. Richardson obtained these equations for the reduced level-independent BCS model starting from a variational approach [11,35], Gaudin started from integrability constraints [6], Zhou *et al.* derived these results starting from the algebraic Bethe ansatz [36], and Ortiz *et al.* used a commutator scheme based on the Gaudin algebra [2].

Once the RG equations have been solved and the rapidities have been determined, the eigenvalues of the constants of

motion $R_i$ can be evaluated as

$$r_i = d_i \left( -1 + g \sum_{k \neq i}^{n} Z_{ik} d_k + g \sum_{\beta=1}^{N} Z_{\beta i} \right), \quad \forall i = 1, \ldots, n.$$

(13)

Although the set of RG equations seems to be linear in the elements $Z_{i\alpha}$ and $Z_{\alpha\beta}$, these variables are still coupled through the Gaudin equations, leading to a set of coupled and nonlinear equations. As an example, the RG equations for the doubly degenerate ($d_i = 1/2, \forall i$) XXX model are given by

$$1 + \frac{g}{2} \sum_{i=1}^{n} \frac{1}{\epsilon_i - x_\alpha} - g \sum_{\beta \neq \alpha}^{N} \frac{1}{x_\beta - x_\alpha} = 0, \quad \forall \alpha = 1, \ldots, N.$$

(14)

These equations need to be solved for the set of rapidities $\{x_\alpha\}$ and are coupled and highly nonlinear. It has been shown that singular points arise in these equations at certain values of the coupling constant, where multiple rapidities $x_\alpha$ equal one of the single-particle levels $\epsilon_i$ [13–15]. It can be readily seen that both the second term and the third term in the RG equations contain diverging poles, but it can also be shown that these singularities cancel exactly. Unfortunately, the behavior arising at these so-called singular points hampers straightforward solutions of these equations, so involved numerical methods have to be found to circumvent these points.

### III. AN EIGENVALUE-BASED NUMERICAL METHOD

#### A. Doubly degenerate models

For the XXX spin-$1/2$ model ($d_i = 1/2, \forall i$), it is possible to introduce a new set of variables

$$\Lambda_i \equiv \Lambda(\epsilon_i) = \sum_{\alpha=1}^{N} \frac{1}{\epsilon_i - x_\alpha},$$

(15)

circumventing the singular points in the Richardson-Gaudin equations [20]. A set of equations equivalent to the RG equations can be found for these variables; however, these equations are void of singular behavior. As an illustration, the set of RG equations for the doubly degenerate XXX model [Eq. (14)] is isomorphic to the set of quadratic equations [20]

$$\Lambda_i^2 = -\frac{2}{g} \Lambda_i + \sum_{j \neq i}^{n} \frac{\Lambda_j - \Lambda_i}{\epsilon_j - \epsilon_i}, \quad \forall i = 1, \ldots, n.$$

(16)

A straightforward numerical solution of these equations can easily be implemented, and no singular behavior will arise since no variables occur in the denominator. We wish to extend these results to the class of more general XXZ models, relying on only the Gaudin algebra and not on an explicit rational expression of $X$ and $Z$. Define

$$\Lambda_i \equiv \sum_{\alpha=1}^{N} Z_{i\alpha} = \sum_{\alpha=1}^{N} Z(\epsilon_i, x_\alpha),$$

(17)

which reduces to the previously introduced variables in the XXX model. It is interesting to note that the eigenvalue $r_i$ of

the constant of motion $R_i$ is related to the variable $\Lambda_i$ as

$$r_i = d_i \left( -1 - g\Lambda_i + g \sum_{k \neq i}^{n} Z_{ik} d_k \right),$$

(18)

which has led to the denomination eigenvalue-based variables.

In the following, we will start from the Richardson-Gaudin equations for spin-$1/2$ systems

$$1 + \frac{g}{2} \sum_{i=1}^{n} Z_{i\alpha} = g \sum_{\beta \neq \alpha}^{N} Z_{\beta\alpha}, \quad \forall \alpha = 1, \ldots, N,$$

(19)

and several properties that can be derived from the Gaudin algebra, as previously obtained by Ortiz *et al.* [2]. First, it has been shown that

$$(X_{ij})^2 - (Z_{ij})^2 = \Gamma, \quad \forall i \neq j,$$

(20)

where $\Gamma$ is a constant. The rational model corresponds to $\Gamma = 0$, while positive and negative $\Gamma$ result in the trigonometric and hyperbolic models, respectively. It also follows from the Gaudin algebra that

$$Z_{ij} Z_{jk} - Z_{ik}(Z_{ij} + Z_{jk}) = \Gamma, \quad \forall i \neq j \neq k \neq i,$$

(21)

which is a generalization of the Gaudin equations for the rational model, where $X_{ij} = Z_{ij}$ and $\Gamma = 0$. Using only these equations, it can straightforwardly be shown that

$$\Lambda_i^2 = N(n - N)\Gamma - \frac{2}{g} \Lambda_i + \sum_{j \neq i}^{n} Z_{ji}(\Lambda_j - \Lambda_i),$$

$$\forall \ i = 1, \ldots, n,$$

(22)

with $N$ being the total number of excitations and $n$ being the number of single-particle levels. The full derivation can be found in Appendix A. Unlike the case of the rational model, these equations depend explicitly on the number of excitations $N$. When solving these equations numerically, it was found that the total number of solutions exceeds the dimension of the Hilbert space for $N$ excitations distributed over $n$ levels. Therefore, these equations necessarily support unphysical solutions, not corresponding to any eigenstate, implying that this new set of equations is not equivalent to the original set of RG equations [Eq. (12)]. In order to obtain a set of equations equivalent to the original equations, additional constraints for the $\Lambda_i$ are needed.

It is clear from the $\Gamma = 0$ case (XXX) that the new set of equations cannot distinguish between the different excitation sectors $N$. This can be imposed by noting that the sum of all constants of motion is given by the operator counting the number of excitations

$$\sum_{i=1}^{n} R_i = \sum_{i=1}^{n} S_i^0.$$

(23)

Writing out the eigenvalues of the constants of motion in the new variables results in

$$-\frac{g}{2} \sum_{i=1}^{n} \Lambda_i = N.$$

(24)

In the following section it will be shown that the full set of equations (22) and (24) has as many solutions as the dimension

of the Hilbert space, so introducing this additional equation leads to a system of equations fully equivalent to the original set of RG equations. These eigenvalue-based equations do not show singular behavior and can easily be solved numerically, as will be shown in the following.

### B. The weak-coupling limit

In order to obtain some insight into the behavior of the solutions of these equations, an approximate solution can be found in the weak-coupling limit (small $|g|$). This can be done by proposing a series expansion in $g$ for the solutions of these equations and solving the equations at each order. For small $g$, a series expansion of $\Lambda_i$ in $g$ can be proposed up to $O(g)$, keeping only the two dominant terms

$$\Lambda_i = \frac{\lambda_i^{(-1)}}{g} + \lambda_i^{(0)} + O(g). \tag{25}$$

By plugging this expansion in the equations, we obtain

$$\frac{1}{g^2}\lambda_i^{(-1)}\big(\lambda_i^{(-1)} + 2\big)$$

$$+ \frac{1}{g}\left\{ 2\lambda_i^{(0)}\big(1 + \lambda_i^{(-1)}\big) - \sum_{j\neq i}^{n} Z_{ji}\big[\lambda_j^{(-1)} - \lambda_i^{(-1)}\big] \right\}$$

$$+ O(g^0) = 0. \tag{26}$$

This results in

$$\lambda_i^{(-1)} = 0 \ \text{ or } -2 \tag{27}$$

and

$$\lambda_i^{(0)} = \frac{1}{2\lambda_i^{(-1)} + 2} \sum_{j\neq i}^{n} Z_{ji}\big(\lambda_j^{(-1)} - \lambda_i^{(-1)}\big). \tag{28}$$

In order to satisfy Eq. (24), the number of dominant terms different from zero has to equal the number of excitations $N$, resulting in $\binom{n}{N}$ solutions. The total number of solutions then equals the dimension of the Hilbert space for $N$ excitations distributed over $n$ doubly degenerate levels. Any solution in the weak-coupling limit can be adiabatically connected to a solution for arbitrary coupling, indicating that all possible solutions are always found and no unphysical solutions are present.

