Tensor networks and large-scale eigenvalue problems

Dissertation

zur Erlangung des Grades
Doktor der Naturwissenschaften

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität
in Mainz

vorgelegt von

Martin Johannes Dönges
geboren in Mainz

Mainz, im August 2024

Tag der mündlichen Prüfung: 17. Dezember 2024

Abstract

The present thesis is concerned with the numerical solution of large-scale eigenvalue problems via methods based on tensor networks. We address the computation of an eigenvector associated with the minimal eigenvalue of a Hamilton operator modeling a quantum mechanical many-particle spin system. If d denotes the number of particles and q the number of degrees of freedom for each particle, then the size of this symmetric eigenvalue problem scales exponentially in d to the base q. A well-known strategy with regard to the numerical treatment is to reshape those components of a method whose size scales exponentially into a tensor, i.e. a higher-dimensional array of numbers. This opens up the possibility of representation in a low-rank tensor format with non-exponentially growing ranks. In this way, a numerical method may be formulated by means of tensor networks which are a convenient tool when working with low-rank tensor formats. Our focus is on the q-XYZ and the q-Potts model. We introduce sets of vectors with a certain sparsity pattern whose definition is based on the q-ary representation of the indices of the potential nonzero entries. We show that for each eigenspace of a Hamilton operator of the q-XYZ or q-Potts model, there exists an orthonormal basis all of whose elements have a sparsity pattern suitable for the model. Moreover we show that for the q-XYZ model, the characterization of the sparsity pattern depends on the coupling parameters determining the Hamilton operator. We consider the hierarchical Tucker format and the tensor train format and describe how to represent the given Hamilton operators in these formats. In order to solve the eigenvalue problem numerically, we employ on the one hand the locally optimal conjugate gradient method expressed in the hierarchical Tucker format and on the other hand the modified alternating linear scheme, also called density matrix renormalization group, built on the representation in the tensor train format. A key aspect of this thesis is the construction of an initial guess for a numerical eigensolver. We propose a strategy which utilizes information about the relation of solutions for two different small problem sizes to construct an approximation of a solution for the original large problem which then may be used as an initial guess in an iterative method. This particular way of construction is enabled by the specific sparsity patterns of the eigenvectors and we show that the result of the construction has a sparsity pattern suitable for an eigenvector as well. In this context we discuss for the q-XYZ model two alternative construction strategies depending on the coupling parameters. We explain how to carry out the construction of the initial guess in the hierarchical Tucker format and in the tensor train format and that the ranks of the initial guess are independent of d or scale only linearly in d. To demonstrate the effect of our contribution on the computational cost, we conduct extensive numerical tests. We compare the constructed initial guesses with random initial guesses in terms of an acceleration of the methods.

Zusammenfassung

Die vorliegende Arbeit behandelt die numerische Lösung von groß-skaligen Eigenwertproblemen durch Methoden, welche auf Tensornetzwerken basieren. Wir beschäftigen uns mit der Bestimmung eines Eigenvektors zum minimalen Eigenwert eines Hamilton-Operators, der ein quantenmechanisches Vielteilchen-Spinsystem modelliert. Wenn d die Anzahl der Teilchen und q die Anzahl der Freiheitsgrade für jedes Teilchen beschreibt, skaliert die Größe dieses symmetrischen Eigenwertproblems exponentiell in d zur Basis q. Im Hinblick auf die numerische Behandlung ist es eine wohlbekannte Strategie, diejenigen Komponenten eines Verfahrens, deren Größe exponentiell skaliert, jeweils in einen Tensor, das heißt in ein höherdimensionales Zahlenfeld, umzusortieren. Dies eröffnet die Möglichkeit der Darstellung in einem Niedrigrang-Tensorformat mit nicht exponentiell wachsenden Rängen. Auf diese Art und Weise kann ein numerisches Verfahren mithilfe von Tensornetzwerken, die ein praktisches Werkzeug sind, um mit Niedrigrang-Tensorformaten zu arbeiten, formuliert werden. Unser Fokus liegt auf dem q-XYZ- und dem q-Potts-Modell. Wir führen Mengen von Vektoren mit einer gewissen Besetzungsstruktur ein, deren Definition auf der q-nären Darstellung der Indizes der potenziellen Nicht-Null-Einträge beruht. Wir zeigen, dass für jeden Eigenraum eines Hamilton-Operators des q-XYZ- oder q-Potts-Modells eine Orthonormalbasis existiert, deren Elemente alle eine zum Modell passende Besetzungsstruktur haben. Zudem zeigen wir, dass für das q-XYZ-Modell die Charakterisierung der möglichen Besetzungsstrukturen von der Wahl der den Hamilton-Operator festlegenden Kopplungsparameter abhängt. Wir betrachten das hierarchische Tucker-Format und das Tensor-Train-Format und beschreiben, wie sich die vorliegenden Hamilton-Operatoren in diesen Formaten darstellen lassen. Zur numerischen Lösung des Eigenwertproblems ziehen wir zum einen das lokal optimale konjugierte-Gradienten-Verfahren in einer Formulierung im hierarchischen Tucker-Format und zum anderen das modifizierte alternierende lineare Schema, auch Dichtematrix-Renormierungsgruppe genannt, das auf der Darstellung im Tensor-Train-Format aufbaut, heran. Einen Schwerpunkt dieser Arbeit bildet die Konstruktion einer Startlösung für einen numerischen Eigenlöser. Wir schlagen eine Strategie vor, die Informationen über die Beziehung von Lösungen für zwei verschiedene kleine Problemgrößen nutzt, um daraus eine Approximation einer Lösung für das ursprüngliche große Problem zu konstruieren, welche dann als Startlösung in einem iterativen Verfahren verwendet werden kann. Die konkrete Art und Weise dieser Konstruktion wird einerseits durch die speziellen Besetzungsstrukturen der Eigenvektoren ermöglicht, andererseits zeigen wir, dass das Resultat der Konstruktion eine zu den Eigenvektoren passende Besetzungsstruktur hat. In diesem Zusammenhang diskutieren wir für das q-XYZ-Modell zwei alternative Konstruktionsstrategien in Abhängigkeit der Kopplungsparameter. Wir erklären, wie sich die Konstruktion der Startlösung im hierarchischen Tucker-Format und im Tensor-Train-Format durchführen lässt und dass die Ränge der Startlösung unabhängig von d sind oder nur linear in d skalieren. Um den Effekt unseres Beitrags auf die Laufzeitkosten zu demonstrieren, führen wir umfangreiche numerische Tests durch. Dabei vergleichen wir die konstruierten Startlösungen mit zufällig gewählten Startlösungen hinsichtlich einer Beschleunigung der Verfahren.

To all my supporters

—

Allen meinen Unterstützern

Contents

1. Introduction **1**
1.1. Motivation and task . . . 1
1.2. Structure of the thesis . . . 4

2. Hamilton operators of spin systems **5**
2.1. XYZ model . . . 5
2.2. Potts model . . . 23

3. Tensors and tensor formats **31**
3.1. Basic concepts . . . 31
3.2. Tensor networks and tensor contractions . . . 34
3.3. Hierarchical Tucker format . . . 36
3.3.1. Construction . . . 37
3.3.2. Low-rank property . . . 42
3.3.3. Representation of Hamilton operators . . . 42
3.3.4. Arithmetical operations . . . 46
3.4. Tensor train format . . . 48

4. Eigensolvers in tensor formats **51**
4.1. Locally optimal conjugate gradient method . . . 51
4.2. Modified alternating linear scheme . . . 56

5. Construction of an initial guess **63**
5.1. Full vector format . . . 64
5.1.1. 2-XYZ model, $A \neq B$ . . . 64
5.1.2. 3-XYZ model, $A \neq B$ . . . 72
5.1.3. 3-Potts model . . . 77
5.2. Linear HT format . . . 85
5.3. Balanced HT format . . . 102
5.4. TT format . . . 106
5.5. XYZ model, $A = B$ . . . 106

6. Numerical tests **111**
6.1. Truncation parameters . . . 112
6.2. HT format, $A \neq B$ . . . 122
6.3. TT format, $A \neq B$ . . . 136
6.4. HT format, $A = B$ . . . 150
6.5. TT format, $A = B$ . . . 158

7. Conclusion **163**
7.1. Summary . . . 163
7.2. Outlook . . . 167

A. Dirac bra-ket notation **169**

B. Sparsity patterns of vectors **171**

C. Invertibility of reshaped eigenvectors for $d = 2$ **177**

List of algorithms **183**

List of figures **187**

Bibliography **189**

1. Introduction

1.1. Motivation and task

Eigenvalue problems are a major topic in numerical linear algebra [GVL13], [HJ13], [TB97]. Especially the case of large [Saa11] or symmetric [Par98] eigenvalue problems is an important field of mathematical research.

Both properties are combined by a matrix that represents a Hamilton operator modeling the energy in a quantum mechanical many-particle system and contains only real entries. Following the fundamental principles of quantum mechanics [Bal15, Chapter 2], to an observable quantity of one particle or component of the system there is associated in general a Hermitian matrix. The possible results of a measurement of the particle observable are the eigenvalues of the matrix and the result of the measurement equals a specific eigenvalue with probability 1 if and only if the particle is in a state that corresponds to an eigenvector associated with the respective eigenvalue. If we consider a many-particle system composed of single particles or components, then the size of the matrix representing the Hamilton operator, which in turn corresponds to an observable quantity of the composite system, equals by [Bal15, Section 3.5] the product of the matrix sizes for each component of the system. Hence the matrix corresponding to the composite many-particle system quickly gets large if the number of particles increases. Besides [Bal15], we also recommend [Gri14], [Tow00], or [NO08, Chapter 2] for an exposition of the underlying rules of quantum mechanics.

In condensed matter physics, a wide branch of research is the investigation of quantum mechanical spin systems, see [PF10], [VLRK03], or [VMC08, Section 1]. An important task, cf. [Nac04, Section 5] or [STG+19, Section 1.5], is to determine an eigenvector associated with the minimal eigenvalue of a Hamilton operator. Such an eigenvector describes the so-called ground state of the system in which the energy is minimal.

Our focus is on one-dimensional systems where d particles are arranged along a line. The energy in this system is modeled in such a way that only the interactions between the spins of two neighboring particles and additionally the interactions of the spins of the single particles with an external magnetic field are taken into account. We impose so-called open boundary conditions, which means that the two outermost particles do not interact with each other, in contrast to periodic boundary conditions, where the arrangement of the particles is considered as a ring. The number of degrees of freedom for a single particle induced by the spin is equal for all particles and we denote this quantity by q, concentrating on the cases $q = 2$ or $q = 3$. Then the Hamilton operator is a sum of Kronecker products of d factors each of which is a $q \times q$ matrix. The interaction of two specific neighboring particles enters the Hamilton operator as a summand where only the two, itself neighboring, Kronecker factors related to the two particles are not identity matrices while all other factors are $q \times q$ identities. The interaction of one specific particle with the external magnetic field enters as a summand where only one factor is a non-identity. Hence the Hamilton operator is of size $q^d \times q^d$ and in the cases we consider it is real symmetric. Accordingly, an eigenvector has q^d elements. This value grows exponentially with the number of particles d, an issue which is called the curse of dimensionality. So, without further considerations, the problem of finding

an eigenvector associated with the minimal eigenvalue is not tractable for d greater than around 25, as the objects to compute with do not fit into the memory.

In order to be able to treat the eigenvalue problem also for larger d numerically via an iterative method, it is necessary to represent the Hamilton operator as well as the iterates of an eigenvector in a format in which the memory and computational demand does not scale exponentially in d. We impose the assumption that this representation is possible in an exact way for the Hamilton operator. Going along with that, we make the ansatz that an eigenvector and also its iterates may be represented at least approximately in such a format with a non-exponential amount of parameters, cf. [VMC08, Section 2].

By [Ose10], the key approach is to regard the single elements of the Hamilton operator respectively the eigenvector and its iterates not as being arranged in a matrix respectively a vector but rather in a higher-dimensional array of numbers which is called a tensor, see [Hac19] for a comprehensive treatment. This different point of view opens up more possibilities to represent the handled objects in a particular tensor format, cf. [Orú19]. Such formats generalize ideas related to the singular value decomposition of matrices, especially the notion of the rank as a number of how many parameters are necessary to represent a given matrix or tensor consisting of single elements. We focus on the hierarchical Tucker (HT) format [HK09] and the tensor train (TT) format [Ose11]. Both formats are instances of a tensor network [Orú14], a concept and perspective very beneficial in the discussion of tensors and their numerical treatment. Indeed, for the spin systems we consider, the number of parameters to represent the tensorized versions of their Hamilton operators scales only linearly in d, cf. [Tob12, Example 3.8].

Concerning the computation of an eigenvector associated with the minimal eigenvalue, there are essentially two classes of algorithms, see [Bac23, Section 6.2]. One approach, like in [Tob12, Section 6.1] with respect to the LOPCG method from [Kny01], is to translate an iterative method designed originally for the classical case of matrices and vectors to their representations in a tensor format and to perform the iterations in the arithmetics of this tensor format. The second approach, called DMRG in [Sch11] or (M)ALS in [HRS12], exploits more directly the structure of the tensor format and updates the single characteristic components of the related tensor network in a specific order.

To the best of our knowledge, an issue which is not addressed in detail so far is the choice of an initial guess for the methods. In [HB09, Chapter V] the necessity of a well-suited initial guess for an eigensolver in general is mentioned, while [Sch11, Chapter 6] formulates this demand for the DMRG algorithm. The perhaps most simple option would be to choose an initial guess randomly which is possible in all the cases we consider. This approach however does not take into account any information about the problem at hand. So, our main concern in this thesis is to develop a more sophisticated construction of an initial guess tailored better to the concrete setting. The main benefit we expect from employing a constructed initial guess is a speed-up of the numerical method in the sense that, compared to a random initial guess, less iteration steps are needed to reach a certain level of approximation of the exact solution. In our opinion, three properties are essential to meet that task.

(P1) The computational cost to construct the initial guess should be small compared to the iteration itself. It may be larger than for setting up a random initial guess but its share in the overall amount of work should be negligible.

(P2) The constructed initial guess should match a low-rank tensor format. As the considered eigenvalue problem only becomes tractable by formulating it with the help of tensors with low ranks, the same should hold for the initial guess.

(P3) The initial guess should approximate an exact eigenvector associated with the minimal eigenvalue. It appears reasonable to assume that an iterative numerical method approaches an exact solution faster if an initial guess is in an appropriate sense closer to this solution.

We borrow a concept that is known from the field of multigrid methods [Hac85]. This means to compute in an exact way, or at least with very high accuracy, a solution of a problem that is regarded as a restriction of the overall problem to a small size where it can be handled at low cost, and then to prolongate this solution to the larger original problem size. One expects that this procedure yields a significant improvement for the large problem size where computations are more expensive. We notice that in the context of the DMRG method, a technique to transfer the prolongation operation to the computation of ground states was developed in [LMJ18] and that the general idea to combine information about small parts of a many-particle system to obtain an initial guess for the overall system appears in [HWSH13a, Section 5], but the respective approaches are very different from ours.

The strategy we propose is to compute up to machine accuracy an eigenvector associated with the minimal eigenvalue for two different small problem sizes which are characterized by two small, say one-digit, numbers $d_1 < d_2$ of particles in the spin system. All other parameters defining the Hamilton operator are the same as for the larger number of d particles. The prolongation operator is set up in a specific way with the requirement that it maps the computed eigenvector for problem size d_1 to the eigenvector corresponding to d_2. Based on this relation, a prolongation to larger problem sizes is defined. We then construct the initial guess by mapping the eigenvector corresponding to d_2 with the obtained prolongation operator up to the original overall problem size d. Hence, the prolongation strategy is built on the idea that the relation between eigenvectors for two small problem sizes approximates the relation also between eigenvectors for larger problems.

In order to put that strategy to use, we have to decide how the relation between eigenvectors corresponding to d_1 and d_2 is exploited precisely. An understanding of the structure of the eigenvectors for different problem sizes helps to design a structure of the prolongation operator. It also enables us to realize whether there are situations in which the prolongation strategy is not feasible and, in such a case, what type of adaption could be indicated. Thus, one of the important tasks is to gain knowledge about the structure of the eigenvectors.

Regarding the requirement of (P1), the undoubtedly cheap computation of eigenvectors for small d_1 and d_2 has to be supplemented with a prolongation operator whose application causes no relevant computational cost. In accordance with (P2), we have to investigate in which way the prolongation may be formulated in a tensor format with low ranks. It is inherent to the arithmetics in the considered tensor formats that the ranks grow significantly during repeated arithmetical operations like applications of an operator. Our task is to think about how to deal with that problem. Concerning (P3), we have to check the approximation quality of the constructed initial guess. This is in turn closely related to the question whether the result of the prolongation is consistent with the structure of an exact eigenvector.

After having described and discussed the construction of an initial guess, another important part of our considerations is to examine the actual behavior of such an initial guess in a numerical method, in particular whether it yields the targeted reduction of necessary iteration steps. We have to compare the performance of a constructed initial guess to that of a random initial guess by extensive numerical tests. To reach a conclusion about the usefulness of the proposed constructions, our task is to name advantages and disadvantages in various scenarios.

1.2. Structure of the thesis

This thesis contains seven chapters and three appendices.

Chapter 1 introduces the subject matter, motivates its consideration, lists problems and tasks, and describes the structure of the thesis.

In Chapter 2 we define the Hamilton operators we are concerned with, namely those of the q-XYZ model and the q-Potts model. We prove that for each eigenspace of these Hamilton operators there exists an orthonormal basis whose elements are vectors with a specific sparsity pattern, hence that especially in eigenvectors associated with simple eigenvalues nonzero entries may occur only at certain positions.

Chapter 3 gives an overview of the notion of tensors, tensor networks, and tensor formats, in particular the hierarchical Tucker format and the tensor train format. The concept of rank is described as well as the characteristics of arithmetical operations within these tensor formats. It is stated how the Hamilton operators may be efficiently represented.

In Chapter 4 we explain the two types of algorithms which are employed to compute an eigenvector associated with the minimal eigenvalue. We mention the adaption of the locally optimal conjugate gradient method to iterates in the hierarchical Tucker format and the modified alternating linear scheme formulated in terms of the tensor train format.

Chapter 5 contains our main contribution, the construction of an initial guess. Referring to the results concerning the sparsity pattern of eigenvectors in Chapter 2, at first we propose how to set up a prolongation operator based on information about eigenvectors in two small problem sizes and how to utilize this prolongation operator in order to construct vectors for larger problem sizes. We discuss in full vector format the feasibility of this construction scheme in terms of reshaping eigenvectors into matrices. We investigate in which way a prolongated vector approximates an exact eigenvector for the large problem size. Then we transfer the construction procedure to the hierarchical Tucker and the tensor train format. We demonstrate that the construction yields a tensor represented with ranks independent of the problem size up to which the prolongation is carried out. For a situation where the construction by prolongation is not feasible due to the non-invertibility of the matricization of an eigenvector, we propose an alternative again based on the possible sparsity patterns of eigenvectors. This alternative initial guess may be represented with ranks scaling linearly with the problem size.

In Chapter 6 various numerical tests are performed. After having decided how to choose the parameters controlling the truncation of ranks during the iteration, we investigate to what extent the choice of the initial guess influences the course and the result of the method. We split the discussion due to the four combinations of the two different types of algorithms from Chapter 4 and the two different types of initial guesses constructed in Chapter 5.

Chapter 7 summarizes the findings obtained in this thesis and relates them to the tasks formulated in Chapter 1. We close our contribution to the subject by pointing to possible further research.

The main text is supplemented by three appendices. Appendix A collects basic facts about the Dirac bra-ket notation which is particularly useful in the context of Kronecker product structures and interacts well with the discussed sparsity patterns of vectors. In Appendix B several examples for these sparsity patterns are given. Appendix C contains some statements about the invertibility of eigenvectors reshaped into a matrix which are referred to during the construction of the initial guess.

2. Hamilton operators of spin systems

We consider Hamilton operators modeling one-dimensional systems of particles with spin $s \in \{\frac{1}{2}, 1, \frac{3}{2}, 2, \dots\}$. The number of different eigenstates of a particle with spin s is denoted by q, so $q = 2s + 1 \in \{2, 3, 4, 5, \dots\}$ due to [Bal15, Sect. 7.4]. In the sequel we use the Dirac notation with the *ket* $|j\rangle = \mathbf{e}_{j+1} \in \mathbb{R}^q$, $0 \leq j \leq q-1$, and the *bra* $\langle j| = \mathbf{e}_{j+1}^\top$ whose properties relevant in our context are summarized in Appendix A.

We are concerned with two different types of Hamilton operators, namely the *XYZ model* in Section 2.1 and the *Potts model* in Section 2.2. Besides the definition of these Hamilton operators, the main subject of the present chapter is a discussion of the *sparsity pattern* of the eigenvectors of the Hamilton operator, by which we mean the position of potential nonzero entries in the eigenvectors. In addition, we determine in some special cases the minimal eigenvalue and associated eigenvectors analytically.

2.1. XYZ model

Definition 2.1. Let $d, q \in \mathbb{N}$ with $q \geq 2$.

(i) The *spin matrices* $\mathbf{S}_{x,q} \in \mathbb{R}^{q\times q}$, $\mathbf{S}_{y,q} \in \mathbb{C}^{q\times q}$, $\mathbf{S}_{z,q} \in \mathbb{R}^{q\times q}$ are

$$\begin{aligned}
\mathbf{S}_{x,q} &:= \sum_{j=1}^{q-1} c_j \left(|j-1\rangle \langle j| + |j\rangle \langle j-1| \right), \\
\mathbf{S}_{y,q} &:= -\mathrm{i} \sum_{j=1}^{q-1} c_j \left(|j-1\rangle \langle j| - |j\rangle \langle j-1| \right), \\
\mathbf{S}_{z,q} &:= \sum_{j=0}^{q-1} \frac{q-1-2j}{2} |j\rangle \langle j|,
\end{aligned}$$

with $c_j := \frac{1}{2}\sqrt{\frac{q-1}{2}\frac{q+1}{2} - \left(j - \frac{q-1}{2}\right)\left(j - \frac{q-1}{2} - 1\right)} = \frac{1}{2}\sqrt{j(q-j)}$, cf. [Gri14, Sect. 4.4].

(ii) For $A, B, \Delta, h \in \mathbb{R}$, the Hamilton operator $\mathbf{H}_{d,q}^{\mathrm{XYZ}} \in \mathbb{R}^{q^d \times q^d}$ of the *q-XYZ model* for a system of d particles is

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}} := A \sum_{i=1}^{d-1} \mathbf{S}_{x,q}^{(i)} \mathbf{S}_{x,q}^{(i+1)} + B \sum_{i=1}^{d-1} \mathbf{S}_{y,q}^{(i)} \mathbf{S}_{y,q}^{(i+1)} + \Delta \sum_{i=1}^{d-1} \mathbf{S}_{z,q}^{(i)} \mathbf{S}_{z,q}^{(i+1)} + h \sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)}, \tag{2.1}$$

where the notation

$$\mathbf{S}^{(i)}\mathbf{S}^{(i+1)} := \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q \otimes \mathbf{S} \otimes \mathbf{S} \otimes \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q$$

respectively

$$\mathbf{S}^{(i)} := \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q \otimes \mathbf{S} \otimes \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q$$

means that the matrix $\mathbf{S} \in \{\mathbf{S}_{x,q}, \mathbf{S}_{y,q}, \mathbf{S}_{z,q}\}$ is the i-th and $(i+1)$-th respectively i-th factor in the Kronecker product with the identity matrix $\mathbf{I}_q \in \mathbb{R}^{q\times q}$.

Remark and Example 2.2. (i) We omit the prefix q- in situations when there is no need to refer to it. The name "XYZ model" appears for example in [Bax07, Sect. 10.14] and [Eck19, Sect. 13.1]. Depending on the parameters, the model may be called

(a) $A = B$: XXZ model,

(b) $\Delta = 0$: XY model,

(c) $A = B,\ \Delta = 0$: XX model,

(d) $B = \Delta = 0$: Ising model.

Another commonly used name for this class of models, with varying conventions on the mutual relations of A, B, or Δ, and the possibility of h being nonzero, is *Heisenberg model*, dating back to [Hei28].

(ii) The spin matrix $\mathbf{S}_{x,q}$ is symmetric since both $|j-1\rangle\langle j|$ and $|j\rangle\langle j-1|$ have the same coefficient c_j. Due to the prefactor i and the different signs of $|j-1\rangle\langle j|$ and $|j\rangle\langle j-1|$, $\mathbf{S}_{y,q}$ is Hermitian. $\mathbf{S}_{z,q}$ is diagonal, as only $|j\rangle\langle j|$ occurs.

(iii) All three spin matrices have the eigenvalues $\lambda_i = -\frac{q-1-2(i-1)}{2}$, $1 \le i \le q$, cf. [Tow00, Sect. 3.8]. Eigenvectors associated with λ_1 and λ_q are listed in Proposition 2.11(i)-(iii).

(iv) It is $\mathbf{S}_{\omega,2} = \frac{1}{2}\boldsymbol{\sigma}_\omega$, $\omega \in \{x, y, z\}$, with the *Pauli matrices*

$$\boldsymbol{\sigma}_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

For $q = 3$, that means for spin $s = 1$, we obtain

$$\mathbf{S}_{x,3} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{S}_{y,3} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \quad \mathbf{S}_{z,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(v) The parameters $A, B, \Delta \in \mathbb{R}$ are regarded as the strength of the coupling of two neighboring spins in the respective direction and $h \in \mathbb{R}$ models an external magnetic field, cf. [LRV04, Sect. III & IV].

(vi) $\mathbf{H}_{d,q}^{\mathrm{XYZ}} \in \mathbb{R}^{q^d \times q^d}$ is symmetric since the Kronecker product of symmetric resp. Hermitian matrices is symmetric resp. Hermitian by [HJ91, Sect. 4.2] and the imaginary unit i in $\mathbf{S}_{y,q}$ is squared when forming $\mathbf{S}_{y,q} \otimes \mathbf{S}_{y,q}$.

We prove a certain structure of the eigenvectors of these Hamilton operators. To be precise, a statement on the existence of eigenvectors with a maximal number of nonzero entries and their position in the eigenvector is made. Before stating the result on the structure of the eigenvectors, we prove a relation between invariant subspaces and eigenvectors.

Lemma 2.3. *Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be Hermitian. Let $U_i \subset \mathbb{C}^n$, $1 \le i \le k$ with $1 \le k \le n$, be an $\mathbf{A}$-invariant subspace, which means $\mathbf{Ax} \in U_i$ for all $\mathbf{x} \in U_i$. Let further $U_1 \oplus \cdots \oplus U_k = \mathbb{C}^n$ and $U_i \perp U_j$, which means orthogonality with respect to the standard inner product characterized by $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^*\mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, for all $1 \le i < j \le k$. Let λ be an eigenvalue of $\mathbf{A}$ with multiplicity μ and let V be the associated eigenspace.*
Then there exists an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_\mu\} \subset \mathbb{C}^n$ of V such that for each $1 \le \alpha \le \mu$ there exists an index $1 \le i(\alpha) \le k$ with $\mathbf{v}_\alpha \in U_{i(\alpha)}$.

Proof. Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_\mu\}$ be an orthonormal basis of the μ-dimensional eigenspace V associated with λ and let $\{\mathbf{u}_1^{(i)}, \ldots, \mathbf{u}_{d_i}^{(i)}\}$ with $d_i := \dim U_i$ be an orthonormal basis of U_i for $1 \leq i \leq k$. Step by step, we will construct the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_\mu\}$ satisfying the claimed properties with the help of the elements of $\{\mathbf{w}_1, \ldots, \mathbf{w}_\mu\}$. Define

$$\mathbf{U}_i := \begin{pmatrix} \mathbf{u}_1^{(i)} & \cdots & \mathbf{u}_{d_i}^{(i)} \end{pmatrix} \in \mathbb{C}^{n \times d_i}.$$

Since U_i is $\mathbf{A}$-invariant, there exists a matrix $\mathbf{C}_i := \left(c_{j,l}^{(i)}\right)_{1 \leq j \leq d_i,\, 1 \leq l \leq d_i}$ such that

$$\mathbf{A}\mathbf{u}_l^{(i)} = \sum_{j=1}^{d_i} c_{j,l}^{(i)} \mathbf{u}_j^{(i)}$$

for each $1 \leq l \leq d_i$, which implies

$$\mathbf{A}\mathbf{U}_i = \mathbf{U}_i \mathbf{C}_i.$$

Due to the Hermiticity of $\mathbf{A}$ it is

$$\mathbf{C}_i = \mathbf{U}_i^* \mathbf{A} \mathbf{U}_i = \mathbf{U}_i^* \mathbf{A}^* \mathbf{U}_i = (\mathbf{U}_i^* \mathbf{A} \mathbf{U}_i)^* = \mathbf{C}_i^*$$

and further

$$\mathbf{C}_i \mathbf{U}_i^* = \mathbf{C}_i^* \mathbf{U}_i^* = (\mathbf{U}_i \mathbf{C}_i)^* = (\mathbf{A}\mathbf{U}_i)^* = \mathbf{U}_i^* \mathbf{A}^* = \mathbf{U}_i^* \mathbf{A}.$$

We start arguing with $\mathbf{w}_1$ and consider the orthogonal projection of $\mathbf{w}_1$ onto the complement of U_1,

$$P_{U_1^\perp}(\mathbf{w}_1) = (\mathbf{I}_n - \mathbf{U}_1 \mathbf{U}_1^*)\mathbf{w}_1.$$

If on the one hand $P_{U_1^\perp}(\mathbf{w}_1) = \mathbf{0}$, then $\mathbf{w}_1 \in U_1$.
If on the other hand $P_{U_1^\perp}(\mathbf{w}_1) \neq \mathbf{0}$, assume as a first case that also $P_{U_1}(\mathbf{w}_1) = \mathbf{U}_1 \mathbf{U}_1^* \mathbf{w}_1 \neq \mathbf{0}$. Then

$$\mathbf{A} P_{U_1}(\mathbf{w}_1) = \mathbf{A}\mathbf{U}_1 \mathbf{U}_1^* \mathbf{w}_1 = \mathbf{U}_1 \mathbf{C}_1 \mathbf{U}_1^* \mathbf{w}_1 = \mathbf{U}_1 \mathbf{U}_1^* \mathbf{A} \mathbf{w}_1 = \lambda \mathbf{U}_1 \mathbf{U}_1^* \mathbf{w}_1 = \lambda P_{U_1}(\mathbf{w}_1), \tag{2.2}$$

thus $P_{U_1}(\mathbf{w}_1) \in U_1$ is an eigenvector of λ. So assume as a second case that $P_{U_1}(\mathbf{w}_1) = \mathbf{0}$, hence $\mathbf{w}_1 \in U_1^\perp = U_2 \oplus \cdots \oplus U_k$. Repeating the argument for $i = 2, \ldots, k-1$ either yields $\mathbf{w}_1 \in U_i$ or the existence of an eigenvector $P_{U_i}(\mathbf{w}_1) \in U_i$ or in the end $\mathbf{w}_1 \in U_k$ since $\mathbf{w}_1 \neq \mathbf{0}$. By the above procedure we obtain an index $1 \leq i(1) \leq k$ for the appropriately defined and normalized eigenvector $\mathbf{v}_1$ of λ such that $\mathbf{v}_1 := \mathbf{w}_1 \in U_{i(1)}$ or $\mathbf{v}_1 := P_{U_{i(1)}}(\mathbf{w}_1) \in U_{i(1)}$.
Now we project all spaces U_i and V onto $(\mathrm{span}\{\mathbf{v}_1\})^\perp$. This projection leaves the basis vectors of all U_i with $i \neq i(1)$ unchanged since $U_i \perp U_{i(1)}$. For the space $U_{i(1)} \cap (\mathrm{span}\{\mathbf{v}_1\})^\perp$ of dimension $d_{i(1)} - 1$ there exists an orthonormal basis $\{\widetilde{\mathbf{u}}_1^{(i(1))}, \ldots, \widetilde{\mathbf{u}}_{d_{i(1)}-1}^{(i(1))}\}$. Also for $V \cap (\mathrm{span}\{\mathbf{v}_1\})^\perp$ there may be found an orthonormal basis $\{\widetilde{\mathbf{w}}_2, \ldots, \widetilde{\mathbf{w}}_\mu\}$ with $\widetilde{\mathbf{w}}_2 := P_{(\mathrm{span}\{\mathbf{v}_1\})^\perp}(\mathbf{w}_2)$.
We continue by arguing with $\widetilde{\mathbf{w}}_2$. It is $\widetilde{\mathbf{w}}_2 \neq \mathbf{0}$ since otherwise $\mathbf{w}_2 \in \mathrm{span}\{\mathbf{v}_1\}$ which is by construction not orthogonal to $\mathbf{w}_1$, in contradiction to the orthogonality of $\mathbf{w}_2$ and $\mathbf{w}_1$. Performing the above procedure again yields $\widetilde{\mathbf{w}}_2 \in U_{i(2)}$ for an $1 \leq i(2) \leq k$ or the existence of an eigenvector $P_{U_{i(2)}}(\widetilde{\mathbf{w}}_2) \in U_{i(2)}$. So, by setting $\mathbf{v}_2 := \widetilde{\mathbf{w}}_2$ or $\mathbf{v}_2 := P_{U_{i(2)}}(\widetilde{\mathbf{w}}_2)$ and normalizing, we obtain an eigenvector $\mathbf{v}_2 \in U_{i(2)}$ for some $i(2)$. As the whole procedure took place in $(\mathrm{span}\{\mathbf{v}_1\})^\perp$, it is $\mathbf{v}_2 \perp \mathbf{v}_1$.

For the next step all spaces U_i and V, especially $\widetilde{\mathbf{w}}_3$, are projected onto $(\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\})^\perp$. We repeat the construction procedure for all (in each step adapted) $\widetilde{\mathbf{w}}_\alpha$ up to and including $\alpha = \mu$ and obtain in each step a normalized eigenvector $\mathbf{v}_\alpha$ of λ which is an element of one $U_{i(\alpha)}$ and is orthogonal to all $\mathbf{v}_\beta$ with $\beta \neq \alpha$. Thus the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_\mu\}$ is an orthonormal basis of V. □

Corollary 2.4. *Under the assumptions of Lemma 2.3, an eigenvector* $\mathbf{v}$ *associated with a* simple *eigenvalue* λ *of* $\mathbf{A}$ *is contained in one of the* U_i*.*

Proof. All eigenvectors of a simple eigenvalue λ only differ by a scalar multiple and so are contained in the same subspace $U_{i(\mu)}$ as that eigenvector $\mathbf{v}_{i(\mu)}$ whose existence is stated by Lemma 2.3 for $\mu = 1$. □

Remark 2.5. (i) The argumentation surrounding (2.2) is possible for an arbitrary subspace U_i, $1 \leq i \leq k$, fulfilling the supposition of Lemma 2.3 and an arbitrary eigenvector $\mathbf{w}$ of $\mathbf{A}$. Hence, if $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$, then also

$$\mathbf{A}\mathbf{U}_i\mathbf{U}_i^*\mathbf{w} = \lambda\mathbf{U}_i\mathbf{U}_i^*\mathbf{w}$$

with $\mathbf{U}_i \in \mathbb{C}^{n \times \dim U_i}$ containing the elements of an orthonormal basis of U_i in its columns.

(ii) Given a decomposition of $\mathbb{C}^n$ into mutually orthogonal $\mathbf{A}$-invariant subspaces for a Hermitian $\mathbf{A} \in \mathbb{C}^{n \times n}$, there might exist an eigenvector associated with a *multiple* eigenvalue λ of $\mathbf{A}$ not belonging to any of these particular $\mathbf{A}$-invariant subspaces. Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathrm{i}/\sqrt{2} & -\mathrm{i}/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{i}/\sqrt{2} & -\mathrm{i}/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^* .$$

The spaces

$$U_1 = \operatorname{span}\left\{\begin{pmatrix} \mathrm{i} \\ 1 \\ 0 \end{pmatrix}\right\}, \quad U_2 = \operatorname{span}\left\{\begin{pmatrix} -\mathrm{i} \\ 1 \\ 0 \end{pmatrix}\right\}, \quad U_3 = \operatorname{span}\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$$

are invariant under $\mathbf{A}$, mutually orthogonal, and span $\mathbb{C}^3$, but

$$\mathbf{v} = \begin{pmatrix} -\mathrm{i} \\ 1 \\ 1 \end{pmatrix} \notin U_j \quad \text{for any } j = 1, 2, 3$$

is an eigenvector associated with the double eigenvalue $\lambda = 1$.

(iii) In case of a symmetric real matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, which is our main concern in the sequel, all quantities in the statement and the proof of Lemma 2.3 may be replaced by their real analogs.

We focus in the further course of the thesis on the most common case $s = \frac{1}{2}$ ($q = 2$) and sometimes $s = 1$ ($q = 3$), but the following statements may also be formulated for general s or q. We introduce sets of vectors with a certain *sparsity pattern* of potential nonzero entries. These sets are subspaces of $\mathbb{R}^{q^d}$ and we show their invariance under the Hamilton operator of the XYZ model, which in turn prepares a statement about the sparsity pattern of the eigenvectors of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$.

Definition 2.6. Let $d, q \in \mathbb{N}$ with $q \geq 2$.

(i) Let $\mathcal{E}_{d,q}^{\text{even}} \subset \mathbb{R}^{q^d}$ respectively $\mathcal{E}_{d,q}^{\text{odd}} \subset \mathbb{R}^{q^d}$ denote the set of vectors whose entries are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation whose sum of digits is even respectively odd.

(ii) Let $\mathcal{E}_{d,q}^{(k)} \subset \mathbb{R}^{q^d}$ denote the set of vectors whose entries are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation whose sum of digits equals $k \in \{0, \ldots, (q-1)d\}$.

We illustrate the meaning of Definition 2.6 for small d and q in Example B.1.

Remark 2.7. In case $q = 2$, the sets

(i) $\mathcal{E}_{d,2}^{\text{even}}$ respectively $\mathcal{E}_{d,2}^{\text{odd}}$,

(ii) $\mathcal{E}_{d,2}^{(k)}$

may also be characterized as these sets of vectors in $\mathbb{R}^{2^d}$ whose entries are nonzero at most at those positions (beginning counting with 0) which have a binary representation with

(i) an even number of ones respectively an odd number of ones,

(ii) exactly $k \in \{0, \ldots, d\}$ ones.

Lemma 2.8. *(i) $\mathcal{E}_{d,q}^{\text{even}}$ and $\mathcal{E}_{d,q}^{\text{odd}}$ are invariant under $\mathbf{H}_{d,q}^{\text{XYZ}}$.*

(ii) If additionally $A = B$, then $\mathcal{E}_{d,q}^{(k)}$ is invariant under $\mathbf{H}_{d,q}^{\text{XYZ}}$ for all $k \in \{0, \ldots, (q-1)d\}$.

Proof. We carry out in parallel the argumentation

(i) for arbitrary A and B,

(ii) for the special case $A = B$.

It is

(i) $\mathcal{E}_{d,q}^{\text{even [odd]}} = \text{span}\{\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_{p(d,q)}}\}$, where $i_1 - 1, \ldots, i_{p(d,q)} - 1$ are the

$$p(d,q) = \begin{cases} q^d/2, & q \text{ even} \\ (q^d+1)/2,\ [(q^d-1)/2], & q \text{ odd} \end{cases}$$

numbers between 0 and $q^d - 1$ which have a q-ary representation whose sum of digits is even respectively odd,

(ii) $\mathcal{E}_{d,q}^{(k)} = \text{span}\{\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_{t(d,q,k)}}\}$, where $i_1 - 1, \ldots, i_{t(d,q,k)} - 1$ are the

$$t(d,q,k) := \sum_{(\tau_0,\tau_1,\ldots,\tau_{q-1}) \in T(d,q,k)} \frac{(\tau_0 + \tau_1 + \ldots + \tau_{q-1})!}{\tau_0!\,\tau_1! \cdots \tau_{q-1}!}, \tag{2.3a}$$

$$T(d,q,k) := \left\{ (\tau_0, \tau_1, \ldots, \tau_{q-1}) \in \{0,1,\ldots,d\}^q \colon \sum_{j=0}^{q-1} (\tau_j \cdot j) = k,\ \sum_{j=0}^{q-1} \tau_j = d \right\}, \tag{2.3b}$$

numbers between 0 and $q^d - 1$ which have a q-ary representation whose sum of digits is $k \in \{0, \ldots, (q-1)d\}$. The summands in the definition of $t(d,q,k)$ are also called *multinomial coefficients*, cf. [AS65, Sect. 24.1].

Observe that the involved $\mathbf{e}_i$ are exactly the vectors $|a_1 \ \dots \ a_d\rangle$, $a_j \in \{0, \dots, q-1\}$, such that the sum of the a_j in $|\cdot\rangle$

(i) is even respectively odd,

(ii) equals k.

Therefore we have to study the effect of applying $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ to such a vector $\mathbf{y} := |a_1 \ \dots \ a_d\rangle$. So, let

(i) $\mathbf{w} \in \mathcal{E}_{d,q}^{\text{even [odd]}}$ be an arbitrary linear combination of the $\mathbf{e}_i \in \mathcal{E}_{d,q}^{\text{even [odd]}}$,

(ii) $\mathbf{w} \in \mathcal{E}_{d,q}^{(k)}$ be an arbitrary linear combination of the $\mathbf{e}_i \in \mathcal{E}_{d,q}^{(k)}$.

It is

$$\begin{aligned}
\mathbf{S}_{x,q} \otimes \mathbf{S}_{x,q} &= \left(\sum_{j=1}^{q-1} c_j \left(|j-1\rangle \langle j| + |j\rangle \langle j-1|\right)\right) \otimes \left(\sum_{j=1}^{q-1} c_j \left(|j-1\rangle \langle j| + |j\rangle \langle j-1|\right)\right) \\
&= \sum_{j=1}^{q-1} c_1 c_j \left[\left(|j-1\rangle \langle j| + |j\rangle \langle j-1|\right) \otimes \left(|0\rangle \langle 1| + |1\rangle \langle 0|\right)\right] \\
&\quad + \dots + \\
&\quad \sum_{j=1}^{q-1} c_{q-1} c_j \left[\left(|j-1\rangle \langle j| + |j\rangle \langle j-1|\right) \otimes \left(|q-2\rangle \langle q-1| + |q-1\rangle \langle q-2|\right)\right] \\
&= \sum_{j=1}^{q-1} c_1 c_j \big(|j-1\ 0\rangle \langle j\ 1| + |j-1\ 1\rangle \langle j\ 0| + |j\ 0\rangle \langle j-1\ 1| + |j\ 1\rangle \langle j-1\ 0| \big) \\
&\quad + \dots + \\
&\quad \sum_{j=1}^{q-1} c_{q-1} c_j \big(|j-1\ q-2\rangle \langle j\ q-1| + |j-1\ q-1\rangle \langle j\ q-2| \\
&\qquad + |j\ q-2\rangle \langle j-1\ q-1| + |j\ q-1\rangle \langle j-1\ q-2| \big) \\
&= \sum_{l=1}^{q-1} \sum_{j=1}^{q-1} c_l c_j \big(|j-1\ l-1\rangle \langle j\ l| + |j-1\ l\rangle \langle j\ l-1| \\
&\qquad + |j\ l-1\rangle \langle j-1\ l| + |j\ l\rangle \langle j-1\ l-1| \big).
\end{aligned}$$

In the same way we obtain

$$\begin{aligned}
\mathbf{S}_{y,q} \otimes \mathbf{S}_{y,q} = -\sum_{l=1}^{q-1} \sum_{j=1}^{q-1} c_l c_j \big(&|j-1\ l-1\rangle \langle j\ l| - |j-1\ l\rangle \langle j\ l-1| \\
&- |j\ l-1\rangle \langle j-1\ l| + |j\ l\rangle \langle j-1\ l-1| \big).
\end{aligned}$$

Furthermore,

$$\mathbf{S}_{z,q} \otimes \mathbf{S}_{z,q} = \sum_{l=1}^{q-1} \sum_{j=1}^{q-1} \frac{(q-1-2l)(q-1-2j)}{4} |j\ l\rangle \langle j\ l| .$$

This yields

$$\begin{aligned}
&A\left(\mathbf{S}_{x,q}\otimes\mathbf{S}_{x,q}\right)+B\left(\mathbf{S}_{y,q}\otimes\mathbf{S}_{y,q}\right)+\Delta\left(\mathbf{S}_{z,q}\otimes\mathbf{S}_{z,q}\right)\\
&=\sum_{l=1}^{q-1}\sum_{j=1}^{q-1}\Big[c_l c_j\Big((A-B)\left|j-1\ l-1\right\rangle\left\langle j\ l\right|+(A+B)\left|j-1\ l\right\rangle\left\langle j\ l-1\right|\\
&\qquad\qquad+(A+B)\left|j\ l-1\right\rangle\left\langle j-1\ l\right|+(A-B)\left|j\ l\right\rangle\left\langle j-1\ l-1\right|\Big)\\
&\qquad\qquad+\frac{(q-1-2l)(q-1-2j)}{4}\left|j\ l\right\rangle\left\langle j\ l\right|\Big].
\end{aligned}$$

For all

(i) five,

(ii) three

most inner summands $c\left|\alpha_1\ \alpha_2\right\rangle\left\langle\beta_1\ \beta_2\right|$, c the respective coefficient, with fixed j and l it holds

(i) $(\alpha_1+\alpha_2)-(\beta_1+\beta_2)\in\{-2,0,2\}$,

(ii) $(\alpha_1+\alpha_2)-(\beta_1+\beta_2)=0$.

Hence, as

$$\left|\ldots\ \alpha_i\ \alpha_{i+1}\ \ldots\right\rangle\left\langle\ldots\ \beta_i\ \beta_{i+1}\ \ldots\right|\ldots\ a_i\ a_{i+1}\ \ldots\rangle=\delta_{\beta_i,a_i}\delta_{\beta_{i+1},a_{i+1}}\left|\ldots\ \alpha_i\ \alpha_{i+1}\ \ldots\right\rangle,$$

the application of one summand

$$\begin{aligned}
&A(\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q\otimes\mathbf{S}_{x,q}\otimes\mathbf{S}_{x,q}\otimes\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q)\\
+\ &B(\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q\otimes\mathbf{S}_{y,q}\otimes\mathbf{S}_{y,q}\otimes\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q)\\
+\ &\Delta(\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q\otimes\mathbf{S}_{z,q}\otimes\mathbf{S}_{z,q}\otimes\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q)
\end{aligned}$$

or

$$h(\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q\otimes\mathbf{S}_{z,q}\otimes\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q)$$

of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$, and by linearity then also the application of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ itself, does not change

(i) the parity,

(ii) the value

of the sum of the entries in $\mathbf{y}=\left|a_1\ \ldots\ a_d\right\rangle$. This particularly holds true for each of the basis vectors $\mathbf{e}_i$ contributing to $\mathbf{w}$. So, the spaces

(i) $\mathcal{E}_{d,q}^{\mathrm{even}}$ and $\mathcal{E}_{d,q}^{\mathrm{odd}}$,

(ii) $\mathcal{E}_{d,q}^{(k)}$

are invariant under $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$. □

After this preparation, we come to the structural property of the eigenvectors of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$, a statement on the maximal number of nonzero entries and their position in an eigenvector.

Theorem 2.9. *Let λ be an eigenvalue of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ with multiplicity μ and let V be the associated eigenspace.*

(i) Then there exists an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_\mu\} \subset \mathbb{R}^{q^d}$ of V such that for each $1 \leq \alpha \leq \mu$ either $\mathbf{v}_\alpha \in \mathcal{E}_{d,q}^{\mathrm{even}}$ or $\mathbf{v}_\alpha \in \mathcal{E}_{d,q}^{\mathrm{odd}}$.

(ii) If additionally $A = B$, then for each $1 \leq \alpha \leq \mu$ there exists $0 \leq k(\alpha) \leq (q-1)d$ such that $\mathbf{v}_\alpha \in \mathcal{E}_{d,q}^{(k(\alpha))}$.

Proof. We carry out in parallel the argumentation

(i) for arbitrary A and B,

(ii) for the special case $A = B$.

By Lemma 2.8, the spaces

(i) $\mathcal{E}_{d,q}^{\mathrm{even}}$ and $\mathcal{E}_{d,q}^{\mathrm{odd}}$,

(ii) $\mathcal{E}_{d,q}^{(k)}$ for all $0 \leq k \leq (q-1)d$

are invariant under $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$. Furthermore,

(i) $\mathcal{E}_{d,q}^{\mathrm{even}} \oplus \mathcal{E}_{d,q}^{\mathrm{odd}} = \mathbb{R}^{q^d}$ and $\mathcal{E}_{d,q}^{\mathrm{even}} \perp \mathcal{E}_{d,q}^{\mathrm{odd}}$,

(ii) $\mathcal{E}_{d,q}^{(0)} \oplus \ldots \oplus \mathcal{E}_{d,q}^{((q-1)d)} = \mathbb{R}^{q^d}$ and $\mathcal{E}_{d,q}^{(k_1)} \perp \mathcal{E}_{d,q}^{(k_2)}$ for all $0 \leq k_1 < k_2 \leq (q-1)d$,

so Lemma 2.3 implies the claim. □

Mostly for reference we formulate the statement of Theorem 2.9 in the important case $\mu = 1$.

Corollary 2.10. *Let $\mathbf{v}$ be an eigenvector of a simple eigenvalue of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$.*

(i) Then the entries of $\mathbf{v}$ are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation whose sum of digits is even respectively odd.

(ii) If additionally $A = B$, then the entries of $\mathbf{v}$ are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation with a fixed sum of digits.

In the special cases where only one of A, B, Δ, h is nonzero, we determine the minimal eigenvalue and associated eigenvectors analytically.

Proposition 2.11. *(i)*

$$\mathbf{v}_{x,q,1} := \frac{1}{\sqrt{2^{q-1}}} \sum_{j=0}^{q-1} (-1)^j \sqrt{\binom{q-1}{j}} \, |j\rangle$$

is a normalized eigenvector of $\mathbf{S}_{x,q}$ associated with $\lambda_{\min} = \lambda_1 = -\frac{q-1}{2}$ and

$$\mathbf{v}_{x,q,q} := \frac{1}{\sqrt{2^{q-1}}} \sum_{j=0}^{q-1} \sqrt{\binom{q-1}{j}} \, |j\rangle$$

is a normalized eigenvector of $\mathbf{S}_{x,q}$ associated with $\lambda_{\max} = \lambda_q = \frac{q-1}{2}$.

(ii)

$$\mathbf{v}_{y,q,1} := \frac{1}{\sqrt{2^{q-1}}} \sum_{j=0}^{q-1} (-\mathrm{i})^j \sqrt{\binom{q-1}{j}} \, |j\rangle$$

is a normalized eigenvector of $\mathbf{S}_{y,q}$ *associated with* $\lambda_{\min} = \lambda_1 = -\frac{q-1}{2}$ *and*

$$\mathbf{v}_{y,q,q} := \frac{1}{\sqrt{2^{q-1}}} \sum_{j=0}^{q-1} \mathrm{i}^j \sqrt{\binom{q-1}{j}} \, |j\rangle$$

is a normalized eigenvector of $\mathbf{S}_{y,q}$ *associated with* $\lambda_{\max} = \lambda_q = \frac{q-1}{2}$.

(iii)

$$\mathbf{v}_{z,q,1} := |q-1\rangle$$

is a normalized eigenvector of $\mathbf{S}_{z,q}$ *associated with* $\lambda_{\min} = \lambda_1 = -\frac{q-1}{2}$ *and*

$$\mathbf{v}_{z,q,q} := |0\rangle$$

is a normalized eigenvector of $\mathbf{S}_{z,q}$ *associated with* $\lambda_{\max} = \lambda_q = \frac{q-1}{2}$.

(iv) For $\omega = x$ *resp.* $\omega = y$ *resp.* $\omega = z$ *consider the case* $B = \Delta = h = 0$ *resp.* $A = \Delta = h = 0$ *resp.* $A = B = h = 0$ *where*

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}} = A_\omega \sum_{i=1}^{d-1} \mathbf{S}_{\omega,q}^{(i)} \mathbf{S}_{\omega,q}^{(i+1)}$$

and assume $A_\omega := A \neq 0$ *resp.* $A_\omega := B \neq 0$ *resp.* $A_\omega := \Delta \neq 0$. *Then*

$$\lambda_{\min}(\mathbf{H}_{d,q}^{\mathrm{XYZ}}) = \begin{cases} -A_\omega (d-1) \frac{(q-1)^2}{4} & , \ A_\omega > 0 \\ A_\omega (d-1) \frac{(q-1)^2}{4} & , \ A_\omega < 0 \end{cases}$$

and the eigenspace associated with $\lambda_{\min}$ *is two-dimensional and equals*

$$V_{\min} = \begin{cases} \mathrm{span}\{\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d}, \\ \qquad \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,b_d}\} & , \ A_\omega > 0 \\ \mathrm{span}\{\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,1}, \\ \qquad \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,q}\} & , \ A_\omega < 0 \end{cases}$$

where $\mathbf{v}_{\omega,q,1}$ *and* $\mathbf{v}_{\omega,q,q}$ *are defined in part (i) resp. (ii) resp. (iii) and*

$$a_d = \begin{cases} q & , \ d \text{ even} \\ 1 & , \ d \text{ odd} \end{cases}, \qquad b_d = \begin{cases} 1 & , \ d \text{ even} \\ q & , \ d \text{ odd} \end{cases}.$$

Moreover, in case $\omega = y$, *the vectors*

$$\begin{gathered} \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,1} \otimes \cdots \otimes \mathbf{v}_{y,q,a_d} + \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,q} \otimes \cdots \otimes \mathbf{v}_{y,q,b_d}, \\ \mathrm{i}\big(\mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,1} \otimes \cdots \otimes \mathbf{v}_{y,q,a_d} - \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,q} \otimes \cdots \otimes \mathbf{v}_{y,q,b_d}\big), \\ \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,1} \otimes \cdots \otimes \mathbf{v}_{y,q,1} + \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,q} \otimes \cdots \otimes \mathbf{v}_{y,q,q}, \\ \mathrm{i}\big(\mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,1} \otimes \mathbf{v}_{y,q,1} \otimes \cdots \otimes \mathbf{v}_{y,q,1} - \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,q} \otimes \mathbf{v}_{y,q,q} \otimes \cdots \otimes \mathbf{v}_{y,q,q}\big) \end{gathered}$$

are real.

(v) *Consider the matrix* $\left(z_{m,n}^{(d,q)}\right)_{1\leq m\leq q^d,\,1\leq n\leq q^d} := \mathbf{Z}_{d,q} := \sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)} \in \mathbb{R}^{q^d\times q^d}$. $\mathbf{Z}_{d,q}$ *is a diagonal matrix and*

$$z_{m,m}^{(d,q)} = \frac{d(q-1)}{2} - k,$$

where k *is the sum of digits of the* q*-ary representation of* $m-1$ *with* $m \in \{1,\dots,q^d\}$. *In particular, for* $A = B = \Delta = 0$ *and* $h \neq 0$ *where*

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}} = h\mathbf{Z}_{d,q},$$

it is

$$\lambda_{\min}(\mathbf{H}_{d,q}^{\mathrm{XYZ}}) = \begin{cases} -hd\frac{q-1}{2} & ,\ h>0 \\ hd\frac{q-1}{2} & ,\ h<0 \end{cases}$$

and the eigenspace associated with $\lambda_{\min}$ *is*

$$V_{\min} = \begin{cases} \mathrm{span}\{|q-1\ \dots\ q-1\rangle\} & ,\ h>0 \\ \mathrm{span}\{|0\ \dots\ 0\rangle\} & ,\ h<0 \end{cases}.$$

Proof. (i) $\mathbf{v}_{x,q,1}$ respectively $\mathbf{v}_{x,q,q}$ is normalized as

$$\begin{aligned}
\mathbf{v}_{x,q,1}^{*}\mathbf{v}_{x,q,1} &= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}(-1)^k\sqrt{\binom{q-1}{k}}\langle k|\right)\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}(-1)^j\sqrt{\binom{q-1}{j}}|j\rangle\right) \\
&= \frac{1}{\sqrt{2^{q-1}}}\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}(-1)^j(-1)^j\sqrt{\binom{q-1}{j}\binom{q-1}{j}}\langle j|j\rangle \\
&= \frac{1}{2^{q-1}}\sum_{j=0}^{q-1}\binom{q-1}{j} = \frac{1}{2^{q-1}}\cdot 2^{q-1} = 1
\end{aligned}$$

respectively

$$\begin{aligned}
\mathbf{v}_{x,q,q}^{*}\mathbf{v}_{x,q,q} &= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}\sqrt{\binom{q-1}{k}}\langle k|\right)\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}\sqrt{\binom{q-1}{j}}|j\rangle\right) \\
&= \frac{1}{\sqrt{2^{q-1}}}\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}\sqrt{\binom{q-1}{j}\binom{q-1}{j}}\langle j|j\rangle \\
&= \frac{1}{2^{q-1}}\sum_{j=0}^{q-1}\binom{q-1}{j} = \frac{1}{2^{q-1}}\cdot 2^{q-1} = 1.
\end{aligned}$$

Furthermore,

$$
\begin{aligned}
&\mathbf{v}^*_{x,q,1}\mathbf{S}_{x,q}\mathbf{v}_{x,q,1} \\
&= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}(-1)^k\sqrt{\binom{q-1}{k}}\langle k|\right)\left(\sum_{j=1}^{q-1}c_j\left(|j-1\rangle\langle j| + |j\rangle\langle j-1|\right)\right) \\
&\qquad\qquad\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{l=0}^{q-1}(-1)^l\sqrt{\binom{q-1}{l}}|l\rangle\right) \\
&= \frac{1}{2^{q-1}}\sum_{j=1}^{q-1}c_j\left((-1)^{j-1}\sqrt{\binom{q-1}{j-1}}(-1)^j\sqrt{\binom{q-1}{j}} + (-1)^j\sqrt{\binom{q-1}{j}}(-1)^{j-1}\sqrt{\binom{q-1}{j-1}}\right) \\
&= \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}(-1)^{j-1}(-1)^{j-1}c_j\sqrt{\binom{q-1}{j-1}}\sqrt{\binom{q-1}{j}} \\
&= \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}c_j\sqrt{\binom{q-1}{j-1}\binom{q-1}{j}} = \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}c_j\sqrt{\frac{j}{q-j}\binom{q-1}{j}^2} \\
&= \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}\frac{1}{2}\sqrt{j(q-j)}\sqrt{\frac{j}{q-j}}\binom{q-1}{j} = \frac{-1}{2^{q-1}}\sum_{j=1}^{q-1}j\binom{q-1}{j} \\
&= \frac{-1}{2^{q-1}}\cdot(q-1)2^{q-2} = -\frac{q-1}{2}
\end{aligned}
$$

and in a very similar way

$$\mathbf{v}^*_{x,q,q}\mathbf{S}_{x,q}\mathbf{v}_{x,q,q} = \frac{q-1}{2}.$$

(ii) $\mathbf{v}_{y,q,1}$ respectively $\mathbf{v}_{y,q,q}$ is normalized as

$$
\begin{aligned}
\mathbf{v}^*_{y,q,1}\mathbf{v}_{y,q,1} &= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}\mathrm{i}^k\sqrt{\binom{q-1}{k}}\langle k|\right)\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}(-\mathrm{i})^j\sqrt{\binom{q-1}{j}}|j\rangle\right) \\
&= \frac{1}{\sqrt{2^{q-1}}}\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}\mathrm{i}^j(-\mathrm{i})^j\sqrt{\binom{q-1}{j}\binom{q-1}{j}}\langle j|j\rangle \\
&= \frac{1}{2^{q-1}}\sum_{j=0}^{q-1}\binom{q-1}{j} = \frac{1}{2^{q-1}}\cdot 2^{q-1} = 1
\end{aligned}
$$

respectively

$$
\begin{aligned}
\mathbf{v}^*_{y,q,q}\mathbf{v}_{y,q,q} &= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}(-\mathrm{i})^k\sqrt{\binom{q-1}{k}}\langle k|\right)\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}\mathrm{i}^j\sqrt{\binom{q-1}{j}}|j\rangle\right) \\
&= \frac{1}{\sqrt{2^{q-1}}}\frac{1}{\sqrt{2^{q-1}}}\sum_{j=0}^{q-1}(-\mathrm{i})^j\mathrm{i}^j\sqrt{\binom{q-1}{j}\binom{q-1}{j}}\langle j|j\rangle \\
&= \frac{1}{2^{q-1}}\sum_{j=0}^{q-1}\binom{q-1}{j} = \frac{1}{2^{q-1}}\cdot 2^{q-1} = 1.
\end{aligned}
$$

Furthermore,

$$\begin{aligned}
&\mathbf{v}_{y,q,1}^{*}\mathbf{S}_{y,q}\mathbf{v}_{y,q,1}\\
&= \left(\frac{1}{\sqrt{2^{q-1}}}\sum_{k=0}^{q-1}\mathrm{i}^{k}\sqrt{\binom{q-1}{k}}\langle k|\right)\left(-\mathrm{i}\sum_{j=1}^{q-1}c_j\left(|j-1\rangle\langle j|-|j\rangle\langle j-1|\right)\right)\\
&\qquad\qquad\left(\frac{1}{\sqrt{2^{q-1}}}\sum_{l=0}^{q-1}(-\mathrm{i})^{l}\sqrt{\binom{q-1}{l}}|l\rangle\right)\\
&= \frac{-\mathrm{i}}{2^{q-1}}\sum_{j=1}^{q-1}c_j\left(\mathrm{i}^{j-1}\sqrt{\binom{q-1}{j-1}}(-\mathrm{i})^{j}\sqrt{\binom{q-1}{j}}-\mathrm{i}^{j}\sqrt{\binom{q-1}{j}}(-\mathrm{i})^{j-1}\sqrt{\binom{q-1}{j-1}}\right)\\
&= \frac{-\mathrm{i}}{2^{q-1}}\sum_{j=1}^{q-1}(-1)^{j-1}c_j\left((-\mathrm{i})^{j-1}\sqrt{\binom{q-1}{j-1}}(-\mathrm{i})^{j}\sqrt{\binom{q-1}{j}}\right.\\
&\qquad\left.+(-\mathrm{i})^{j}\sqrt{\binom{q-1}{j}}(-\mathrm{i})^{j-1}\sqrt{\binom{q-1}{j-1}}\right)\\
&= \frac{-2\mathrm{i}}{2^{q-1}}\sum_{j=1}^{q-1}(-1)^{j-1}(-\mathrm{i})^{j-1}(-\mathrm{i})^{j}c_j\sqrt{\binom{q-1}{j-1}}\sqrt{\binom{q-1}{j}}\\
&= \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}c_j\sqrt{\binom{q-1}{j-1}\binom{q-1}{j}} = \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}c_j\sqrt{\frac{j}{q-j}\binom{q-1}{j}^{2}}\\
&= \frac{-2}{2^{q-1}}\sum_{j=1}^{q-1}\frac{1}{2}\sqrt{j(q-j)}\sqrt{\frac{j}{q-j}}\binom{q-1}{j} = \frac{-1}{2^{q-1}}\sum_{j=1}^{q-1}j\binom{q-1}{j}\\
&= \frac{-1}{2^{q-1}}\cdot(q-1)2^{q-2} = -\frac{q-1}{2}
\end{aligned}$$

and in a very similar way

$$\mathbf{v}_{y,q,q}^{*}\mathbf{S}_{y,q}\mathbf{v}_{y,q,q} = \frac{q-1}{2}.$$

(iii) $|q-1\rangle$ and $|0\rangle$ are unit vectors, thus normalized. Furthermore,

$$\mathbf{S}_{z,q}\mathbf{v}_{z,q,1} = \sum_{j=0}^{q-1}\frac{q-1-2j}{2}|j\rangle\langle j|q-1\rangle = \frac{q-1-2(q-1)}{2}|q-1\rangle = -\frac{q-1}{2}|q-1\rangle$$

and

$$\mathbf{S}_{z,q}\mathbf{v}_{z,q,q} = \sum_{j=0}^{q-1}\frac{q-1-2j}{2}|j\rangle\langle j|0\rangle = \frac{q-1-2\cdot 0}{2}|0\rangle = \frac{q-1}{2}|0\rangle.$$

(iv) Due to [HJ91, Thm. 4.2.12], the minimal eigenvalue of

$$\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)} = \mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q\otimes\mathbf{S}_{\omega,q}\otimes\mathbf{S}_{\omega,q}\otimes\mathbf{I}_q\otimes\cdots\otimes\mathbf{I}_q$$

equals $-\frac{(q-1)^2}{4}$ and the maximal eigenvalue of $\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}$ equals $\frac{(q-1)^2}{4}$. By [HJ13, Cor. 4.3.15], it is

$$\lambda_{\min}\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right) \geq -(d-1)\frac{(q-1)^2}{4} \tag{2.4}$$

and

$$\lambda_{\max}\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right) \leq (d-1)\frac{(q-1)^2}{4}. \tag{2.5}$$

Equality in (2.4) respectively (2.5) follows from

$$\begin{aligned}
&\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})\\
&=\sum_{i=1}^{d-1}\left(\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})\\
&=\sum_{i=1}^{d-1}-\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})\\
&=-(d-1)\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})
\end{aligned}$$

respectively

$$\begin{aligned}
&\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,1})\\
&=\sum_{i=1}^{d-1}\left(\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,1})\\
&=\sum_{i=1}^{d-1}\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,1})\\
&=(d-1)\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,1}).
\end{aligned}$$

Just as well it holds

$$\begin{aligned}
&\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d})\\
&=-(d-1)\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d})
\end{aligned}$$

and

$$\begin{aligned}
&\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)(\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,q})\\
&=(d-1)\frac{(q-1)^2}{4}(\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,q}).
\end{aligned}$$

The fact $\lambda_{\min}(\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}) = -\frac{(q-1)^2}{4}$ together with the demonstrated equality in (2.4) implies that every vector not being an eigenvector associated with $\lambda_{\min}(\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)})$ for some $1 \leq i \leq d-1$ is not an eigenvector associated with $\lambda_{\min}\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)$. As $\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}$ is a Kronecker product with all factors being Hermitian, $\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}$ is Hermitian as well and there exists an orthonormal basis of $\mathbb{C}^{q^d}$ consisting of eigenvectors each of which is a Kronecker product of eigenvectors of the respective factors of $\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}$ by [HJ91, Thm. 4.2.12]. In fact every normalized product vector $\mathbf{w} = \mathbf{w}_1 \otimes \cdots \otimes \mathbf{w}_d$, $\mathbf{w}_i \in \mathbb{C}^q$, being orthogonal to $\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d}$ and $\mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,b_d}$ has at least for one $1 \leq j \leq d-1$ two successive factors $\mathbf{w}_j \otimes \mathbf{w}_{j+1}$ not being equal to $\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q}$ or $\mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1}$. But then

$$\mathbf{w}^*(\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)})\mathbf{w} > -\frac{(q-1)^2}{4}$$

which yields

$$\mathbf{w}^*\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)\mathbf{w} > -(d-1)\frac{(q-1)^2}{4}.$$

Hence the eigenspace associated with $\lambda_{\min}\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)$ equals

$$\operatorname{span}\{\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d},\ \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,b_d}\}.$$

This space is two-dimensional since the two spanning vectors are orthogonal due to

$$\begin{aligned}&(\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d})^*(\mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,b_d})\\ &= \mathbf{v}_{\omega,q,1}^*\mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q}^*\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1}^*\mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d}^*\mathbf{v}_{\omega,q,b_d}\\ &= 0 \cdot \cdots \cdot 0 = 0\end{aligned}$$

as $\mathbf{v}_{\omega,q,1}$ and $\mathbf{v}_{\omega,q,q}$ are eigenvectors associated with different eigenvalues of a Hermitian matrix. An analogous argumentation shows that the eigenspace associated with $\lambda_{\max}\left(\sum_{i=1}^{d-1}\mathbf{S}_{\omega,q}^{(i)}\mathbf{S}_{\omega,q}^{(i+1)}\right)$ equals

$$\operatorname{span}\{\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,1},\ \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,q}\}$$

and is two-dimensional as well. Depending on the sign of A_ω, we obtain the stated form of the minimal eigenvalue and the associated eigenspace of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$. In case $\omega = y$, the eigenvectors

$$\begin{aligned}&\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,a_d}, \quad \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,b_d},\\ &\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,1}, \quad \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,q}\end{aligned}$$

of the real matrix $\sum_{i=1}^{d-1}\mathbf{S}_{y,q}^{(i)}\mathbf{S}_{y,q}^{(i+1)}$ are not real. However, with

$$(\underline{d}, \overline{d}) := \begin{cases}(d-1, d) & ,\ d \text{ even}\\ (d, d-1) & ,\ d \text{ odd}\end{cases},$$

we compute

$$\begin{aligned}
&\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d}\\
&=(2^{q-1})^{-d/2}\sum_{j_1=0}^{q-1}\sum_{j_2=0}^{q-1}\cdots\sum_{j_d=0}^{q-1}(-\mathrm{i})^{j_1}(+\mathrm{i})^{j_2}\cdots\big((-1)^d\mathrm{i}\big)^{j_d}\\
&\qquad\qquad\sqrt{\binom{q-1}{j_1}\binom{q-1}{j_2}\cdots\binom{q-1}{j_d}}\,|j_1\ j_2\ \ldots\ j_d\rangle\\
&=(2^{q-1})^{-d/2}\sum_{j_1=0}^{q-1}\sum_{j_2=0}^{q-1}\cdots\sum_{j_d=0}^{q-1}(-1)^{j_1+j_3+\ldots+j_{\underline{d}}}\,\mathrm{i}^{j_1+j_2+j_3+\ldots+j_d}\\
&\qquad\qquad\sqrt{\binom{q-1}{j_1}\binom{q-1}{j_2}\cdots\binom{q-1}{j_d}}\,|j_1\ j_2\ \ldots\ j_d\rangle
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d}\\
&=(2^{q-1})^{-d/2}\sum_{j_1=0}^{q-1}\sum_{j_2=0}^{q-1}\cdots\sum_{j_d=0}^{q-1}(+\mathrm{i})^{j_1}(-\mathrm{i})^{j_2}\cdots\big((-1)^{d-1}\mathrm{i}\big)^{j_d}\\
&\qquad\qquad\sqrt{\binom{q-1}{j_1}\binom{q-1}{j_2}\cdots\binom{q-1}{j_d}}\,|j_1\ j_2\ \ldots\ j_d\rangle\\
&=(2^{q-1})^{-d/2}\sum_{j_1=0}^{q-1}\sum_{j_2=0}^{q-1}\cdots\sum_{j_d=0}^{q-1}(-1)^{j_2+j_4+\ldots+j_{\overline{d}}}\,\mathrm{i}^{j_1+j_2+j_3+\ldots+j_d}\\
&\qquad\qquad\sqrt{\binom{q-1}{j_1}\binom{q-1}{j_2}\cdots\binom{q-1}{j_d}}\,|j_1\ j_2\ \ldots\ j_d\rangle\,.
\end{aligned}$$

So the entries of $\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d}$ and $\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d}$ are equal up to a prefactor ±1 or $\pm\mathrm{i}$. We get a sequence of equivalent statements

for the respective entry $(\cdot)_m$, $m\in\{1,\ldots,q^d\}$, where $(j_1\ \ldots\ j_d)_q+1=m$, it holds
$(\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})_m=(\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d})_m$,
$\Leftrightarrow\ (-1)^{j_1+j_3+\ldots+j_{\underline{d}}}=(-1)^{j_2+j_4+\ldots+j_{\overline{d}}}$,
$\Leftrightarrow$ the parity of $j_1+j_3+\ldots+j_{\underline{d}}$ and $j_2+j_4+\ldots+j_{\overline{d}}$ is equal,
$\Leftrightarrow\ j_1+j_2+j_3+j_4+\ldots+j_d$ is even,
$\Leftrightarrow\ \mathrm{i}^{j_1+j_2+\ldots+j_d}\in\mathbb{R}$,
$\Leftrightarrow\ (\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d})_m$ is real.

Hence, if at a certain position the entries of $\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\cdots\otimes\mathbf{v}_{\omega,q,a_d}$ and $\mathbf{v}_{\omega,q,q}\otimes\mathbf{v}_{\omega,q,1}\otimes\mathbf{v}_{\omega,q,q}\otimes\cdots\otimes\mathbf{v}_{\omega,q,b_d}$ are real, they are equal and if they are complex, they are conjugate. Thus

$$\begin{aligned}
&\mathbf{v}_{y,q,1}\otimes\mathbf{v}_{y,q,q}\otimes\mathbf{v}_{y,q,1}\otimes\cdots\otimes\mathbf{v}_{y,q,a_d}+\mathbf{v}_{y,q,q}\otimes\mathbf{v}_{y,q,1}\otimes\mathbf{v}_{y,q,q}\otimes\cdots\otimes\mathbf{v}_{y,q,b_d}\in\mathbb{R}^{q^d},\\
&\mathrm{i}\big(\mathbf{v}_{y,q,1}\otimes\mathbf{v}_{y,q,q}\otimes\mathbf{v}_{y,q,1}\otimes\cdots\otimes\mathbf{v}_{y,q,a_d}-\mathbf{v}_{y,q,q}\otimes\mathbf{v}_{y,q,1}\otimes\mathbf{v}_{y,q,q}\otimes\cdots\otimes\mathbf{v}_{y,q,b_d}\big)\in\mathbb{R}^{q^d}.
\end{aligned}$$

A similar line of reasoning may be carried out for $\mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \mathbf{v}_{\omega,q,1} \otimes \cdots \otimes \mathbf{v}_{\omega,q,1}$ and $\mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \mathbf{v}_{\omega,q,q} \otimes \cdots \otimes \mathbf{v}_{\omega,q,q}$.

(v) $\mathbf{Z}_{d,q}$ is a sum of Kronecker products of diagonal matrices and therefore diagonal itself. Let $(a_1 \ \ldots \ a_d)$ be the q-ary representation of $m-1$. Since

$$\mathbf{I}_q = \sum_{j=0}^{q-1} |j\rangle \langle j| ,$$

we have

$$\mathbf{S}_{z,q}^{(i)} = \sum_{j_1=0}^{q-1} \cdots \sum_{j_i=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} \frac{q-1-2j_i}{2} |j_1 \ \ldots \ j_i \ \ldots \ j_d\rangle \langle j_1 \ \ldots \ j_i \ \ldots \ j_d| ,$$

hence

$$\sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)} = \sum_{i=1}^{d} \sum_{j_1=0}^{q-1} \cdots \sum_{j_i=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} \frac{q-1-2j_i}{2} |j_1 \ \ldots \ j_i \ \ldots \ j_d\rangle \langle j_1 \ \ldots \ j_i \ \ldots \ j_d| .$$

In this multiple sum, a summand comprising $|a_1 \ \ldots \ a_d\rangle \langle a_1 \ \ldots \ a_d|$, which is a $q^d \times q^d$ matrix containing exactly one 1 at position (m,m) and 0 elsewhere, occurs d times, once for each $i \in \{1,\ldots,d\}$. The prefactor then takes the value $(q-1-2a_i)/2$ each time. The value of $z_{m,m}^{(d,q)}$ in turn equals the sum of all these prefactors of $|a_1 \ \ldots \ a_d\rangle \langle a_1 \ \ldots \ a_d|$. We obtain

$$z_{m,m}^{(d,q)} = \sum_{i=1}^{d} \frac{q-1-2a_i}{2} = \sum_{i=1}^{d} \frac{q-1}{2} - \sum_{i=1}^{d} a_i = \frac{d(q-1)}{2} - k.$$

Hence $z_{m,m}^{(d,q)}$ gets maximal if and only if $k=0$, thus if and only if $m-1 = (0 \ldots 0)_q = 0_{10}$. Conversely $z_{m,m}^{(d,q)}$ gets minimal if and only if $k = d(q-1)$, thus if and only if $m-1 = (q-1 \ldots q-1)_q = (q^d-1)_{10}$. We conclude the stated form of the minimal eigenvalue and associated eigenvectors of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$.

□

Now we are ready to prove a statement about the invariance of the set of eigenvectors of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ under a change of h in case of $A = B$.

Theorem 2.12. *Let $A = B$ and let for arbitrary $k \in \{0,\ldots,(q-1)d\}$ be $\mathbf{v} \in \mathcal{E}_{d,q}^{(k)}$ an eigenvector associated with an eigenvalue λ of $\mathbf{H}_{d,q,h}^{\mathrm{XYZ}}$, where the extra subscript emphasizes the dependence on the parameter h.*
Then $\mathbf{v}$ is also an eigenvector of $\mathbf{H}_{d,q,\widetilde{h}}^{\mathrm{XYZ}}$ for all other $\widetilde{h} \in \mathbb{R}$. Let $\widetilde{\lambda}$ be the associated eigenvalue. There is the relation

$$\widetilde{\lambda} - \lambda = \left(\frac{d(q-1)}{2} - k \right) (\widetilde{h} - h).$$

Proof. Without loss of generality assume $\|\mathbf{v}\|_2 = 1$ and define the diagonal matrix $\mathbf{Z}_{d,q} := \sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)}$. Consider the residual

$$\begin{aligned}
&\mathbf{H}_{d,q,\widetilde{h}}^{\mathrm{XYZ}}\mathbf{v} - \left\langle \mathbf{v}, \mathbf{H}_{d,q,\widetilde{h}}^{\mathrm{XYZ}}\mathbf{v} \right\rangle \mathbf{v} \\
&= \mathbf{H}_{d,q,h}^{\mathrm{XYZ}}\mathbf{v} + \widetilde{h}\mathbf{Z}_{d,q}\mathbf{v} - h\mathbf{Z}_{d,q}\mathbf{v} - \left\langle \mathbf{v}, \mathbf{H}_{d,q,h}^{\mathrm{XYZ}}\mathbf{v} \right\rangle \mathbf{v} - \left\langle \mathbf{v}, \widetilde{h}\mathbf{Z}_{d,q}\mathbf{v} \right\rangle \mathbf{v} + \langle \mathbf{v}, h\mathbf{Z}_{d,q}\mathbf{v} \rangle \mathbf{v} \\
&= (\widetilde{h} - h)\mathbf{Z}_{d,q}\mathbf{v} - \left\langle \mathbf{v}, (\widetilde{h} - h)\mathbf{Z}_{d,q}\mathbf{v} \right\rangle \mathbf{v} \\
&= (\widetilde{h} - h)(\mathbf{Z}_{d,q}\mathbf{v} - \langle \mathbf{v}, \mathbf{Z}_{d,q}\mathbf{v} \rangle \mathbf{v}).
\end{aligned}$$

To justify that the last expression equals $\mathbf{0} \in \mathbb{R}^{q^d}$, we examine the structure of the diagonal matrix $(z_{m,n}^{(d,q)})_{1 \le m \le q^d,\, 1 \le n \le q^d} := \mathbf{Z}_{d,q} = \sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)}$. By Proposition 2.11(v), it is

$$z_{m,m}^{(d,q)} = \frac{d(q-1)}{2} - k,$$

where k denotes the sum of digits in the q-ary representation of $m-1$. For fixed k, so for constant $z_{m,m}^{(d,q)}$, these m are exactly the positions where v_m might be nonzero which implies

$$\mathbf{Z}_{d,q}\mathbf{v} = \left(\frac{d(q-1)}{2} - k \right) \mathbf{v},$$

and hence the aforementioned residual vanishes. Additionally we have

$$\begin{aligned}
\widetilde{\lambda} - \lambda &= \left\langle \mathbf{v}, \mathbf{H}_{d,q,\widetilde{h}}^{\mathrm{XYZ}}\mathbf{v} \right\rangle - \left\langle \mathbf{v}, \mathbf{H}_{d,q,h}^{\mathrm{XYZ}}\mathbf{v} \right\rangle = \left\langle \mathbf{v}, \left(\mathbf{H}_{d,q,\widetilde{h}}^{\mathrm{XYZ}} - \mathbf{H}_{d,q,h}^{\mathrm{XYZ}} \right) \mathbf{v} \right\rangle = \left\langle \mathbf{v}, (\widetilde{h} - h)\mathbf{Z}_{d,q}\mathbf{v} \right\rangle \\
&= (\widetilde{h} - h) \left\langle \mathbf{v}, \left(\frac{d(q-1)}{2} - k \right) \mathbf{v} \right\rangle = \left(\frac{d(q-1)}{2} - k \right) (\widetilde{h} - h).
\end{aligned}$$

□

In the cases $A = B$ and $A = -B$, some unit vectors are eigenvectors.

Remark 2.13. (i) Let $A = B$. For $\mathbf{y}_1 := |0\ 0\ \ldots\ 0\rangle = \mathbf{e}_1$ it is

$$\begin{aligned}
\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_1 = \quad & Ac_1^2\,|1\ 1\ 0\ 0\ \ldots\ 0\ 0\rangle - Bc_1^2\,|1\ 1\ 0\ 0\ \ldots\ 0\ 0\rangle + \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + Ac_1^2\,|0\ 1\ 1\ 0\ \ldots\ 0\ 0\rangle - Bc_1^2\,|0\ 1\ 1\ 0\ \ldots\ 0\ 0\rangle + \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + \ldots \\
& + Ac_1^2\,|0\ 0\ 0\ 0\ \ldots\ 1\ 1\rangle - Bc_1^2\,|0\ 0\ 0\ 0\ \ldots\ 1\ 1\rangle + \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + h\frac{q-1}{2}\mathbf{y}_1 + h\frac{q-1}{2}\mathbf{y}_1 + \ldots + h\frac{q-1}{2}\mathbf{y}_1 + h\frac{q-1}{2}\mathbf{y}_1 \\
= & \left(\frac{(d-1)(q-1)^2}{4}\Delta + \frac{d(q-1)}{2}h \right) \mathbf{y}_1.
\end{aligned}$$

For $\mathbf{y}_2 := |q-1\ q-1\ \ldots\ q-1\rangle = \mathbf{e}_{q^d}$ we obtain in a similar way

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_2 = \left(\frac{(d-1)(q-1)^2}{4}\Delta - \frac{d(q-1)}{2}h \right) \mathbf{y}_2.$$

(ii) Let $A = -B$ and let d be even. Consider the two ket vectors $\mathbf{y}_1 := |0\ q-1\ 0\ q-1\ \ldots\ 0\ q-1\rangle$ and $\mathbf{y}_2 := |q-1\ 0\ q-1\ 0\ \ldots\ q-1\ 0\rangle$ with alternating 0's and $q-1$'s. Since

$$\begin{aligned}
\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_1 = \quad & Ac_1c_{q-1}\,|1\ q-2\ 0\ q-1\ \ldots\ 0\ q-1\rangle \\
& + Bc_1c_{q-1}\,|1\ q-2\ 0\ q-1\ \ldots\ 0\ q-1\rangle - \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + Ac_1c_{q-1}\,|0\ q-2\ 1\ q-1\ \ldots\ 0\ q-1\rangle \\
& + Bc_1c_{q-1}\,|0\ q-2\ 1\ q-1\ \ldots\ 0\ q-1\rangle - \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + \ldots \\
& + Ac_1c_{q-1}\,|0\ q-1\ 0\ q-1\ \ldots\ 1\ q-2\rangle \\
& + Bc_1c_{q-1}\,|0\ q-1\ 0\ q-1\ \ldots\ 1\ q-2\rangle - \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + h\frac{q-1}{2}\mathbf{y}_1 - h\frac{q-1}{2}\mathbf{y}_1 + \ldots + h\frac{q-1}{2}\mathbf{y}_1 - h\frac{q-1}{2}\mathbf{y}_1 \\
= & -\frac{(d-1)(q-1)^2}{4}\Delta\mathbf{y}_1
\end{aligned}$$

and analogously

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_2 = -\frac{(d-1)(q-1)^2}{4}\Delta\mathbf{y}_2,$$

there are two eigenvectors associated with the eigenvalue $\lambda = -\frac{(d-1)(q-1)^2}{4}\Delta$ which have product structure and are in fact unit vectors.

(iii) Let $A = -B$ and let now d be odd. Consider again the two ket vectors $\mathbf{y}_1 := |0\ q-1\ 0\ q-1\ \ldots\ q-1\ 0\rangle$ and $\mathbf{y}_2 := |q-1\ 0\ q-1\ 0\ \ldots\ 0\ q-1\rangle$ with alternating 0's and $q-1$'s. As d is odd, the number of the digits 0 and $q-1$ is not equal in $\mathbf{y}_1$ respectively in $\mathbf{y}_2$. We have

$$\begin{aligned}
\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_1 = \quad & Ac_1c_{q-1}\,|1\ q-2\ 0\ q-1\ \ldots\ q-1\ 0\rangle \\
& + Bc_1c_{q-1}\,|1\ q-2\ 0\ q-1\ \ldots\ q-1\ 0\rangle - \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + \ldots \\
& + Ac_1c_{q-1}\,|0\ q-1\ 0\ q-1\ \ldots\ q-2\ 1\rangle \\
& + Bc_1c_{q-1}\,|0\ q-1\ 0\ q-1\ \ldots\ q-2\ 1\rangle - \Delta\frac{(q-1)^2}{4}\mathbf{y}_1 \\
& + h\frac{q-1}{2}\mathbf{y}_1 - h\frac{q-1}{2}\mathbf{y}_1 + \ldots + h\frac{q-1}{2}\mathbf{y}_1 - h\frac{q-1}{2}\mathbf{y}_1 + h\frac{q-1}{2}\mathbf{y}_1 \\
= & \left(-\frac{(d-1)(q-1)^2}{4}\Delta + \frac{q-1}{2}h\right)\mathbf{y}_1
\end{aligned}$$

and likewise

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}}\mathbf{y}_2 = \left(-\frac{(d-1)(q-1)^2}{4}\Delta - \frac{q-1}{2}h\right)\mathbf{y}_2.$$

The statements obtained in the present section are used in Chapter 5 to set up a construction scheme for an initial guess to be employed in an iterative numerical method which computes an eigenvector associated with the minimal eigenvalue of a Hamilton operator. Especially in Subsections 5.1.1 and 5.1.2, the XYZ model is considered.

2.2. Potts model

Besides the XYZ model, we consider a second class of Hamilton operators, cf. [SP81, Eq. (2.10)], dating back to [Pot52].