This series expansion can also be connected to the series expansion obtained for the rapidities [2]. In the limit $g \to 0$ we obtain a noninteracting model, where the rapidities $x_\alpha$ converge to the parameters $\epsilon_i$, depending on the corresponding distribution of excitations over energy levels. A rapidity converging to $\epsilon_i$ then corresponds to an excited level $i$ in the noninteracting limit. For finite but small $g$, dominant corrections of $O(g)$ are present, which are proportional to the roots of orthogonal polynomials via a Heine-Stieltjes connection [37,38]. For $x_\alpha$ converging to $\epsilon_i$, this results in $Z_{i\alpha} = Z(\epsilon_i, x_\alpha)$ diverging as $1/g$ in the weak-coupling limit, where the proportionality factor can be found to be $-2$ from the Heine-Stieltjes connection for $d_i = 1/2$. For all other levels $j \neq i$, $Z_{j\alpha}$ remains finite, and the dominant order is $O(g^0)$.

A dominant order of $O(1/g)$ in $\Lambda_i$ then corresponds to an excited state $i$ in the noninteracting limit, while a

dominant order of $O(g^0)$ results in a nonexcited state $i$. This behavior can be generalized through a connection of $\Lambda_i$ to occupation numbers, as will be shown in the section about form factors. Note that the divergence in $g \to 0$ can pose numerical problems, so this suggests the use of $g\Lambda_i$ as variables instead of $\Lambda_i$.

### C. Solving the equations

Due to the similarity of Eq. (22) to the set of equations found for the rational model [20], it is straightforward to extend the solution method for the rational model to our equations. General sets of nonlinear equations have to be solved by an iterative approach starting from an initial guess, such as the Newton-Raphson method. This method converges quadratically to the solution if the initial guess lies in the basin of attraction. An efficient numerical approach can be implemented based on this method once we have access to a sufficiently good initial guess for the solution.

As in [20], we start from an approximation of the solution in the weak-coupling limit (small $|g|$). A solution at any value of the coupling constant can then be found by adiabatically varying $g$ starting from the weak-coupling limit and using the solution at the previous step as the starting point for an iterative solution at the current step. This initial guess can be improved by using a Taylor expansion of the solutions at the previous step,

$$\Lambda_i(g + \delta g) \approx \Lambda_i(g) + \delta g \frac{\partial \Lambda_i}{\partial g} \tag{29}$$

since the derivatives of the $\Lambda_i$ can be found by solving a linear system once the set of $\Lambda_i$ is known, similar to the procedure followed in [20]. The equations for the derivatives can be found by deriving the set of equations to $g$, leading to

$$\Lambda_i \frac{\partial \Lambda_i}{\partial g} = \frac{\Lambda_i}{g^2} - \frac{1}{g}\frac{\partial \Lambda_i}{\partial g} + \frac{1}{2}\sum_{j\neq i}^{n} Z_{ji}\left(\frac{\partial \Lambda_j}{\partial g} - \frac{\partial \Lambda_i}{\partial g}\right). \tag{30}$$

By taking the higher derivatives of the original set of equations, linear equations can be found for higher-order derivatives of $\Lambda_i$, which allows a Taylor expansion up to arbitrary order. For an efficient numerical implementation, combining the Newton-Raphson method with a Taylor approximation up to first order already offers a remarkable increase in speed.

### D. Inverting the transformation

In the previous sections, it was shown how, instead of solving the RG equations (12) for the rapidities, it is possible to solve an equivalent set of equations for the eigenvalue-based variables. If the rapidities are known, the set of $\Lambda_i$ can immediately be determined from the definition (17). However, the inverse is more cumbersome.

In order to be as general as possible, we propose a method to obtain the rapidities that does not rely on the explicit expression for $X$ and $Z$ except in the very last step. We introduce an auxiliary index $r$, corresponding to an auxiliary single-particle level or an additional uncoupled rapidity $\epsilon_r$, such that the Gaudin algebra can be extended. Following Ortiz *et al.* [2] and

Eq. (21), each $Z_{i\alpha}$ can be written as

$$Z_{i\alpha} = \frac{\Gamma + Z_{ri}Z_{r\alpha}}{Z_{ri} - Z_{r\alpha}}. \tag{31}$$

Once an explicit expression for $Z_{ij}$ is known, $Z_{ri} = Z(\epsilon_r, \epsilon_i)$ can easily be calculated after choosing $\epsilon_r$. Instead of solving for the rapidities $\{x_\alpha\}$, it is now possible to find a transformation that allows us to determine the set of Gaudin algebra elements $\{Z_{r\alpha}\}$. If these are known, the rapidities can always be found by solving each separate equation $Z_{r\alpha} = Z(\epsilon_r, x_\alpha)$ for $x_\alpha$.

This expression can be used to show that

$$\Lambda_i = \sum_{\alpha=1}^{N} Z_{i\alpha} = \sum_{\alpha=1}^{N} \frac{\Gamma + Z_{ri}Z_{r\alpha}}{Z_{ri} - Z_{r\alpha}}$$

$$= -NZ_{ri} + \left(\Gamma + Z_{ri}^2\right) \sum_{\alpha=1}^{N} \frac{1}{Z_{ri} - Z_{r\alpha}}. \tag{32}$$

We can now define a polynomial with the full set ($\alpha = 1, \ldots, N$) of $Z_{r\alpha}$ as roots

$$P(z) = \prod_{\alpha=1}^{N} (z - Z_{r\alpha}) = \sum_{m=0}^{N} P_{N-m} z^m, \tag{33}$$

with $P_0 = 1$. Once the coefficients $P_{N-m}$ are known, the roots of this polynomial can be determined to find the variables. This method arises naturally for different problems in the theory of integrable systems, such as the Heine-Stieltjes connection [37,38], the numerical methods by Guan *et al.* [18] and Rombouts *et al.* [15], and the weak-coupling limit in RG models [2].

The definition of $P(z)$ can now be used to consider

$$\frac{P'(z)}{P(z)} = \frac{\sum_{m=0}^{N} m P_{N-m} z^{m-1}}{\sum_{m=0}^{N} P_{N-m} z^m} = \sum_{\alpha=1}^{N} \frac{1}{z - Z_{r\alpha}}. \tag{34}$$

Evaluating this equality in $z = Z_{ri}$, we find that

$$\Lambda_i = -NZ_{ri} + \left(\Gamma + Z_{ri}^2\right) \frac{P'(Z_{ri})}{P(Z_{ri})}, \tag{35}$$

resulting in

$$\frac{P'(Z_{ri})}{P(Z_{ri})} = \sum_{\alpha=1}^{N} \frac{1}{Z_{ri} - Z_{r\alpha}} = \frac{\Lambda_i + NZ_{ri}}{\Gamma + Z_{ri}^2}. \tag{36}$$

The coefficients of the polynomial $P(z)$ can then be found by solving a linear problem

$$\sum_{m=0}^{N-1} P_{N-m} \left[m\Gamma Z_{ri}^{m-1} - \Lambda_i Z_{ri}^m + (m - N)Z_{ri}^{m+1}\right]$$

$$= \Lambda_i Z_{ri}^N - N\Gamma Z_{ri}^{N-1}. \tag{37}$$

We obtain $n$ equations for $N < n$ variables, but if the set of $\Lambda_i$ variables can be written as Eq. (32), these are linearly dependent, and we are free to choose $N$ equations from the full set and solve these. Once these coefficients are known, the variables $Z_{r\alpha}$ can be determined using a root-finding algorithm such as Laguerre's method. Although easy to implement, this method has the disadvantage of being prone to numerical errors. Indeed, it is well known that the roots of a polynomial

are highly sensitive to changes in the coefficients, sometimes even changing the solutions at the qualitative level (a set of complex conjugate roots can be found numerically instead of two separate real roots). These considerations are not pressing for a limited number of excitations but become more and more important for an increasing number of excitations.

This problem was also encountered for the rational model [20], after which an improved method was proposed by decomposing $P(z)$ in the basis of Lagrange polynomials [21]. This led to an increased accuracy and allowed the method to tackle problems with a large number of excitations (up to a few hundred). It is expected that these results can be generalized towards our problem because of the similarity of all necessary equations. We refer the reader to [21] for a detailed analysis of the method.

The polynomial representation in Eq. (36) also suggests an alternative way of solving the equations. From the RG equations it is possible to obtain a Heine-Stieltjes equation for $P(z)$, which can be used to numerically determine the polynomial in the monomial basis [18,39]. In this method it would not be necessary to start from the weak-coupling limit when numerically solving the equations, and only two coupled matrix equations need to be solved. This method does not allow for the targeting of specific states, such as the ground state, but is useful when all eigenstates need to be efficiently determined for an arbitrary coupling constant, such as for the evaluation of partition functions.

As an illustration, the eigenvalue-based variables and the related rapidities have been plotted for the different models in Fig. 1 as a function of the coupling constant. For each model, the levels $\epsilon_i$ are given by a picket-fence model [40] with equal level spacing: $\epsilon_i = i, \forall i$. For each system singular points occur, where multiple real rapidities combine and continue as a pair of complex conjugate variables. As is clear from Fig. 1, the RG equations [Eq. (12)] become highly singular, whereas the eigenvalue-based variables vary very smoothly.

### E. Degenerate models

The discussion has been limited to only $d_i = 1/2$ models thus far. This is sufficient for the majority of interacting (quasi)spin systems; however, in situations with higher symmetries arbitrary degeneracies $d_i > 1/2$ may occur. For the rational model, the set of equations for the eigenvalue-based variables were derived starting from a differential equation, where each equation for $\Lambda_i$ is obtained by evaluating this equation at a different level $\epsilon_i$. For degenerate levels, additional equations could be obtained by taking derivatives of the differential equation and evaluating these at the different values $\epsilon_i$ [20,21]. This necessitated the introduction of new variables

$$\Lambda_i^{(n)} \sim \sum_{\alpha=1}^{N} \frac{1}{(\epsilon_i - x_\alpha)^n}, \tag{38}$$

which are highly reminiscent of the set of variables introduced by Rombouts *et al.* [15]. The total number of variables is determined by the number of single-particle levels weighted by their degeneracies. For each single-particle level $i$ with degeneracy $2d_i + 1$, taking the first $2d_i - 1$ derivatives of the differential equation and evaluating each equation at $\epsilon_i$ then
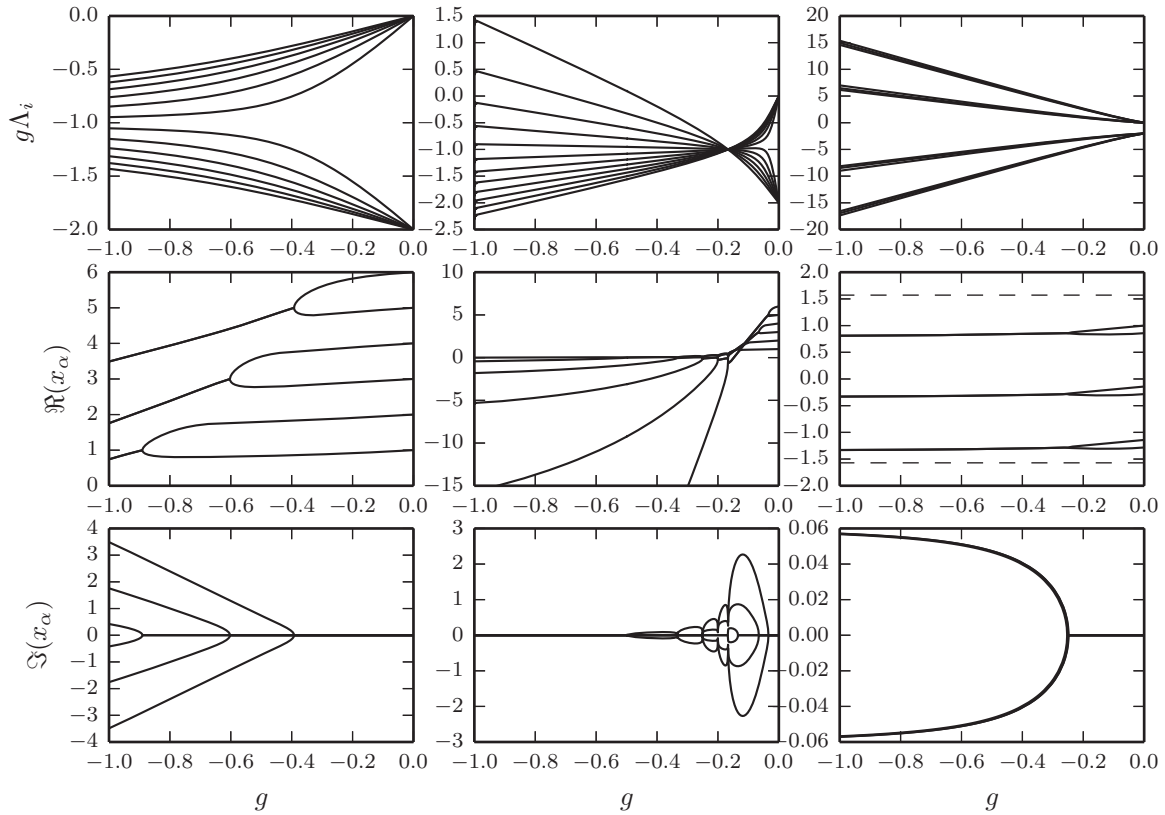
FIG. 1. Evolution of the variables for the ground state of a picket-fence model with six excitations in 12 doubly degenerate levels [1,2,3,...,12]. The rational, the hyperbolic, and the trigonometric models are organized in columns, while the rows depict the set of eigenvalue-based variables $g\Lambda_i$ and the real and imaginary parts of the rapidities as a function of the coupling constant $g = -1,\ldots,0$. The parametrizations used are given by Eqs. (5), (9), and (6), respectively. Note the Moore-Read point at $g = -2/n = -0.167$ for the hyperbolic model [5], where all rapidities condense to zero and the eigenvalue-based variables all become equal to $-1$. For the trigonometric model the rapidities are determined up to only a shift of $\pi$ due to the appearance of the tangent function, where $x_\alpha$ and $x_\alpha + k\pi, k \in \mathbb{N}$ describe the same state. This has been used to shift the real part of the rapidities to the interval $[-\pi/2, \pi/2]$. Note that the values of $g$ where the singularities arise and the complex pairs come in for the trigonometric model are not exactly equal, although they coincide visually in the third column.

results in a closed set of equations for the set of variables $\Lambda_i, \Lambda_i^{(2)}, \ldots, \Lambda_i^{(2d_i)}, \forall i$.

The outlined procedure is slightly more involved for the class of XXZ RG models; however, it is analogous to the rational case and remains, therefore, tractable. The main idea is identical: starting from a continuous representation of the equations, it is possible to obtain any number of equations for the total set of variables. The continuous representation of the variables is given by

$$\Lambda(z) = \sum_{\alpha=1}^{N} Z(z, x_\alpha), \tag{39}$$

where $\Lambda(\epsilon_i) \equiv \Lambda_i$. A derivation similar to that for the $d_i = 1/2$ model in Appendix A results in a continuous equation,

$$[\Lambda(z)]^2 = \Gamma N \left( 1 - N + 2 \sum_{j \neq i} d_j \right) - \frac{2}{g} \Lambda(z)$$
$$+ 2 \sum_{j \neq i} d_j Z(z, \epsilon_j)[\Lambda(z) - \Lambda(\epsilon_j)]$$
$$+ \sum_\alpha Z(z, x_\alpha)[Z(z, x_\alpha) - 2d_i Z(\epsilon_i, x_\alpha)], \tag{40}$$

which holds if $z \neq \epsilon_j$ for all $j \neq i$. The evaluation of these equations at $z = \epsilon_i$ results in the known set of equations for $\Lambda_i$ if $d_i = 1/2$, but for larger degeneracies the set of equations also depends on

$$\Lambda_i^{(2)} \equiv \sum_{\alpha=1}^{N} Z_{i\alpha}^2. \tag{41}$$

By taking the derivative of Eq. (40), additional equations can be obtained linking these higher-order variables to the original variables. The derivation of this method is given in Appendix B. It is remarkable that the derivative of $\Lambda(z)$, a summation over $Z(z, x_\alpha)$, can still be related to $\Lambda^{(2)}(z)$, a summation over $Z(z, x_\alpha)^2$.