Definition 2.14. Let $d, q \in \mathbb{N}$ with $q \geq 2$ and let

$$\mathbf{\Gamma}_q := \begin{pmatrix} \mathbf{0}_{q-1} & \mathbf{I}_{q-1} \\ 1 & \mathbf{0}_{q-1}^\top \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix} = \sum_{j=0}^{q-1} |j-1\rangle \langle j| ,$$

with the convention that the entries in $|\cdot\rangle$ respectively $\langle\cdot|$ have to be regarded *modulo* q. Let $\omega := \mathrm{e}^{2\pi \mathrm{i}/q}$ be a q-th root of unity and let further

$$\mathbf{\Omega}_q := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \omega & 0 & & \vdots \\ 0 & 0 & \omega^2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \omega^{q-1} \end{pmatrix}.$$

For $A, h \in \mathbb{R}$, $A \geq 0$, the Hamilton operator $\mathbf{H}_{d,q}^{\mathrm{Potts}} \in \mathbb{R}^{q^d \times q^d}$ of the *q-Potts model* for a system of d particles is defined via

$$\mathbf{H}_{d,q}^{\mathrm{Potts}} := -A \sum_{i=1}^{d-1} \sum_{m=1}^{q-1} (\mathbf{\Gamma}_q^m)^{(i)} (\mathbf{\Gamma}_q^{q-m})^{(i+1)} - h \sum_{i=1}^{d} \sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)}, \tag{2.6}$$

where the notation

$$(\mathbf{\Gamma}_q^m)^{(i)} (\mathbf{\Gamma}_q^{q-m})^{(i+1)} := \mathbf{I}_q \otimes \dots \otimes \mathbf{I}_q \otimes \mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{I}_q \otimes \dots \otimes \mathbf{I}_q$$

means that the matrix $\mathbf{\Gamma}_q^m$ is the i-th factor and $\mathbf{\Gamma}_q^{q-m}$ is the $(i+1)$-th factor in the Kronecker product, analogously for

$$(\mathbf{\Omega}_q^m)^{(i)} := \mathbf{I}_q \otimes \dots \otimes \mathbf{I}_q \otimes \mathbf{\Omega}_q^m \otimes \mathbf{I}_q \otimes \dots \otimes \mathbf{I}_q.$$

It holds

$$\sum_{m=1}^{q-1} \mathbf{\Omega}_q^m = \begin{pmatrix} q-1 & 0 & \dots & 0 \\ 0 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 \end{pmatrix} \tag{2.7}$$

since for each $l \in \{1, \dots, q-1\}$ we have

$$\sum_{j=1}^{q-1} \left(\mathrm{e}^{2\pi \mathrm{i} l/q}\right)^j = \frac{1 - \left(\mathrm{e}^{2\pi \mathrm{i} l/q}\right)^q}{1 - \mathrm{e}^{2\pi \mathrm{i} l/q}} - 1 = -1.$$

As additionally the first inner sum in (2.6) for each $i \in \{1, \dots, d-1\}$ may be written

$$\begin{aligned}&\sum_{m=1}^{q-1} (\mathbf{\Gamma}_q^m)^{(i)} (\mathbf{\Gamma}_q^{q-m})^{(i+1)} \\ &= \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q \otimes \left(\left(\sum_{m=1}^{\lfloor \frac{q-1}{2} \rfloor} (\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m) \right) + \mathbf{G} \right) \otimes \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q,\end{aligned}$$

where

$$\mathbf{G} := \begin{cases} \mathbf{\Gamma}_q^{q/2} \otimes \mathbf{\Gamma}_q^{q/2} & , \ q \text{ even} \\ \mathbf{0}_{q^2 \times q^2} & , \ q \text{ odd} \end{cases},$$

and $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m$, $m \in \{1, \dots, \lfloor \frac{q-1}{2} \rfloor\}$, is symmetric due to $(\mathbf{\Gamma}_q^m)^\top = \mathbf{\Gamma}_q^{q-m}$ and moreover in case of even q,

$$\begin{aligned}\mathbf{\Gamma}_q^{q/2} = \sum_{j=0}^{q-1} |j - q/2\rangle \langle j| &= \sum_{j=0}^{q/2-1} (|j - q/2\rangle \langle j| + |j\rangle \langle j + q/2|) \\ &= \sum_{j=0}^{q/2-1} (|j + q/2\rangle \langle j| + |j\rangle \langle j + q/2|)\end{aligned}$$

is symmetric, we obtain that $\mathbf{H}_{d,q}^{\text{Potts}} \in \mathbb{R}^{q^d \times q^d}$ is symmetric as well.

Note that for $q = 2$ it is $\mathbf{\Gamma}_2 = \boldsymbol{\sigma}_x$ and $\mathbf{\Omega}_2 = \boldsymbol{\sigma}_z$, hence $\mathbf{H}_{d,2}^{\text{Potts}}$ equals the Hamilton operator of the Ising model with parameters $-4A$ and $-2h$. So the smallest case where the q-Potts model is different from the q-XYZ model with $B = \Delta = 0$ is $q = 3$ which is in turn that value of q we will focus on mostly. When there is no need to refer to it, we omit the prefix q- from the name of the model.

Like Theorem 2.9, we prove a certain structure of the eigenvectors of $\mathbf{H}_{d,q}^{\text{Potts}}$. The following concept turns out to be reasonable.

Definition 2.15. For $d, q \in \mathbb{N}$ with $q \geq 2$, let $\mathcal{E}_{d,q}^{[k]} \subset \mathbb{R}^{q^d}$, $k \in \{0, \dots, q-1\}$, denote the set of vectors whose entries are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation whose sum of digits is congruent to k modulo q.

Recalling Definition 2.6, it is $\mathcal{E}_{d,2}^{[0]} = \mathcal{E}_{d,2}^{\text{even}}$ and $\mathcal{E}_{d,2}^{[1]} = \mathcal{E}_{d,2}^{\text{odd}}$. For some small $q \geq 3$ we illustrate the meaning of Definition 2.15 in Example B.2.

Lemma 2.16. *$\mathcal{E}_{d,q}^{[k]}$ is invariant under $\mathbf{H}_{d,q}^{\text{Potts}}$ for all $k \in \{0, \dots, q-1\}$.*

Proof. It is $\mathcal{E}_{d,q}^{[k]} = \text{span}\{\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_{q^{d-1}}}\}$, where $i_1 - 1, \dots, i_{q^{d-1}} - 1$ are the q^{d-1} numbers between 0 and $q^d - 1$ which have a q-ary representation whose sum of digits is k modulo q. Observe that the involved $\mathbf{e}_i$ are exactly the vectors $|a_1 \ \dots \ a_d\rangle$, $a_j \in \{0, \dots, q-1\}$, such that the sum of the a_j in $|\cdot\rangle$ is k modulo q. Therefore we have to study the effect of applying $\mathbf{H}_{d,q}^{\text{Potts}}$ to such a vector $\mathbf{y} := |a_1 \ \dots \ a_d\rangle$. So, let $\mathbf{w} \in \mathcal{E}_{d,q}^{[k]}$ be an arbitrary linear combination of the $\mathbf{e}_i \in \mathcal{E}_{d,q}^{[k]}$. Keeping in mind the convention that the entries in $|\cdot\rangle$ and $\langle\cdot|$ have to be regarded modulo q, it is

$$\mathbf{\Gamma}_q^m = \sum_{j=0}^{q-1} |j - m\rangle \langle j| .$$

This yields

$$\begin{aligned}
\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} &= \sum_{j=0}^{q-1} |j-m\rangle\langle j| \otimes \sum_{j=0}^{q-1} |j-(q-m)\rangle\langle j| \\
&= \sum_{j=0}^{q-1} |j-m\rangle\langle j| \otimes |0-(q-m)\rangle\langle 0| + \sum_{j=0}^{q-1} |j-m\rangle\langle j| \otimes |1-(q-m)\rangle\langle 1| \\
&\quad + \ldots + \\
&\quad \sum_{j=0}^{q-1} |j-m\rangle\langle j| \otimes |(q-1)-(q-m)\rangle\langle q-1| \\
&= \sum_{j=0}^{q-1} |j-m \quad -(q-m)\rangle\langle j\ 0| + \sum_{j=0}^{q-1} |j-m\ \ 1-(q-m)\rangle\langle j\ 1| \\
&\quad + \ldots + \\
&\quad \sum_{j=0}^{q-1} |j-m\ \ (q-1)-(q-m)\rangle\langle j\ q-1| \\
&= \sum_{l=0}^{q-1}\sum_{j=0}^{q-1} |j-m\ \ l-(q-m)\rangle\langle j\ l|\,.
\end{aligned}$$

For the most inner summands $|\alpha_1\ \alpha_2\rangle\langle\beta_1\ \beta_2| = |j-m\ \ l-(q-m)\rangle\langle j\ l|$, with fixed j and l and arbitrary m it holds

$$(\alpha_1+\alpha_2)-(\beta_1+\beta_2) = j-m+l-(q-m)-(j+l) = -q \equiv 0 \bmod q.$$

By (2.7) it is additionally

$$\sum_{m=1}^{q-1} \mathbf{\Omega}_q^m = (q-1)\,|0\rangle\langle 0| - \sum_{j=1}^{q-1} |j\rangle\langle j|\,.$$

Hence, as

$$|\ldots\ \alpha_i\ \alpha_{i+1}\ \ldots\rangle\langle\ldots\ \beta_i\ \beta_{i+1}\ \ldots|\ldots\ a_i\ a_{i+1}\ \ldots\rangle = \delta_{\beta_i,a_i}\delta_{\beta_{i+1},a_{i+1}}\,|\ldots\ \alpha_i\ \alpha_{i+1}\ \ldots\rangle\,,$$

the application of one summand

$$-A\left(\mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q \otimes (\mathbf{\Gamma}_q \otimes \mathbf{\Gamma}_q^{q-1} + \ldots + \mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \ldots + \mathbf{\Gamma}_q^{q-1} \otimes \mathbf{\Gamma}_q) \otimes \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q\right)$$

or

$$-h\left(\mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q \otimes (\mathbf{\Omega}_q + \ldots + \mathbf{\Omega}_q^m + \ldots + \mathbf{\Omega}_q^{q-1}) \otimes \mathbf{I}_q \otimes \cdots \otimes \mathbf{I}_q\right)$$

of $\mathbf{H}_{d,q}^{\text{Potts}}$, and by linearity also the application of $\mathbf{H}_{d,q}^{\text{Potts}}$ itself, does not change the value modulo q of the sum of the entries in $\mathbf{y} = |a_1\ \ldots\ a_d\rangle$. This particularly holds true for each of the basis vectors $\mathbf{e}_i$ contributing to $\mathbf{w}$. So, the space $\mathcal{E}_{d,q}^{[k]}$ for $k \in \{0,\ldots,q-1\}$ fixed is invariant under $\mathbf{H}_{d,q}^{\text{Potts}}$. □

The next two statements are the counterparts of Theorem 2.9 and Corollary 2.10 for the Potts model.

Theorem 2.17. *Let λ be an eigenvalue of $\mathbf{H}_{d,q}^{\text{Potts}}$ with multiplicity μ and let V be the associated eigenspace.*
Then there exists an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_\mu\} \subset \mathbb{R}^{q^d}$ of V such that for each $1 \leq \alpha \leq \mu$ there exists $0 \leq k(\alpha) \leq q-1$ with $\mathbf{v}_\alpha \in \mathcal{E}_{d,q}^{[k(\alpha)]}$.

Proof. By Lemma 2.16, the spaces $\mathcal{E}_{d,q}^{[0]}, \dots, \mathcal{E}_{d,q}^{[q-1]}$ are invariant under $\mathbf{H}_{d,q}^{\text{Potts}}$. Furthermore, $\mathcal{E}_{d,q}^{[0]} \oplus \dots \oplus \mathcal{E}_{d,q}^{[q-1]} = \mathbb{R}^{q^d}$ and $\mathcal{E}_{d,q}^{[k_1]} \perp \mathcal{E}_{d,q}^{[k_2]}$ for all $0 \leq k_1 < k_2 \leq q-1$, so Lemma 2.3 implies the claim. □

Corollary 2.18. *Let $\mathbf{v}$ be an eigenvector of a simple eigenvalue of $\mathbf{H}_{d,q}^{\text{Potts}}$. Then the entries of $\mathbf{v}$ are nonzero at most at those positions (beginning counting with 0) which have a q-ary representation whose sum of digits has a fixed value modulo q.*

In the two special cases $A = 0$ and $h = 0$ we can determine the minimal eigenvalue and associated eigenvectors analytically.

Proposition 2.19. *(i) Consider the case $A = 0$ where*

$$\mathbf{H}_{d,q}^{\text{Potts}} = -h \sum_{i=1}^{d} \sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)}$$

and assume $h \neq 0$. The matrix

$$\left(\omega_{k,l}^{(d,q)}\right)_{1 \leq k \leq q^d,\, 1 \leq l \leq q^d} := \sum_{i=1}^{d} \sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)} \in \mathbb{R}^{q^d \times q^d}$$

is diagonal and

$$\omega_{k,k}^{(d,q)} = (q-1) \cdot |\{1 \leq i \leq d \colon a_i = 0\}| - |\{1 \leq i \leq d \colon a_i \neq 0\}|,$$

where $(a_1 \dots a_d)$ is the q-ary representation of $k-1$ with $k \in \{1, \dots, q^d\}$. In particular this implies

$$\lambda_{\min}(\mathbf{H}_{d,q}^{\text{Potts}}) = \begin{cases} -hd(q-1) & , \ h > 0 \\ hd & , \ h < 0 \end{cases}$$

and

$$\mathbf{v}_{\min} \begin{cases} = \mathbf{e}_1 & , \ h > 0 \\ \in \operatorname{span}\{\mathbf{e}_k \colon (k-1)_{10} = (a_1 \dots a_d)_q \text{ and } a_i \neq 0 \text{ for all } 1 \leq i \leq d\} & , \ h < 0 \end{cases}.$$

(ii) Consider the case $h = 0$ where

$$\mathbf{H}_{d,q}^{\text{Potts}} = -A \sum_{i=1}^{d-1} \sum_{m=1}^{q-1} (\mathbf{\Gamma}_q^m)^{(i)} (\mathbf{\Gamma}_q^{q-m})^{(i+1)}$$

and assume $A > 0$. Then

$$\lambda_{\min}(\mathbf{H}_{d,q}^{\text{Potts}}) = -A(d-1)(q-1)$$

and the vectors

$$\mathbf{1}_{d,q}^{[0]} \in \mathcal{E}_{d,q}^{[0]}, \dots, \mathbf{1}_{d,q}^{[q-1]} \in \mathcal{E}_{d,q}^{[q-1]}$$

containing as each potential nonzero entry a 1 are associated eigenvectors.

Proof. (i) The matrix $\left(\omega_{k,l}^{(d,q)}\right)_{k,l}$ is a sum of Kronecker products of diagonal matrices and therefore diagonal itself. Let $(a_1 \ \dots \ a_d)$ be the q-ary representation of $k-1$. Since

$$\mathbf{I}_q = \sum_{j=0}^{q-1} |j\rangle\langle j| \quad \text{and} \quad \sum_{m=1}^{q-1} \mathbf{\Omega}_q^m = (q-1)\,|0\rangle\langle 0| - \sum_{j=1}^{q-1} |j\rangle\langle j|$$

by (2.7), we have

$$\begin{aligned}
\sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)} &= \sum_{j_1=0}^{q-1} \cdots \sum_{j_{i-1}=0}^{q-1} |j_1 \ \dots \ j_{i-1}\rangle\langle j_1 \ \dots \ j_{i-1}| \otimes \left((q-1)\,|0\rangle\langle 0| - \sum_{j_i=1}^{q-1} |j_i\rangle\langle j_i| \right) \\
&\quad \otimes \sum_{j_{i+1}=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} |j_{i+1} \ \dots \ j_d\rangle\langle j_{i+1} \ \dots \ j_d| \\
&= (q-1) \sum_{j_1=0}^{q-1} \cdots \sum_{j_{i-1}=0}^{q-1} \sum_{j_{i+1}=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} |j_1 \ \dots \ j_{i-1} \ 0 \ j_{i+1} \ \dots \ j_d\rangle \\
&\qquad\qquad \langle j_1 \ \dots \ j_{i-1} \ 0 \ j_{i+1} \ \dots \ j_d| \\
&\quad - \sum_{j_1=0}^{q-1} \cdots \sum_{j_{i-1}=0}^{q-1} \sum_{j_i=1}^{q-1} \sum_{j_{i+1}=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} |j_1 \ \dots \ j_d\rangle\langle j_1 \ \dots \ j_d| \,,
\end{aligned}$$

hence

$$\begin{aligned}
\sum_{i=1}^{d} \sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)} = \sum_{i=1}^{d} \Bigg(& (q-1) \sum_{j_1=0}^{q-1} \cdots \sum_{j_{i-1}=0}^{q-1} \sum_{j_{i+1}=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} |j_1 \ \dots \ j_{i-1} \ 0 \ j_{i+1} \ \dots \ j_d\rangle \\
& \qquad\qquad \langle j_1 \ \dots \ j_{i-1} \ 0 \ j_{i+1} \ \dots \ j_d| \\
& - \sum_{j_1=0}^{q-1} \cdots \sum_{j_{i-1}=0}^{q-1} \sum_{j_i=1}^{q-1} \sum_{j_{i+1}=0}^{q-1} \cdots \sum_{j_d=0}^{q-1} |j_1 \ \dots \ j_d\rangle\langle j_1 \ \dots \ j_d| \Bigg) .
\end{aligned}$$

In this multiple sum, a summand comprising $|a_1 \ \dots \ a_d\rangle\langle a_1 \ \dots \ a_d|$, which is a $q^d \times q^d$ matrix containing exactly one 1 at position (k,k) and 0 elsewhere, occurs d times, once for each $i \in \{1, \dots, d\}$. The prefactor each time then takes the value $q-1$ if $a_i = 0$ and -1 if $a_i \neq 0$. The value of $\omega_{k,k}^{(d,q)}$ in turn equals the sum of all these prefactors of $|a_1 \ \dots \ a_d\rangle\langle a_1 \ \dots \ a_d|$. We obtain

$$\omega_{k,k}^{(d,q)} = (q-1) \cdot |\{1 \le i \le d \colon a_i = 0\}| - |\{1 \le i \le d \colon a_i \neq 0\}| \,.$$

Hence $\omega_{k,k}^{(d,q)}$ gets maximal if and only if all $a_i = 0$, thus if and only if $k-1 = (0 \ \dots \ 0)_q = 0_{10}$. Conversely $\omega_{k,k}^{(d,q)}$ gets minimal if and only if all $a_i \neq 0$, which is the case for $(q-1)^d$ different values of $k-1$, e.g. if $k-1 = (q-1 \ \dots \ q-1)_q = (q^d-1)_{10}$. We conclude the stated form of the minimal eigenvalue and eigenvectors of $\mathbf{H}_{d,q}^{\text{Potts}}$.

(ii) The matrix $\mathbf{\Gamma}_q$ has simple eigenvalues $1, \omega, \omega^2, \dots, \omega^{q-1}$, where $\omega := \mathrm{e}^{2\pi\mathrm{i}/q}$ is a q-th root of unity. Hence for each $1 \le m \le q-1$, the matrix $\mathbf{\Gamma}_q^m$ has eigenvalues $1, \omega^m, \omega^{2m}, \dots, \omega^{(q-1)m}$ out of which $t := \gcd(m,q)$ due to the cyclic behavior of powers of ω are respectively equal and therefore each eigenvalue has multiplicity t. Likewise,

$\mathbf{\Gamma}_q^{q-m}$ has eigenvalues $1, \omega^{q-m}, \omega^{2(q-m)}, \ldots, \omega^{(q-1)(q-m)}$ each with multiplicity $\gcd(q-m,q) = \gcd(m,q) = t$. It is

$$\omega^{r(q-m)} = \omega^{rq}\omega^{-rm} = \omega^{-rm} = \omega^{qm}\omega^{-rm} = \omega^{(q-r)m}, \quad 1 \leq r \leq q-1,$$

so $\mathbf{\Gamma}_q^m$ and $\mathbf{\Gamma}_q^{q-m}$ have the same eigenvalues. An eigenvector associated with $\lambda = 1$ is $\mathbf{1}_{1,q} := (1, \ldots, 1)^\top \in \mathbb{R}^q$. By [HJ91, Thm. 4.2.12], the eigenvalues of $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m}$ are

$$\omega^{rm}\omega^{sm} = \omega^{(r+s)m}, \quad 0 \leq r, s \leq q-1,$$

which contains again $\lambda = 1$. Furthermore, $\mathbf{1}_{2,q} := \mathbf{1}_{1,q} \otimes \mathbf{1}_{1,q} = (1, \ldots, 1)^\top \in \mathbb{R}^{q^2}$ is an eigenvector of $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m}$ associated with $\lambda = 1$. Following [HJ91, exercise to Thm. 4.2.12], the multiplicity of $\lambda = 1$ equals ct^2, where c is the number of ways to write 1 as the product of distinct eigenvalues of $\mathbf{\Gamma}_q^m$ and $\mathbf{\Gamma}_q^{q-m}$ which equals in turn the number of ways that $(r+s)m \equiv 0 \mod q$ with $0 \leq r, s \leq \frac{q}{t} - 1$. Since $\frac{q}{t} \cdot m = \operatorname{lcm}(m,q)$, those $(r+s)m$ which are congruent to 0 modulo q are those with $r + s \in \{0, \frac{q}{t}\}$. This situation occurs for

$$r = s = 0, \quad r = 1 \wedge s = \frac{q}{t} - 1, \quad r = 2 \wedge s = \frac{q}{t} - 2, \quad \ldots, \quad r = \frac{q}{t} - 1 \wedge s = 1,$$

which implies $c = \frac{q}{t}$ and therefore $ct^2 = qt$. Consider the t vectors $\mathbf{w}_l \in \mathbb{C}^q$, $0 \leq l \leq t-1$, having only $\frac{q}{t}$ nonzero entries

$$(\mathbf{w}_l)_u := \omega^{nrm}$$

at the evenly distributed positions u, beginning counting these positions with 0, where $l + nm \equiv u \mod q$ for $0 \leq n \leq \frac{q}{t} - 1$. Due to their sparsity pattern, $\mathbf{w}_0, \ldots, \mathbf{w}_{t-1}$ are linearly independent eigenvectors of $\mathbf{\Gamma}_q^m$ associated with $\lambda = \omega^{rm}$ since an application of $\mathbf{\Gamma}_q^m$ shifts modulo q by m positions upwards and so elementwise

$$\mathbf{\Gamma}_q^m \colon \omega^{nrm} \mapsto \omega^{(n+1)rm} = \omega^{rm}\omega^{nrm} = \lambda\omega^{nrm}.$$

Thus $\mathbf{\Gamma}_q^m$ is also diagonalizable in case $t > 1$. Analogously, $\mathbf{\Gamma}_q^{q-m}$ is diagonalizable and by [HJ91, Probl. 4.3.15], $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m}$ is diagonalizable, too. As $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m}$ and $\mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m$ commute, they are simultaneously diagonalizable by [HJ13, Thm. 1.3.12]. So there exists $\mathbf{V} \in \mathbb{C}^{q^2 \times q^2}$ and diagonal $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \in \mathbb{C}^{q^2 \times q^2}$ with

$$\mathbf{V}^{-1}(\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m})\mathbf{V} = \mathbf{\Lambda}_1 \quad \text{and} \quad \mathbf{V}^{-1}(\mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m)\mathbf{V} = \mathbf{\Lambda}_2,$$

hence

$$\mathbf{V}^{-1}(\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m)\mathbf{V} = \mathbf{\Lambda}_1 + \mathbf{\Lambda}_2$$

contains on the diagonal a sum of two complex numbers each of absolute value ≤ 1 and, since $\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m$ is symmetric, these elements of $\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2$ are real. Thus $\lambda_{\max}(\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m) \leq 2$ and due to

$$(\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m)\mathbf{1}_{2,q} = (\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m})\mathbf{1}_{2,q} + (\mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m)\mathbf{1}_{2,q} = \mathbf{1}_{2,q} + \mathbf{1}_{2,q},$$

we obtain $\lambda_{\max}(\mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m} + \mathbf{\Gamma}_q^{q-m} \otimes \mathbf{\Gamma}_q^m) = 2$. In case of even q, the sum $\sum_{m=1}^{q-1} \mathbf{\Gamma}_q^m \otimes \mathbf{\Gamma}_q^{q-m}$ comprises the symmetric summand $\mathbf{\Gamma}_q^{q/2} \otimes \mathbf{\Gamma}_q^{q/2}$ with maximal eigenvalue $\lambda = 1$ and an associated eigenvector $\mathbf{1}_{2,q}$. Setting

$$\mathbf{G} := \begin{cases} \mathbf{\Gamma}_q^{q/2} \otimes \mathbf{\Gamma}_q^{q/2} & , \ q \text{ even} \\ \mathbf{0}_{q^2 \times q^2} & , \ q \text{ odd} \end{cases},$$

it is

$$\begin{aligned}\lambda_{\max}\left(\sum_{m=1}^{q-1}\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}\right) &= \lambda_{\max}\left(\sum_{m=1}^{\lfloor\frac{q-1}{2}\rfloor}(\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}+\mathbf{\Gamma}_q^{q-m}\otimes\mathbf{\Gamma}_q^m)+\mathbf{G}\right)\\ &\leq \lambda_{\max}(\mathbf{G})+\sum_{m=1}^{\lfloor\frac{q-1}{2}\rfloor}\lambda_{\max}(\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}+\mathbf{\Gamma}_q^{q-m}\otimes\mathbf{\Gamma}_q^m)\\ &= 1-\delta_{\lfloor\frac{q-1}{2}\rfloor,\frac{q-1}{2}}+2\cdot\left\lfloor\frac{q-1}{2}\right\rfloor = q-1\end{aligned}$$

by [HJ13, Cor. 4.3.15] and from

$$\left(\sum_{m=1}^{q-1}\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}\right)\mathbf{1}_{2,q}=\sum_{m=1}^{q-1}\left(\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}\right)\mathbf{1}_{2,q}=(q-1)\mathbf{1}_{2,q}$$

we conclude

$$\lambda_{\max}\left(\sum_{m=1}^{q-1}\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}\right)=q-1.$$

Along the same lines we show

$$\begin{aligned}&\lambda_{\max}\left(\sum_{i=1}^{d-1}\sum_{m=1}^{q-1}(\mathbf{\Gamma}_q^m)^{(i)}(\mathbf{\Gamma}_q^{q-m})^{(i+1)}\right)\\ &=\lambda_{\max}\left(\sum_{i=1}^{d-1}\left(\mathbf{I}_q^{\otimes(i-1)}\otimes\Big(\sum_{m=1}^{q-1}\mathbf{\Gamma}_q^m\otimes\mathbf{\Gamma}_q^{q-m}\Big)\otimes\mathbf{I}_q^{\otimes(d-i-1)}\right)\right)=(d-1)(q-1)\end{aligned}$$

and an associated eigenvector is $\mathbf{1}_{d,q}:=\mathbf{1}_{1,q}^{\otimes d}=(1,\ldots,1)^\top\in\mathbb{R}^{q^d}$. Remembering Remark 2.5(i), then also the projected vectors

$$\mathbf{1}_{d,q}^{[0]}:=P_{\mathcal{E}_{d,q}^{[0]}}(\mathbf{1}_{d,q}),\ \ldots\ ,\mathbf{1}_{d,q}^{[q-1]}:=P_{\mathcal{E}_{d,q}^{[q-1]}}(\mathbf{1}_{d,q})$$

containing as each potential nonzero entry a 1 are eigenvectors associated with $\lambda_{\max}=(d-1)(q-1)$. Just as well, $\mathbf{1}_{d,q}^{[0]},\ \ldots\ ,\mathbf{1}_{d,q}^{[q-1]}$ are eigenvectors of $\mathbf{H}_{d,q}^{\text{Potts}}$ associated with

$$\begin{aligned}\lambda_{\min}(\mathbf{H}_{d,q}^{\text{Potts}}) &= \lambda_{\min}\left(-A\sum_{i=1}^{d-1}\sum_{m=1}^{q-1}(\mathbf{\Gamma}_q^m)^{(i)}(\mathbf{\Gamma}_q^{q-m})^{(i+1)}\right)\\ &= -A\cdot\lambda_{\max}\left(\sum_{i=1}^{d-1}\sum_{m=1}^{q-1}(\mathbf{\Gamma}_q^m)^{(i)}(\mathbf{\Gamma}_q^{q-m})^{(i+1)}\right)=-A(d-1)(q-1).\end{aligned}$$

□

As for the last section, also the results of the present one are utilized in Chapter 5 to discuss the construction of an initial guess for an iterative eigensolver. The Potts model is considered particularly in Subsection 5.1.3.

3. Tensors and tensor formats

The objects we are concerned with in this thesis are mainly eigenvectors of symmetric matrices representing a Hamilton operator. These eigenvectors have q^d entries, where d is the number of particles/sites of the system under consideration and q is the number of degrees of freedom at each of the d sites. The value of q is typically chosen out of $\{2, 3\}$ but nevertheless the size of the eigenvalue problem at hand grows exponentially in d. This is called the *curse of dimensionality*. Even when d is only 100, such an eigenvector with $2^{100} \approx 1.2677 \cdot 10^{30}$ entries exceeds the capacity of a computer memory by far when we attempt to store each single entry. In order to make computations with such objects like the eigenvectors tractable, one has to assume some kind of structure of these objects. This allows to encode, at least up to a certain accuracy, the information about the object in some set of parameters whose quantity is much lower than the original size of the object. The concept of a *tensor* plays a central role for this approach. A standard reference is [Hac19], see also [KB09] and the german lecture notes [Raa17].

In Section 3.1 we introduce the basic terms and notions and elementary operations with tensors. As tensors can be seen as higher-dimensional generalizations of vectors and matrices, Section 3.2 deals with types of multiplication operations which are in turn generalizations of products of matrices or vectors. In most situations it is impossible to store each single entry of a tensor explicitly since their number scales exponentially with the dimension of the tensor. Techniques to represent a tensor with much lower complexity, under appropriate assumptions and perhaps only approximately, are the topic of Sections 3.3 and 3.4.

3.1. Basic concepts

We begin by defining the central term.

Definition 3.1. Let $d \in \mathbb{N}$ and let $n_j \in \mathbb{N}$, $1 \leq j \leq d$. In this thesis, a *tensor* is defined as a d-dimensional array of real numbers

$$\mathbf{T} = (T_{i_1,\dots,i_d})_{1 \leq i_j \leq n_j,\, 1 \leq j \leq d} \in \mathbb{R}^{n_1 \times \cdots \times n_d}.$$

The number d is the *order* of the tensor.

As two special cases we obtain for $d = 1$ a vector and for $d = 2$ a matrix, so the concept of a tensor may be seen as a higher-order generalization of those. The single indices $j \in \{1, \dots, d\}$ are called *modes* of the tensor $\mathbf{T}$ but the terms j-th *direction, position, dimension, axis, site* are also used. We define the multi-index $\mathbf{i} := (i_1, \dots, i_d)$ with $i_j \in I_j := \{1, \dots, n_j\}$. The index set of $\mathbf{i}$ is the Cartesian product $\mathbf{I} := I_1 \times \cdots \times I_d$. The set of all tensors with modes having size n_j, $1 \leq j \leq d$, is denoted by $\mathbb{R}^{\mathbf{I}} := \mathbb{R}^{n_1 \times \cdots \times n_d}$. For a single entry of $\mathbf{T}$ there are the equivalent notations

$$T_{i_1,\dots,i_d} = T_{\mathbf{i}} = \mathbf{T}(\mathbf{i}) = \mathbf{T}(i_1, \dots, i_d),$$

so we may write $\mathbf{T} = (\mathbf{T}(\mathbf{i}))_{\mathbf{i}\in\mathbf{I}}$. If we fix all indices $l \neq j$ except one index j which runs through its index set I_j we obtain a vector, the *mode-j fiber*. Employing a MATLAB-like notation ":" for the varying index, such a fiber is written as

$$\mathbf{T}(i_1, \ldots, i_{j-1}, :, i_{j+1}, \ldots, i_d) := (\mathbf{T}(\mathbf{i}))_{i_j \in I_j} \in \mathbb{R}^{n_j}.$$

If we instead fix all but two indices (j, k) with $j < k$, we obtain a subset of entries of order two, a matrix, which is called the *mode-(j,k) slice* and analogously written as

$$\mathbf{T}(i_1, \ldots, i_{j-1}, :, i_{j+1}, \ldots, i_{k-1}, :, i_{k+1}, \ldots, i_d) := (\mathbf{T}(\mathbf{i}))_{i_j \in I_j, i_k \in I_k} \in \mathbb{R}^{n_j \times n_k}.$$

In the same fashion, higher-order subsets of tensors may be defined.

The set $\mathbb{R}^{n_1 \times \cdots \times n_d} = \mathbb{R}^{\mathbf{I}}$ of all tensors with mode sizes n_j constitutes a real vector space with the entry-wise addition

$$(\mathbf{S} + \mathbf{T})(\mathbf{i}) := \mathbf{S}(\mathbf{i}) + \mathbf{T}(\mathbf{i}), \quad \mathbf{i} \in \mathbf{I}, \quad \mathbf{S}, \mathbf{T} \in \mathbb{R}^{\mathbf{I}},$$

and the entry-wise scalar multiplication

$$(\lambda\mathbf{T})(\mathbf{i}) := \lambda\mathbf{T}(\mathbf{i}), \quad \mathbf{i} \in \mathbf{I}, \quad \mathbf{T} \in \mathbb{R}^{\mathbf{I}}, \quad \lambda \in \mathbb{R}.$$

For the dimension of $\mathbb{R}^{\mathbf{I}}$ it holds

$$\dim \mathbb{R}^{\mathbf{I}} = \prod_{j=1}^{d} n_j.$$

On $\mathbb{R}^{\mathbf{I}}$ we may define the inner product

$$\langle \mathbf{S}, \mathbf{T} \rangle := \sum_{\mathbf{i}\in\mathbf{I}} \mathbf{S}(\mathbf{i})\mathbf{T}(\mathbf{i}) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} S_{i_1,\ldots,i_d} T_{i_1,\ldots,i_d}, \quad \mathbf{S}, \mathbf{T} \in \mathbb{R}^{\mathbf{I}},$$

which generalizes the Euclidean inner product on $\mathbb{R}^{n_j}$ to the space of tensors. The inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{\mathbf{I}}$ induces a norm, the so-called *Frobenius norm*

$$\|\mathbf{T}\| := \|\mathbf{T}\|_{\mathrm{F}} := \langle \mathbf{T}, \mathbf{T} \rangle^{1/2} = \sqrt{\sum_{\mathbf{i}\in\mathbf{I}} \mathbf{T}(\mathbf{i})^2} = \sqrt{\sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} T_{i_1,\ldots,i_d}^2}, \quad \mathbf{T} \in \mathbb{R}^{\mathbf{I}}.$$

The *Kronecker product* of two tensors $\mathbf{S} \in \mathbb{R}^{m_1 \times \cdots \times m_d}$ and $\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is denoted by

$$\mathbf{S} \otimes \mathbf{T} \in \mathbb{R}^{m_1 n_1 \times \cdots \times m_d n_d}$$

and defined via its entries

$$(\mathbf{S} \otimes \mathbf{T})_{(i_1-1)n_1+j_1,\ldots,(i_d-1)n_d+j_d} := S_{i_1,\ldots,i_d} T_{j_1,\ldots,j_d}, \quad 1 \le i_k \le m_k,\ 1 \le j_k \le n_k,\ 1 \le k \le d.$$

For $d = 2$, the Kronecker product of $\mathbf{A} \in \mathbb{R}^{m_1 \times m_2}$ and $\mathbf{B} \in \mathbb{R}^{n_1 \times n_2}$ is the block matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{1,1}\mathbf{B} & \cdots & a_{1,m_2}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m_1,1}\mathbf{B} & \cdots & a_{m_1,m_2}\mathbf{B} \end{pmatrix} \in \mathbb{R}^{m_1 n_1 \times m_2 n_2},$$

consistent with the definition of the same term in textbooks on matrix analysis like [HJ91]. Regarding the attribution of this particular product to Kronecker, nowadays being standard, see [HPS83].

Next we discuss the possibility to identify tensors in $\mathbb{R}^{n_1\times\cdots\times n_d}$ with vectors, matrices, or other tensors which have also $\prod_{i=1}^d n_i$ single entries. Let us start with the identification of tensors and vectors and define for fixed mode size $\mathbf{n} := (n_1,\dots,n_d)$ the bijective *reverse-lexicographical index ordering function*

$$\begin{aligned}&\operatorname{col}(\cdot,\mathbf{n})\colon \{1,\dots,n_1\}\times\cdots\times\{1,\dots,n_d\}\to\{1,\dots,n_1\cdots n_d\},\\ &\operatorname{col}(\mathbf{i},\mathbf{n}) := i_1+(i_2-1)n_1+(i_3-1)n_1n_2+\ldots+(i_d-1)n_1\cdots n_{d-1}.\end{aligned}$$

It holds

$$\{1,\dots,n_1\}\times\cdots\times\{1,\dots,n_d\} = \{\underbrace{\operatorname{col}(\cdot,\mathbf{n})^{-1}(1)}_{=(1,1,\dots,1)},\underbrace{\operatorname{col}(\cdot,\mathbf{n})^{-1}(2)}_{=(2,1,\dots,1)},\dots,\underbrace{\operatorname{col}(\cdot,\mathbf{n})^{-1}(n_1\cdots n_d)}_{=(n_1,n_2,\dots,n_d)}\}.$$

This corresponds to a *column-major order*. For $d=2$ we obtain that the entries of a matrix $\mathbf{T}\in\mathbb{R}^{n_1\times n_2}$ are stored column-wise in a vector, so for $n_1=2$ and $n_2=3$ it is

$$(T_{\operatorname{col}(\cdot,\mathbf{n})^{-1}(k)})_{1\le k\le 6} = (T_{1,1},T_{2,1},T_{1,2},T_{2,2},T_{1,3},T_{2,3})^\top.$$

Applying the ordering of the indices by $\operatorname{col}(\cdot,\mathbf{n})$, we define the *vectorization* of a tensor $\mathbf{T}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ via the bijective linear operator

$$\operatorname{vec}\colon \mathbb{R}^{n_1\times\cdots\times n_d}\to\mathbb{R}^{n_1\cdots n_d},\quad \operatorname{vec}(\mathbf{T})_k := T_{\operatorname{col}(\cdot,\mathbf{n})^{-1}(k)},\quad 1\le k\le n_1\cdots n_d.$$

The inverse operator, the transformation from a vector to a tensor, is called *tensorization* and denoted by

$$\operatorname{tens}_{n_1\times\cdots\times n_d}\colon \mathbb{R}^{n_1\cdots n_d}\to\mathbb{R}^{n_1\times\cdots\times n_d},\quad \big(\operatorname{tens}_{n_1\times\cdots\times n_d}(\mathbf{v})\big)_{i_1,\dots,i_d} := v_{\operatorname{col}((i_1,\dots,i_d),(n_1,\dots,n_d))}.$$

Especially the case $d=2$ yielding a matrix as the result is included.

Recombinations of modes where the order of both input and output tensor is arbitrary may be described by a composition of vec and tens. This operation is referred to as a *reshaping*, motivated by the MATLAB command `reshape` which implements such types of operations and is in fact heavily used when computationally working with tensors. To indicate that a tensor $\mathbf{T}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ is reshaped into a $\widetilde{d}$-order tensor $\widetilde{\mathbf{T}}$ with mode sizes $(m_1,\dots,m_{\widetilde{d}})$, we use the notation

$$\widetilde{\mathbf{T}} = \operatorname{resh}_{m_1\times\cdots\times m_{\widetilde{d}}}(\mathbf{T}) := \operatorname{tens}_{m_1\times\cdots\times m_{\widetilde{d}}}(\operatorname{vec}(\mathbf{T})),$$

assuming that $\prod_{i=1}^d n_i = \prod_{j=1}^{\widetilde{d}} m_j$.

Remark 3.2. For $\mathbf{v}\in\mathbb{R}^{n\times 1}$ and $\mathbf{w}\in\mathbb{R}^{m\times 1}$ it holds

$$\mathbf{v}\otimes\mathbf{w}^\top = \mathbf{v}\mathbf{w}^\top, \tag{3.1a}$$

$$\mathbf{v}\otimes\mathbf{w} = \operatorname{vec}(\mathbf{w}\mathbf{v}^\top), \tag{3.1b}$$

where $\mathbf{v}\mathbf{w}^\top$ and $\mathbf{w}\mathbf{v}^\top$ have to be understood as the standard matrix-matrix product applied to vectors.

Furthermore, we define the *permutation* operator

$$\mathrm{perm}_{p_1,\dots,p_d} \colon \mathbb{R}^{n_1\times\cdots\times n_d} \to \mathbb{R}^{n_{p_1}\times\cdots\times n_{p_d}}, \quad \left(\mathrm{perm}_{p_1,\dots,p_d}(\mathbf{T})\right)_{i_{p_1},\dots,i_{p_d}} := T_{i_1,\dots,i_d},$$

with $(p_1,\dots,p_d)$ being a permutation of $(1,\dots,d)$, to permute the modes without changing the order of the tensor.

A particular combination of permutation and reshaping which we call *matricization* occurs frequently in the sequel. For $1 \le k \le d$ let $t = \{t_1,\dots,t_k\} \subset \{1,\dots,d\}$ be a subset of modes and $s = \{1,\dots,d\} \setminus t$, so $\{1,\dots,d\} = t \dot\cup s$. Then we define

$$\mathrm{mat}_t(\mathbf{T}) := \mathbf{T}^{(t)} \in \mathbb{R}^{(n_{t_1}\cdots n_{t_k})\times(n_{s_1}\cdots n_{s_{d-k}})} \quad \text{via} \quad \left(\mathbf{T}^{(t)}\right)_{\mathrm{col}(\mathbf{i}_t,\mathbf{n}_t),\mathrm{col}(\mathbf{i}_s,\mathbf{n}_s)} := T_{i_1,\dots,i_d},$$

where $\mathbf{i}_t := (i_{t_1},\dots,i_{t_k})$, $\mathbf{n}_t := (n_{t_1},\dots,n_{t_k})$ and $\mathbf{i}_s := (i_{s_1},\dots,i_{s_{d-k}})$, $\mathbf{n}_s := (n_{s_1},\dots,n_{s_{d-k}})$. It holds

$$\mathrm{mat}_t(\mathbf{T}) = \mathrm{tens}_{(n_{t_1}\cdots n_{t_k})\times(n_{s_1}\cdots n_{s_{d-k}})}\left(\mathrm{vec}\left(\mathrm{perm}_{t_1,\dots,t_k,s_1,\dots,s_{d-k}}(\mathbf{T})\right)\right).$$

3.2. Tensor networks and tensor contractions

A very useful tool to visualize tensors and especially some types of arithmetical operations involving tensors is given by a *diagrammatic notation* which was originally introduced in the physics community and is explained in detail in the survey article [Orú14]. In this diagrammatic notation, a tensor is depicted by some (round or angled) shape having d legs, where d is the order of the tensor. So, a vector as an element of $\mathbb{R}^{n_1}$ has one leg, a matrix has two legs, a third-order tensor from $\mathbb{R}^{n_1\times n_2\times n_3}$ has three legs, and so on. The actual size of n_j cannot be read off from the purely graphical object but it may be attached to the respective leg. It depends on what is better for understanding whether or not a leg corresponding to a mode j with $n_j = 1$ is left out. A scalar value is typically depicted by a shape without legs, see Figure 3.1 for an illustration.

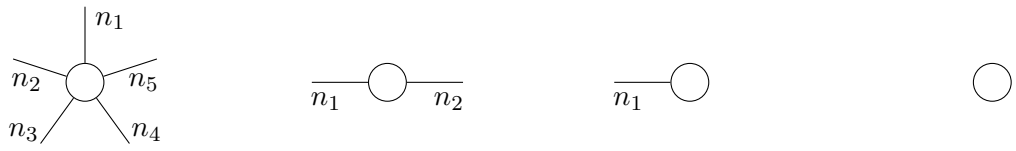


Figure 3.1.: From left to right: Diagrammatic notation of a tensor of order $d = 5$, a matrix ($d = 2$), a vector ($d = 1$), and a scalar ($d = 1$, $n = 1$).

Definition 3.3. For $d_A, d_B \in \mathbb{N}$ let two tensors $\mathbf{A} \in \mathbb{R}^{n_1\times\cdots\times n_{d_A}}$, $\mathbf{B} \in \mathbb{R}^{m_1\times\cdots\times m_{d_B}}$ be given. Assume there are $1 \le k \le \min\{d_A, d_B\}$ modes $t_1,\dots,t_k$ of $\mathbf{A}$ and $s_1,\dots,s_k$ of $\mathbf{B}$ with $n_{t_i} = m_{s_i}$ for $1 \le i \le k$. Then we define the *contraction* of $\mathbf{A}$ and $\mathbf{B}$ along the modes $t_1,\dots,t_k$ and $s_1,\dots,s_k$ entrywise via

$$\begin{aligned}\boxdot_{t_1,\dots,t_k}^{s_1,\dots,s_k} \colon\ & \mathbb{R}^{n_1\times\cdots\times n_{d_A}} \times \mathbb{R}^{m_1\times\cdots\times m_{d_B}} \\ & \to \left(\bigtimes_{p=1,\dots,d_A;\ p\notin\{t_1,\dots,t_k\}} \mathbb{R}^{n_p}\right) \times \left(\bigtimes_{q=1,\dots,d_B;\ q\notin\{s_1,\dots,s_k\}} \mathbb{R}^{m_q}\right), \\ \mathbf{C} := \mathbf{A} \boxdot_{t_1,\dots,t_k}^{s_1,\dots,s_k} \mathbf{B}, \quad & c_{\widehat{(\mathbf{i},\mathbf{j})}} := \sum_{j_{s_1}=i_{t_1}=1}^{n_{t_1}} \cdots \sum_{j_{s_k}=i_{t_k}=1}^{n_{t_k}} a_{i_1,\dots,i_{d_A}} b_{j_1,\dots,j_{d_B}},\end{aligned}$$

where $(\widehat{\mathbf{i}},\widehat{\mathbf{j}}) := \Big((i_p)_{p=1,\dots,d_A;\ p\notin\{t_1,\dots,t_k\}}, (j_q)_{q=1,\dots,d_B;\ q\notin\{s_1,\dots,s_k\}}\Big)$.

The usual product of two matrices $\mathbf{M} \in \mathbb{R}^{m\times n}$ and $\mathbf{N} \in \mathbb{R}^{n\times k}$ reads in this terminology as $\mathbf{MN} = \mathbf{M} \boxdot_2^1 \mathbf{N}$ and with $\mathbf{L} \in \mathbb{R}^{m\times n}$ it holds $\operatorname{trace}(\mathbf{M}^\top\mathbf{L}) = \mathbf{M} \boxdot_{1,2}^{1,2} \mathbf{L}$. This generalizes to the inner product of two tensors $\mathbf{S}, \mathbf{T} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ which may be regarded as a contraction via

$$\langle \mathbf{S}, \mathbf{T}\rangle = \mathbf{S} \boxdot_{1,\dots,d}^{1,\dots,d} \mathbf{T}.$$

A contraction is visualized in the diagrammatic notation by connecting the respective legs of the two involved tensors corresponding to the modes along which the contraction is performed. Using the variables from Definition 3.3, the tensors $\mathbf{A}$ and $\mathbf{B}$ are connected by k legs. The remaining $d_A - k$ legs of $\mathbf{A}$ which are not connected are also called *dangling legs* as well as the $d_B - k$ legs of $\mathbf{B}$. The sum of the number of these dangling legs illustrates that $\mathbf{C} = \mathbf{A} \boxdot_{t_1,\dots,t_k}^{s_1,\dots,s_k} \mathbf{B}$ is a tensor of order $d_A + d_B - 2k$. A small example is given in Figure 3.2.

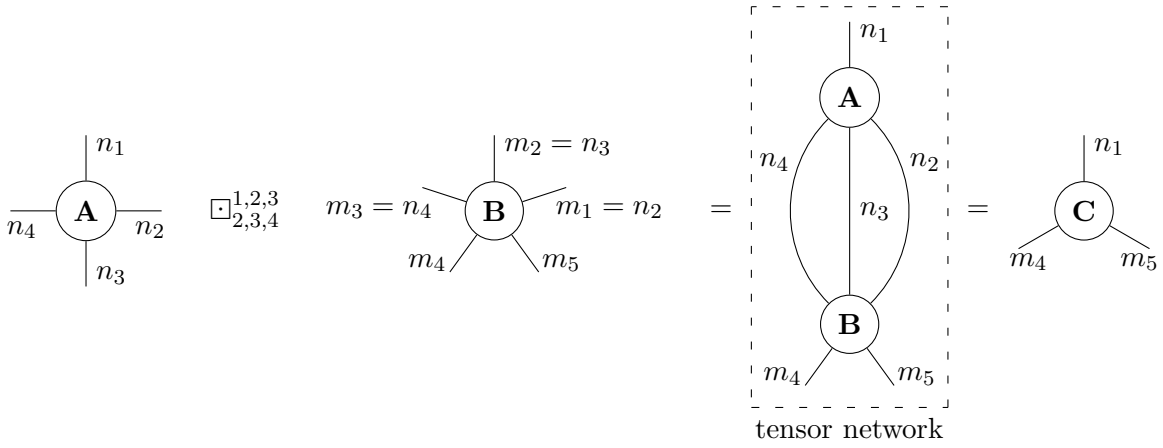


Figure 3.2.: Contraction of fourth-order tensor $\mathbf{A}$ with fifth-order tensor $\mathbf{B}$ along three modes, resulting in third-order tensor $\mathbf{C}$.

An ensemble of tensors together with contractions that involve all tensors of the ensemble is called a *tensor network*. Even for a small number of tensors and contractions in the network, the precise algebraic description of the performed multiplications and additions might become confusing. If one is satisfied with a rather qualitative perspective that displays the structural meaning and behavior of the network, the diagrammatic notation is a very convenient and powerful approach.

As already stated, the matrix-vector multiplication is a certain type of a contraction. So, with regard to the entries of the result, there is no difference if these entries either are computed in the usual way as the sum of products of matrix entries and vector entries, or the matrix and the vector are reshaped into tensors and these two tensors are contracted in an appropriate way. To obtain again a vector as the result, a final reshaping is necessary. Depending on the actual size of the operands, there are many possibilities how to reshape them. Consider $\mathbf{A} \in \mathbb{R}^{16\times 16}$ and $\mathbf{v} \in \mathbb{R}^{16}$. Due to the factorizability of 16, the matrix $\mathbf{A}$ may be reshaped for example into a tensor $\mathbf{A}' \in \mathbb{R}^{4\times4\times4\times4}$, $\mathbf{A}'' \in \mathbb{R}^{2\times2\times2\times2\times2\times2\times2\times2}$, or $\mathbf{A}''' \in \mathbb{R}^{16\times8\times2}$, just to name a few possibilities. In order to perform a contraction, the reshaping of $\mathbf{v}$ has to match the reshaping of $\mathbf{A}$. To illustrate this further, consider the

even smaller case $\mathbf{A} \in \mathbb{R}^{4\times 4}$ and $\mathbf{v} \in \mathbb{R}^4$. Reshape $\mathbf{A}$ into $\widetilde{\mathbf{A}} \in \mathbb{R}^{2\times 2\times 2\times 2}$ according to the column-major order

$$\begin{aligned}
&\widetilde{a}_{1,1,1,1} := a_{1,1}, \quad \widetilde{a}_{1,1,2,1} := a_{1,2}, \quad \widetilde{a}_{1,1,1,2} := a_{1,3}, \quad \widetilde{a}_{1,1,2,2} := a_{1,4},\\
&\widetilde{a}_{2,1,1,1} := a_{2,1}, \quad \widetilde{a}_{2,1,2,1} := a_{2,2}, \quad \widetilde{a}_{2,1,1,2} := a_{2,3}, \quad \widetilde{a}_{2,1,2,2} := a_{2,4},\\
&\widetilde{a}_{1,2,1,1} := a_{3,1}, \quad \widetilde{a}_{1,2,2,1} := a_{3,2}, \quad \widetilde{a}_{1,2,1,2} := a_{3,3}, \quad \widetilde{a}_{1,2,2,2} := a_{3,4},\\
&\widetilde{a}_{2,2,1,1} := a_{4,1}, \quad \widetilde{a}_{2,2,2,1} := a_{4,2}, \quad \widetilde{a}_{2,2,1,2} := a_{4,3}, \quad \widetilde{a}_{2,2,2,2} := a_{4,4}
\end{aligned}$$

and $\mathbf{v}$ into $\widetilde{\mathbf{v}} \in \mathbb{R}^{2\times 2}$ by

$$\begin{aligned}
\widetilde{v}_{1,1} &:= v_1,\\
\widetilde{v}_{2,1} &:= v_2,\\
\widetilde{v}_{1,2} &:= v_3,\\
\widetilde{v}_{2,2} &:= v_4.
\end{aligned}$$

This corresponds to the results of the MATLAB commands `Atilde=reshape(A,[2,2,2,2])` and `vtilde=reshape(v,[2,2])`. With $\widetilde{\mathbf{c}} := \widetilde{\mathbf{A}} \boxdot_{3,4}^{1,2} \widetilde{\mathbf{v}} \in \mathbb{R}^{2\times 2}$ we have

$$\widetilde{c}_{i,j} = \sum_{k=1}^{2}\sum_{l=1}^{2} \widetilde{a}_{i,j,k,l}\widetilde{v}_{k,l} = \sum_{\kappa=1}^{4} a_{\mathrm{col}((i,j),(2,2)),\kappa} v_\kappa$$

which yields

$$\mathbf{c} = \mathbf{A}\mathbf{v} = \mathrm{vec}\left(\widetilde{\mathbf{A}} \boxdot_{3,4}^{1,2} \widetilde{\mathbf{v}}\right) \in \mathbb{R}^4.$$

Another particular type of reshaping a matrix representing a linear mapping into a tensor will be used in the sequel, namely

$$\begin{aligned}
\widetilde{\mathbf{A}} &:= \Psi(\mathbf{A})\\
&:= \mathrm{resh}_{m_1 n_1 \times \cdots \times m_d n_d}\left(\mathrm{perm}_{1,d+1,2,d+2,\ldots,d-1,2d-1,d,2d}\left(\mathrm{resh}_{m_1\times\cdots\times m_d\times n_1\times\cdots\times n_d}(\mathbf{A})\right)\right)
\end{aligned} \tag{3.2}$$

for $\mathbf{A} \in \mathbb{R}^{(m_1\cdots m_d)\times(n_1\cdots n_d)}$.

Other operations like the matrix-matrix product $\mathbf{AB}$ or the Rayleigh quotient $\mathbf{v}^\top\mathbf{A}\mathbf{v}$ in case $\|\mathbf{v}\| = 1$ may be expressed similarly in the set of higher-order tensors via reshaping and contraction. So, many problems involving these operations may be formulated and solved for tensors instead of matrices and vectors. From a numerical point of view, this observation is particularly useful when a tensorized version of a matrix or vector admits a representation with much less data than the original matrix or vector would do. In that case one may expect that the computational complexity as well as the storage requirements can be reduced significantly.

3.3. Hierarchical Tucker format

After having introduced the concept of a tensor, the next step is to describe how to efficiently *represent* a tensor when an explicit storage of each single entry is not feasible due to memory restrictions as the number of entries grows exponentially with the order d. Also the numerical treatment of tensors may greatly benefit from storing only a hopefully much smaller set of parameters which describe the tensor sufficiently precise in view of the need to manipulate

much less quantities. This transition from the full tensor of entries to a set of perhaps much more relevant components is called *tensor decomposition*. Different strategies are known to accomplish this, many of them extending the singular value decomposition (SVD) of matrices [HJ13, Thm. 2.6.3] each in their own way. We recommend the survey articles [KB09] and [GKT13].

Depending on what particular type of decomposition is employed, we say that the result is in a specific *tensor format*. A tensor, as well as a matrix or vector, just consisting explicitly of the single entries, is called to be given in *full format*. In [Hac19, Sect. 7.1.3] a conceptual difference between the terms "decomposition" and "representation" is pointed out insofar, as decomposition describes the process of passing from a tensor to a set of parameters that encode information about the tensor, while representation means to have that set of parameters and to regard its specific aggregation as being a tensor.

The type of tensor format we are mainly interested in is the *hierarchical Tucker (HT) format* which was introduced in [HK09]. A comprehensive treatment of HT may be found in [Hac19, Chap. 11]. We also follow the exposition in [Tob12, Chap. 3].

3.3.1. Construction

First, we need a hierarchical splitting of the modes $1, \dots, d$.

Definition 3.4. A binary tree $\mathcal{T}$ where each node represents a subset of $\{1, \dots, d\}$ is called a *dimension tree* if

- the root node is $\{1, \dots, d\}$,
- each leaf node is a set with exactly one element,
- each parent node is the disjoint union of its two children nodes.

We denote by

- $\mathcal{L}(\mathcal{T}) := \{\{1\}, \dots, \{d\}\}$ the set of all leaf nodes,
- $\mathcal{N}(\mathcal{T})$ the set of all non-leaf nodes, $\mathcal{N}(\mathcal{T}) := \mathcal{T} \setminus \mathcal{L}(\mathcal{T})$.

For simplicity we assume that each element in the left child t_l of a node $t \in \mathcal{T}$ is smaller than any element in the right child t_r of t. By an appropriate reordering of the modes this assumption can always be satisfied.

Remark 3.5 (cf. [Hac19, Remark 11.3]). For the number of elements in a dimension tree $\mathcal{T}$ associated with $\{1, \dots, d\}$ it holds

$$|\mathcal{T}| = 2d - 1,$$

so $|\mathcal{L}(\mathcal{T})| = d$ implies that

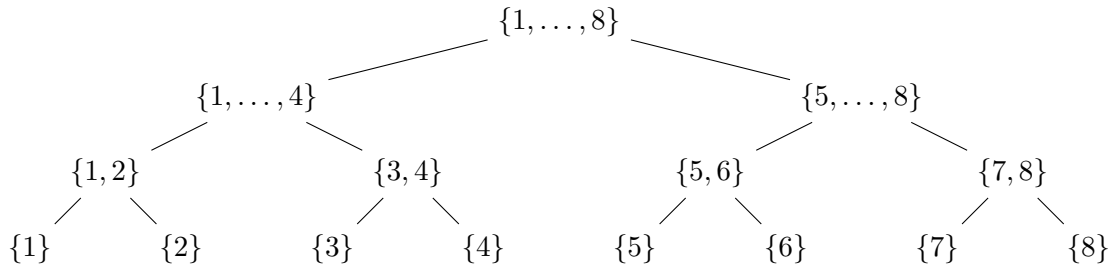
$$|\mathcal{N}(\mathcal{T})| = d - 1.$$

Definition 3.6. A dimension tree with root node $\{1, \dots, d\}$ is called

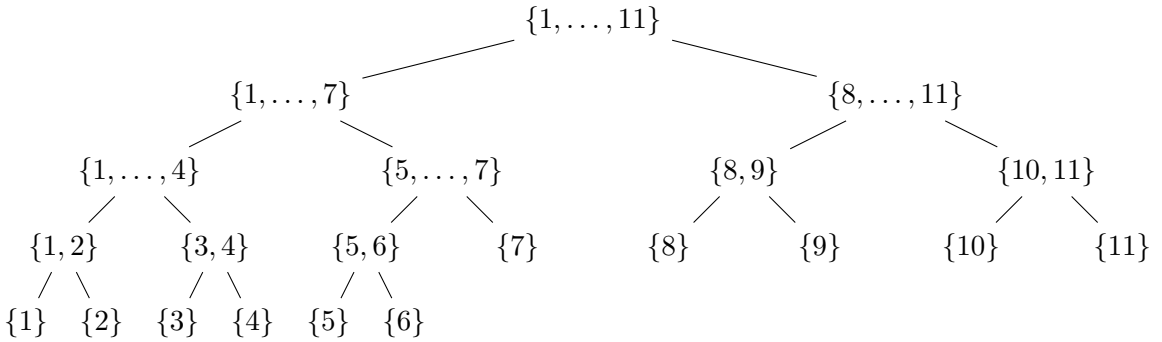
(i) *balanced*, if the $k = 2d - 2^{\lceil \log_2(d) \rceil}$ consecutive leaf nodes $\{1\}, \dots, \{k\}$ all have distance δ to the root node and the other $d - k$, if $d = 2^l$ zero, consecutive leaf nodes all have distance $\delta - 1$ to the root node,

(ii) *linear*, if only the left child or no child of every non-leaf node is a non-leaf node.

Example 3.7. For $d = 8 = 2^3$ the balanced dimension tree equals

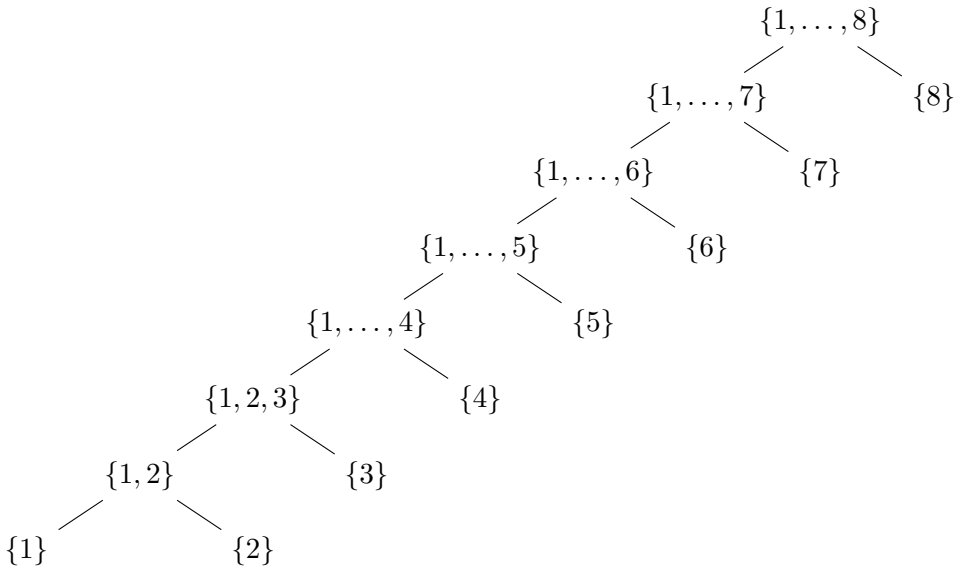


while for $d = 11$, not being a power of 2, it looks like



.

The linear dimension tree for $d = 8$ has the form



.

The hierarchy of the modes prescribed by the dimension tree gives rise to a hierarchy of the different matricizations of the tensor $\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$. The properties of these matricizations in relation to each other are the subject of the following statement.

Lemma 3.8 ([Tob12, Lemma 3.1]). *Let* $\mathbf{T} \in \mathbb{R}^{n_1 \times \ldots \times n_d}$ *and for* $1 \leq i_l \leq i_m \leq i_r \leq d$ *let* $t = t_l \cup t_r$ *with* $t_l = \{i_l, i_l + 1, \ldots, i_m\}$ *and* $t_r = \{i_m + 1, \ldots, i_r\}$. *For the respective matricizations it holds*

$$\operatorname{span}\left(\mathbf{T}^{(t)}\right) \subset \operatorname{span}\left(\mathbf{T}^{(t_r)} \otimes \mathbf{T}^{(t_l)}\right),$$

where $\operatorname{span}(\mathbf{A}) = \operatorname{span}\{a_{:,j_2} \colon 1 \leq j_2 \leq m_2\}$ *denotes the span of the columns of a matrix* $\mathbf{A} = (a_{j_1,j_2})_{1 \leq j_1 \leq m_1,\, 1 \leq j_2 \leq m_2} \in \mathbb{R}^{m_1 \times m_2}$.

For $\tau \in \{t, t_l, t_r\}$, let $\left\{\mathbf{u}_i^{(\tau)} \colon 1 \leq i \leq r_\tau\right\}$ be a basis of $\operatorname{span}\left(\mathbf{T}^{(\tau)}\right)$ with $r_\tau = \operatorname{rank}\left(\mathbf{T}^{(\tau)}\right)$. Due to [Hac19, Lemma 3.13(a)], the elements of $\left\{\mathbf{u}_j^{(t_r)} \otimes \mathbf{u}_i^{(t_l)} \colon 1 \leq j \leq r_{t_r}, 1 \leq i \leq r_{t_l}\right\}$ are basis vectors of $\operatorname{span}\left(\mathbf{T}^{(t_r)} \otimes \mathbf{T}^{(t_l)}\right)$. By Lemma 3.8, each $\mathbf{u}_q^{(t)}$ may be represented by a linear combination of the $\mathbf{u}_j^{(t_r)} \otimes \mathbf{u}_i^{(t_l)}$, so there exist coefficients $b_{i,j,q}^{(t)} \in \mathbb{R}$ with

$$\mathbf{u}_q^{(t)} = \sum_{i=1}^{r_{t_l}} \sum_{j=1}^{r_{t_r}} \left(\mathbf{u}_j^{(t_r)} \otimes \mathbf{u}_i^{(t_l)}\right) b_{i,j,q}^{(t)}, \quad 1 \leq q \leq r_t. \tag{3.3}$$

We collect these coefficients in a tensor $\mathbf{B}_t := \left(b_{i,j,q}^{(t)}\right)_{1 \leq i \leq r_{t_l},\, 1 \leq j \leq r_{t_r},\, 1 \leq q \leq r_t} \in \mathbb{R}^{r_{t_l} \times r_{t_r} \times r_t}$, which is called the *transfer tensor* at the node t.

A transfer tensor $\mathbf{B}_t$ exists for each node $t \in \mathcal{N}(\mathcal{T})$. At the node $t = \text{root} := \{1, \ldots, d\}$, the transfer tensor is in fact an $r_{t_l} \times r_{t_r}$ matrix.

Definition 3.9. For a dimension tree $\mathcal{T}$, let a tuple $(k_t)_{t \in \mathcal{T}} \in \mathbb{N}_0^{\mathcal{T}}$ be given, which assigns to each node $t \in \mathcal{T}$ a nonnegative integer $k_t \in \mathbb{N}_0$.

(i) The set of *hierarchical Tucker tensors of hierarchical rank at most* $(k_t)_{t \in \mathcal{T}}$ is defined as

$$\mathcal{H}\text{-Tucker}\big((k_t)_{t \in \mathcal{T}}\big) := \left\{\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d} \colon \operatorname{rank}\left(\mathbf{T}^{(t)}\right) \leq k_t \text{ for all } t \in \mathcal{T}\right\}.$$

(ii) A tensor $\mathbf{T} \in \mathcal{H}\text{-Tucker}\big((k_t)_{t \in \mathcal{T}}\big)$ is represented in the *hierarchical Tucker format* if

- at each leaf node $t = \{j\} \in \mathcal{L}(\mathcal{T})$, a matrix $\mathbb{R}^{n_j \times r_t} \ni \mathbf{U}_t = \left(\mathbf{u}_i^{(t)}\right)_{1 \leq i \leq r_t}$ whose columns are basis vectors of $\operatorname{span}\left(\mathbf{T}^{(t)}\right)$ with $r_t := \operatorname{rank}\left(\mathbf{T}^{(t)}\right) \leq k_t$ is stored,
- at each non-leaf parent node $t \in \mathcal{N}(\mathcal{T})$ with children t_l and t_r, the third-order transfer tensor $\mathbf{B}_t \in \mathbb{R}^{r_{t_l} \times r_{t_r} \times r_t}$ satisfying (3.3) with $r_{\text{root}} = 1$ is stored.

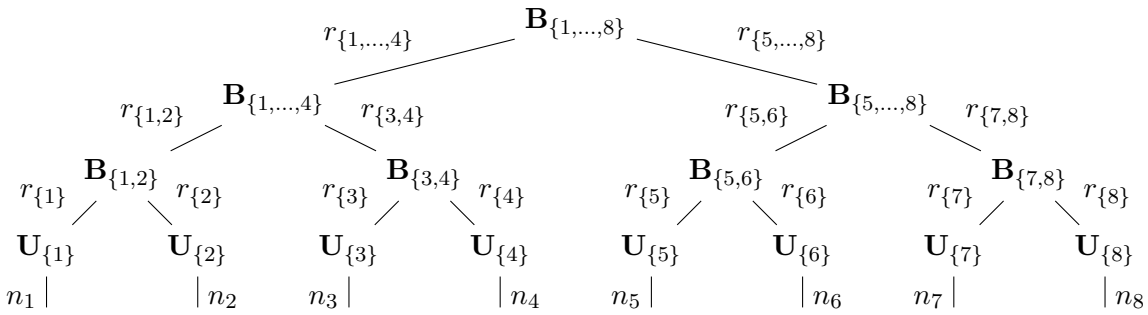
The tuple

$$\mathfrak{H}_{\mathbf{T}} := \left[\mathcal{T}, (\mathbf{B}_t)_{t \in \mathcal{N}(\mathcal{T})}, (\mathbf{U}_t)_{t \in \mathcal{L}(\mathcal{T})}\right]$$

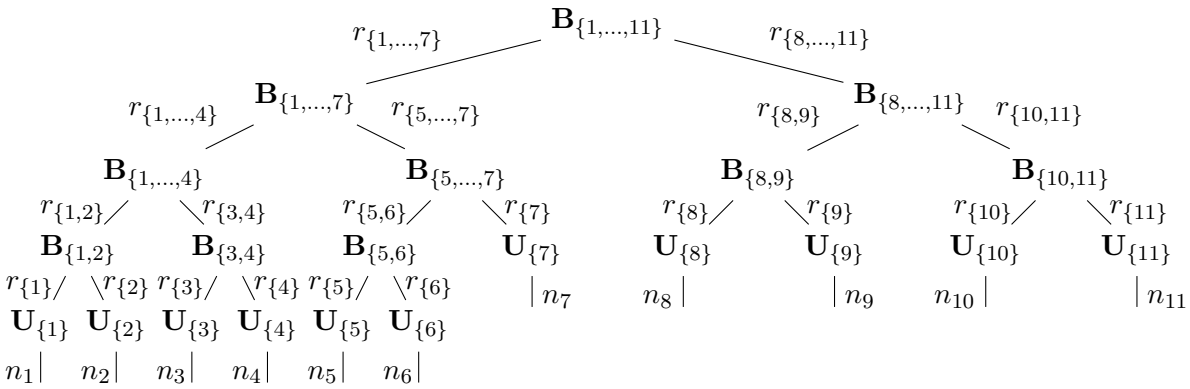
is called a *hierarchical Tucker (HT) representative* of $\mathbf{T}$.

To put it in one rather informal phrase: The hierarchical Tucker format represents recursively the vectorization of a tensor as a linear combination of Kronecker products where the two factors itself are represented as linear combinations of Kronecker products, the shape of the matricizations whose factorizations are considered going along with the hierarchy prescribed by the dimension tree, and the recursion in each branch stops if one arrives at the respective leaf of the tree. One might argue whether in the previous sentence "recurs-" may be replaced by "iterat-" and the direction is reversed to starting at a leaf.

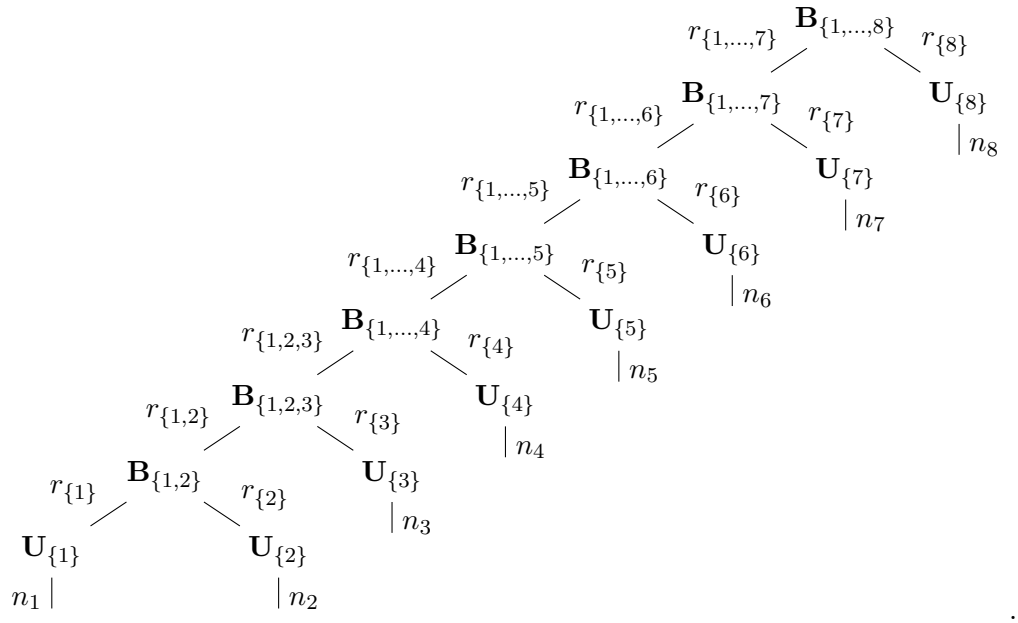
Remark and Example 3.10. (i) It is often beneficial to visualize an HT representative as a tensor network. For this purpose, the leaf matrices $\mathbf{U}_t$, $t \in \mathcal{L}(\mathcal{T})$, and the transfer tensors $\mathbf{B}_t$, $t \in \mathcal{N}(\mathcal{T})$, are just inserted at the respective nodes of the dimension tree. The graphical surround of the content at each node like in Figure 3.2 may be omitted. With the dimension trees from Example 3.7 we obtain



and



as well as



.

The latter special case of a tensor network based on a linear dimension tree is in case $\mathbf{U}_{\{j\}} = \mathbf{I}_{n_j}$, $2 \leq j \leq d$, closely related to a *tensor train* which is an object interesting to study on its own terms, see [Ose11]. We return to this topic in Section 3.4.

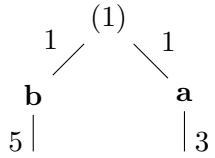
(ii) The somewhat unintuitive fact that in (3.3) the basis vector $\mathbf{u}^{(t_r)}$ corresponding to the right child appears as the left Kronecker factor and vice versa may be illustrated as follows. Let $\mathbf{a} = (a_1\ a_2\ a_3)^\top \in \mathbb{R}^{3\times 1}$ and $\mathbf{b} = (b_1\ b_2\ b_3\ b_4\ b_5)^\top \in \mathbb{R}^{5\times 1}$ and consider the Kronecker product

$$\mathbf{a} \otimes \mathbf{b} = \begin{pmatrix} a_1\mathbf{b} \\ a_2\mathbf{b} \\ a_3\mathbf{b} \end{pmatrix}.$$

Let further the dimension tree $\mathcal{T}$ be such that each element in the left child is smaller than any element in the right child, hence

$$\mathcal{T} = \begin{array}{ccc} & \{1,2\} & \\ \swarrow & & \searrow \\ \{1\} & & \{2\} \end{array}.$$

Then, if the representative $\Big[\mathcal{T}, \mathbf{B}_{\{1,2\}} = (1), \mathbf{U}_{\{1\}} = \mathbf{b}, \mathbf{U}_{\{2\}} = \mathbf{a}\Big]$ or equally as a tensor network



is contracted, see Definition 3.3, to a tensor $\mathbf{b} \boxdot_2^2 \mathbf{a} \in \mathbb{R}^{5\times 3}$ via

$$(\mathbf{b} \boxdot_2^2 \mathbf{a})_{i,j} = b_i a_j,$$

then it holds $\operatorname{vec}(\mathbf{b} \boxdot_2^2 \mathbf{a}) = \mathbf{a} \otimes \mathbf{b}$ when the vectorization is carried out column-wise according to our convention. In essence this is the identity $\operatorname{vec}(\mathbf{b}\mathbf{a}^\top) = \mathbf{a} \otimes \mathbf{b}$ from (3.1b).

Remark 3.11. As long as they constitute a generating system for $\operatorname{span}\big(\operatorname{mat}_t(\mathbf{T})\big)$, also linearly dependent vectors $\mathbf{u}_i^{(t)} \in \mathbb{R}^{n_j}$ may be contained at a leave $t = \{j\} \in \mathcal{L}(\mathcal{T})$. In that case, the transfer tensor $\mathbf{B}_\tau$ is not uniquely determined, where τ is the parent node of t, cf. [Hac19, Chapter 11, Footnote 9].

According to Remark 3.5 there are d leaf matrices $\mathbf{U}_t$ and $d-1$ transfer tensors $\mathbf{B}_t$, one of which is a matrix at the root node of $\mathcal{T}$. Therefore, if $r = \max_{t\in\mathcal{T}}\{r_t\}$ and $n = \max\{n_1, \ldots, n_d\}$, we have a maximal storage requirement of

$$dnr + (d-2)r^3 + r^2.$$

How this number scales with respect to d for a specific tensor is determined by how the maximal rank r depends on d.

3.3.2. Low-rank property

Every tensor $\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ can be trivially represented in the HT format by choosing the transfer tensor at the root node of $\mathcal{T}$ as $\mathbf{B}_{\mathrm{root}} = \mathrm{mat}_{\mathrm{root}_l}(\mathbf{T})$ and the transfer tensors at lower levels in the tree as well as bases in the leaves as concatenations of matricizations of standard basis vectors of appropriate size. To illustrate this, let the 8-order tensor $\mathbf{A} \in \mathbb{R}^{2\times2\times2\times2\times2\times2\times2\times2}$ be given. We choose the balanced dimension tree $\mathcal{T}$ from Example 3.7 and set $\mathbf{U}_{\{j\}} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for all $1 \leq j \leq 8$ and $\mathbf{B}_k = \mathrm{resh}_{2\times2\times4}(\mathbf{I}_4)$ for all $k \in \{\{1,2\},\{3,4\},\{5,6\},\{7,8\}\}$ which yields

$$(\mathbf{B}_k)_{:,:,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (\mathbf{B}_k)_{:,:,2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (\mathbf{B}_k)_{:,:,3} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad (\mathbf{B}_k)_{:,:,4} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

If we set further

$$\mathbf{B}_i = \mathrm{resh}_{4\times4\times16}(\mathbf{I}_{16})$$

for $i \in \{\{1,2,3,4\},\{5,6,7,8\}\}$ and $\mathbf{B}_{\{1,2,3,4,5,6,7,8\}} = \mathrm{mat}_{\{1,2,3,4\}}(\mathbf{A})$, we obtain a representation of $\mathbf{A}$ in HT. So it is $r_t = r_{t_l} r_{t_r}$, $t \in \mathcal{N}(\mathcal{T})$, and $r_{\{j\}} = n_j$, $j \in \{1,\ldots,d\}$. The behavior

$$r_t = \min\left\{\prod\nolimits_{j\in t} n_j, \prod\nolimits_{j\in\{1,\ldots,d\}\setminus t} n_j\right\}, \quad t \in \mathcal{T},$$

provides a natural upper bound for the hierarchical ranks and it turns out that with probability one, the hierarchical ranks necessary to represent a random tensor attain these bounds, see [Hac19, Remk. 11.16]. This is unfavourable from a numerical point of view since it leads again to an exponential scaling with respect to d of the objects to handle. Therefore we assume that the tensors we are concerned with in the sequel obey a *low-rank property* which is known for example from [Bac23] and is in our opinion best understood as

> "The largest one of the three ranks r_t, r_{t_l}, r_{t_r}
> is much smaller than the product of the other two for most of the t
> so that $\max\{r_t : t \in \mathcal{T}\}$ does not grow exponentially in d."

When formulating this criterion we note that for a given representative with ranks r_t, we can reduce these ranks at no accuracy loss insofar as for each $t \in \mathcal{T}$, the largest one of r_t, r_{t_l}, r_{t_r} does not exceed the product of the other two. For a practical illustration we point to the transition from (5.41) to (5.42).

3.3.3. Representation of Hamilton operators

A particular example of tensors which admit a low-rank decomposition in HT format are the Hamilton operators considered in Chapter 2. Recall that $\mathbf{H}_{d,q} \in \mathbb{R}^{q^d \times q^d}$ may be reshaped, including a permutation of modes, into $\widetilde{\mathbf{H}_{d,q}} \in \mathbb{R}^{(q^2)^{\times d}}$ via (3.2). Consider

$$\mathbf{H}_{d,q}^{\mathrm{XYZ}} = A \sum_{i=1}^{d-1} \mathbf{S}_{x,q}^{(i)} \mathbf{S}_{x,q}^{(i+1)} + B \sum_{i=1}^{d-1} \mathbf{S}_{y,q}^{(i)} \mathbf{S}_{y,q}^{(i+1)} + \Delta \sum_{i=1}^{d-1} \mathbf{S}_{z,q}^{(i)} \mathbf{S}_{z,q}^{(i+1)} + h \sum_{i=1}^{d} \mathbf{S}_{z,q}^{(i)}$$

from (2.1) with associated $\widetilde{\mathbf{H}_{d,q}^{\mathrm{XYZ}}} \in \mathbb{R}^{(q^2)^{\times d}}$. Given a dimension tree $\mathcal{T}$, the tensor $\widetilde{\mathbf{H}_{d,q}^{\mathrm{XYZ}}}$ is represented by

$$\left[\mathcal{T}, (\mathbf{B}_t)_{t\in\mathcal{N}(\mathcal{T})}, (\mathbf{U}_t)_{t\in\mathcal{L}(\mathcal{T})}\right],$$

where

$$\mathbf{U}_t = \Big(\mathrm{vec}(\mathbf{I}_q) \quad \mathrm{vec}(\mathbf{S}_{x,q}) \quad \mathrm{vec}(\mathrm{i}\mathbf{S}_{y,q}) \quad \mathrm{vec}(\mathbf{S}_{z,q})\Big) \tag{3.4}$$

for all $t \in \mathcal{L}(\mathcal{T})$. The additional prefactor of $\mathbf{S}_{y,q}$ may be included to obtain, as $\mathrm{i}\mathbf{S}_{y,q} \in \mathbb{R}^{q\times q}$, only real entries in $\mathbf{U}_t$ since $\mathrm{i}\mathbf{S}_{y,q} \otimes \mathrm{i}\mathbf{S}_{y,q} = -(\mathbf{S}_{y,q} \otimes \mathbf{S}_{y,q})$ and the opposite sign is then compensated in the respective entries of $\mathbf{B}_t$ containing B. For the value of $\mathbf{B}_t, t \in \mathcal{N}(\mathcal{T}) \setminus \{\text{root}\}$, we have to distinguish four cases depending on the position of t in the dimension tree, whether none, the left, the right, or both children of t are leaves.

- $t_l, t_r \in \mathcal{N}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{8\times 8\times 8}$$

$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \quad (1,2,2),(3,1,3),(1,4,4),(5,1,5),(1,6,6),(7,1,7), \\ & \quad (8,1,8),(1,8,8)\} \\ A, & (i,j,k) = (2,3,8) \\ -B, & (i,j,k) = (4,5,8) \\ \Delta, & (i,j,k) = (6,7,8) \\ 0, & \text{otherwise} \end{cases}$$

- $t_l \in \mathcal{L}(\mathcal{T})$, $t_r \in \mathcal{N}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{4\times 8\times 8}$$

$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \quad (1,2,2),(2,1,3),(1,4,4),(3,1,5),(1,6,6),(4,1,7), \\ & \quad (1,8,8)\} \\ A, & (i,j,k) = (2,3,8) \\ -B, & (i,j,k) = (3,5,8) \\ \Delta, & (i,j,k) = (4,7,8) \\ h, & (i,j,k) = (4,1,8) \\ 0, & \text{otherwise} \end{cases}$$

- $t_l \in \mathcal{N}(\mathcal{T})$, $t_r \in \mathcal{L}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{8\times 4\times 8}$$

$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \quad (1,2,2),(3,1,3),(1,3,4),(5,1,5),(1,4,6),(7,1,7), \\ & \quad (8,1,8)\} \\ A, & (i,j,k) = (2,2,8) \\ -B, & (i,j,k) = (4,3,8) \\ \Delta, & (i,j,k) = (6,4,8) \\ h, & (i,j,k) = (1,4,8) \\ 0, & \text{otherwise} \end{cases}$$

- $t_l, t_r \in \mathcal{L}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{4\times 4\times 8}$$
$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \qquad (1,2,2),(2,1,3),(1,3,4),(3,1,5),(1,4,6),(4,1,7)\} \\ A, & (i,j,k) = (2,2,8) \\ -B, & (i,j,k) = (3,3,8) \\ \Delta, & (i,j,k) = (4,4,8) \\ h, & (i,j,k) \in \{(4,1,8),(1,4,8)\} \\ 0, & \text{otherwise} \end{cases}$$

Furthermore, $\mathbf{B}_{\text{root}} = (\mathbf{B}_t)_{:,:,8}$, where $\mathbf{B}_t$ has to be chosen from one of the cases above depending on the character of the children of $t = \text{root}$. Note that in case that some of the parameters A, B, Δ, h equal 0, the corresponding parts of $\mathbf{U}_t$ and $\mathbf{B}_t$ may be removed. Actually the mode sizes r_t and, if not already smaller, r_{t_l} respectively r_{t_r} of the transfer tensors most left respectively right at each level of the tree, may be further reduced from 8 to 5. This also means that $\mathbf{B}_{\text{root}}$ has at most 5 rows and columns.