As an example, the evaluation of the first derivative results in

$$\Lambda_i \left( N\Gamma + \Lambda_i^{(2)} \right)$$
$$= -\frac{1}{g} \left( N\Gamma + \Lambda_i^{(2)} \right) + \sum_{j \neq i} d_j \left( \Gamma + Z_{ij}^2 \right) (\Lambda_i - \Lambda_j)$$
$$+ \sum_{j \neq i} d_j Z_{ij} \left( \Lambda_i^{(2)} + \Gamma \right) + \Gamma (1 - d_i) \Lambda_i + (1 - d_i) \Lambda_i^{(3)}. \tag{42}$$
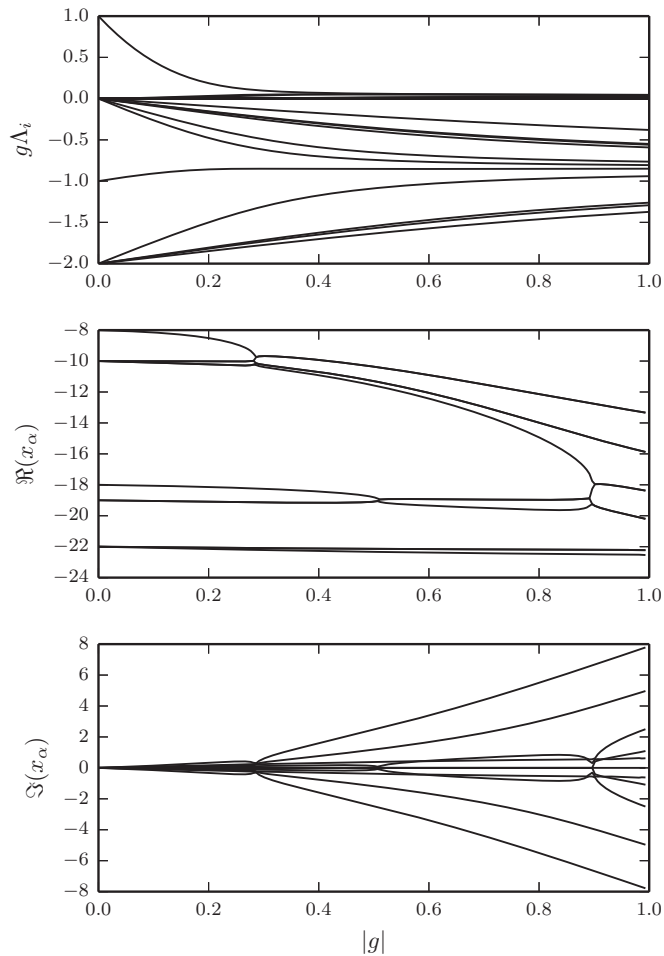
FIG. 2. Evolution of the variables for the ground state containing 11 excitations in ten levels for a varying (negative) coupling constant $g$, describing the pairing of neutrons in $^{56}$Fe by means of a level-independent reduced BCS model. The parameters of the system are taken from [15], where the set of $d_i$ is given by [3/2,1,1/2,2,1,3/2,1/2,1/2,1,5/2], with a maximal degeneracy of $2d_i + 1 = 6$. The first plot shows the evolution of the eigenvalue-based variables, while the second and third plots correspond to the real and imaginary parts of the rapidities, respectively.

For $d_i = 1$, we have obtained a closed set of equations in $\{\Lambda_i, \Lambda_i^{(2)}\}$ for each level. For arbitrary $d_i$, the first $2d_i - 1$ derivatives results in a closed set of equations. An additional equation for the total number of excitations can easily be determined as

$$N = -\sum_{i=1}^{n} g d_i \Lambda_i. \tag{43}$$

As an illustration, the results for a level-independent reduced BCS model describing neutron superfluidity [15] in $^{56}$Fe are given in Fig. 2. The energy levels are determined as the levels of a spherical symmetrical Woods-Saxon potential, and $d_i$, which are dependent on the angular momentum quantum numbers, vary from 1/2 to 5/2.

## IV. FORM FACTORS

### A. Overlap with Slater determinants

When performing calculations with Bethe ansatz states, it is customary to expand them in a basis set of uncorrelated wave functions. These expansion coefficients are given by permanents of matrices in the general case, following from the theory of multivariate generating functions [41]. However, these expressions are not practical for computational purposes since the evaluation of a permanent scales exponentially with the matrix size. Compared to the determinant, which can be evaluated in a polynomial time, permanents require a factorial scaling computational time and therefore severely limit the size of the systems that can be considered.

Luckily, multiple formulas exist for the rational model linking the permanents to determinants, allowing numerically efficient expansions of the Bethe ansatz state. Here we present several such results for $d_i = 1/2$ XXZ systems as a generalization of the results previously obtained for the XXX model [23] and the inhomogeneous Dicke model [24]. Starting from these expressions, determinant expressions for systems with arbitrary $d_i$ can also be obtained as a limiting case where several levels $\epsilon_i$ coincide, but this will not be discussed here.

Multiple determinant formulas exist for the overlap of a Bethe ansatz state in the rational model with a Slater determinant,

$$|\{i_a\}\rangle = |\{i_1 \cdots i_N\}\rangle = \prod_{a=1}^{N} S_{i_a}^{\dagger} |\downarrow \cdots \downarrow\rangle, \tag{44}$$

which is an uncorrelated wave function and an eigenstate in the uncoupled limit ($g = 0$). The Bethe ansatz state (11) can be expanded in this basis as

$$|\psi_N\rangle = \prod_{\alpha=1}^{N} \left( \sum_{i=1}^{n} X_{i\alpha} S_i^{\dagger} \right) |\theta\rangle$$

$$= \sum_{\{i_a\}} |\{i_a\}\rangle \langle \{i_a\} | \psi_N \rangle \tag{45}$$

$$= \sum_{\{i_a\}} |\{i_a\}\rangle \, \mathrm{per}([X_{i_a\alpha}]), \tag{46}$$

where $\{i_a\}$ runs over the partitioning of $n$ levels over $N$ different excitations modes and

$$\langle \{i_a\} | \psi_N \rangle = \mathrm{per}([X_{i_a\alpha}]), \tag{47}$$

which is the permanent of an ($N \times N$) matrix with matrix elements given by $X_{i_a\alpha}$. This is a formula independent of the integrability or the explicit expression for $X$ and is simply a result of the structure of the wave function. However, by imposing the Gaudin algebra on the matrix elements, it is always possible to rewrite this permanent as a determinant.

$$\mathrm{per}\left([X_{i_a\alpha}]\right) = \frac{\left(\prod_{a>b} X_{i_a i_b}\right)\left(\prod_{\beta>\alpha} X_{\alpha\beta}\right)}{\prod_{a,\alpha} X_{i_a\alpha}} \det\left[X_{i_a\alpha} * X_{i_a\alpha}\right], \tag{48}$$

where the Hadamard product of two matrices is introduced, defined as $(A * B)_{ij} = A_{ij} B_{ij}$. This is a generalization of the Borchardt or Izergin determinant representation for the rational

model [42–44] and has a structure similar to some results presented previously for the trigonometric model [45].