Especially for the linear dimension tree, where at each level of the tree only one transfer tensor is present, we obtain

$$\mathbf{B}_{\{1,2\}} \in \mathbb{R}^{4\times 4\times 5}$$
$$\left(\mathbf{B}_{\{1,2\}}\right)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \qquad (1,2,2),(1,3,3),(1,4,4)\} \\ A, & (i,j,k) = (2,2,5) \\ -B, & (i,j,k) = (3,3,5) \\ \Delta, & (i,j,k) = (4,4,5) \\ h, & (i,j,k) \in \{(4,1,5),(1,4,5)\} \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

as well as

$$\mathbf{B}_{\{1,3\}} = \ldots = \mathbf{B}_{\{1,\ldots,d-2\}} = \mathbf{B}^* \in \mathbb{R}^{5\times 4\times 5}$$
$$(\mathbf{B}^*)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1), \\ & \qquad (1,2,2),(1,3,3),(1,4,4), \\ & \qquad (5,1,5)\} \\ A, & (i,j,k) = (2,2,5) \\ -B, & (i,j,k) = (3,3,5) \\ \Delta, & (i,j,k) = (4,4,5) \\ h, & (i,j,k) = (1,4,5) \\ 0, & \text{otherwise} \end{cases} \tag{3.6}$$

and

$$\mathbf{B}_{\{1,\dots,d-1\}} = \mathbf{B}^* \boxdot_3^1 \begin{pmatrix} 0 & 0 & 0 & h \\ 0 & A & 0 & 0 \\ 0 & 0 & -B & 0 \\ 0 & 0 & 0 & \Delta \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{B}_{\mathrm{root}} = \mathbf{I}_4, \tag{3.7}$$

noticing that the additional contraction yields $r_{\{1,\dots,d-1\}} = r_{\{d\}} = 4$.

So, for general $\mathcal{T}$ we have $r_t \leq 8$ for all $t \in \mathcal{T}$ independently of d. Due to $n_j = q^2$, $1 \leq j \leq d$, we obtain that the number of single entries to represent $\widetilde{\mathbf{H}_{d,q}^{\mathrm{XYZ}}}$ is bounded by $4q^2d + 8^3(d-2) + 8^2$ and depends linearly on d.

Also the Hamilton operator of the Potts model

$$\mathbf{H}_{d,q}^{\mathrm{Potts}} = -A \sum_{i=1}^{d-1} \sum_{m=1}^{q-1} (\mathbf{\Gamma}_q^m)^{(i)} (\mathbf{\Gamma}_q^{q-m})^{(i+1)} - h \sum_{i=1}^{d} \sum_{m=1}^{q-1} (\mathbf{\Omega}_q^m)^{(i)}$$

from (2.6) admits a low-rank decomposition in HT. For a dimension tree $\mathcal{T}$, the tensorized version $\widetilde{\mathbf{H}_{d,q}^{\mathrm{Potts}}} = \Psi(\mathbf{H}_{d,q}^{\mathrm{Potts}}) \in \mathbb{R}^{(q^2)^{\times d}}$ of the Hamilton operator is represented by

$$\left[\mathcal{T}, (\mathbf{B}_t)_{t \in \mathcal{N}(\mathcal{T})}, (\mathbf{U}_t)_{t \in \mathcal{L}(\mathcal{T})}\right],$$

where

$$\mathbf{U}_t = \begin{pmatrix} \mathrm{vec}(\mathbf{I}_q) & \mathrm{vec}(\mathbf{\Gamma}_q) & \mathrm{vec}(\mathbf{\Gamma}_q^2) & \dots & \mathrm{vec}(\mathbf{\Gamma}_q^{q-1}) & \mathrm{vec}(h \sum_{m=1}^{q-1} \mathbf{\Omega}_q^m) \end{pmatrix}$$

for all $t \in \mathcal{L}(\mathcal{T})$ and

- $t_l, t_r \in \mathcal{N}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{2q \times 2q \times 2q}$$

$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1)\} \\ & \quad \cup \{(1,s,s) \colon 2 \leq s \leq q\} \\ & \quad \cup \{(s,1,s) \colon q+1 \leq s \leq 2q-1\} \\ & \quad \cup \{(1,2q,2q), (2q,1,2q)\} \\ A, & (i,j,k) \in \{(s, 2q+1-s, 2q) \colon 2 \leq s \leq q\} \\ 0, & \text{otherwise} \end{cases}$$

- $t_l \in \mathcal{L}(\mathcal{T})$, $t_r \in \mathcal{N}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{(q+1) \times 2q \times 2q}$$

$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1)\} \\ & \quad \cup \{(1,s,s) \colon 2 \leq s \leq q\} \\ & \quad \cup \{(s-q+1,1,s) \colon q+1 \leq s \leq 2q-1\} \\ & \quad \cup \{(1,2q,2q), (q+1,1,2q)\} \\ A, & (i,j,k) \in \{(s, 2q+1-s, 2q) \colon 2 \leq s \leq q\} \\ 0, & \text{otherwise} \end{cases}$$

- $t_l \in \mathcal{N}(\mathcal{T})$, $t_r \in \mathcal{L}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{2q\times(q+1)\times 2q}$$
$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1)\} \\ & \quad \cup \{(1,s,s)\colon 2 \le s \le q\} \\ & \quad \cup \{(s,1,s)\colon q+1 \le s \le 2q-1\} \\ & \quad \cup \{(1,q+1,2q),(2q,1,2q)\} \\ A, & (i,j,k) \in \{(s,q+2-s,2q)\colon 2 \le s \le q\} \\ 0, & \text{otherwise} \end{cases}$$

- $t_l, t_r \in \mathcal{L}(\mathcal{T})$:

$$\mathbf{B}_t \in \mathbb{R}^{(q+1)\times(q+1)\times 2q}$$
$$(\mathbf{B}_t)_{i,j,k} = \begin{cases} 1, & (i,j,k) \in \{(1,1,1)\} \\ & \quad \cup \{(1,s,s)\colon 2 \le s \le q\} \\ & \quad \cup \{(s-q+1,1,s)\colon q+1 \le s \le 2q-1\} \\ & \quad \cup \{(1,q+1,2q),(q+1,1,2q)\} \\ A, & (i,j,k) \in \{(s,q+2-s,2q)\colon 2 \le s \le q\} \\ 0, & \text{otherwise} \end{cases}$$

for $t \in \mathcal{N}(\mathcal{T})$. Furthermore, $\mathbf{B}_{\text{root}} = -(\mathbf{B}_t)_{:,:,2q}$, where $\mathbf{B}_t$ has to be chosen from one of the cases above depending on the character of the children of $t = \text{root}$.

Analogously to the representation for the XYZ model, certain ranks / mode sizes of the transfer tensors most left and right at a level of the tree, including r_{root_l} and r_{root_r}, may be reduced from $2q$ to $q+1$. In the same way, the ranks of a representative based on the linear dimension tree are bounded by $q+1$.

3.3.4. Arithmetical operations

Next we describe how the addition and the application of a linear operator, which may also be viewed as a contraction, affects the hierarchical ranks, see [Hac19, Sect. 11.5] and [Tob12, Sect. 3.3.2 & 3.8]. Fix a dimension tree $\mathcal{T}$. Let $\mathbf{S}, \mathbf{T} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ be two tensors represented by $\mathfrak{H}_{\mathbf{S}} = \left[\mathcal{T}, (\mathbf{B}_t^{\mathbf{S}})_{t\in\mathcal{N}(\mathcal{T})}, (\mathbf{U}_t^{\mathbf{S}})_{t\in\mathcal{L}(\mathcal{T})}\right]$ and $\mathfrak{H}_{\mathbf{T}} = \left[\mathcal{T}, (\mathbf{B}_t^{\mathbf{T}})_{t\in\mathcal{N}(\mathcal{T})}, (\mathbf{U}_t^{\mathbf{T}})_{t\in\mathcal{L}(\mathcal{T})}\right]$ respectively. The sum $\mathbf{S}+\mathbf{T}$ may be represented by $(\mathbf{U}_t)_{t\in\mathcal{L}(\mathcal{T})}$, where

$$\mathbf{U}_t = \begin{pmatrix} \mathbf{U}_t^{\mathbf{S}} & \mathbf{U}_t^{\mathbf{T}} \end{pmatrix} \in \mathbb{R}^{n_j\times(r_t^{\mathbf{S}}+r_t^{\mathbf{T}})}$$

with $t = \{j\}$, $1 \le j \le d$, and $(\mathbf{B}_t)_{t\in\mathcal{N}(\mathcal{T})}$, where

$$\mathbf{B}_{\text{root}} = \begin{pmatrix} \mathbf{B}_{\text{root}}^{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{\text{root}}^{\mathbf{T}} \end{pmatrix}$$

and for $t \in \mathcal{N}(\mathcal{T}) \setminus \{\text{root}\}$, $\mathbf{B}_t \in \mathbb{R}^{(r_{t_l}^{\mathbf{S}}+r_{t_l}^{\mathbf{T}})\times(r_{t_r}^{\mathbf{S}}+r_{t_r}^{\mathbf{T}})\times(r_t^{\mathbf{S}}+r_t^{\mathbf{T}})}$ is a block diagonal tensor with

$$(\mathbf{B}_t)_{:,:,1:r_t^{\mathbf{S}}} = \begin{pmatrix} \mathbf{B}_t^{\mathbf{S}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad (\mathbf{B}_t)_{:,:,r_t^{\mathbf{S}}+1:r_t^{\mathbf{S}}+r_t^{\mathbf{T}}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_t^{\mathbf{T}} \end{pmatrix}.$$

In this way, the hierarchical ranks r_t are added for each $t \in \mathcal{T}$.

Let now $A\colon \mathbb{R}^{n_1 \cdots n_d} \to \mathbb{R}^{m_1 \cdots m_d}$ be a linear mapping and assume that its coefficient matrix $\mathbf{A} \in \mathbb{R}^{(m_1 \cdots m_d) \times (n_1 \cdots n_d)}$ is reshaped into a tensor $\widetilde{\mathbf{A}} \in \mathbb{R}^{m_1 n_1 \times \cdots \times m_d n_d}$ according to (3.2). Let $\widetilde{\mathbf{A}}$ be represented in hierarchical Tucker format by $\mathfrak{H}_{\widetilde{\mathbf{A}}} = \left[\mathcal{T}, (\mathbf{B}_t^{\widetilde{\mathbf{A}}})_{t \in \mathcal{N}(\mathcal{T})}, (\mathbf{U}_t^{\widetilde{\mathbf{A}}})_{t \in \mathcal{L}(\mathcal{T})}\right]$ and let further $\mathfrak{H}_{\widetilde{\mathbf{v}}} = \left[\mathcal{T}, (\mathbf{B}_t^{\widetilde{\mathbf{v}}})_{t \in \mathcal{N}(\mathcal{T})}, (\mathbf{U}_t^{\widetilde{\mathbf{v}}})_{t \in \mathcal{L}(\mathcal{T})}\right]$ be a representative of some $\widetilde{\mathbf{v}} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $\mathbf{v} := \mathrm{vec}(\widetilde{\mathbf{v}})$. Then $\widetilde{\mathbf{w}} := \mathrm{tens}_{m_1 \times \cdots \times m_d}(\mathbf{A}\mathbf{v})$ may be represented by $(\mathbf{U}_t)_{t \in \mathcal{L}(\mathcal{T})}$, where

$$\mathbf{U}_t = \left(\mathrm{tens}_{m_j \times n_j}\left((\mathbf{U}_t^{\widetilde{\mathbf{A}}})_{:,1}\right)\mathbf{U}_t^{\widetilde{\mathbf{v}}} \quad \ldots \quad \mathrm{tens}_{m_j \times n_j}\left((\mathbf{U}_t^{\widetilde{\mathbf{A}}})_{:,r_t^{\widetilde{\mathbf{A}}}}\right)\mathbf{U}_t^{\widetilde{\mathbf{v}}}\right) \in \mathbb{R}^{m_j \times (r_t^{\widetilde{\mathbf{A}}} r_t^{\widetilde{\mathbf{v}}})}$$

with $t = \{j\}$, $1 \le j \le d$, and $(\mathbf{B}_t)_{t \in \mathcal{N}(\mathcal{T})}$, where

$$\mathbf{B}_t = \mathbf{B}_t^{\widetilde{\mathbf{A}}} \otimes \mathbf{B}_t^{\widetilde{\mathbf{v}}} \in \mathbb{R}^{(r_{t_l}^{\widetilde{\mathbf{A}}} r_{t_l}^{\widetilde{\mathbf{v}}}) \times (r_{t_r}^{\widetilde{\mathbf{A}}} r_{t_r}^{\widetilde{\mathbf{v}}}) \times (r_t^{\widetilde{\mathbf{A}}} r_t^{\widetilde{\mathbf{v}}})}.$$

So, the hierarchical ranks r_t are multiplied for each $t \in \mathcal{T}$.

The described addition respectively multiplication of the hierarchical ranks causes a major difficulty when working in an algorithmic way with hierarchical tensors. The ranks grow rapidly and depending on the problem size, only after a few iteration steps the computing resource would be exhausted. Therefore, the ranks have to be reduced regularly in the course of the iteration. A procedure of approximating a hierarchical tensor by another one with smaller ranks is called *truncation*.

For matrices, the singular value decomposition (SVD) is known to yield the best rank-k approximation when truncating the contributions of all singular values except the first k. This result dates back to [Sch07], cf. the historical remark [Ste93, Sect. 5]. In case the distance between two matrices is expressed by the Frobenius norm, it is as well attributed to [EY36] and appears for an arbitrary unitarily invariant norm in [Mir60].

Also for tensors in hierarchical Tucker format, truncation methods based on the SVD exist. The *higher-order singular value decomposition (HOSVD)* [DLDMV00] and the further adaption to the HT format as a *hierarchical SVD* [Gra10] play a central role. Different algorithmic strategies especially concerning the order of traversing the dimension tree are known. We refer to [Hac19, Sect. 11.4] and [Tob12, Sect. 3.6] for a formulation of the algorithms, also for the case of approximating a full tensor in hierarchical Tucker format.

With these techniques, a truncated tensor $\widehat{\mathbf{T}} \in \mathcal{H}\text{-Tucker}\big((k_t)_{t \in \mathcal{T}}\big)$ may be found that approximates a given tensor $\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and satisfies a *quasi-best approximation property* [Gra10, Thm. 3.11 & Remk. 3.12], [Hac19, Thm. 11.66],

$$\|\mathbf{T} - \widehat{\mathbf{T}}\| \le \sqrt{\sum_{t \in \mathcal{T}'} \sum_{i=k_t+1}^{\overline{n_t}} \sigma_i\big(\mathrm{mat}_t(\mathbf{T})\big)^2} \le \sqrt{2d-3}\,\|\mathbf{T} - \mathbf{T}_{\mathrm{best}}\|, \tag{3.8}$$

where $\mathcal{T}' := \mathcal{T} \setminus \{\mathrm{root}, c\}$, c is one child of root, $\overline{n_t} := \min\{\prod_{j \in t} n_j, \prod_{j \in \mathrm{root} \setminus t} n_j\}$, and $\mathbf{T}_{\mathrm{best}}$ is a minimizer of

$$\min_{\mathbf{S} \in \mathcal{H}\text{-Tucker}((k_t)_{t \in \mathcal{T}})} \|\mathbf{T} - \mathbf{S}\|.$$

Note that in the matrix case $d = 2$, the estimate (3.8) coincides with the best approximation property [Hac19, Concl. 2.36] of the classical SVD for a matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$, since $\mathcal{T}'$ then only contains $\{1\}$ resp. $\{2\}$ and so $\mathrm{mat}_t(\mathbf{A})$, $t \in \mathcal{T}'$, equals $\mathbf{A}$ resp. $\mathbf{A}^\top$.

3.4. Tensor train format

Besides the hierarchical Tucker format discussed so far, we also consider the tensor train (TT) format introduced in [Ose11]. TT may be seen as a special case of HT for a linear dimension tree, but due to the structural simplicity it is an object interesting to study on its own terms. We briefly present some basic facts from [Ose11]. The relation between TT and HT is detailed further in [Hac19, Chap. 12].

Definition 3.12. A tensor $\mathbf{T} = \big(\mathbf{T}(i_1,\dots,i_d)\big)_{1\le i_j\le n_j,\,1\le j\le d} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ is said to be represented in the *tensor train (TT) format* by the representative

$$\mathfrak{T}_{\mathbf{T}} := \Big[\mathbf{G}^{(1)},\mathbf{G}^{(2)},\dots,\mathbf{G}^{(d)}\Big]$$

with the *core tensors*

$$\begin{aligned}\mathbf{G}^{(1)} &= \Big(\mathbf{G}^{(1)}(i_1)\Big)_{1\le i_1\le n_1} \in \mathbb{R}^{n_1\times r_1}, \quad \mathbf{G}^{(d)} = \Big(\mathbf{G}^{(d)}(i_d)\Big)_{1\le i_d\le n_d} \in \mathbb{R}^{r_{d-1}\times n_d},\\ \mathbf{G}^{(j)} &= \Big(\mathbf{G}^{(j)}(i_j)\Big)_{1\le i_j\le n_j} \in \mathbb{R}^{r_{j-1}\times n_j\times r_j},\ 2\le j\le d-1,\end{aligned}$$

if it holds entry-wise

$$\begin{aligned}&\mathbf{T}(i_1,\dots,i_d)\\ &= \sum_{k_1=1}^{r_1}\cdots\sum_{k_{d-1}=1}^{r_{d-1}} \Big(\mathbf{G}^{(1)}(i_1)\Big)_{k_1}\Big(\mathbf{G}^{(2)}(i_2)\Big)_{k_1,k_2}\cdots\Big(\mathbf{G}^{(d-1)}(i_{d-1})\Big)_{k_{d-2},k_{d-1}}\Big(\mathbf{G}^{(d)}(i_d)\Big)_{k_{d-1}},\end{aligned} \tag{3.9}$$

understanding $\mathbf{G}^{(j)}(i_j)$, $1\le j\le d$, as $r_{j-1}\times r_j$ matrices. The nonnegative integers $r_1,\dots,r_{d-1}$ are called the *TT ranks* of $\mathbf{T}$ and are supplemented by the convention $r_0 = r_d = 1$.

For fixed $(i_1,\dots,i_d)$ we may rewrite (3.9) as

$$\mathbf{T}(i_1,\dots,i_d) = \mathbf{G}^{(1)}(i_1)\mathbf{G}^{(2)}(i_2)\cdots\mathbf{G}^{(d-1)}(i_{d-1})\mathbf{G}^{(d)}(i_d), \tag{3.10}$$

which motivates the term *matrix product state (MPS)* used in physics literature like [Orú14].

Let $\mathcal{T}$ be the linear dimension tree and let

$$\Big[\mathcal{T},(\mathbf{B}_t)_{t\in\mathcal{N}(\mathcal{T})},(\mathbf{U}_t)_{t\in\mathcal{L}(\mathcal{T})}\Big]$$

be an HT representative of the tensor $\mathbf{T}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ with HT ranks r_t^{HT}, $t\in\mathcal{T}$. Then

$$\begin{aligned}\mathbf{G}^{(1)} &:= \mathbf{U}_{\{1\}}, \quad \mathbf{G}^{(d)} := \mathbf{B}_{\mathrm{root}}\boxdot_2^2\mathbf{U}_{\{d\}},\\ \mathbf{G}^{(j)} &:= \mathrm{perm}_{1,3,2}\Big(\mathbf{B}_{\{1,\dots,j\}}\boxdot_2^2\mathbf{U}_{\{j\}}\Big),\ 2\le j\le d-1,\end{aligned} \tag{3.11}$$

constitutes a TT representative of $\mathbf{T}$ with TT ranks $r_j = r^{\mathrm{HT}}_{\{1,\dots,j\}}$ for all $1\le j\le d-1$. Concerning the positioning of $\mathbf{B}_{\mathrm{root}}$, we may also choose

$$\mathbf{G}^{(d-1)} := \Big(\mathbf{B}_{\{1,\dots,d-1\}}\boxdot_2^2\mathbf{U}_{\{d-1\}}\Big)\boxdot_2^1\mathbf{B}_{\mathrm{root}}, \quad \mathbf{G}^{(d)} := \mathbf{U}_{\{d\}}^{\top},$$

while the other $\mathbf{G}^{(1)},\dots,\mathbf{G}^{(d-2)}$ from (3.11) stay the same, then yielding $r_{d-1} = r^{\mathrm{HT}}_{\{d\}}$.

Going in the opposite direction, from the TT representative $\left[\mathbf{G}^{(1)}, \mathbf{G}^{(2)}, \dots, \mathbf{G}^{(d)}\right]$ with TT ranks r_j^{TT}, $1 \le j \le d-1$, we obtain an HT representative $\left[\mathcal{T}, (\mathbf{B}_t)_{t\in\mathcal{N}(\mathcal{T})}, (\mathbf{U}_t)_{t\in\mathcal{L}(\mathcal{T})}\right]$ via

$$\begin{aligned}\mathbf{U}_{\{1\}} &:= \mathbf{G}^{(1)}, \quad \mathbf{U}_{\{d\}} := (\mathbf{G}^{(d)})^\top, \quad \mathbf{B}_{\mathrm{root}} := \mathbf{I}_{r_{d-1}^{\mathrm{TT}}},\\ \mathbf{U}_{\{j\}} &:= \mathbf{I}_{n_j}, \quad \mathbf{B}_{\{1,\dots,j\}} := \mathbf{G}^{(j)}, \quad 2 \le j \le d-1.\end{aligned}$$

Typically, a representative in the tensor train format is depicted in diagrammatic notation as a horizontal alignment of the cores $\mathbf{G}^{(1)}, \dots, \mathbf{G}^{(d)}$, that is to say

$$\begin{array}{ccccccccc}\mathbf{G}^{(1)} & \overset{r_1}{\text{---}} & \mathbf{G}^{(2)} & \overset{r_2}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{G}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{G}^{(d)}\\ |\, n_1 & & |\, n_2 & & & & |\, n_{d-1} & & |\, n_d\end{array},$$

inspiring the notion "train". A TT representative contains $d-2$ third-order core tensors $\mathbf{G}^{(2)}, \dots, \mathbf{G}^{(d-1)}$ plus the second-order cores $\mathbf{G}^{(1)}$ and $\mathbf{G}^{(d)}$. Therefore, if $r = \max\{r_0, \dots, r_d\}$ and $n = \max\{n_1, \dots, n_d\}$, we have a maximal storage requirement of

$$(d-2)nr^2 + 2nr.$$

How this number scales with respect to d for a specific tensor is determined by how the maximal rank r depends on d.

Also the tensor $\Psi(\mathbf{A}) \in \mathbb{R}^{m_1 n_1 \times \cdots \times m_d n_d}$ from (3.2) associated with the matrix $\mathbf{A} \in \mathbb{R}^{m_1 \cdots m_d \times n_1 \cdots n_d}$ of a linear mapping may be represented in TT. In this situation, it is a matter of choice and might be advantageous to regard the TT cores rather as fourth-order tensors $\mathbf{G}^{(j)}$ of size $r_{j-1} \times m_j \times n_j \times r_j$, $1 \le j \le d$, hence

$$\begin{aligned}&\Big(\mathrm{resh}_{m_1\times n_1\times\cdots\times m_d\times n_d}\big(\Psi(\mathbf{A})\big)\Big)(l_1, i_1, \dots, l_d, i_d)\\ &= \sum_{k_1=1}^{r_1} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} \Big(\mathbf{G}^{(1)}(l_1, i_1)\Big)_{k_1} \Big(\mathbf{G}^{(2)}(l_2, i_2)\Big)_{k_1,k_2}\\ &\qquad\qquad \cdots \Big(\mathbf{G}^{(d-1)}(l_{d-1}, i_{d-1})\Big)_{k_{d-2},k_{d-1}} \Big(\mathbf{G}^{(d)}(l_d, i_d)\Big)_{k_{d-1}}.\end{aligned} \tag{3.12}$$

As in (3.10), for fixed $(l_1, i_1, \dots, l_d, i_d)$, and now understanding $\mathbf{G}^{(j)}(l_j, i_j)$ as $r_{j-1} \times r_j$ matrices, we may rewrite (3.12) as

$$\begin{aligned}&\Big(\mathrm{resh}_{m_1\times n_1\times\cdots\times m_d\times n_d}\big(\Psi(\mathbf{A})\big)\Big)(l_1, i_1, \dots, l_d, i_d)\\ &= \mathbf{G}^{(1)}(l_1, i_1)\mathbf{G}^{(2)}(l_2, i_2)\cdots\mathbf{G}^{(d-1)}(l_{d-1}, i_{d-1})\mathbf{G}^{(d)}(l_d, i_d).\end{aligned}$$

Such a type of representation is also called *matrix product operator (MPO)* in physics and we draw it as

$$\begin{array}{ccccccccc}|\, n_1 & & |\, n_2 & & & & |\, n_{d-1} & & |\, n_d\\ \mathbf{G}^{(1)} & \overset{r_1}{\text{---}} & \mathbf{G}^{(2)} & \overset{r_2}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{G}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{G}^{(d)}\\ |\, m_1 & & |\, m_2 & & & & |\, m_{d-1} & & |\, m_d\end{array}.$$

The low-rank HT representations of the Hamilton operators under our consideration like in (3.5) to (3.7) for the linear dimension tree with leaves (3.4) are transferred to the TT format via the contractions (3.11) followed by, if required, a reshaping of the TT cores to size $r_{j-1} \times m_j \times n_j \times r_j$. Therefore, also the TT ranks of these Hamilton operators are bounded independently of d.

Yet another feature of HT translates to TT, namely that in the course of the addition of two HT tensors respectively the HT version of a matrix-vector product, the ranks are added respectively multiplied, see [Ose11, Sect. 4]. If TT decompositions

$$\mathbf{S} = \left(\mathbf{S}^{(1)}(i_1) \cdots \mathbf{S}^{(d)}(i_d)\right)_{1 \le i_j \le n_j,\, 1 \le j \le d}, \quad \mathbf{T} = \left(\mathbf{T}^{(1)}(i_1) \cdots \mathbf{T}^{(d)}(i_d)\right)_{1 \le i_j \le n_j,\, 1 \le j \le d}$$

of $\mathbf{S}, \mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ are given, then

$$\mathbf{S} + \mathbf{T} =: \mathbf{W} = \left(\mathbf{W}^{(1)}(i_1) \cdots \mathbf{W}^{(d)}(i_d)\right)_{1 \le i_j \le n_j,\, 1 \le j \le d}$$

with

$$\mathbf{W}^{(1)}(i_1) = \begin{pmatrix} \mathbf{S}^{(1)}(i_1) & \mathbf{T}^{(1)}(i_1) \end{pmatrix}, \quad \mathbf{W}^{(d)}(i_d) = \begin{pmatrix} \mathbf{S}^{(d)}(i_d) \\ \mathbf{T}^{(d)}(i_d) \end{pmatrix},$$
$$\mathbf{W}^{(j)}(i_j) = \begin{pmatrix} \mathbf{S}^{(j)}(i_j) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{(j)}(i_j) \end{pmatrix}, \quad 2 \le j \le d-1.$$

If moreover $\mathbf{A} \in \mathbb{R}^{m_1 \cdots m_d \times n_1 \cdots n_d}$ and

$$\mathrm{resh}_{m_1 \times n_1 \times \cdots \times m_d \times n_d}(\Psi(\mathbf{A})) = \left(\mathbf{A}^{(1)}(l_1, i_1) \cdots \mathbf{A}^{(d)}(l_d, i_d)\right)_{1 \le l_j \le m_j,\, 1 \le i_j \le n_j,\, 1 \le j \le d},$$

then

$$\mathrm{tens}_{m_1 \times \cdots \times m_d}(\mathbf{A}\mathrm{vec}(\mathbf{S})) =: \mathbf{Y} = \left(\mathbf{Y}^{(1)}(l_1) \cdots \mathbf{Y}^{(d)}(l_d)\right)_{1 \le l_j \le m_j,\, 1 \le j \le d}$$

with

$$\mathbf{Y}^{(j)}(l_j) = \sum_{i_j=1}^{n_j} \left(\mathbf{A}^{(j)}(l_j, i_j) \otimes \mathbf{S}^{(j)}(i_j)\right).$$

4. Eigensolvers in tensor formats

In this chapter we describe two methods to compute an eigenvector associated with the minimal eigenvalue of a large symmetric matrix in the situation when this matrix is only available by a representation of its tensorization in a tensor format. So, a minimum requirement is that the method does not need access to the single matrix entries, a property that is called *matrix-free*.

The first method, see Section 4.1, is an adaption of an iterative scheme originally developed in full vector format to the hierarchical Tucker format. Each arithmetical operation is carried out in HT, cf. Subsection 3.3.4, followed regularly by truncations to avoid unlimited growth of the ranks.

Section 4.2 deals with a method that exploits the specific structure of the representation of the iterate more directly by updating the single components of the representative one after another, or at most two of them at a time, while leaving the other major number of components unchanged at this time. The precise realization of such a scheme depends on the particular tensor format the tensorization of the matrix and the searched eigenvector is represented in and we focus on the most evident case of the tensor train format.

4.1. Locally optimal conjugate gradient method

We want to find an eigenvector associated with the minimal eigenvalue of a given Hamilton operator $\mathbf{H} := \mathbf{H}_{d,q} \in \mathbb{R}^{q^d \times q^d}$ with $q \in \{2, 3\}$ and $d \in \mathbb{N}$. Throughout this thesis we assume that this minimal eigenvalue $\lambda_{\min}^{(d)}$ is simple. Since $\mathbf{H}$ is a symmetric matrix, it is useful to consider the *Rayleigh quotient*

$$\rho(\mathbf{v}) := \rho_{\mathbf{H}}(\mathbf{v}) := \frac{\mathbf{v}^\top \mathbf{H}\mathbf{v}}{\mathbf{v}^\top \mathbf{v}}, \quad \mathbf{v} \neq \mathbf{0}.$$

By [HJ13, Thm. 4.2.2] it holds

$$\lambda_{\min}^{(d)} = \min_{\mathbf{0}\neq\mathbf{v}\in\mathbb{R}^{q^d}} \rho_{\mathbf{H}}(\mathbf{v})$$

and for the maximal eigenvalue $\lambda_{\max}^{(d)}$ also

$$\lambda_{\max}^{(d)} = \max_{\mathbf{0}\neq\mathbf{v}\in\mathbb{R}^{q^d}} \rho_{\mathbf{H}}(\mathbf{v}).$$

A minimizer resp. maximizer of the Rayleigh quotient is an associated eigenvector $\mathbf{v}_{\min}^{(d)}$ resp. $\mathbf{v}_{\max}^{(d)}$. Except in some trivial cases it is not possible to determine $\lambda_{\min}^{(d)}$ or $\mathbf{v}_{\min}^{(d)}$ analytically. So we have to employ a numerical method. A typical strategy is to construct a sequence of iterates $\{\mathbf{v}_k\}_{k=1,2,\ldots} \subset \mathbb{R}^{q^d}$ such that

$$\rho(\mathbf{v}_{k+1}) \leq \rho(\mathbf{v}_k)$$

for all $k = 0, 1, 2, \ldots$, where $\mathbf{v}_0$ is some initial guess. We may hope that $\mathbf{v}_k$ then converges to $\mathbf{v}_{\min}$ as $k \to \infty$.

The perhaps most simple idea of constructing a sequence of iterates with decreasing Rayleigh quotient is via *gradient descent*. The negative gradient $-\nabla f(\mathbf{x})$ of a function $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ evaluated at a point $\mathbf{x} \in D$ is the direction where f decreases most unless $\mathbf{x}$ is a critical point where $\nabla f(\mathbf{x}) = \mathbf{0}$. So, given a current iterate $\mathbf{v}_k$ and under the assumption that $\mathbf{v}_k \neq \mathbf{0}$ is not any eigenvector, we set $\mathbf{v}_{k+1} := \mathbf{v}_k - \tau_k \mathbf{g}_k$, where

$$\mathbf{g}_k := \frac{1}{2}\nabla\rho(\mathbf{v}_k) = \frac{1}{\mathbf{v}_k^\top \mathbf{v}_k}(\mathbf{H}\mathbf{v}_k - \rho(\mathbf{v}_k)\mathbf{v}_k)$$

is the gradient of the Rayleigh quotient evaluated at $\mathbf{v}_k$, multiplied by a factor $\frac{1}{2}$ for convenience, and $\tau_k > 0$ is a step size. This step size is chosen as the minimizer of

$$\min_{\tau_k \in \mathbb{R}} \rho(\mathbf{v}_k - \tau_k \mathbf{g}_k). \tag{4.1}$$

A more refined scheme for minimizing the Rayleigh quotient is the *locally optimal preconditioned conjugate gradient (LOPCG)* method proposed in [Kny01]. We also point to the lecture notes [Arb16] as a reference. In each iteration step, the Rayleigh quotient is minimized in a three-dimensional search space

$$\operatorname{span}\{\mathbf{v}_k, -\mathbf{g}_k, \mathbf{p}_{k-1}\},$$

where $\mathbf{v}_k$ is the current iterate, $\mathbf{g}_k$ is half the gradient of the Rayleigh quotient evaluated at $\mathbf{v}_k$, and $\mathbf{p}_{k-1}$ is recursively defined as an appropriately weighted linear combination of gradients $\nabla\rho$ evaluated at previous iterates and used in previous iteration steps. To be precise, we set the $(k+1)$-th iterate as

$$\mathbf{v}_{k+1} := \alpha_k \mathbf{v}_k - \beta_k \mathbf{g}_k + \gamma_k \mathbf{p}_{k-1},$$

where $\alpha_k, \beta_k, \gamma_k \in \mathbb{R}$ are simultaneously chosen to be a minimizer of

$$\min_{\alpha_k, \beta_k, \gamma_k \in \mathbb{R}} \rho(\alpha_k \mathbf{v}_k - \beta_k \mathbf{g}_k + \gamma_k \mathbf{p}_{k-1}), \tag{4.2}$$

and we recursively define $\mathbf{p}_{k-1} := -\beta_{k-1}\mathbf{g}_{k-1} + \gamma_{k-1}\mathbf{p}_{k-2}$ with $\mathbf{p}_{-1} = \mathbf{0} \in \mathbb{R}^{q^d}$. This implies that $\rho(\mathbf{v}_{k+1})$ is the minimal value of ρ in $\operatorname{span}\{\mathbf{v}_k, -\mathbf{g}_k, \mathbf{p}_{k-1}\}$. It is

$$\rho_{\mathbf{H}}(\alpha_k \mathbf{v}_k - \beta_k \mathbf{g}_k + \gamma_k \mathbf{p}_{k-1}) = \frac{\begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix}^\top \begin{pmatrix}\mathbf{v}_k^\top \mathbf{H}\mathbf{v}_k & -\mathbf{v}_k^\top \mathbf{H}\mathbf{g}_k & \mathbf{v}_k^\top \mathbf{H}\mathbf{p}_{k-1}\\ -\mathbf{g}_k^\top \mathbf{H}\mathbf{v}_k & \mathbf{g}_k^\top \mathbf{H}\mathbf{g}_k & -\mathbf{g}_k^\top \mathbf{H}\mathbf{p}_{k-1}\\ \mathbf{p}_{k-1}^\top \mathbf{H}\mathbf{v}_k & -\mathbf{p}_{k-1}^\top \mathbf{H}\mathbf{g}_k & \mathbf{p}_{k-1}^\top \mathbf{H}\mathbf{p}_{k-1}\end{pmatrix} \begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix}}{\begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix}^\top \begin{pmatrix}\mathbf{v}_k^\top \mathbf{v}_k & -\mathbf{v}_k^\top \mathbf{g}_k & \mathbf{v}_k^\top \mathbf{p}_{k-1}\\ -\mathbf{g}_k^\top \mathbf{v}_k & \mathbf{g}_k^\top \mathbf{g}_k & -\mathbf{g}_k^\top \mathbf{p}_{k-1}\\ \mathbf{p}_{k-1}^\top \mathbf{v}_k & -\mathbf{p}_{k-1}^\top \mathbf{g}_k & \mathbf{p}_{k-1}^\top \mathbf{p}_{k-1}\end{pmatrix} \begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix}}$$

which is the Rayleigh quotient associated with the 3×3 generalized eigenvalue problem

$$\widetilde{\mathbf{H}} \begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix} = \mu \widetilde{\mathbf{B}} \begin{pmatrix}\alpha_k\\ \beta_k\\ \gamma_k\end{pmatrix},$$

where

$$\widetilde{\mathbf{H}} := \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}^\top \mathbf{H} \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix},$$
$$\widetilde{\mathbf{B}} := \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}^\top \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}$$

are the projections of $\mathbf{H}$ respectively the identity on $\operatorname{span}\{\mathbf{v}_k, -\mathbf{g}_k, \mathbf{p}_{k-1}\}$. For the minimal generalized eigenvalue $\mu_{\min}$ it holds $\rho_{\mathbf{H}}(\mathbf{v}_{k+1}) = \mu_{\min}$. So, the parameters $\alpha_k, \beta_k, \gamma_k$ minimizing (4.2) have to be chosen as the components of a (normalized) generalized eigenvector associated with $\mu_{\min}$. In an analogous way, the minimizing parameter τ_k in (4.1) is determined via solving a 2×2 generalized eigenvalue problem.

The method can be improved by a *preconditioning* of the gradient. Instead of employing $\mathbf{g}_k$ as a search direction, we may replace $\mathbf{g}_k$ by $\widehat{\mathbf{g}_k} := \mathbf{Q}\mathbf{g}_k$, where $\mathbf{Q} \in \mathbb{R}^{q^d \times q^d}$, perhaps even with $\mathbf{Q} = \mathbf{Q}_k$ varying in k. The goal is to enhance the search directions in order to accelerate convergence of the method. Assuming $\mathbf{v}_k$ is not an eigenvector of $\mathbf{H}$, since otherwise $\mathbf{g}_k = \mathbf{0}$, a theoretically ideal choice of this preconditioner $\mathbf{Q}_k$ would be

$$\mathbf{Q}_k = \left(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I}\right)^{-1} \left(\mathbf{I} - \frac{\mathbf{v}_{\min}\mathbf{v}_{\min}^\top}{\mathbf{v}_{\min}^\top \mathbf{v}_{\min}}\right),$$

as can be justified as follows, see also [Hac85, Sect. 12.3.1]. Suppose that the current iterate $\mathbf{v}_k$ is normalized and splitted into $\mathbf{v}_k = \mathbf{v}_{\min} + \mathbf{e}_k$, where $\mathbf{e}_k \in \mathbb{R}^{q^d}$ is some error with $\mathbf{e}_k \in \operatorname{span}\{\mathbf{v}_{\min}\}^\perp$, so $\langle \mathbf{v}_{\min}, \mathbf{e}_k \rangle = 0$. Then it holds

$$\begin{aligned} \mathbf{g}_k &= \mathbf{H}(\mathbf{v}_{\min} + \mathbf{e}_k) - \rho(\mathbf{v}_{\min} + \mathbf{e}_k)(\mathbf{v}_{\min} + \mathbf{e}_k) \\ &= (\mathbf{H} - \rho(\mathbf{v}_{\min} + \mathbf{e}_k)\mathbf{I})\mathbf{v}_{\min} + (\mathbf{H} - \rho(\mathbf{v}_{\min} + \mathbf{e}_k)\mathbf{I})\mathbf{e}_k \\ &= (\lambda_{\min} - \rho(\mathbf{v}_k))\mathbf{v}_{\min} + (\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I})\mathbf{e}_k. \end{aligned}$$

It follows

$$\begin{aligned} \mathbf{Q}_k\mathbf{g}_k &= (\lambda_{\min} - \rho(\mathbf{v}_k)) \left(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I}\right)^{-1} \left(\mathbf{I} - \frac{\mathbf{v}_{\min}\mathbf{v}_{\min}^\top}{\mathbf{v}_{\min}^\top \mathbf{v}_{\min}}\right) \mathbf{v}_{\min} \\ &\quad + \left(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I}\right)^{-1} \left(\mathbf{I} - \frac{\mathbf{v}_{\min}\mathbf{v}_{\min}^\top}{\mathbf{v}_{\min}^\top \mathbf{v}_{\min}}\right) (\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I})\mathbf{e}_k \\ &= \mathbf{0} + \left(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I}\right)^{-1} (\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I})\mathbf{e}_k = \mathbf{e}_k, \end{aligned}$$

since $\mathbf{e}_k \perp \mathbf{v}_{\min}$ and hence also $(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I})\mathbf{e}_k \perp \mathbf{v}_{\min}$. So we obtain $\operatorname{span}\{\mathbf{v}_{\min} + \mathbf{e}_k, \mathbf{e}_k, \mathbf{p}_{k-1}\}$ as the search space and with $\alpha_k \neq 0$, $\beta_k = -\alpha_k$, $\gamma_k = 0$ it is

$$\mathbf{v}_{k+1} = \alpha_k(\mathbf{v}_{\min} + \mathbf{e}_k) + \beta_k\mathbf{e}_k + \gamma_k\mathbf{p}_{k-1} = \alpha_k\mathbf{v}_{\min}.$$

The described choice of the preconditioner surely is unrealistic. One might think of an approximation of

$$\left(\mathbf{H} - \rho(\mathbf{v}_k)\mathbf{I}\right)^{-1} = \sum_{i=1}^{q^d} \frac{1}{\lambda_i - \rho(\mathbf{v}_k)} \mathbf{u}_i\mathbf{u}_i^\top,$$

where $\{\mathbf{u}_i\}_{i=1}^{q^d}$ now denotes an orthonormal set of eigenvectors of $\mathbf{H}$ to avoid confusion with the iterates $\{\mathbf{v}_k\}_{k=1}^{k_{\max}}$. Then, information especially about a good approximation of $\mathbf{u}_1 = \mathbf{v}_{\min}$

is required, due to the proximity of $\lambda_1 = \lambda_{\min}$ to $\rho(\mathbf{v}_k)$ in the relevant stage of the iteration. In fact, the current iterate $\mathbf{v}_k$ is already an approximation of $\mathbf{v}_{\min}$ and it seems questionable wherefrom information about a better approximation should be available.

This line of thought, namely to incorporate approximative information about $\mathbf{v}_{\min}$ directly into the iterates, is central to our idea of starting the iteration already with a best possible initial guess, which is our major concern for the remainder of this thesis. Preconditioning in the LOPCG method however is not discussed further. Therefore we drop the letter "P" and refer in the sequel to the *locally optimal conjugate gradient (LOCG)* method. We summarize it in Algorithm 4.1.

Algorithm 4.1 LOCG in full vector format

Input: symmetric $\mathbf{H} \in \mathbb{R}^{q^d \times q^d}$
initial guess $\mathbf{v}_0 \in \mathbb{R}^{q^d}$
number of iteration steps $k_{\max}$
Output: normalized approximation $\mathbf{v}_{k_{\max}}$ of an eigenvector associated with the minimal eigenvalue of $\mathbf{H}$

1: $\mathbf{v}_0 := \mathbf{v}_0 / \|\mathbf{v}_0\|$
2: $\mathbf{x}_0 := \mathbf{H}\mathbf{v}_0$
3: $\rho_0 := \mathbf{v}_0^\top \mathbf{x}_0$
4: $\mathbf{p}_{-1} := \mathbf{0} \in \mathbb{R}^{q^d}$
5: **for** $k = 0, \ldots, k_{\max} - 1$ **do**
6: $\quad \mathbf{g}_k := \mathbf{x}_k - \rho_k \mathbf{v}_k$
7: $\quad \widetilde{\mathbf{H}} := \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}^\top \mathbf{H} \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}$
8: $\quad \widetilde{\mathbf{B}} := \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}^\top \begin{pmatrix} \mathbf{v}_k & -\mathbf{g}_k & \mathbf{p}_{k-1} \end{pmatrix}$
9: $\quad$ find a normalized generalized eigenvector $(\alpha_k, \beta_k, \gamma_k)^\top := \mathbf{w}$ associated with the minimal generalized eigenvalue $\rho_{k+1} := \mu_{\min}$ of the problem $\widetilde{\mathbf{H}}\mathbf{w} = \mu\widetilde{\mathbf{B}}\mathbf{w}$
10: $\quad \mathbf{p}_k := -\beta_k \mathbf{g}_k + \gamma_k \mathbf{p}_{k-1}$
11: $\quad \mathbf{v}_{k+1} := \alpha_k \mathbf{v}_k + \mathbf{p}_k$
12: $\quad \mathbf{v}_{k+1} := \mathbf{v}_{k+1} / \|\mathbf{v}_{k+1}\|$
13: $\quad \mathbf{x}_{k+1} := \mathbf{H}\mathbf{v}_{k+1}$
14: **end for**

To ensure symmetry up to machine precision of the generalized eigenvalue problem $\widetilde{\mathbf{H}}\mathbf{w} = \mu\widetilde{\mathbf{B}}\mathbf{w}$, which turns out to be beneficial when calling `eig` in MATLAB for this problem, we may add $\widetilde{\mathbf{H}} := (\widetilde{\mathbf{H}} + \widetilde{\mathbf{H}}^\top)/2$ to Line 7 respectively $\widetilde{\mathbf{B}} := (\widetilde{\mathbf{B}} + \widetilde{\mathbf{B}}^\top)/2$ to Line 8. Note further that omitting the vector $\mathbf{p}_k$ in each iteration step and therefore reducing $\widetilde{\mathbf{H}}$ and $\widetilde{\mathbf{B}}$ to the size 2×2 yields the gradient descent method.

In principle, every algorithm of numerical linear algebra which incorporates vectors and matrices as building blocks, and the basic arithmetical operations are sums and products of them, can be formulated in the hierarchical Tucker format. For that purpose, certain vector or matrix valued components of the numerical method have to be regarded as a tensor via an appropriate reshaping, possibly including a permutation of modes. Depending on low-rank properties of the data and the user's preferences concerning accuracy as well as time and memory consumption, these tensors are represented exactly or approximatively in HT. The HT analogs of vector addition and matrix-vector multiplication are easily implemented but lead to a quick growth of the hierarchical ranks, see Subsection 3.3.4. As a remedy, after a

certain number of basic arithmetical operations, a truncation of the HT tensors is necessary to obtain smaller ranks. Here again it is a matter of choice and inherent properties of the problem at hand how much reduction of the ranks is feasible. To what extent the terminal value of the iterate after reaching a stopping criterion is a good approximation of the result of the method when carried out in full format is mostly hard to answer.

In Algorithm 4.2 we state an HT variant of the LOCG method 4.1, cf. [Tob12, Alg. 12]. For $\mathbf{A} \in \mathbb{R}^{m_1 n_1 \times \cdots \times m_d n_d}$ and $\mathbf{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ both represented in HT format with the same dimension tree, we introduce the notation

$$\Phi_{\mathbf{A}}(\mathbf{X}) := \mathrm{tens}_{m_1 \times \cdots \times m_d}\left(\Psi^{-1}(\mathbf{A})\mathrm{vec}(\mathbf{X})\right) = \mathrm{resh}_{m_1 \times n_1 \times \cdots \times m_d \times n_d}(\mathbf{A}) \boxdot_{2,4,\ldots,2d}^{1,2,\ldots,d} \mathbf{X} \tag{4.3}$$

to indicate that a linear operator which is regarded as a tensor via $\Psi(\cdot)$ from (3.2) is applied to another tensor of appropriate size and the result is again a representative in HT format. Concerning an efficient realization in HT format of the inner product $\langle \cdot, \cdot \rangle$ resp. $\langle \cdot, \Phi_{\mathbf{S}}(\cdot) \rangle$ for a symmetric matrix $\Psi^{-1}(\mathbf{S})$, see [Tob12, Sect. 3.5.1 resp. 3.8].

Algorithm 4.2 LOCG for tensors in HT format

Input: $\Phi_{\mathbf{H}}(\cdot)$ for $\mathbf{H}$ given in HT format such that $\Psi^{-1}(\mathbf{H})$ is a symmetric matrix
initial guess $\mathbf{V}_0$ in HT format
number of iteration steps $k_{\max}$
truncation operator Θ

Output: approximation $\mathbf{V}_{k_{\max}}$ in HT of a tensorized normalized eigenvector associated with the minimal eigenvalue of $\Psi^{-1}(\mathbf{H})$

$\mathbf{V}_0 := \mathbf{V}_0 / \|\mathbf{V}_0\|$
$\mathbf{X}_0 := \Phi_{\mathbf{H}}(\mathbf{V}_0)$
$\rho_0 := \langle \mathbf{V}_0, \mathbf{X}_0 \rangle$
$\mathbf{P}_{-1} := \mathbf{0}$ % matching the HT representation of $\mathbf{V}_0$
for $k = 0, \ldots, k_{\max} - 1$ **do**
  $\mathbf{G}_k := \Theta(\mathbf{X}_k - \rho_k \mathbf{V}_k)$
  $\mathbf{G}_k := \mathbf{G}_k / \|\mathbf{G}_k\|$
  $\widetilde{\mathbf{H}} := \begin{pmatrix} \frac{1}{2}\langle \mathbf{V}_k, \Phi_{\mathbf{H}}(\mathbf{V}_k) \rangle & \langle \mathbf{V}_k, \Phi_{\mathbf{H}}(-\mathbf{G}_k) \rangle & \langle \mathbf{V}_k, \Phi_{\mathbf{H}}(\mathbf{P}_{k-1}) \rangle \\ 0 & \frac{1}{2}\langle -\mathbf{G}_k, \Phi_{\mathbf{H}}(-\mathbf{G}_k) \rangle & \langle -\mathbf{G}_k, \Phi_{\mathbf{H}}(\mathbf{P}_{k-1}) \rangle \\ 0 & 0 & \frac{1}{2}\langle \mathbf{P}_{k-1}, \Phi_{\mathbf{H}}(\mathbf{P}_{k-1}) \rangle \end{pmatrix}$
  $\widetilde{\mathbf{B}} := \begin{pmatrix} \frac{1}{2}\langle \mathbf{V}_k, \mathbf{V}_k \rangle & \langle \mathbf{V}_k, -\mathbf{G}_k \rangle & \langle \mathbf{V}_k, \mathbf{P}_{k-1} \rangle \\ 0 & \frac{1}{2}\langle -\mathbf{G}_k, -\mathbf{G}_k \rangle & \langle -\mathbf{G}_k, \mathbf{P}_{k-1} \rangle \\ 0 & 0 & \frac{1}{2}\langle \mathbf{P}_{k-1}, \mathbf{P}_{k-1} \rangle \end{pmatrix}$
  $\widetilde{\mathbf{H}} := \widetilde{\mathbf{H}} + \widetilde{\mathbf{H}}^\top, \quad \widetilde{\mathbf{B}} := \widetilde{\mathbf{B}} + \widetilde{\mathbf{B}}^\top$ % as $\widetilde{\mathbf{H}}$ and $\widetilde{\mathbf{B}}$ are symmetric
  find a normalized generalized eigenvector $(\alpha_k, \beta_k, \gamma_k)^\top := \mathbf{w}$ associated with the minimal generalized eigenvalue $\rho_{k+1} := \mu_{\min}$ of the problem $\widetilde{\mathbf{H}}\mathbf{w} = \mu \widetilde{\mathbf{B}}\mathbf{w}$
  $\mathbf{P}_k := \Theta(-\beta_k \mathbf{G}_k + \gamma_k \mathbf{P}_{k-1})$
  $\mathbf{V}_{k+1} := \Theta(\alpha_k \mathbf{V}_k + \mathbf{P}_k)$
  $\mathbf{V}_{k+1} := \mathbf{V}_{k+1} / \|\mathbf{V}_{k+1}\|$
  $\mathbf{P}_k := \mathbf{P}_k / \|\mathbf{P}_k\|$
  if $k \neq k_{\max} - 1$ **then**
    $\mathbf{X}_{k+1} := \Theta(\Phi_{\mathbf{H}}(\mathbf{V}_{k+1}))$
  end if
end for

4.2. Modified alternating linear scheme

The method we described in the previous section is obtained by taking an iterative method originally designed for full matrices and vectors, replacing these components of the method by representatives in a low-rank tensor format, and then performing the iterations with the arithmetics of this tensor format, keeping the ranks low by regular truncations. In each iteration step, the whole representative is updated and in the present case of the HT format, in principle each transfer tensor and each leaf matrix might be changed.

A conceptually different approach is to update the single component tensors of a representative one after another by some particular rule, which is called a *micro iteration step*, until having updated each of them, and then to repeat this step-by-step treatment of the single components. A widely used method implementing this concept is the *(modified) alternating linear scheme (ALS/MALS)* [HRS12] in an application to eigenvalue problems, known in the physics community as the *density matrix renormalization group (DMRG)* algorithm [Whi92], [Sch11]. It is formulated in the tensor train format, and the order in which the cores are updated proceeds from a particular core to its immediate neighbor. In this section we give an overview of the method, based on and referring for details to [HRS12].

Our task is to find a minimizer of the Rayleigh quotient

$$\rho_{\mathbf{A}}(\mathbf{v}) := \frac{\mathbf{v}^\top \mathbf{A}\mathbf{v}}{\mathbf{v}^\top \mathbf{v}}$$

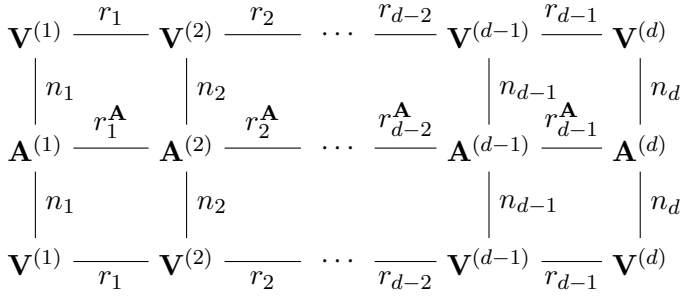
subject to $\mathbf{v} \neq \mathbf{0}$ with a given symmetric $\mathbf{A} \in \mathbb{R}^{n_1 \cdots n_d \times n_1 \cdots n_d}$. If

$$\begin{aligned}&\Big(\mathrm{resh}_{n_1 \times \cdots \times n_d}(\mathbf{v})\Big)(i_1, \ldots, i_d) =: \mathbf{V} \\ &= \sum_{k_1=1}^{r_1} \cdots \sum_{k_{d-1}=1}^{r_{d-1}} \Big(\mathbf{V}^{(1)}(i_1)\Big)_{k_1} \Big(\mathbf{V}^{(2)}(i_2)\Big)_{k_1,k_2} \cdots \Big(\mathbf{V}^{(d-1)}(i_{d-1})\Big)_{k_{d-2},k_{d-1}} \Big(\mathbf{V}^{(d)}(i_d)\Big)_{k_{d-1}}\end{aligned}$$

and

$$\begin{aligned}&\Big(\mathrm{resh}_{n_1 \times n_1 \times \cdots \times n_d \times n_d}(\Psi(\mathbf{A}))\Big)(l_1, i_1, \ldots, l_d, i_d) \\ &= \sum_{k_1=1}^{r_1^{\mathbf{A}}} \cdots \sum_{k_{d-1}=1}^{r_{d-1}^{\mathbf{A}}} \Big(\mathbf{A}^{(1)}(l_1, i_1)\Big)_{k_1} \Big(\mathbf{A}^{(2)}(l_2, i_2)\Big)_{k_1,k_2} \\ &\qquad \cdots \Big(\mathbf{A}^{(d-1)}(l_{d-1}, i_{d-1})\Big)_{k_{d-2},k_{d-1}} \Big(\mathbf{A}^{(d)}(l_d, i_d)\Big)_{k_{d-1}}\end{aligned}$$

are TT representations and provided $\mathbf{v}^\top \mathbf{v} = 1$, then $\rho_{\mathbf{A}}(\mathbf{v})$ reads as



.

For the first micro iteration step we regard $\mathbf{V}^{(2)}, \dots, \mathbf{V}^{(d)}$ being fixed and determine

$$\widetilde{\mathbf{V}}^{(1)} := \arg\min\{\rho_{\mathbf{A}} \circ P_{1,1}(\mathbf{X}^{(1)}) \colon \mathbf{X}^{(1)} \in \mathbb{R}^{r_0 \times n_1 \times r_1}\},$$

where

$$P_{j,1} := P_{j,1,\mathbf{V}} \colon \mathbb{R}^{r_{j-1} \times n_j \times r_j} \to \mathbb{R}^{n_1 \times \cdots \times n_d},$$
$$P_{j,1}(\mathbf{X}^{(j)}) = \left(\mathbf{V}^{(1)}(i_1) \cdots \mathbf{V}^{(j-1)}(i_{j-1}) \mathbf{X}^{(j)}(i_j) \mathbf{V}^{(j+1)}(i_{j+1}) \cdots \mathbf{V}^{(d)}(i_d)\right)_{1 \le i_j \le n_j,\, 1 \le j \le d}$$

is a *one-component retraction operator*. A minimizer is given by

$$\widetilde{\mathbf{V}}^{(1)} = \mathrm{tens}_{r_0 \times n_1 \times r_1}(\mathbf{x}^{(1)}_{\min}), \tag{4.4}$$

where $\mathbf{x}^{(1)}_{\min}$ is a normalized generalized eigenvector associated with the minimal generalized eigenvalue $\lambda^{(1)}_{\min}$ of

$$\widetilde{\mathbf{A}}^{(1)} \mathbf{x} = \lambda \widetilde{\mathbf{B}}^{(1)} \mathbf{x} \tag{4.5}$$

with

$$\widetilde{\mathbf{A}}^{(1)} := \mathrm{resh}_{r_0 n_1 r_1 \times r_0 n_1 r_1} \left(\begin{array}{ccccccccc} & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ \Big| & n_1 & \Big| & n_2 & & & & \Big| \; n_{d-1} & \Big| \; n_d \\ \mathbf{A}^{(1)} & \overline{\qquad} & \mathbf{A}^{(2)} & \overline{\qquad} & & \cdots & \overline{\qquad} \; \mathbf{A}^{(d-1)} & \overline{\qquad} & \mathbf{A}^{(d)} \\ \Big| & n_1 & \Big| & n_2 & & & & \Big| \; n_{d-1} & \Big| \; n_d \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \end{array} \right)$$

and

$$\widetilde{\mathbf{B}}^{(1)} := \mathrm{resh}_{r_0 n_1 r_1 \times r_0 n_1 r_1} \left(\begin{array}{ccccccccc} & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ \Big| & n_1 & \Big| & n_2 & & & & \Big| \; n_{d-1} & \Big| \; n_d \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \end{array} \right).$$

The generalized eigenproblem (4.5) becomes a standard one if

$$\mathrm{resh}_{r_0 n_1 r_1 \times r_0 n_1 r_1} \left(\begin{array}{ccccccccc} & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ \Big| & n_1 & \Big| & n_2 & & & & \Big| \; n_{d-1} & \Big| \; n_d \\ & \overline{\qquad} & \mathbf{V}^{(2)} & \overline{\qquad} & & & \overline{\qquad} \; \mathbf{V}^{(d-1)} & \overline{\qquad} & \mathbf{V}^{(d)} \\ & r_1 & & r_2 & & \cdots & r_{d-2} & r_{d-1} & \end{array} \right) = \mathbf{I}_{r_0 n_1 r_1},$$

which is true if the *right unfolding*

$$\underline{\mathbf{V}^{(j)}}_{\longrightarrow} := \mathrm{mat}_{\{3,2\}}(\mathbf{V}^{(j)}) = \mathrm{resh}_{r_j n_j \times r_{j-1}}\big(\mathrm{perm}_{3,2,1}(\mathbf{V}^{(j)})\big)$$

has orthonormal columns, in which case the core $\mathbf{V}^{(j)}$ is also called *right orthonormal*, for all $2 \leq j \leq d$. Right orthonormality of $\mathbf{V}^{(2)}, \ldots, \mathbf{V}^{(d)}$ can be achieved by successive economy size QR decompositions as a preprocessing step. Decomposing

$$\underrightarrow{\mathbf{V}^{(d)}} = \mathbf{Q}^{(d)}\mathbf{R}^{(d)},$$

overwriting

$$\mathbf{V}^{(d)} := \mathrm{resh}_{r_{d-1} \times n_d \times r_d}((\mathbf{Q}^{(d)})^\top),$$

and shifting the non-orthonormal part $\mathbf{R}^{(d)}$ to the core at position, also called *site*, $d-1$ via replacing $\mathbf{V}^{(d-1)}$ by

$$\widehat{\mathbf{V}}^{(d-1)} := \mathbf{V}^{(d-1)} \boxdot_3^1 (\mathbf{R}^{(d)})^\top,$$

we obtain a right orthonormal core at position d, while the whole represented tensor $\mathbf{V}$ is not altered. This orthogonalization procedure is continued by decomposing

$$\underrightarrow{\widehat{\mathbf{V}}^{(d-1)}} = \mathbf{Q}^{(d-1)}\mathbf{R}^{(d-1)},$$

overwriting $\widehat{\mathbf{V}}^{(d-1)}$ by

$$\mathbf{V}^{(d-1)} := \mathrm{resh}_{r_{d-2} \times n_{d-1} \times r_{d-1}}((\mathbf{Q}^{(d-1)})^\top),$$

shifting $\mathbf{R}^{(d-1)}$ to site $d-2$, and so on until all cores from site 2 to d are right orthonormal, the non-orthonormal part of $\mathbf{V}$ sitting at site 1 which is updated due to (4.4) anyway.

To prepare the next micro iteration step, the *left unfolding*

$$\underleftarrow{\mathbf{V}^{(j)}} := \mathrm{mat}_{\{1,2\}}(\mathbf{V}^{(j)}) = \mathrm{resh}_{r_{j-1} n_j \times r_j}(\mathbf{V}^{(j)})$$

for $j = 1$ is QR decomposed as

$$\underleftarrow{\widetilde{\mathbf{V}}^{(1)}} = \mathbf{Q}^{(1)}\mathbf{R}^{(1)}$$

and the core at position 1 is overwritten by $\mathbf{V}^{(1)} := \mathrm{resh}_{r_0 \times n_1 \times r_1}(\mathbf{Q}^{(1)})$, yielding that the left unfolding of core 1 then has orthonormal columns, in which case we call that core *left orthonormal*.

Now core 2 is updated by

$$\widetilde{\mathbf{V}}^{(2)} := \mathrm{tens}_{r_1 \times n_2 \times r_2}(\mathbf{x}_{\min}^{(2)}), \tag{4.6}$$

where $\mathbf{x}_{\min}^{(2)}$ is a normalized eigenvector associated with the minimal eigenvalue $\lambda_{\min}^{(2)}$ of

$$\widetilde{\mathbf{A}}^{(2)}\mathbf{x} = \lambda\mathbf{x}$$

with

$$\widetilde{\mathbf{A}}^{(2)} := \mathrm{resh}_{r_1 n_2 r_2 \times r_1 n_2 r_2} \left(\begin{array}{ccccccccccc} \mathbf{V}^{(1)} & \overset{r_1}{\text{---}} & & \overset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \overset{r_3}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \\ & \vert & & \vert & & \vert & & & \vert & & \vert \\ & n_1 & & n_2 & & n_3 & & & n_{d-1} & & n_d \\ & \vert & & \vert & & \vert & & & \vert & & \vert \\ \mathbf{A}^{(1)} & \overset{r_1^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(2)} & \overset{r_2^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(3)} & \overset{r_3^{\mathbf{A}}}{\text{---}} & \cdots & \overset{r_{d-2}^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(d-1)} & \overset{r_{d-1}^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(d)} \\ & \vert & & \vert & & \vert & & & \vert & & \vert \\ & n_1 & & n_2 & & n_3 & & & n_{d-1} & & n_d \\ & \vert & & \vert & & \vert & & & \vert & & \vert \\ \mathbf{V}^{(1)} & \overset{r_1}{\text{---}} & & \overset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \overset{r_3}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \end{array} \right).$$

Notice that

$$\operatorname{resh}_{r_1 n_2 r_2 \times r_1 n_2 r_2}\left(\begin{array}{ccccccccccc} \mathbf{V}^{(1)} & \overset{r_1}{\text{---}} & & \overset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \overset{r_3}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \\ \big| & n_1 & & \big| & n_2 & \big| & n_3 & & & \big| & n_{d-1} & \big| & n_d \\ \mathbf{V}^{(1)} & \underset{r_1}{\text{---}} & & \underset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \underset{r_3}{\text{---}} & \cdots & \underset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \underset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \end{array}\right)$$

$$= \mathbf{I}_{r_1 n_2 r_2},$$

since all cores located left of the core to be optimized are left orthonormal and all cores located right are right orthonormal. In fact, with (4.6), a normalized choice of $\mathbf{x}_{\min}^{(2)}$ yields

$$P_{2,1}(\widetilde{\mathbf{V}}^{(2)})^\top P_{2,1}(\widetilde{\mathbf{V}}^{(2)}) = \begin{array}{ccccccccccc} \mathbf{V}^{(1)} & \overset{r_1}{\text{---}} & \widetilde{\mathbf{V}}^{(2)} & \overset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \overset{r_3}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \\ \big| & n_1 & & \big| & n_2 & \big| & n_3 & & & \big| & n_{d-1} & \big| & n_d \\ \mathbf{V}^{(1)} & \underset{r_1}{\text{---}} & \widetilde{\mathbf{V}}^{(2)} & \underset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \underset{r_3}{\text{---}} & \cdots & \underset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \underset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \end{array}$$

$$= r_1 \left[\begin{array}{c} \widetilde{\mathbf{V}}^{(2)} \\ \big| \; n_2 \\ \widetilde{\mathbf{V}}^{(2)} \end{array}\right] r_2 = \|\widetilde{\mathbf{V}}^{(2)}\|^2 = \|\mathbf{x}_{\min}^{(2)}\|^2 = 1,$$

and for $\lambda_{\min}^{(2)}$, which is both the minimum of the micro eigenproblem as well as the current Rayleigh quotient of the overall problem $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ in $\mathbb{R}^{n_1 \cdots n_d}$, it holds

$$\lambda_{\min}^{(2)} = \rho_{\mathbf{A}}\big(P_{2,1}(\widetilde{\mathbf{V}}^{(2)})\big) = \big(P_{2,1}(\widetilde{\mathbf{V}}^{(2)})\big)^\top \mathbf{A}\big(P_{2,1}(\widetilde{\mathbf{V}}^{(2)})\big)$$

$$= \begin{array}{ccccccccccc} \mathbf{V}^{(1)} & \overset{r_1}{\text{---}} & \widetilde{\mathbf{V}}^{(2)} & \overset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \overset{r_3}{\text{---}} & \cdots & \overset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \overset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \\ \big| & n_1 & \big| & n_2 & \big| & n_3 & & & \big| & n_{d-1} & \big| & n_d \\ \mathbf{A}^{(1)} & \overset{r_1^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(2)} & \overset{r_2^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(3)} & \overset{r_3^{\mathbf{A}}}{\text{---}} & \cdots & \overset{r_{d-2}^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(d-1)} & \overset{r_{d-1}^{\mathbf{A}}}{\text{---}} & \mathbf{A}^{(d)} \\ \big| & n_1 & \big| & n_2 & \big| & n_3 & & & \big| & n_{d-1} & \big| & n_d \\ \mathbf{V}^{(1)} & \underset{r_1}{\text{---}} & \widetilde{\mathbf{V}}^{(2)} & \underset{r_2}{\text{---}} & \mathbf{V}^{(3)} & \underset{r_3}{\text{---}} & \cdots & \underset{r_{d-2}}{\text{---}} & \mathbf{V}^{(d-1)} & \underset{r_{d-1}}{\text{---}} & \mathbf{V}^{(d)} \end{array}.$$

Again we decompose

$$\underset{\longleftarrow}{\widetilde{\mathbf{V}}^{(2)}} = \mathbf{Q}^{(2)}\mathbf{R}^{(2)}$$

and overwrite the core at site 2 by $\mathbf{V}^{(2)} := \operatorname{resh}_{r_1 \times n_2 \times r_2}(\mathbf{Q}^{(2)})$ to prepare the optimization at site 3.

This interplay in a micro iteration step of solving a local eigenvalue problem of size $r_{j-1}n_j r_j \times r_{j-1}n_j r_j$, updating the core at site j by the orthonormal part of the unfolded normalized eigenvector associated with the minimal eigenvalue, and moving on to the next site is repeated until the right end of the TT tensor at site d is reached. It is referred to as a *half sweep* (from left to right). The same procedure is subsequently performed in the opposite direction from right to left, hence from site d to site 1, and again the left orthonormality of all cores left and the right orthonormality of all cores right to the core being currently

optimized is maintained in each step. This second half sweep completes one full sweep which in turn may be iterated itself, explaining the name "alternating linear scheme".

During the whole sweeping procedure, the TT ranks r_j of $\mathbf{V}$ are not changed. Since the TT ranks of $\mathrm{tens}_{n_1\times\cdots\times n_d}(\mathbf{v}_{\min})$, or at least of a good approximation, are not known a priori in general, this might have a negative effect on the overall convergence. If rank adaptivity is desired, the method may be altered to the *modified alternating linear scheme MALS* insomuch as in one micro step *two* neighboring cores at site j and $j+1$ are regarded as being contracted, and a local eigenproblem of size $r_{j-1}n_jn_{j+1}r_{j+1}\times r_{j-1}n_jn_{j+1}r_{j+1}$ is considered. A reshaped normalized eigenvector $\widetilde{\mathbf{V}}^{(j,j+1)} := \mathrm{resh}_{r_{j-1}\times n_j\times n_{j+1}\times r_{j+1}}(\mathbf{x}_{\min}^{(j,j+1)})$ subject to

$$\widetilde{\mathbf{V}}^{(j,j+1)} = \arg\min\{\rho_{\mathbf{A}}\circ P_{j,2}(\mathbf{X}^{(j,j+1)})\colon \mathbf{X}^{(j,j+1)}\in\mathbb{R}^{r_{j-1}\times n_j\times n_{j+1}\times r_{j+1}}\},$$

where

$$\begin{aligned}&P_{j,2} := P_{j,2,\mathbf{V}}\colon \mathbb{R}^{r_{j-1}\times n_j\times n_{j+1}\times r_{j+1}}\to\mathbb{R}^{n_1\times\cdots\times n_d},\\ &P_{j,2}(\mathbf{X}^{(j,j+1)}) = \Big(\mathbf{V}^{(1)}(i_1)\cdots\mathbf{V}^{(j-1)}(i_{j-1})\mathbf{X}^{(j,j+1)}(i_j,i_{j+1})\mathbf{V}^{(j+2)}(i_{j+2})\\ &\qquad\qquad\cdots\mathbf{V}^{(d)}(i_d)\Big)_{1\le i_j\le n_j,\,1\le j\le d}\end{aligned}$$

is a *two-component retraction operator*, is in turn reshaped into a matrix and decomposed by an SVD,

$$\mathrm{resh}_{r_{j-1}n_j\times n_{j+1}r_{j+1}}(\widetilde{\mathbf{V}}^{(j,j+1)}) = \mathbf{Y}^{(j,j+1)}\mathbf{\Sigma}^{(j,j+1)}(\mathbf{Z}^{(j,j+1)})^\top,$$

and we may truncate these singular values due to some rule. Hence, if we keep only the $\widetilde{r}$ largest singular values, we update, assuming that the current half sweep is left-to-right, the core at site j by the left orthonormal

$$\mathbf{V}^{(j)} := \mathrm{resh}_{r_{j-1}\times n_j\times\widetilde{r}}((\mathbf{Y}^{(j,j+1)})_{:,1:\widetilde{r}}),$$

therefore overwriting $r_j := \widetilde{r}$, and continue with the site pair $(j+1, j+2)$.

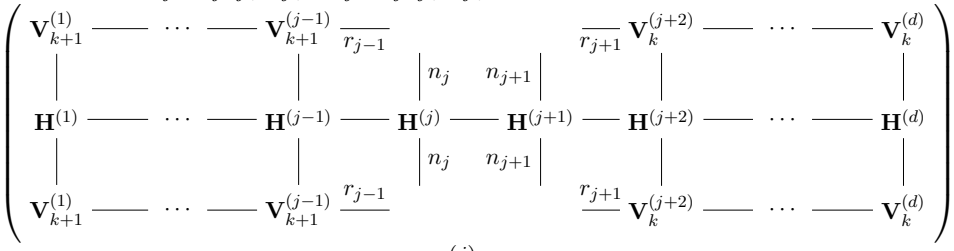
We summarize the MALS, also called two-site DMRG, in Algorithm 4.3 and change the variable $\mathbf{A}$ into $\mathbf{H}$ in order to better match our objective application.

Algorithm 4.3 MALS for eigenvalue problems

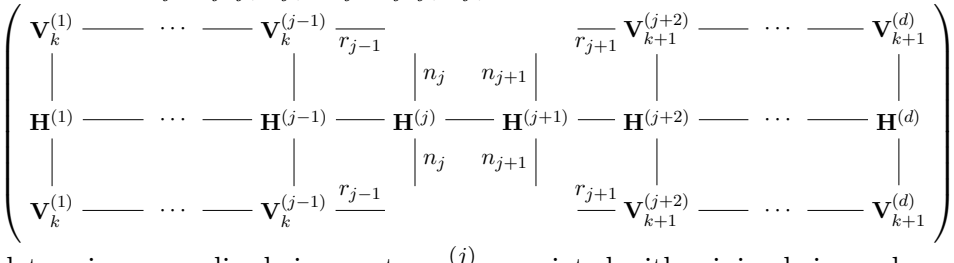
Input: TT representative $[\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(d)}]$ of a tensorized symmetric $\mathbf{H} \in \mathbb{R}^{n_1 \cdots n_d \times n_1 \cdots n_d}$
TT representative $[\mathbf{V}_0^{(1)}, \dots, \mathbf{V}_0^{(d)}]$ with TT ranks r_j of an initial guess $\mathbf{V}_0$
number of half sweeps $k_{\max}$
strategy to determine truncation parameter $r_k^{(j)}$ for the TT ranks

Output: TT representative $[\mathbf{V}_{k_{\max}}^{(1)}, \dots, \mathbf{V}_{k_{\max}}^{(d)}]$ of an approximation $\mathbf{V}_{k_{\max}}$ of a tensorized normalized eigenvector associated with the minimal eigenvalue of $\mathbf{H}$

for $j = d, \dots, 2$ **do** % initial right orthonormalization
  determine QR decomposition $\operatorname{resh}_{r_j n_j \times r_{j-1}}(\operatorname{perm}_{3,2,1}(\mathbf{V}_0^{(j)})) =: \mathbf{Q}^{(j)}\mathbf{R}^{(j)}$
  $\mathbf{V}_0^{(j)} := \operatorname{resh}_{r_{j-1} \times n_j \times r_j}((\mathbf{Q}^{(j)})^\top)$
  $\mathbf{V}_0^{(j-1)} := \mathbf{V}_0^{(j-1)} \boxdot_3^1 (\mathbf{R}^{(j)})^\top$
end for
for $k = 0, \dots, k_{\max} - 1$ **do**
  if $k \equiv 0 \mod 2$ **then** % left-to-right half sweep
    if $k = k_{\max} - 1$ **then** $\widehat{j} := d - 1$ **else** $\widehat{j} := d - 2$ **end if**
    for $j = 1, 2, \dots, \widehat{j}$ **do**
      $\widetilde{\mathbf{H}}_k^{(j)} := \operatorname{resh}_{r_{j-1}n_j n_{j+1} r_{j+1} \times r_{j-1} n_j n_{j+1} r_{j+1}}$



      determine normalized eigenvector $\mathbf{x}_k^{(j)}$ associated with minimal eigenvalue of $\widetilde{\mathbf{H}}_k^{(j)}$
      determine SVD $\operatorname{resh}_{r_{j-1}n_j \times n_{j+1} r_{j+1}}(\mathbf{x}_k^{(j)}) =: \mathbf{Y}^{(j)}\mathbf{\Sigma}^{(j)}(\mathbf{Z}^{(j)})^\top$
      determine truncation parameter $r_k^{(j)}$ and set $r_j := r_k^{(j)}$
      $\mathbf{V}_{k+1}^{(j)} := \operatorname{resh}_{r_{j-1} \times n_j \times r_j}((\mathbf{Y}^{(j)})_{:,1:r_j})$
    end for
    if $k = k_{\max} - 1$ **then** $\mathbf{V}_{k+1}^{(d)} := (\mathbf{\Sigma}^{(d-1)})_{1:r_{d-1},1:r_{d-1}}((\mathbf{Z}^{(d-1)})^\top)_{1:r_{d-1},:}$ **end if**
  else % right-to-left half sweep
    if $k = k_{\max} - 1$ **then** $\widehat{j} := 1$ **else** $\widehat{j} := 2$ **end if**
    for $j = d - 1, d - 2, \dots, \widehat{j}$ **do**
      $\widetilde{\mathbf{H}}_k^{(j)} := \operatorname{resh}_{r_{j-1}n_j n_{j+1} r_{j+1} \times r_{j-1} n_j n_{j+1} r_{j+1}}$



      determine normalized eigenvector $\mathbf{x}_k^{(j)}$ associated with minimal eigenvalue of $\widetilde{\mathbf{H}}_k^{(j)}$
      determine SVD $\operatorname{resh}_{r_{j-1}n_j \times n_{j+1} r_{j+1}}(\mathbf{x}_k^{(j)}) =: \mathbf{Y}^{(j)}\mathbf{\Sigma}^{(j)}(\mathbf{Z}^{(j)})^\top$
      determine truncation parameter $r_k^{(j)}$ and set $r_j := r_k^{(j)}$
      $\mathbf{V}_{k+1}^{(j+1)} := \operatorname{resh}_{r_j \times n_{j+1} \times r_{j+1}}(((\mathbf{Z}^{(j)})^\top)_{1:r_j,:})$
    end for
    if $k = k_{\max} - 1$ **then** $\mathbf{V}_{k+1}^{(1)} := (\mathbf{Y}^{(1)})_{:,1:r_1}(\mathbf{\Sigma}^{(1)})_{1:r_1,1:r_1}$ **end if**
  end if
end for

5. Construction of an initial guess

Most iterative numerical algorithms need as an input a point where the iteration starts. This is what we call an *initial guess*. In many cases the choice of an initial guess strongly influences the behavior of the iteration, but a detailed analysis of this dependence often is quite difficult.

A common strategy is to choose the initial guess randomly. For our problem at hand, namely computing an eigenvector associated with the minimal eigenvalue of a Hamilton operator $\mathbf{H}_{d,q} \in \mathbb{R}^{q^d \times q^d}$, and the algorithms we employ, this random choice of the initial guess is in principle possible but there might be better ways.

In the present chapter we discuss possibilities to improve the performance of an iterative eigensolver by a more sophisticated choice of an initial guess. As investigated in Chapter 6, this improvement in many cases results in a reduced number of iteration steps necessary until convergence. We can assume that this necessary number of iteration steps gets smaller if the initial guess is in some sense "closer" to the exact solution. So, the number of iteration steps may be significantly reduced if the initial guess is already an approximation of the sought eigenvector instead of some random vector. We will restrict ourselves to the cases $q = 2$ and $q = 3$ and generally assume that the minimal eigenvalue is simple.

The Hamilton operator $\mathbf{H}_{d,q}$ of the XYZ respectively the Potts model with given coupling parameters is defined for each $d \in \mathbb{N}$ in an analogous way as only the number of summands and their respective number of contained Kronecker factors varies. Despite the crucial fact that $\mathbf{H}_{d,q}$ is not just a Kronecker product of $\mathbf{H}_{d',q}$ and $\mathbf{H}_{d'',q}$ for $d' + d'' = d$, $\mathbf{H}_{d,q}$ is composed of the constituents of $\mathbf{H}_{d',q}$ and $\mathbf{H}_{d'',q}$ plus some portion that links together the two corresponding subsystems of particles.

In the context of multigrid methods, see [Hac85], one encounters the situation of several resolutions of a problem on several nested discretizing grids, from the finest to the coarsest grid, which are related by *restriction* and *prolongation* operators. Typically, the solution is more expensive for a larger, here finer, problem than for a smaller, here coarser, one. Hence, an inherent idea of multigrid methods is to solve the coarsest problem up to an arbitrary accuracy at low cost, then to prolongate this solution to a finer grid respectively larger problem size, and to continue the computation for this finer problem based on the prolongated solution of the coarser/smaller problem.

In our opinion, there is a certain analogy to the characteristics of the problem we are concerned with in this thesis. Therefore, in order to construct a suitable starting point of an iteration, we opt for translating the general strategy of prolongating information from the small to the large also to our problem of computing an eigenvector associated with the minimal eigenvalue of $\mathbf{H}_{d,q}$.

For small d, say for a one-digit d, we are able to compute the eigenvectors up to an arbitrary accuracy in a negligible amount of time. The central idea now is to use information about the eigenvectors for small d and especially the relation between eigenvectors for different small values of d to construct some approximation for larger values of d where each iteration step becomes costly, and any reduction of iteration steps would be helpful and demanded.

In Section 5.1 we detail this idea in the situation when the Hamilton operator as well

as its eigenvectors are given in full matrix/vector format. We consider coupling parameters $A \neq B$ if the Hamilton operator represents an XYZ model which differs from the case $A = B$ in terms of the sparsity patterns of the eigenvectors, cf. Section 2.1. In order to be able to apply our strategy also when computations are only feasible in a low-rank tensor format, we describe in Section 5.2 the construction of an initial guess in HT format based on a linear dimension tree, which we call for short *linear HT format*. The transition from linear to balanced HT format, i.e. with a balanced dimension tree, respectively to TT format is the subject of Section 5.3 respectively 5.4. For the structurally different situation of coupling parameters $A = B$ in the XYZ model, we discuss an alternative strategy in Section 5.5.

5.1. Full vector format

Although the proposed construction of an initial guess is in principle the same for the 2-XYZ, the 3-XYZ, and the 3-Potts model, there are some peculiarities concerning its feasibility arising from the different sparsity patterns of the eigenvectors for the three mentioned types of Hamilton operators. Therefore, we spread the discussion of the respective cases over Subsections 5.1.1 to 5.1.3.

Since it might be more intuitive in the present context where we only deal with vectors and matrices and not with higher-dimensional tensors, we introduce for a vector $\mathbf{v} \in \mathbb{R}^{mn}$ the notation

$$\mathrm{mat}_{m\times n}(\mathbf{v}) := \mathrm{resh}_{m\times n}(\mathbf{v}) \in \mathbb{R}^{m\times n}$$

and refer to it as a *matricization*, not to be confused with the same term defined in Section 3.1 which was intended to have tensors in its domain.

5.1.1. 2-XYZ model, $A \neq B$

We consider first the q-XYZ model (2.1) in the case $q = 2$ and write $\mathbf{H}_d := \mathbf{H}_{d,2}^{\mathrm{XYZ}}$. The parameters A, B, Δ, h are fixed for all d. Let us assume from now on that we are in the case $A \neq B$, so Corollary 2.10 states that the eigenvectors associated with simple eigenvalues for $d = 2$ either have a maximal nonzero pattern

$$\begin{pmatrix} * & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{\mathrm{even}} \quad \text{or} \quad \begin{pmatrix} 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{\mathrm{odd}},$$

and for $d = 3$ either

$$\begin{pmatrix} * & 0 & 0 & * & 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{\mathrm{even}} \quad \text{or} \quad \begin{pmatrix} 0 & * & * & 0 & * & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{\mathrm{odd}}.$$

Let $(v_i^{(2)})_{1\leq i\leq 2^2} := \mathbf{v}_{\min}^{(2)} \in \mathbb{R}^{2^2}$ resp. $(v_i^{(3)})_{1\leq i\leq 2^3} := \mathbf{v}_{\min}^{(3)} \in \mathbb{R}^{2^3}$ be a normalized eigenvector associated with the minimal eigenvalue $\lambda_{\min}^{(2)}$ of $\mathbf{H}_2$ resp. $\lambda_{\min}^{(3)}$ of $\mathbf{H}_3$. We assume for a moment that $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,2}^{\mathrm{even}}$ and $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,2}^{\mathrm{even}}$. The description is similar in the "odd" or in the mixed case. We construct a matrix $\mathbf{M} \in \mathbb{R}^{8\times 4}$ such that

$$\mathbf{M}\mathbf{v}_{\min}^{(2)} = \mathbf{v}_{\min}^{(3)}.$$

One could just define $\mathbf{M} := \mathbf{v}_{\min}^{(3)}(\mathbf{v}_{\min}^{(2)})^\top$ but it turns out that for our purposes, it is better to make the ansatz

$$\mathbf{M} = \mathbf{I}_2 \otimes \mathbf{N}$$

with some $\mathbf{N} \in \mathbb{R}^{4\times 2}$. This particular ansatz is possible due to the sparsity pattern of the eigenvectors, as the first column of $\mathbf{N}$ may be constructed to map the first entry $v_1^{(2)}$ of $\mathbf{v}_{\min}^{(2)}$ to the upper half $(v_1^{(3)}, \dots, v_4^{(3)})^\top$ of $\mathbf{v}_{\min}^{(3)}$ while the second column of $\mathbf{N}$ maps the last entry $v_4^{(2)}$ of $\mathbf{v}_{\min}^{(2)}$ to the lower half $(v_5^{(3)}, \dots, v_8^{(3)})^\top$ of $\mathbf{v}_{\min}^{(3)}$. This means

$$\mathbf{N} = \begin{pmatrix} v_1^{(3)} & v_5^{(3)} \\ v_2^{(3)} & v_6^{(3)} \\ v_3^{(3)} & v_7^{(3)} \\ v_4^{(3)} & v_8^{(3)} \end{pmatrix} \begin{pmatrix} v_1^{(2)} & v_3^{(2)} \\ v_2^{(2)} & v_4^{(2)} \end{pmatrix}^{-1} = \mathrm{mat}_{4\times 2}(\mathbf{v}_{\min}^{(3)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1}. \tag{5.1}$$

The reshaped eigenvector $\mathbf{v}_{\min}^{(2)}$, namely $\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})$, is invertible since $v_1^{(2)} \neq 0 \neq v_4^{(2)}$ by Lemma C.1(i) and $v_2^{(2)} = 0 = v_3^{(2)}$. In case that the eigenspace of $\lambda_{\min}^{(2)}$ is contained in $\mathcal{E}_{2,2}^{\mathrm{odd}}$, then by Lemma C.1(ii) there may also be chosen an eigenvector $\mathbf{v}_{\min}^{(2)}$ such that $\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})$ is invertible.

The matrix $\mathbf{M}$ contains information on how $\mathbf{v}_{\min}^{(2)}$ is related to $\mathbf{v}_{\min}^{(3)}$. This information is now used to construct a vector $\widetilde{\mathbf{v}}^{(4)} \in \mathbb{R}^{2^4}$ which corresponds to $d = 4$. We map the two blocks

$$\begin{pmatrix} * \\ 0 \\ 0 \\ * \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ * \\ * \\ 0 \end{pmatrix}$$

of $\mathbf{v}_{\min}^{(3)}$ by $\mathbf{M}$ and then compose the two resulting vectors of length 8 into one vector $\widetilde{\mathbf{v}}^{(4)} \in \mathbb{R}^{16}$. In other words,

$$\widetilde{\mathbf{v}}^{(4)} := (\mathbf{I}_2 \otimes \mathbf{M})\mathbf{v}_{\min}^{(3)}$$

and we call this a *prolongation* of the eigenvector $\mathbf{v}_{\min}^{(3)}$. This may be interpreted as treating the two blocks of $\mathbf{v}_{\min}^{(3)}$ as ground states in $d = 2$ and applying on them the particular rule expressed by $\mathbf{M}$ that maps a ground state in $d = 2$ to a ground state in $d = 3$.

Omitting the assumption $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,2}^{\mathrm{even}}$ and $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,2}^{\mathrm{even}}$, the sparsity structure of the vector $\widetilde{\mathbf{v}}^{(4)}$ in the general case depends on that of $\mathbf{v}_{\min}^{(2)}$ and $\mathbf{v}_{\min}^{(3)}$. If we take for example $A = 1$, $B = \Delta = 0$, $h = 1$ which corresponds to the Ising model, we have that $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,2}^{\mathrm{even}}$ and $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,2}^{\mathrm{odd}}$. Then, due to the resulting sparsity pattern of

$$\begin{aligned} \mathbf{N} &= \mathrm{mat}_{4\times 2}(\mathbf{v}_{\min}^{(3)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1} \\ &\approx \begin{pmatrix} 0 & -0.1229 \\ -0.1229 & 0 \\ 0.0298 & 0 \\ 0 & 0.9843 \end{pmatrix} \begin{pmatrix} 0.1222 & 0 \\ 0 & -0.9925 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0 & 0.1239 \\ -1.0061 & 0 \\ 0.2439 & 0 \\ 0 & -0.9918 \end{pmatrix}, \end{aligned}$$

we obtain $\widetilde{\mathbf{v}}^{(4)} \in \mathcal{E}_{4,2}^{\mathrm{even}}$ in agreement with the exact eigenvector $\mathbf{v}_{\min}^{(4)}$. We come back to this issue in Theorem 5.3.

In the same way interesting is the quality of the prolongated vector $\widetilde{\mathbf{v}}^{(4)}$ in the sense that it approximates $\mathbf{v}_{\min}^{(4)}$. After normalizing (it is $\|\widetilde{\mathbf{v}}^{(4)}\| \approx 1.000039$) we get

$$\left\| \widetilde{\mathbf{v}}^{(4)} / \|\widetilde{\mathbf{v}}^{(4)}\| - \mathbf{v}_{\min}^{(4)} \right\| \approx 2.0 \cdot 10^{-3}.$$

Considering instead the Rayleigh quotient (it is $\lambda_{\min}^{(4)} \approx -2.0942$) we even obtain

$$\frac{(\widetilde{\mathbf{v}}^{(4)})^\top \mathbf{H}_4 \widetilde{\mathbf{v}}^{(4)}}{(\widetilde{\mathbf{v}}^{(4)})^\top \widetilde{\mathbf{v}}^{(4)}} - \lambda_{\min}^{(4)} \approx 7.1 \cdot 10^{-6}.$$

The matrix $\mathbf{M}$ may now be used to create approximations to the "minimal" eigenvector also for larger values of d. With the help of $\mathbf{M}$, we just prolongate the approximation $\widetilde{\mathbf{v}}^{(4)}$ to a vector $\widetilde{\mathbf{v}}^{(5)} \in \mathbb{R}^{32}$ which corresponds to $d = 5$. Again, the idea is to map the (now four) blocks of length four of $\widetilde{\mathbf{v}}^{(4)}$ by $\mathbf{M}$ and to compose the resulting vectors of length eight into one vector $\widetilde{\mathbf{v}}^{(5)} \in \mathbb{R}^{32}$. That means

$$\widetilde{\mathbf{v}}^{(5)} := (\mathbf{I}_{2^2} \otimes \mathbf{M})\widetilde{\mathbf{v}}^{(4)} = (\mathbf{I}_{2^2} \otimes \mathbf{M})(\mathbf{I}_2 \otimes \mathbf{M})\mathbf{v}_{\min}^{(3)}.$$

After normalizing (it is $\|\widetilde{\mathbf{v}}^{(5)}\| \approx 1.000080$) we get

$$\left\| \widetilde{\mathbf{v}}^{(5)} / \|\widetilde{\mathbf{v}}^{(5)}\| - \mathbf{v}_{\min}^{(5)} \right\| \approx 3.5 \cdot 10^{-3}.$$

Considering instead the Rayleigh quotient (it is $\lambda_{\min}^{(5)} \approx -2.6260$) we obtain

$$\frac{(\widetilde{\mathbf{v}}^{(5)})^\top \mathbf{H}_5 \widetilde{\mathbf{v}}^{(5)}}{(\widetilde{\mathbf{v}}^{(5)})^\top \widetilde{\mathbf{v}}^{(5)}} - \lambda_{\min}^{(5)} \approx 2.0 \cdot 10^{-5}.$$

In principle, this prolongation strategy may be used to construct approximations $\widetilde{\mathbf{v}}^{(d)}$ of eigenvectors $\mathbf{v}_{\min}^{(d)}$ for arbitrary d via

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{d-3} (\mathbf{I}_{2^i} \otimes \mathbf{M}) \right) \mathbf{v}_{\min}^{(3)},$$

but in practice d is limited by the available memory. The matrices $\mathbf{I}_{2^i} \otimes \mathbf{M}$ do not have to be stored explicitly, but even this cannot avoid the exponential growth in d of $\widetilde{\mathbf{v}}^{(d)}$. Nevertheless, still considering as an example $A = 1$, $B = \Delta = 0$, $h = 1$, we get for $d = 16$ as the absolute error in the Rayleigh quotient (with $\lambda_{\min}^{(16)} \approx -8.4755$)

$$\frac{(\widetilde{\mathbf{v}}^{(16)})^\top \mathbf{H}_{16} \widetilde{\mathbf{v}}^{(16)}}{(\widetilde{\mathbf{v}}^{(16)})^\top \widetilde{\mathbf{v}}^{(16)}} - \lambda_{\min}^{(16)} \approx 1.8 \cdot 10^{-4}. \tag{5.2}$$

For $d = 22$ resp. $d = 23$ which is actually the largest even resp. odd d tractable for us when storing $\mathbf{H}_d$ as a sparse matrix, we obtain (with $\lambda_{\min}^{(22)} \approx -11.6661$, $\lambda_{\min}^{(23)} \approx -12.1979$)

$$\frac{(\widetilde{\mathbf{v}}^{(22)})^\top \mathbf{H}_{22} \widetilde{\mathbf{v}}^{(22)}}{(\widetilde{\mathbf{v}}^{(22)})^\top \widetilde{\mathbf{v}}^{(22)}} - \lambda_{\min}^{(22)} \approx 2.7 \cdot 10^{-4} \tag{5.3}$$

respectively

$$\frac{(\widetilde{\mathbf{v}}^{(23)})^\top \mathbf{H}_{23} \widetilde{\mathbf{v}}^{(23)}}{(\widetilde{\mathbf{v}}^{(23)})^\top \widetilde{\mathbf{v}}^{(23)}} - \lambda_{\min}^{(23)} \approx 2.8 \cdot 10^{-4}. \tag{5.4}$$

This is quite remarkable since we constructed $\widetilde{\mathbf{v}}^{(d)}$ with this high approximation quality for $d = 22$ or $d = 23$ only with knowledge that we extracted from the cases $d = 2$ and $d = 3$.

For the difference between the smallest and the second smallest eigenvalue it holds

$$\lambda_2^{(d)} - \lambda_{\min}^{(d)} \approx \begin{cases} 0.5150, & d = 16 \\ 0.5085, & d = 22 \\ 0.5078, & d = 23 \end{cases},$$

which as an example for $d \in \{16, 22, 23\}$ implies that

$$\lambda_{\min}^{(d)} < \frac{(\widetilde{\mathbf{v}}^{(d)})^\top \mathbf{H}_d \widetilde{\mathbf{v}}^{(d)}}{(\widetilde{\mathbf{v}}^{(d)})^\top \widetilde{\mathbf{v}}^{(d)}} < \lambda_2^{(d)}.$$

So, when starting an iteration where descent in the Rayleigh quotient in each step is guaranteed with $\widetilde{\mathbf{v}}^{(d)}$ as initial guess, there is no danger that the iteration gets stuck in a local minimum larger than the smallest eigenvalue, at least in the current example case.

Remark 5.1. In each scenario considered in this thesis, the minimal eigenvalue $\lambda_{\min}^{(d)}$ is negative with absolute value roughly around d, cf. Proposition 2.11(iv)-(v) and Proposition 2.19 for the special cases as well as the concrete values and the figures in the present section. Indeed, $\lambda_{\min}^{(d)}$ is scaled by the dominant value of A, B, Δ, h, but this value is around 1 in our tests. Hence $|\lambda_{\min}^{(d)}|$ is approximately of order 10^1, so a relative error with respect to $\lambda_{\min}^{(d)}$ and the corresponding absolute error are quite similar when considering them logarithmically in a usual range between 1 and the machine epsilon. Therefore we document in the sequel only absolute errors, especially in Chapter 6 about the various numerical tests.

The described prolongation strategy may also be built up of $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ for $d_1 < d_2$ larger than 2 or 3 when information about both of them is still easily available. In case $(d_1, d_2) = (2, 4)$, the matrix $\mathbf{M} = \mathbf{I}_2 \otimes \mathbf{N} \in \mathbb{R}^{16\times 4}$ with

$$\mathbf{M}\mathbf{v}_{\min}^{(2)} = \mathbf{v}_{\min}^{(4)}$$

is defined by, with example values for $A = 1$, $B = \Delta = 0$, $h = 1$,

$$\begin{aligned} \mathbf{N} &:= \mathrm{mat}_{8\times 2}(\mathbf{v}_{\min}^{(4)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1} \qquad (5.5) \\ &\approx \begin{pmatrix} 0.0155 & 0 \\ 0 & -0.1238 \\ 0 & 0.0304 \\ -0.1220 & 0 \\ 0 & -0.1220 \\ 0.0304 & 0 \\ -0.0091 & 0 \\ 0 & 0.9761 \end{pmatrix} \begin{pmatrix} 0.1222 & 0 \\ 0 & -0.9925 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.1265 & 0 \\ 0 & 0.1247 \\ 0 & -0.0307 \\ -0.9985 & 0 \\ 0 & 0.1229 \\ 0.2490 & 0 \\ -0.0746 & 0 \\ 0 & -0.9835 \end{pmatrix}. \end{aligned}$$

So we may construct approximations $\widetilde{\mathbf{v}}^{(d)}$ of eigenvectors $\mathbf{v}_{\min}^{(d)}$ for arbitrary *even* d via

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{\frac{d-4}{2}} (\mathbf{I}_{2^{2i}} \otimes \mathbf{M}) \right) \mathbf{v}_{\min}^{(4)}$$

as d is increased by 2 in each prolongation step. Determining $\widetilde{\mathbf{v}}^{(16)}$ and $\widetilde{\mathbf{v}}^{(22)}$ this way, we obtain

$$\frac{(\widetilde{\mathbf{v}}^{(16)})^\top \mathbf{H}_{16} \widetilde{\mathbf{v}}^{(16)}}{(\widetilde{\mathbf{v}}^{(16)})^\top \widetilde{\mathbf{v}}^{(16)}} - \lambda_{\min}^{(16)} \approx 8.3 \cdot 10^{-5} \qquad (5.6)$$

and

$$\frac{(\widetilde{\mathbf{v}}^{(22)})^\top \mathbf{H}_{22}\widetilde{\mathbf{v}}^{(22)}}{(\widetilde{\mathbf{v}}^{(22)})^\top \widetilde{\mathbf{v}}^{(22)}} - \lambda_{\min}^{(22)} \approx 1.3 \cdot 10^{-4}, \tag{5.7}$$

which is a bit smaller than the respective Rayleigh quotient errors (5.2) and (5.3) resulting from $(d_1, d_2) = (2, 3)$. Still with $d_1 = 2$, for arbitrary d_2 the definition (5.5) of $\mathbf{N}$ reads

$$\mathbf{N} := \mathrm{mat}_{2^{d_2-1}\times 2}(\mathbf{v}_{\min}^{(d_2)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1}$$

and successive prolongation steps $i = 1, 2, 3, \ldots$ governed by the application of

$$\mathbf{I}_{2^{(d_2-2)i}} \otimes (\mathbf{I}_2 \otimes \mathbf{N}) \tag{5.8}$$

increase d in each step by $d_2 - 2$.

For both $d_2 > d_1$ arbitrary with d_1 even which is necessary to obtain a quadratic invertible matricization $\mathrm{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})$, we generalize the definition of $\mathbf{N}$ via

$$\mathbf{N} := \mathrm{mat}_{2^{d_2-d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_2)}) \left(\mathrm{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\right)^{-1} \tag{5.9}$$

and set

$$\mathbf{M} := \mathbf{I}_{2^{d_1/2}} \otimes \mathbf{N} \tag{5.10}$$

which yields

$$\mathbf{M}\mathbf{v}_{\min}^{(d_1)} = \mathbf{v}_{\min}^{(d_2)}. \tag{5.11}$$

Provided $d - d_2$ is divisible by $d_2 - d_1$, we may then construct a prolongated vector

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{(d-d_2)/(d_2-d_1)} (\mathbf{I}_{2^{(d_2-d_1)i}} \otimes \mathbf{M})\right) \mathbf{v}_{\min}^{(d_2)}, \tag{5.12}$$

and we notice that the prolongation technique (5.12) is not restricted to even d_1 once a matrix $\mathbf{M}$ satisfying (5.11) is given.

If we choose $d_1 = 3$ we have to adjust the definition of $\mathbf{M}$ slightly as d_1 is odd and a quadratic matricization of $\mathbf{v}_{\min}^{(3)}$ is not possible. In order that

$$\mathbf{M}\mathbf{v}_{\min}^{(3)} = \mathbf{v}_{\min}^{(d_2)},$$

the idea is not to impose $\mathbf{M} = \mathbf{I}_{2^j} \otimes \mathbf{N}$ for some j but instead to split the problem in two subproblems resembling the case $d_1 = 2$. Therefore we make the ansatz to map the upper resp. lower half part of $\mathbf{v}_{\min}^{(3)}$ to the upper resp. lower half part of $\mathbf{v}_{\min}^{(d_2)}$. Hence, exploiting that the sparsity patterns also occur in the subproblems and assuming invertibility of the relevant matricizations, we set

$$\mathbf{M} := \begin{pmatrix} \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{I})} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{II})} \end{pmatrix}, \tag{5.13a}$$

$$\mathbf{N}^{(\mathrm{I})} := \mathrm{mat}_{2^{d_2-2}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{1:2^{d_2-1}}\big) \left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{1:4}\big)\right)^{-1}, \tag{5.13b}$$

$$\mathbf{N}^{(\mathrm{II})} := \mathrm{mat}_{2^{d_2-2}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{2^{d_2-1}+1:2^{d_2}}\big) \left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{5:8}\big)\right)^{-1}. \tag{5.13c}$$

We may then construct $\widetilde{\mathbf{v}}^{(d)}$ via (5.12) setting $d_1 = 3$, provided $d - d_2$ is divisible by $d_2 - 3$. In the test case $A = 1$, $B = \Delta = 0$, $h = 1$ with $d_2 = 4$ this yields

$$\frac{(\widetilde{\mathbf{v}}^{(16)})^\top \mathbf{H}_{16} \widetilde{\mathbf{v}}^{(16)}}{(\widetilde{\mathbf{v}}^{(16)})^\top \widetilde{\mathbf{v}}^{(16)}} - \lambda_{\min}^{(16)} \approx 8.1 \cdot 10^{-5}$$

and

$$\frac{(\widetilde{\mathbf{v}}^{(22)})^\top \mathbf{H}_{22} \widetilde{\mathbf{v}}^{(22)}}{(\widetilde{\mathbf{v}}^{(22)})^\top \widetilde{\mathbf{v}}^{(22)}} - \lambda_{\min}^{(22)} \approx 1.2 \cdot 10^{-4},$$

which is again a bit smaller than the respective Rayleigh quotient errors (5.6) and (5.7) resulting from $d_1 = 2$, $d_2 = 4$. Also

$$\frac{(\widetilde{\mathbf{v}}^{(23)})^\top \mathbf{H}_{23} \widetilde{\mathbf{v}}^{(23)}}{(\widetilde{\mathbf{v}}^{(23)})^\top \widetilde{\mathbf{v}}^{(23)}} - \lambda_{\min}^{(23)} \approx 1.3 \cdot 10^{-4} \tag{5.14}$$

in turn is smaller than (5.4).

If $d_1 = 4$ is even we may apply (5.9)-(5.12) and with $d_2 = 5$ we obtain

$$\frac{(\widetilde{\mathbf{v}}^{(23)})^\top \mathbf{H}_{23} \widetilde{\mathbf{v}}^{(23)}}{(\widetilde{\mathbf{v}}^{(23)})^\top \widetilde{\mathbf{v}}^{(23)}} - \lambda_{\min}^{(23)} \approx 3.1 \cdot 10^{-7}. \tag{5.15}$$

Instead we could use the idea from the case $d_1 = 3$ and split the vectors $\mathbf{v}_{\min}^{(4)}$ and $\mathbf{v}_{\min}^{(d_2)}$. In contrast to $d_1 = 3$ we split both vectors in four parts of equal size since now $d_1 - 2$ is even. Hence

$$\mathbf{M} := \begin{pmatrix} \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{I})} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{II})} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{III})} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{IV})} \end{pmatrix},$$

$$\begin{aligned}
\mathbf{N}^{(\mathrm{I})} &:= \mathrm{mat}_{2^{d_2-3}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{1:2^{d_2-2}}\big)\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(4)})_{1:4}\big)\right)^{-1}, \\
\mathbf{N}^{(\mathrm{II})} &:= \mathrm{mat}_{2^{d_2-3}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{2^{d_2-2}+1:2^{d_2-1}}\big)\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(4)})_{5:8}\big)\right)^{-1}, \\
\mathbf{N}^{(\mathrm{III})} &:= \mathrm{mat}_{2^{d_2-3}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{2^{d_2-1}+1:2^{d_2-1}+2^{d_2-2}}\big)\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(4)})_{9:12}\big)\right)^{-1}, \\
\mathbf{N}^{(\mathrm{IV})} &:= \mathrm{mat}_{2^{d_2-3}\times 2}\big((\mathbf{v}_{\min}^{(d_2)})_{2^{d_2-1}+2^{d_2-2}+1:2^{d_2}}\big)\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(4)})_{13:16}\big)\right)^{-1}.
\end{aligned}$$

With this $\mathbf{M}$ and $\widetilde{\mathbf{v}}^{(d)}$ defined by (5.12) for $d_2 = 5$, we obtain

$$\frac{(\widetilde{\mathbf{v}}^{(23)})^\top \mathbf{H}_{23} \widetilde{\mathbf{v}}^{(23)}}{(\widetilde{\mathbf{v}}^{(23)})^\top \widetilde{\mathbf{v}}^{(23)}} - \lambda_{\min}^{(23)} \approx 2.2 \cdot 10^{-4}. \tag{5.16}$$

In fact this error in the Rayleigh quotient is larger than (5.15), so it appears advantageous to construct the matrix $\mathbf{M}$ governing the prolongation by regarding the exact eigenvectors as a whole rather than to utilize relations between the corresponding blocks of a splitted vector. Additionally we note that (5.16) is larger than

$$\frac{(\widetilde{\mathbf{v}}^{(23)})^\top \mathbf{H}_{23} \widetilde{\mathbf{v}}^{(23)}}{(\widetilde{\mathbf{v}}^{(23)})^\top \widetilde{\mathbf{v}}^{(23)}} - \lambda_{\min}^{(23)} \approx 7.2 \cdot 10^{-5}$$

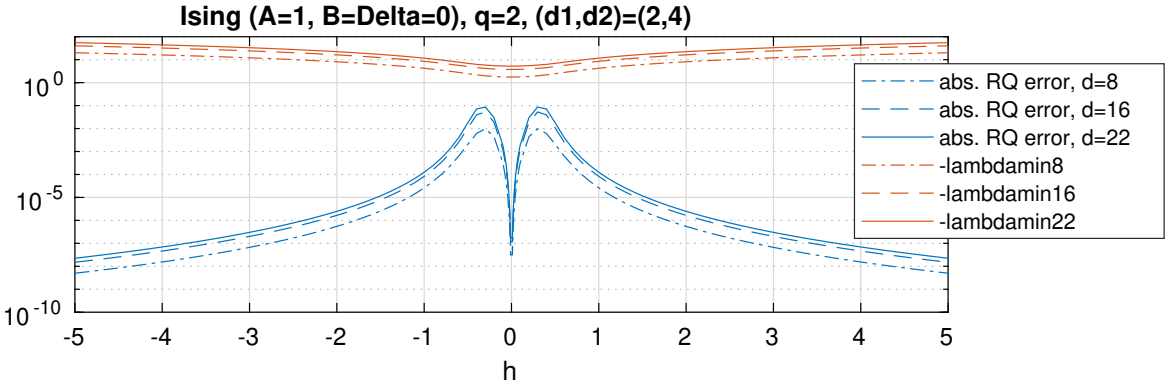


Figure 5.1.: Approximation quality of prolongated vector for 2-Ising model with varying $h \in [-5, -0.01] \cup [0.01, 5]$, value of $-\lambda_{\min}^{(d)}$.

if $\widetilde{\mathbf{v}}^{(23)}$ is constructed based on $(d_1, d_2) = (3, 5)$ using (5.13) and (5.12), and even slightly larger than (5.14) obtained from $(d_1, d_2) = (3, 4)$.

So far, to demonstrate the quality of the prolongation technique, we considered as an example the quantum Ising model ($A = 1$, $B = \Delta = 0$) with $h = 1$. In Figure 5.1 the absolute error in the Rayleigh quotient

$$\frac{(\widetilde{\mathbf{v}}^{(d)})^\top \mathbf{H}_d \widetilde{\mathbf{v}}^{(d)}}{(\widetilde{\mathbf{v}}^{(d)})^\top \widetilde{\mathbf{v}}^{(d)}} - \lambda_{\min}^{(d)}$$

is depicted for $d \in \{8, 16, 22\}$ and $h \in [-5, -0.01] \cup [0.01, 5]$. We see that this error gets small when h is close to 0. Additionally we plot the respective value of $\lambda_{\min}^{(d)}$ with changed sign in order to match the logarithmic axis. As stated in Remark 5.1, the size of $|\lambda_{\min}^{(d)}|$ is around 10^1. If $h = 0$, the minimal eigenvalue of $\mathbf{H}_d$ is $\lambda_{\min}^{(d)} = -\frac{1}{4}(d-1)$ by Proposition 2.11(iv) and has multiplicity 2. The eigenspace is spanned by the two vectors

$$\mathbf{v}_{\min,1}^{(d)} := \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\-1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\-1\end{pmatrix} \otimes \ldots \otimes \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\-1\end{pmatrix}$$

and

$$\mathbf{v}_{\min,2}^{(d)} := \begin{pmatrix}1\\-1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} \otimes \begin{pmatrix}1\\-1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix} \otimes \ldots \otimes \begin{pmatrix}1\\-1\end{pmatrix} \otimes \begin{pmatrix}1\\1\end{pmatrix},$$

where each Kronecker product has d factors. The matrices $\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min,1}^{(2)})$ and $\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min,2}^{(2)})$ are singular, but for example

$$\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min,1}^{(2)} + \mathbf{v}_{\min,2}^{(2)}) = \begin{pmatrix}2 & 0\\0 & -2\end{pmatrix}.$$

Therefore we may set

$$\mathbf{N} := \mathrm{mat}_{8\times 2}(\mathbf{v}_{\min,1}^{(4)} + \mathbf{v}_{\min,2}^{(4)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min,1}^{(2)} + \mathbf{v}_{\min,2}^{(2)})\right)^{-1} = \begin{pmatrix}1 & 0\\0 & 1\\0 & -1\\-1 & 0\\0 & 1\\1 & 0\\-1 & 0\\0 & -1\end{pmatrix}.$$

d	2	4	6	8	10	12	14
$\operatorname{cond}\left(\mathbf{V}_{\min}^{(d)}\right)$	8.1	$4.9 \cdot 10^3$	$2.1 \cdot 10^8$	$5.7 \cdot 10^{14}$	$\approx 10^{17}$	$\approx 10^{17}$	$\approx 10^{18}$
$\operatorname{rank}\left(\mathbf{V}_{\min}^{(d)}\,;\,10^{-6}\right)$	2	4	6	6	6	6	6
$\operatorname{rank}\left(\mathbf{V}_{\min}^{(d)}\,;\,10^{-12}\right)$	2	4	8	13	17	20	20
$\operatorname{rank}\left(\mathbf{V}_{\min}^{(d)}\,;\,10^{-15}\right)$	2	4	8	16	22	27	30

Table 5.1.: Condition number and ε-rank of the matricized "minimal" eigenvector $\mathbf{V}_{\min}^{(d)} := \operatorname{mat}_{2^{d/2}\times 2^{d/2}}(\mathbf{v}_{\min}^{(d)})$ for $A = h = 1$, $B = \Delta = 0$.

By this choice of $\mathbf{N}$, and setting $\mathbf{M} := \mathbf{I}_2 \otimes \mathbf{N}$, we obtain

$$\left(\prod_{i=1}^{\frac{d-4}{2}} (\mathbf{I}_{2^{2i}} \otimes \mathbf{M})\right) (\mathbf{v}_{\min,1}^{(4)} + \mathbf{v}_{\min,2}^{(4)}) = \mathbf{v}_{\min,1}^{(d)} + \mathbf{v}_{\min,2}^{(d)},$$

so the prolongation strategy yields an exact eigenvector for problem size d if $h = 0$.

One may ask whether the quality of the approximation obtained via the prolongation might be improved further if one constructs $\mathbf{M}$ based on exact eigenvectors $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$, where d_1 and d_2 are getting still larger than considered so far. Concentrating on the case $d_2 = d_1 + 2$ and even d_1, the construction of $\mathbf{N}$ and $\mathbf{M}$ might be carried out analogously to (5.5) via

$$\mathbf{N} := \operatorname{mat}_{2^{(d_1+4)/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1+2)}) \left(\operatorname{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\right)^{-1}, \quad \mathbf{M} := \mathbf{I}_{2^{d_1/2}} \otimes \mathbf{N}. \tag{5.17}$$

In fact, $\operatorname{mat}_{2^{d/2}\times 2^{d/2}}(\mathbf{v}_{\min}^{(d)})$ quickly gets ill-conditioned for growing d, see Table 5.1, noticing values in the order of or larger than the reciprocal of the machine precision have to be regarded as approximative. Moreover, the value of the *ε-rank*

$$\operatorname{rank}\left(\operatorname{mat}_{2^{d/2}\times 2^{d/2}}(\mathbf{v}_{\min}^{(d)})\,;\,\varepsilon\right) := \max\left\{1 \le j \le 2^{d/2} \colon \sigma_j > \varepsilon\right\},$$

where $\{\sigma_j\}_{j=1}^{2^{d/2}}$ are the singular values of $\operatorname{mat}_{2^{d/2}\times 2^{d/2}}(\mathbf{v}_{\min}^{(d)})$, is for several of the considered d and ε only a portion of $2^{d/2}$, so we encounter some form of low-rank property of $\mathbf{v}_{\min}^{(d)}$.

In view of the existence of very small singular values, we propose to replace the inverse of the matricization of $\mathbf{v}_{\min}^{(d_1)}$ in (5.17) by a *truncated pseudoinverse*. To be precise, let $\mathbf{U}\mathbf{\Sigma}\mathbf{W}^\top = \operatorname{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})$ be an SVD and let $j_\varepsilon := \operatorname{rank}\left(\operatorname{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\,;\,\varepsilon\right)$ for a given $\varepsilon \ge 0$. Then the truncated pseudoinverse with respect to a tolerance ε is given by

$$\operatorname{pinv}\left(\operatorname{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\,;\,\varepsilon\right) := \mathbf{W}_{:,1:j_\varepsilon} \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_{j_\varepsilon}^{-1} \end{pmatrix} (\mathbf{U}_{:,1:j_\varepsilon})^\top. \tag{5.18}$$

It is $\operatorname{pinv}(\mathbf{A}\,;\,0) = \mathbf{A}^+$ with the usual pseudoinverse $\mathbf{A}^+$ of some $\mathbf{A} \in \mathbb{R}^{m\times n}$, see [GVL13, Sect. 5.5.2]. So, if we test what effect it has on the approximation property of the prolongated vectors when the prolongation is based on (d_1, d_2), $d_2 = d_1 + 2$, $d_1 \in \{2, 4, 6, 8, 10, 12, 14\}$, we set

$$\mathbf{N} := \operatorname{mat}_{2^{(d_1+4)/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1+2)}) \left(\operatorname{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\right)^{-1} \quad \text{for } d_1 \in \{2, 4, 6, 8\}$$

and

$$\mathbf{N} := \mathrm{mat}_{2^{(d_1+4)/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1+2)}) \cdot \mathrm{pinv}\left(\mathrm{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\,;\, 10^{-6}\right)$$

for $d_1 \in \{6,8,10,12,14\}$. In the cases $d_1 \in \{6,8\}$, where the matricized eigenvector is actually invertible, both scenarios using inverse and truncated pseudoinverse are tested in order to check whether rank $\left(\mathrm{mat}_{2^{d_1/2}\times 2^{d_1/2}}(\mathbf{v}_{\min}^{(d_1)})\,;\, 10^{-6}\right) < 2^{d_1/2}$ has some influence. Figure 5.2 depicts the values of

$$\frac{(\widetilde{\mathbf{v}}^{(d)})^\top \mathbf{H}_d \widetilde{\mathbf{v}}^{(d)}}{(\widetilde{\mathbf{v}}^{(d)})^\top \widetilde{\mathbf{v}}^{(d)}} - \lambda_{\min}^{(d)}.$$

We observe that in each of the two different situations distinguished by the construction of $\mathbf{N}$ via classical or pseudoinverse, the larger the "starting point" of the prolongation gets, the more exact the approximation property measured in the Rayleigh quotient becomes. For $d_1 = 6$, the solid and the dashed yellow line are identical. However, when the pseudoinverse is employed for $d_1 \in \{8,10,12,14\}$, the resulting absolute error in the Rayleigh quotient is larger than for $d_1 = 8$ with classical inverse, but still in a satisfactory range.

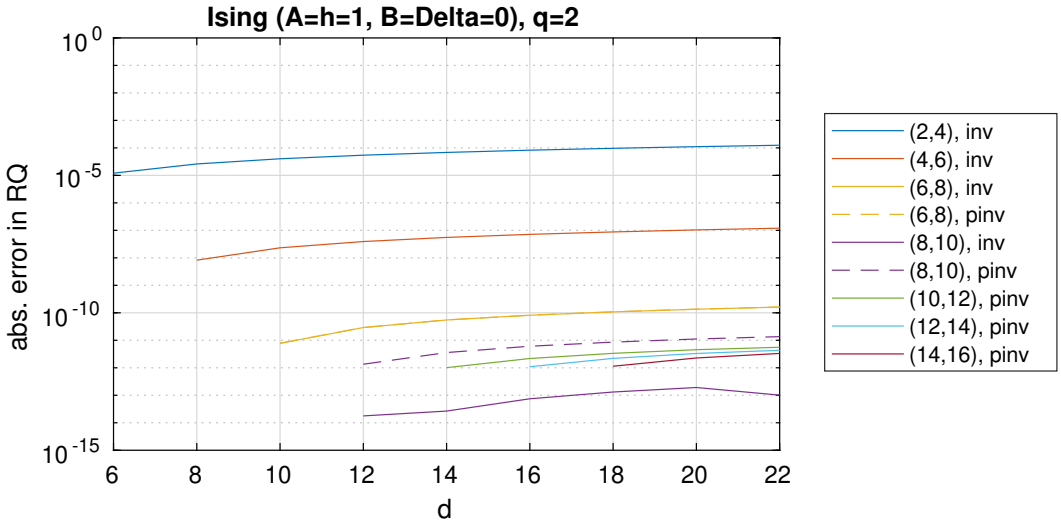


Figure 5.2.: Approximation quality of prolongated vector for 2-Ising model with varying (d_1, d_2) and classical inverse or truncated pseudoinverse.

5.1.2. 3-XYZ model, $A \neq B$

Next we consider the q-XYZ model (2.1) for $q = 3$ and write in this subsection $\mathbf{H}_d := \mathbf{H}_{d,3}^{\mathrm{XYZ}}$. Due to Corollary 2.10 and assuming like in the previous subsection $A \neq B$, the eigenvector $\mathbf{v}_{\min}^{(2)}$ has a maximal nonzero pattern

$$\text{either} \quad \begin{pmatrix} * & 0 & * & 0 & * & 0 & * & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,3}^{\mathrm{even}} \tag{5.19a}$$

$$\text{or} \quad \begin{pmatrix} 0 & * & 0 & * & 0 & * & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,3}^{\mathrm{odd}}. \tag{5.19b}$$

For $\mathrm{mat}_{3\times3}(\mathbf{v}_{\min}^{(2)})$ to be regular it is necessary that $\mathbf{v}_{\min}^{(2)}$ belongs to the first case since

$$\mathrm{mat}_{3\times3}\left(\begin{pmatrix}0 & * & 0 & * & 0 & * & 0 & * & 0\end{pmatrix}^\top\right) = \begin{pmatrix}0 & * & 0\\ * & 0 & *\\ 0 & * & 0\end{pmatrix}$$

for sure is singular.

For this condition to be sufficient we need further information about the different entries of $\mathbf{v}_{\min}^{(2)}$. Define $\mathbf{P}_{(1,3,5,7,9)} \in \mathbb{R}^{9\times5}$ as the matrix having a one in the entries $(1,1)$, $(3,2)$, $(5,3)$, $(7,4)$, $(9,5)$ and zeros elsewhere. Let further

$$\mathbf{H}_2^{(1,3,5,7,9)} := \mathbf{P}_{(1,3,5,7,9)}^\top \mathbf{H}_2 \mathbf{P}_{(1,3,5,7,9)} \in \mathbb{R}^{5\times5} \tag{5.20}$$

be the projection of $\mathbf{H}_2$ onto its rows and columns $1,3,5,7,9$. Let analogously $\mathbf{H}_2^{(2,4,6,8)} \in \mathbb{R}^{4\times4}$ be the projection of $\mathbf{H}_2$ onto its rows and columns $2,4,6,8$. We have to show that

$$\lambda_{\min}(\mathbf{H}_2^{(1,3,5,7,9)}) < \lambda_{\min}(\mathbf{H}_2^{(2,4,6,8)})$$

in order to conclude that $\mathbf{v}_{\min}^{(2)}$ has the sparsity (5.19a). At least in the case $A \neq 0 \neq h$, $B = \Delta = 0$, we are able to do that. By Lemma C.2 it is

$$\lambda_1(\mathbf{H}_2^{(1,3,5,7,9)}) = -\sqrt{\frac{A^2}{2} + 2h^2 + \sqrt{\frac{A^4}{4} + 4h^4}}, \tag{5.21a}$$

$$\lambda_2(\mathbf{H}_2^{(1,3,5,7,9)}) = -\sqrt{\frac{A^2}{2} + 2h^2 - \sqrt{\frac{A^4}{4} + 4h^4}}, \tag{5.21b}$$

$$\lambda_3(\mathbf{H}_2^{(1,3,5,7,9)}) = 0, \tag{5.21c}$$

$$\lambda_4(\mathbf{H}_2^{(1,3,5,7,9)}) = +\sqrt{\frac{A^2}{2} + 2h^2 - \sqrt{\frac{A^4}{4} + 4h^4}}, \tag{5.21d}$$

$$\lambda_5(\mathbf{H}_2^{(1,3,5,7,9)}) = +\sqrt{\frac{A^2}{2} + 2h^2 + \sqrt{\frac{A^4}{4} + 4h^4}}, \tag{5.21e}$$

and

$$\lambda_1(\mathbf{H}_2^{(2,4,6,8)}) = -\sqrt{\frac{A^2}{2} + h^2 + \sqrt{\frac{A^4}{4} + A^2h^2}}, \tag{5.22a}$$

$$\lambda_2(\mathbf{H}_2^{(2,4,6,8)}) = -\sqrt{\frac{A^2}{2} + h^2 - \sqrt{\frac{A^4}{4} + A^2h^2}}, \tag{5.22b}$$

$$\lambda_3(\mathbf{H}_2^{(2,4,6,8)}) = +\sqrt{\frac{A^2}{2} + h^2 - \sqrt{\frac{A^4}{4} + A^2h^2}}, \tag{5.22c}$$

$$\lambda_4(\mathbf{H}_2^{(2,4,6,8)}) = +\sqrt{\frac{A^2}{2} + h^2 + \sqrt{\frac{A^4}{4} + A^2h^2}}. \tag{5.22d}$$

We have $\frac{A^2}{2} + h^2 - \sqrt{\frac{A^4}{4} + A^2h^2} > 0$ since $\lambda_2(\mathbf{H}_2^{(2,4,6,8)}) \in \mathbb{R}$, as $\mathbf{H}_2^{(2,4,6,8)}$ is symmetric. This yields

$$\begin{aligned}\left(\lambda_1(\mathbf{H}_2^{(1,3,5,7,9)})\right)^2 - \left(\lambda_1(\mathbf{H}_2^{(2,4,6,8)})\right)^2 &= \frac{A^2}{2} + 2h^2 + \sqrt{\frac{A^4}{4} + 4h^4} - \frac{A^2}{2} - h^2 - \sqrt{\frac{A^4}{4} + A^2h^2} \\ &= -\frac{A^2}{2} + \sqrt{\frac{A^4}{4} + 4h^4} + \frac{A^2}{2} + h^2 - \sqrt{\frac{A^4}{4} + A^2h^2} \\ &> -\frac{A^2}{2} + \sqrt{\frac{A^4}{4} + 4h^4} > -\frac{A^2}{2} + \sqrt{\frac{A^4}{4}} = 0.\end{aligned}$$

Since $\lambda_1(\mathbf{H}_2^{(1,3,5,7,9)}) < 0$ and $\lambda_1(\mathbf{H}_2^{(2,4,6,8)}) < 0$, we obtain $\lambda_1(\mathbf{H}_2^{(1,3,5,7,9)}) < \lambda_1(\mathbf{H}_2^{(2,4,6,8)})$. So

$$\mathbf{v}_{\min}^{(2)} = \begin{pmatrix} * & 0 & * & 0 & * & 0 & * & 0 & * \end{pmatrix}^\top.$$

As stated in Lemma C.3,

$$\mathbf{v} = \begin{pmatrix} w_1 & 0 & w_2 & 0 & w_3 & 0 & w_4 & 0 & w_5 \end{pmatrix}^\top = \begin{pmatrix} \frac{-A}{4h-2\lambda} & 0 & \frac{A}{2\lambda} & 0 & 1 & 0 & \frac{A}{2\lambda} & 0 & \frac{A}{4h+2\lambda} \end{pmatrix}^\top$$

is an eigenvector especially associated with $\lambda = \lambda_{\min}(\mathbf{H}_2) = \lambda_1(\mathbf{H}_2^{(1,3,5,7,9)}) \neq 0$. This implies

$$\begin{aligned}\det\left(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right) &= \det\begin{pmatrix} w_1 & 0 & w_2 \\ 0 & w_3 & 0 \\ w_2 & 0 & w_5 \end{pmatrix} = w_3(w_1w_5 - w_2^2) \\ &= \frac{-A^2}{16h^2 - 4\lambda^2} - \frac{A^2}{4\lambda^2} = \frac{-A^2\lambda^2 - A^2(4h^2 - \lambda^2)}{4(4h^2 - \lambda^2)\lambda^2} = \frac{-A^2h^2}{(4h^2 - \lambda^2)\lambda^2} \neq 0,\end{aligned}$$

and hence the invertibility of $\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})$.

To investigate the sparsity pattern of $\mathbf{v}_{\min}^{(2)}$ for arbitrary $A \neq B$, Δ, h, we choose the four parameters randomly and check whether $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\mathrm{even}}$ or $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\mathrm{odd}}$. We repeat this procedure 10,000 times, and 1049 times thereof we obtain $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\mathrm{odd}}$, noticing that this value might differ slightly when the test is carried out once again. Each time if $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\mathrm{odd}}$, there is a marker in the 100×100 array traversed row-wise on the left of Figure 5.3. The right plot depicts $\mathrm{cond}\left(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)$ for the 8951 fold occurence of $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\mathrm{even}}$, where the single condition numbers are sorted for better readability. It is

$$\begin{cases} 1981 \\ 192 \\ 16 \\ 1 \end{cases} \quad \text{times that} \quad \mathrm{cond}\left(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right) > \begin{cases} 10^2 \\ 10^4 \\ 10^6 \\ 10^8 \end{cases}.$$

Also in the case $q = 3$ it is possible to construct an approximation $\widetilde{\mathbf{v}}^{(d)}$ of the eigenvector $\mathbf{v}_{\min}^{(d)}$ associated with the minimal eigenvalue $\lambda_{\min}^{(d)}$ based on information about the relation between $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ for small $d_2 > d_1$. The strategy is analogous to the situation $q = 2$ but one has to take into account that now all the sizes of the involved objects are related to

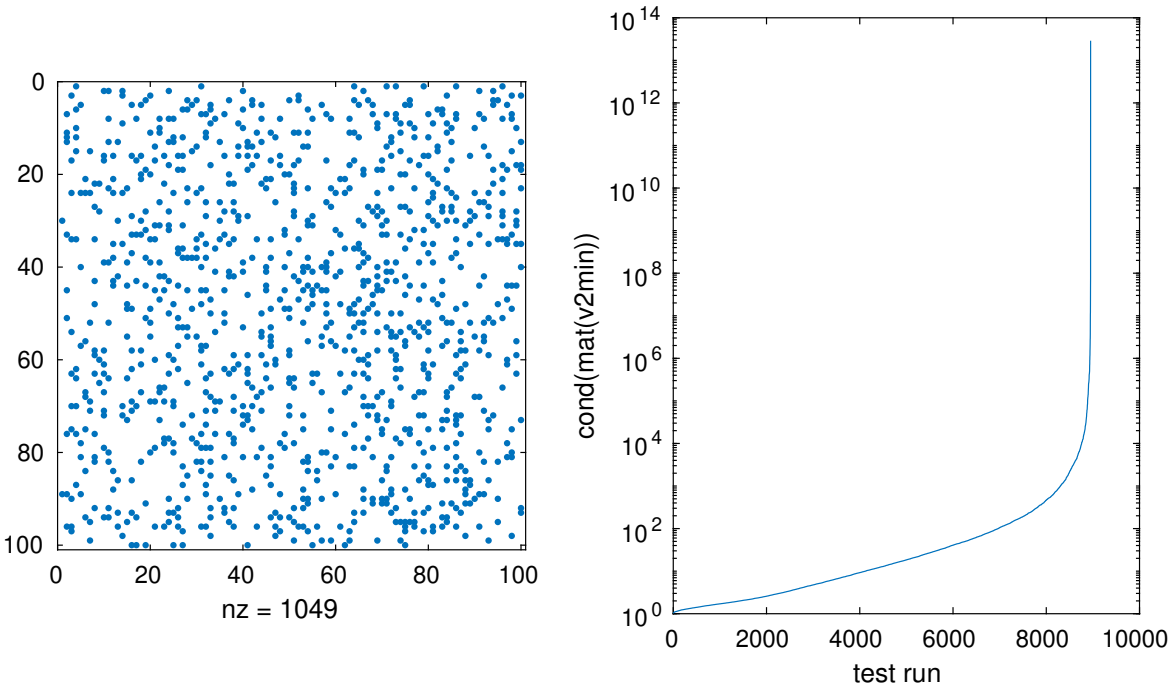


Figure 5.3.: Sparsity pattern of $\mathbf{v}_{\min}^{(2)}$ for 3-XYZ model with 10,000 times randomly chosen $A \neq B$, Δ, h.
Left: blue marker if $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$, no marker if $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$.
Right: Condition number of $\text{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})$ in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$ (sorted in ascending order).

a power of 3 instead of 2. Concentrating on the case $d_1 = 2$ and assuming that $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$ is such that $\text{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})$ is invertible, we set

$$\mathbf{N} := \text{mat}_{3^{d_2-1}\times 3}(\mathbf{v}_{\min}^{(d_2)}) \left(\text{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^{-1} \tag{5.23}$$

as well as

$$\mathbf{M} := \mathbf{I}_3 \otimes \mathbf{N} \tag{5.24}$$

in order that

$$\mathbf{M}\mathbf{v}_{\min}^{(2)} = \mathbf{v}_{\min}^{(d_2)}.$$

The *prolongation* to d, where $d - d_2$ is divisible by $d_2 - 2$, may be performed like in (5.8) via

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{(d-d_2)/(d_2-2)} (\mathbf{I}_{3^{(d_2-2)i}} \otimes \mathbf{M})\right) \mathbf{v}_{\min}^{(d_2)}. \tag{5.25}$$

In Figure 5.4 the approximation quality of $\widetilde{\mathbf{v}}^{(d)}$ is visualized with $d_2 = 4$ in case $A = 1$, $B = \Delta = 0$, the analog to the Ising model. For $h \in [-5, -0.01] \cup [0.01, 5]$, the value of $-\lambda_{\min}^{(d)}$ as

well as the absolute error concerning the Rayleigh quotient

$$\frac{(\widetilde{\mathbf{v}}^{(d)})^\top \mathbf{H}_d \widetilde{\mathbf{v}}^{(d)}}{(\widetilde{\mathbf{v}}^{(d)})^\top \widetilde{\mathbf{v}}^{(d)}} - \lambda_{\min}^{(d)} \tag{5.26}$$

is depicted for $d = 10$ and $d = 14$ which is actually the largest d tractable exactly (compare $3^{14} = 4{,}782{,}969$ with $2^{22} = 4{,}194{,}304$). This error gets small when h is close to 0, an observation we already made for the Ising model with $q = 2$, cf. Figure 5.1.

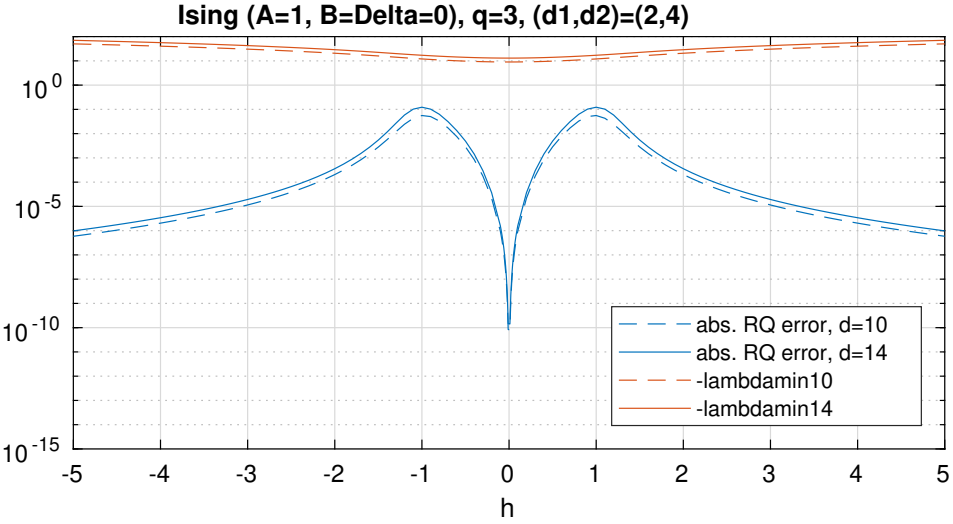


Figure 5.4.: Approximation quality of prolongated vector for 3-Ising model with varying $h \in [-5, -0.01] \cup [0.01, 5]$, value of $-\lambda_{\min}^{(d)}$.

If $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$, the construction (5.23) of $\mathbf{N}$ is not directly possible since $\text{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})$ is not invertible. We might employ instead the pseudoinverse $\cdot^+$, hence

$$\mathbf{N} := \text{mat}_{3^{d_2-1}\times 3}(\mathbf{v}_{\min}^{(d_2)}) \left(\text{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^+. \tag{5.27}$$

As an alternative, we observe that each time $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$, then $\mathbf{v}_2^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$, where $\mathbf{v}_2^{(2)}$ denotes an eigenvector associated with the second smallest eigenvalue, so in this situation we might utilize $\mathbf{v}_2^{(2)}$ to set up $\mathbf{N}$. We have to decide whether $\mathbf{N}$ should be constructed in order that $\mathbf{M} = \mathbf{I}_3 \otimes \mathbf{N}$ maps $\mathbf{v}_2^{(2)}$ to $\mathbf{v}_{\min}^{(d_2)}$ or to $\mathbf{v}_2^{(d_2)}$, which means

$$\mathbf{N} := \text{mat}_{3^{d_2-1}\times 3}(\mathbf{v}_{\min}^{(d_2)}) \left(\text{mat}_{3\times 3}(\mathbf{v}_2^{(2)})\right)^{-1} \quad \text{or} \quad \mathbf{N} := \text{mat}_{3^{d_2-1}\times 3}(\mathbf{v}_2^{(d_2)}) \left(\text{mat}_{3\times 3}(\mathbf{v}_2^{(2)})\right)^{-1}. \tag{5.28}$$

Furthermore, both choices of $\mathbf{N}$ may be employed to prolongate either $\mathbf{v}_{\min}^{(d_2)}$ or $\mathbf{v}_2^{(d_2)}$ via (5.25) to some $\widetilde{\mathbf{v}}^{(d)}$. In the left part of Figure 5.5 we compare the resulting absolute errors in the Rayleigh quotient (5.26) with $d_2 = 3$ and $d = 12$, where $\mathbf{N}$ is set up via the pseudoinverse (5.27) or via $\mathbf{v}_2^{(2)}$ (5.28). In the latter case, the legend entry (a, b) for $a, b \in \{\min, 2\}$ has to be understood as

$$\mathbf{N} := \text{mat}_{3^2\times 3}(\mathbf{v}_a^{(3)}) \left(\text{mat}_{3\times 3}(\mathbf{v}_2^{(2)})\right)^{-1}, \quad \widetilde{\mathbf{v}}^{(12)} := \left(\prod_{i=1}^{(12-3)/(3-2)} (\mathbf{I}_{3^{(3-2)i}} \otimes (\mathbf{I}_3 \otimes \mathbf{N}))\right) \mathbf{v}_b^{(3)}.$$

We sort the results due to their magnitude in the pseudoinverse case. We randomly choose 100 times parameters A, B, Δ, h that yield $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$ and keep these parameters fixed for the five scenarios in one repetition. For the right part of Figure 5.5 we take 100 randomly chosen parameter sets that yield $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$, set

$$\mathbf{N} := \operatorname{mat}_{3^2\times 3}(\mathbf{v}_{\min}^{(3)}) \left(\operatorname{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^{-1}, \quad \widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{(d-3)/(3-2)} \left(\mathbf{I}_{3^{(3-2)i}} \otimes (\mathbf{I}_3 \otimes \mathbf{N})\right) \right) \mathbf{v}_{\min}^{(3)},$$

evaluate (5.26) for $d \in \{8, 12\}$, and plot these absolute errors, sorted due to their magnitude in case $d = 8$. Additionally we plot in both pictures the respective current value of $-\lambda_{\min}^{(d)}$. We observe that in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$, the definition of $\mathbf{N}$ via the pseudoinverse by (5.27) yields in average a slightly better approximation. Also the size of the error is comparable with that in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$.

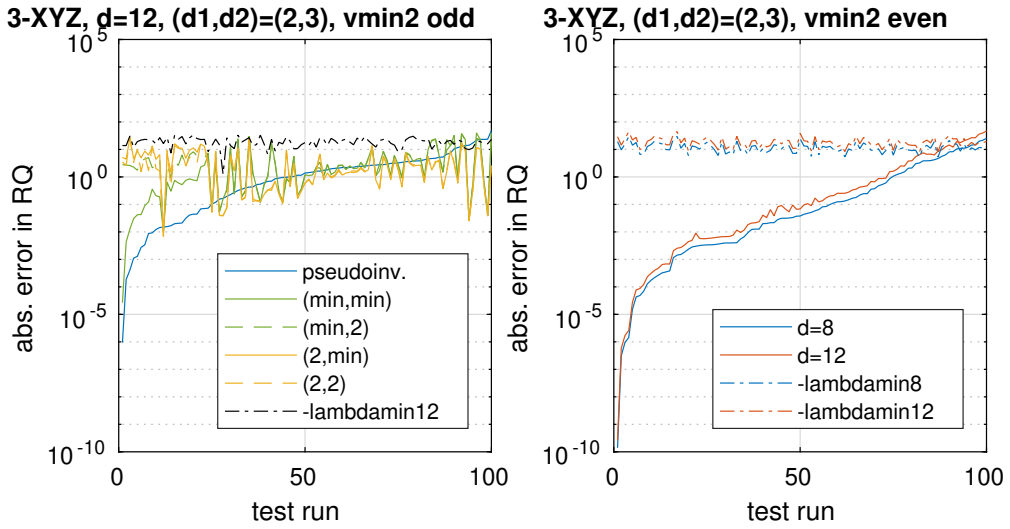


Figure 5.5.: Left: Approximation quality of prolongated vector constructed via $\left(\operatorname{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^{+}$ or $\left(\operatorname{mat}_{3\times 3}(\mathbf{v}_{2}^{(2)})\right)^{-1}$ in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$.
Right: For reference, approximation quality in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{even}}$.

5.1.3. 3-Potts model

Now we turn towards the 3-Potts model. As it is relevant for the construction of the prolongation operator, we discuss again the invertibility of the matricization of $\mathbf{v}_{\min}^{(2)}$. Following example B.2, the possible shape of $\mathbf{v}_{\min}^{(2)}$ is

$$\begin{aligned} \text{either} \quad & \begin{pmatrix} * & 0 & 0 & 0 & 0 & * & 0 & * & 0 \end{pmatrix}^{\top} \in \mathcal{E}_{2,3}^{[0]} \\ \text{or} \quad & \begin{pmatrix} 0 & * & 0 & * & 0 & 0 & 0 & 0 & * \end{pmatrix}^{\top} \in \mathcal{E}_{2,3}^{[1]} \\ \text{or} \quad & \begin{pmatrix} 0 & 0 & * & 0 & * & 0 & * & 0 & 0 \end{pmatrix}^{\top} \in \mathcal{E}_{2,3}^{[2]}. \end{aligned}$$

We justify that $\mathbf{v}_{\min}^{(2)}$ has the first shape and indeed each $*$ is nonzero. This implies the invertibility of

$$\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)}) = \begin{pmatrix} * & 0 & 0 \\ 0 & 0 & * \\ 0 & * & 0 \end{pmatrix}.$$

For this purpose, it is sufficient to consider a via $A = 1$ "normalized" Hamilton operator $\mathbf{H}_{d,3}^{\mathrm{Potts}}$ of the Potts model from (2.6) which we denote in this subsection by

$$\mathbf{H}_d := -\sum_{i=1}^{d-1} \left((\mathbf{\Gamma}_3)^{(i)} (\mathbf{\Gamma}_3^2)^{(i+1)} + (\mathbf{\Gamma}_3^2)^{(i)} (\mathbf{\Gamma}_3)^{(i+1)} \right) - h \sum_{i=1}^{d} \left((\mathbf{\Omega}_3)^{(i)} + (\mathbf{\Omega}_3^2)^{(i)} \right).$$

We focus on the case $h \neq 0$ since otherwise the minimal eigenvalue of $\mathbf{H}_d$ and an associated eigenvector would already be known by Proposition 2.19(ii). Defining $\mathbf{P}_{(k_1,k_2,k_3)} \in \mathbb{R}^{9\times 3}$ as the matrix having a one in the entries $(k_1,1),(k_2,2),(k_3,3)$ and zeros elsewhere, we set

$$\begin{aligned}
\mathbf{H}_2^{(1,6,8)} &:= \mathbf{P}_{(1,6,8)}^\top \mathbf{H}_2 \mathbf{P}_{(1,6,8)} = -\begin{pmatrix} 4h & 1 & 1 \\ 1 & -2h & 1 \\ 1 & 1 & -2h \end{pmatrix}, \\
\mathbf{H}_2^{(2,4,9)} &:= \mathbf{P}_{(2,4,9)}^\top \mathbf{H}_2 \mathbf{P}_{(2,4,9)} = -\begin{pmatrix} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & -2h \end{pmatrix}, \\
\mathbf{H}_2^{(3,5,7)} &:= \mathbf{P}_{(3,5,7)}^\top \mathbf{H}_2 \mathbf{P}_{(3,5,7)} = -\begin{pmatrix} h & 1 & 1 \\ 1 & -2h & 1 \\ 1 & 1 & h \end{pmatrix}.
\end{aligned}$$

Since

$$\mathbf{H}_2^{(3,5,7)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \mathbf{H}_2^{(2,4,9)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{5.29}$$

the eigenvalues of $\mathbf{H}_2^{(2,4,9)}$ and $\mathbf{H}_2^{(3,5,7)}$ are identical due to [HJ13, Cor. 1.3.4]. For one component of an eigenvector associated with, say, $\lambda_k(\mathbf{H}_2^{(1,6,8)})$, $k \in \{1,2,3\}$, to be zero, it is necessary that there exists a principal 2×2 submatrix $\mathbf{H}_2^{(1,6)}$, $\mathbf{H}_2^{(1,8)}$, or $\mathbf{H}_2^{(6,8)}$, respectively, of $\mathbf{H}_2^{(1,6,8)}$ (formed by deleting the respectively third, second, or first row and column, cf. [HJ13, Sect. 0.7.1]) which has an eigenvalue equal to $\lambda_k(\mathbf{H}_2^{(1,6,8)})$. This can be inferred from Proposition C.4, see also Corollary C.5.

The characteristic polynomial of $\mathbf{H}_2^{(1,6,8)}$ is

$$\begin{aligned}
\chi^{(1,6,8)}(\lambda) := \det\left(\lambda \mathbf{I}_3 - \mathbf{H}_2^{(1,6,8)}\right) &= \det \begin{pmatrix} \lambda + 4h & 1 & 1 \\ 1 & \lambda - 2h & 1 \\ 1 & 1 & \lambda - 2h \end{pmatrix} \\
&= \lambda^3 - 3(4h^2+1)\lambda + 2(8h^3+1).
\end{aligned}$$

For the 2×2 submatrices

$$\mathbf{H}_2^{(1,6)} := \mathbf{H}_2^{(1,8)} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}^\top \mathbf{H}_2^{(1,6,8)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}^\top \mathbf{H}_2^{(1,6,8)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = -\begin{pmatrix} 4h & 1 \\ 1 & -2h \end{pmatrix}$$

and

$$\mathbf{H}_2^{(6,8)} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}^\top \mathbf{H}_2^{(1,6,8)} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = - \begin{pmatrix} -2h & 1 \\ 1 & -2h \end{pmatrix}$$

of $\mathbf{H}_2^{(1,6,8)}$ it holds

$$\begin{aligned} \lambda_1(\mathbf{H}_2^{(1,6)}) &= -h - \sqrt{9h^2+1} = \lambda_1(\mathbf{H}_2^{(1,8)}), \\ \lambda_2(\mathbf{H}_2^{(1,6)}) &= -h + \sqrt{9h^2+1} = \lambda_2(\mathbf{H}_2^{(1,8)}), \\ \lambda_1(\mathbf{H}_2^{(6,8)}) &= 2h-1, \\ \lambda_2(\mathbf{H}_2^{(6,8)}) &= 2h+1. \end{aligned}$$

It is

$$\chi^{(1,6,8)}\left(\lambda_1(\mathbf{H}_2^{(1,6)})\right) = 2 + 2\sqrt{9h^2+1} \geq 4, \tag{5.30a}$$

$$\chi^{(1,6,8)}\left(\lambda_2(\mathbf{H}_2^{(1,6)})\right) = 2 - 2\sqrt{9h^2+1} \begin{cases} < 0 & , \ h \neq 0 \\ = 0 & , \ h = 0 \end{cases}, \tag{5.30b}$$

$$\chi^{(1,6,8)}\left(\lambda_1(\mathbf{H}_2^{(6,8)})\right) = 4, \tag{5.30c}$$

$$\chi^{(1,6,8)}\left(\lambda_2(\mathbf{H}_2^{(6,8)})\right) = 0. \tag{5.30d}$$

Motivated by (5.30d) we calculate

$$\begin{aligned} \chi^{(1,6,8)}(\lambda) &= (\lambda - (2h+1))\left(\lambda^2 + (2h+1)\lambda - 2(4h^2 - 2h + 1)\right) \\ &= (\lambda - (2h+1))\left(\lambda - \left(-\left(h+\frac{1}{2}\right) - 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right) \\ &\quad \cdot \left(\lambda - \left(-\left(h+\frac{1}{2}\right) + 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right). \end{aligned}$$

Due to

$$2h-1 < 2h+1, \quad -\left(h+\frac{1}{2}\right) - 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}} < -\left(h+\frac{1}{2}\right) + 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}},$$

and the eigenvalue interlacing property

$$\lambda_i(\mathbf{A}) \leq \lambda_i(\mathbf{B}) \leq \lambda_{i+1}(\mathbf{A}),$$

where $\mathbf{B} \in \mathbb{C}^{n-1 \times n-1}$ is a principal submatrix of a Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, cf. [HJ13, Thm. 4.3.28 & Cor. 1.3.4], we obtain

$$\lambda_{\min}(\mathbf{H}_2^{(1,6,8)}) = -\left(h+\frac{1}{2}\right) - 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}.$$

Hence we conclude from (5.30) that no eigenvalue of any principal 2×2 submatrix of $\mathbf{H}_2^{(1,6,8)}$ equals $\lambda_{\min}(\mathbf{H}_2^{(1,6,8)})$.

An analogous argumentation may be carried out for $\mathbf{H}_2^{(2,4,9)}$, and thanks to (5.29), the results also apply to $\mathbf{H}_2^{(3,5,7)}$. It is

$$\begin{aligned}\chi^{(2,4,9)}(\lambda) := \det\left(\lambda\mathbf{I}_3 - \mathbf{H}_2^{(2,4,9)}\right) &= \det\begin{pmatrix}\lambda+h & 1 & 1\\ 1 & \lambda+h & 1\\ 1 & 1 & \lambda-2h\end{pmatrix}\\ &= \lambda^3 - 3(h^2+1)\lambda + 2(-h^3+1)\end{aligned}$$

and for the 2×2 submatrices $\mathbf{H}_2^{(2,4)}$, $\mathbf{H}_2^{(2,9)}$, $\mathbf{H}_2^{(4,9)}$ of $\mathbf{H}_2^{(2,4,9)}$ resp. $\mathbf{H}_2^{(3,5)}$, $\mathbf{H}_2^{(3,7)}$, $\mathbf{H}_2^{(5,7)}$ of $\mathbf{H}_2^{(3,5,7)}$ it holds

$$\begin{aligned}\lambda_1(\mathbf{H}_2^{(2,4)}) &= -h-1 = \lambda_1(\mathbf{H}_2^{(3,7)}),\\ \lambda_2(\mathbf{H}_2^{(2,4)}) &= -h+1 = \lambda_2(\mathbf{H}_2^{(3,7)}),\\ \lambda_1(\mathbf{H}_2^{(2,9)}) &= \frac{1}{2}\left(h - \sqrt{9h^2+4}\right) = \lambda_1(\mathbf{H}_2^{(4,9)}) = \lambda_1(\mathbf{H}_2^{(3,5)}) = \lambda_1(\mathbf{H}_2^{(5,7)}),\\ \lambda_2(\mathbf{H}_2^{(2,9)}) &= \frac{1}{2}\left(h + \sqrt{9h^2+4}\right) = \lambda_2(\mathbf{H}_2^{(4,9)}) = \lambda_2(\mathbf{H}_2^{(3,5)}) = \lambda_2(\mathbf{H}_2^{(5,7)}).\end{aligned}$$

Now

$$\chi^{(2,4,9)}\left(\lambda_1(\mathbf{H}_2^{(2,4)})\right) = 4, \tag{5.31a}$$

$$\chi^{(2,4,9)}\left(\lambda_2(\mathbf{H}_2^{(2,4)})\right) = 0, \tag{5.31b}$$

$$\chi^{(2,4,9)}\left(\lambda_1(\mathbf{H}_2^{(6,8)})\right) = 2 + \sqrt{9h^2+4} \geq 4, \tag{5.31c}$$

$$\chi^{(2,4,9)}\left(\lambda_2(\mathbf{H}_2^{(6,8)})\right) = 2 - \sqrt{9h^2+4}\begin{cases}<0 & ,\ h\neq 0\\ =0 & ,\ h=0\end{cases}, \tag{5.31d}$$

and therefore

$$\begin{aligned}\chi^{(2,4,9)}(\lambda) &= (\lambda-(-h+1))\left(\lambda^2 + (-h+1)\lambda - 2(h^2+h+1)\right)\\ &= (\lambda-(-h+1))\left(\lambda - \frac{1}{2}\left(-(-h+1) - 3\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}\right)\right)\\ &\quad\cdot\left(\lambda - \frac{1}{2}\left(-(-h+1) + 3\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}\right)\right).\end{aligned}$$

So we obtain

$$\lambda_{\min}(\mathbf{H}_2^{(2,4,9)}) = \frac{1}{2}\left(-(-h+1) - 3\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}\right),$$

and we conclude from (5.31) that no eigenvalue of any principal 2×2 submatrix of $\mathbf{H}_2^{(2,4,9)}$ equals $\lambda_{\min}(\mathbf{H}_2^{(2,4,9)})$. Finally we set

$$\begin{aligned}\widehat{\chi}^{(1,6,8)}(\lambda) &:= \frac{\chi^{(1,6,8)}(\lambda)}{\lambda-(2h+1)} = \lambda^2 + (2h+1)\lambda - 2(4h^2-2h+1)\\ &= \left(\lambda - \left(-\left(h+\frac{1}{2}\right) - 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right)\left(\lambda - \left(-\left(h+\frac{1}{2}\right) + 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right)\end{aligned}$$

and Lemma C.6 verifies

$$\widehat{\chi}^{(1,6,8)}\left(\lambda_{\min}(\mathbf{H}_2^{(2,4,9)})\right)\begin{cases}<0 & , \ h\neq 0\\ =0 & , \ h=0\end{cases}. \tag{5.32}$$

Since $\widehat{\chi}^{(1,6,8)}$ is a polynomial function of degree two with positive leading coefficient and the smaller one of its two zeros equals $\lambda_{\min}(\mathbf{H}_2^{(1,6,8)})$, (5.32) implies

$$\lambda_{\min}(\mathbf{H}_2^{(1,6,8)}) < \lambda_{\min}(\mathbf{H}_2^{(2,4,9)}) \quad \text{if } h\neq 0.$$

To summarize,

$$\lambda_{\min}(\mathbf{H}_2) = \lambda_{\min}(\mathbf{H}_2^{(1,6,8)}), \quad \mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{[0]},$$

and due to the inequality of both eigenvalues of each principal 2×2 submatrix of $\mathbf{H}_2^{(1,6,8)}$ to $\lambda_{\min}(\mathbf{H}_2^{(1,6,8)})$, the three potential nonzero elements of $\mathbf{v}_{\min}^{(2)}$ are indeed all nonzero by Corollary C.5. This shows the invertibility of

$$\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)}) = \begin{pmatrix} * & 0 & 0\\ 0 & 0 & *\\ 0 & * & 0\end{pmatrix}.$$

The construction (5.23)-(5.25) of an initial guess for the 3-XYZ model may be performed in the same manner for the 3-Potts model. Figure 5.6 presents the absolute error concerning the Rayleigh quotient

$$\frac{(\widetilde{\mathbf{v}}^{(d)})^\top \mathbf{H}_d \widetilde{\mathbf{v}}^{(d)}}{(\widetilde{\mathbf{v}}^{(d)})^\top \widetilde{\mathbf{v}}^{(d)}} - \lambda_{\min}^{(d)}$$

in case of the 3-Potts model with $A=1$, $h\in[-5,-0.01]\cup[0.01,5]$ for $d\in\{10,14\}$ and $(d_1,d_2)=(2,4)$ and in addition the respective value of $-\lambda_{\min}^{(d)}$. As for the 3-XYZ model, cf. Figure 5.4, the error gets small when h is close to 0.

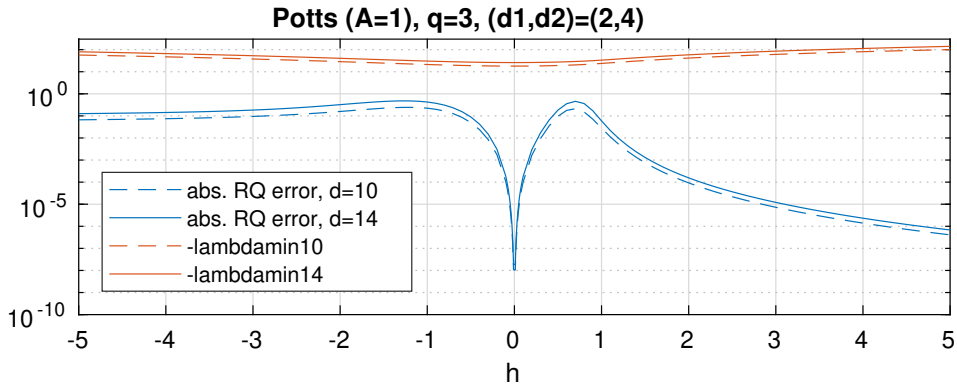


Figure 5.6.: Approximation quality of prolongated vector for 3-Potts model with varying $h\in[-5,-0.01]\cup[0.01,5]$, value of $-\lambda_{\min}^{(d)}$.

We conclude Section 5.1 with a statement that the described construction of an initial guess yields a vector that has a sparsity pattern suitable for eigenvectors of the Hamilton operators considered so far. We formulate this statement in the context of the spaces $\mathcal{E}_{d,q}^{(k)}$, $0\leq k\leq$

$(q-1)d$, but it is easily translated to $\mathcal{E}_{d,q}^{\text{even}}$ or $\mathcal{E}_{d,q}^{\text{odd}}$ respectively $\mathcal{E}_{d,q}^{[k]}$, $0 \leq k \leq q-1$. Therefore we employ for generality in the construction of the matrix $\mathbf{M}$, cf. (5.9) and (5.10), the pseudoinverse which coincides with the classical inverse in case of an invertible matrix. As a preparation, we show that the sparsity pattern of potential nonzero entries is identical for the quadratic matricization of an eigenvector and its pseudoinverse.

Lemma 5.2. *For $d \geq 2$ even, $q \geq 2$, and $0 \leq k \leq (q-1)d$, let $\mathbf{v}^{(d)} \in \mathcal{E}_{d,q}^{(k)}$ and let $\cdot^+$ denote the pseudoinverse. Then*

$$\operatorname{vec}\left(\left(\operatorname{mat}_{q^{d/2}\times q^{d/2}}(\mathbf{v}^{(d)})\right)^+\right) \in \mathcal{E}_{d,q}^{(k)}.$$

Proof. Since $\mathbf{v}^{(d)} \in \mathcal{E}_{d,q}^{(k)}$, it is

$$\mathbf{A} := \operatorname{mat}_{q^{d/2}\times q^{d/2}}(\mathbf{v}^{(d)}) = \sum_{\alpha_1=0}^{q-1} \cdots \sum_{\alpha_d=0}^{q-1} a_{\alpha_1,\ldots,\alpha_d} \left|\alpha_{d/2+1}\ \ldots\ \alpha_d\right\rangle \left\langle\alpha_1\ \ldots\ \alpha_{d/2}\right|,$$

where $a_{\alpha_1,\ldots,\alpha_d} \neq 0$ only if $\alpha_1 + \ldots + \alpha_d = k$. So, for a column $\mathbf{A}_{:,j}$ with $1 \leq j \leq q^{d/2}$ such that the sum of digits of the q-ary representation of $j-1$ equals $\kappa \in K := \{\max\{0, k-(q-1)d/2\}, \ldots, \min\{(q-1)d/2, k\}\}$, it is $\mathbf{A}_{:,j} \in \mathcal{E}_{d/2,q}^{(k-\kappa)}$. For an analogously characterized row $\mathbf{A}_{j,:}$ it is $(\mathbf{A}_{j,:})^\top \in \mathcal{E}_{d/2,q}^{(k-\kappa)}$, likewise. Hence the sparsity pattern of $\operatorname{mat}_{q^{d/2}\times q^{d/2}}(\mathbf{v}^{(d)})$ is symmetric with respect to the main diagonal. We notice that the set K contains all possible sums of $d/2$ numbers each between 0 and $q-1$ that are additionally a subset of d numbers each between 0 and $q-1$ whose sum equals k. Remembering $t(\cdot,\cdot,\cdot)$ from (2.3), let

$$\mathbf{P}^{(\kappa)} := \begin{pmatrix} \mathbf{e}_{i_1} & \cdots & \mathbf{e}_{i_{t(d/2,q,\kappa)}} \end{pmatrix} \in \mathbb{R}^{q^{d/2}\times t(d/2,q,\kappa)},$$

where $i_1 - 1, \ldots, i_{t(d/2,q,\kappa)} - 1$ are the $t(d/2,q,\kappa)$ numbers between 0 and $q^{d/2}-1$ which have a q-ary representation whose sum of digits is κ. We define

$$\mathbf{A}^{(\kappa)} := (\mathbf{P}^{(\kappa)})^\top \mathbf{A} \mathbf{P}^{(k-\kappa)} \in \mathbb{R}^{t(d/2,q,\kappa)\times t(d/2,q,k-\kappa)}.$$

For each $\kappa \in K$, there exists a *thin* SVD, cf. [GVL13, Sect. 2.4.3], $\mathbf{A}^{(\kappa)} = \mathbf{U}^{(\kappa)}\mathbf{\Sigma}^{(\kappa)}(\mathbf{W}^{(\kappa)})^\top$ with quadratic $\mathbf{\Sigma}^{(\kappa)} \in \mathbb{R}^{m(\kappa)\times m(\kappa)}$, $m(\kappa) := \min\{t(d/2,q,\kappa), t(d/2,q,k-\kappa)\}$, and $\mathbf{U}^{(\kappa)}$ and $\mathbf{W}^{(\kappa)}$ both have $m(\kappa)$ columns. Due to the sparsity pattern of $\mathbf{A}$, it is

$$\begin{aligned}
\mathbf{A} &= \sum_{\kappa\in K} \mathbf{P}^{(\kappa)}(\mathbf{P}^{(\kappa)})^\top \mathbf{A} \mathbf{P}^{(k-\kappa)}(\mathbf{P}^{(k-\kappa)})^\top \\
&= \sum_{\kappa\in K} \mathbf{P}^{(\kappa)}\mathbf{A}^{(\kappa)}(\mathbf{P}^{(k-\kappa)})^\top \\
&= \sum_{\kappa\in K} \mathbf{P}^{(\kappa)}\mathbf{U}^{(\kappa)}\mathbf{\Sigma}^{(\kappa)}(\mathbf{W}^{(\kappa)})^\top(\mathbf{P}^{(k-\kappa)})^\top \\
&= \sum_{\kappa\in K} \mathbf{P}^{(\kappa)}\mathbf{U}^{(\kappa)}\mathbf{\Sigma}^{(\kappa)}(\mathbf{P}^{(k-\kappa)}\mathbf{W}^{(\kappa)})^\top.
\end{aligned}$$

Since the spaces $\mathcal{E}^{(\kappa_1)}_{d/2,q}$ and $\mathcal{E}^{(\kappa_2)}_{d/2,q}$ are orthogonal for $\kappa_1 \neq \kappa_2$, with $M := \sum_{\kappa\in K} m(\kappa)$ and

$$\begin{aligned}
\mathbf{U} &:= \begin{pmatrix} \mathbf{P}^{(\min(K))}\mathbf{U}^{(\min(K))} & \dots & \mathbf{P}^{(\max(K))}\mathbf{U}^{(\max(K))} \end{pmatrix} \in \mathbb{R}^{q^{d/2}\times M},\\
\mathbf{\Sigma} &:= \begin{pmatrix} \mathbf{\Sigma}^{(\min(K))} & & \\ & \ddots & \\ & & \mathbf{\Sigma}^{(\max(K))} \end{pmatrix} \in \mathbb{R}^{M\times M},\\
\mathbf{W} &:= \begin{pmatrix} \mathbf{P}^{(k-\min(K))}\mathbf{W}^{(\min(K))} & \dots & \mathbf{P}^{(k-\max(K))}\mathbf{W}^{(\max(K))} \end{pmatrix} \in \mathbb{R}^{q^{d/2}\times M},
\end{aligned}$$

a thin SVD of $\mathbf{A}$ is given by $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^\top$, where we notice that the columns of $\mathbf{U}$ and $\mathbf{W}$ have to be rearranged so that the diagonal entries of $\mathbf{\Sigma}$ are sorted in decreasing order. Thus, for each $1 \le i \le M$, it is $\mathbf{u}_i := \mathbf{U}_{:,i} \in \mathcal{E}^{(\kappa(i))}_{d/2,q}$ for some $\kappa(i) \in K$ and additionally $\mathbf{w}_i := \mathbf{W}_{:,i} \in \mathcal{E}^{(k-\kappa(i))}_{d/2,q}$. As by [GVL13, Sect. 5.5.2]

$$\mathbf{A}^+ = \sum_{1\le i\le M\colon\, \sigma_i\neq 0} \frac{1}{\sigma_i}\mathbf{w}_i\mathbf{u}_i^\top,$$

the result follows. □

Next we show that the prolongation procedure, when set up with vectors satisfying a specific sparsity pattern, yields a result again with a particular sparsity pattern which in turn can be determined by the provided formula.

Theorem 5.3. *For $d > d_2 > d_1 \ge 2$ and $q \ge 2$ with d_1 even and $d - d_2$ divisible by $d_2 - d_1$, let $\mathbf{v}^{(d_1)} \in \mathcal{E}^{(k_1)}_{d_1,q}$ and $\mathbf{v}^{(d_2)} \in \mathcal{E}^{(k_2)}_{d_2,q}$. Let*

$$\mathbf{M} := \mathbf{I}_{q^{d_1/2}} \otimes \mathrm{mat}_{q^{d_2-d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_2)}) \left(\mathrm{mat}_{q^{d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_1)})\right)^+,$$

where $\cdot^+$ denotes the pseudoinverse, and let

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{(d-d_2)/(d_2-d_1)} (\mathbf{I}_{q^{(d_2-d_1)i}} \otimes \mathbf{M})\right) \mathbf{w}^{(d_2)} \tag{5.33}$$

with $\mathbf{w}^{(d_2)} \in \mathcal{E}^{(K_2)}_{d_2,q}$. Then $\widetilde{\mathbf{v}}^{(d)} \in \mathcal{E}^{(\widetilde{k})}_{d,q}$, where

$$\widetilde{k} = K_2 + \frac{d-d_2}{d_2-d_1}(k_2 - k_1). \tag{5.34}$$

Proof. By Lemma 5.2, the sparsity patterns of potential nonzero entries of $\mathrm{mat}_{q^{d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_1)})$ and $\left(\mathrm{mat}_{q^{d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_1)})\right)^+$ are identical. So

$$\left(\mathrm{mat}_{q^{d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_1)})\right)^+ = \sum_{\alpha_1=0}^{q-1}\cdots\sum_{\alpha_{d_1}=0}^{q-1} \widehat{a}_{\alpha_1,\dots,\alpha_{d_1}} |\alpha_1\ \dots\ \alpha_{d_1/2}\rangle\langle\alpha_{d_1/2+1}\ \dots\ \alpha_{d_1}|,$$

where $\widehat{a}_{\alpha_1,\dots,\alpha_{d_1}} \neq 0$ only if $\alpha_1 + \dots + \alpha_{d_1} = k_1$. Just as well

$$\mathrm{mat}_{q^{d_2-d_1/2}\times q^{d_1/2}}(\mathbf{v}^{(d_2)}) = \sum_{\beta_1=0}^{q-1}\cdots\sum_{\beta_{d_2}=0}^{q-1} b_{\beta_1,\dots,\beta_{d_2}} |\beta_{d_1/2+1}\ \dots\ \beta_{d_2}\rangle\langle\beta_1\ \dots\ \beta_{d_1/2}|,$$

where $b_{\beta_1,\ldots,\beta_{d_2}} \neq 0$ only if $\beta_1 + \ldots + \beta_{d_2} = k_2$. Furthermore

$$\mathbf{I}_{q^{d_1/2}} = \sum_{\gamma_1=0}^{q-1} \cdots \sum_{\gamma_{d_1/2}=0}^{q-1} |\gamma_1 \ \ldots \ \gamma_{d_1/2}\rangle \langle \gamma_1 \ \ldots \ \gamma_{d_1/2}|,$$

$$\mathbf{I}_{q^{d_2-d_1}} = \sum_{\delta_1=0}^{q-1} \cdots \sum_{\delta_{d_2-d_1}=0}^{q-1} |\delta_1 \ \ldots \ \delta_{d_2-d_1}\rangle \langle \delta_1 \ \ldots \ \delta_{d_2-d_1}|,$$

and

$$\mathbf{w}^{(d_2)} = \sum_{\mu_1=0}^{q-1} \cdots \sum_{\mu_{d_2}=0}^{q-1} m_{\mu_1,\ldots,\mu_{d_2}} |\mu_1 \ \ldots \ \mu_{d_2}\rangle,$$

where $m_{\mu_1,\ldots,\mu_{d_2}} \neq 0$ only if $\mu_1 + \ldots + \mu_{d_2} = K_2$. Hence the result of the first prolongation step, $i = 1$ in (5.33), is

$$\begin{aligned}
\widetilde{\mathbf{v}}^{(d_2+d_2-d_1)} := & \sum_{\delta_1=0}^{q-1} \cdots \sum_{\delta_{d_2-d_1}=0}^{q-1} \sum_{\gamma_1=0}^{q-1} \cdots \sum_{\gamma_{d_1/2}=0}^{q-1} \sum_{\beta_1=0}^{q-1} \cdots \sum_{\beta_{d_2}=0}^{q-1} \sum_{\alpha_1=0}^{q-1} \cdots \sum_{\alpha_{d_1}=0}^{q-1} \sum_{\mu_1=0}^{q-1} \cdots \sum_{\mu_{d_2}=0}^{q-1} \\
& b_{\beta_1,\ldots,\beta_{d_2}} \widehat{a}_{\alpha_1,\ldots,\alpha_{d_1}} m_{\mu_1,\ldots,\mu_{d_2}} |\delta_1 \ \ldots \ \delta_{d_2-d_1} \ \gamma_1 \ \ldots \ \gamma_{d_1/2} \ \beta_{d_1/2+1} \ \ldots \ \beta_{d_2}\rangle \\
& \langle \beta_1 \ \ldots \ \beta_{d_1/2} | \alpha_1 \ \ldots \ \alpha_{d_1/2}\rangle \\
& \langle \delta_1 \ \ldots \ \delta_{d_2-d_1} \ \gamma_1 \ \ldots \ \gamma_{d_1/2} \ \alpha_{d_1/2+1} \ \ldots \ \alpha_{d_1} | \mu_1 \ \ldots \ \mu_{d_2}\rangle.
\end{aligned}$$

In order that $|\delta_1 \ \ldots \ \delta_{d_2-d_1} \ \gamma_1 \ \ldots \ \gamma_{d_1/2} \ \beta_{d_1/2+1} \ \ldots \ \beta_{d_2}\rangle$ contributes to $\widetilde{\mathbf{v}}^{(d_2+d_2-d_1)}$, it is necessary that

$$\langle \beta_1 \ \ldots \ \beta_{d_1/2} | \alpha_1 \ \ldots \ \alpha_{d_1/2}\rangle = 1$$

and

$$\langle \delta_1 \ \ldots \ \delta_{d_2-d_1} \ \gamma_1 \ \ldots \ \gamma_{d_1/2} \ \alpha_{d_1/2+1} \ \ldots \ \alpha_{d_1} | \mu_1 \ \ldots \ \mu_{d_2}\rangle = 1,$$

which is equivalent to

$$(\beta_1, \ldots, \beta_{d_1/2}) = (\alpha_1, \ldots, \alpha_{d_1/2})$$

and

$$\begin{aligned}
(\delta_1, \ldots, \delta_{d_2-d_1}, \gamma_1, \ldots, \gamma_{d_1/2}) &= (\mu_1, \ldots, \mu_{d_2-d_1/2}), \\
(\alpha_{d_1/2+1}, \ldots, \alpha_{d_1}) &= (\mu_{d_2-d_1/2+1}, \ldots, \mu_{d_2}).
\end{aligned}$$

We obtain that the sum of the indices for those $|\delta_1 \ \ldots \ \delta_{d_2-d_1} \ \gamma_1 \ \ldots \ \gamma_{d_1/2} \ \beta_{d_1/2+1} \ \ldots \ \beta_{d_2}\rangle$ which contribute to the result of the first prolongation step equals

$$\begin{aligned}
&\delta_1 + \ldots + \delta_{d_2-d_1} + \gamma_1 + \ldots + \gamma_{d_1/2} + \beta_{d_1/2+1} + \ldots + \beta_{d_2} \\
&= \mu_1 + \ldots + \mu_{d_2-d_1/2} + \beta_{d_1/2+1} + \ldots + \beta_{d_2} \\
&= \mu_1 + \ldots + \mu_{d_2-d_1/2} + k_2 - (\beta_1 + \ldots + \beta_{d_1/2}) \\
&= \mu_1 + \ldots + \mu_{d_2-d_1/2} + k_2 - (\alpha_1 + \ldots + \alpha_{d_1/2}) \\
&= \mu_1 + \ldots + \mu_{d_2-d_1/2} + k_2 - \big(k_1 - (\alpha_{d_1/2+1} + \ldots + \alpha_{d_1})\big) \\
&= \mu_1 + \ldots + \mu_{d_2-d_1/2} + \mu_{d_2-d_1/2+1} + \ldots + \mu_{d_2} + k_2 - k_1 \\
&= K_2 + k_2 - k_1.
\end{aligned}$$

In each subsequent prolongation step $i = 2, \dots, (d-d_2)/(d_2-d_1)$, there is analogously added $k_2 - k_1$ to the sum of digits of the q-ary representation of the contributing indices. □

Remark 5.4. (i) In the case of coupling parameters $A \neq B$ for the Hamilton operator $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$, the eigenvectors associated with simple eigenvalues are elements of $\mathcal{E}_{d,q}^{\mathrm{even}}$ or $\mathcal{E}_{d,q}^{\mathrm{odd}}$, see Corollary 2.10. With adapted suppositions in the formulation of Lemma 5.2 and Theorem 5.3, the corresponding proofs may be carried out analogously. To evaluate the formula in (5.34), one has to apply the calculation rules "even ± even = even", "odd ± odd = even", "even ± odd = odd".