In Ref. [23], a case was made for a RG theory that would require only the eigenvalue-based variables instead of the rapidities. This would be desirable from a numerical point of view as the eigenvalue-based variables are free of singularities, as opposed to the singularity-prone rapidities. In addition, this would enable us to skip the inversion step, which is the main bottleneck of this method.

For the rational model, it was shown by Faribault and Schuricht [23] that

$$\text{per}\left(\left[\frac{1}{\epsilon_{i_a} - x_\alpha}\right]\right) = \det J, \quad (49)$$

where the left-hand side is the permanent of an $N \times N$ Cauchy matrix and the right-hand side is the determinant of a matrix $J$ defined as

$$J_{ab} = \begin{cases} \sum_{\alpha=1}^N \frac{1}{\epsilon_{i_a} - x_\alpha} - \sum_{c \neq a}^N \frac{1}{\epsilon_{i_a} - \epsilon_{i_c}} & \text{if } a = b, \\ \frac{1}{\epsilon_{i_a} - \epsilon_{i_b}} & \text{if } a \neq b, \end{cases} \quad (50)$$

where the only dependence on the rapidities is through the diagonal elements, which can be expressed in terms of the eigenvalue-based variables. The ingenious proof of this expression is based on a recursion relation which was found to hold for both sides of Eq. (50). We will show that these results can be easily generalized to XXZ models.

The Gaudin equations can be used to rewrite the permanent in the XXZ model in a structure similar to that of a Cauchy matrix. We once again introduce an auxiliary level $r$ and use the Gaudin equations to write

$$X_{i\alpha} = \frac{X_{ri} X_{r\alpha}}{Z_{ri} - Z_{r\alpha}}, \quad Z_{i\alpha} = \frac{\Gamma + Z_{ri} Z_{r\alpha}}{Z_{ri} - Z_{r\alpha}}. \quad (51)$$

Rewriting all matrix elements in the permanent $\phi$ allows us to take

$$\begin{aligned} \text{per}\left(\left[X_{i_a\alpha}\right]\right) &= \text{per}\left(\left[\frac{X_{ri_a} X_{r\alpha}}{Z_{ri_a} - Z_{r\alpha}}\right]\right) \\ &= \left(\prod_a X_{ri_a}\right)\left(\prod_\alpha X_{r\alpha}\right) \text{per}\left(\left[\frac{1}{Z_{ri_a} - Z_{r\alpha}}\right]\right), \end{aligned} \quad (52)$$

resulting in a Cauchy matrix, which we can rewrite as a determinant

$$\text{per}\left(\left[X_{i_a\alpha}\right]\right) = \left(\prod_a X_{ri_a}\right)\left(\prod_\alpha X_{r\alpha}\right) \det J, \quad (53)$$

with $J$ redefined as

$$J_{ab} = \begin{cases} \sum_{\alpha=1}^N \frac{1}{Z_{ri_a} - Z_{r\alpha}} - \sum_{c \neq a}^N \frac{1}{Z_{ri_a} - Z_{ri_c}} & \text{if } a = b, \\ \frac{1}{Z_{ri_a} - Z_{ri_b}} & \text{if } a \neq b. \end{cases} \quad (54)$$

By multiplying each row and column $c$ with $X_{ri_c}$ and compensating for these factors in the prefactor, this can be written as

$$\text{per}\left(\left[X_{i_a\alpha}\right]\right) = \frac{\prod_\alpha X_{r\alpha}}{\prod_a X_{ri_a}} \det J, \quad (55)$$

with

$$J_{ab} = \begin{cases} \sum_{\alpha=1}^N \frac{X_{ri_a}^2}{Z_{ri_a} - Z_{r\alpha}} - \sum_{c \neq a}^N \frac{X_{ri_a}^2}{Z_{ri_a} - Z_{ri_c}} & \text{if } a = b, \\ \frac{X_{ri_a} X_{ri_b}}{Z_{ri_a} - Z_{ri_b}} & \text{if } a \neq b, \end{cases} \quad (56)$$

where we recognize $X_{i_a i_b}$ in the off-diagonal elements and the diagonal elements can be rewritten by using the Gaudin algebra until

$$J_{ab} = \begin{cases} \sum_{\alpha=1}^N Z_{i_a\alpha} - \sum_{c \neq a}^N Z_{i_a i_c} + Z_{ri_a} & \text{if } a = b, \\ X_{i_a i_b} & \text{if } a \neq b. \end{cases} \quad (57)$$

The only explicit dependency on the rapidities is now found in the prefactor and in the diagonal elements, which depend on only the eigenvalue-based variables. Fortunately, the prefactor can be absorbed in the definition of the state without loss of generality, which will still allow the calculation of all necessary normalizations and prefactors. Note the appearance of $Z_{ri}$ in the diagonal elements, which can be easily evaluated and links the matrix to the prefactor for the Bethe ansatz state through our choice of the additional level $r$.

To recapitulate, the overlap of a state

$$|\{x_\alpha\}\rangle = \prod_{\alpha=1}^N \left(\sum_{i=1}^n \frac{X_{i\alpha}}{X_{r\alpha}} S_i^\dagger\right) |\downarrow \cdots \downarrow\rangle \quad (58)$$

with a Slater determinant

$$|\{i_a\}\rangle = |\{i_1 \cdots i_N\}\rangle = \prod_{j=1}^N S_{i_j}^\dagger |\downarrow \cdots \downarrow\rangle \quad (59)$$

is given by

$$\langle\{i_1 \cdots i_N\}|\{x_\alpha\}\rangle = \frac{\det J}{\prod_a X_{ri_a}}, \quad (60)$$

with

$$J_{ab} = \begin{cases} \Lambda_{i_a} - \sum_{c \neq a}^N Z_{i_a i_c} + Z_{ri_a} & \text{if } a = b, \\ X_{i_a i_b} & \text{if } a \neq b. \end{cases} \quad (61)$$

### B. Dual representations

So far, all Bethe ansatz states have been created by acting with creation operators on the vacuum state, destroyed by all annihilation operators. Due to the particle-hole symmetry of the RG models, every Bethe ansatz state can also be constructed by acting on the fully filled state, annihilated by all creation operators, with a set of generalized annihilation operators, the so-called dual representation.

A renormalized Bethe ansatz state is given in its normal and dual representation as

$$|\{x_\alpha\}\rangle = \prod_{\alpha=1}^N \left(\sum_{i=1}^n \frac{X_{i\alpha}}{X_{r\alpha}} S_i^\dagger\right) |\downarrow \cdots \downarrow\rangle, \quad (62)$$

$$|\{x_{\alpha'}\}\rangle = \prod_{\alpha'=1}^{n-N} \left(\sum_{i=1}^n \frac{X_{i\alpha'}}{X_{r\alpha'}} S_i\right) |\uparrow \cdots \uparrow\rangle. \quad (63)$$

The RG equations for the dual state are given by

$$-1 + \frac{g}{2} \sum_i Z_{j\alpha'} = g \sum_{\beta' \neq \alpha'}^{n-N} Z_{\beta'\alpha'}, \quad \forall \alpha' = 1, \ldots, n - N, \tag{64}$$

or, written in the eigenvalue-based variables,

$$[\Lambda_i']^2 = N(n - N)\Gamma + \frac{2}{g} \Lambda_i' + \sum_{j \neq i} Z_{ji}(\Lambda_j' - \Lambda_i'),$$

$$\forall i = 1, \ldots, n. \tag{65}$$

Both representations describe, up to the normalization, the same state. So by comparing the eigenvalues of the constants of motion

$$r_i = \frac{1}{2}\left(-1 - g\Lambda_i + \frac{g}{2} \sum_{k \neq i}^n Z_{ik}\right)$$

$$= \frac{1}{2}\left(1 - g\Lambda_i' + \frac{g}{2} \sum_{k \neq i}^n Z_{ik}\right), \tag{66}$$

it can be seen that the dual representation of an eigenstate is determined by a set of dual eigenvalue-based variables given by

$$\Lambda_i' = \Lambda_i + \frac{2}{g}, \quad \forall i = 1, \ldots, n, \tag{67}$$

which can be verified by simply substituting these variables in the dual eigenvalue-based equations. The existence of the dual state allows one to write the overlap of a normal state and a dual state as the overlap of a Bethe ansatz state with the fully filled state. This will be exploited in the following section to obtain the normalization of the Bethe ansatz states and several form factors. Note the link with Slavnov's theorem [46] for the rational model, which presents a determinant expression for the overlap of an (on-shell) Bethe ansatz state that is an eigenstate and an (off-shell) Bethe ansatz state that is not an eigenstate.