(ii) Concerning the applicability of Lemma 5.2 and Theorem 5.3 to the q-Potts model, where eigenvectors associated with simple eigenvalues of $\mathbf{H}_{d,q}^{\mathrm{Potts}}$ are elements of $\mathcal{E}_{d,q}^{[k]}$, see Corollary 2.18, we notice that it causes no problems to regard the relevant quantities describing the sum of digits in the proofs modulo q.

(iii) The formula in (5.34) also applies to the situation of odd values of d_1 where the matrix $\mathbf{M}$ governing the prolongation is set up block-wise, cf. (5.13). In that case, the effective values of d_1 and d_2 are both reduced by 1 and the argumentation may also be carried out block-wise. With the convention that $\mathcal{E}_{d,q}^{(\kappa)}$ for $\kappa < 0$ contains only the zero vector in $\mathbb{R}^{q^d}$, we have that if $\mathbf{v}^{(d_1)} \in \mathcal{E}_{d_1,q}^{(k_1)}$ and $\mathbf{v}^{(d_2)} \in \mathcal{E}_{d_2,q}^{(k_2)}$, then for $1 \leq \eta \leq q$ it is

$$\left(\mathbf{v}^{(d_1)}\right)_{(\eta-1)q^{d_1-1}+1:\eta q^{d_1-1}} \in \mathcal{E}_{d_1-1,q}^{(k_1-\eta)}$$

and

$$\left(\mathbf{v}^{(d_2)}\right)_{(\eta-1)q^{d_2-1}+1:\eta q^{d_2-1}} \in \mathcal{E}_{d_2-1,q}^{(k_2-\eta)},$$

since the first digit in the q-ary representation of an index $(\eta-1)q^{d-1}+1, \dots, \eta q^{d-1}$, after a subtraction of 1, equals η. Thus, as $(k_2-\eta)-(k_1-\eta) = k_2-k_1$, the relevant quantity $k_2 - k_1$ in (5.34) does not change when we consider appropriate blocks in vectors.

5.2. Linear HT format

In Section 5.1 we described the construction of an initial guess in the full matrix/vector format. Now we transfer this procedure to the HT format. We first focus on the case $q = 2$ and assume the prolongation to be based on eigenvectors $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ for $(d_1, d_2) = (2, 4)$.

As in (5.5), set

$$\mathbf{N} = (n_{i,j})_{1\leq i\leq 8,\, 1\leq j\leq 2} := \mathrm{mat}_{8\times 2}(\mathbf{v}_{\min}^{(4)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1} \in \mathbb{R}^{8\times 2}.$$

With

$$\widehat{\mathbf{N}} := \begin{pmatrix} n_{1,1} & n_{5,1} \\ \vdots & \vdots \\ n_{4,1} & n_{8,1} \\ n_{1,2} & n_{5,2} \\ \vdots & \vdots \\ n_{4,2} & n_{8,2} \end{pmatrix} = \mathrm{resh}_{8\times 2}\left(\mathrm{resh}_{4\times 4}(\mathbf{N}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right),$$

consider the SVD

$$\widehat{\mathbf{N}} = \widehat{\mathbf{U}}^{(0)}\mathbf{\Sigma}^{(0)}(\mathbf{W}^{(0)})^\top = \begin{pmatrix}\widehat{\mathbf{u}}_1^{(0)} & \widehat{\mathbf{u}}_2^{(0)}\end{pmatrix}\begin{pmatrix}\sigma_1^{(0)} & 0\\ 0 & \sigma_2^{(0)}\end{pmatrix}\begin{pmatrix}\mathbf{w}_1^{(0)} & \mathbf{w}_2^{(0)}\end{pmatrix}^\top,$$

where $\widehat{\mathbf{u}}_i^{(0)} \in \mathbb{R}^8$, $\mathbf{w}_i^{(0)} \in \mathbb{R}^2$, $i \in \{1,2\}$. Setting

$$\mathbf{u}_i^{(0)} := \mathrm{resh}_{4\times 2}\left(\widehat{\mathbf{u}}_i^{(0)}\right),$$

it is

$$\mathbf{N} = \sigma_1^{(0)}\mathbf{w}_1^{(0)} \otimes \mathbf{u}_1^{(0)} + \sigma_2^{(0)}\mathbf{w}_2^{(0)} \otimes \mathbf{u}_2^{(0)}. \tag{5.35}$$

In the same manner we may further decompose $\mathbf{u}_i^{(0)}$. With

$$\begin{aligned}\mathrm{resh}_{4\times 2}\left(\mathrm{resh}_{2\times 4}\left(\mathbf{u}_i^{(0)}\right)\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}\right) &= \widehat{\mathbf{U}}^{(i)}\mathbf{\Sigma}^{(i)}(\mathbf{W}^{(i)})^\top\\ &= \begin{pmatrix}\widehat{\mathbf{u}}_1^{(i)} & \widehat{\mathbf{u}}_2^{(i)}\end{pmatrix}\begin{pmatrix}\sigma_1^{(i)} & 0\\ 0 & \sigma_2^{(i)}\end{pmatrix}\begin{pmatrix}\mathbf{w}_1^{(i)} & \mathbf{w}_2^{(i)}\end{pmatrix}^\top\end{aligned}$$

for $i \in \{1,2\}$, and by setting

$$\mathbf{u}_j^{(i)} := \mathrm{resh}_{2\times 2}\left(\widehat{\mathbf{u}}_j^{(i)}\right)$$

for $j \in \{1,2\}$, we obtain

$$\begin{aligned}\mathbf{N} = \quad & \sigma_1^{(0)}\mathbf{w}_1^{(0)} \otimes \left(\sigma_1^{(1)}\mathbf{w}_1^{(1)} \otimes \mathbf{u}_1^{(1)} + \sigma_2^{(1)}\mathbf{w}_2^{(1)} \otimes \mathbf{u}_2^{(1)}\right)\\ & + \sigma_2^{(0)}\mathbf{w}_2^{(0)} \otimes \left(\sigma_1^{(2)}\mathbf{w}_1^{(2)} \otimes \mathbf{u}_1^{(2)} + \sigma_2^{(2)}\mathbf{w}_2^{(2)} \otimes \mathbf{u}_2^{(2)}\right),\end{aligned} \tag{5.36}$$

where $\mathbf{w}_{1,2}^{(0),(1),(2)} \in \mathbb{R}^{2\times 1}$, $\mathbf{u}_{1,2}^{(1),(2)} \in \mathbb{R}^{2\times 2}$.

Due to the specific sparsity pattern of $\mathbf{N}$ caused by that of $\mathbf{v}_{\min}^{(2)}$ and $\mathbf{v}_{\min}^{(4)}$, we may for the decomposition (5.36) as well choose

$$\mathbf{w}_1^{(0),(1),(2)} = \begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{w}_2^{(0),(1),(2)} = \begin{pmatrix}0\\1\end{pmatrix}, \tag{5.37}$$

$$\begin{aligned}
\sigma_1^{(0)} &= \left\|\begin{pmatrix}n_{1,1} & n_{1,2}\\ n_{2,1} & n_{2,2}\\ n_{3,1} & n_{3,2}\\ n_{4,1} & n_{4,2}\end{pmatrix}\right\|, \quad \sigma_2^{(0)} = \left\|\begin{pmatrix}n_{5,1} & n_{5,2}\\ n_{6,1} & n_{6,2}\\ n_{7,1} & n_{7,2}\\ n_{8,1} & n_{8,2}\end{pmatrix}\right\|,\\
\sigma_1^{(1)} &= \left\|\begin{pmatrix}n_{1,1} & n_{1,2}\\ n_{2,1} & n_{2,2}\end{pmatrix}/\sigma_1^{(0)}\right\|, \quad \mathbf{u}_1^{(1)} = \begin{pmatrix}n_{1,1} & n_{1,2}\\ n_{2,1} & n_{2,2}\end{pmatrix}/(\sigma_1^{(0)}\sigma_1^{(1)}),\\
\sigma_2^{(1)} &= \left\|\begin{pmatrix}n_{3,1} & n_{3,2}\\ n_{4,1} & n_{4,2}\end{pmatrix}/\sigma_1^{(0)}\right\|, \quad \mathbf{u}_2^{(1)} = \begin{pmatrix}n_{3,1} & n_{3,2}\\ n_{4,1} & n_{4,2}\end{pmatrix}/(\sigma_1^{(0)}\sigma_2^{(1)}),\\
\sigma_1^{(2)} &= \left\|\begin{pmatrix}n_{5,1} & n_{5,2}\\ n_{6,1} & n_{6,2}\end{pmatrix}/\sigma_2^{(0)}\right\|, \quad \mathbf{u}_1^{(2)} = \begin{pmatrix}n_{5,1} & n_{5,2}\\ n_{6,1} & n_{6,2}\end{pmatrix}/(\sigma_2^{(0)}\sigma_1^{(2)}),\\
\sigma_2^{(2)} &= \left\|\begin{pmatrix}n_{7,1} & n_{7,2}\\ n_{8,1} & n_{8,2}\end{pmatrix}/\sigma_2^{(0)}\right\|, \quad \mathbf{u}_2^{(2)} = \begin{pmatrix}n_{7,1} & n_{7,2}\\ n_{8,1} & n_{8,2}\end{pmatrix}/(\sigma_2^{(0)}\sigma_2^{(2)}),
\end{aligned}$$

where $\mathbf{w}_1^{(i)}$ and $\mathbf{w}_2^{(i)}$ respectively $\mathbf{u}_1^{(i)}$ and $\mathbf{u}_2^{(i)}$ are also orthonormal for each fixed $i \in \{0,1,2\}$ respectively $i \in \{1,2\}$. The corresponding HT representative is given by

$$\mathfrak{H}_{\Psi(\mathbf{N})} = \Big[\mathcal{T}, (\mathbf{B}_t)_{t\in\mathcal{N}(\mathcal{T})}, (\mathbf{U}_t)_{t\in\mathcal{L}(\mathcal{T})}\Big]$$

with

$$\mathcal{T} = \begin{array}{ccccc} & & \{1,2,3\} & & \\ & \swarrow & & \searrow & \\ & \{1,2\} & & & \{3\} \\ \swarrow & & \searrow & & \\ \{1\} & & & \{2\} & \end{array},$$

$$\mathbf{B}_{\{1,2,3\}} = \begin{pmatrix} \sigma_1^{(0)} & 0 \\ 0 & \sigma_2^{(0)} \end{pmatrix},$$

$$(\mathbf{B}_{\{1,2\}})_{:,:,1} = \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad (\mathbf{B}_{\{1,2\}})_{:,:,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix},$$

$$\mathbf{U}_{\{1\}} = \begin{pmatrix} \mathrm{vec}(\mathbf{u}_1^{(1)}) & \mathrm{vec}(\mathbf{u}_2^{(1)}) & \mathrm{vec}(\mathbf{u}_1^{(2)}) & \mathrm{vec}(\mathbf{u}_2^{(2)}) \end{pmatrix},$$

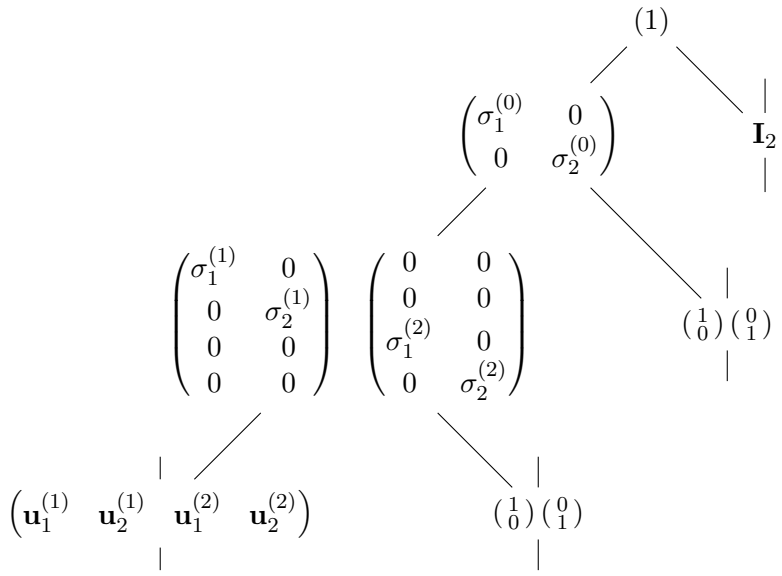
$$\mathbf{U}_{\{2\}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{U}_{\{3\}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $\Psi(\cdot)$ from (3.2), so in the present case $\Psi(\mathbf{N}) \in \mathbb{R}^{4\times 2\times 2}$.

Concerning the visualization of the prolongation, hence the product of some $\mathbf{I} \otimes \mathbf{M}$ with some vector $\mathbf{v}$, as a contraction of tensors, it is advantageous to draw the leaves of the HT representative of $\Psi(\mathbf{M})$ or $\Psi(\mathbf{N})$ not as matrices as it is inherent in the HT format, but rather to regard the r_t single columns at a leave $t = \{j\} \in \mathcal{L}(\mathcal{T})$ as r_t matrices of size $m_j \times n_j$. This way, there are two dangling legs at each leaf in the tensor network depicting $\mathfrak{H}_{\Psi(\mathbf{N})}$, one of which may be joined/contracted with the corresponding dangling leg of an HT representative of tens($\mathbf{v}$) for some $\mathbf{v}$ to be prolongated. With σ and $\mathbf{u}$ from (5.37) and the convention that the slices $(\mathbf{B}_t)_{:,:,1}, \dots, (\mathbf{B}_t)_{:,:,r_t}$ of a transfer tensor for $t \in \mathcal{N}(\mathcal{T})$ and the matricized columns of $\mathbf{U}_t$, $t \in \mathcal{L}(\mathcal{T})$, are stacked or written side by side at a node $t \in \mathcal{T}$ and perhaps are surrounded by additional brackets, we obtain the tensor network

$$\begin{array}{ccccc} & & \begin{pmatrix} \sigma_1^{(0)} & 0 \\ 0 & \sigma_2^{(0)} \end{pmatrix} & & \\ & \swarrow & & \searrow & \\ \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & & \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} & & \begin{matrix} | \\ \binom{1}{0}\binom{0}{1} \\ | \end{matrix} \\ | \quad \swarrow & & & \searrow \quad | & \\ \begin{pmatrix} \mathbf{u}_1^{(1)} & \mathbf{u}_2^{(1)} & \mathbf{u}_1^{(2)} & \mathbf{u}_2^{(2)} \end{pmatrix} & & & & \binom{1}{0}\binom{0}{1} \\ | & & & & | \end{array}, \tag{5.38}$$

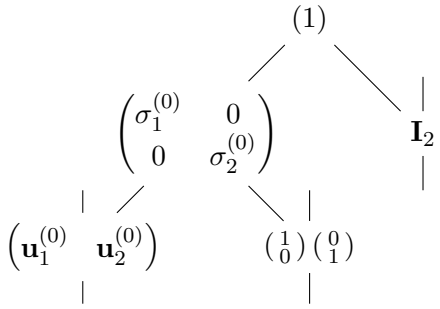
the upwards pointing legs of the leaves visualizing the action of $\Phi_{\Psi(\mathbf{N})}$, cf. (4.3). In order to keep the terminology simple, we call a tensor network like (5.38) with two dangling legs at each leaf node again an HT representative of $\Psi(\mathbf{N})$. Now, it follows that for $\mathbf{M} = \mathbf{I}_2 \otimes \mathbf{N}$ an HT representative of $\Psi(\mathbf{M})$ is given by



. (5.39)

So, the additional Kronecker factor $\mathbf{I}_2$ is appended at the top of the representative (5.38) of $\Psi(\mathbf{N})$.

If, instead of $(d_1, d_2) = (2, 4)$, the prolongation is based on $(d_1, d_2) = (2, 3)$ and $\mathbf{N} \in \mathbb{R}^{4\times 2}$ is defined by (5.1), already the decomposition (5.35) yields an HT representation of $\Psi(\mathbf{M}) \in \mathbb{R}^{4\times 2\times 4}$, $\mathbf{M} = \mathbf{I}_2 \otimes \mathbf{N}$, with smallest possible mode size since $\mathbf{w}_{1,2}^{(0)} \in \mathbb{R}^{2\times 1}$, $\mathbf{u}_{1,2}^{(0)} \in \mathbb{R}^{2\times 2}$. Hence, in this case, for $\Psi(\mathbf{M})$ we may choose the representative



,

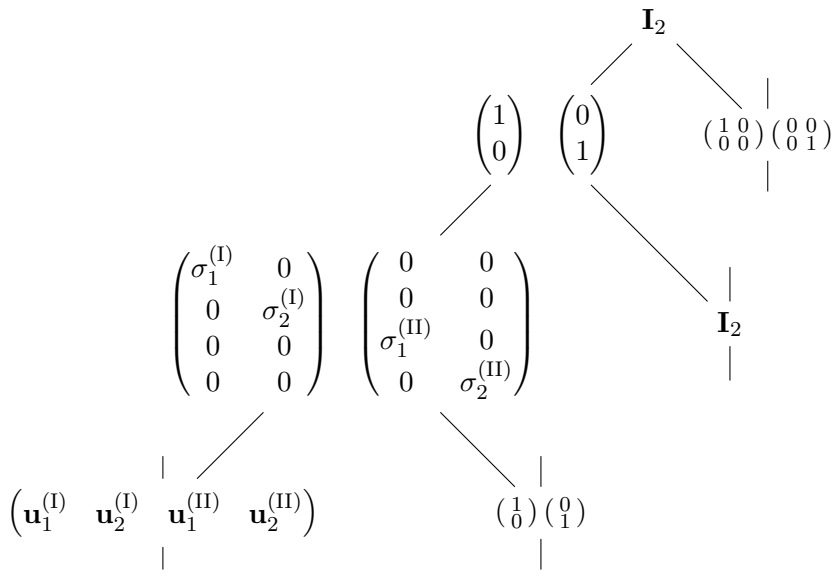
$$\sigma_1^{(0)} = \left\| \begin{pmatrix} n_{1,1} & n_{1,2} \\ n_{2,1} & n_{2,2} \end{pmatrix} \right\|, \quad \mathbf{u}_1^{(0)} = \begin{pmatrix} n_{1,1} & n_{1,2} \\ n_{2,1} & n_{2,2} \end{pmatrix} / \sigma_1^{(0)},$$
$$\sigma_2^{(0)} = \left\| \begin{pmatrix} n_{3,1} & n_{3,2} \\ n_{4,1} & n_{4,2} \end{pmatrix} \right\|, \quad \mathbf{u}_2^{(0)} = \begin{pmatrix} n_{3,1} & n_{3,2} \\ n_{4,1} & n_{4,2} \end{pmatrix} / \sigma_2^{(0)}.$$

As another example where the HT representation of $\Psi(\mathbf{M})$ is just read off the single entries

of $\mathbf{M}$, we consider the case $(d_1, d_2) = (3, 4)$ and, remembering (5.13), write

$$\begin{aligned}
\mathbf{M} = &\begin{pmatrix} \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{I})} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{II})} \end{pmatrix} \\
= &\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{I})} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{II})} \\
= &\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \mathbf{I}_2 \otimes \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} n_{1,1}^{(\mathrm{I})} & n_{1,2}^{(\mathrm{I})} \\ n_{2,1}^{(\mathrm{I})} & n_{2,2}^{(\mathrm{I})} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} n_{3,1}^{(\mathrm{I})} & n_{3,2}^{(\mathrm{I})} \\ n_{4,1}^{(\mathrm{I})} & n_{4,2}^{(\mathrm{I})} \end{pmatrix} \right) \\
&+ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \mathbf{I}_2 \otimes \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} n_{1,1}^{(\mathrm{II})} & n_{1,2}^{(\mathrm{II})} \\ n_{2,1}^{(\mathrm{II})} & n_{2,2}^{(\mathrm{II})} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} n_{3,1}^{(\mathrm{II})} & n_{3,2}^{(\mathrm{II})} \\ n_{4,1}^{(\mathrm{II})} & n_{4,2}^{(\mathrm{II})} \end{pmatrix} \right),
\end{aligned}$$

and obtain as a representative



, (5.40)

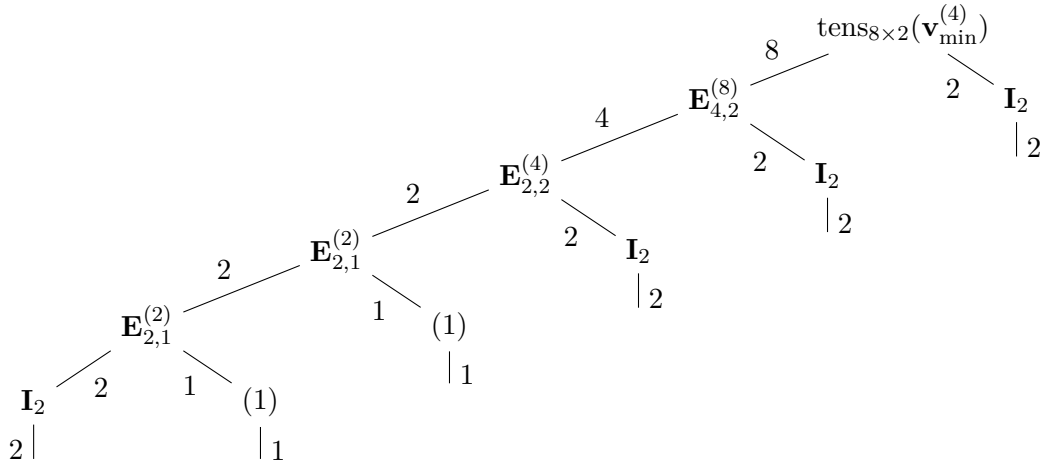
$$\begin{aligned}
\sigma_1^{(\mathrm{I})} &= \left\| \begin{pmatrix} n_{1,1}^{(\mathrm{I})} & n_{1,2}^{(\mathrm{I})} \\ n_{2,1}^{(\mathrm{I})} & n_{2,2}^{(\mathrm{I})} \end{pmatrix} \right\|, \quad \mathbf{u}_1^{(\mathrm{I})} = \begin{pmatrix} n_{1,1}^{(\mathrm{I})} & n_{1,2}^{(\mathrm{I})} \\ n_{2,1}^{(\mathrm{I})} & n_{2,2}^{(\mathrm{I})} \end{pmatrix} / \sigma_1^{(\mathrm{I})}, \\
\sigma_2^{(\mathrm{I})} &= \left\| \begin{pmatrix} n_{3,1}^{(\mathrm{I})} & n_{3,2}^{(\mathrm{I})} \\ n_{4,1}^{(\mathrm{I})} & n_{4,2}^{(\mathrm{I})} \end{pmatrix} \right\|, \quad \mathbf{u}_2^{(\mathrm{I})} = \begin{pmatrix} n_{3,1}^{(\mathrm{I})} & n_{3,2}^{(\mathrm{I})} \\ n_{4,1}^{(\mathrm{I})} & n_{4,2}^{(\mathrm{I})} \end{pmatrix} / \sigma_2^{(\mathrm{I})}, \\
\sigma_1^{(\mathrm{II})} &= \left\| \begin{pmatrix} n_{1,1}^{(\mathrm{II})} & n_{1,2}^{(\mathrm{II})} \\ n_{2,1}^{(\mathrm{II})} & n_{2,2}^{(\mathrm{II})} \end{pmatrix} \right\|, \quad \mathbf{u}_1^{(\mathrm{II})} = \begin{pmatrix} n_{1,1}^{(\mathrm{II})} & n_{1,2}^{(\mathrm{II})} \\ n_{2,1}^{(\mathrm{II})} & n_{2,2}^{(\mathrm{II})} \end{pmatrix} / \sigma_1^{(\mathrm{II})}, \\
\sigma_2^{(\mathrm{II})} &= \left\| \begin{pmatrix} n_{3,1}^{(\mathrm{II})} & n_{3,2}^{(\mathrm{II})} \\ n_{4,1}^{(\mathrm{II})} & n_{4,2}^{(\mathrm{II})} \end{pmatrix} \right\|, \quad \mathbf{u}_2^{(\mathrm{II})} = \begin{pmatrix} n_{3,1}^{(\mathrm{II})} & n_{3,2}^{(\mathrm{II})} \\ n_{4,1}^{(\mathrm{II})} & n_{4,2}^{(\mathrm{II})} \end{pmatrix} / \sigma_2^{(\mathrm{II})}.
\end{aligned}$$

For general (d_1, d_2), the tensorization $\Psi(\mathbf{M}) \in \mathbb{R}^{m_1 n_1 \times \cdots \times m_{d_2} n_{d_2}}$ of the matrix $\mathbf{M} \in \mathbb{R}^{2^{d_2} \times 2^{d_1}}$ with $\mathbf{M}\mathbf{v}_{\min}^{(d_1)} = \mathbf{v}_{\min}^{(d_2)}$ may also be represented in HT format according to its block structure and just using the single entries. Thereby $m_j = 2$ for all $1 \le j \le d_2$, but as $\mathbf{M}$ is not quadratic, the actual value of $n_j \in \{1, 2\}$ depends on the particular construction of $\mathbf{M}$. For example in the case d_1 even and $\mathbf{M} = \mathbf{I}_{2^{d_1/2}} \otimes \mathbf{N}$, $\mathbf{N} \in \mathbb{R}^{2^{d_2 - d_1/2} \times 2^{d_1/2}}$, it is

$n_1 = \ldots = n_{d_1/2} = 2$, $n_{d_1/2+1} = \ldots = n_{d_1/2+(d_2-d_1)} = 1$, $n_{d_2-d_1/2+1} = \ldots = n_{d_2} = 2$. In the most simple case, the underlying dimension tree $\mathcal{T}$ is such that each transfer tensor including the root node has at most only one, say always the left, child which is itself a transfer tensor, the other (right) child being a leaf. This yields the linear tree, see Definition 3.6(ii).

Given the HT representation of $\Psi(\mathbf{M})$, the tensorizations of the matrices $(\mathbf{I}_{2^{(d_2-d_1)i}} \otimes \mathbf{M})$, $1 \le i \le (d-d_2)/(d_2-d_1)$, cf. (5.12), needed to prolongate to the desired problem size d, are represented by appending at the top of the tensor tree $(d_2-d_1)i$ identity matrices $\mathbf{I}_2$ as leaves and an appropriate number of transfer tensors of size $1 \times 1 \times 1$ with entry 1.

Returning to the example $(d_1, d_2) = (2, 4)$, in order to perform the first prolongation step from d_2 to $d_2 + (d_2 - d_1) = 4 + 2 = 6$ in HT format via an application of (5.39), we need to represent the tensorization $\mathrm{tens}_{n_1 \times \cdots \times n_6}(\mathbf{v}^{(4)}_{\min})$, where $n_1 = n_4 = n_5 = n_6 = 2$ and $n_2 = n_3 = 1$, with the same linear dimension tree $\mathcal{T}$ like $\Psi(\mathbf{M})$. The perhaps most obvious way would be to set the root transfer tensor $\mathbf{B}_{\mathrm{root}} = \mathrm{tens}_{n_1 \cdots n_5 \times n_6}(\mathbf{v}^{(4)}_{\min})$, the leaves $\mathbf{U}_{\{j\}} = \mathbf{I}_{n_j}$, and the remaining transfer tensors as appropriately tensorized identity matrices of suitable size, so to consider



, (5.41)

where

$$\mathbf{E}^{(8)}_{4,2} := \mathrm{resh}_{4\times 2\times 8}(\mathbf{I}_8) = \begin{array}{cccccc} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} & \cdots & \\ \cdots & \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} & \end{array} \in \mathbb{R}^{4\times 2\times 8}$$

and

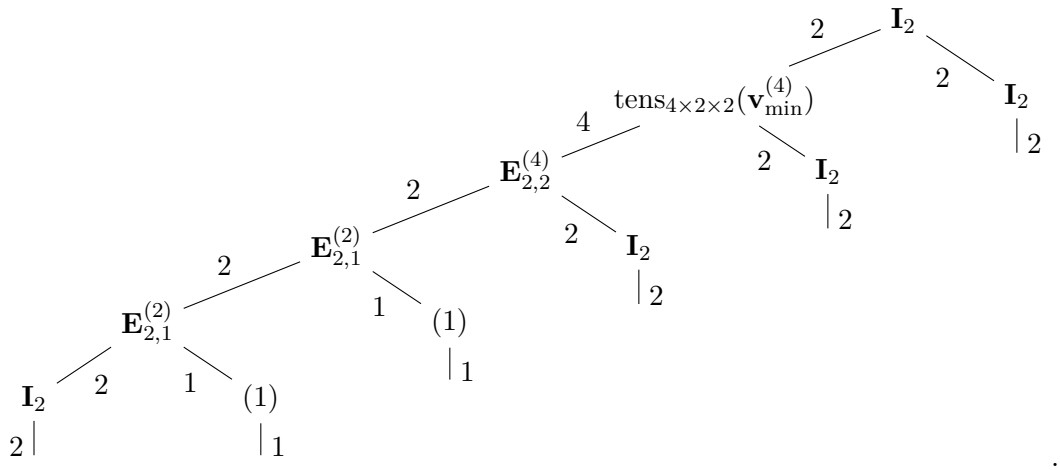
$$\mathbf{E}^{(4)}_{2,2} := \mathrm{resh}_{2\times 2\times 4}(\mathbf{I}_4) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{E}^{(2)}_{2,1} := \mathrm{resh}_{2\times 1\times 2}(\mathbf{I}_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

But, since

$$\mathbf{E}_{4,2}^{(8)} \boxdot_3^1 \mathrm{tens}_{8\times 2}(\mathbf{v}_{\min}^{(4)}) = \mathrm{tens}_{4\times 2\times 2}(\mathbf{v}_{\min}^{(4)}) = \mathrm{tens}_{4\times 2\times 2}(\mathbf{v}_{\min}^{(4)}) \boxdot_3^1 \mathbf{I}_2,$$

as a result reducing $r_{\{1,\dots,5\}}$ from 8 to 2, we choose for $\mathrm{tens}_{2\times 1\times 1\times 2\times 2\times 2}(\mathbf{v}_{\min}^{(4)})$ the representative

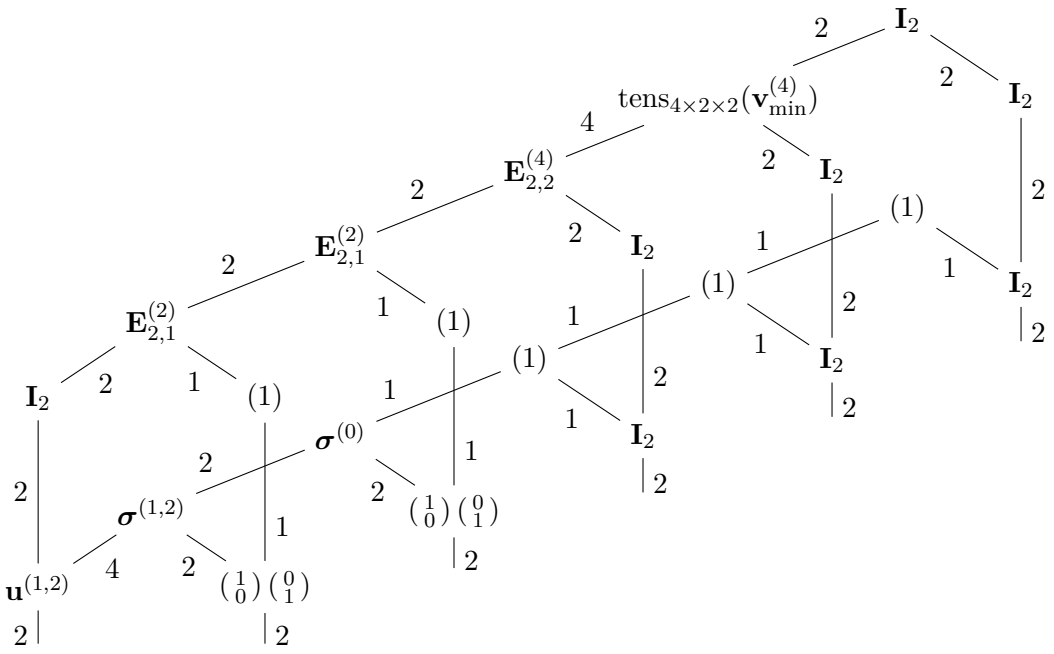


(5.42)

Now, the first prolongation step

$$\Phi_{\Psi(\mathbf{I}_4\otimes\mathbf{M})}\big(\mathrm{tens}_{2\times 1\times 1\times 2\times 2\times 2}(\mathbf{v}_{\min}^{(4)})\big)$$

is realized by connecting the respective dangling legs which yields

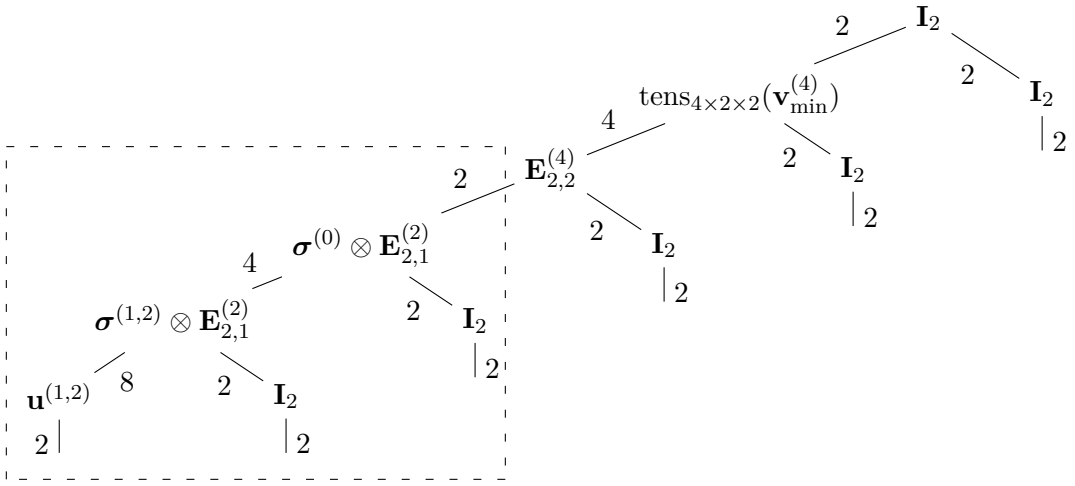


(5.43)

where we abbreviate

$$\boldsymbol{\sigma}^{(0)} := \begin{pmatrix} \sigma_1^{(0)} & 0 \\ 0 & \sigma_2^{(0)} \end{pmatrix}, \quad \boldsymbol{\sigma}^{(1,2)} := \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix},$$
$$\mathbf{u}^{(1,2)} := \begin{pmatrix} \mathbf{u}_1^{(1)} & \mathbf{u}_2^{(1)} & \mathbf{u}_1^{(2)} & \mathbf{u}_2^{(2)} \end{pmatrix},$$

and write the mode sizes at the dangling legs as well as the HT ranks at the connecting legs. Remembering Subsection 3.3.4 on arithmetical operations with HT tensors, a representative of (5.43) is given by



, (5.44)

noticing

$$\boldsymbol{\sigma}^{(1,2)} \otimes \mathbf{E}_{2,1}^{(2)} = \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \sigma_1^{(1)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(1)} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \sigma_1^{(2)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} \in \mathbb{R}^{8\times2\times4}$$

and

$$\boldsymbol{\sigma}^{(0)} \otimes \mathbf{E}_{2,1}^{(2)} = \begin{pmatrix} \sigma_1^{(0)} & 0 \\ 0 & \sigma_2^{(0)} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sigma_1^{(0)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(0)} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \sigma_1^{(0)} & 0 \\ 0 & 0 \\ 0 & \sigma_2^{(0)} \end{pmatrix} \in \mathbb{R}^{4\times2\times2}.$$

Let us investigate in detail the content of the dashed box in (5.44). Since

$$\begin{aligned}
&\sigma_1^{(1)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \left(\mathbf{u}_1^{(1)}\right)_{:,1} + \sigma_2^{(1)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \left(\mathbf{u}_2^{(1)}\right)_{:,1} \\
&= \quad \sigma_1^{(1)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \left(\left(\mathbf{u}_1^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left(\mathbf{u}_1^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\
&\quad + \sigma_2^{(1)} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \left(\left(\mathbf{u}_2^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left(\mathbf{u}_2^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\
&= \quad \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{2,1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&\quad + \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{1,1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned} \tag{5.45}$$

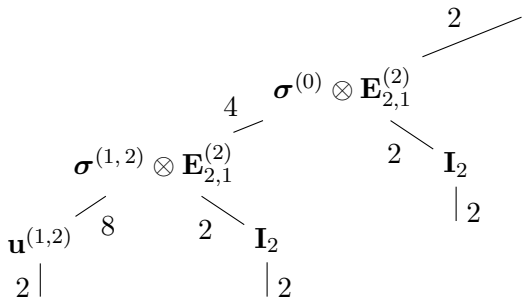
as well as

$$\begin{pmatrix} \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{1,1} & \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{2,1} \\ \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{1,1} & \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{2,1} \end{pmatrix}^\top = \begin{pmatrix} \left(\mathbf{u}_1^{(1)}\right)_{:,1} & \left(\mathbf{u}_2^{(1)}\right)_{:,1} \end{pmatrix} \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \end{pmatrix}, \tag{5.46}$$

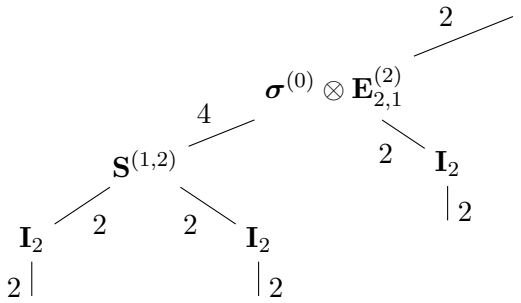
and by analogous computations for the coefficients in the other three slices of $\boldsymbol{\sigma}^{(1,2)} \otimes \mathbf{E}_{2,1}^{(2)}$, we may, setting

$$\mathbf{S}^{(1,2)} := \begin{pmatrix} \left(\begin{pmatrix} \left(\mathbf{u}_1^{(1)}\right)_{:,1} & \left(\mathbf{u}_2^{(1)}\right)_{:,1} \end{pmatrix} \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \end{pmatrix} \right) \\ \left(\begin{pmatrix} \left(\mathbf{u}_1^{(1)}\right)_{:,2} & \left(\mathbf{u}_2^{(1)}\right)_{:,2} \end{pmatrix} \begin{pmatrix} \sigma_1^{(1)} & 0 \\ 0 & \sigma_2^{(1)} \end{pmatrix} \right) \\ \left(\begin{pmatrix} \left(\mathbf{u}_1^{(2)}\right)_{:,1} & \left(\mathbf{u}_2^{(2)}\right)_{:,1} \end{pmatrix} \begin{pmatrix} \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} \right) \\ \left(\begin{pmatrix} \left(\mathbf{u}_1^{(2)}\right)_{:,2} & \left(\mathbf{u}_2^{(2)}\right)_{:,2} \end{pmatrix} \begin{pmatrix} \sigma_1^{(2)} & 0 \\ 0 & \sigma_2^{(2)} \end{pmatrix} \right) \end{pmatrix} \in \mathbb{R}^{2\times 2\times 4}, \tag{5.47}$$

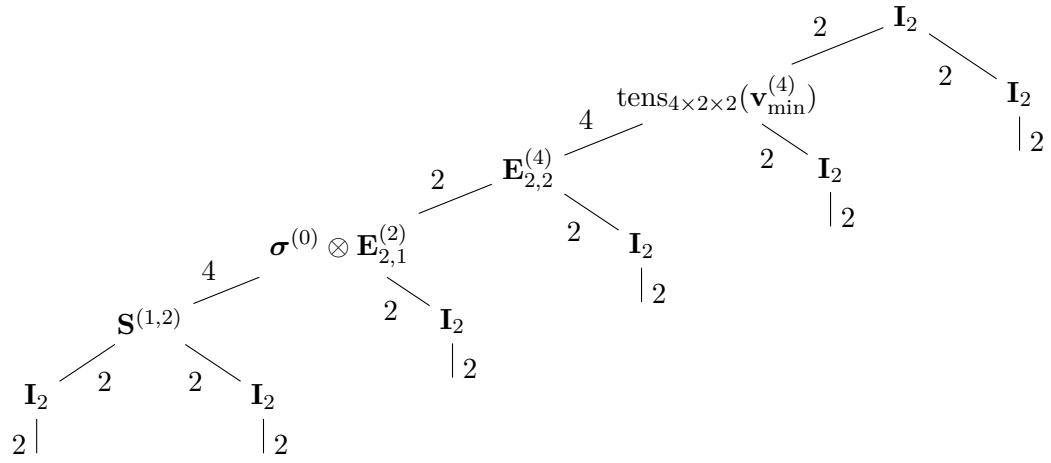
replace



in (5.44) by



and obtain the representative

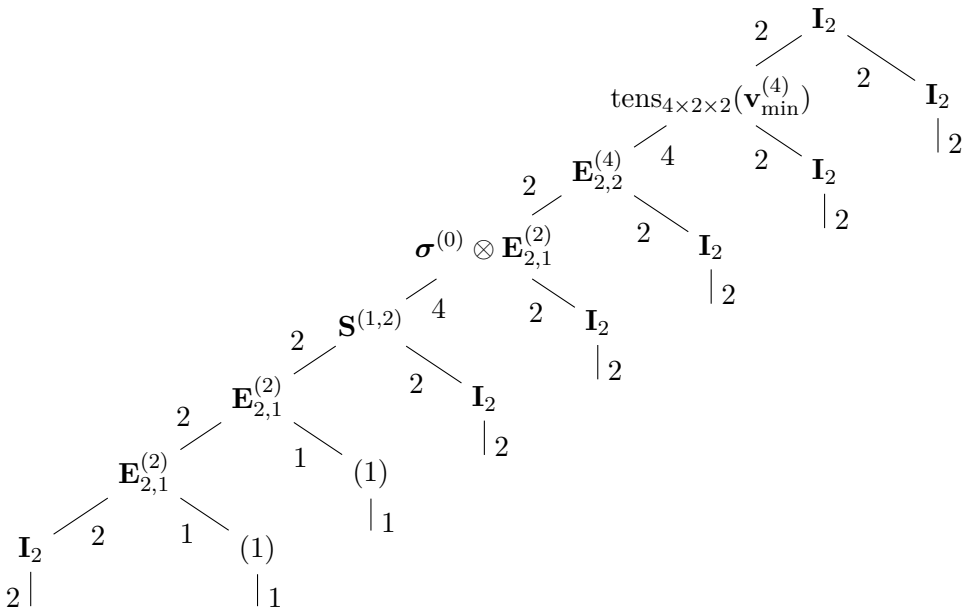


(5.48)

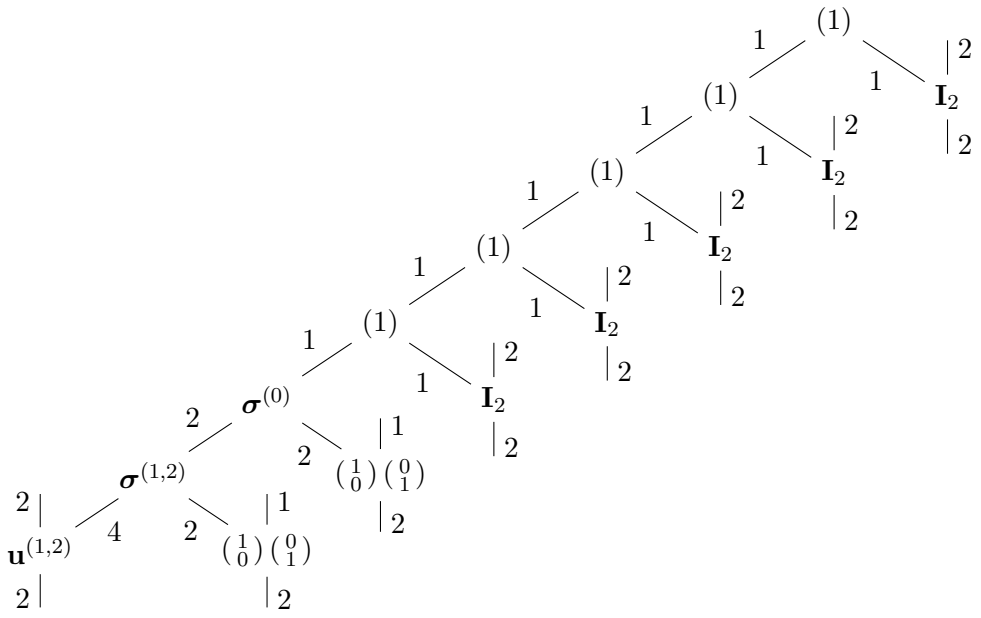
of

$$\widetilde{\mathbf{V}}^{(6)} := \Phi_{\Psi(\mathbf{I}_4 \otimes \mathbf{M})}\big(\mathrm{tens}_{2\times1\times1\times2\times2\times2}(\mathbf{v}_{\min}^{(4)})\big).$$

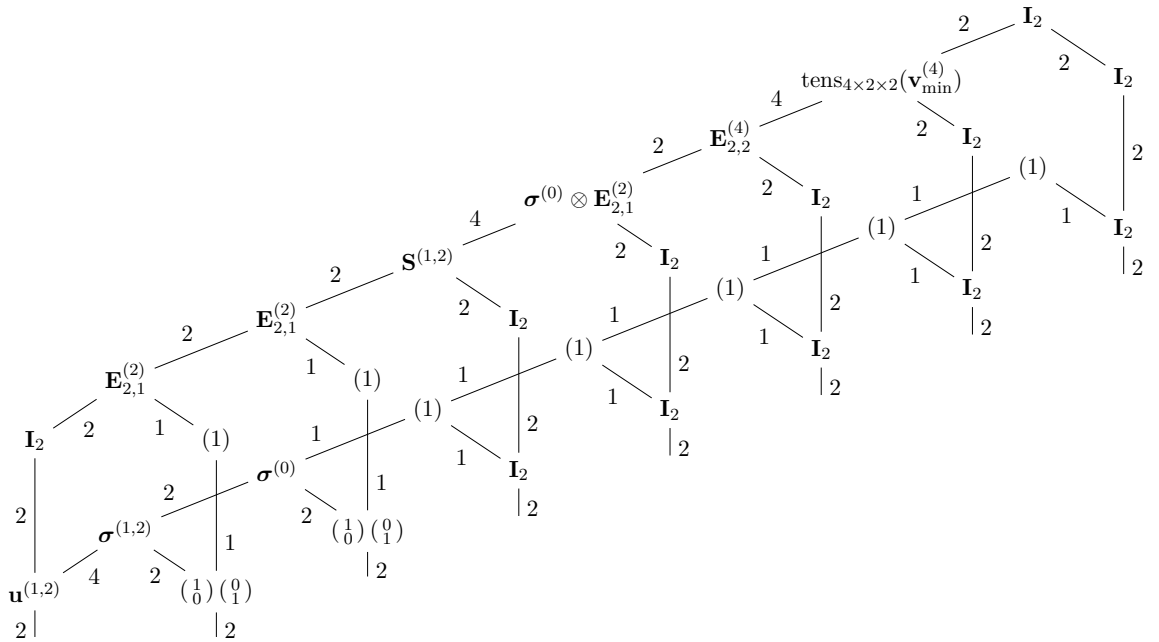
Now we comment on the benefit of this observation. If, for the next prolongation step to problem size $d_2 + 2(d_2 - d_1) = 8$, the representative (5.48) of $\widetilde{\mathbf{V}}^{(6)}$ is prepared via



in order to be contracted with the representative

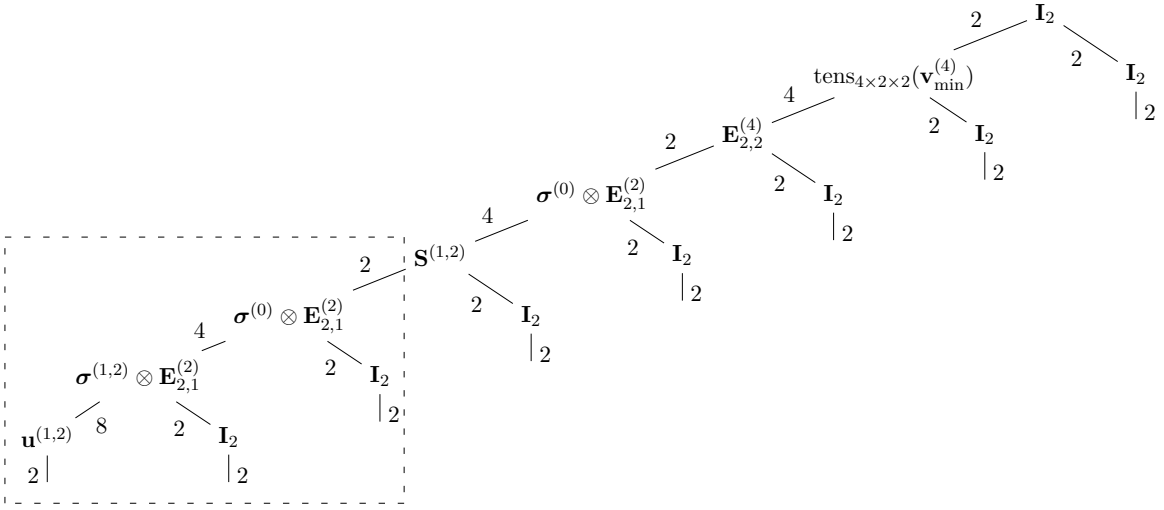


of $\Psi(\mathbf{I}_{16} \otimes \mathbf{M})$, yielding



,

which equals

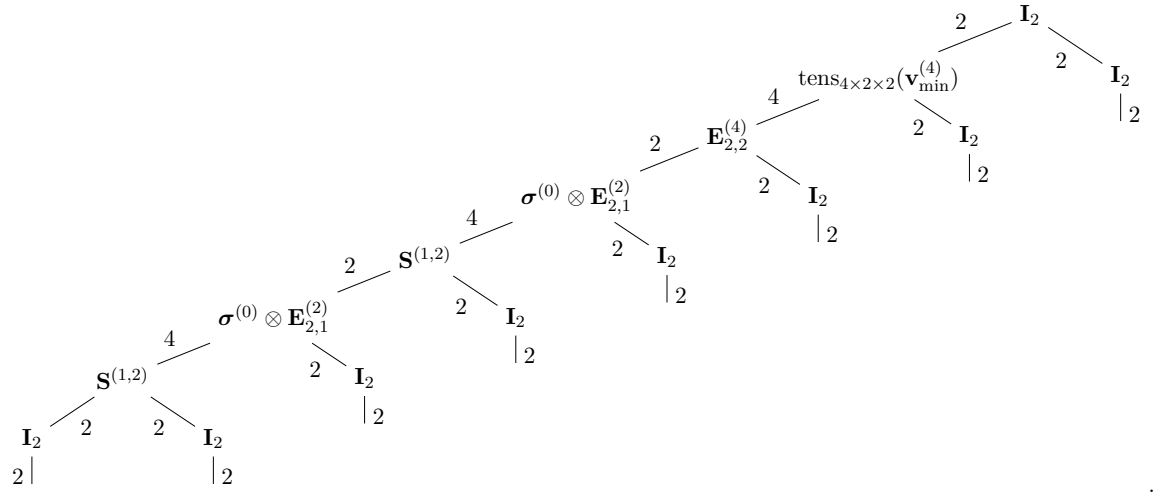


(5.49)

then the dashed box in (5.49) has the same content as the box in (5.44), and hence with (5.45) and (5.46), we may represent

$$\widetilde{\mathbf{V}}^{(8)} := \Phi_{\Psi(\mathbf{I}_{16}\otimes\mathbf{M})}\big(\widetilde{\mathbf{V}}^{(6)}\big)$$

by



.

Iterating this procedure, a representative of

$$\mathrm{tens}_{2\times d}\big((\mathbf{I}_{2^{d-4}}\otimes\mathbf{M})(\mathbf{I}_{2^{d-6}}\otimes\mathbf{M})\cdots(\mathbf{I}_4\otimes\mathbf{M})(\mathbf{I}_2\otimes\mathbf{M})\mathbf{v}^{(4)}_{\min}\big)$$

for arbitrary even d is obtained with all leaves $\mathbf{U}_{\{1\}},\dots,\mathbf{U}_{\{d\}}$ equal to $\mathbf{I}_2$ and transfer tensors

$$\begin{aligned}
\mathbf{B}_{\{1,\dots,d\}} &= \mathbf{I}_2, \quad \mathbf{B}_{\{1,\dots,d-1\}} = \mathrm{tens}_{4\times2\times2}(\mathbf{v}^{(4)}_{\min}), \quad \mathbf{B}_{\{1,\dots,d-2\}} = \mathbf{E}^{(4)}_{2,2},\\
\mathbf{B}_{\{1,\dots,d-3\}} &= \mathbf{B}_{\{1,\dots,d-5\}} = \dots = \mathbf{B}_{\{1,\dots,3\}} = \boldsymbol{\sigma}^{(0)}\otimes\mathbf{E}^{(2)}_{2,1},\\
\mathbf{B}_{\{1,\dots,d-4\}} &= \mathbf{B}_{\{1,\dots,d-6\}} = \dots = \mathbf{B}_{\{1,2\}} = \mathbf{S}^{(1,2)},
\end{aligned} \tag{5.50}$$

so the HT ranks are given by

$$r_{\{1\}} = \ldots = r_{\{d\}} = 2,$$
$$r_{\{1,\ldots,d-1\}} = r_{\{1,\ldots,d-3\}} = \ldots = r_{\{1,\ldots,3\}} = 2, \quad r_{\{1,\ldots,d-2\}} = r_{\{1,\ldots,d-4\}} = \ldots = r_{\{1,2\}} = 4.$$

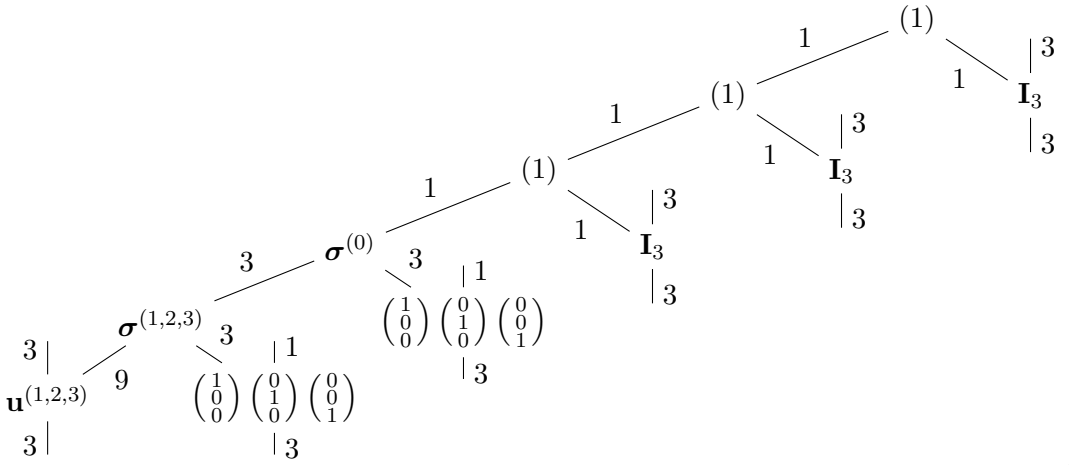
The whole described procedure, with suitable modifications, translates to the case $q = 3$. We stay at the example $(d_1, d_2) = (2, 4)$. It is $\mathbf{M} = \mathbf{I}_3 \otimes \mathbf{N}$ with

$$\mathbf{N} = \mathrm{mat}_{27\times 3}(\mathbf{v}^{(4)}_{\min}) \left(\mathrm{mat}_{3\times 3}(\mathbf{v}^{(2)}_{\min})\right)^{-1} \in \mathbb{R}^{27\times 3}.$$

If $1 \le j \le 3$ and

$$\begin{aligned}
\sigma^{(0)}_1 &= \|\mathbf{N}_{1:9,:}\|, \quad \sigma^{(0)}_2 = \|\mathbf{N}_{10:18,:}\|, \quad \sigma^{(0)}_3 = \|\mathbf{N}_{19:27,:}\| \\
\sigma^{(1)}_j &= \left\|\mathbf{N}_{3(j-1)+1:3(j-1)+3,:}/\sigma^{(0)}_1\right\|, \quad \mathbf{u}^{(1)}_j = \mathbf{N}_{3(j-1)+1:3(j-1)+3,:}/(\sigma^{(0)}_1\sigma^{(1)}_j), \\
\sigma^{(2)}_j &= \left\|\mathbf{N}_{3(j+2)+1:3(j+2)+3,:}/\sigma^{(0)}_2\right\|, \quad \mathbf{u}^{(2)}_j = \mathbf{N}_{3(j+2)+1:3(j+2)+3,:}/(\sigma^{(0)}_2\sigma^{(2)}_j), \\
\sigma^{(3)}_j &= \left\|\mathbf{N}_{3(j+5)+1:3(j+5)+3,:}/\sigma^{(0)}_3\right\|, \quad \mathbf{u}^{(3)}_j = \mathbf{N}_{3(j+5)+1:3(j+5)+3,:}/(\sigma^{(0)}_3\sigma^{(3)}_j),
\end{aligned}$$

an HT representative of $\Psi(\mathbf{I}_9 \otimes \mathbf{M})$ is given by



,

where

$$\boldsymbol{\sigma}^{(0)} := \begin{pmatrix} \sigma^{(0)}_1 & 0 & 0 \\ 0 & \sigma^{(0)}_2 & 0 \\ 0 & 0 & \sigma^{(0)}_3 \end{pmatrix},$$

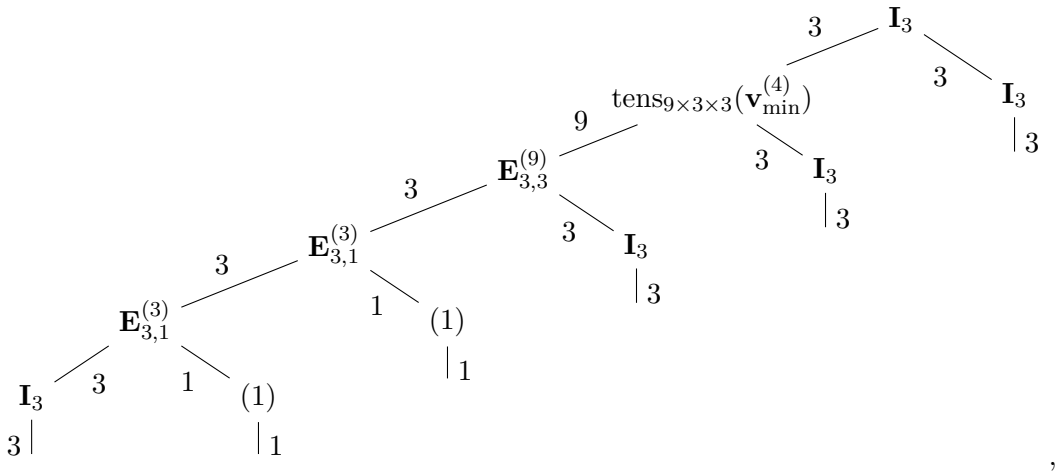
$$\boldsymbol{\sigma}^{(1,2,3)} := \begin{pmatrix} \sigma^{(1)}_1 & 0 & 0 \\ 0 & \sigma^{(1)}_2 & 0 \\ 0 & 0 & \sigma^{(1)}_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sigma^{(2)}_1 & 0 & 0 \\ 0 & \sigma^{(2)}_2 & 0 \\ 0 & 0 & \sigma^{(2)}_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sigma^{(3)}_1 & 0 & 0 \\ 0 & \sigma^{(3)}_2 & 0 \\ 0 & 0 & \sigma^{(3)}_3 \end{pmatrix},$$

$$\mathbf{u}^{(1,2,3)} := \begin{pmatrix} \mathbf{u}^{(1)}_1 & \mathbf{u}^{(1)}_2 & \mathbf{u}^{(1)}_3 & \mathbf{u}^{(2)}_1 & \mathbf{u}^{(2)}_2 & \mathbf{u}^{(2)}_3 & \mathbf{u}^{(3)}_1 & \mathbf{u}^{(3)}_2 & \mathbf{u}^{(3)}_3 \end{pmatrix}.$$

In order to realize the first prolongation step from $d_2 = 4$ to $d_2 + (d_2 - d_1) = 6$ via

$$\widetilde{\mathbf{V}}^{(6)} := \Phi_{\Psi(\mathbf{I}_9 \otimes \mathbf{M})}\big(\mathrm{tens}_{3\times1\times1\times3\times3\times3}(\mathbf{v}_{\min}^{(4)})\big),$$

the representative of $\mathrm{tens}_{3\times1\times1\times3\times3\times3}(\mathbf{v}_{\min}^{(4)})$ reads



,

where $\mathbf{E}_{3,3}^{(9)} := \mathrm{resh}_{3\times3\times9}(\mathbf{I}_9)$ and $\mathbf{E}_{3,1}^{(3)} := \mathrm{resh}_{3\times1\times3}(\mathbf{I}_3)$. Analogously to the discussion of the content of the dashed box in (5.44) by means of (5.45)-(5.48), we observe exemplarily that for the first of the 9 slices of $\boldsymbol{\sigma}^{(1,2,3)} \otimes \mathbf{E}_{3,1}^{(3)} \in \mathbb{R}^{27\times3\times9}$ it is

$$\big(\boldsymbol{\sigma}^{(1,2,3)} \otimes \mathbf{E}_{3,1}^{(3)}\big)_{:,:,1} = \begin{pmatrix} \sigma_1^{(1)} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sigma_2^{(1)} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_3^{(1)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix},$$

and we compute

$$
\begin{aligned}
&\sigma_1^{(1)} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \left(\mathbf{u}_1^{(1)}\right)_{:,1} + \sigma_2^{(1)} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \left(\mathbf{u}_2^{(1)}\right)_{:,1} + \sigma_3^{(1)} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \left(\mathbf{u}_3^{(1)}\right)_{:,1} \\
&= \quad \sigma_1^{(1)} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \left(\left(\mathbf{u}_1^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left(\mathbf{u}_1^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \left(\mathbf{u}_1^{(1)}\right)_{3,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&\quad + \sigma_2^{(1)} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \left(\left(\mathbf{u}_2^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left(\mathbf{u}_2^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \left(\mathbf{u}_2^{(1)}\right)_{3,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&\quad + \sigma_3^{(1)} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \left(\left(\mathbf{u}_3^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left(\mathbf{u}_3^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \left(\mathbf{u}_3^{(1)}\right)_{3,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&= \quad \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{1,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{2,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{3,1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&\quad + \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{1,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{3,1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&\quad + \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{1,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{2,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{3,1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}
$$

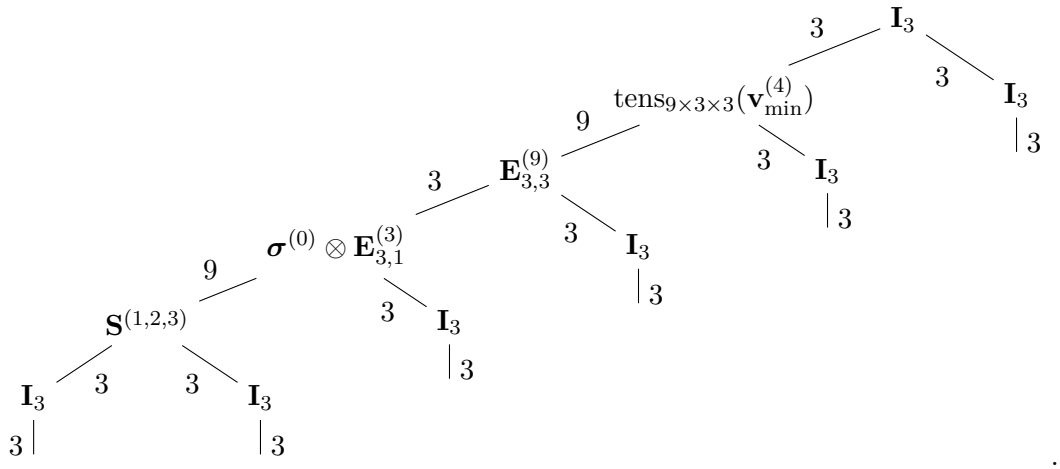
as well as

$$
\begin{aligned}
&\begin{pmatrix} \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{1,1} & \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{2,1} & \sigma_1^{(1)} \left(\mathbf{u}_1^{(1)}\right)_{3,1} \\ \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{1,1} & \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{2,1} & \sigma_2^{(1)} \left(\mathbf{u}_2^{(1)}\right)_{3,1} \\ \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{1,1} & \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{2,1} & \sigma_3^{(1)} \left(\mathbf{u}_3^{(1)}\right)_{3,1} \end{pmatrix}^\top \\
&= \begin{pmatrix} \left(\mathbf{u}_1^{(1)}\right)_{:,1} & \left(\mathbf{u}_2^{(1)}\right)_{:,1} & \left(\mathbf{u}_3^{(1)}\right)_{:,1} \end{pmatrix} \begin{pmatrix} \sigma_1^{(1)} & 0 & 0 \\ 0 & \sigma_2^{(1)} & 0 \\ 0 & 0 & \sigma_3^{(1)} \end{pmatrix}.
\end{aligned}
$$

Setting

$$
\begin{aligned}
&\mathbf{S}^{(1,2,3)} \in \mathbb{R}^{3\times 3\times 9}, \\
&\left(\mathbf{S}^{(1,2,3)}\right)_{:,:,3(j-1)+i} := \begin{pmatrix} \left(\mathbf{u}_1^{(j)}\right)_{:,i} & \left(\mathbf{u}_2^{(j)}\right)_{:,i} & \left(\mathbf{u}_3^{(j)}\right)_{:,i} \end{pmatrix} \begin{pmatrix} \sigma_1^{(j)} & 0 & 0 \\ 0 & \sigma_2^{(j)} & 0 \\ 0 & 0 & \sigma_3^{(j)} \end{pmatrix}, \quad 1 \le i, j \le 3,
\end{aligned}
$$

we obtain that a representative of $\widetilde{\mathbf{V}}^{(6)}$ is given by



By iterating this way, we arrive at the transfer tensors

$$\begin{aligned}\mathbf{B}_{\{1,\dots,d\}} &= \mathbf{I}_3, \quad \mathbf{B}_{\{1,\dots,d-1\}} = \mathrm{tens}_{9\times3\times3}(\mathbf{v}^{(4)}_{\min}), \quad \mathbf{B}_{\{1,\dots,d-2\}} = \mathbf{E}^{(9)}_{3,3},\\ \mathbf{B}_{\{1,\dots,d-3\}} &= \mathbf{B}_{\{1,\dots,d-5\}} = \dots = \mathbf{B}_{\{1,\dots,3\}} = \boldsymbol{\sigma}^{(0)}\otimes\mathbf{E}^{(3)}_{3,1},\\ \mathbf{B}_{\{1,\dots,d-4\}} &= \mathbf{B}_{\{1,\dots,d-6\}} = \dots = \mathbf{B}_{\{1,2\}} = \mathbf{S}^{(1,2,3)},\end{aligned}$$

cf. (5.50), and the leaf matrices $\mathbf{U}_{\{1\}} = \dots = \mathbf{U}_{\{d\}} = \mathbf{I}_3$, constituting for arbitrary even d an HT representative of

$$\mathrm{tens}_{3\times d}\big((\mathbf{I}_{3^{d-4}}\otimes\mathbf{M})(\mathbf{I}_{3^{d-6}}\otimes\mathbf{M})\cdots(\mathbf{I}_9\otimes\mathbf{M})(\mathbf{I}_3\otimes\mathbf{M})\mathbf{v}^{(4)}_{\min}\big).$$

For other choices of (d_1, d_2) similar observations may be made. Having once determined the transfer tensors in a representative of a prolongated $\widetilde{\mathbf{V}}^{(d_2+k(d_2-d_1))}$ for a large enough k, a representative of $\widetilde{\mathbf{V}}^{(d_2+n(d_2-d_1))}$ for arbitrary n is obtained by repeating a suitable set of $d_2 - d_1$ consecutive transfer tensors.

Remark 5.5. We refrain from describing the construction procedure of the prolongated initial guess in HT format with linear dimension tree for other (d_1, d_2) than $(2, 4)$ in detail. Instead we provide a list with the HT ranks of the respective representatives. We mention those (d_1, d_2) and d which are used in the numerical tests of Chapter 6. Due to $\mathbf{U}_t = \mathbf{I}_q$ for $t \in \mathcal{L}(\mathcal{T})$, it is $r_{\{1\}} = \dots = r_{\{d\}} = q$ in each scenario. So we list only r_t for $t \in \mathcal{N}(\mathcal{T}) \setminus \{1, \dots, d\}$, supplemented for reasons of symmetry by $r_{\{1\}}$, and abbreviate

$$\mathbf{r}^{(d)} := \Big(r_{\{1\}}, r_{\{1,2\}}, r_{\{1,2,3\}}, \dots, r_{\{1,\dots,d-2\}}, r_{\{1,\dots,d-1\}}\Big).$$

In addition, we write down the matrices $\mathbf{M}$ governing the prolongation, cf. Section 5.1, where submatrices of $\mathbf{M}$ which are not specified only contain zeros. Especially in case $q = 3$, invertibility of the matricizations of $\mathbf{v}^{(d_1)}_{\min}$ or parts thereof might not be given and we refer to the discussion in Subsection 5.1.2.

(i) $q = 2$, $(d_1, d_2) = (2, 3)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_2 \otimes \mathrm{mat}_{4\times2}(\mathbf{v}^{(3)}_{\min})\Big(\mathrm{mat}_{2\times2}(\mathbf{v}^{(2)}_{\min})\Big)^{-1},\\ \mathbf{r}^{(16)} &= (2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2),\\ \mathbf{r}^{(22)} &= (2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2).\end{aligned}$$

(ii) $q = 2$, $(d_1, d_2) = (2, 4)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_2 \otimes \mathrm{mat}_{8\times 2}(\mathbf{v}_{\min}^{(4)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1},\\ \mathbf{r}^{(16)} &= (2,4,2,4,2,4,2,4,2,4,2,4,2,4,2),\\ \mathbf{r}^{(22)} &= (2,4,2,4,2,4,2,4,2,4,2,4,2,4,2,4,2,4,2,4,2).\end{aligned}$$

(iii) $q = 2$, $(d_1, d_2) = (3, 4)$:

$$\begin{aligned}\mathbf{M}_{1:8,1:4} &= \mathbf{I}_2 \otimes \mathrm{mat}_{4\times 2}((\mathbf{v}_{\min}^{(4)})_{1:8}) \left(\mathrm{mat}_{2\times 2}((\mathbf{v}_{\min}^{(3)})_{1:4})\right)^{-1},\\ \mathbf{M}_{9:16,5:8} &= \mathbf{I}_2 \otimes \mathrm{mat}_{4\times 2}((\mathbf{v}_{\min}^{(4)})_{9:16}) \left(\mathrm{mat}_{2\times 2}((\mathbf{v}_{\min}^{(3)})_{5:8})\right)^{-1},\\ \mathbf{r}^{(16)} &= (2,4,8,8,8,8,8,8,8,8,8,8,8,8,4,2),\\ \mathbf{r}^{(22)} &= (2,4,8,8,8,8,8,8,8,8,8,8,8,8,8,8,8,8,4,2).\end{aligned}$$

(iv) $q = 2$, $(d_1, d_2) = (2, 9)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_2 \otimes \mathrm{mat}_{256\times 2}(\mathbf{v}_{\min}^{(9)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)})\right)^{-1},\\ \mathbf{r}^{(16)} &= (2,4,8,16,16,8,4,2,4,8,16,16,8,4,2),\\ \mathbf{r}^{(23)} &= (2,4,8,16,16,8,4,2,4,8,16,16,8,4,2,4,8,16,16,8,4,2).\end{aligned}$$

(v) $q = 2$, $(d_1, d_2) = (8, 9)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_{16} \otimes \mathrm{mat}_{32\times 16}(\mathbf{v}_{\min}^{(9)}) \left(\mathrm{mat}_{16\times 16}(\mathbf{v}_{\min}^{(8)})\right)^{-1},\\ \mathbf{r}^{(16)} &= (2,4,8,16,16,16,16,16,16,16,16,16,16,8,4,2),\\ \mathbf{r}^{(23)} &= (2,4,8,16,16,16,16,16,16,16,16,16,16,16,16,16,16,16,16,8,4,2).\end{aligned}$$

(vi) $q = 3$, $(d_1, d_2) = (2, 3)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_3 \otimes \mathrm{mat}_{9\times 3}(\mathbf{v}_{\min}^{(3)}) \left(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^{-1},\\ \mathbf{r}^{(10)} &= (3,3,3,3,3,3,3,3,3),\\ \mathbf{r}^{(14)} &= (3,3,3,3,3,3,3,3,3,3,3,3,3,3).\end{aligned}$$

(vii) $q = 3$, $(d_1, d_2) = (2, 4)$:

$$\begin{aligned}\mathbf{M} &= \mathbf{I}_3 \otimes \mathrm{mat}_{27\times 3}(\mathbf{v}_{\min}^{(4)}) \left(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\right)^{-1},\\ \mathbf{r}^{(10)} &= (3,9,3,9,3,9,3,9,3),\\ \mathbf{r}^{(14)} &= (3,9,3,9,3,9,3,9,3,9,3,9,3).\end{aligned}$$

(viii) $q = 3$, $(d_1, d_2) = (3, 4)$:

$$\begin{aligned}
\mathbf{M}_{1:27,1:9} &= \mathbf{I}_3 \otimes \mathrm{mat}_{9\times 3}\big((\mathbf{v}_{\min}^{(4)})_{1:27}\big)\,\Big(\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{1:9}\big)\Big)^{-1},\\
\mathbf{M}_{28:54,10:18} &= \mathbf{I}_3 \otimes \mathrm{mat}_{9\times 3}\big((\mathbf{v}_{\min}^{(4)})_{28:54}\big)\,\Big(\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{10:18}\big)\Big)^{-1},\\
\mathbf{M}_{55:81,19:27} &= \mathbf{I}_3 \otimes \mathrm{mat}_{9\times 3}\big((\mathbf{v}_{\min}^{(4)})_{55:81}\big)\,\Big(\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{19:27}\big)\Big)^{-1},\\
\mathbf{r}^{(10)} &= (3, 9, 27, 27, 27, 27, 27, 9, 3),\\
\mathbf{r}^{(14)} &= (3, 9, 27, 27, 27, 27, 27, 27, 27, 27, 27, 9, 3).
\end{aligned}$$

(ix) $q = 3$, $(d_1, d_2) = (2, 6)$:

$$\begin{aligned}
\mathbf{M} &= \mathbf{I}_3 \otimes \mathrm{mat}_{243\times 3}(\mathbf{v}_{\min}^{(6)})\,\Big(\mathrm{mat}_{3\times 3}(\mathbf{v}_{\min}^{(2)})\Big)^{-1},\\
\mathbf{r}^{(10)} &= (3, 9, 27, 9, 3, 9, 27, 9, 3),\\
\mathbf{r}^{(14)} &= (3, 9, 27, 9, 3, 9, 27, 9, 3, 9, 27, 9, 3).
\end{aligned}$$

(x) $q = 3$, $(d_1, d_2) = (5, 6)$:

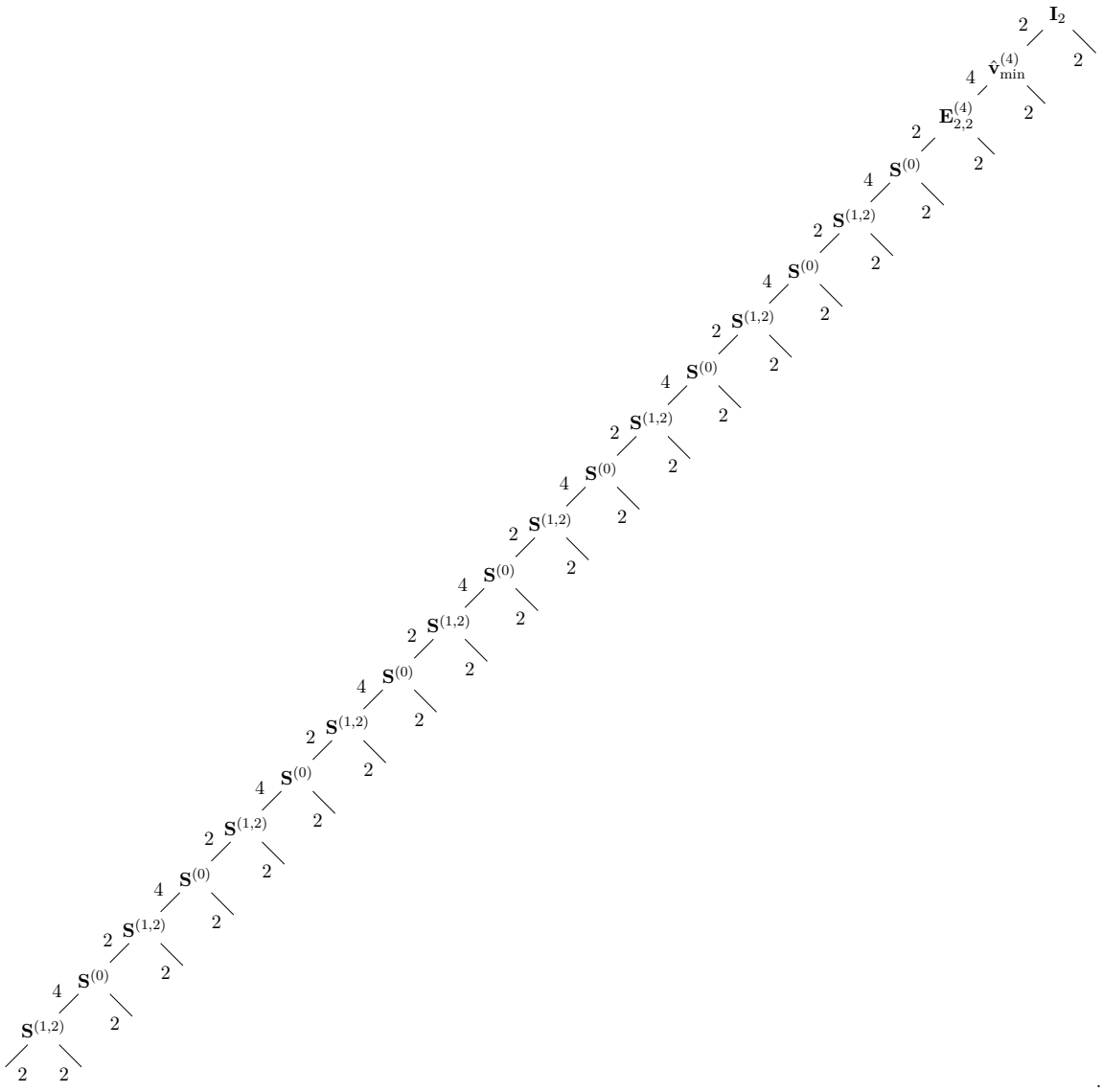
$$\begin{aligned}
\mathbf{M}_{1:243,1:81} &= \mathbf{I}_9 \otimes \mathrm{mat}_{27\times 9}\big((\mathbf{v}_{\min}^{(6)})_{1:243}\big)\,\Big(\mathrm{mat}_{9\times 9}\big((\mathbf{v}_{\min}^{(5)})_{1:81}\big)\Big)^{-1},\\
\mathbf{M}_{244:486,82:162} &= \mathbf{I}_9 \otimes \mathrm{mat}_{27\times 9}\big((\mathbf{v}_{\min}^{(6)})_{244:486}\big)\,\Big(\mathrm{mat}_{9\times 9}\big((\mathbf{v}_{\min}^{(5)})_{82:162}\big)\Big)^{-1},\\
\mathbf{M}_{487:729,163:243} &= \mathbf{I}_9 \otimes \mathrm{mat}_{27\times 9}\big((\mathbf{v}_{\min}^{(6)})_{487:729}\big)\,\Big(\mathrm{mat}_{9\times 9}\big((\mathbf{v}_{\min}^{(5)})_{163:243}\big)\Big)^{-1},\\
\mathbf{r}^{(10)} &= (3, 9, 27, 81, 243, 81, 27, 9, 3),\\
\mathbf{r}^{(14)} &= (3, 9, 27, 81, 243, 243, 243, 243, 243, 81, 27, 9, 3).
\end{aligned}$$

5.3. Balanced HT format

We may convert the representative of $\widetilde{\mathbf{V}}^{(d)}$ based on a linear dimension tree into a representative based on a balanced tree via a sequence of contractions, reshapings, and factorizations, or insertions of identities. Let as before exemplarily $(d_1, d_2) = (2, 4)$ and consider $d = 22$ in case $q = 2$. Writing for short

$$\mathbf{S}^{(0)} := \boldsymbol{\sigma}^{(0)} \otimes \mathbf{E}_{2,1}^{(2)}, \quad \hat{\mathbf{v}}_{\min}^{(4)} := \mathrm{tens}_{4\times 2\times 2}(\mathbf{v}_{\min}^{(4)}),$$

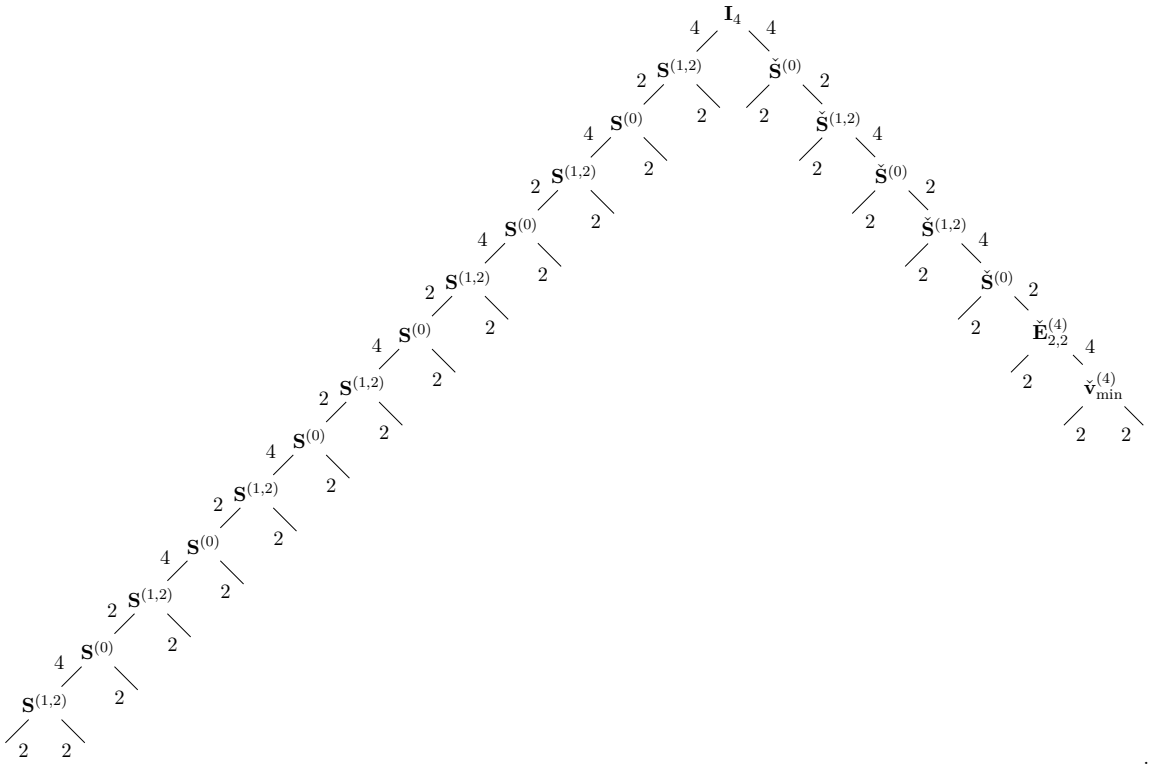
remembering $\mathbf{S}^{(1,2)}$ from (5.47), and omitting any leaf matrix $\mathbf{I}_2$ for better readability, a linear representative of $\widetilde{\mathbf{V}}^{(22)}$ is given by



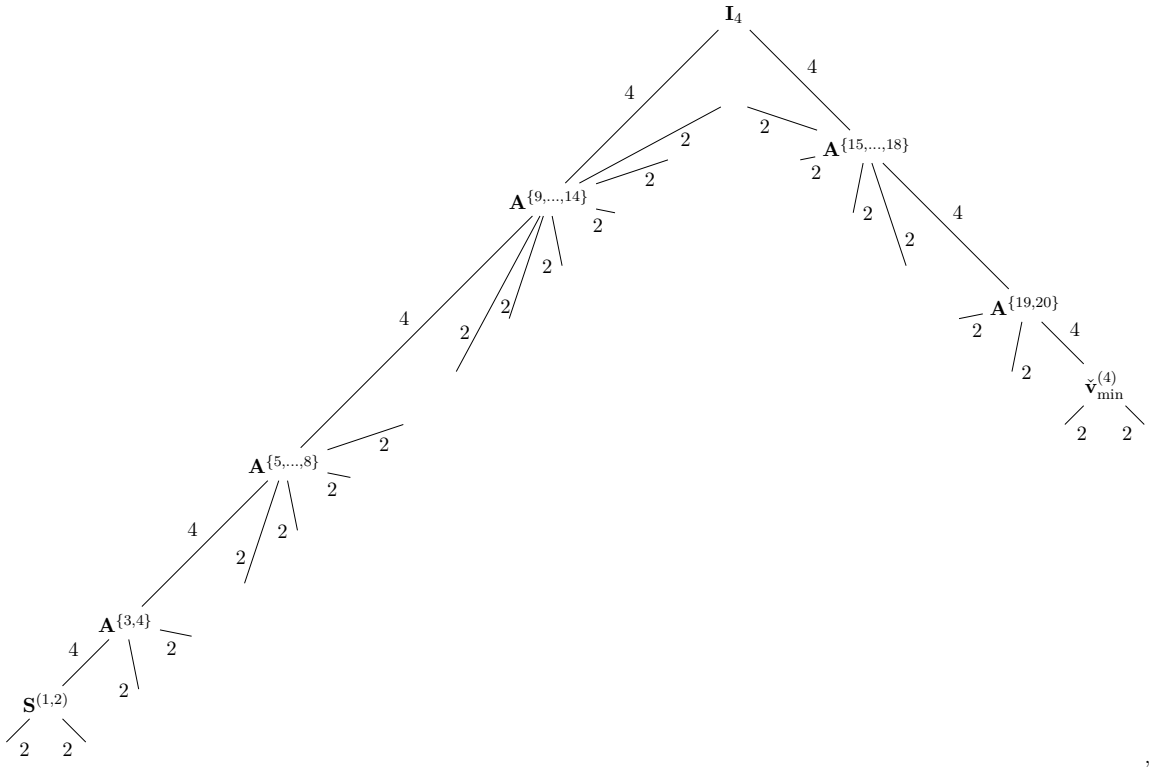
The first step is to fold the tree structure between $\mathbf{B}_{\{1,\dots,14\}} = \mathbf{S}^{(1,2)}$ and $\mathbf{B}_{\{1,\dots,15\}} = \mathbf{S}^{(0)}$ to get at the right side of the new root a number of $8 = 2^3$ modes. Writing

$$\check{\mathbf{v}}^{(4)}_{\min} := \mathrm{perm}_{2,3,1}(\hat{\mathbf{v}}^{(4)}_{\min} \boxdot^1_3 \mathbf{I}_2) = \mathrm{perm}_{2,3,1}(\hat{\mathbf{v}}^{(4)}_{\min})$$

and indicating by $\check{\cdot}$ also the application of $\mathrm{perm}_{2,3,1}$ to other transfer tensors, we obtain



Next we contract groups of 2, 4, and 6 adjacent transfer tensors to collect the respective number of modes in subtrees. As an intermediate result, this yields a non-binary tree structure



where

$$\begin{aligned}
\mathbf{A}^{\{3,4\}} &:= \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)},\\
\mathbf{A}^{\{5,\ldots,8\}} &:= \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)} \boxdot_3^1 \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)},\\
\mathbf{A}^{\{9,\ldots,14\}} &:= \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)} \boxdot_3^1 \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)} \boxdot_3^1 \mathbf{S}^{(0)} \boxdot_3^1 \mathbf{S}^{(1,2)},\\
\mathbf{A}^{\{15,\ldots,18\}} &:= \mathrm{perm}_{1,3,4,5,6,2}(\check{\mathbf{S}}^{(0)} \boxdot_2^3 \check{\mathbf{S}}^{(1,2)} \boxdot_2^3 \check{\mathbf{S}}^{(0)} \boxdot_2^3 \check{\mathbf{S}}^{(1,2)}),\\
\mathbf{A}^{\{19,20\}} &:= \mathrm{perm}_{1,3,4,2}(\check{\mathbf{S}}^{(0)} \boxdot_2^3 \check{\mathbf{E}}^{(4)}_{2,2}).
\end{aligned}$$

Now we reshape and factorize these higher-than-three-order tensors $\mathbf{A}^{\alpha}$ in order to get back the binary and additionally the desired balanced structure. In general it depends on those particular mode sizes of the contracted transfer tensors which are connected, hence on the HT ranks

$$r_{\tau_i}, \quad (\tau_i)_{1\le i\le 6} := \big(\{1,2\},\{1,\ldots,4\},\{1,\ldots,8\},\{1,\ldots,14\},\{1,\ldots,18\},\{1,\ldots,20\}\big),$$

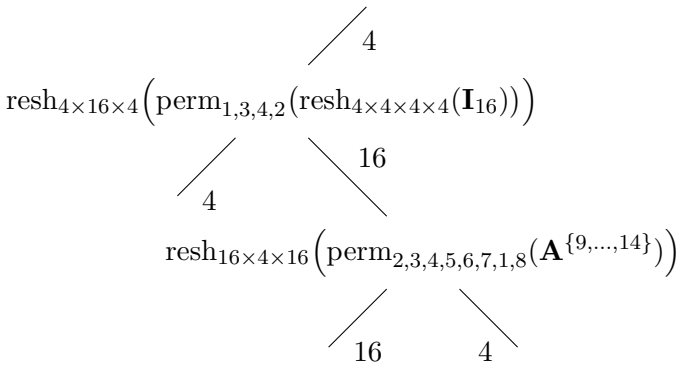
in the linear representative, whether it is advantageous to insert a suitably reshaped identity at the right resp. left child node of

$$\mathrm{resh}_{r_{\tau_i}\times 2^{|\tau_{i+1}\setminus\tau_i|}\times r_{\tau_{i+1}}}(\mathbf{A}^{\tau_{i+1}\setminus\tau_i}),\ i=1,2,3 \quad \text{resp.} \quad \mathrm{resh}_{2^{|\tau_{i+1}\setminus\tau_i|}\times r_{\tau_{i+1}}\times r_{\tau_i}}(\mathbf{A}^{\tau_{i+1}\setminus\tau_i}),\ i=4,5$$

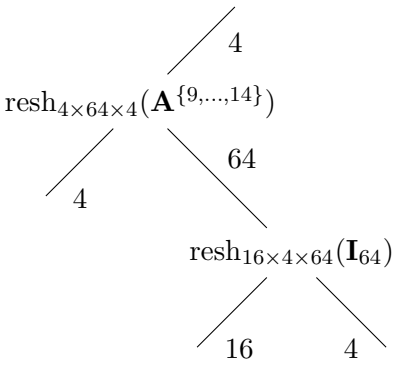
or to propagate $\mathbf{A}^{\tau_{i+1}\setminus\tau_i}$ itself into its right resp. left child node and to insert an identity at the vacant parent node. Since in the present case $r_{\tau_i}=4$ for all $1\le i\le 6$ and so

$$r_{\tau_i} r_{\tau_{i+1}} < 2^{|\tau_{i+1}\setminus\tau_i|}$$

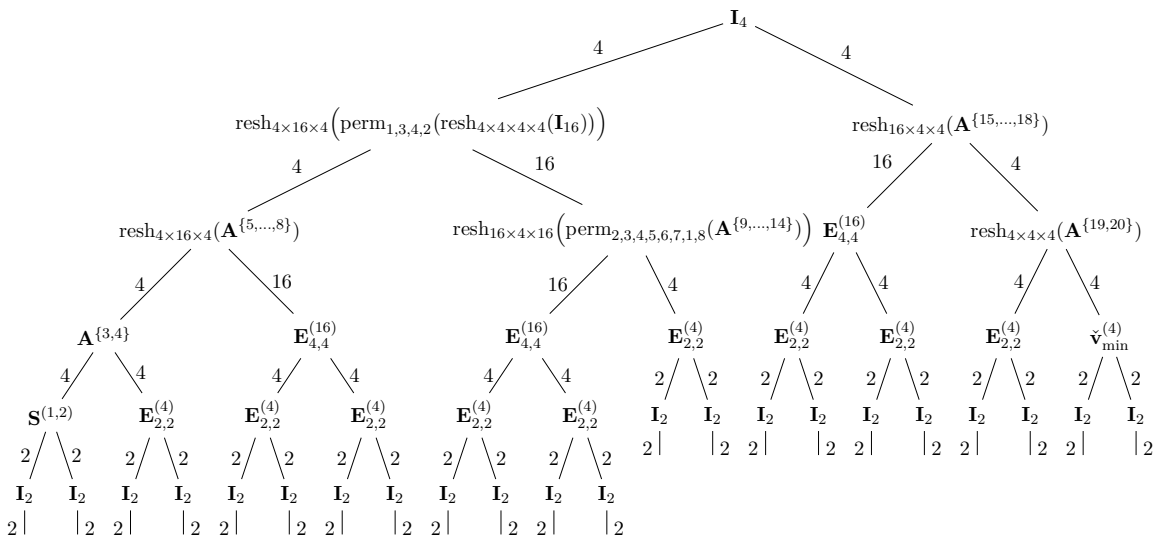
only for $i=3$, hence for $\alpha=\tau_4\setminus\tau_3=\{9,\ldots,14\}$, we notice that $\mathbf{A}^{\{9,\ldots,14\}}$ should rather be represented by



than by



and that the other $\mathbf{A}^{\beta}$ for $\beta = \{5, \dots, 8\}, \{15, \dots, 18\}$ may remain and for $\beta = \{3, 4\}, \{19, 20\}$ should remain at their places in the tree structure. It is also possible to perform an SVD of an appropriate matricization of $\mathbf{A}^{\alpha}$ to reveal inherent smaller ranks in the balanced representation or to truncate the ranks due to some criterion. Summing up, and introducing $\mathbf{E}^{(16)}_{4,4} := \mathrm{resh}_{4\times4\times16}(\mathbf{I}_{16})$, we choose



as an HT representative of $\widetilde{\mathbf{V}}^{(22)}$ with balanced dimension tree.

5.4. TT format

We expressed the MALS, Algorithm 4.3, in the tensor train format consistently with the predominant part of the literature. Hence, also the initial guess has to be in TT format. Since $\mathbf{U}_t = \mathbf{I}_q$, $t \in \mathcal{L}(\mathcal{T})$, for the initial guess constructed in Section 5.2, we may set the core tensors $\mathbf{G}^{(j)}$, $2 \le j \le d$, of a TT representative as the transfer tensors $\mathbf{B}_{\{1,\dots,j\}}$ of an HT representative based on the linear dimension tree supplemented by $\mathbf{G}^{(1)} = \mathbf{U}_{\{1\}}$, cf. (3.11).

In the exemplarily discussed case $(d_1, d_2) = (2, 4)$ with $q = 2$ we obtain, due to (5.50),

$$\begin{aligned}
\mathbf{G}^{(2)} = \mathbf{G}^{(4)} = \ldots = \mathbf{G}^{(d-4)} &= \mathbf{S}^{(1,2)},\\
\mathbf{G}^{(3)} = \mathbf{G}^{(5)} = \ldots = \mathbf{G}^{(d-3)} &= \boldsymbol{\sigma}^{(0)} \otimes \mathbf{E}^{(2)}_{2,1},\\
\mathbf{G}^{(d-2)} = \mathbf{E}^{(4)}_{2,2}, \quad \mathbf{G}^{(d-1)} &= \mathrm{tens}_{4\times2\times2}(\mathbf{v}^{(4)}_{\min}), \quad \mathbf{G}^{(d)} = \mathbf{I}_2,
\end{aligned}$$

and the TT ranks are given by

$$r_1 = r_3 \ldots = r_{d-1} = 2, \quad r_2 = r_4 = \ldots = r_{d-2} = 4.$$

Concerning the TT ranks of the representatives of the prolongated initial guess for other (d_1, d_2) or q, we refer to Remark 5.5.

5.5. XYZ model, $A = B$

The possibility to construct an initial guess via prolongation, as it is described in the previous sections of this chapter, relies on the invertibility of the quadratic matricization of $\mathbf{v}^{(d_1)}_{\min}$ or,

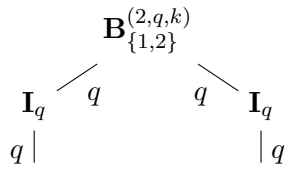
especially if d_1 is odd, all suitably sized blocks thereof. In the case of an XYZ model with coupling parameters $A = B$, for a simple minimal eigenvalue $\lambda_{\min}^{(d_1)}$ it is $\mathbf{v}_{\min}^{(d_1)} \in \mathcal{E}_{d_1,q}^{(k)}$ for an appropriate $0 \leq k \leq (q-1)d_1$ by Corollary 2.10(ii), and therefore these matricizations are in general singular. One might pass to the pseudoinverse, but as will be discussed in Section 6.4, the behavior of such an initial guess in the LOCG method is not satisfactory.

To propose an alternative in that situation, we consider as initial guess a vector $\widetilde{\mathbf{v}}_{\text{const}}^{(k)} \in \mathbb{R}^{q^d}$ which has the sparsity pattern of $\mathcal{E}_{d,q}^{(k)}$ for some $k \in \{0, \ldots, (q-1)d\}$ and all of its nonzero entries are constant, calling it *constant initial guess*. Setting the value of these constant nonzero entries equal to the reciprocal of the square root of the number of nonzero entries, $\widetilde{\mathbf{v}}_{\text{const}}^{(k)}$ is normalized. If we choose $k = k^*$ such that $\mathbf{v}_{\min}^{(d)} \in \mathcal{E}_{d,q}^{(k^*)}$, then we might expect that the initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$ may lead to faster convergence than a random initial guess would, since it has the same sparsity pattern as the vector $\mathbf{v}_{\min}^{(d)}$ to be approximated. In general we do not know the correct value of k^*. But, in case we can start several instances of the iterative numerical method in parallel, we could start for each $k \in \{0, \ldots, (q-1)d\}$ an instance of the method with $\widetilde{\mathbf{v}}_{\text{const}}^{(k)} \in \mathcal{E}_{d,q}^{(k)}$ and pick in the course of the iteration that instance which converges fastest.

The constant initial guess may be constructed in hierarchical Tucker format with ranks scaling linearly in d. Let $\mathbf{e}^{(d,q,k)}$ be that element of $\mathcal{E}_{d,q}^{(k)}$ with all potential nonzero entries equal to 1, thus $\mathbf{e}^{(d,q,k)}$ is the non-normalized constant initial guess and normalization may be done by dividing the entries in the resulting transfer tensor at $t = \text{root}$ by $\|\mathbf{e}^{(d,q,k)}\|$. We find it convenient to employ in the description of the construction the bra-ket notation, see Appendix A, and start with the case of the linear dimension tree $\mathcal{T}$. It is

$$\mathbf{e}^{(d,q,k)} = \sum_{j_1+\ldots+j_d=k} |j_1 \ \ldots \ j_d\rangle\,, \quad j_1, \ldots, j_d \in \{0, \ldots, q-1\}.$$

Hence, if $d = 2$, a representative of $\text{tens}_{q\times q}(\mathbf{e}^{(2,q,k)}) = \text{tens}_{q\times q}\left(\sum_{j_1+j_2=k} |j_1\ j_2\rangle\right)$ is given by



with

$$\left(\mathbf{B}_{\{1,2\}}^{(2,q,k)}\right)_{i_1,i_2} = \begin{cases} 1, & (i_1-1)+(i_2-1) = k \\ 0, & \text{otherwise} \end{cases}.$$

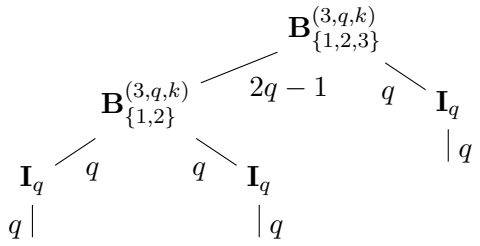
We notice as an example

$$\mathbf{B}_{\{1,2\}}^{(2,3,0)} = \left(\begin{smallmatrix} 1&0&0\\0&0&0\\0&0&0 \end{smallmatrix}\right), \quad \mathbf{B}_{\{1,2\}}^{(2,3,1)} = \left(\begin{smallmatrix} 0&1&0\\1&0&0\\0&0&0 \end{smallmatrix}\right), \quad \mathbf{B}_{\{1,2\}}^{(2,3,2)} = \left(\begin{smallmatrix} 0&0&1\\0&1&0\\1&0&0 \end{smallmatrix}\right),$$
$$\mathbf{B}_{\{1,2\}}^{(2,3,3)} = \left(\begin{smallmatrix} 0&0&0\\0&0&1\\0&1&0 \end{smallmatrix}\right), \quad \mathbf{B}_{\{1,2\}}^{(2,3,4)} = \left(\begin{smallmatrix} 0&0&0\\0&0&0\\0&0&1 \end{smallmatrix}\right).$$

If $d = 3$, a representative of

$$\text{tens}_{q\times q\times q}(\mathbf{e}^{(3,q,k)}) = \text{tens}_{q\times q\times q}\left(\sum_{j_1+j_2+j_3=k} |j_1\ j_2\ j_3\rangle\right) = \text{tens}_{q\times q\times q}\left(\sum_{j_1+l=k} |j_1\rangle \sum_{j_2+j_3=l} |j_2\ j_3\rangle\right)$$

is given by



with

$$\begin{aligned}\left(\mathbf{B}^{(3,q,k)}_{\{1,2\}}\right)_{:,:,l+1} &= \mathbf{B}^{(2,q,l)}_{\{1,2\}}, \quad 0 \le l \le 2q-2,\\ \left(\mathbf{B}^{(3,q,k)}_{\{1,2,3\}}\right)_{i_1,i_2} &= \begin{cases}1, & (i_1-1)+(i_2-1)=k\\ 0, & \text{otherwise}\end{cases}.\end{aligned}$$

As an illustration we notice

$$\mathbf{B}^{(3,3,0)}_{\{1,2,3\}} = \begin{pmatrix}1&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\end{pmatrix}, \quad \mathbf{B}^{(3,3,1)}_{\{1,2,3\}} = \begin{pmatrix}0&1&0\\1&0&0\\0&0&0\\0&0&0\\0&0&0\end{pmatrix}, \quad \mathbf{B}^{(3,3,2)}_{\{1,2,3\}} = \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\\0&0&0\\0&0&0\end{pmatrix}, \quad \mathbf{B}^{(3,3,3)}_{\{1,2,3\}} = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\\1&0&0\\0&0&0\end{pmatrix},$$
$$\mathbf{B}^{(3,3,4)}_{\{1,2,3\}} = \begin{pmatrix}0&0&0\\0&0&0\\0&0&1\\0&1&0\\1&0&0\end{pmatrix}, \quad \mathbf{B}^{(3,3,5)}_{\{1,2,3\}} = \begin{pmatrix}0&0&0\\0&0&0\\0&0&0\\0&0&1\\0&1&0\end{pmatrix}, \quad \mathbf{B}^{(3,3,6)}_{\{1,2,3\}} = \begin{pmatrix}0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1\end{pmatrix}.$$

For general d, let $\mathbf{m} \in \mathbb{R}^d$, $m_i := i(q-1)+1$, and observe that $m_\delta = m_{\delta-1} + q - 1$ for $2 \le \delta \le d$ is a recursion for the number of different anti-diagonals of a $m_{\delta-1} \times q$ matrix with $m_1 = q$. Then a representative of $\mathrm{tens}_{q\times d}(\mathbf{e}^{(d,q,k)})$ is

$$\left[\mathcal{T}, (\mathbf{B}^{(d,q,k)}_t)_{t\in\mathcal{N}(\mathcal{T})}, \mathbf{U}_{\{1\}} = \ldots = \mathbf{U}_{\{d\}} = \mathbf{I}_q\right] \tag{5.51}$$

with

$$\begin{aligned}\mathbf{B}^{(d,q,k)}_{\{1,\ldots,\delta\}} &\in \mathbb{R}^{m_{\delta-1}\times q\times m_\delta}, \quad 2 \le \delta \le d-1,\\ \left(\mathbf{B}^{(d,q,k)}_{\{1,\ldots,\delta\}}\right)_{i_1,i_2,i_3} &= \begin{cases}1, & (i_1-1)+(i_2-1)=i_3-1\\ 0, & \text{otherwise}\end{cases},\end{aligned} \tag{5.52}$$

hence

$$\mathbf{B}^{(d,q,k)}_{\{1,\ldots,\delta\}} = \begin{pmatrix}1&&\\&&\\&&\end{pmatrix}\begin{pmatrix}1&1&\\&&\\&&\end{pmatrix}\cdots\begin{pmatrix}&&1\\1&1&\\&&\end{pmatrix}\cdots\begin{pmatrix}&&\\&&1\\1&1&\end{pmatrix}\cdots\begin{pmatrix}&&\\&&\\&1&1\end{pmatrix}\begin{pmatrix}&&\\&&\\&&1\end{pmatrix},$$

where void entries are 0, and furthermore

$$\mathbf{B}^{(d,q,k)}_{\{1,\ldots,d\}} \in \mathbb{R}^{m_{d-1}\times q}, \quad \left(\mathbf{B}^{(d,q,k)}_{\{1,\ldots,d\}}\right)_{i_1,i_2} = \begin{cases}1, & (i_1-1)+(i_2-1)=k\\ 0, & \text{otherwise}\end{cases}. \tag{5.53}$$

In the case that $\mathrm{tens}_{q\times d}(\mathbf{e}^{(d,q,k)})$ is represented based on the balanced dimension tree $\mathcal{T}$, also $\mathbf{U}_t = \mathbf{I}_q$, $t \in \mathcal{L}(\mathcal{T})$, and for the transfer tensors $\mathbf{B}_t \in \mathbb{R}^{r_{t_l}\times r_{t_r}\times r_t}$, $t \in \mathcal{N}(\mathcal{T}) \setminus \{1,\dots,d\}$, it is in turn $r_t = r_{t_l} + r_{t_r} - 1$ as well as

$$(\mathbf{B}_t)_{i_1,i_2,i_3} = \begin{cases} 1, & (i_1-1)+(i_2-1) = i_3 - 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$\left(\mathbf{B}_{\{1,\dots,d\}}\right)_{i_1,i_2} = \begin{cases} 1, & (i_1-1)+(i_2-1) = k \\ 0, & \text{otherwise} \end{cases}.$$

A TT representative $\left[\mathbf{G}^{(1)}, \mathbf{G}^{(2)}, \dots, \mathbf{G}^{(d)}\right]$ of $\mathrm{tens}_{q\times d}(\mathbf{e}^{(d,q,k)})$ is directly obtained from the HT representative (5.51)-(5.53) based on the linear dimension tree, cf. (3.11), by

$$\mathbf{G}^{(1)} = \mathbf{U}_{\{1\}} = \mathbf{I}_q, \quad \mathbf{G}^{(j)} = \mathbf{B}^{(d,q,k)}_{\{1,\dots,j\}},\ 2 \le j \le d.$$

The ranks may be significantly reduced by removing those slices of the transfer tensors and columns of the leaf matrices which are not addressed by a nonzero coefficient in the respective parent transfer tensor. This way we obtain, focusing for simplicity on the linear dimension tree,

$$\begin{aligned} r_j^{\mathrm{TT}} &= r^{\mathrm{HT}}_{\{1,\dots,j\}} \\ &= \min\left\{k+1, (q-1)d-k+1, (q-1)j+1, (q-1)(d-j)+1\right\}, \quad 1 \le j \le d-1. \end{aligned}$$

Hence, in each situation considered in the present section, the maximal rank scales linearly in d.

6. Numerical tests

Our main objective in this chapter is to investigate the usefulness of the initial guess constructed by prolongation in comparison with a random initial guess when employed in an iterative numerical method. All tests are performed in a tensor format whose type will be clear from the settings of the respective test. We focus on the same types of Hamilton operators we considered so far, namely the 2-XYZ model, the 3-XYZ model, and the 3-Potts model. In each of these three classes, there are further parameters to be chosen:

1. The method itself and the underlying tensor format. In Sections 6.2 and 6.4 we mainly test the HT variant of LOCG from Algorithm 4.2 and only to a minor extent gradient descent which is obtained by omitting all $\mathbf{P}_k$ in Algorithm 4.2, therefore reducing the size of the generalized eigenvalue problem to be solved in each iteration step from 3×3 to 2×2. The MALS formulated in TT format is considered in Sections 6.3 and 6.5.

2. The problem sizes d for which we actually compute $\mathbf{v}_{\min}^{(d)}$ and (d_1, d_2) wherefrom information about the exact eigenvectors associated with $\lambda_{\min}$ is utilized to set up the prolongation. For $q = 2$ we consider $d \in \{16, 22\}$ with $(d_1, d_2) \in \{(2,3), (2,4), (3,4)\}$ and additionally $d \in \{16, 23\}$ with $(d_1, d_2) \in \{(2,9), (8,9)\}$. For $q = 3$ we consider $d \in \{10, 14\}$ with $(d_1, d_2) \in \{(2,3), (2,4), (3,4), (2,6), (5,6)\}$. Notice that $2^{16} = 65{,}536$ and $3^{10} = 59{,}049$ as well as $2^{22} = 4{,}194{,}304$ and $3^{14} = 4{,}782{,}969$ have a similar size.

3. The type of the dimension tree, linear or balanced, cf. Definition 3.6 and Example 3.7. This distinction applies only to LOCG in HT format.

4. The coupling parameters A, B, Δ, h for the XYZ model respectively A, h for the Potts model. In Sections 6.2 and 6.3 we deal with the case $A \neq B$ where prolongation based on problem sizes (d_1, d_2) is feasible. The performance of the constant initial guess proposed for $A = B$ is discussed in Sections 6.4 and 6.5.

5. The parameters for the truncation of the ranks during the iteration. Although the precise meaning of this term is different for LOCG in HT and MALS in TT, in both cases we have the possibility to impose a relative error bound and an upper bound of the rank when truncating the respective set of singular values. In Section 6.1 we discuss suitable choices in case of LOCG in HT and we apply our findings also in the subsequent tests with MALS/TT, due to the similarity of HT with linear dimension tree to TT.

We do not examine all possible combinations of these parameters but rather choose a set of various exemplary showcases. "Exact" eigenvectors $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$, utilized to set up the prolongation, are computed in full vector format by the function `eig` in MATLAB for (d_1, d_2) mentioned in Point 2 in the above list. As needed for reference, the "exact" minimal eigenvalue $\lambda_{\min}^{(d)}$ for $d \in \{16, 22, 23\}$ if $q = 2$ respectively $d \in \{10, 14\}$ if $q = 3$ is obtained from the MATLAB function `eigs` with `opts.tol=1e-15` storing $\mathbf{H}_d$ as a sparse matrix. Likewise, this holds in a few scenarios for some larger eigenvalues of $\mathbf{H}_d$ or associated eigenvectors. Representatives of random initial guesses are set up with the help of `randn`.

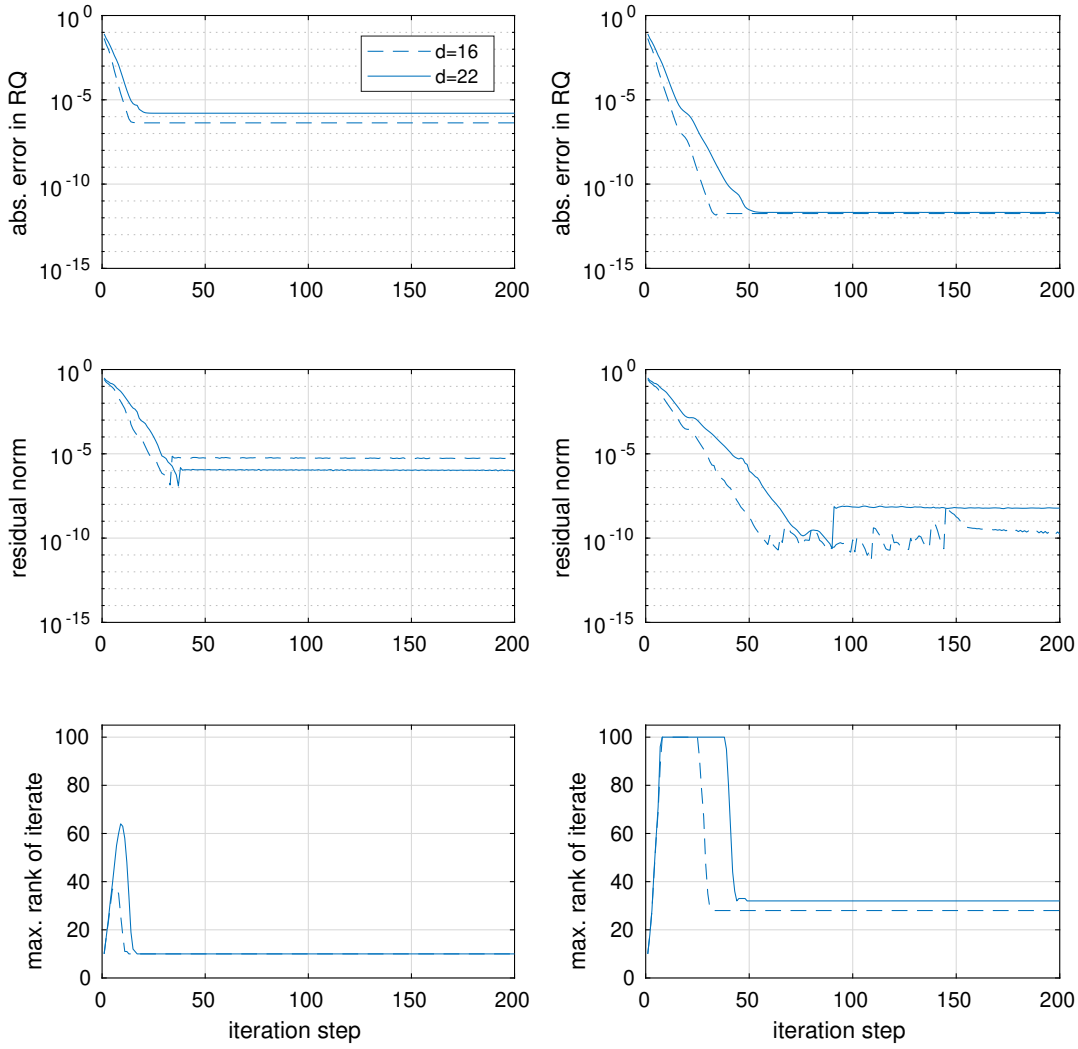


Figure 6.1.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, lin. tree, $r_{\max} = 100$, *left:* $\varepsilon_{\text{rel}} = 10^{-3}$, *right:* $\varepsilon_{\text{rel}} = 10^{-6}$

The computations with tensors in hierarchical Tucker format are based on the MATLAB toolbox `htucker` [KT14]. When computing with tensor trains, we employ the `TT-Toolbox` [ODK+12], also written for MATLAB. In addition, we use `ncon.m` [PESV15] that implements the contraction of tensors.

All numerical experiments are performed in MATLAB R2017a (9.2.0.556344) 64-bit. The computer is equipped with an Intel Xeon E5-1620 CPU (4 cores, 3.6 GHz), 64 GB RAM, and Debian GNU/Linux 9.

6.1. Truncation parameters

Concerning Point 5 in the above list, as a first step of the numerical tests, we perform a short parameter study in order to investigate how large the hierarchical ranks respectively the relative error in each truncation step should be chosen. Due to [Tob12, Remk. 3.6], when

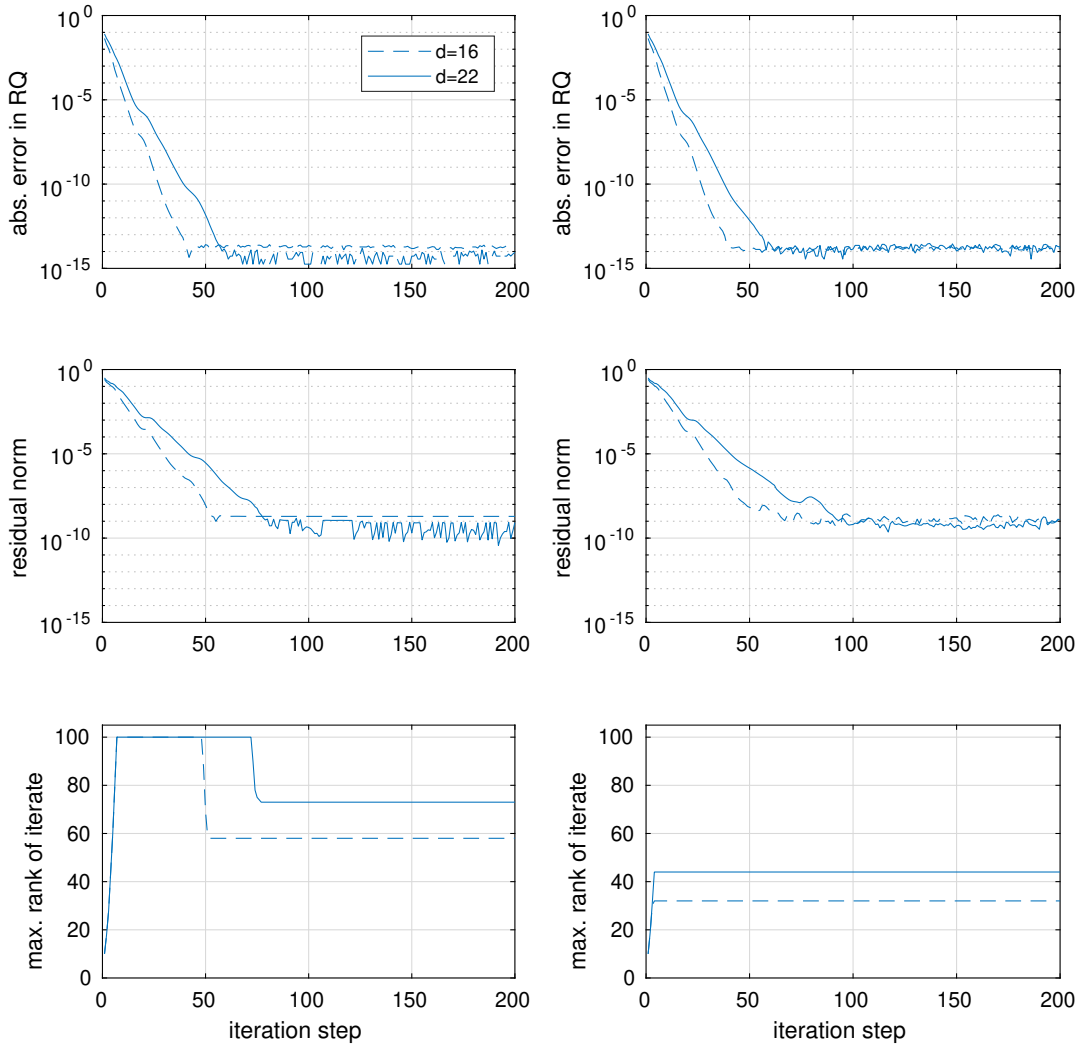


Figure 6.2.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, lin. tree, *left:* $r_{\max} = 100$, $\varepsilon_{\mathrm{rel}} = 10^{-9}$, *right:* $r_{\max} = qd$, $\varepsilon_{\mathrm{rel}} = 10^{-12}$

requiring

$$\|\mathbf{T} - \widehat{\mathbf{T}}\| \leq \varepsilon_{\mathrm{rel}} \|\mathbf{T}\| \tag{6.1}$$

for a truncated tensor $\widehat{\mathbf{T}} \in \mathcal{H}\text{-Tucker}\big((r_t)_{t \in \mathcal{T}}\big)$, we may choose r_t such that

$$\sqrt{\sum_{i=r_t+1}^{\overline{n_t}} \sigma_i\big(\mathrm{mat}_t(\mathbf{T})\big)^2} \leq \frac{\varepsilon_{\mathrm{rel}} \|\mathbf{T}\|}{\sqrt{2d-3}} \tag{6.2}$$

for all $t \in \mathcal{T} \setminus \{\mathrm{root}, c\}$, c one child of root. Confer also (3.8). Those functions in the toolbox `htucker` which implement truncation procedures, allow to specify an upper bound $r_{\max}$ of the HT rank that is certainly not exceeded for any $t \in \mathcal{T}$, and additionally the relative error bound $\varepsilon_{\mathrm{rel}}$ which, due to the precedence of $r_{\max}$, is probably not fulfilled in each iteration step.

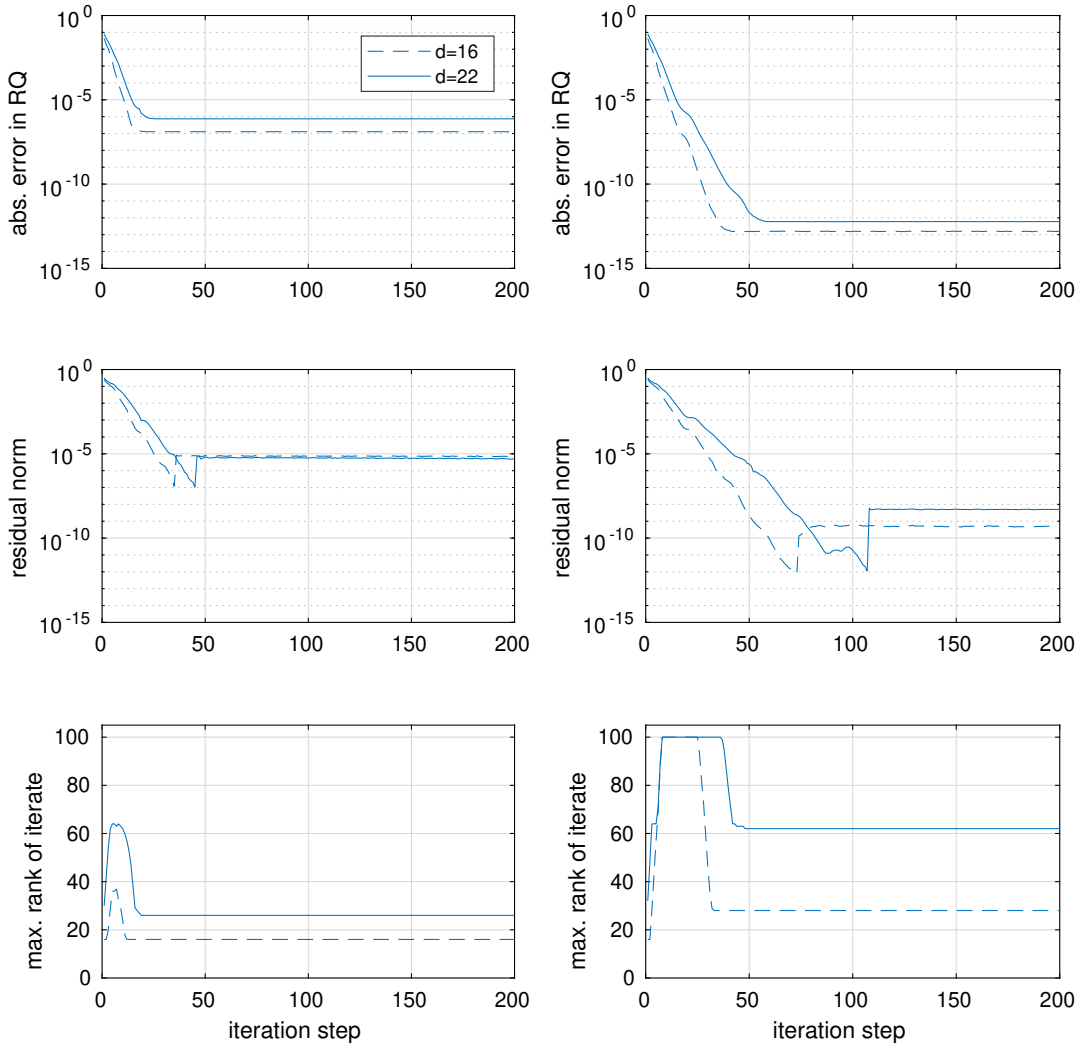


Figure 6.3.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, bal. tree, $r_{\max} = 100$, *left:* $\varepsilon_{\text{rel}} = 10^{-3}$, *right:* $\varepsilon_{\text{rel}} = 10^{-6}$

We start with LOCG for an XYZ model with $q = 2$, set $A = 1.9$, $B = 0.4$, $\Delta = -1.1$, $h = 0.2$ as well as $(d_1, d_2) = (2, 4)$, $d \in \{16, 22\}$ and consider the linear dimension tree. Furthermore it is $r_{\max} = 100$ and we compare the behavior of the method, executed with the prolongated initial guess, for the values $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}, 10^{-9}\}$ in the first three columns of Figure 6.1 together with Figure 6.2 which have to be regarded as one composite graphic just spread on two pages. In the first row we plot the absolute error

$$\left| \langle \mathbf{V}_k, \Phi_{\mathbf{H}}(\mathbf{V}_k) \rangle - \lambda_{\min}^{(d)} \right| \tag{6.3}$$

in the Rayleigh quotient, also called "RQ error" in the sequel, further noticing $\|\mathbf{V}_k\| = 1$ at the end of an iteration step. We consider here the absolute value since in practice $\lambda_{\min}^{(d)}$ is, as a result of `eigs` in MATLAB, itself only an approximation with an absolute error near the machine epsilon, and the difference in (6.3) might become negative not being visible in

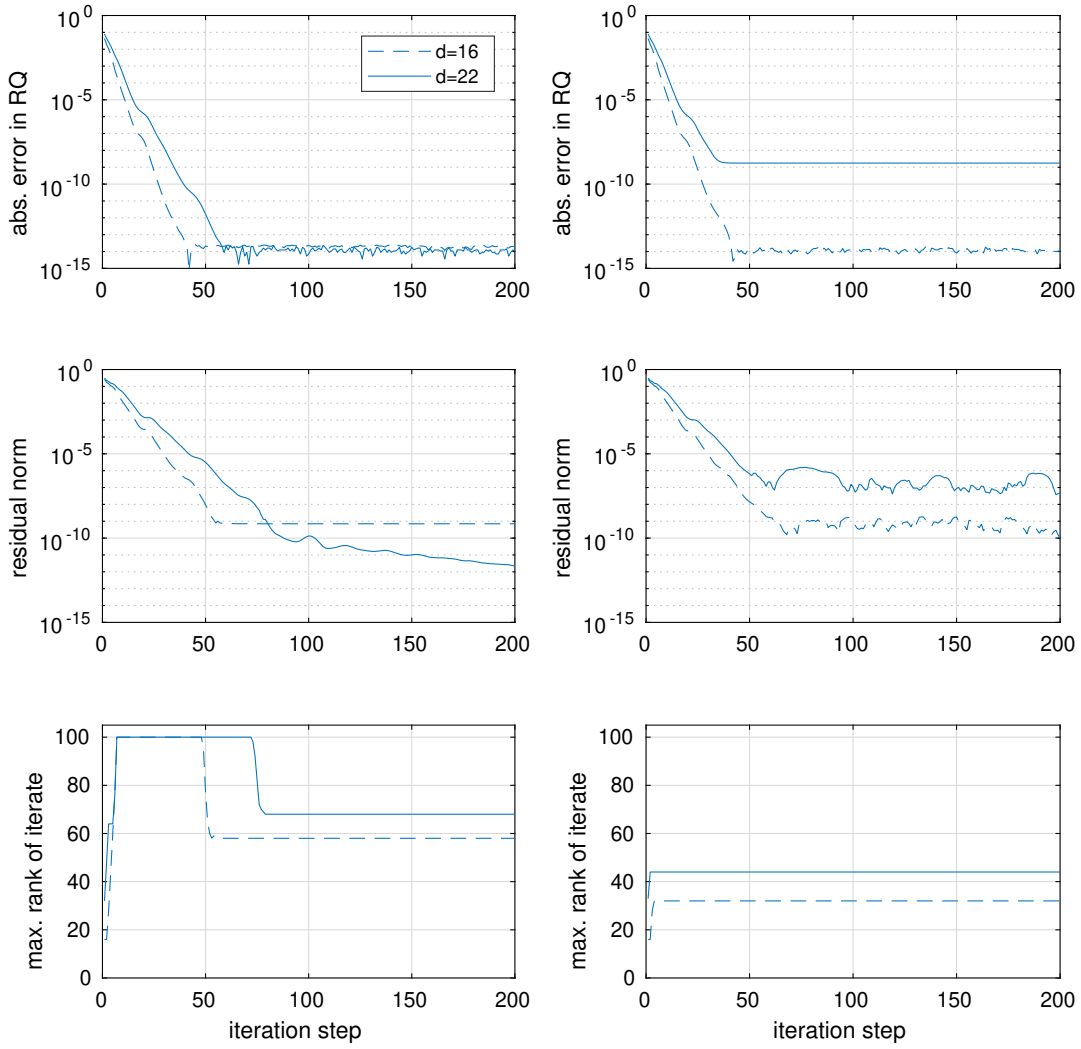


Figure 6.4.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, bal. tree, *left:* $r_{\max} = 100$, $\varepsilon_{\mathrm{rel}} = 10^{-9}$, *right:* $r_{\max} = qd$, $\varepsilon_{\mathrm{rel}} = 10^{-12}$

a semilogarithmic plot. The second row depicts the residual norm

$$\left\|\Phi_{\mathbf{H}}(\mathbf{V}_k) - \langle \mathbf{V}_k, \Phi_{\mathbf{H}}(\mathbf{V}_k)\rangle \mathbf{V}_k\right\|, \tag{6.4}$$

and the third row the maximal hierarchical rank of $\mathbf{V}_k$ as a result of the truncation in Line 13 in Algorithm 4.2. The rightmost column contains the same three quantities in case $r_{\max} = qd$, so $r_{\max} \in \{32, 44\}$, and $\varepsilon_{\mathrm{rel}} = 10^{-12}$ to cut off singular values near the machine epsilon. To avoid confusion, we emphasize that one has to distinguish between the parameter $r_{\max}$ which is an upper bound on each single HT rank r_t for all $t \in \mathcal{T}$ during the truncation procedure, and the actual maximal value $\max\{r_t : t \in \mathcal{T}\}$ of the HT ranks of the iterate which is indicated by the y-axis of the plot in the third row.

We observe for $r_{\max} = 100$, $\varepsilon_{\mathrm{rel}} \in \{10^{-3}, 10^{-6}, 10^{-9}\}$ that the ranks increase at the beginning of the iteration and around the point where the RQ error stops to improve, the ranks again decrease and stabilize at a specific value $r_{\mathrm{final}}^{(d,\varepsilon_{\mathrm{rel}})}$. This final maximal rank becomes the

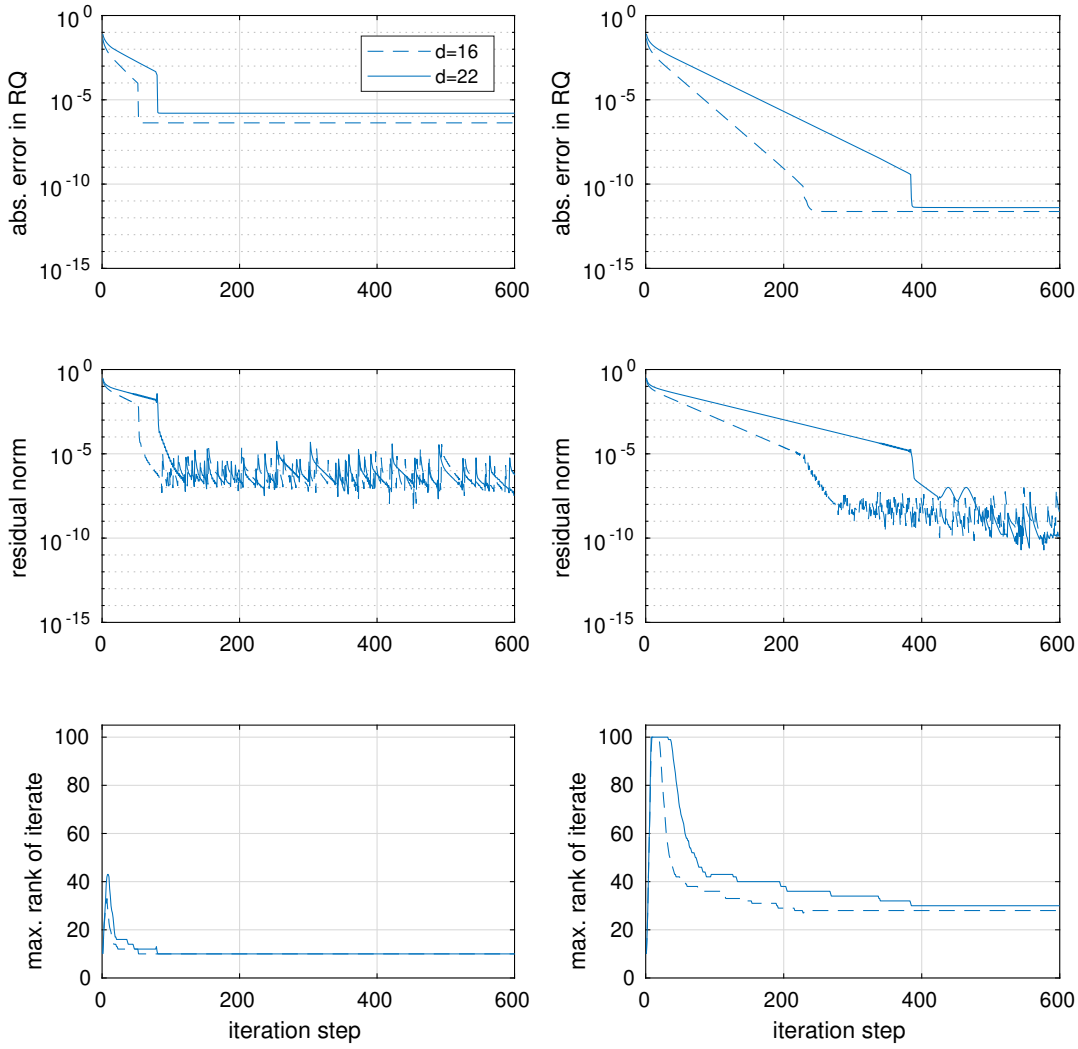


Figure 6.5.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, grad. desc., lin. tree, $r_{\max} = 100$, *left:* $\varepsilon_{\text{rel}} = 10^{-3}$, *right:* $\varepsilon_{\text{rel}} = 10^{-6}$

larger and occurs at a larger number of steps, the smaller ε_{rel} gets. That goes also along with a smaller final RQ error and a smaller residual norm. For $d = 16$, the convergence is slightly faster and the RQ error as well as the ranks are slightly smaller. The behavior in the case $r_{\max} = qd$ and $\varepsilon_{\text{rel}} = 10^{-12}$ is comparable with the behavior in the case $r_{\max} = 100$ and $\varepsilon_{\text{rel}} = 10^{-9}$.

We repeat the same test for a balanced dimension tree and depict the results in Figures 6.3 and 6.4. For each $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}, 10^{-9}\}$ with $r_{\max} = 100$, it is again $r_{\text{final}}^{(d,\varepsilon_{\text{rel}})} < r_{\max}$. Compared to the case of a linear tree, $r_{\text{final}}^{(d,\varepsilon_{\text{rel}})}$ is in some cases larger for fixed d, especially for $\varepsilon_{\text{rel}} = 10^{-6}$ and $d = 22$. The RQ error and the residual norm have the same magnitudes as for the linear tree. Regarding the case $r_{\max} = 44$, $\varepsilon_{\text{rel}} = 10^{-12}$, so for $d = 22$, we observe $r_{\text{final}}^{(22,10^{-3})} < 44 < r_{\text{final}}^{(22,10^{-6})}$ and hence the final RQ error as well as the residual norm stabilize between these respective values for $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ at a level around 10^{-8} which is larger than for the linear dimension tree. In the case $d = 16$ we observe this effect with

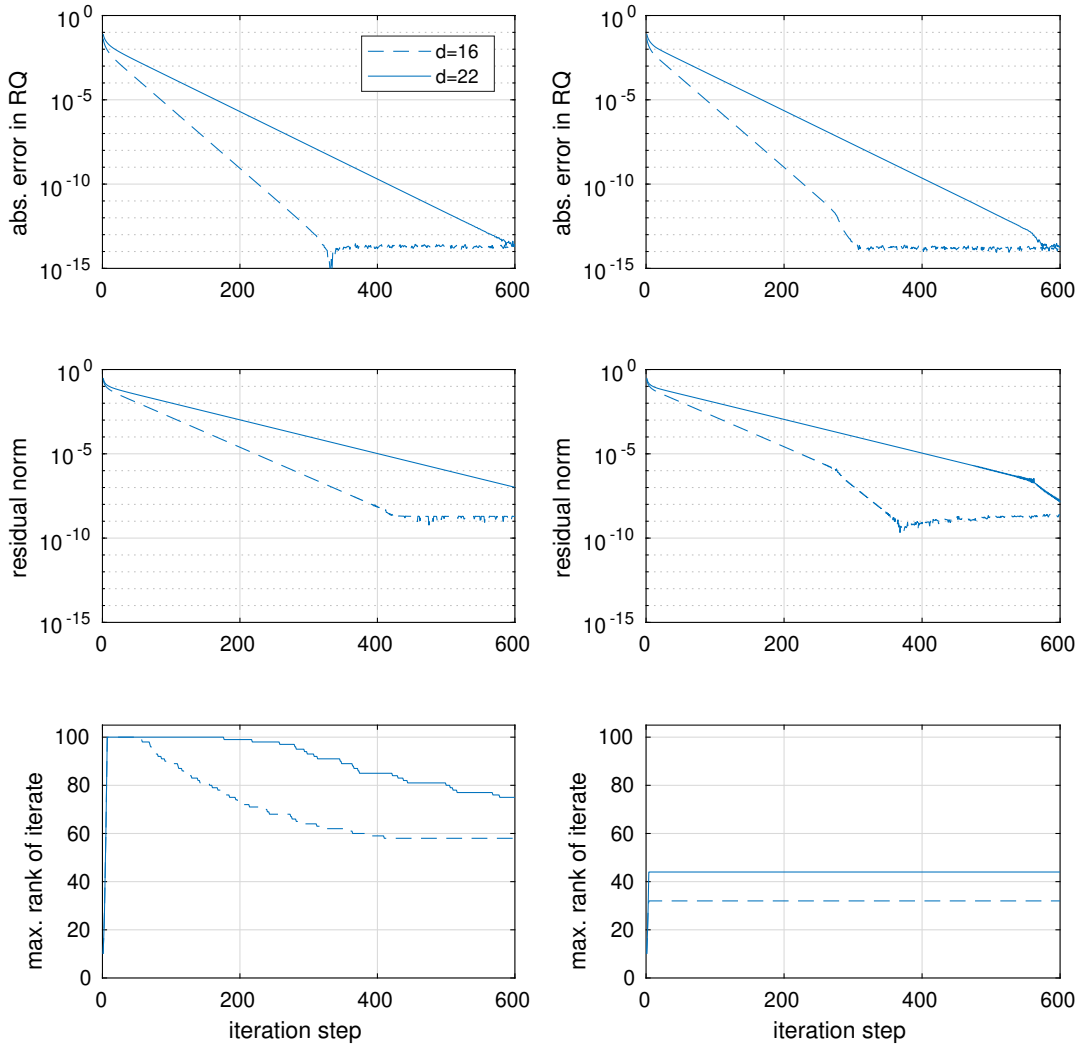


Figure 6.6.: 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, grad. desc., lin. tree, *left:* $r_{\max} = 100, \varepsilon_{\text{rel}} = 10^{-9}$, *right:* $r_{\max} = qd, \varepsilon_{\text{rel}} = 10^{-12}$

$r_{\text{final}}^{(16,10^{-6})} < 32 < r_{\text{final}}^{(16,10^{-9})}$, but nevertheless we recover the convergence behavior of RQ error and residual norm like for the linear tree.

The next test employs gradient descent instead of LOCG, now again for the linear dimension tree. As we read off from Figures 6.5 and 6.6, the convergence is much slower. We reach similar final values as with LOCG, but we need approximately five up to ten times as many iteration steps.

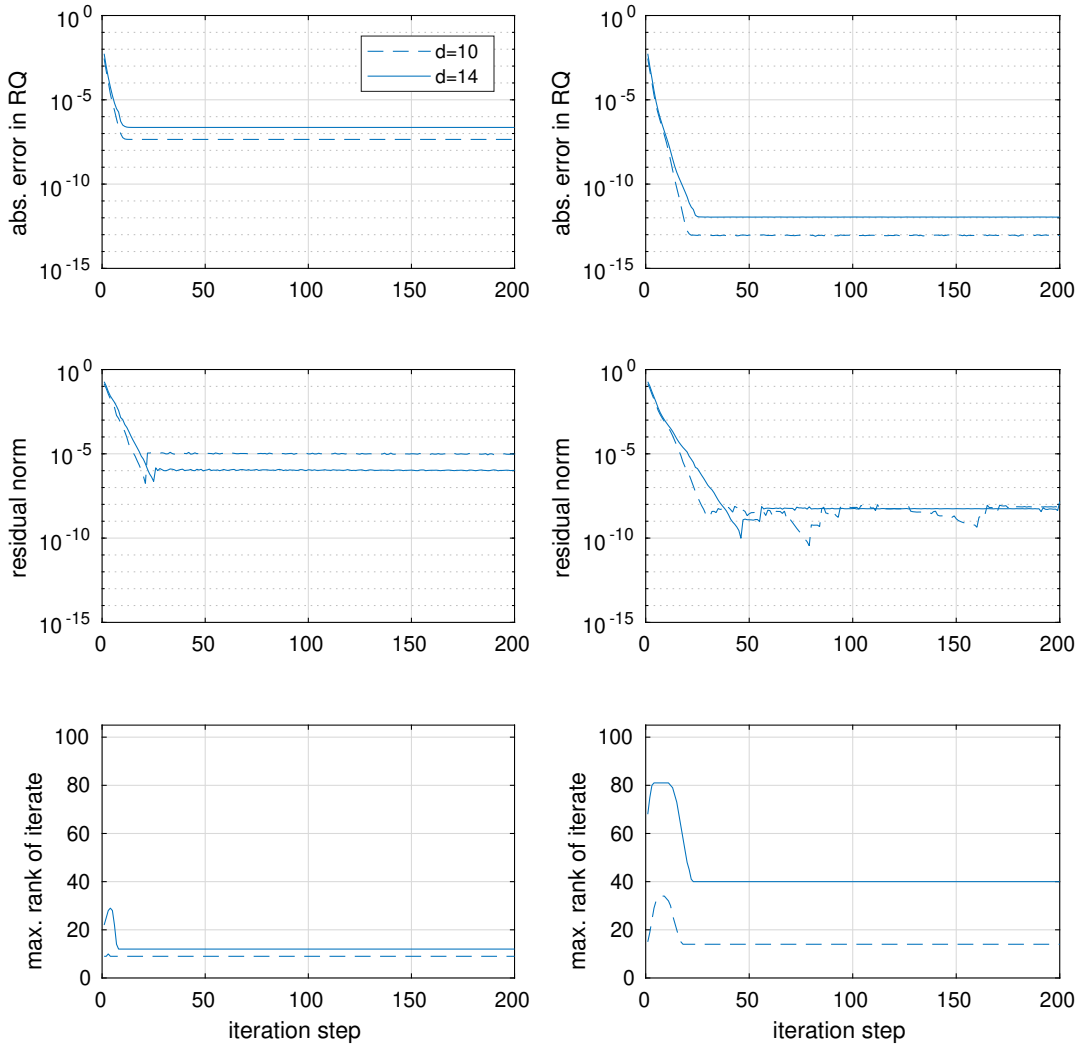


Figure 6.7.: 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (1, 0, 0, -0.7)$, $(d_1, d_2) = (2, 4)$, LOCG, bal. tree, $r_{\max} = 100$, *left:* $\varepsilon_{\text{rel}} = 10^{-3}$, *right:* $\varepsilon_{\text{rel}} = 10^{-6}$

For yet another test based on the balanced dimension tree, we consider an XYZ model for $q = 3$ with $A = 1$, $B = \Delta = 0$, $h = -0.7$, more precisely an Ising model. We set $d \in \{10, 14\}$ and still $(d_1, d_2) = (2, 4)$, employ LOCG, and compare again the parameters $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}, 10^{-9}\}$, $r_{\max} = 100$ with $\varepsilon_{\text{rel}} = 10^{-12}$, $r_{\max} = qd$. The results visible in Figures 6.7 and 6.8 concerning final RQ error and residual norm resemble those from the other test with balanced tree when we compare $q = 3$, $d = 10$ with $q = 2$, $d = 16$ and $q = 3$, $d = 14$ with $q = 2$, $d = 22$, noticing the proximity of qd in both cases. To name a difference, for the present Ising model we need in general less iteration steps as for the situation in Figures 6.3 and 6.4, and the final ranks $r_{\text{final}}^{(d,\varepsilon_{\text{rel}})}$ are smaller, except in case $\varepsilon_{\text{rel}} = 10^{-9}$, $d \in \{14, 22\}$. We observe further that for $d = 14$, the value $r_{\max} = qd = 42$ is a bit too small so that the RQ error could reach the otherwise minimal possible level between 10^{-13} and 10^{-14}, cf. a similar behavior in Figure 6.4.

As a last test concerning the truncation parameters, we examine a 3-Potts model with

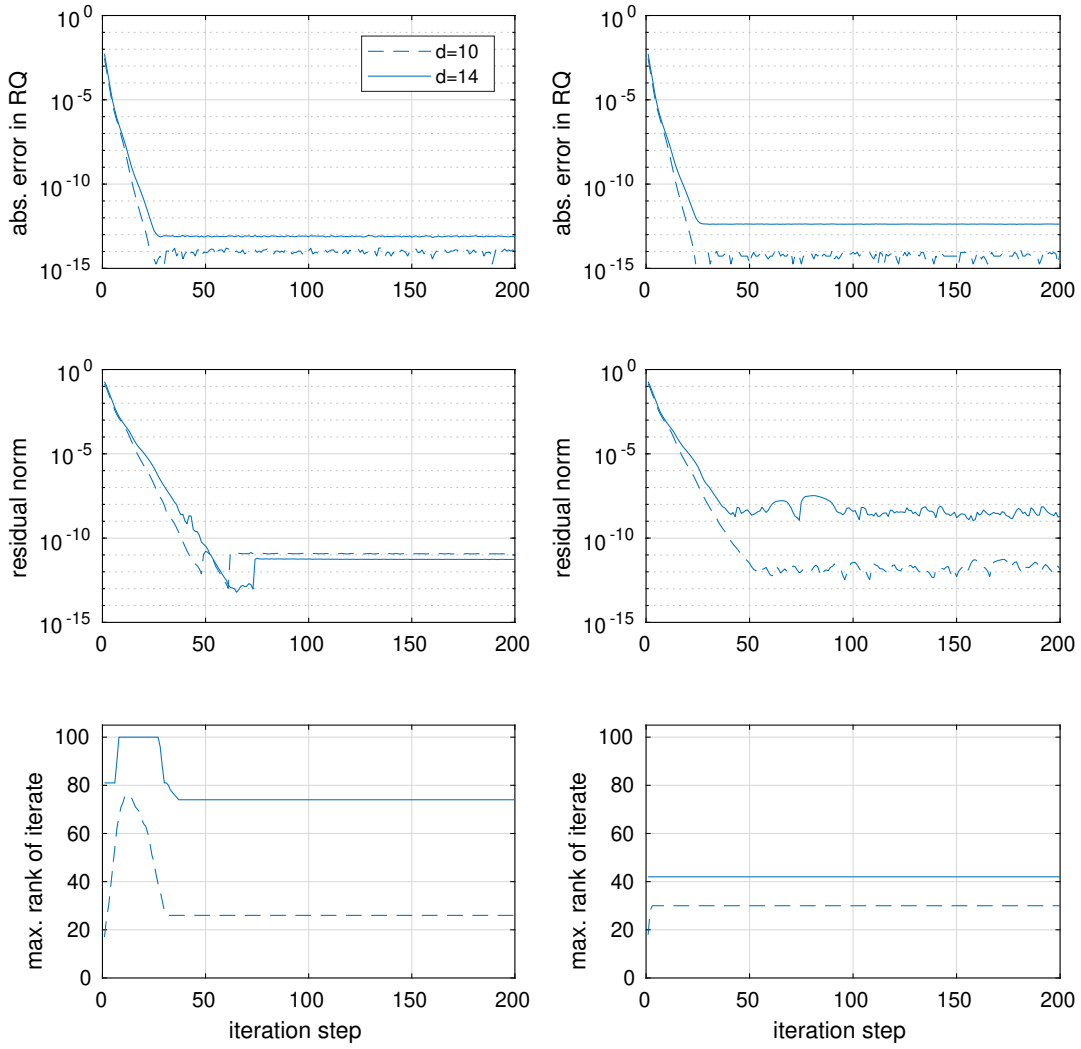


Figure 6.8.: 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (1, 0, 0, -0.7)$, $(d_1, d_2) = (2, 4)$, LOCG, bal. tree, *left:* $r_{\max} = 100$, $\varepsilon_{\text{rel}} = 10^{-9}$, *right:* $r_{\max} = qd$, $\varepsilon_{\text{rel}} = 10^{-12}$

$A = 1.6$, $h = -0.3$, $d \in \{10, 14\}, (d_1, d_2) = (2, 4)$, LOCG, and linear dimension tree. We collect the results in Figures 6.9 and 6.10. Compared to the XYZ models discussed so far, convergence is faster and already for $\varepsilon_{\text{rel}} = 10^{-6}$ and with $r_{\text{final}}^{(d,10^{-6})} = 15$, an RQ error around 10^{-13} is realized. $\varepsilon_{\text{rel}} = 10^{-9}$ yields a larger final rank and for $d = 14$ a smaller residual norm. The only advantage we observe for the choice $r_{\max} = qd$ is that the residual norm is also in the order of 10^{-11} or 10^{-13}.

To sum up, in each of the test cases above, the choice $r_{\max} = 100$ and $\varepsilon_{\text{rel}} = 10^{-9}$ is appropriate in the sense that the error in the Rayleigh quotient converges down to the observed minimal possible order of 10^{-13}. The maximal rank $r_{\text{final}}^{(d,10^{-9})}$ of the final iterates is bounded by a value between 60 and 80, for the investigated 3-Potts model only around 30. A prescribed value of $r_{\max} = qd$ turned out also to be sufficient in the case of a linear dimension tree, being however a bit too small in some settings with a balanced tree and a bit too large for the Potts model with $q = 3, d = 14$. Hence we decide to utilize for the

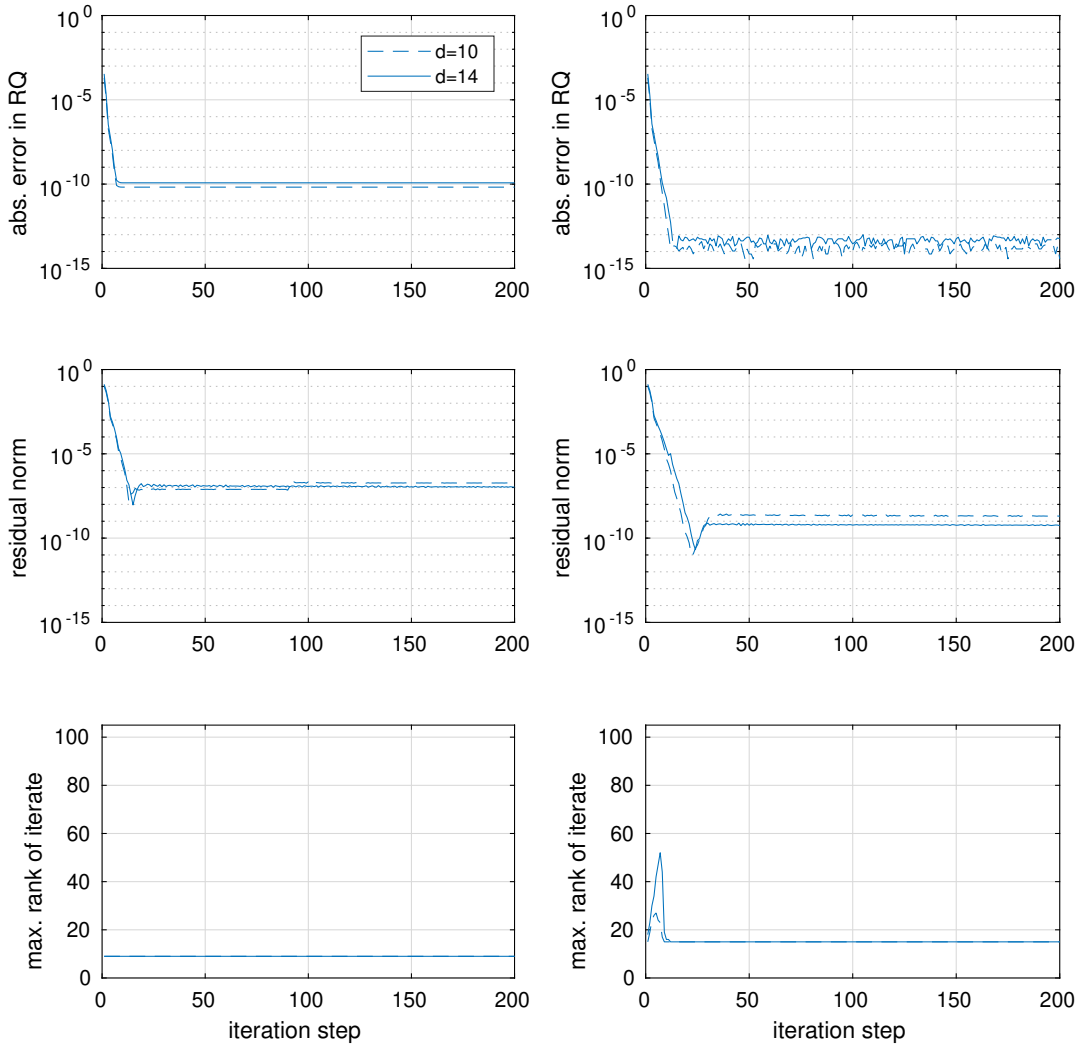


Figure 6.9.: 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.6, -0.3)$, $(d_1, d_2) = (2, 4)$, LOCG, lin. tree, $r_{\max} = 100$, *left:* $\varepsilon_{\text{rel}} = 10^{-3}$, *right:* $\varepsilon_{\text{rel}} = 10^{-6}$

various tests to be performed in the rest of this chapter $r_{\max} = 100$ and $\varepsilon_{\text{rel}} = 10^{-9}$ as the parameters governing the truncations.

These parameters are also applied in the tests with MALS formulated in TT format. For a given tensor $\mathbf{T} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, recall the quasi-best approximation property from [Ose11, Sect. 2],

$$\|\mathbf{T} - \widehat{\mathbf{T}}\| \leq \sqrt{\sum_{j=1}^{d-1} \sum_{i=r_j+1}^{\min\{n_1 \cdots n_j, n_{j+1} \cdots n_d\}} \sigma_i \left(\text{resh}_{n_1 \cdots n_j \times n_{j+1} \cdots n_d}(\mathbf{T})\right)^2} \leq \sqrt{d-1} \|\mathbf{T} - \mathbf{T}_{\text{best}}\|,$$

where $\widehat{\mathbf{T}}$ is represented in TT format with TT ranks r_j, $1 \leq j \leq d-1$, as a result of the truncation procedure [Ose11, Alg. 1], and $\mathbf{T}_{\text{best}}$ is a minimizer of $\|\mathbf{T} - \mathbf{S}\|$ among all tensors $\mathbf{S}$ represented with TT ranks r_j. Similar to (6.1) and (6.2), the rank update $r_k^{(j)}$ in Line 13

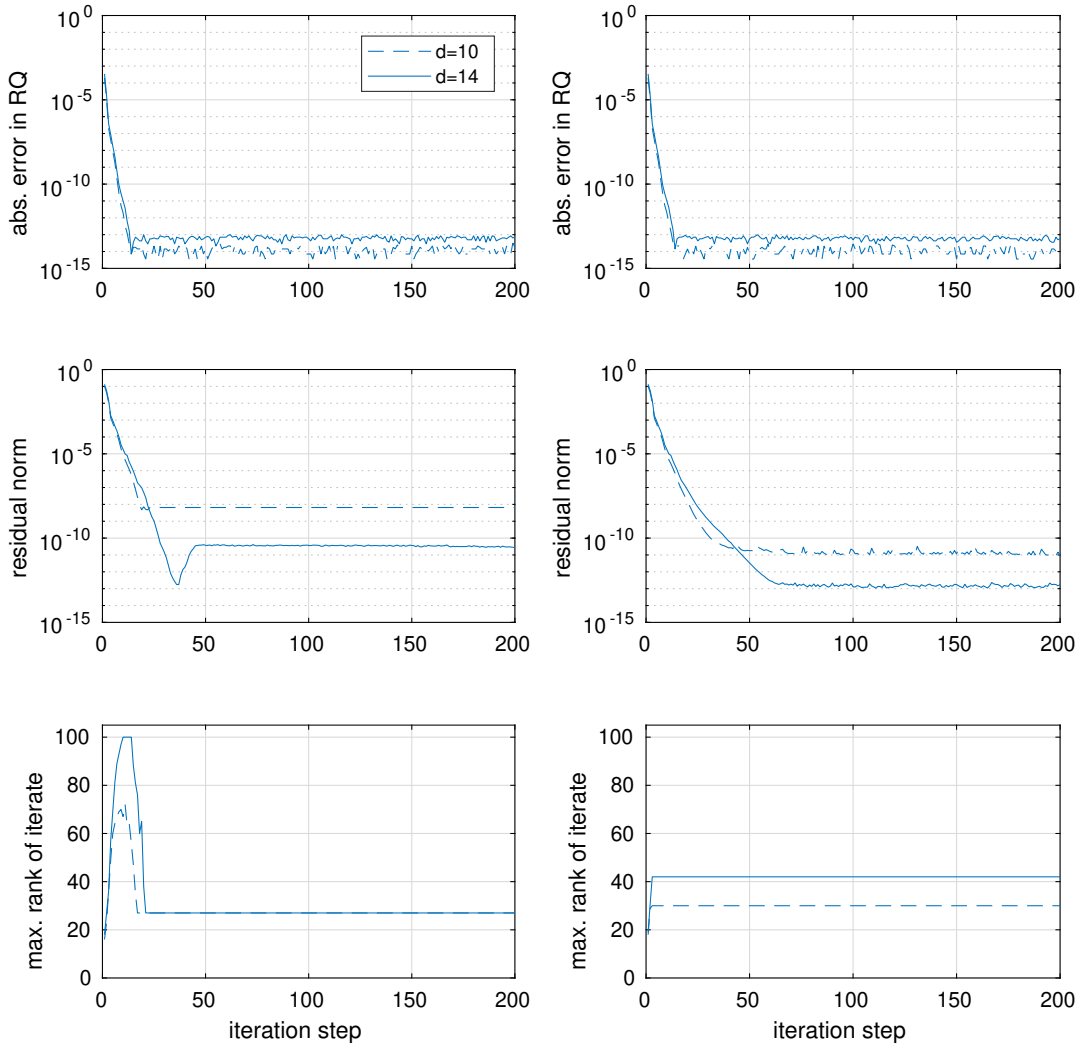


Figure 6.10.: 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.6, -0.3)$, $(d_1, d_2) = (2, 4)$, LOCG, lin. tree, *left:* $r_{\max} = 100$, $\varepsilon_{\mathrm{rel}} = 10^{-9}$, *right:* $r_{\max} = qd$, $\varepsilon_{\mathrm{rel}} = 10^{-12}$

or 23 in Algorithm 4.3 is chosen such that

$$\sqrt{\sum_{i \geq r_k^{(j)}+1} \left((\mathbf{\Sigma}^{(j)})_{i,i}\right)^2} \leq \frac{\varepsilon_{\mathrm{rel}} \|\mathbf{\Sigma}^{(j)}\|}{\sqrt{d-1}}.$$

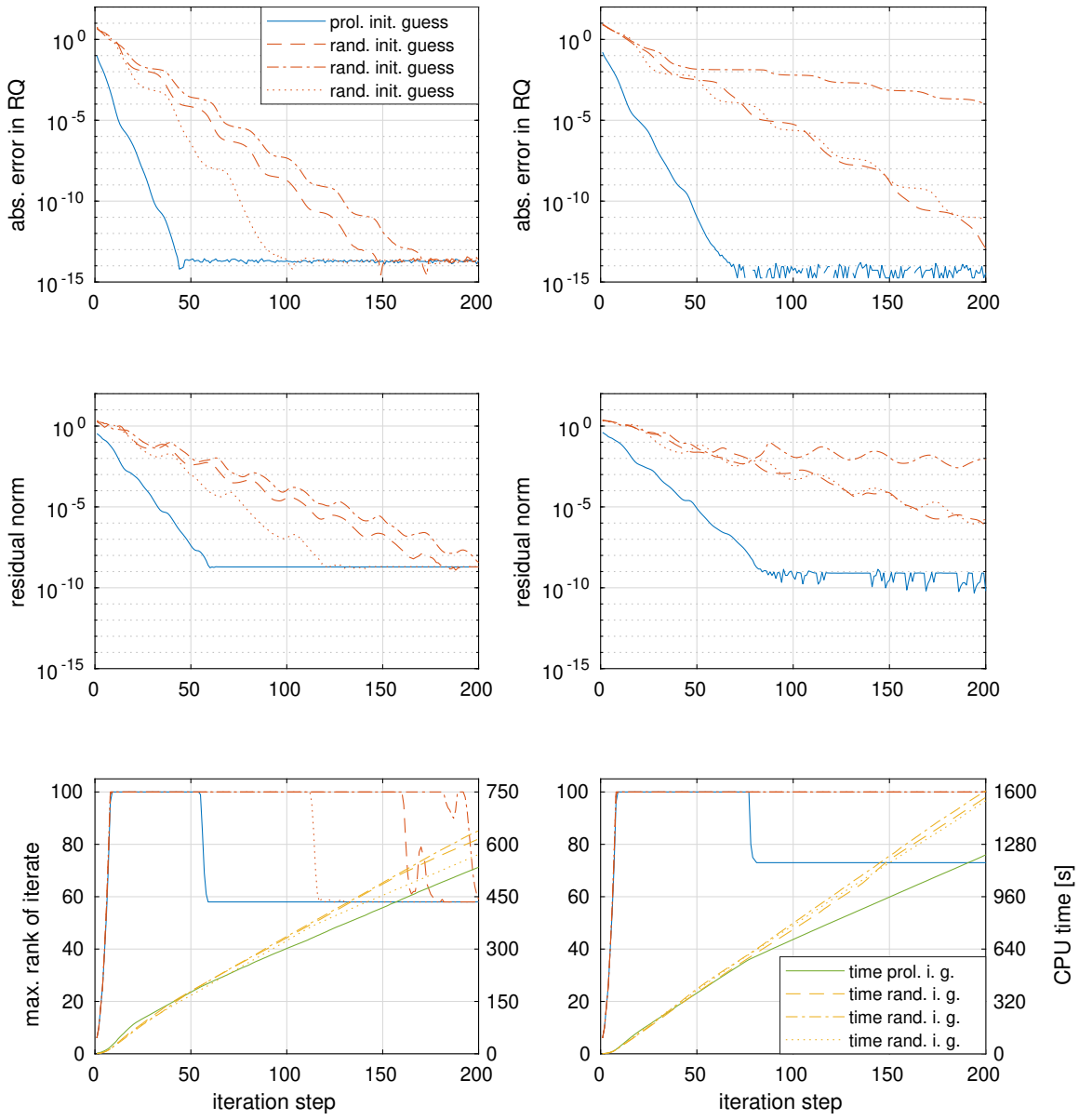


Figure 6.11.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 3)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 22$

6.2. HT format, $A \neq B$

Like in Section 6.1 and in the remainder of this chapter, the presentation of the tests is organized in such a way that the three plots in one of the two columns in a figure belong to one specific scenario which is described by the respective caption. The three plotted quantities are the absolute error in the Rayleigh quotient (6.3), the residual norm (6.4), and the maximal rank of the representative of the current iterate, where the upper bound on the ranks in the truncation equals the largest number labeling the y-axis, mostly 100,

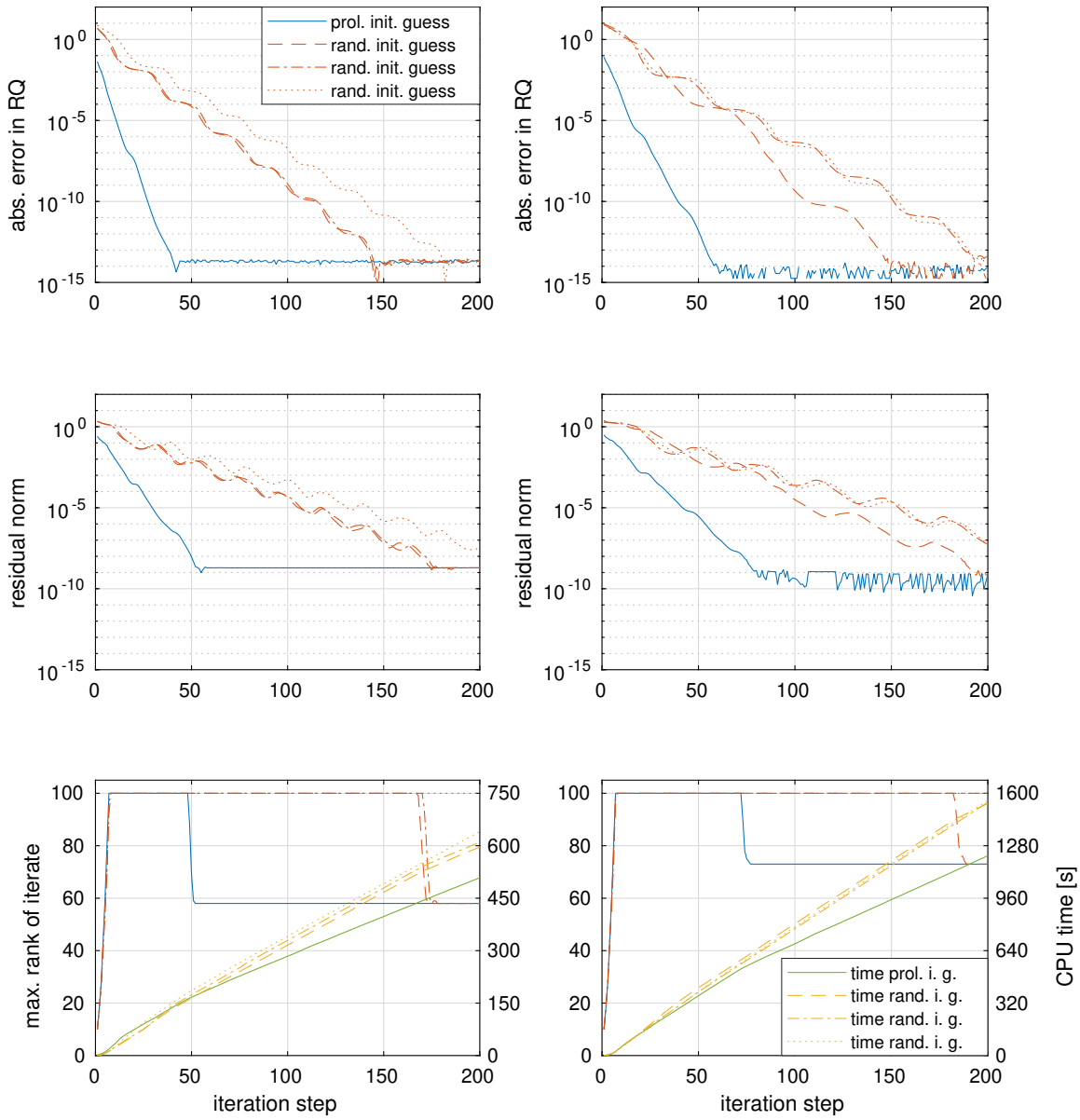


Figure 6.12.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 22$

unless otherwise stated. In a number of test cases, we additionally report on the CPU time consumed by the numerical method. The two columns in one figure usually differ only by one different parameter of the problem setting, for example the value of d, the shape of the dimension tree, or the used algorithm. Typically, several consecutive figures belong to the same value of coupling parameters A, B, Δ, h respectively A, h and differ by the value of (d_1, d_2). In a single plot, the blue line, or in some situations two blue lines, belong to the case when the initial guess is constructed by prolongation based on problem sizes (d_1, d_2) and the, mostly three, red lines result from the performance of a random initial guess. These random

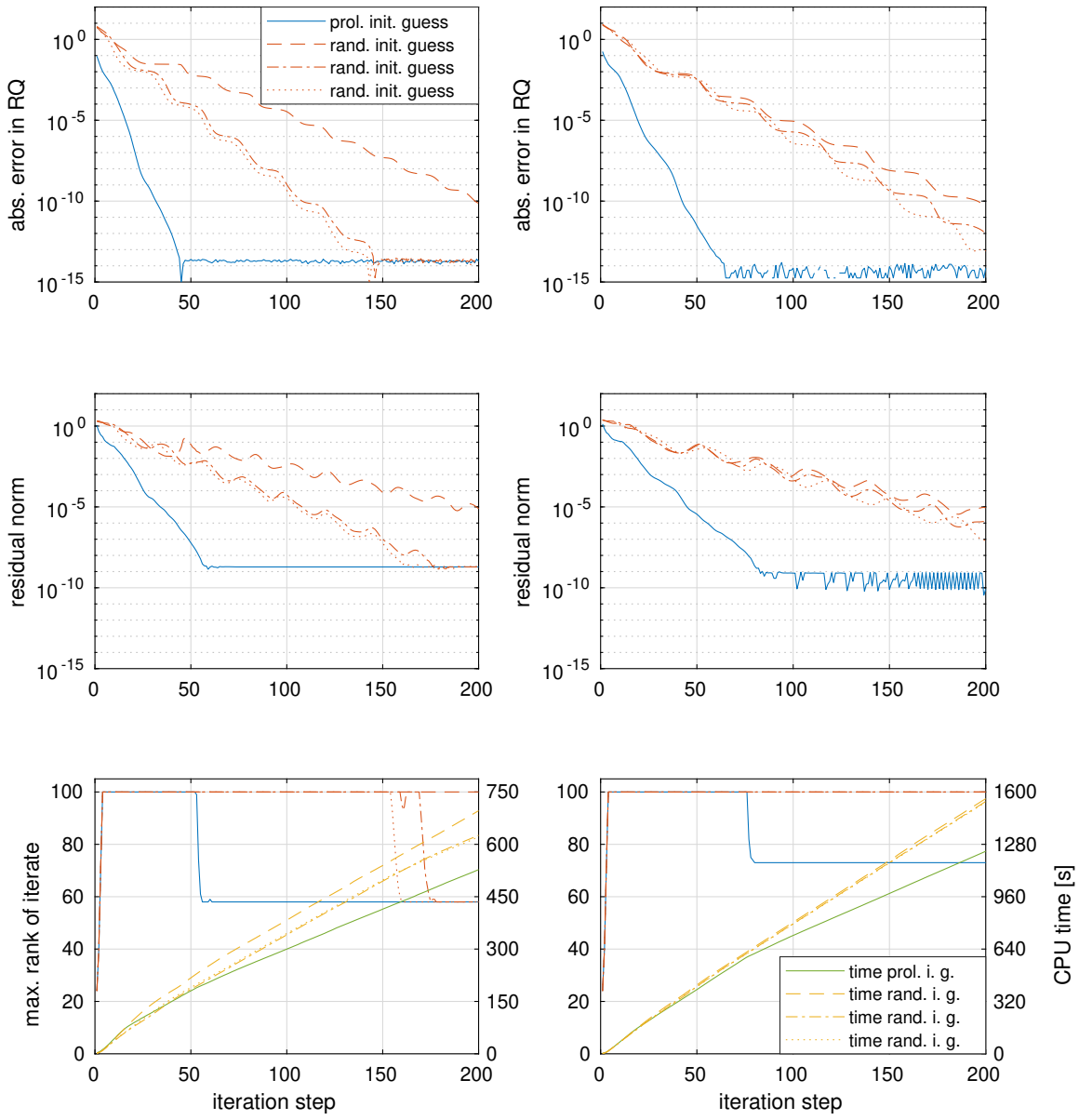


Figure 6.13.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (3, 4)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 22$

initial guesses are set up with ranks equal to those of the prolongated initial guess which in turn depend on the particular construction with exact representation in the tensor format, hence on (d_1, d_2) and q, see Section 5.2, especially Remark 5.5. If present, and indicated by the y-axis at the right edge, the green solid line in the plot in the third row visualizes the CPU time for the prolongated initial guess, and the yellow dashed and dotted lines the respective value for the random initial guesses.