### C. Normalization and form factors

Due to the dual representation, we have two different representations of each Bethe ansatz state. These two states have different normalizations and can be written as

$$|\{x_\alpha\}\rangle = \prod_{\alpha=1}^N \left(\sum_{i=1}^n \frac{X_{i\alpha}}{X_{r\alpha}} S_i^\dagger\right)|\downarrow \cdots \downarrow\rangle = N_\alpha |\{x_\alpha\}\rangle_n,$$

$$|\{x_{\alpha'}\}\rangle = \prod_{\alpha'=1}^{n-N} \left(\sum_{i=1}^n \frac{X_{i\alpha'}}{X_{r\alpha'}} S_i\right)|\uparrow \cdots \uparrow\rangle = N_{\alpha'} |\{x_\alpha\}\rangle_n,$$

where $|\{x_\alpha\}\rangle_n$ is the normalized eigenstate. This can be used to calculate

$$\langle\{x_{\alpha'}\}|\{x_\alpha\}\rangle = N_\alpha N_{\alpha'}$$

$$= \langle\uparrow \cdots \uparrow| \prod_{\alpha'=1}^{n-N} \left(\sum_{i=1}^n \frac{X_{i\alpha'}}{X_{r\alpha'}} S_i^\dagger\right)$$

$$\times \prod_{\alpha=1}^N \left(\sum_{i=1}^n \frac{X_{i\alpha}}{X_{r\alpha}} S_i^\dagger\right)|\downarrow \cdots \downarrow\rangle$$

$$= \frac{\det J}{\prod_{i=1}^n X_{ri}}, \tag{68}$$

with

$$J_{ij} = \begin{cases} 2\Lambda_i + \frac{2}{g} - \sum_{k \neq i}^n Z_{ik} + Z_{ri} & \text{if } i = j, \\ X_{ij} & \text{if } i \neq j, \end{cases} \tag{69}$$

since the overlap between the two states is just the overlap of a Bethe ansatz state determined by $\{x\} = \{x_\alpha\} \cup \{x_{\alpha'}\}$ or by eigenvalue-based variables $\Lambda_i = \Lambda_i + \Lambda_i' = 2\Lambda_i + 2/g$ with $|\uparrow \cdots \uparrow\rangle$. The ratio $N_\alpha/N_{\alpha'}$ can easily be found by taking the overlap of both representations with a single reference state and taking this ratio. Knowing both the ratio and the product of the normalizations, the normalization of both representations is uniquely defined.

In a manner similar to that for the XXX model [23,24], these results can be used to calculate form factors. Since this is a generalization of the results presented in these papers, we will summarize the results for the XXZ model here.

The form factor for the raising operator between a state $\{x_\alpha\}$ with $N$ excitations and a dual state $\{x_{\mu'}\}$ with $N + 1$ excitations is given by

$$\langle\{x_{\mu'}\}|S_k^\dagger|\{x_\alpha\}\rangle = \frac{\det J^k}{\prod_{i \neq k} X_{ri}}, \tag{70}$$

with

$$J_{ab}^k = \begin{cases} \Lambda_i^\alpha + \Lambda_i^\mu + \frac{2}{g} - \sum_{l \neq i, l \neq k}^n Z_{il} + Z_{ri} & \text{if } a = b, \\ X_{ab} & \text{if } a \neq b, \end{cases} \tag{71}$$

and $a, b \neq k$. The form factor for a lowering operator is then given by the Hermitian conjugate of this expression.

The form factor for local operators $S_k^0$ can similarly be found for both diagonal and off-diagonal expectation values. Here we will make the dependency on the coupling constant explicit and define $\{x_\alpha(g)\}$ as an eigenstate at coupling constant $g$.

From the Hellmann-Feynman theorem we obtain for the diagonal expectation values

$$\langle\{x_{\alpha'}(g)\}|S_k^0|\{x_\alpha(g)\}\rangle$$

$$= \frac{1}{2}\left(-1 + g^2 \frac{\partial \Lambda_k^\alpha}{\partial g}\right)\langle\{x_{\alpha'}(g)\}|\{x_\alpha(g)\}\rangle, \tag{72}$$

while the off-diagonal expectation values can be found as

$$
\langle \{x_{\mu'}(g)\}|S_k^0|\{x_\alpha(g)\}\rangle
$$

$$
= \frac{1}{2}\big(g\Lambda_k^\mu - g\Lambda_k^\alpha + 2\big)\frac{\langle \{x_{\mu'}(g)\}|\{x_\alpha(g+dg)\}\rangle}{dg}
$$

$$
= \frac{1}{2}\big(g\Lambda_k^\mu - g\Lambda_k^\alpha + 2\big)\sum_{k=1}^n \frac{\partial \Lambda_k^\alpha}{\partial g}\det \tilde{J}^k, \qquad (73)
$$

with

$$
\tilde{J}_{ab}^k = \begin{cases} \Lambda_i^\alpha + \Lambda_i^\mu + \frac{2}{g} - \sum_{l\neq i}^n Z_{il} + Z_{ri} & \text{if } a = b, \\ X_{ab} & \text{if } a \neq b, \end{cases} \qquad (74)
$$

and $a,b \neq k$. Note that the partial derivatives of the $\Lambda_i$ appearing in Eq. (73) can be easily obtained from the set of $\Lambda_i$ through solving the linear set of equations (30). This allows the form factors to be calculated from the eigenvalue-based variables, so the rapidities do not need to be determined explicitly. However, if these have been determined, other form factors can also be calculated by commuting the operator through the creation operators in the Bethe ansatz state, as has been shown in [36] for the XXX model and [47] for the XXZ model. This allows the computation of form factors such as, e.g., $S_k^\dagger S_l$ as a sum over determinants.

It is interesting to note that the expectation value for $S_i^0$ offers a physical interpretation for $\Lambda_i$ since the derivative of $\Lambda_i$ to $g$ fully determines the filling of the level $i$ as described by $\langle S_i^0 \rangle$. Knowledge of the evolution of $\Lambda_i$ with a changing coupling constant is then equivalent to knowing how the $N$ excitations are distributed over the $n$ levels. This has been illustrated in Fig. 3 for the picket-fence model.

## V. CONCLUSIONS AND OUTLOOK

In the present paper, we presented a numerical solution method for the RG equations for XXZ integrable models while also proposing determinant expressions for the normalization and form factors of the Bethe ansatz states that are dependent on only the new set of eigenvalue-based variables. The availability of an efficient solution method and explicit expressions for the overlaps opens up possibilities for the investigation of integrable and nonintegrable quantum systems.

An efficient numerical expression for overlaps between two Bethe ansatz states allows for a numerically nearly exact investigation of the dynamics resulting from a quantum quench in the reduced BCS pairing model [25], while for models close to integrability the eigenstates of the integrable model are being used to study decoherence [26]. These results were all obtained for the rational XXX model, and it would be interesting to investigate similar dynamics for XXZ models, for which a rich phase diagram is known [5].

Due to the favorable computational scaling of this method it is also possible to numerically approach the thermodynamic limit and investigate the limiting behavior of the system. Such an investigation has been carried out for the reduced BCS Hamiltonian by El Araby and Baeriswyl [48], who compared exact results for large system sizes with the BCS ansatz. Again, it would be worthwhile to compare the results rigorously with the results from BCS mean-field theory [4].
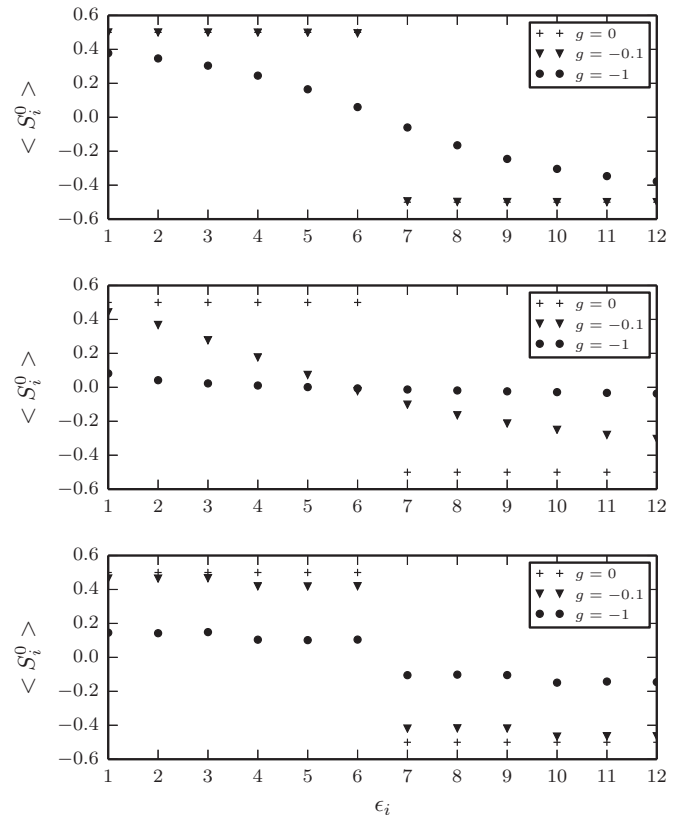


FIG. 3. The plots correspond to the rational model (5), the hyperbolic model (9), and the trigonometric model (6), respectively. All variables correspond to those of Fig. 1, and the expectation value of the spin for each level is presented for different values of the coupling constant $g = 0, -0.1, -1$. For $g = 0$ all models reduce to an uncoupled model with $\langle S_i^0 \rangle = \pm 1/2$, while for $g \neq 0$ the excitations become more spread out.

The overlaps between Bethe ansatz states and Slater determinants have also been used in quantum chemistry for strongly correlated systems. These cannot be accurately described by means of a single Slater determinant, and several wave functions have been proposed to capture the correlation present in the system. One example is the class of geminal-based wave functions, which have a structure highly reminiscent of the Bethe ansatz states found in the Gaudin models. Unfortunately, these wave functions are only computationally feasible if a numerical efficient expression is known for the overlaps with Slater determinants. One such class (APr2G) is based on the rational XXX model [27–29], and the presented results should allow for an extension of these geminal-based wave functions with a class based on the XXZ models. The results presented here should also allow for a variational approach based on the Bethe ansatz states in the XXZ model. By starting from Richardson's general solution for the Gaudin algebra [Eq. (8)] and varying over the free parameters in the wave function, it is possible to scan over all possible Gaudin models (XXX and XXZ), allowing more freedom and as such a better approximation to the energy compared to a variational approach based solely on the XXX model.

It was already noted by Babelon and Talalaev that similar equations could be found for the XXX Heisenberg spin

chain [22], so the question naturally arises if this method can be generalized to other Bethe ansatz equations obtained from the quantum inverse scattering method [49]. The extension of the proposed determinant expressions to degenerate models will be the subject of future research.

## APPENDIX A: DERIVING THE EQUATIONS

In this Appendix we derive the equations for the eigenvalue-based variables by making use of the Richardson-Gaudin equations and the Gaudin algebra. All summations with indices labeled by Greek characters run over the $N$ rapidities, while for indices labeled by Latin characters the summations run over the $n$ energy levels. Starting from the definition of the variable $\Lambda_i$, we can write

$$
\begin{aligned}
\Lambda_i^2 &= \sum_{\alpha,\beta} Z_{i\alpha} Z_{i\beta} = \sum_{\alpha,\beta\neq\alpha} Z_{i\alpha} Z_{i\beta} + \sum_\alpha Z_{i\alpha}^2 \\
&= -\sum_{\alpha,\beta\neq\alpha}[\Gamma + Z_{\alpha\beta}(Z_{\alpha i} + Z_{i\beta})] + \sum_\alpha Z_{i\alpha}^2, \quad \text{(A1)}
\end{aligned}
$$

where we have used Eq. (21) to rewrite $Z_{i\alpha} Z_{i\beta}$. The antisymmetry of $Z_{i\beta}$ and $Z_{\alpha\beta}$ can be used to obtain

$$
\begin{aligned}
\Lambda_i^2 &= -\sum_{\alpha,\beta\neq\alpha}(\Gamma + 2 Z_{\alpha\beta} Z_{\alpha i}) + \sum_\alpha Z_{i\alpha}^2 \\
&= -N(N-1)\Gamma + 2\sum_\alpha Z_{\alpha i}\left(\sum_{\beta\neq\alpha} Z_{\beta\alpha}\right) + \sum_\alpha Z_{i\alpha}^2.
\end{aligned}
$$

Making use of the RG equations [Eq. (12)], we obtain

$$
\begin{aligned}
\Lambda_i^2 &= -N(N-1)\Gamma + 2\sum_\alpha Z_{\alpha i}\left(\frac{1}{g} + \frac{1}{2}\sum_j Z_{j\alpha}\right) \\
&\quad + \sum_\alpha Z_{i\alpha}^2 \\
&= -N(N-1)\Gamma + \frac{2}{g}\sum_\alpha Z_{\alpha i} + \sum_\alpha \sum_{j\neq i} Z_{\alpha i} Z_{j\alpha}, \quad \text{(A2)}
\end{aligned}
$$

which can again be rewritten by making use of Eq. (21).

$$
\begin{aligned}
\Lambda_i^2 &= -N(N-1)\Gamma + \frac{2}{g}\sum_\alpha Z_{\alpha i} \\
&\quad + \sum_\alpha \sum_{j\neq i}[\Gamma + Z_{ji}(Z_{j\alpha} + Z_{\alpha i})] \\
&= N(n-N)\Gamma - \frac{2}{g}\Lambda_i + \sum_{j\neq i} Z_{ji}(\Lambda_j - \Lambda_i), \quad \text{(A3)}
\end{aligned}
$$

resulting in a set of equations closed in the variables $\Lambda_i, i = 1,\ldots,n$.

## APPENDIX B: OBTAINING EQUATIONS FOR DEGENERATE MODELS

Here we show how to obtain additional equations for degenerate models starting from a continuous equation for the eigenvalue-based variables. We define

$$
\Lambda_i^{(p)} \equiv \sum_{\alpha=1}^N Z_{i\alpha}^p, \quad p \in \mathbb{N} \tag{B1}
$$

and

$$
\Lambda^{(p)}(z) \equiv \sum_{\alpha=1}^N Z(z,x_\alpha)^p, \quad p \in \mathbb{N}. \tag{B2}
$$

By taking the derivative with respect to $z$ of Eq. (40), additional equations can be found. Unfortunately, an explicit expression for $Z(z,\epsilon_i)$ is needed if we want to relate the derivative of $\Lambda(z)$ to $\Lambda^{(2)}(z)$ and obtain a closed set of equations. In order to circumvent this problem, we introduce a fixed auxiliary level $\epsilon_r$, so that every $Z$ can now again be parametrized as in Eq. (21),

$$
Z(z,x_\alpha) = \frac{Z(\epsilon_r,z)Z(\epsilon_r,x_\alpha) + \Gamma}{Z(\epsilon_r,z) - Z(\epsilon_r,x_\alpha)}, \tag{B3}
$$

so instead of deriving this equation with respect to the continuous variable $z$, we can derive it with respect to $Z(\epsilon_r,z)$ and multiply with $-X_{rz}^2$, resulting in an equation independent of the explicit expression for $Z$. For notational ease, we take $Z_{rz} = Z(\epsilon_r,z)$. Then

$$
-X_{rz}^2 \frac{d}{dZ_{rz}} Z(z,x_\alpha) = X(z,x_\alpha)^2 = Z(z,x_\alpha)^2 + \Gamma, \tag{B4}
$$

which is independent of the auxiliary level. The action of this operator (derivation and multiplication) is given by

$$
\Lambda^{(p)}(z) \rightarrow p\Lambda^{(p+1)}(z) + p\Gamma\Lambda^{(p-1)}(z) \tag{B5}
$$

and

$$
Z(z,\epsilon_j)^p \rightarrow pZ(z,\epsilon_j)^{(p+1)} + p\Gamma Z(z,\epsilon_j)^{(p-1)}. \tag{B6}
$$