We start in Figures 6.11 up to 6.15 with a 2-XYZ model for $A = 1.9$, $B = 0.4$, $\Delta = -1.1$, $h = 0.2$. Concerning the five different values of (d_1, d_2), we observe that for fixed d_1

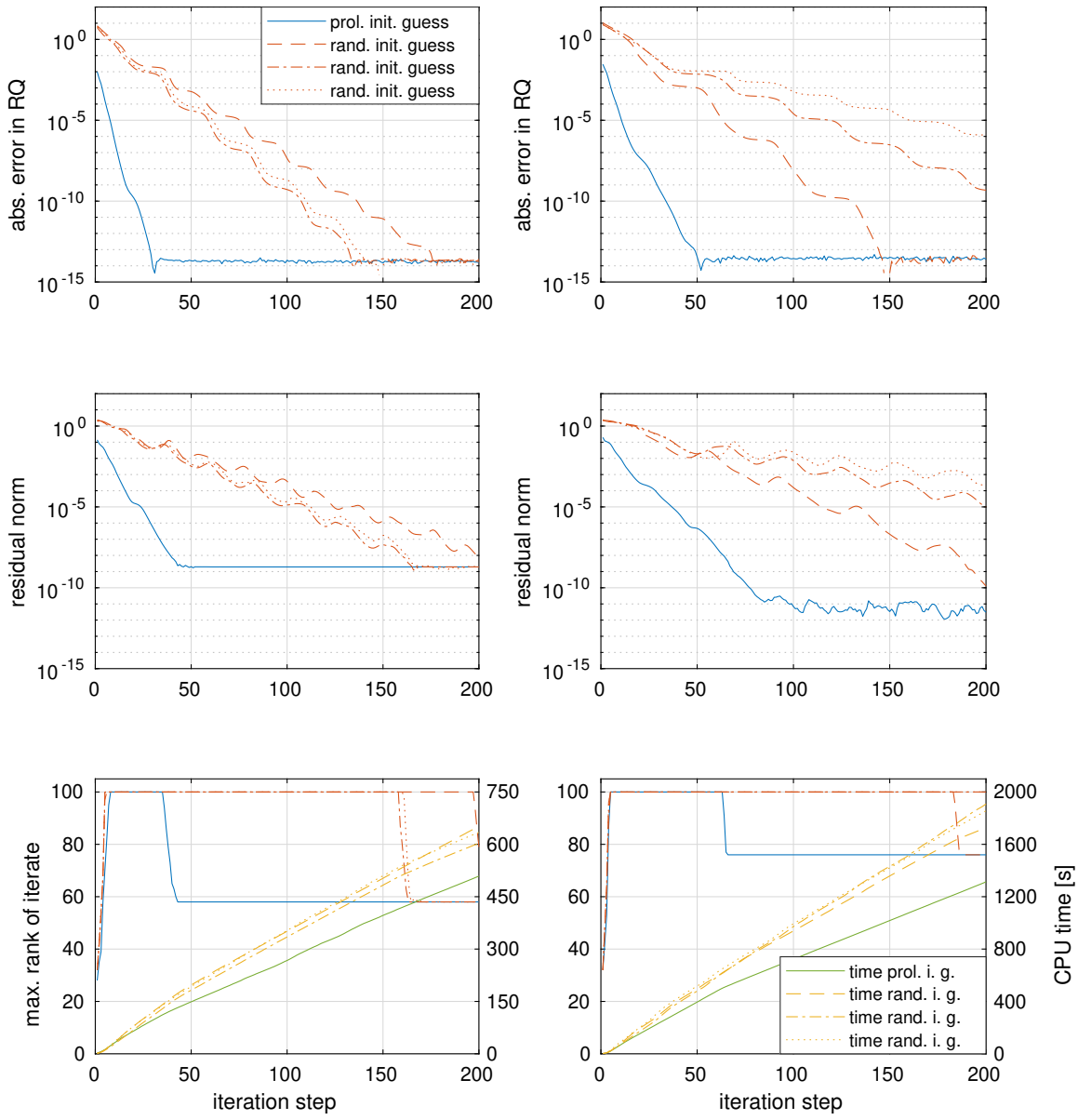


Figure 6.14.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 9)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 23$

a larger d_2 and vice versa yields a faster convergence with the exception that for $(3, 4)$ the convergence is a bit slower than for $(2, 4)$. Convergence of the maximal rank sets in shortly before the residual norm has reached its final value. So, besides the stagnation of the residual norm, this drop of the maximal HT rank, existent if $r_{\max}$ was chosen sufficiently large, may be used as a stopping criterion for the algorithm. In each situation, the prolongated initial guess yields faster convergence than the random initial guess. At least twice up to more than ten times as many iteration steps are needed. This also leads to a smaller CPU time until reaching a low error level for the prolongated initial guess. Moreover, we observe that

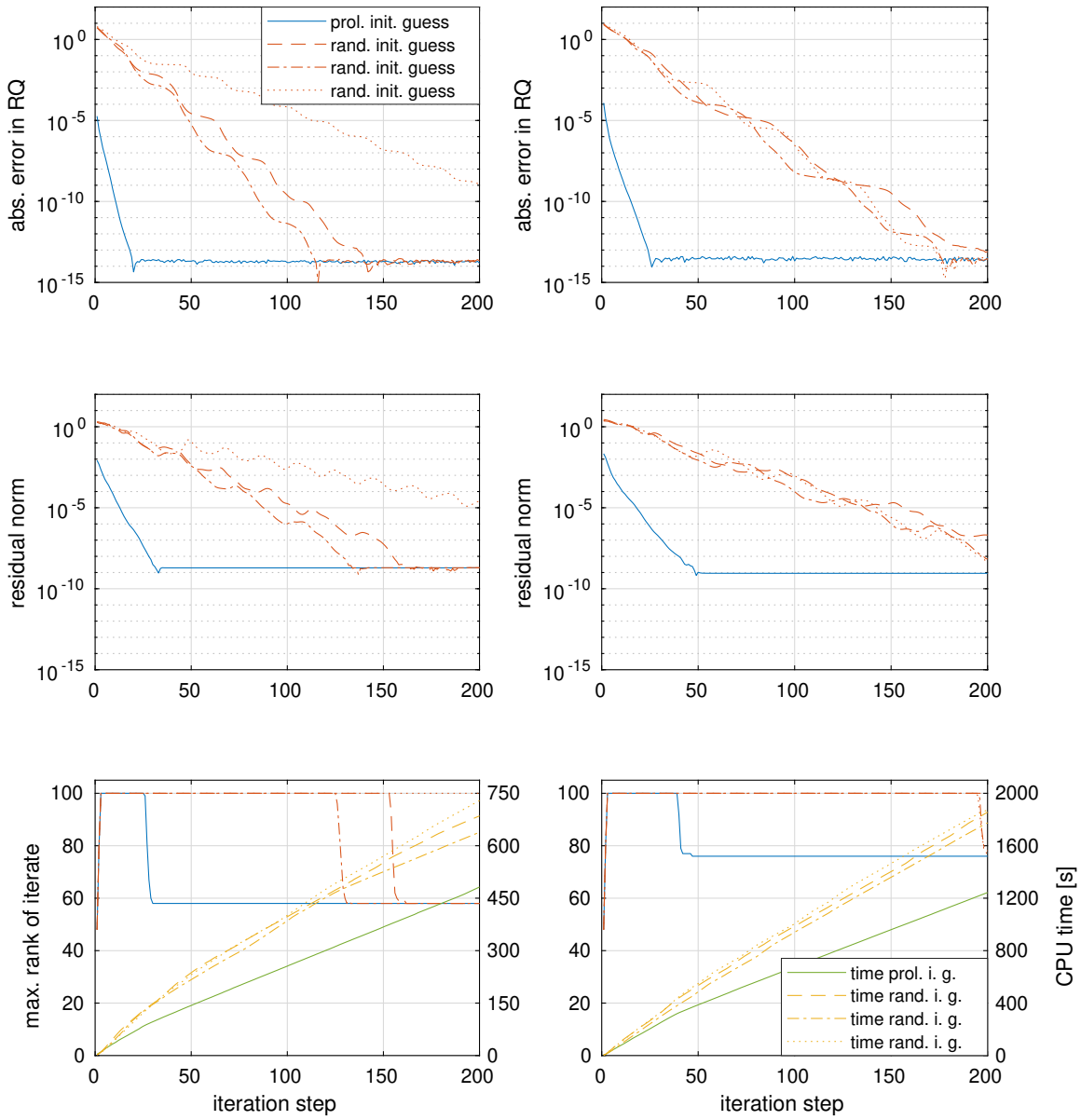


Figure 6.15.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (8, 9)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 23$

the CPU time per iteration step, hence the slope of the green and yellow lines, depends on the maximal rank of the current iterate.

We continue in Figures 6.16 to 6.18 with a 2-XYZ model for $A = -0.9$, $B = 1.6$, $\Delta = 1.7$, $h = 1.2$. If $(d_1, d_2) = (2, 3)$ and if the prolongated initial guess $\widetilde{\mathbf{v}}^{(d)}$ is based on $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$, we face the problem that for both $d \in \{16, 22\}$ the iterates converge, measured in the Rayleigh quotient, to the second smallest eigenvalue $\lambda_2^{(d)}$ since $\lambda_2^{(16)} - \lambda_{\min}^{(16)} \approx 3.1 \cdot 10^{-4}$ and $\lambda_2^{(22)} - \lambda_{\min}^{(22)} \approx 1.2 \cdot 10^{-5}$. This is due to $\mathbf{v}_{\min}^{(d)} \in \mathcal{E}_{d,2}^{\text{odd}}$, but $\widetilde{\mathbf{v}}^{(d)} \in \mathcal{E}_{d,2}^{\text{even}}$ and during the

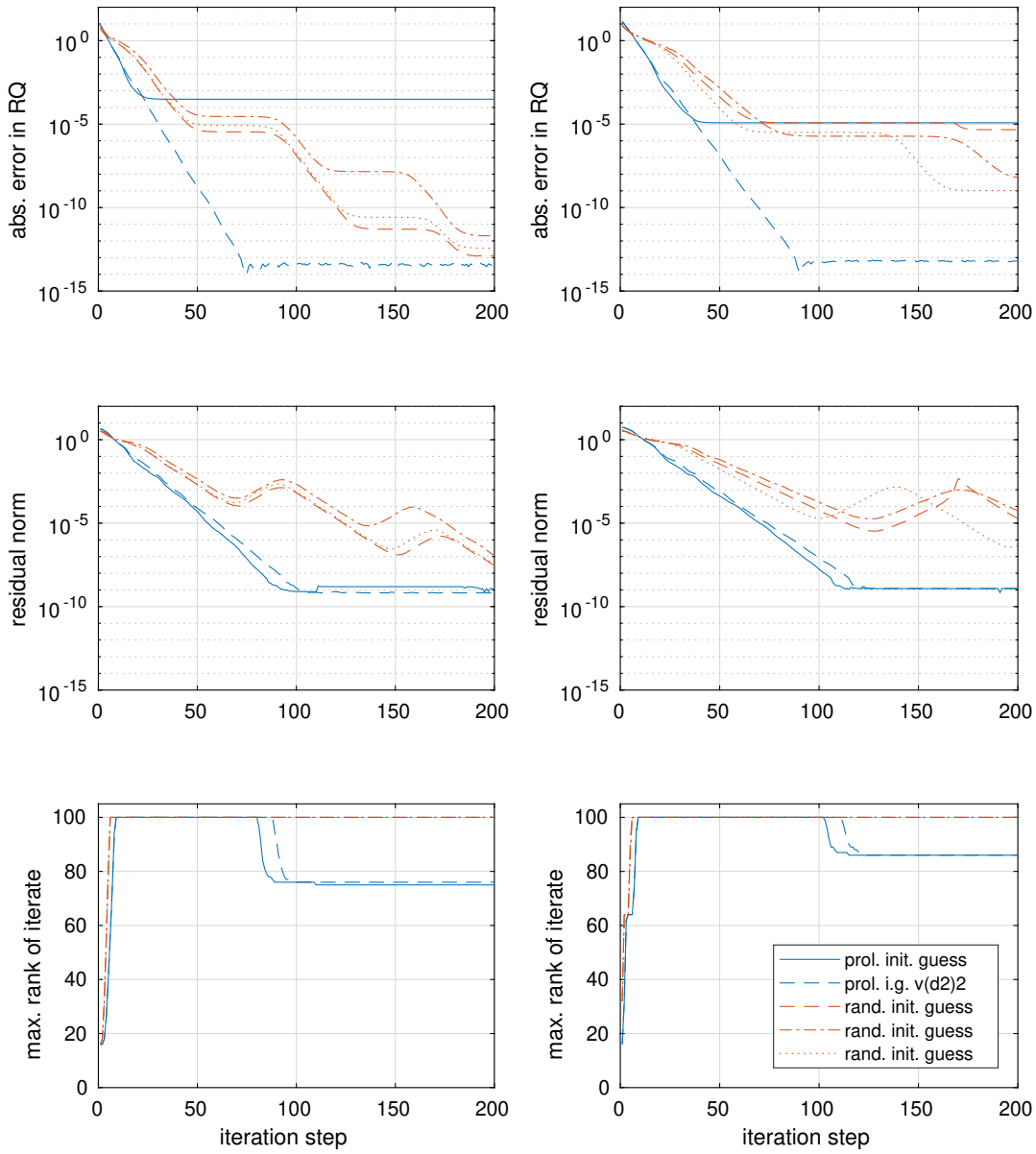


Figure 6.16.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 3)$, LOCG, balanced tree, *left:* $d = 16$, *right:* $d = 22$

iteration, the iterates remain to represent vectors from this set. As a remedy we may choose

$$\widetilde{\mathbf{v}}_2^{(d)} := \left(\prod_{i=1}^{d-3} (\mathbf{I}_{2^i} \otimes \mathbf{M}) \right) \mathbf{v}_2^{(d_2)} \in \mathcal{E}_{d,2}^{\mathrm{odd}} \tag{6.5}$$

as an initial guess having the same sparsity pattern like $\mathbf{v}_{\min}^{(d)}$, where $\mathbf{v}_2^{(d_2)}$ is a normalized eigenvector associated with $\lambda_2^{(d_2)}$, and still

$$\mathbf{M} := \mathbf{I}_2 \otimes \mathrm{mat}_{4\times 2}(\mathbf{v}_{\min}^{(3)}) \left(\mathrm{mat}_{2\times 2}(\mathbf{v}_{\min}^{(2)}) \right)^{-1},$$

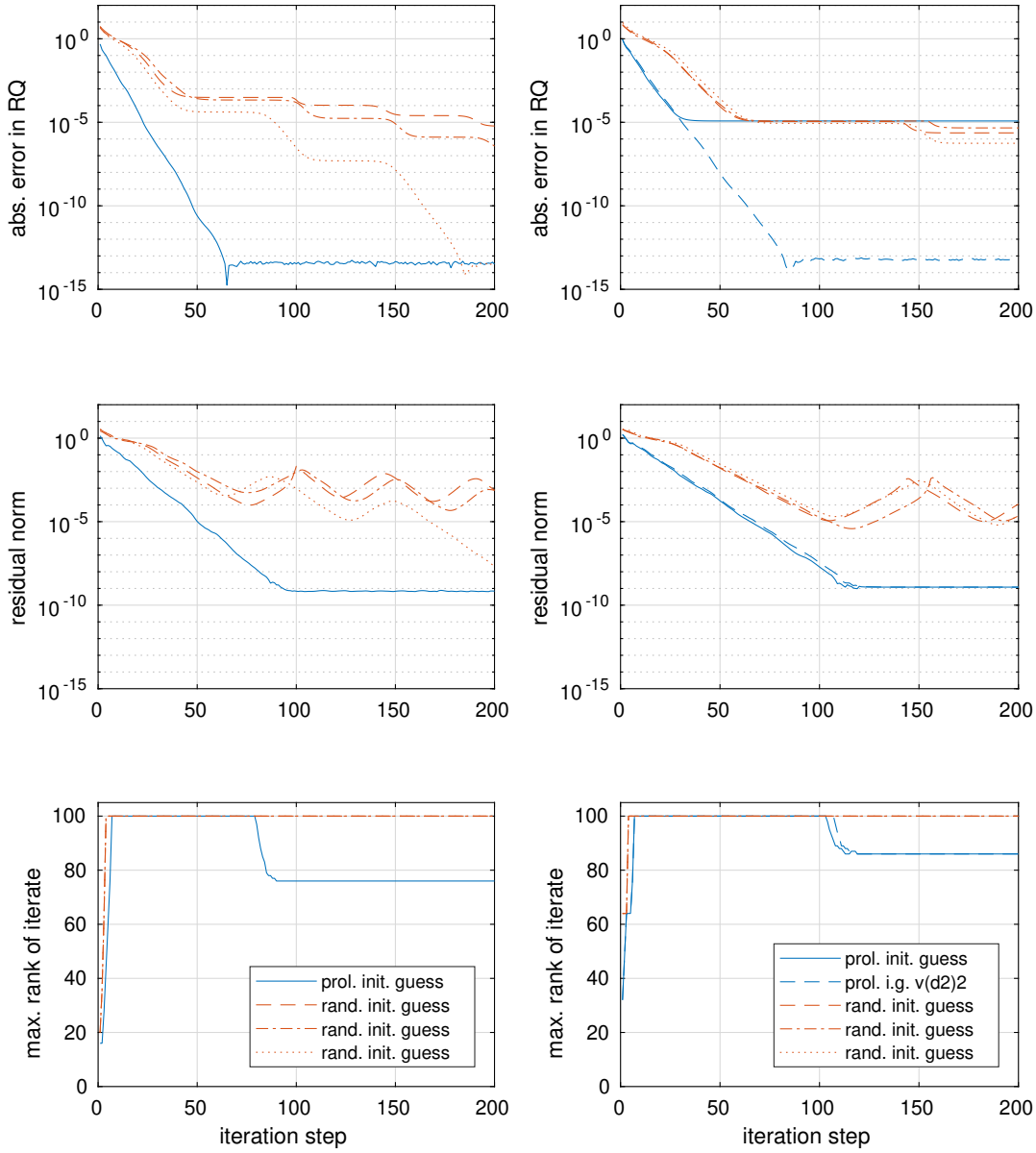


Figure 6.17.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree, *left:* $d = 16$, *right:* $d = 22$

cf. (5.1). The construction of a representative in a tensor format of the tensorization of (6.5) is completely analogous to that described in Section 5.2. As illustrated by the dashed blue line, we observe that $\widetilde{\mathbf{v}}_2^{(d)}$ yields convergence in the Rayleigh quotient to $\lambda_{\min}^{(d)}$. It is again much faster than for the random initial guess. The final maximal ranks of the HT representatives of both $\mathbf{v}_{\min}^{(d)}$ and $\mathbf{v}_2^{(d)}$ are practically equal. Just as well, the decrease of the error in RQ and of the residual norm until convergence is almost equal. Regarding Figure 6.17, for $(d_1, d_2) = (2, 4)$ the initial guess $\widetilde{\mathbf{v}}^{(16)} \in \mathcal{E}_{16,2}^{\mathrm{odd}}$ has the correct sparsity pattern

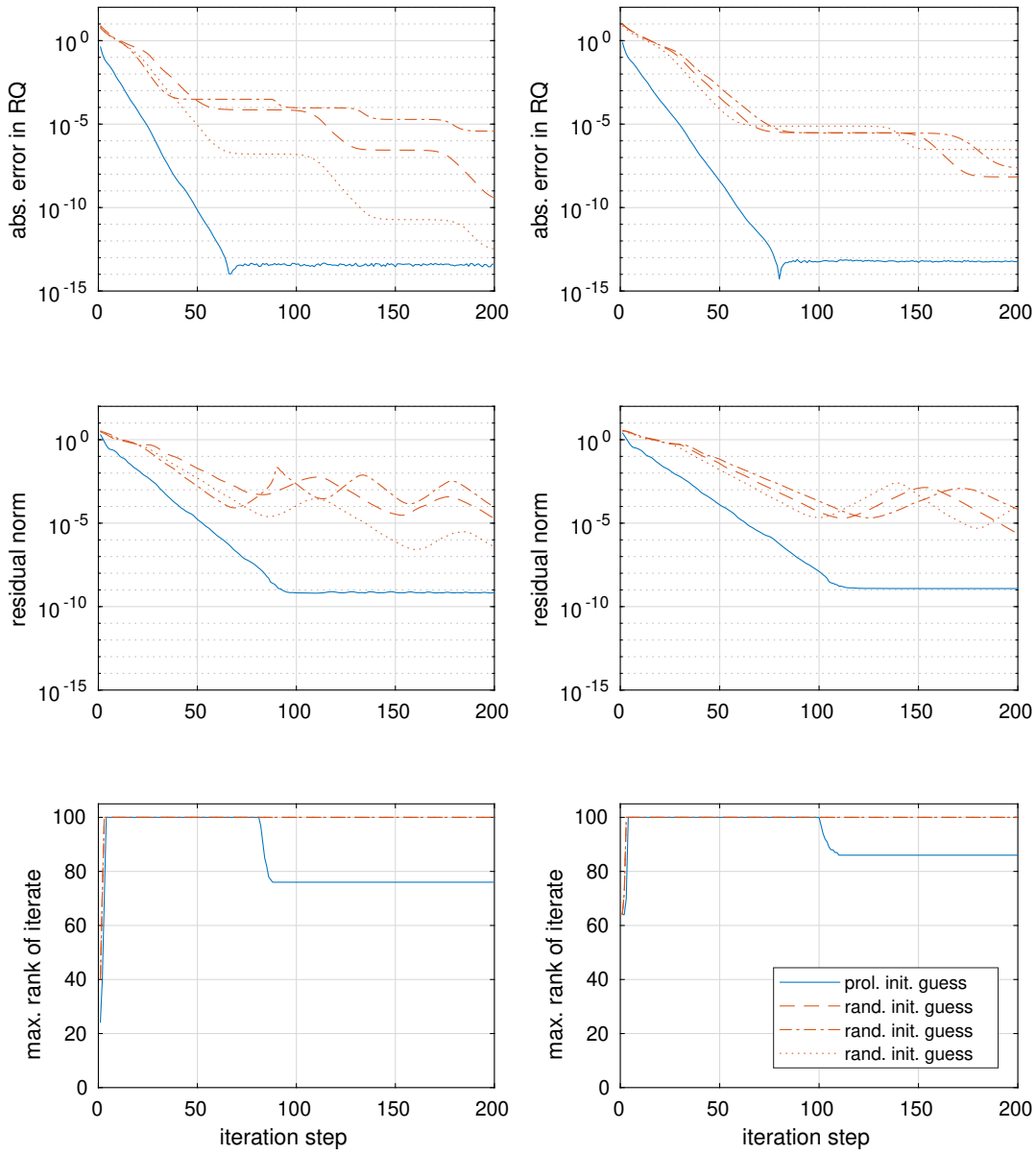


Figure 6.18.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (3, 4)$, LOCG, balanced tree, *left:* $d = 16$, *right:* $d = 22$

and the RQ converges to $\lambda_{\min}^{(16)}$. However, $\widetilde{\mathbf{v}}^{(22)} \in \mathcal{E}_{22,2}^{\text{even}}$ has the wrong pattern and, like for $(d_1, d_2) = (2, 3)$, convergence is to $\lambda_2^{(22)}$. The alternative choice $\widetilde{\mathbf{v}}_2^{(22)}$, defined analogously to (6.5), in turn behaves well and leads to slightly faster convergence than in the case $(d_1, d_2) = (2, 3)$. In the situation $(d_1, d_2) = (3, 4)$ of Figure 6.18, $\widetilde{\mathbf{v}}^{(d)} \in \mathcal{E}_{d,2}^{\text{odd}}$ for both $d \in \{16, 22\}$ has the correct sparsity pattern and we need approximately the same number of iteration steps until convergence as for $(d_1, d_2) = (2, 4)$.

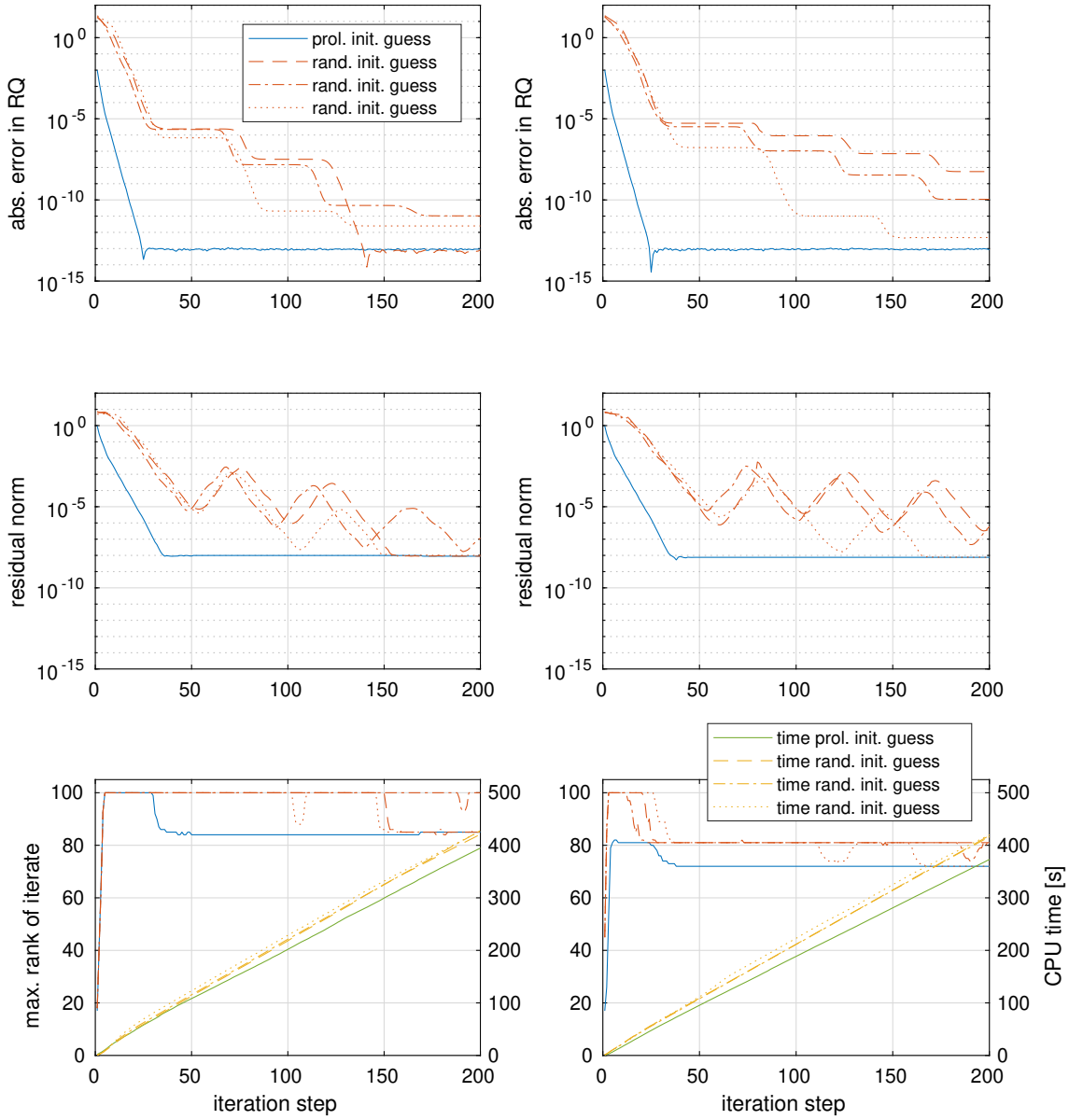


Figure 6.19.: 3-XYZ, $d = 10$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 4)$, LOCG, *left:* linear tree, *right:* balanced tree

We move on to the next test case, now $q = 3$, an XYZ model with $A = 1.6$, $B = -0.3$, $\Delta = 2.9$, $h = -0.8$. Besides the comparison between prolongated and random initial guess, we investigate whether or not the shape of the dimension tree, linear or balanced, has an effect on the convergence behavior. The left resp. right column in Figures 6.19 and 6.20 correspond to the linear resp. balanced tree. In the case $d = 10$, $(d_1, d_2) = (2, 4)$, the evolution of RQ error and residual norm is very similar, only the final rank as well as the consumed CPU time is a bit smaller for the balanced tree. For $d = 14$, $(d_1, d_2) = (2, 6)$ we observe practical equality of RQ error, residual norm, and maximal rank. However, the CPU time is twice as

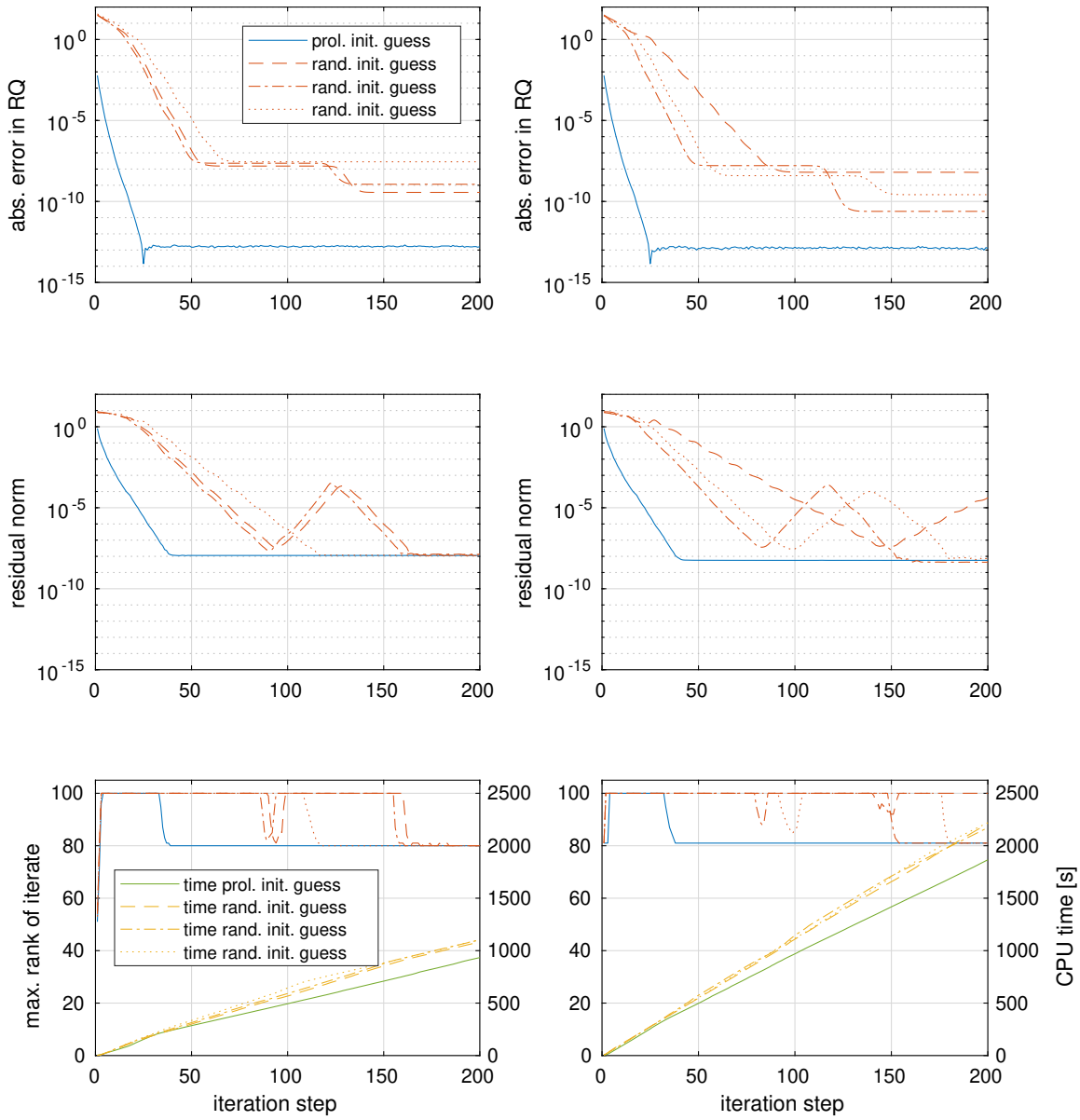


Figure 6.20.: 3-XYZ, $d = 14$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 6)$, LOCG, *left:* linear tree, *right:* balanced tree

large for the balanced dimension tree. Also in an analogous test, Figure 6.21, with coupling parameters $A = 2.6$, $B = 0.7$, $\Delta = -1.9$, $h = -0.3$ and $d = 14$, $(d_1, d_2) = (5, 6)$, there is a similar difference of the CPU time while the other three traced quantities do not deviate with respect to the shape of the dimension tree. Again we observe that a larger rank yields a larger CPU time per iteration step. Notice that the upper bound of the maximal rank is set to 150 in this single test. In each case, a clear benefit of the prolongated initial guess is visible.

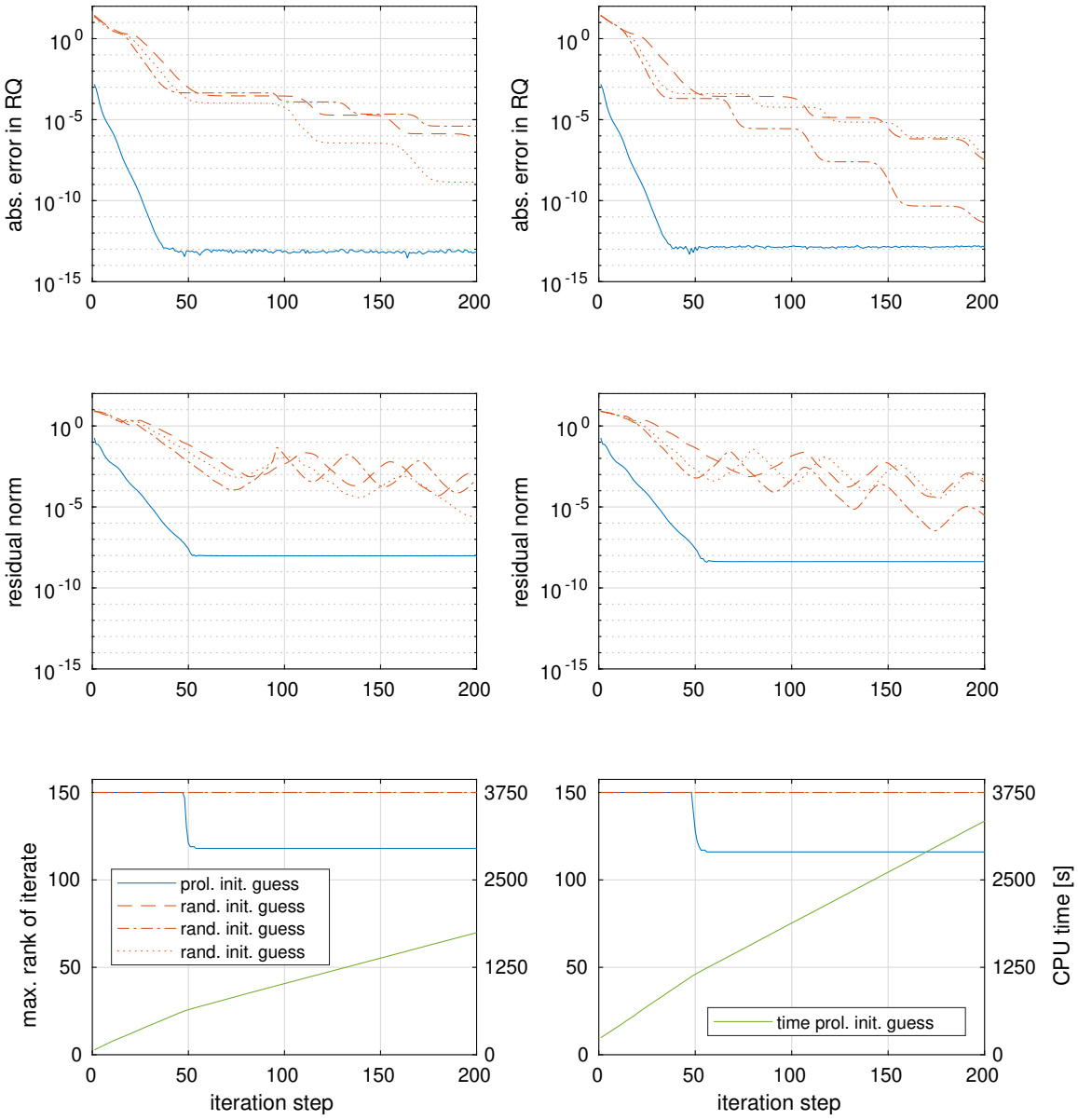


Figure 6.21.: 3-XYZ, $d = 14$, $(A, B, \Delta, h) = (2.6, 0.7, -1.9, -0.3)$, $(d_1, d_2) = (5, 6)$, LOCG, *left:* linear tree, *right:* balanced tree

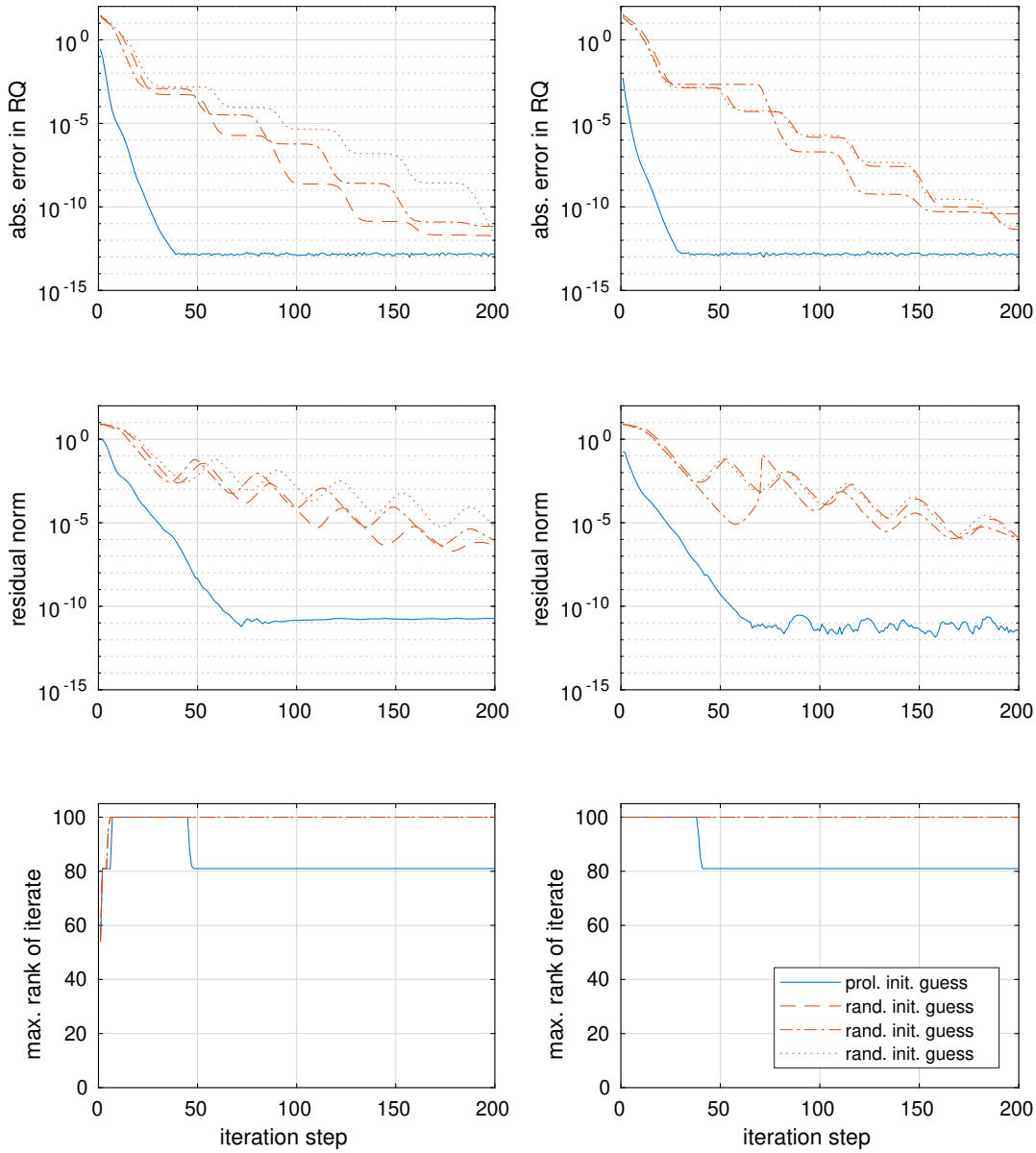


Figure 6.22.: 3-Potts, $d = 14$, $(A, h) = (1.3, 0.9)$, LOCG, balanced tree, *left:* $(d_1, d_2) = (2, 3)$, *right:* $(d_1, d_2) = (5, 6)$

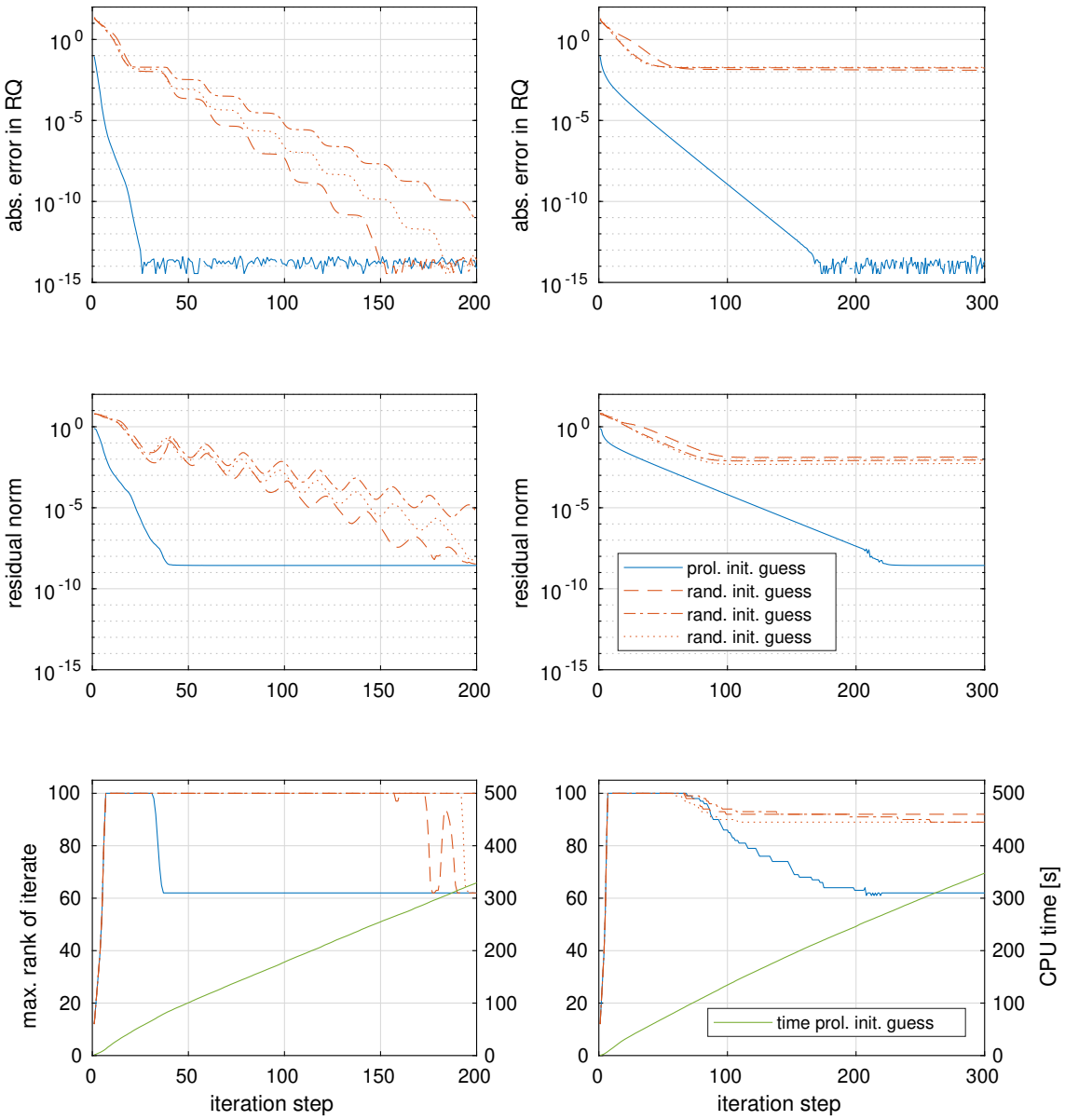


Figure 6.23.: 3-Potts, $d = 10$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (2, 3)$, linear tree, *left:* LOCG, *right:* gradient descent

The last class of Hamilton operators we discuss in this section is a 3-Potts model with $A = 1.3$, $h = 0.9$. In Figure 6.22 we compare two different choices $(d_1, d_2) \in \{(2, 3), (5, 6)\}$ and get a reduction of the necessary number of iteration steps by about one fourth. A test with respect to the used numerical method is depicted in Figures 6.23 and 6.24. For both scenarios $d = 10$, $(d_1, d_2) = (2, 3)$ and $d = 14$, $(d_1, d_2) = (3, 4)$ we observe that about seven times as many iteration steps are needed when employing gradient descent instead of LOCG. Again the prolongated initial guess is beneficial. Especially with gradient descent, the random initial guesses yield stagnation concerning the Rayleigh quotient at some value different

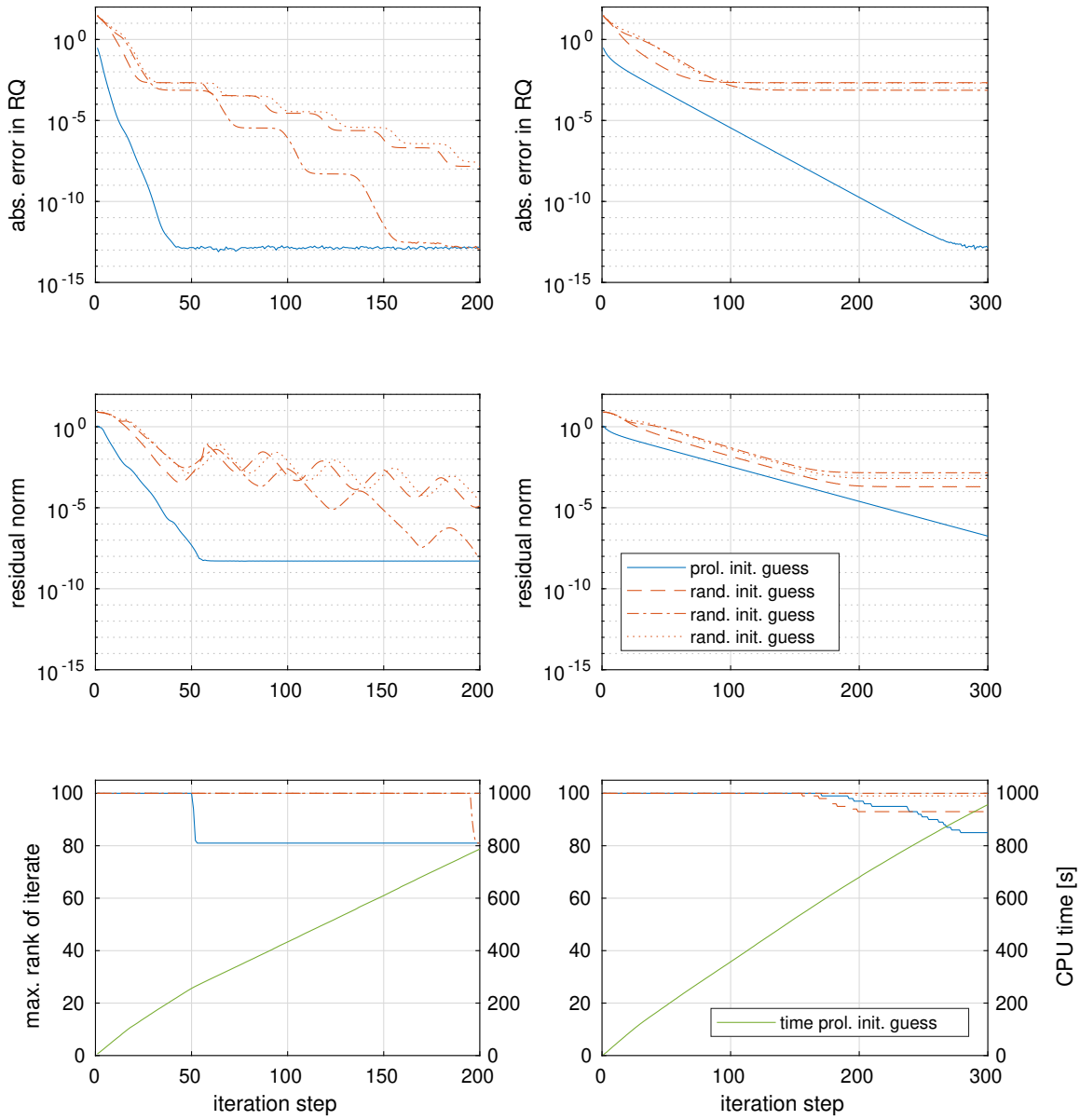


Figure 6.24.: 3-Potts, $d = 14$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (3, 4)$, linear tree, *left:* LOCG, *right:* gradient descent

from an eigenvalue, in fact between $\lambda_{\min}^{(d)}$ and $\lambda_2^{(d)}$, which is also illustrated by a relatively large residual norm. The consumed time for 200 iteration steps is slightly smaller for gradient descent, but due to the much less iteration steps to be performed until convergence, LOCG should be preferred.

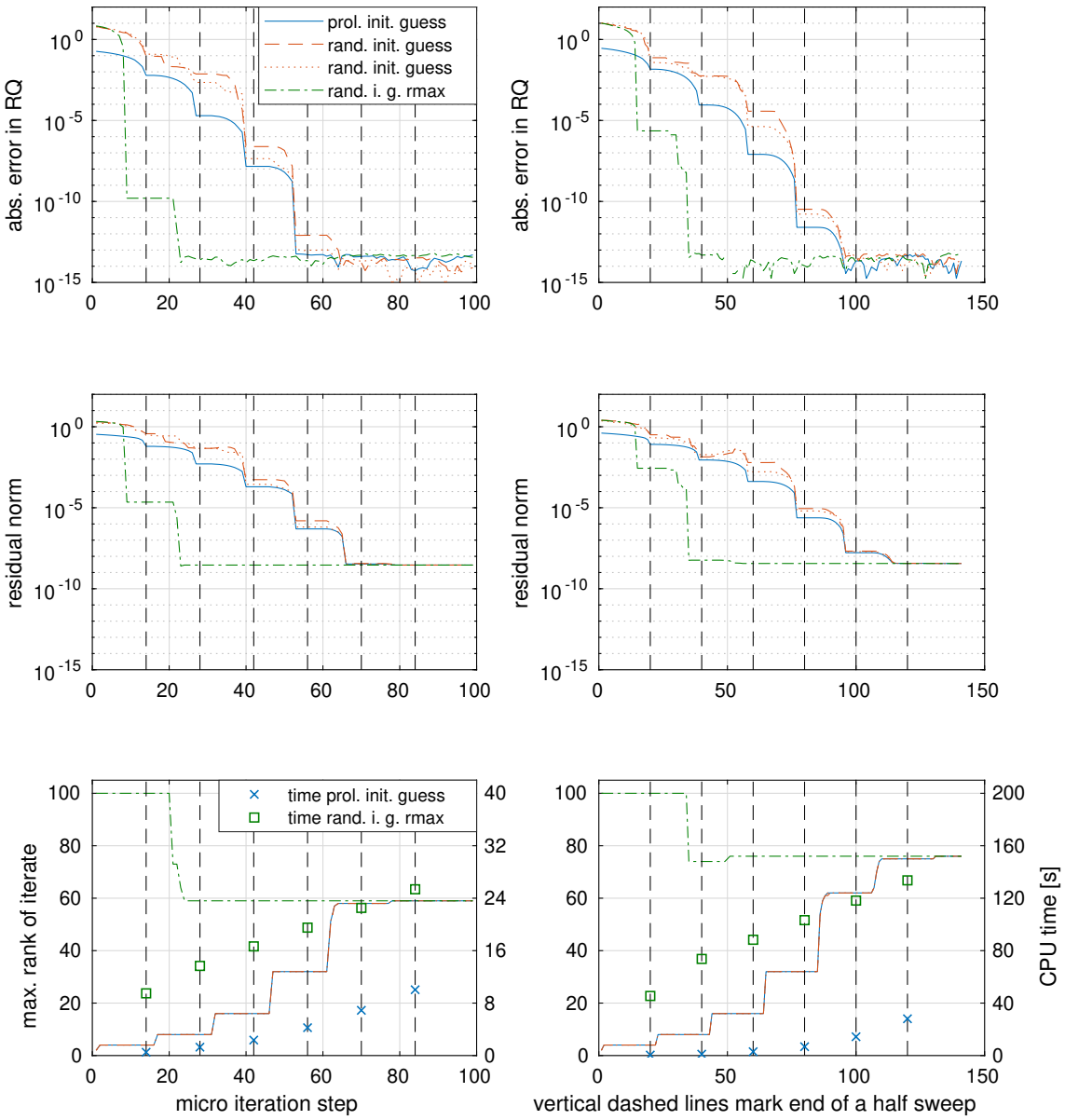


Figure 6.25.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 3)$, MALS, *left:* $d = 16$, *right:* $d = 22$

6.3. TT format, $A \neq B$

In this section we repeat many but not all tests from Section 6.2, now with the MALS, Algorithm 4.3, in the TT format. The graphical presentation of the results is similar as for LOCG with the difference that the x-axis indicates the micro iteration steps, hence the single steps in the for-loop from Lines 9 to 15 respectively Lines 19 to 25 in Algorithm 4.3. After $d-2$ micro iteration steps, one half sweep is completed which is marked by the vertical dashed lines in the plots. As the comparison concerning the shape of the dimension tree and

the utilized numerical method, performed at some places in Section 6.2, does not apply in the present setting, the left and right column in a single figure differ just by the problem size d, with one exception in Figure 6.31 where additionally (d_1, d_2) differs.

Following the implementation in [ODK+12], the eigenvalue problem in each micro iteration step, cf. Lines 11 and 21 in Algorithm 4.3, is solved by an iterative method, namely again LOCG, if $r_{j-1} n_j n_{j+1} r_{j+1} \geq 50$, which holds already for $n_j = n_{j+1} = q = 2$ and $r_{j-1}, r_{j+1} \geq 4$. We notice that in [OD12, Sect. 4.1] it is recommended to solve the micro step eigenvalue problem up to a size of around $r_{j-1} n_j n_{j+1} r_{j+1} = 1000$ exactly, hence via `eig` in MATLAB, but in practice we did not observe any advantage, neither in execution time nor in the amount of iteration steps. Due to [HRS12, Sect. 3.2], as an initial guess for this inner LOCG method there is chosen

$$\operatorname{vec}\left((\mathbf{\Sigma}^{(j-1)})_{1:r_k^{(j-1)},1:r_k^{(j-1)}} \left((\mathbf{Z}^{(j-1)})_{:,1:r_k^{(j-1)}} \right)^\top \operatorname{resh}_{r_j \times n_{j+1} r_{j+1}}(\mathbf{V}_k^{(j+1)}) \right)$$

in case of a left-to-right half sweep, with $\mathbf{\Sigma}^{(j-1)}$ and $\mathbf{Z}^{(j-1)}$ just from the previous micro iteration step at sites $(j-1, j)$, cf. Line 12 of Algorithm 4.3, and the core tensor $\mathbf{V}_k^{(j+1)}$ updated in the previous half sweep. For a right-to-left half sweep, this applies analogously with $\mathbf{Y}^{(j)}$ and $\mathbf{\Sigma}^{(j)}$ from Line 22 of Algorithm 4.3 and $\mathbf{V}_k^{(j)}$. The size of the eigenvalue problem in the micro iteration step equals $r_{j-1} n_j n_{j+1} r_{j+1} \times r_{j-1} n_j n_{j+1} r_{j+1}$, so it scales quadratically with $n_j = n_{j+1} = q$ and the current TT ranks r. Especially if $q = 3$ and $r \geq 81$, which is in the range of ranks detected in Section 6.1, it is $q^2 r^2 \geq 3^{10}$, which is in turn the size of the overall eigenvalue problem for $d = 10$. Due to the matrix-free implementation of LOCG in [ODK+12], this comparatively large size causes no technical problems. And as our main concern is the comparison of different initial guesses rather than the determination of optimal ranks, we do not attach too much importance to this otherwise somewhat questionable fact that the "micro" problem might be larger or at least is of the same order of magnitude as the overall problem. For $d = 14$ this inconsistency does not occur anyway.

We start in Figures 6.25 to 6.27 with a 2-XYZ model for $A = 1.9$, $B = 0.4$, $\Delta = -1.1$, $h = 0.2$ and focus at first on the solid blue resp. dashed or dotted red lines which represent the prolongated resp. random initial guess with TT ranks equal to the prolongated one like in Section 6.2. In case of $(d_1, d_2) = (2, 3)$ in Figure 6.25, although there is a small gap between the error in the Rayleigh quotient or the residual norm for the prolongated and the random initial guess during the iteration, we reach the final level of convergence of these values after the same number of micro iteration steps. We observe further that the maximal TT rank of the iterates is doubled in each of the first few half sweeps and that the jump of the blue and red graphs during one half sweep occurs the later the larger the maximal rank already is. Additionally, the significant decrease of RQ error and residual norm during one half sweep occurs in each of the successive half sweeps a bit earlier. When the maximal rank stops to increase, the residual norm stops to decrease, and this occurs one half sweep after the RQ error has reached its final value, again simultaneously for prolongated and random initial guess. Hence the stagnation of the residual norm or of the maximal TT rank, provided $r_{\max}$ was chosen sufficiently large, may be used as a stopping criterion for the method. It seems that the convergence behavior of RQ error and residual norm does not depend on the particular initial guess, but is rather related to the current value of the ranks of the iterate. The doubling of the ranks occurs since, cf. Lines 10-13 or 20-23 in Algorithm 4.3, the micro iteration step eigenvalue problem leads to a truncated SVD of a matrix of size

$r_{j-1}n_j \times n_{j+1}r_{j+1}$ with $n_j = n_{j+1} = q = 2$, and if actually no truncation takes place when all singular values are above the truncation threshold, r_j is updated by $\min\{r_{j-1}n_j, n_{j+1}r_{j+1}\}$ which at least for some j equals $n_j r_j = 2r_j$. The described shift towards the middle of the whole half sweep of the specific micro iteration step when the maximal rank increases respectively the error quantities significantly decrease may be explained by the fact that also the theoretical upper bounds of the value of the single TT ranks r_j of an iterate grow towards the middle of the TT representative, namely

$$r_j \leq q^{\min\{j,d-j\}}, \quad 0 \leq j \leq d. \tag{6.6}$$

We also refer to the corresponding discussion for HT ranks at the end of Subsection 3.3.2. So, $\max\{r_j : 0 \leq j \leq d\}$ can jump from q^j to q^{j+1} at the earliest in the $(j+1)$-th micro iteration step of a half sweep and at the latest in the $\big(d-(j+1)\big)$-th micro iteration step. Having said this, we conclude that the RQ error and the residual norm decrease most at that particular time during a half sweep when all TT ranks allowed to be increased have been updated.

In order to examine the relation of increase of rank and decrease of error further, we additionally consider a random initial guess which is set up with TT ranks equal to

$$(r_1, r_2, \ldots, r_{d-2}, r_{d-1}) = (2, 4, \ldots, 2^{\lfloor \log_2(r_{\max}) \rfloor}, r_{\max}, \ldots, r_{\max}, 2^{\lfloor \log_2(r_{\max}) \rfloor}, \ldots, 4, 2)$$

with the parameter $r_{\max}$ originally introduced and discussed in Section 6.1 as the upper bound for the TT ranks in the truncation. The ranks smaller than $r_{\max}$ follow the exponentially growing theoretical upper bound (6.6). We choose, as before, $r_{\max} = 100$ and depict the convergence behavior of this "$r_{\max}$-random initial guess" by the dash-dotted green line. To avoid confusion, we emphasize the difference between the global parameter $r_{\max}$ and the actual maximal value of the single TT ranks of an iterate in the course of the iteration visualized in the plots in the third row. We observe that during the first half sweep, the RQ error and the residual norm already decrease rapidly, and during the second half sweep they reach their final level, which equals the corresponding level for the other two types of initial guess. At the same time, the maximal rank of the iterate decreases to its final value, which also equals that value reached when employing the prolongated or random initial guess sharing the same much smaller ranks at the beginning of the iteration. We notice that the final value of the ranks is consistent with the result for the linear HT format in Figure 6.11.

As another step in the discussion of the $r_{\max}$-random initial guess, we compare the CPU time consumed by the MALS when started with this particular random initial guess or the prolongated initial guess. We depict the respective values in the plots in the third row of Figure 6.25 with a marker at the end of each half sweep and indicate them by the y-axis at the right edge of the plots. Due to the much larger TT ranks at the beginning of the iteration, the two half sweeps with the $r_{\max}$-random initial guess necessary for convergence take two to three times as much total CPU time than the five ($d = 16$) or six ($d = 22$) half sweeps with the prolongated initial guess. One may argue the particular value of $r_{\max} = 100$ was quite arbitrarily chosen and is unnecessarily large, but in general the correct value of the TT ranks of the tensorization of the eigenvector is unknown a priori, so it is not feasible to expect being able to set up an initial guess with ranks very close to the correct ranks.

Moreover, we observe that the CPU time consumed by the five or six half sweeps of MALS until convergence with the prolongated initial guess amounts for both $d \in \{16, 22\}$ only to about five percent of the CPU time used by the approximately 50 or 75 iteration steps of LOCG for the equivalent setting recorded by Figure 6.11. With regard to the comparability

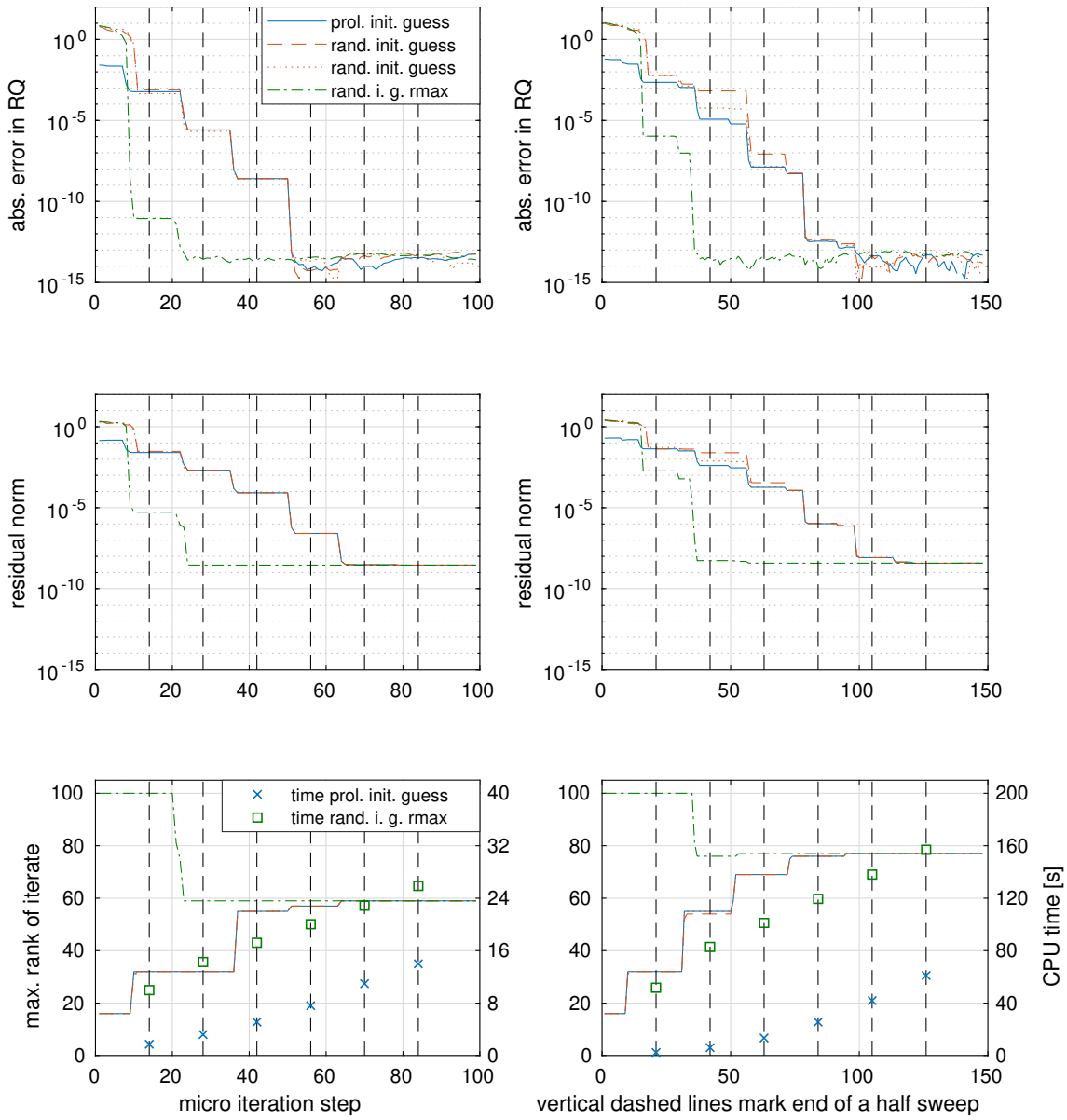


Figure 6.26.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 9)$, MALS, *left:* $d = 16$, *right:* $d = 23$

of the time consumptions and possibilities for improvement, we notice that the respective implementations of the algorithms rely on two different toolboxes [KT14] and [ODK+12].

The next test case, Figure 6.26, with $(d_1, d_2) = (2, 9)$ shows similar behavior. The ranks increase in most of the half sweeps, however a precise doubling is only observable during the first half sweep. The maximal rank of the prolongated initial guess is by construction larger than for $(d_1, d_2) = (2, 3)$, see Remark 5.5, and reaches the level necessary for convergence of RQ error and residual norm after the same number of half sweeps. We notice that for $d = 16$, in the second half sweep no increase of the maximal rank occurs. Again, the $r_{\max}$-

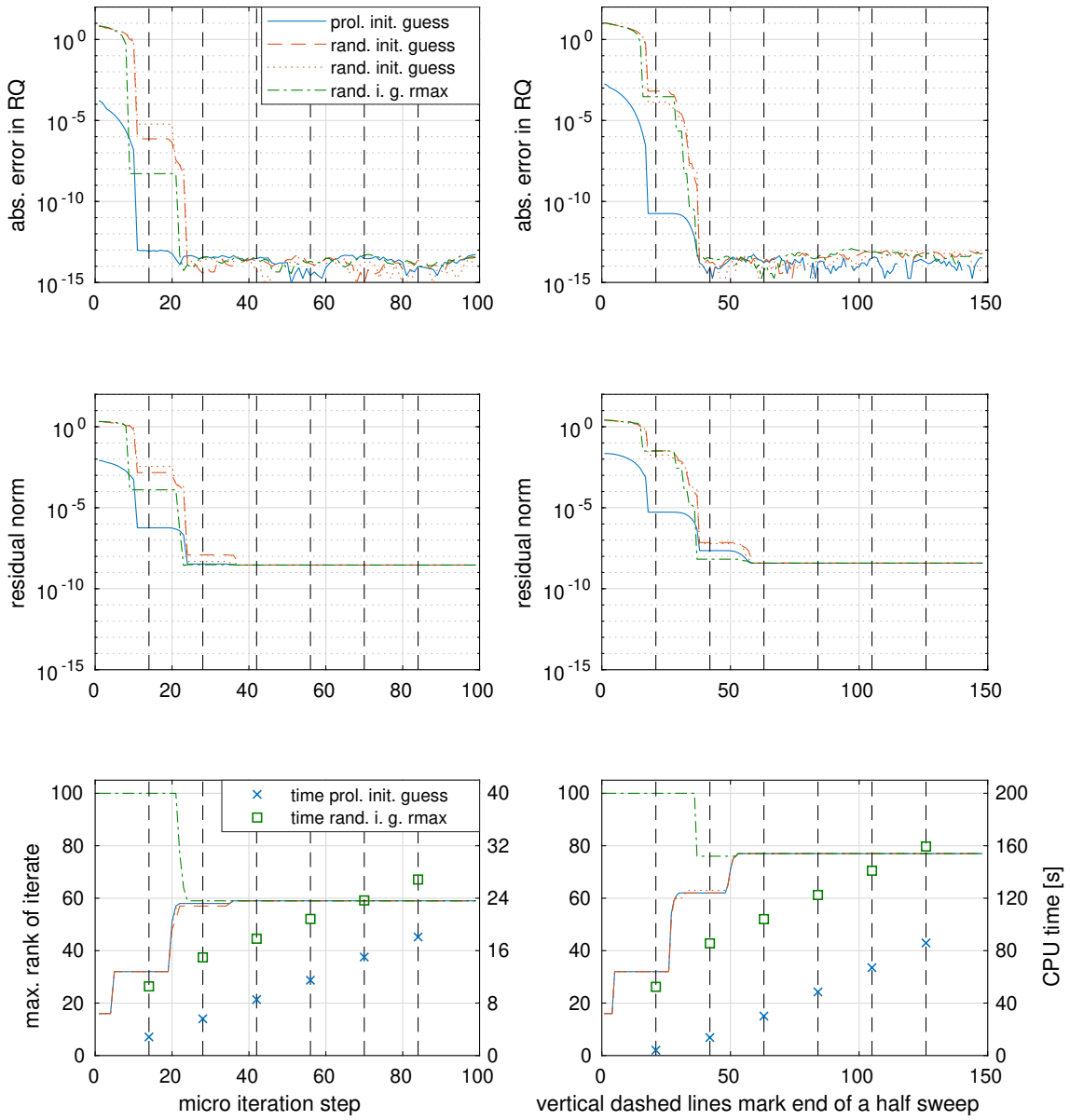


Figure 6.27.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (8, 9)$, MALS, *left:* $d = 16$, *right:* $d = 23$

random initial guess yields convergence after two half sweeps, but the total CPU time until convergence is larger than for the prolongated initial guess.

For $(d_1, d_2) = (8, 9)$, Figure 6.27, the prolongated initial guess now leads to convergence itself only after two half sweeps, and the RQ error and the residual norm are during the iteration smaller than for the two types of random initial guess. The maximal rank of the iterates start, like in Figure 6.26, at a value of 16, but especially in the second half sweep it is almost doubled and hence the final value is reached rather quickly.

In Figure 6.28 we repeat the test with a 2-XYZ model for $A = -0.9$, $B = 1.6$, $\Delta = 1.7$, $h =$

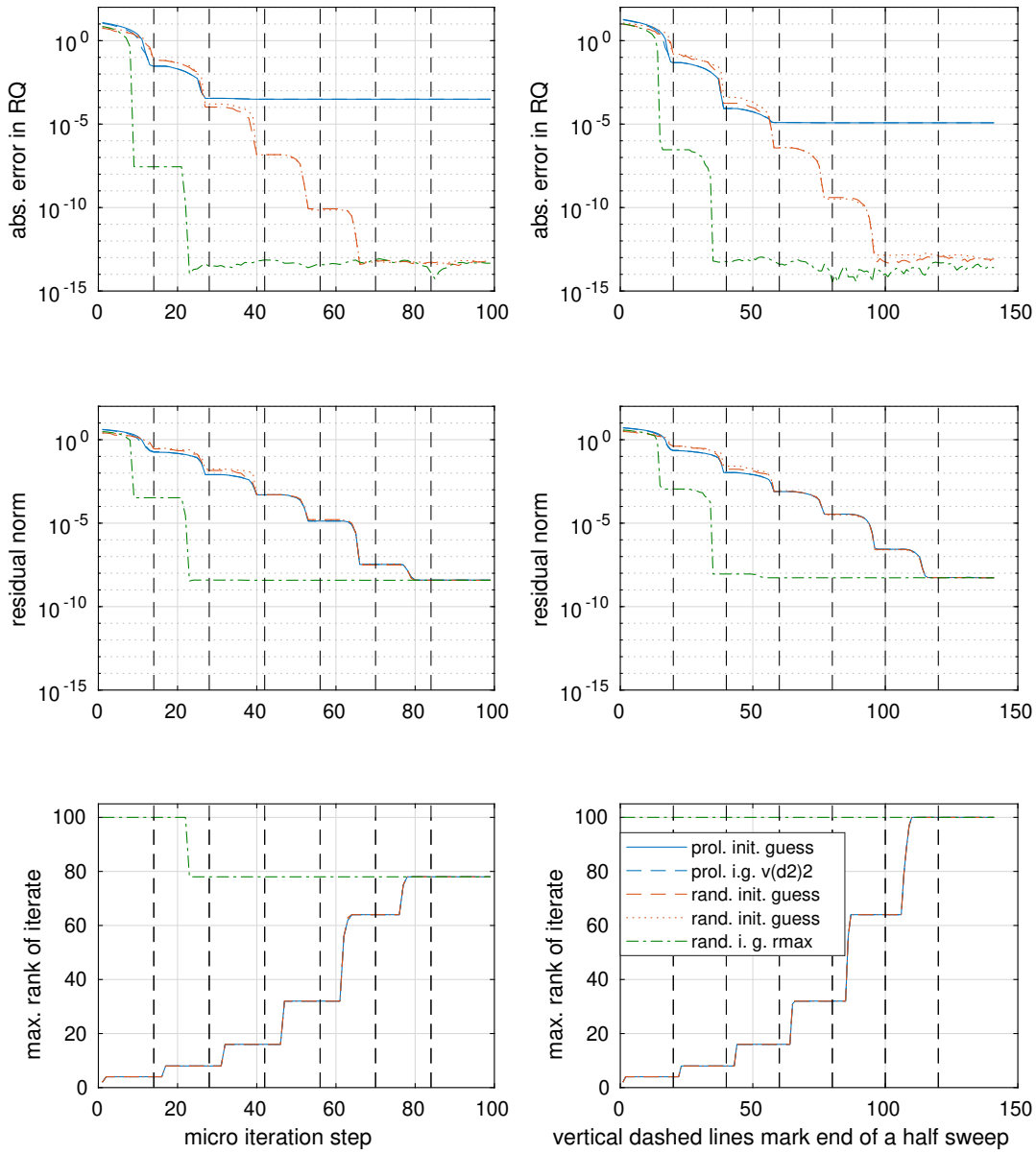


Figure 6.28.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 3)$, MALS, *left:* $d = 16$, *right:* $d = 22$

1.2 and $(d_1, d_2) = (2, 3)$ which was characterized by the fact that for both $d \in \{16, 22\}$, the prolongated initial guess $\widetilde{\mathbf{v}}^{(d)}$ has not the same sparsity pattern as $\mathbf{v}_{\min}^{(d)}$. As opposed to Figure 6.16, the alternative initial guess $\widetilde{\mathbf{v}}_2^{(d)}$ which is constructed by prolongating $\mathbf{v}_2^{(d_2)}$ to problem size d, see (6.5), also yields convergence measured in RQ to $\lambda_2^{(d)}$. For the two types of random initial guess, the iterates converge to $\lambda_{\min}^{(d)}$ with a similar behavior like in the test visualized by Figure 6.25 which was also based on $(d_1, d_2) = (2, 3)$ but for different A, B, Δ, h.

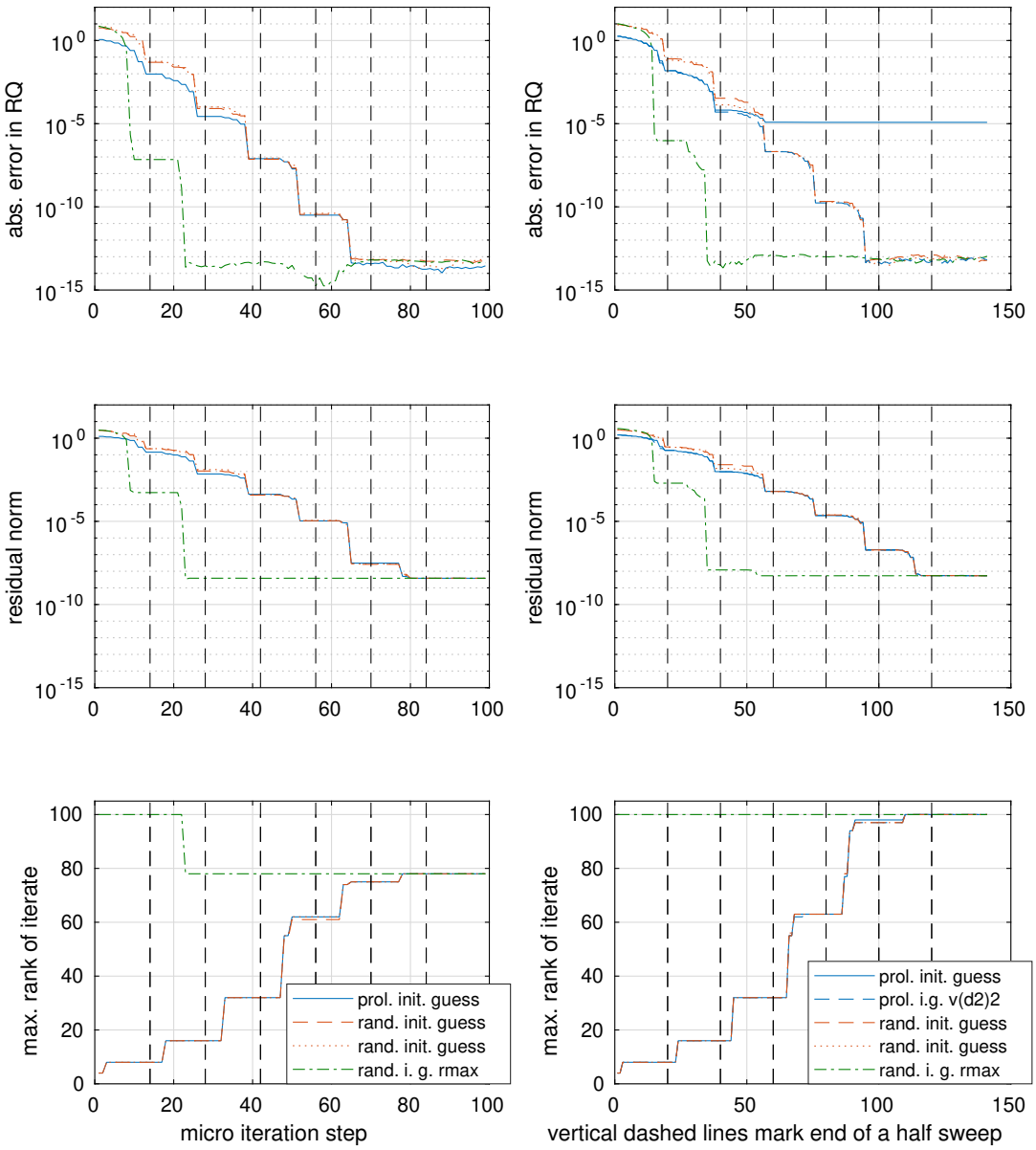


Figure 6.29.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 4)$, MALS, *left:* $d = 16$, *right:* $d = 22$

If we choose instead $(d_1, d_2) = (2, 4)$, in Figure 6.29, then for $d = 16$ and with the prolongated initial guess $\widetilde{\mathbf{v}}^{(d)}$ which has the correct sparsity pattern anyway, the iterates converge, just as for the random initial guess, in six half sweeps to $\lambda_{\min}^{(d)}$. Taking the $r_{\max}$-random initial guess, we need again only two half sweeps. In the more interesting case $d = 22$ we observe, in contrast to Figure 6.28, that the alternative prolongated initial guess $\widetilde{\mathbf{v}}_2^{(d)}$ set up with the same sparsity pattern as $\mathbf{v}_{\min}^{(d)}$ and depicted by the dashed blue line, leads to convergence in RQ to $\lambda_{\min}^{(d)}$, while $\widetilde{\mathbf{v}}^{(d)}$ with sparsity pattern equal to $\mathbf{v}_2^{(d)}$ results in

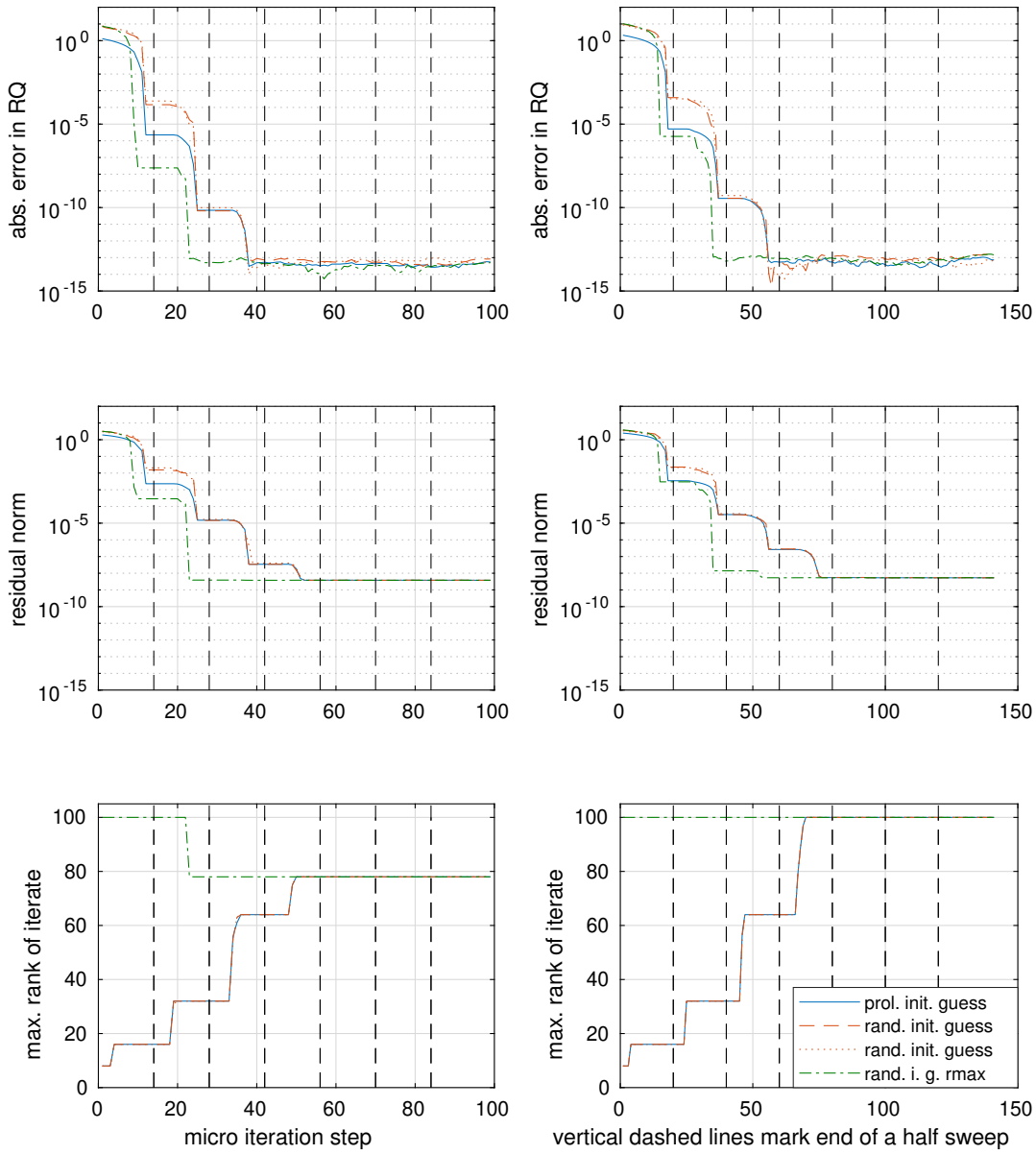


Figure 6.30.: 2-XYZ, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (3, 4)$, MALS, *left:* $d = 16$, *right:* $d = 22$

convergence to $\lambda_2^{(d)}$.

The last test for the current parameters A, B, Δ, h is with $(d_1, d_2) = (3, 4)$ in Figure 6.30. Here, as doubling of the maximal rank, equal to 8 at the beginning, occurs in each half sweep, we need four half sweeps until convergence of both RQ error and residual norm, while we need again two half sweeps with the $r_{\max}$-random initial guess.