By acting with this operator on the equations for $\Lambda(z)$, we obtain

$$
\begin{aligned}
\Lambda(z)&[N\Gamma + \Lambda^{(2)}(z)] \\
&= -\frac{1}{g}[N\Gamma + \Lambda^{(2)}(z)] \\
&\quad + \sum_{j\neq i} d_j[\Gamma + Z(z,\epsilon_j)^2][\Lambda(z) - \Lambda(\epsilon_j)] \\
&\quad + \sum_{j\neq i} d_j Z(z,\epsilon_j)[\Lambda^{(2)}(z) + \Gamma] \\
&\quad + \sum_\alpha Z(z,x_\alpha)^2[Z(z,x_\alpha) - d_i Z(\epsilon_i,x_\alpha)] \\
&\quad + \Gamma[\Lambda(z) - d_i\Lambda(\epsilon_i)], \tag{B7}
\end{aligned}
$$

resulting in Eq. (42) when evaluated in $z = \epsilon_i$. Successive derivatives result in as many additional equations as there are variables, allowing a closed set of equations to be obtained.

[1] J. Dukelsky, S. Pittel, and G. Sierra, Rev. Mod. Phys. **76**, 643 (2004).

[2] G. Ortiz, R. Somma, J. Dukelsky, and S. Rombouts, Nucl. Phys. B **707**, 421 (2005).

[3] M. C. Cambiaggio, A. M. F. Rivas, and M. Saraceno, Nucl. Phys. A **624**, 157 (1997).

[4] M. Ibañez, J. Links, G. Sierra, and S.-Y. Zhao, Phys. Rev. B **79**, 180501 (2009).

[5] S. M. A. Rombouts, J. Dukelsky, and G. Ortiz, Phys. Rev. B **82**, 224510 (2010).

[6] M. Gaudin, J. Phys. **37**, 1087 (1976).

[7] J. Dukelsky, S. Lerma, H. L. M. Robledo, R. Rodriguez-Guzman, and S. M. A. Rombouts, Phys. Rev. C **84**, 061301 (2011).

[8] I. Marquette and J. Links, Nucl. Phys. B **866**, 378 (2013).

[9] J. Dukelsky, G. G. Dussel, C. Esebbag, and S. Pittel, Phys. Rev. Lett. **93**, 050403 (2004).

[10] C. Dunning, P. S. Isaac, J. Links, and S.-Y. Zhao, Nucl. Phys. B **848**, 372 (2011).

[11] R. W. Richardson and N. Sherman, Nucl. Phys. **52**, 221 (1964).

[12] J. Dukelsky, C. Esebbag, and P. Schuck, Phys. Rev. Lett. **87**, 066403 (2001).

[13] R. Richardson, Phys. Rev. **141**, 949 (1966).

[14] F. Domínguez, C. Esebbag, and J. Dukelsky, J. Phys. A **39**, 11349 (2006).

[15] S. Rombouts, D. Van Neck, and J. Dukelsky, Phys. Rev. C **69**, 061303 (2004).

[16] S. De Baerdemacker, Phys. Rev. C **86**, 044332 (2012).

[17] M. Van Raemdonck, S. De Baerdemacker, and D. Van Neck, Phys. Rev. B **89**, 155136 (2014).

[18] X. Guan, K. D. Launey, M. Xie, L. Bao, F. Pan, and J. P. Draayer, Phys. Rev. C **86**, 024313 (2012).

[19] W. V. Pogosov, J. Phys. Condens. Matter **24**, 075701 (2012).

[20] A. Faribault, O. El Araby, C. Sträter, and V. Gritsev, Phys. Rev. B **83**, 235124 (2011).

[21] O. El Araby, V. Gritsev, and A. Faribault, Phys. Rev. B **85**, 115130 (2012).

[22] O. Babelon and D. Talalaev, J. Stat. Mech. (2007) P06013.

[23] A. Faribault and D. Schuricht, J. Phys. A **45**, 485202 (2012).

[24] H. Tschirhart and A. Faribault, J. Phys. A **47**, 405204 (2014).

[25] A. Faribault, P. Calabrese, and J.-S. Caux, J. Stat. Mech. (2009) P03018.

[26] R. van den Berg, G. P. Brandino, O. El Araby, R. M. Konik, V. Gritsev, and J.-S. Caux, Phys. Rev. B **90**, 155117 (2014).

[27] P. A. Johnson, P. W. Ayers, P. A. Limacher, S. De Baerdemacker, D. Van Neck, and P. Bultinck, Comput. Theor. Chem. **1003**, 101 (2013).

[28] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, J. Chem. Theory Comput. **9**, 1394 (2013).

[29] P. Tecmer, K. Boguslawski, P. A. Johnson, P. A. Limacher, M. Chan, T. Verstraelen, and P. W. Ayers, J. Phys. Chem. A **118**, 9058 (2014).

[30] J.-S. Caux and J. Mossel, J. Stat. Mech. (2011) P02023.

[31] J. Links, H.-Q. Zhou, R. H. McKenzie, and M. D. Gould, J. Phys. A **36**, R63 (2003).

[32] I. Talmi, *Simple Models of Complex Nuclei* (Harwood Academic, Chur, Switzerland, 1993).

[33] R. W. Richardson, arXiv:cond-mat/0203512.

[34] With excitations we denote spin-up flips for magnetic (spin) systems, and the number of pairs for pairing (quasispin) systems.

[35] R. W. Richardson, Phys. Lett. **3**, 277 (1963).

[36] H.-Q. Zhou, J. Links, R. H. McKenzie, and M. D. Gould, Phys. Rev. B **65**, 060502 (2002).

[37] T. J. Stieltjes, Acta Math. **6**, 319 (1885).

[38] B. Sriram Shastry and A. Dhar, J. Phys. A **34**, 6197 (2001).

[39] X. Guan, K. D. Launey, M. Xie, L. Bao, F. Pan, and J. P. Draayer, Comput. Phys. Commun. **185**, 2714 (2014).

[40] J. G. Hirsch, A. Mariano, J. Dukelsky, and P. Schuck, Ann. Phys. (N.Y.) **296**, 187 (2002).

[41] J. K. Percus, *Combinatorial Methods*, Applied Mathematical Sciences Vol. 4 (Springer, New York, 1971).

[42] C. W. Borchardt, J. Reine Angew. Math. **1857**, 193 (1857).

[43] A. G. Izergin, D. A. Coker, and V. E. Korepin, J. Phys. A **25**, 4315 (1992).

[44] D. Singer, Electron. J. Combinatorics **11**, R48 (2004).

[45] K. Hao, W.-L. Yang, H. Fan, S.-Y. Liu, K. Wu, Z.-Y. Yang, and Y.-Z. Zhang, Nucl. Phys. B **862**, 835 (2012).

[46] N. Slavnov, Theor. Math. Phys. **79**, 502 (1989).

[47] C. Dunning, M. Ibañez, J. Links, G. Sierra, and S.-Y. Zhao, J. Stat. Mech. (2010) P08025.

[48] O. El Araby and D. Baeriswyl, Phys. Rev. B **89**, 134521 (2014).

[49] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge University Press, Cambridge, 1993).