In Figure 6.31 we consider a 3-XYZ model for $A = 1.6$, $B = -0.3$, $\Delta = 2.9$, $h = -0.8$.

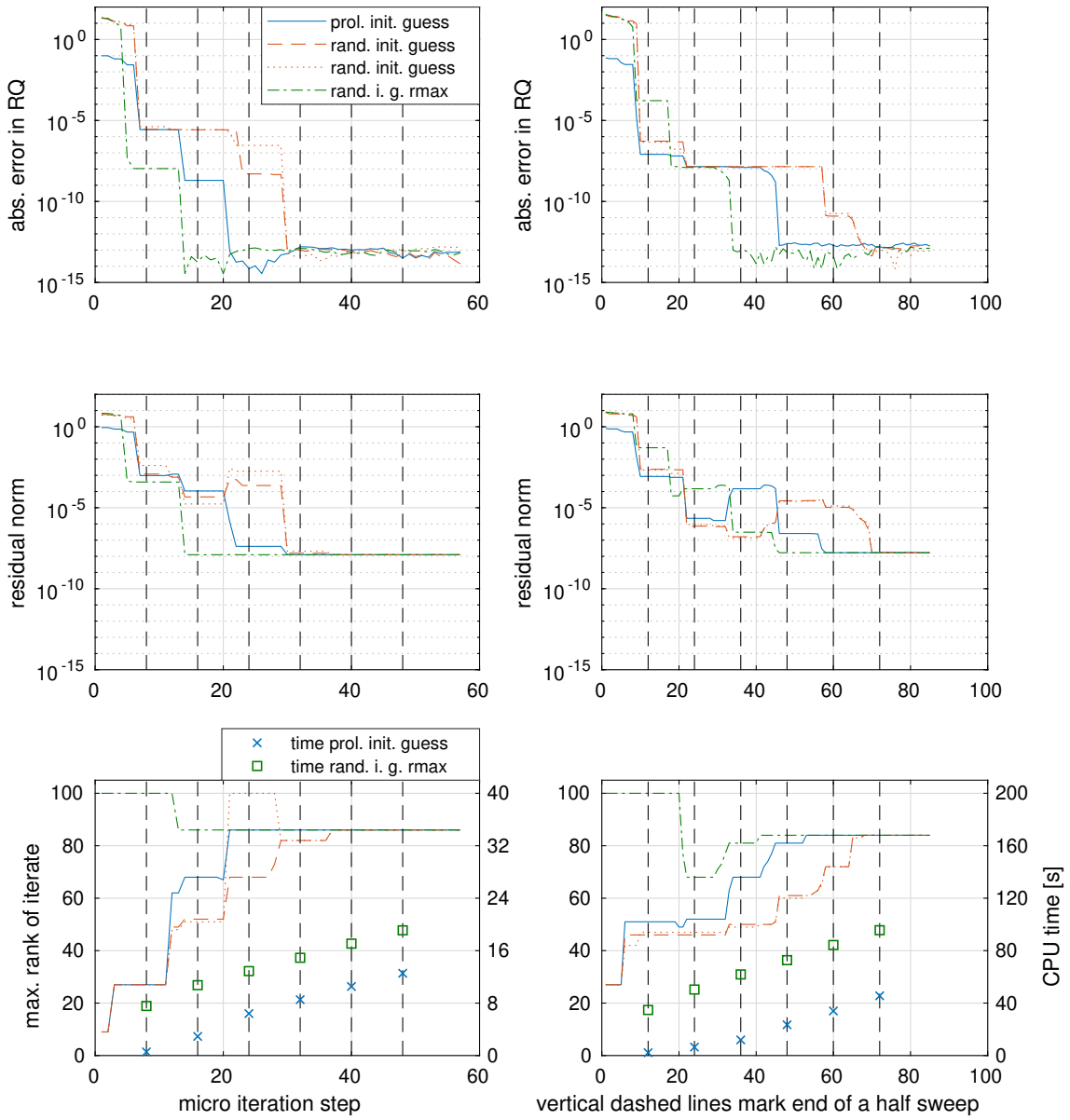


Figure 6.31.: 3-XYZ, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, MALS,
left: $d = 10$, $(d_1, d_2) = (2, 4)$, *right:* $d = 14$, $(d_1, d_2) = (2, 6)$

Due to $q = 3$, the $r_{\max}$-random initial guess is set up with TT ranks

$$(r_1, r_2, \ldots, r_{d-2}, r_{d-1}) = (3, 9, \ldots, 3^{\lfloor \log_3(r_{\max}) \rfloor}, r_{\max}, \ldots, r_{\max}, 3^{\lfloor \log_3(r_{\max}) \rfloor}, \ldots, 9, 3)$$

consistently with (6.6). In the left column, representing $d = 10$ and $(d_1, d_2) = (2, 4)$, with the prolongated initial guess essentially one half sweep more is necessary than for the $r_{\max}$-random initial guess, and one half sweep less than for the random initial guess set up with the same ranks as the prolongated one. The CPU time until convergence is for the $r_{\max}$-random initial guess about twice as much as for the prolongated initial guess. A similar behavior is

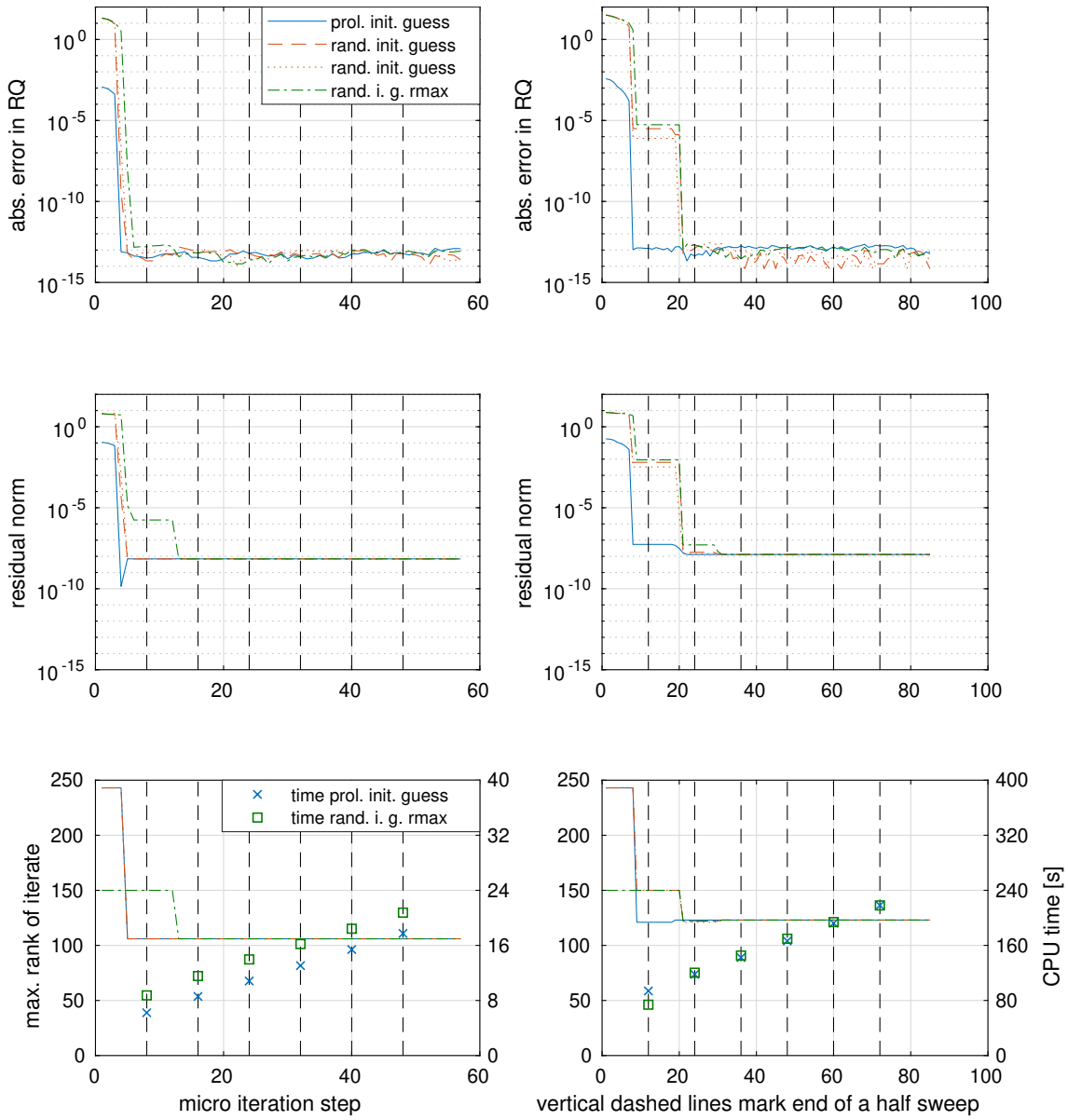


Figure 6.32.: 3-XYZ, $(A, B, \Delta, h) = (2.6, 0.7, -1.9, -0.3)$, $(d_1, d_2) = (5, 6)$, MALS, *left:* $d = 10$, *right:* $d = 14$

given in case of $d = 14$, $(d_1, d_2) = (2, 6)$ depicted in the right column of Figure 6.31 with the specific observation that the RQ error stagnates for the duration of two to three half sweeps at a level equal to $\lambda_2^{(d)} - \lambda_{\min}^{(d)}$ before decreasing near machine epsilon.

In the next test, Figure 6.32, for a 3-XYZ model with $A = 2.6$, $B = 0.7$, $\Delta = -1.9$, $h = -0.3$, and $(d_1, d_2) = (5, 6)$, each of the initial guesses leads to convergence already after at most two half sweeps, the prolongated initial guess even for both $d \in \{10, 14\}$ after only one half sweep. We attribute this to the large maximal rank equal to 243, see Remark 5.5, of the prolongated initial guess at the beginning of the iteration which is in fact larger than

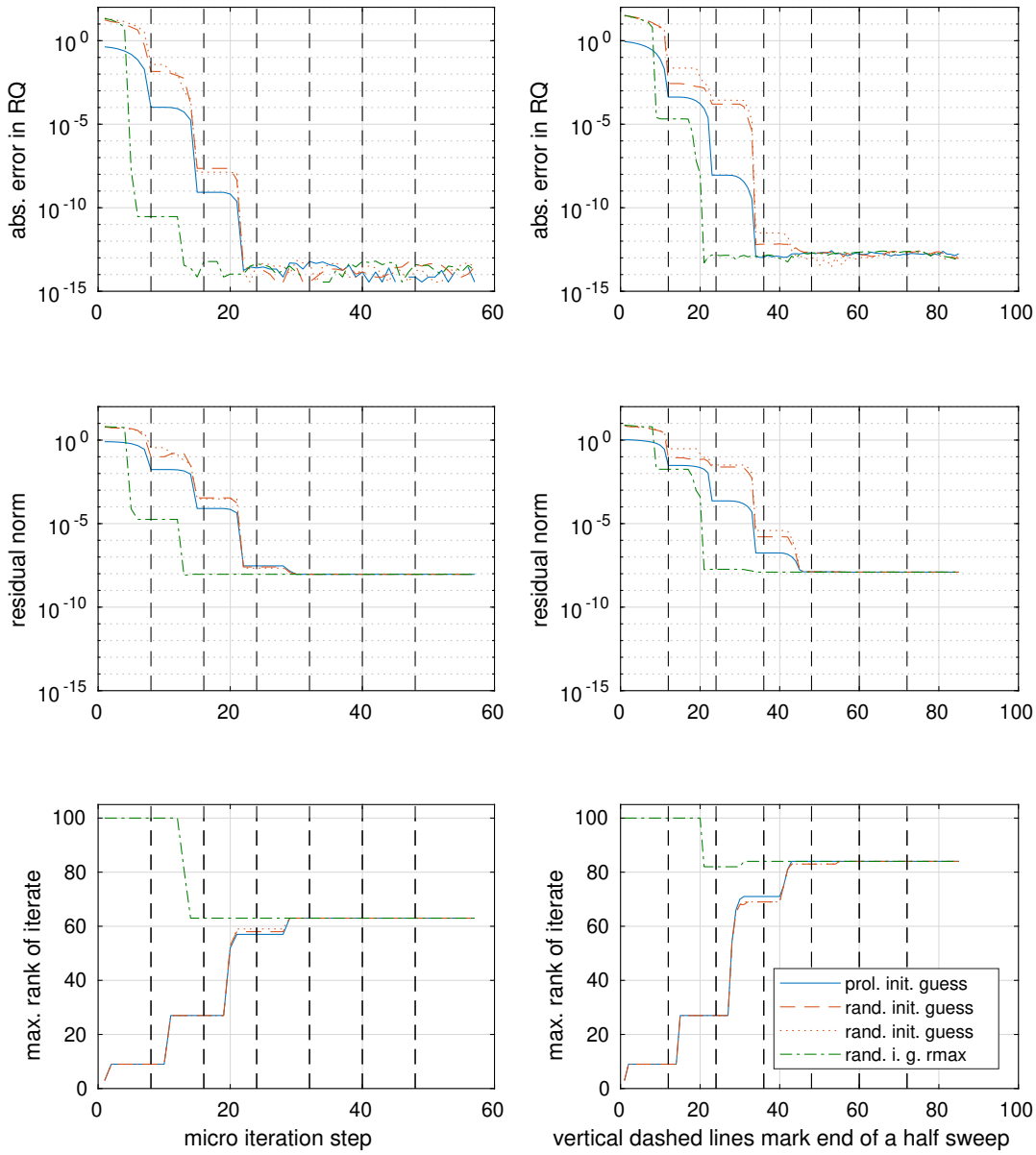


Figure 6.33.: 3-Potts, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (2, 3)$, MALS, *left:* $d = 10$, *right:* $d = 14$

the final rank and also larger than the value of $r_{\max} = 150$. Despite that large rank of the prolongated initial guess, the CPU time in case $d = 10$ for the first half sweep is a bit smaller than for the $r_{\max}$-random initial guess since the maximal value of the ranks of the current iterate drops down to the final level slightly above 100 already during the first half sweep while the maximal rank in case of the $r_{\max}$-random initial guess remains equal to 150 until the middle of the second half sweep. In case of $d = 14$, one half sweep necessary to converge to $\lambda_{\min}^{(d)}$ when employing the prolongated initial guess requires less time than two half sweeps necessary with the random initial guess set up with maximal rank 150.

The tests for a 3-Potts model with $A = 1.3$, $h = 0.9$, whose results are contained in

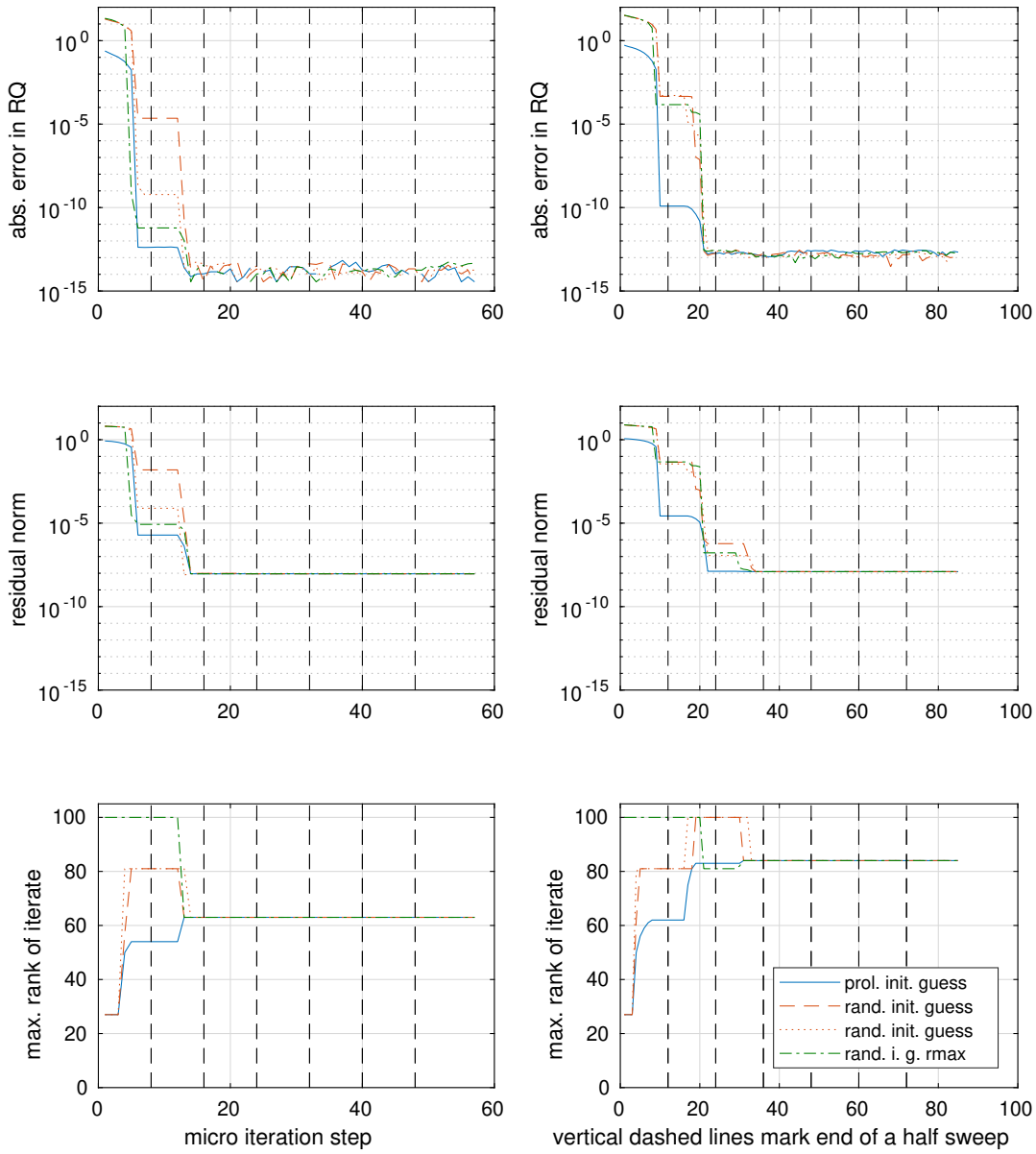


Figure 6.34.: 3-Potts, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (3, 4)$, MALS, *left:* $d = 10$, *right:* $d = 14$

Figures 6.33 and 6.34, are going along with the findings so far. The prolongated initial guess performs at least as well as the random initial guess set up with the same TT ranks. The iteration started with the $r_{\max}$-random initial guess needs in each scenario two half sweeps to converge. Again, the larger ranks of the prolongated initial guess in case $(d_1, d_2) = (3, 4)$ lead to a faster convergence compared to $(d_1, d_2) = (2, 3)$.

We conclude this section with two tests for values of d much larger than considered so far, namely $d \in \{100, 300\}$. For these d it is no longer possible to determine an "exact" value of $\lambda_{\min}^{(d)}$ via the MATLAB function `eigs` applied to a sparse matrix. So the reference value for $\lambda_{\min}^{(d)}$ is now the Rayleigh quotient of the last iterate in the iteration with prolongated

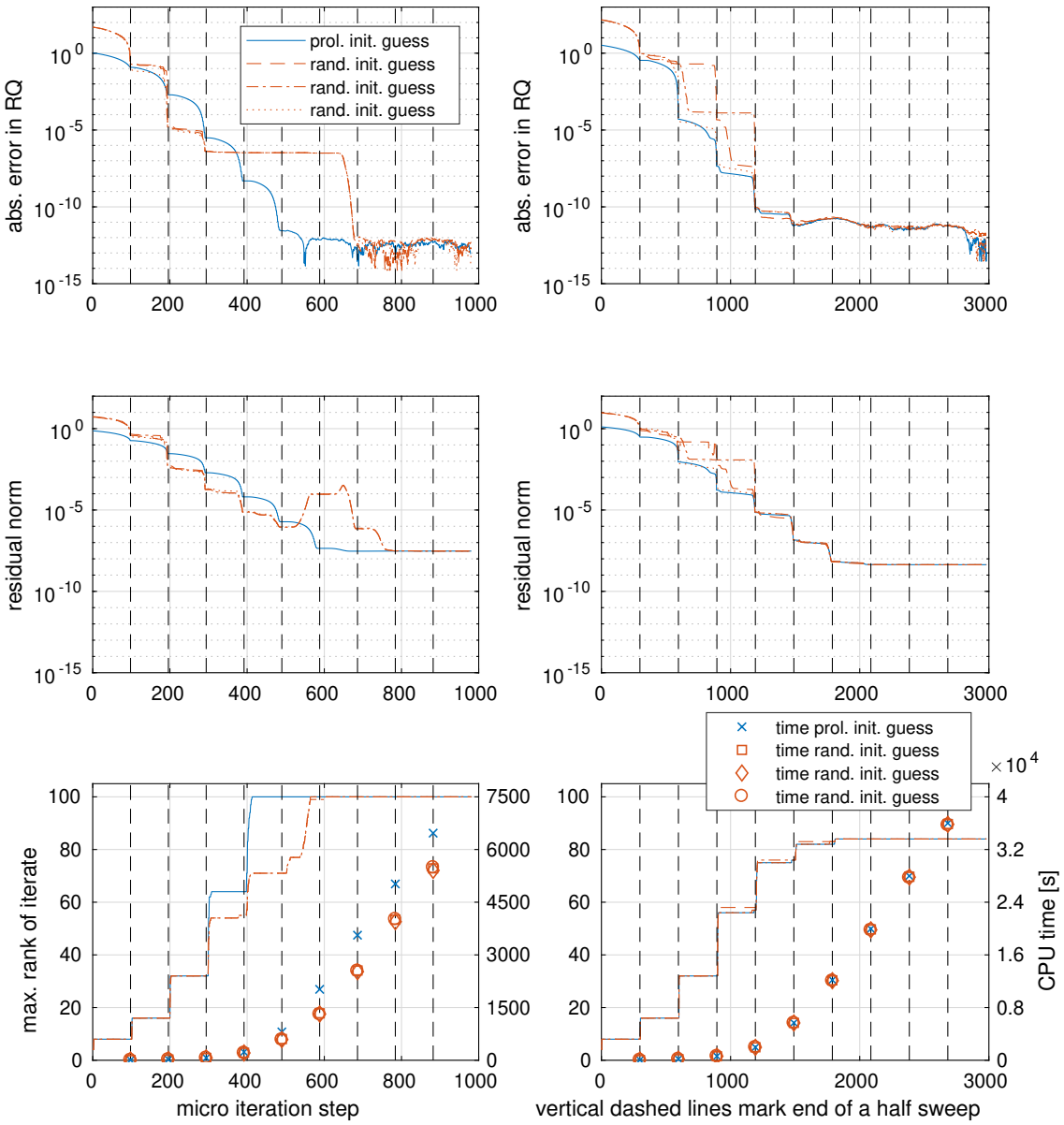


Figure 6.35.: 2-XYZ, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, MALS, *left:* $d = 100$, *right:* $d = 300$

initial guess with the additional requirement that the residual norm should be small for the duration of some half sweeps. The first test, Figure 6.35, is the already considered 2-XYZ model with $A = 1.9$, $B = 0.4$, $\Delta = -1.1$, $h = 0.2$, now for $(d_1, d_2) = (2, 4)$. While for $d = 300$ the prolongated initial guess yields only a slightly better performance during the iteration, in case $d = 100$ all three instances of the random initial guess stagnate for three half sweeps concerning the error in the Rayleigh quotient at a level of 10^{-6} before decreasing in the next half sweep to approximately machine epsilon. As the maximal rank grows faster for the prolongated initial guess in case $d = 100$, also the CPU time is larger, beginning

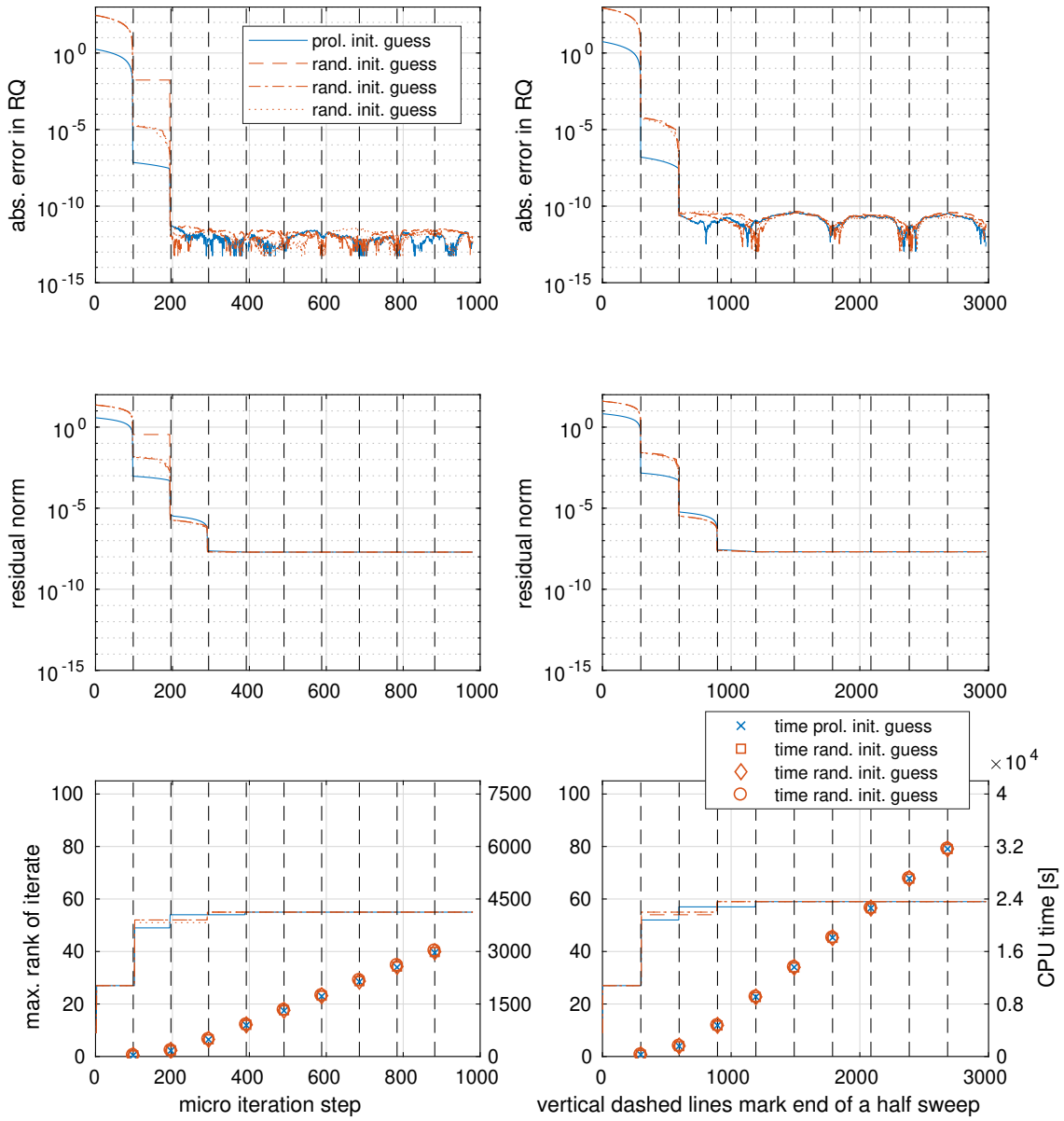


Figure 6.36.: 3-XYZ, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 4)$, MALS, *left:* $d = 100$, *right:* $d = 300$

with the fifth half sweep. But, since one half sweep less is necessary in this scenario, the consumed time until convergence is smaller than for the random initial guess. The second test, Figure 6.36, deals with the 3-XYZ model for $A = 1.6$, $B = -0.3$, $\Delta = 2.9$, $h = -0.8$, again with $(d_1, d_2) = (2, 4)$. Here, for both $d \in \{100, 300\}$, convergence in RQ to the reference eigenvalue is reached after two half sweeps for both types of initial guess simultaneously and the residual norm attains its final value around 10^{-8} after one more half sweep and stays there for remainder of the iteration. The CPU times for prolongated and random initial guess are practically equal.

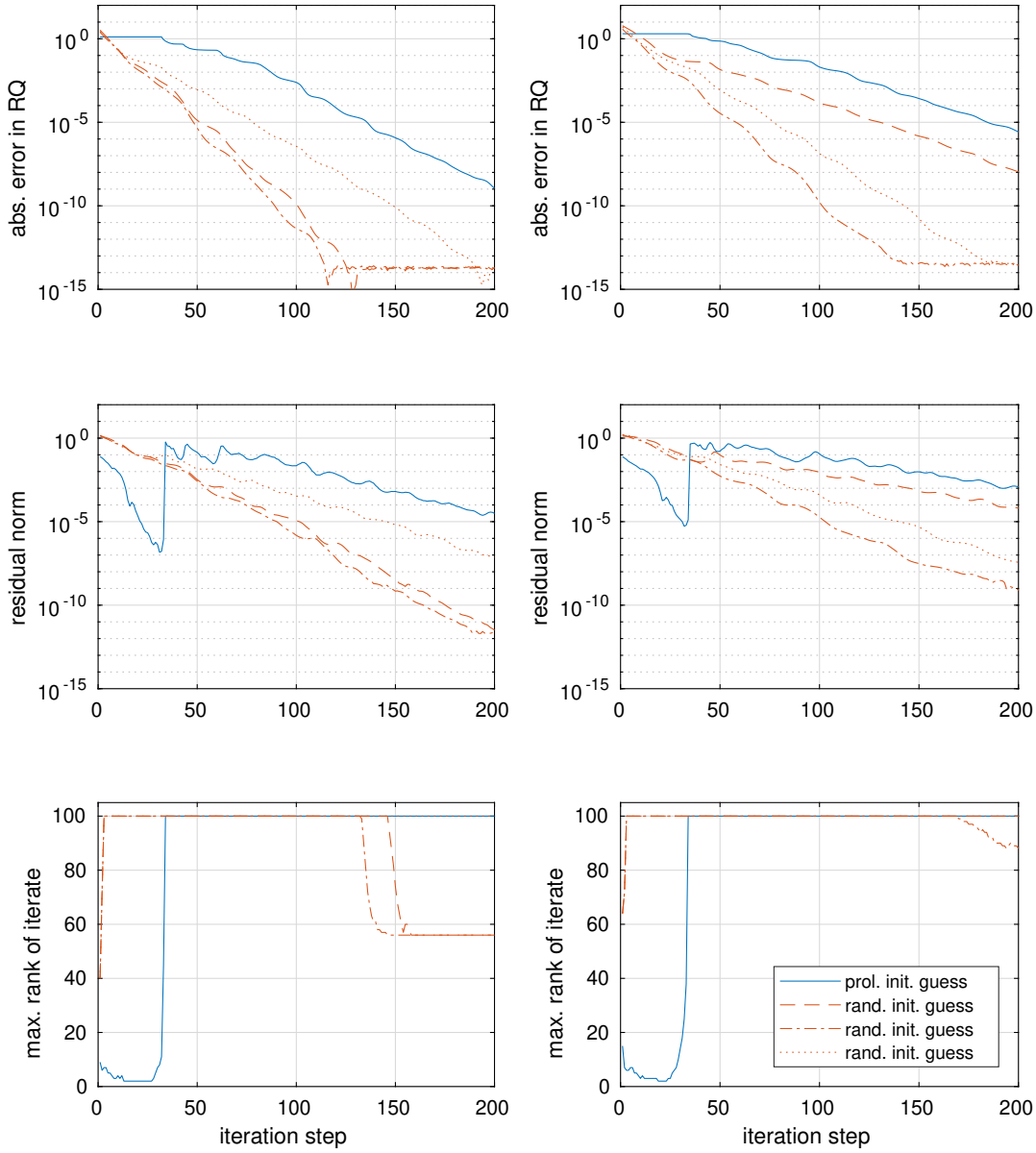


Figure 6.37.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (3, 4)$, LOCG, balanced tree, *left:* $d = 16$, *right:* $d = 22$

6.4. HT format, $A = B$

Until now we performed various tests for different types of Hamilton operators, but the common feature of all tests was that in case of an XYZ model, the coupling parameters satisfied $A \neq B$, so we were able to apply the prolongation strategy discussed in Sections 5.1-5.4 to construct an initial guess. In contrast, the present and the next section is devoted to the case $A = B$ where this prolongation strategy is not directly feasible, since in general the quadratic matricization of $\mathbf{v}_{\min}^{(d_1)}$ or blocks thereof are not invertible as they contain too

many zeros. A possible adaption of the prolongation strategy to this non-invertible scenario is to replace the classical inverse by the pseudoinverse in the construction of the matrix $\mathbf{N}$ governing the prolongation. To test this, we consider a 2-XYZ model with $A = B = 1$, $\Delta = -0.2$, $h = 0.3$. Remembering (5.12) and (5.13) for a first test with $(d_1, d_2) = (3, 4)$ and the definition of the truncated pseudoinverse $\mathrm{pinv}(\mathbf{A}; \varepsilon)$ around (5.18), we set

$$\mathbf{M} := \begin{pmatrix} \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{I})} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \otimes \mathbf{N}^{(\mathrm{II})} \end{pmatrix},$$

$$\mathbf{N}^{(\mathrm{I})} := \mathrm{mat}_{4\times 2}\big((\mathbf{v}_{\min}^{(4)})_{1:8}\big) \cdot \mathrm{pinv}\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{1:4}\big); 10^{-6}\right),$$
$$\mathbf{N}^{(\mathrm{II})} := \mathrm{mat}_{4\times 2}\big((\mathbf{v}_{\min}^{(4)})_{9:16}\big) \cdot \mathrm{pinv}\left(\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{5:8}\big); 10^{-6}\right),$$

and

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{d-4} (\mathbf{I}_{2^i} \otimes \mathbf{M})\right) \mathbf{v}_{\min}^{(4)}.$$

In the present case it is $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,2}^{(2)}$, thus

$$\mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{1:4}\big) = \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix}, \quad \mathrm{mat}_{2\times 2}\big((\mathbf{v}_{\min}^{(3)})_{5:8}\big) = \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix}.$$

Concerning the construction of such an initial guess in HT format, we notice that some of the σ from (5.40) as norms of submatrices of $\mathbf{N}^{(\mathrm{I})}$ or $\mathbf{N}^{(\mathrm{II})}$ may be zero, hence we set $\sigma_1^{(\mathrm{I})} = \sigma_2^{(\mathrm{I})} = \sigma_1^{(\mathrm{II})} = \sigma_2^{(\mathrm{II})} := 1$ and

$$\mathbf{u}_1^{(\mathrm{I})} := \begin{pmatrix} n_{1,1}^{(\mathrm{I})} & n_{1,2}^{(\mathrm{I})} \\ n_{2,1}^{(\mathrm{I})} & n_{2,2}^{(\mathrm{I})} \end{pmatrix}, \quad \mathbf{u}_2^{(\mathrm{I})} := \begin{pmatrix} n_{3,1}^{(\mathrm{I})} & n_{3,2}^{(\mathrm{I})} \\ n_{4,1}^{(\mathrm{I})} & n_{4,2}^{(\mathrm{I})} \end{pmatrix},$$
$$\mathbf{u}_1^{(\mathrm{II})} := \begin{pmatrix} n_{1,1}^{(\mathrm{II})} & n_{1,2}^{(\mathrm{II})} \\ n_{2,1}^{(\mathrm{II})} & n_{2,2}^{(\mathrm{II})} \end{pmatrix}, \quad \mathbf{u}_2^{(\mathrm{II})} := \begin{pmatrix} n_{3,1}^{(\mathrm{II})} & n_{3,2}^{(\mathrm{II})} \\ n_{4,1}^{(\mathrm{II})} & n_{4,2}^{(\mathrm{II})} \end{pmatrix}$$

when we actually construct the HT representative.

Figure 6.37 shows the performance of this prolongated initial guess for $d \in \{16, 22\}$ and a balanced dimension tree. We observe that the error in the Rayleigh quotient decreases more slowly than for all three instances of a random initial guess set up with the same ranks as the prolongated one. The residual norm decreases at the beginning for some iteration steps rather quickly but the RQ error stagnates, which indicates that in this stage the iterates approximate an eigenvector associated with an eigenvalue larger than $\lambda_{\min}^{(d)}$. A closer investigation in case $d = 16$ shows that up to iteration step $l = 30$, the modified RQ error

$$\left|\langle \mathbf{V}_l, \Phi_{\mathbf{H}}(\mathbf{V}_l)\rangle - \lambda_{343}^{(16)}\right|$$

with respect to the 343-th smallest, simple eigenvalue $\lambda_{343}^{(16)}$ decreases to the order of 10^{-14}. The choice $(d_1, d_2) = (3, 4)$ yields for the current $A = B = 1$, $\Delta = -0.2$, $h = 0.3$ that $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,2}^{(2)}$ and $\mathbf{v}_{\min}^{(4)} \in \mathcal{E}_{4,2}^{(3)}$, so $\widetilde{\mathbf{v}}^{(16)} \in \mathcal{E}_{16,2}^{(15)}$ by Theorem 5.3. In addition, $\lambda_{343}^{(16)}$ is the smallest eigenvalue such that an associated eigenvector $\mathbf{v}_{343}^{(16)}$ satisfies $\mathbf{v}_{343}^{(16)} \in \mathcal{E}_{16,2}^{(15)}$, hence

has the same sparsity pattern as the prolongated initial guess. Thus, up to around iteration step $l = 30$, the iterates remain to approximately represent the tensorization of a vector in $\mathcal{E}_{16,2}^{(15)}$. This also explains why the maximal HT rank of the iterates between iteration step 13 and 27 equals 2, reflecting the statement in exact arithmetics that $\mathrm{tens}_{2\times 16}(\mathbf{w})$ with $\mathbf{w} \in \mathcal{E}_{16,2}^{(15)}$ having $\binom{16}{15} = 16$ nonzero entries $w_1, \ldots, w_{16}$ may be represented for a balanced dimension tree by the leaf matrices $\mathbf{U}_{\{1\}} = \ldots = \mathbf{U}_{\{16\}} = \mathbf{I}_2$ and the transfer tensors

$$\mathbf{B}_{\{2i-1,2i\}} = \begin{pmatrix} 0 & w_{16+1-(2i-1)} \\ w_{16+1-2i} & 0 \end{pmatrix}, \quad 1 \le i \le 8,$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{B}_{\{4i-3,\ldots,4i\}} = \mathbf{B}_{\{1,\ldots,8\}} = \mathbf{B}_{\{9,\ldots,16\}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad 1 \le i \le 4,$$

$$\mathbf{B}_{\mathrm{root}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In the further course of the iteration, the usual RQ error from (6.3) measuring the distance to $\lambda_{\min}^{(16)}$ decreases for the prolongated initial guess more slowly than for the random initial guesses. Similar observations are made for $d = 22$, thus the prolongated initial guess constructed by the pseudoinverse seems to be unsuitable.

A second test, Figure 6.38, with the prolongated initial guess based on the pseudoinverse is carried out for $(d_1, d_2) = (8, 9)$ with linear dimension tree, so we set

$$\begin{aligned} \mathbf{M} &:= \mathbf{I}_{16} \otimes \left[\mathrm{mat}_{2^5\times 2^4}(\mathbf{v}_{\min}^{(9)}) \cdot \mathrm{pinv}\left(\mathrm{mat}_{2^4\times 2^4}(\mathbf{v}_{\min}^{(8)}); 10^{-6} \right) \right], \\ \widetilde{\mathbf{v}}^{(d)} &:= \left(\prod_{i=1}^{d-9} (\mathbf{I}_{2^i} \otimes \mathbf{M}) \right) \mathbf{v}_{\min}^{(9)}, \end{aligned} \tag{6.7}$$

noticing for the present setting $\mathbf{v}_{\min}^{(8)} \in \mathcal{E}_{8,2}^{(5)}$ and $\mathrm{rank}\left(\mathrm{mat}_{2^4\times 2^4}(\mathbf{v}_{\min}^{(8)}) \right) = 8$. We observe again, in fact more than once, a stagnation of the RQ error at the beginning of the iteration while the residual norm decreases. If we take a closer look in case $d = 16$, we find that for $(d_1, d_2) = (8, 9)$, besides $\mathbf{v}_{\min}^{(8)} \in \mathcal{E}_{8,2}^{(5)}$ it is $\mathbf{v}_{\min}^{(9)} \in \mathcal{E}_{9,2}^{(6)}$, which implies $\widetilde{\mathbf{v}}^{(16)} \in \mathcal{E}_{16,2}^{(13)}$ by Theorem 5.3. Up to iteration step $l = 40$, the Rayleigh quotient of the iterates is close to the 25-th smallest, simple eigenvalue $\lambda_{25}^{(16)}$, which in turn is the smallest eigenvalue such that an associated eigenvector belongs to $\mathcal{E}_{16,2}^{(13)}$ just like the prolongated initial guess. The second, more narrow plateau of the RQ error occurs since the Rayleigh quotient of the iterates around iteration step $l = 60$ approximates the second smallest eigenvalue $\lambda_2^{(16)}$ up to order 10^{-6}. The sparsity pattern of an eigenvector associated with $\lambda_2^{(16)}$ resp. $\lambda_{\min}^{(16)}$ also differs as $\mathbf{v}_2^{(16)} \in \mathcal{E}_{16,2}^{(11)}$ resp. $\mathbf{v}_{\min}^{(16)} \in \mathcal{E}_{16,2}^{(10)}$. Similar to the previous test, the overall performance of the prolongated initial guess compared with the random ones is not satisfactory.

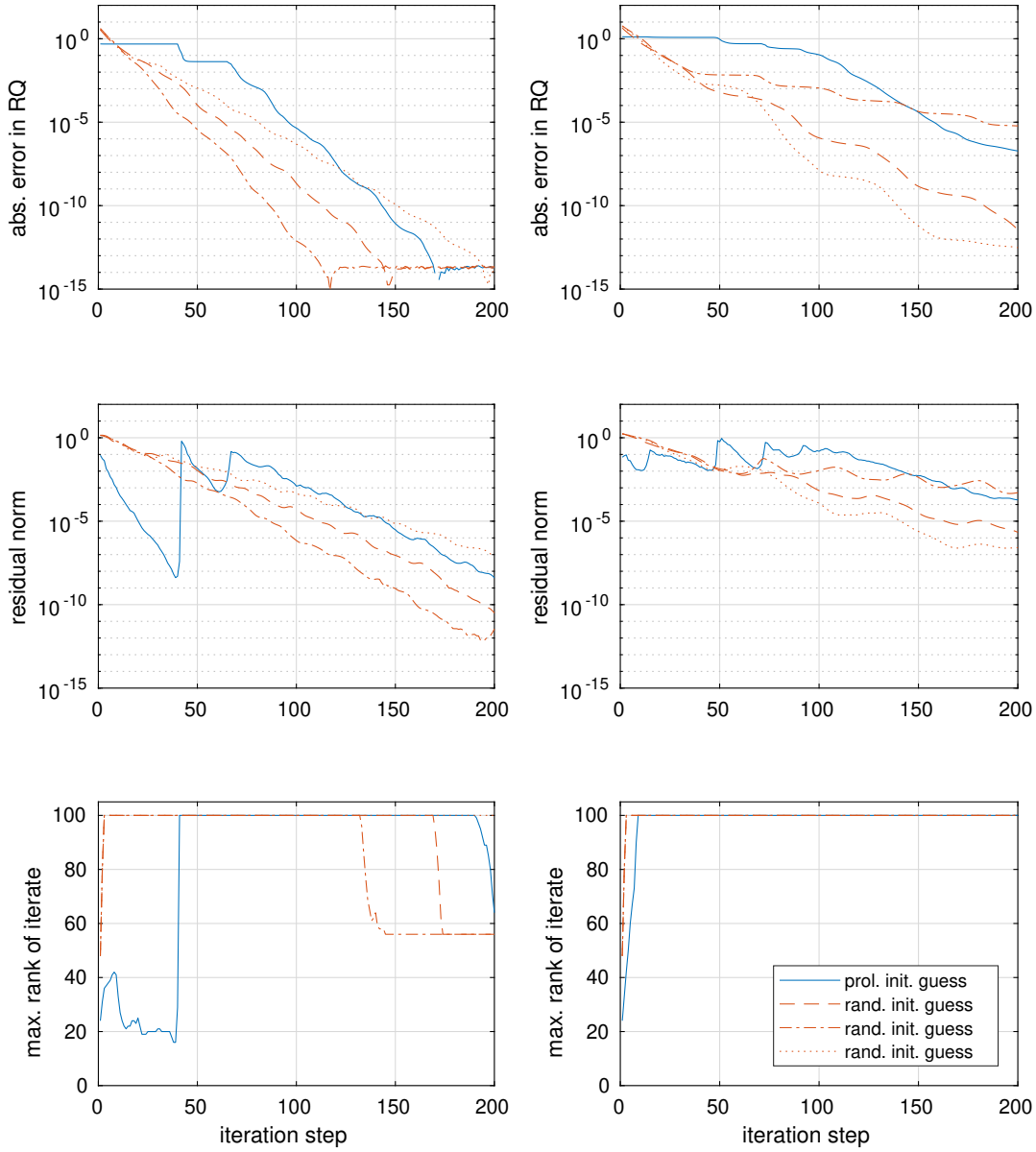


Figure 6.38.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (8, 9)$, LOCG, linear tree, *left:* $d = 16$, *right:* $d = 23$

In Section 5.5 we proposed as an alternative in the case $A = B$ the "constant initial guess". This term stands for a vector $\widetilde{\mathbf{v}}_{\text{const}}^{(k)} \in \mathcal{E}_{d,q}^{(k)}$ with $0 \le k \le (q-1)d$ all of whose nonzero entries are equal and may be chosen such that $\|\widetilde{\mathbf{v}}_{\text{const}}^{(k)}\| = 1$. The construction of this constant initial guess in a tensor format is also detailed in Section 5.5. Our main objective in the next tests is to investigate whether a constant initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$ with k^* such that $\mathbf{v}_{\min}^{(d)} \in \mathcal{E}_{d,q}^{(k^*)}$ is advantageous compared to a constant initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k)}$ for some $k \neq k^*$ or a random initial guess. In the present situation of $q = 2$, $A = B = 1$, $\Delta = -0.2$, $h = 0.3$, it is $k^* = 10$ if $d = 16$

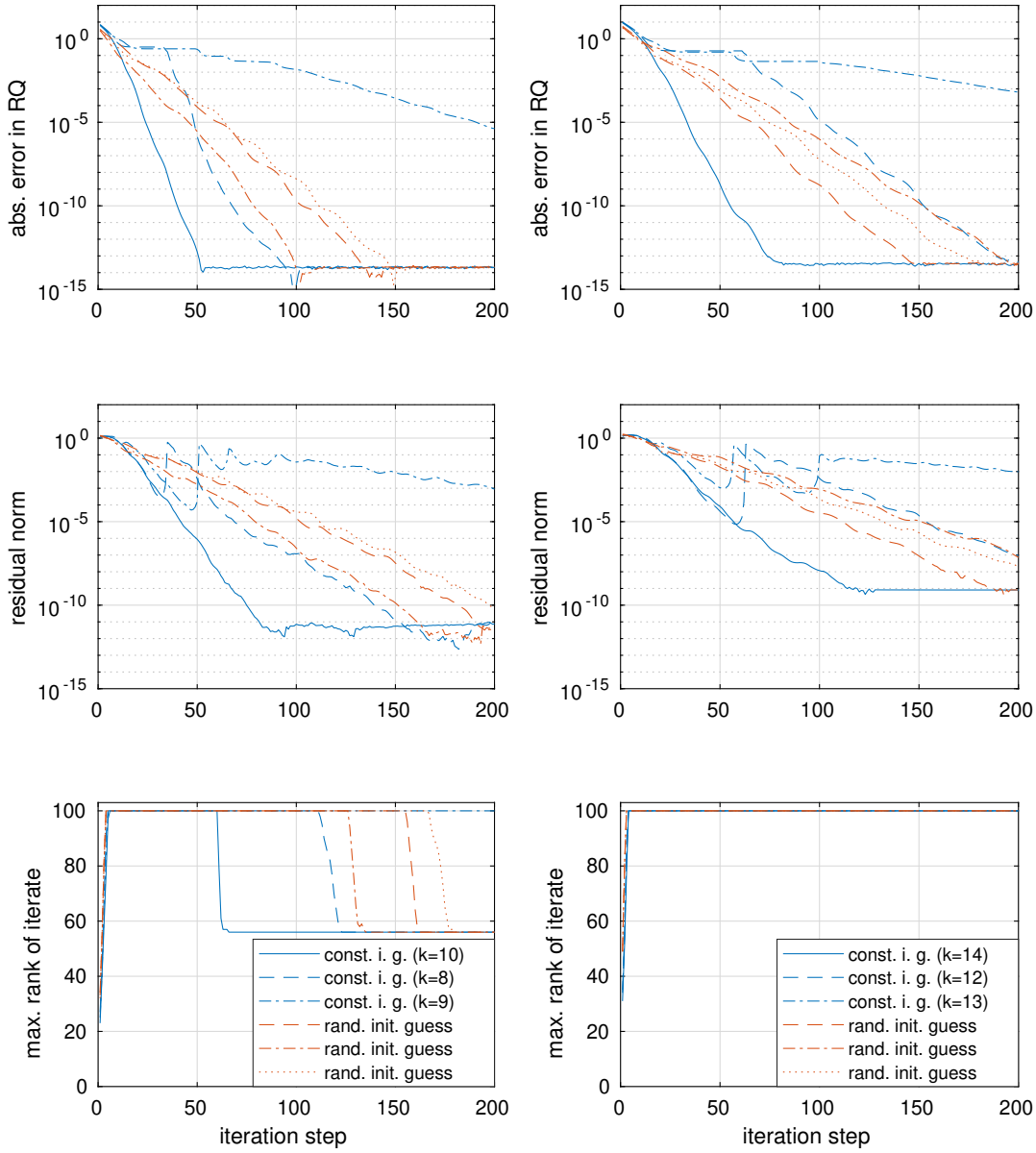


Figure 6.39.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, LOCG, linear tree,
left: $d = 16$, *right:* $d = 22$

and $k^* = 14$ if $d = 22$. For the tests whose results are displayed by Figures 6.39 and 6.40, which are in turn distinguished by linear or balanced dimension tree, we executed the LOCG method with initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k)}$ for each $1 \leq k \leq d-1$, supplemented by three instances of a random initial guess set up with the same ranks as $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$. The vectors $\widetilde{\mathbf{v}}_{\text{const}}^{(0)} = (1, 0, \ldots, 0)^\top$ and $\widetilde{\mathbf{v}}_{\text{const}}^{(d)} = (0, \ldots, 0, 1)^\top$ are actually eigenvectors, see Remark 2.13(i), but associated with an eigenvalue larger than $\lambda_{\min}^{(d)}$. We depict the convergence behavior originating from those $\widetilde{\mathbf{v}}_{\text{const}}^{(k)}$ where the error in the Rayleigh quotient decreases the fastest, the second fastest, and

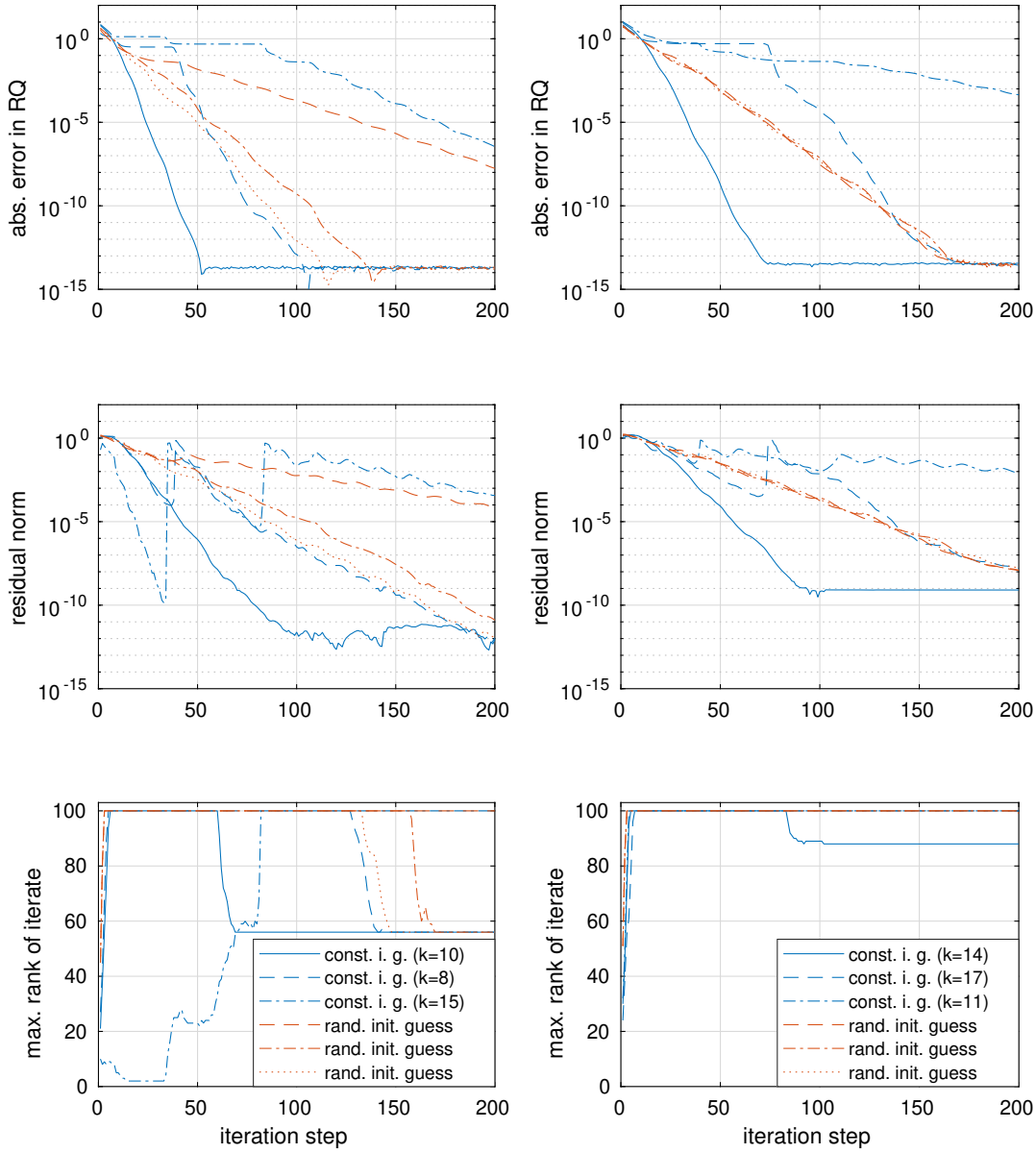


Figure 6.40.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, LOCG, balanced tree,
left: $d = 16$, *right:* $d = 22$

the slowest. In fact, the constant initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$ which has the same sparsity pattern as $\mathbf{v}_{\min}^{(d)}$ yields the fastest convergence which is also at least two times faster than for the random initial guess. The geometry of the dimension tree has practically no effect on the course of the fastest iteration started with $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$ having the correct sparsity pattern. However, the indices k for which we observe second fastest and slowest convergence deviate with respect to the tree structure in three of four cases.

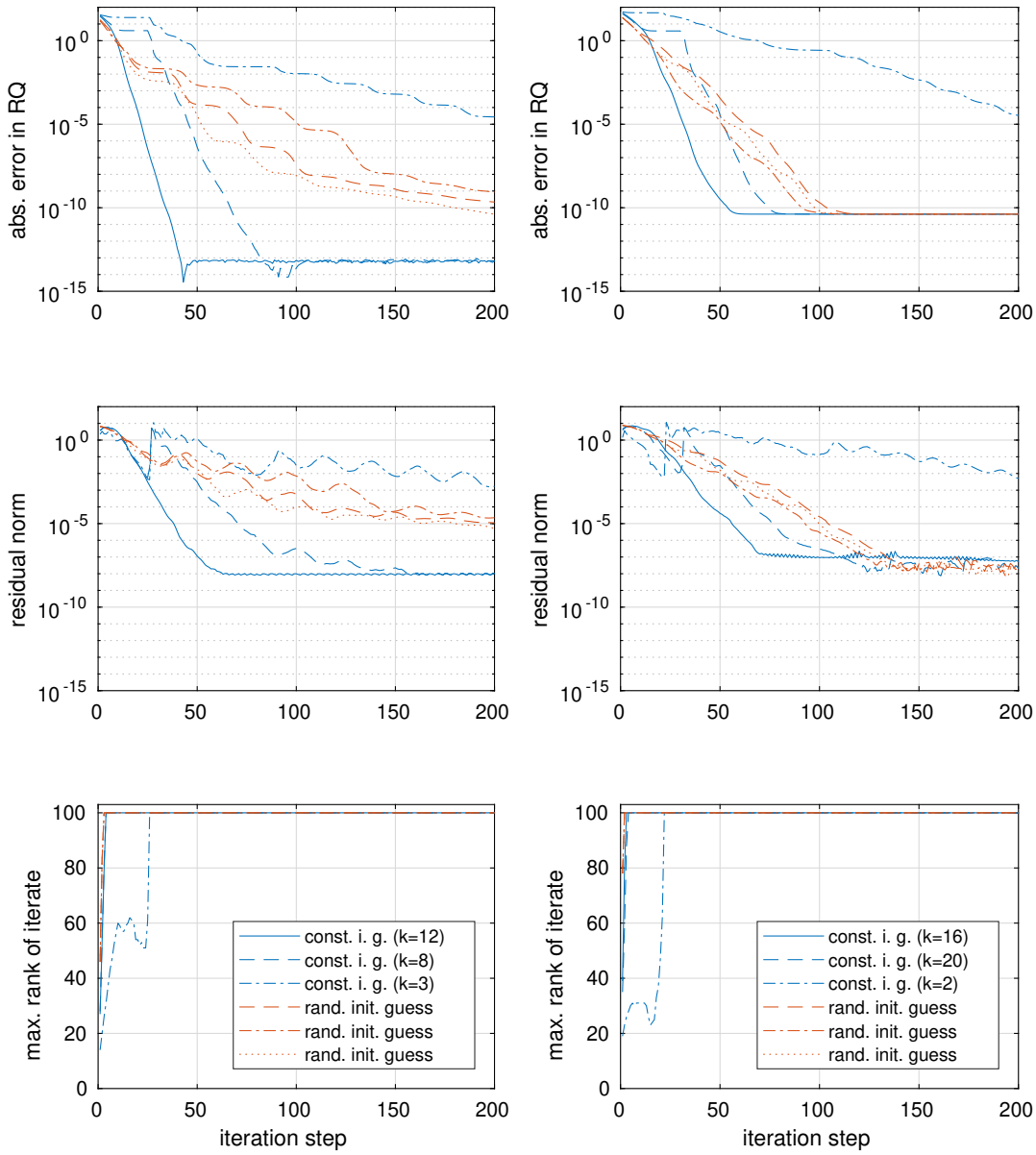


Figure 6.41.: 3-XYZ, $(A, B, \Delta, h) = (2, 2, 1, 1)$, LOCG, linear tree,
left: $d = 10$, *right:* $d = 14$

We perform the same test for a 3-XYZ model with $A = B = 2$, $\Delta = h = 1$ and collect the results in Figures 6.41 and 6.42. For these parameters it is $\mathbf{v}_{\min}^{(10)} \in \mathcal{E}_{10,3}^{(12)}$ and $\mathbf{v}_{\min}^{(14)} \in \mathcal{E}_{14,3}^{(16)}$, and we observe again that the iteration started with $\widetilde{\mathbf{v}}_{\text{const}}^{(12)}$ respectively $\widetilde{\mathbf{v}}_{\text{const}}^{(16)}$ converges the fastest compared to the other constant initial guesses and additionally much faster than for the random initial guess. The influence of the two different dimension trees on the convergence behavior is similar to the previous test.

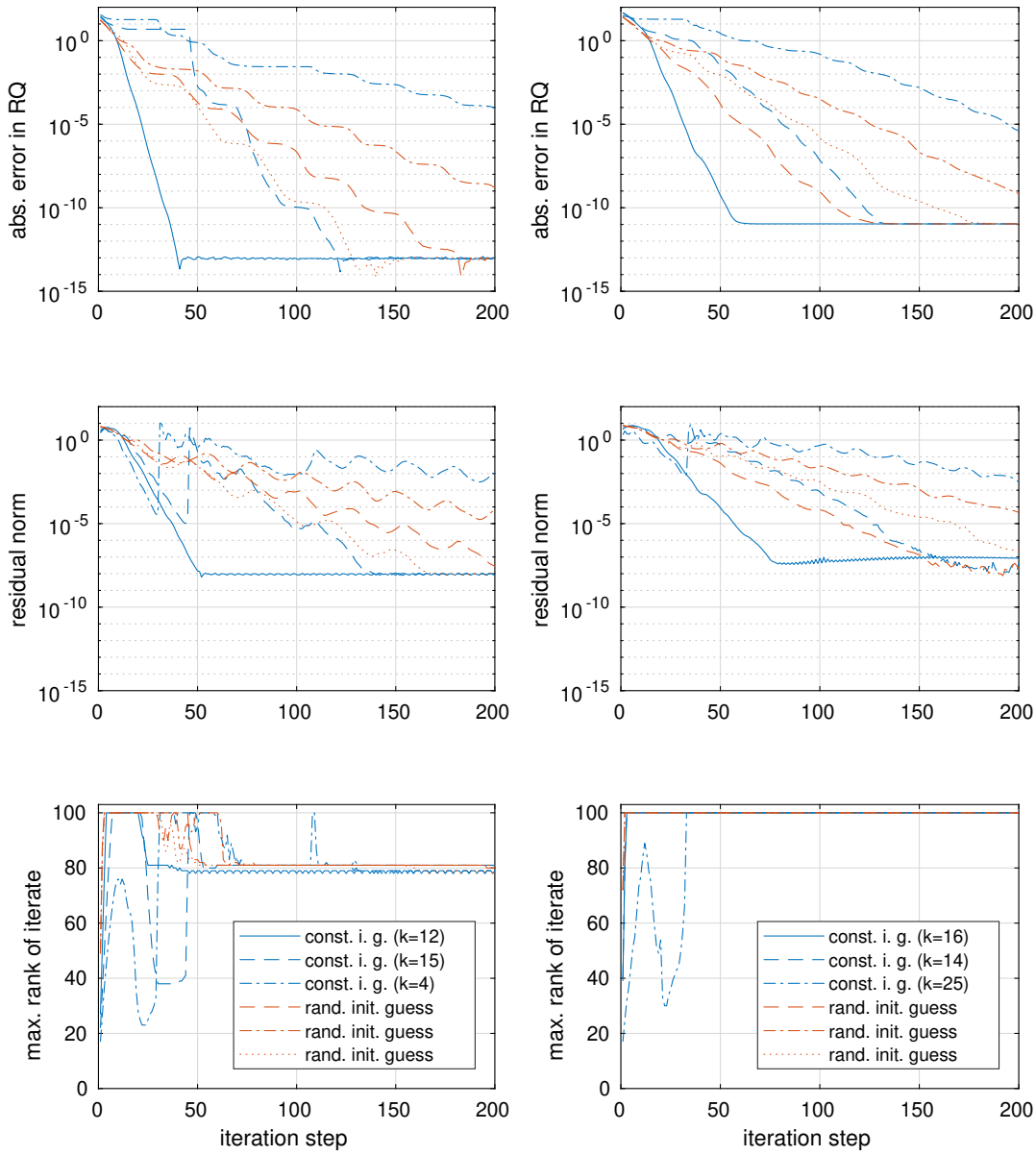


Figure 6.42.: 3-XYZ, $(A, B, \Delta, h) = (2, 2, 1, 1)$, LOCG, balanced tree,
left: $d = 10$, *right:* $d = 14$

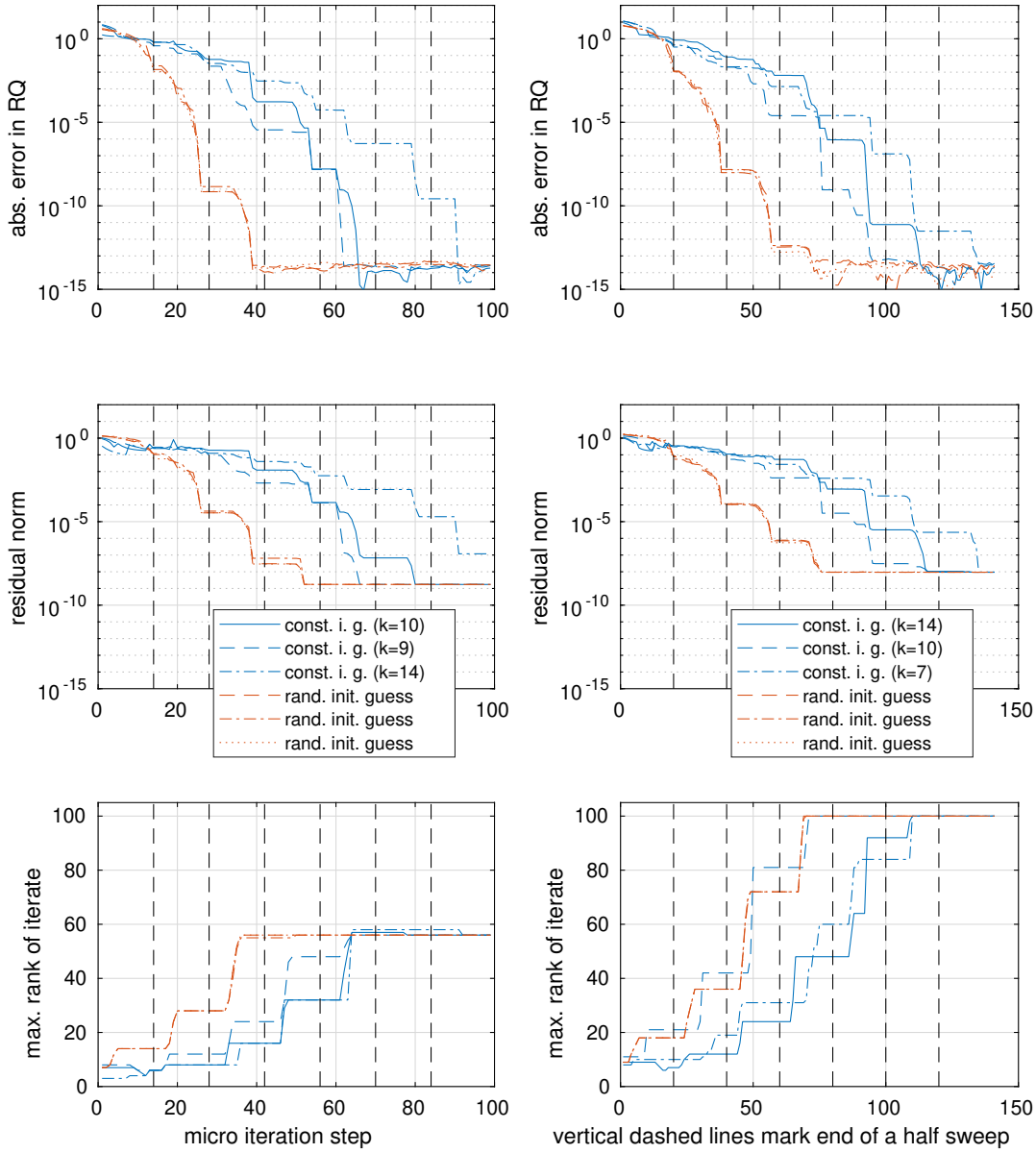


Figure 6.43.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, MALS, *left:* $d = 16$, *right:* $d = 22$

6.5. TT format, $A = B$

In this section we repeat the tests from Section 6.4 concerning the usefulness of the constant initial guess, now for the MALS in the TT format. We remember the graphical presentation of the results of this algorithm in Section 6.3 and consider in Figure 6.43 the 2-XYZ model with $A = B = 1$, $\Delta = -0.2$, $h = 0.3$ where $\mathbf{v}_{\min}^{(d)} \in \mathcal{E}_{d,q}^{(k^*)}$ for $k^* = 10$ if $d = 16$ and $k^* = 14$ if $d = 22$. In contrast to Section 6.4, we observe that the iteration started with the constant initial guess $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$ needs more half sweeps to converge than with the random initial guesses all of which show very similar performance and are again set up with the same TT ranks

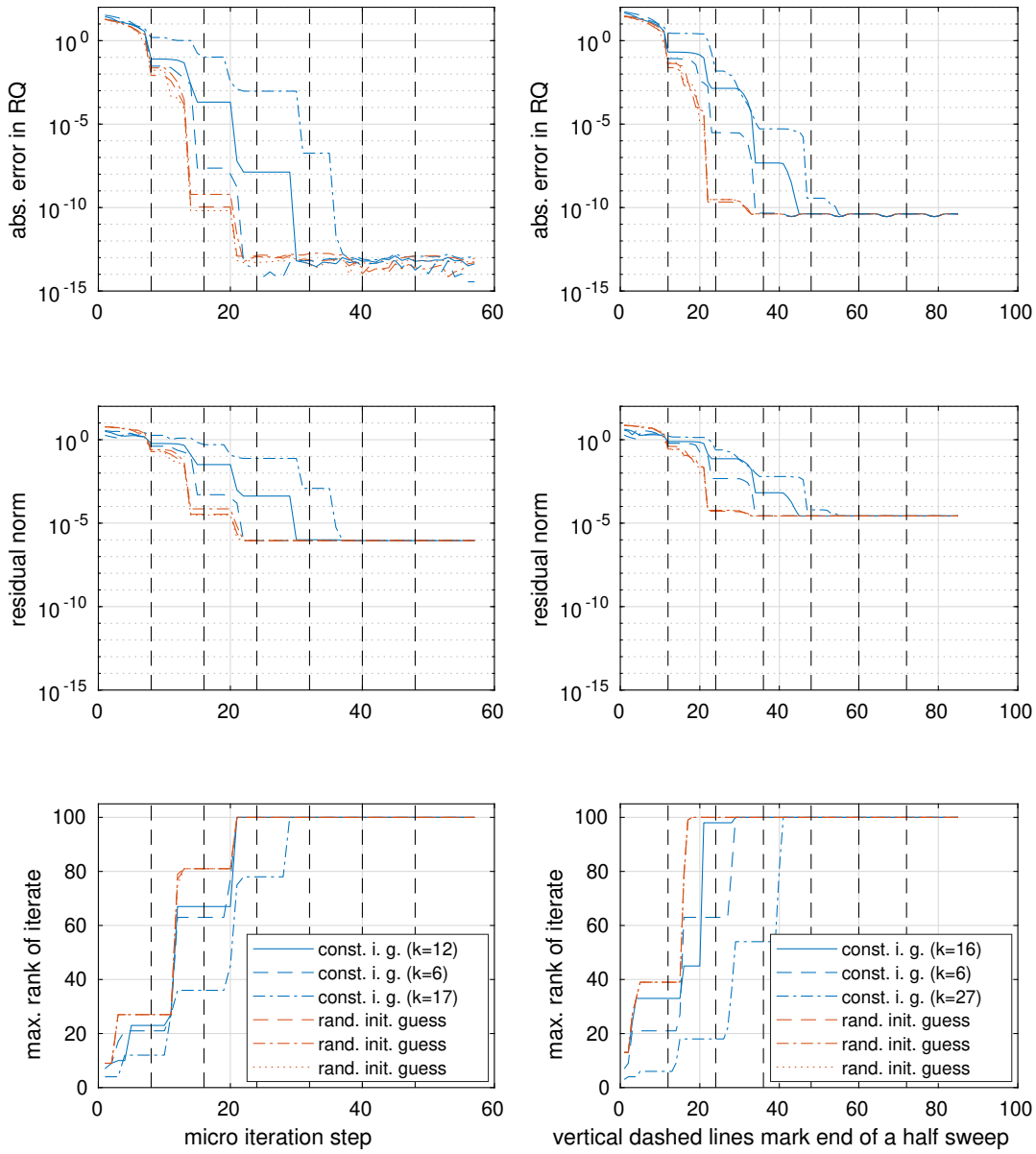


Figure 6.44.: 3-XYZ, $(A, B, \Delta, h) = (2, 2, 1, 1)$, MALS, *left:* $d = 10$, *right:* $d = 14$

as $\widetilde{\mathbf{v}}_{\text{const}}^{(k^*)}$. Additionally, we depict the behavior for that constant initial guess which leads to the fastest resp. the slowest convergence by the dashed resp. dash-dotted blue line. Even for the fastest constant initial guess, we need more half sweeps than for the random one. For the 3-XYZ model with $A = B = 2$, $\Delta = h = 1$, Figure 6.44, we observe a similar behavior, however the fastest constant initial guess shows a performance closer to the random initial guess.

Meanwhile, the prolongated initial guess based on the pseudoinverse performs much better than for the LOCG method in HT format discussed in Section 6.4. In the situation of $q = 2$, $A = B = 1$, $\Delta = -0.2$, $h = 0.3$, $(d_1, d_2) = (8, 9)$, $d \in \{16, 23\}$, depicted in Figure

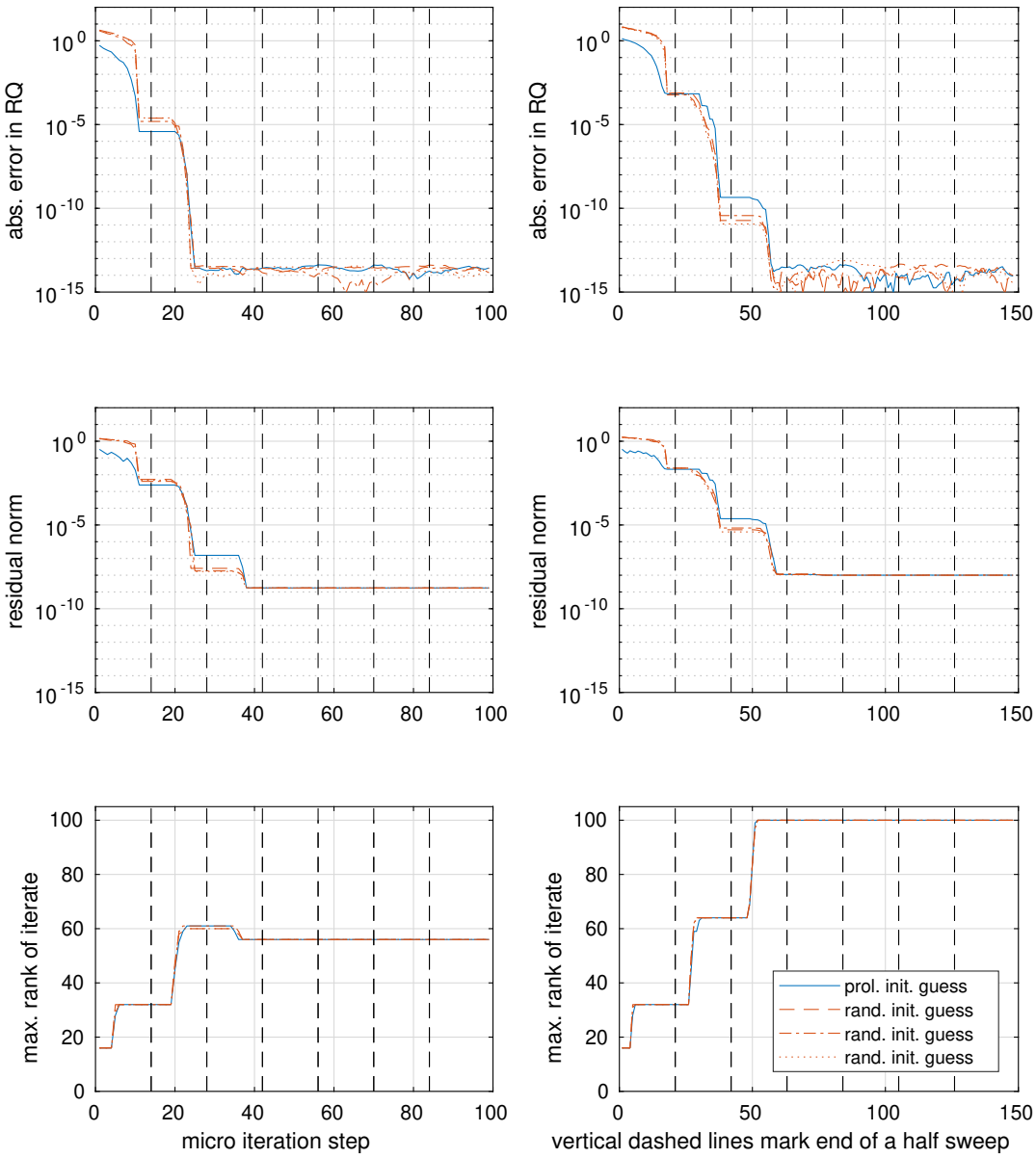


Figure 6.45.: 2-XYZ, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (8, 9)$, MALS, *left:* $d = 16$, *right:* $d = 23$

6.45, the convergence behavior for the TT representative of the prolongated initial guess $\widetilde{\mathbf{v}}^{(d)}$ from (6.7) is practically the same as for the random initial guess. Concerning the 3-XYZ model with $A = B = 2$, $\Delta = h = 1$ and $(d_1, d_2) = (3, 4)$, $d \in \{10, 14\}$, we construct the prolongated initial guess via

$$\mathbf{M} := \begin{pmatrix} \mathbf{I}_3 \otimes \mathbf{N}^{(\mathrm{I})} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_3 \otimes \mathbf{N}^{(\mathrm{II})} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_3 \otimes \mathbf{N}^{(\mathrm{III})} \end{pmatrix},$$

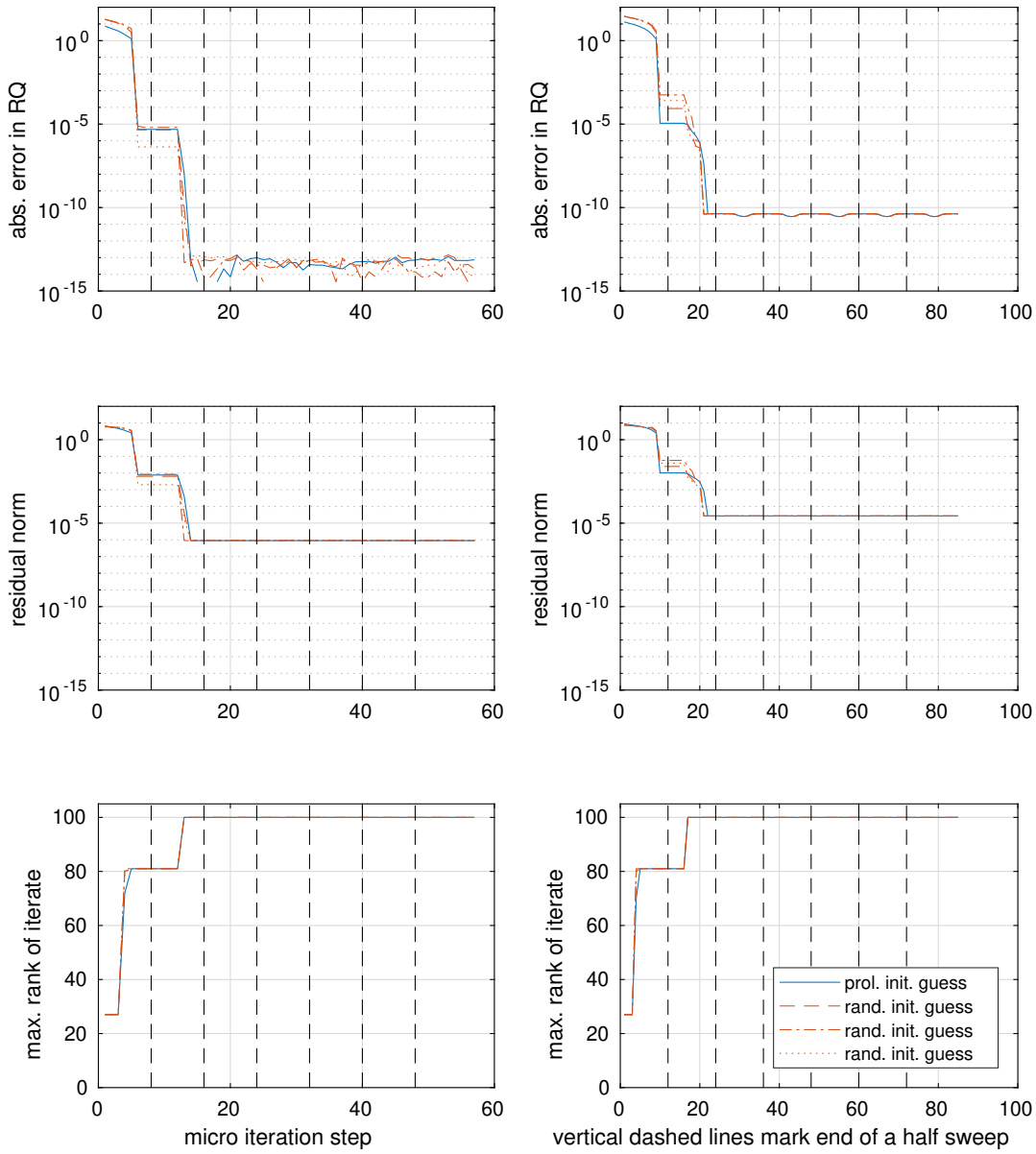


Figure 6.46.: 3-XYZ, $(A, B, \Delta, h) = (2, 2, 1, 1)$, $(d_1, d_2) = (3, 4)$, MALS, *left:* $d = 10$, *right:* $d = 14$

$$\mathbf{N}^{(\mathrm{I})} := \mathrm{mat}_{9\times 3}((\mathbf{v}^{(4)}_{\min})_{1:27}) \cdot \mathrm{pinv}\left(\mathrm{mat}_{3\times 3}((\mathbf{v}^{(3)}_{\min})_{1:9}); 10^{-6}\right),$$
$$\mathbf{N}^{(\mathrm{II})} := \mathrm{mat}_{9\times 3}((\mathbf{v}^{(4)}_{\min})_{28:54}) \cdot \mathrm{pinv}\left(\mathrm{mat}_{3\times 3}((\mathbf{v}^{(3)}_{\min})_{10:18}); 10^{-6}\right),$$
$$\mathbf{N}^{(\mathrm{III})} := \mathrm{mat}_{9\times 3}((\mathbf{v}^{(4)}_{\min})_{55:81}) \cdot \mathrm{pinv}\left(\mathrm{mat}_{3\times 3}((\mathbf{v}^{(3)}_{\min})_{19:27}); 10^{-6}\right),$$

and

$$\widetilde{\mathbf{v}}^{(d)} := \left(\prod_{i=1}^{d-4} (\mathbf{I}_{3^i} \otimes \mathbf{M})\right) \mathbf{v}^{(4)}_{\min},$$

noticing for the present setting $\mathbf{v}_{\min}^{(3)} \in \mathcal{E}_{3,3}^{(4)}$, thus

$$\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{1:9}\big) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & * \end{pmatrix},$$
$$\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{10:18}\big) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & * \\ 0 & * & 0 \end{pmatrix},$$
$$\mathrm{mat}_{3\times 3}\big((\mathbf{v}_{\min}^{(3)})_{19:27}\big) = \begin{pmatrix} 0 & 0 & * \\ 0 & * & 0 \\ * & 0 & 0 \end{pmatrix}.$$

As we read off from Figure 6.46, this prolongated initial guess also behaves practically equally to the random initial guess.

7. Conclusion

In this thesis we have been concerned with the computation of an eigenvector associated with the minimal eigenvalue of a Hamilton operator modeling a quantum mechanical spin system, in other words the ground state of the system. As the number of elements of the eigenvector to be determined grows exponentially with the number of particles in the spin system, we had to represent the involved objects in a format with an amount of parameters that has much lower complexity. Formulating the large-scale eigenvalue problem in terms of tensor networks has turned out to be very useful. A special focus has been laid on the construction of an initial guess in order to accelerate the computations.

We summarize in this last chapter of the thesis our gain in knowledge and relate it to the tasks phrased in Section 1.1. In addition, we name some directions in which further research may be carried out.

7.1. Summary

We have considered two types of Hamilton operators $\mathbf{H}_{d,q} \in \mathbb{R}^{q^d \times q^d}$ modeling the energy in a one-dimensional quantum mechanical spin system with d particles, namely those of the q-XYZ model and the q-Potts model. In Section 2.1, the main contribution was the statement of Theorem 2.9 about the existence of an orthonormal basis of each eigenspace of $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ each of whose basis elements obeys a certain sparsity pattern. We introduced these sets $\mathcal{E}_{d,q}^{\mathrm{even}}$ and $\mathcal{E}_{d,q}^{\mathrm{odd}}$ respectively $\mathcal{E}_{d,q}^{(k)}$, $0 \leq k \leq (q-1)d$, of specifically structured vectors in Definition 2.6 by means of the sum of digits of q-ary representations of the indices of the vector entries with the convention to start counting the indices at 0. It turned out that it depends on the equality respectively inequality of the coupling parameters A and B in $\mathbf{H}_{d,q}^{\mathrm{XYZ}}$ whether the eigenvectors belong to $\mathcal{E}_{d,q}^{(k)}$ respectively $\mathcal{E}_{d,q}^{\mathrm{even}}$ or $\mathcal{E}_{d,q}^{\mathrm{odd}}$. For some special cases of the coupling parameters A, B, Δ, and the parameter h corresponding to an external magnetic field, we determined the minimal eigenvalue and associated eigenvectors in Proposition 2.11 analytically. Furthermore, a relation between eigenvalues for varying h if $A = B$ was the content of Theorem 2.12.

In Section 2.2, similar statements were proven for the Hamilton operator $\mathbf{H}_{d,q}^{\mathrm{Potts}}$ of the q-Potts model. The relevant sparsity patterns of the eigenvectors were characterized by the spaces $\mathcal{E}_{d,q}^{[k]}$, $0 \leq k \leq q-1$, see Definition 2.15, in terms of considering modulo q the sum of digits of the q-ary representation of the indices. This led to Theorem 2.17 as a counterpart of Theorem 2.9 for the Potts model. Proposition 2.19 contained the adaption of Proposition 2.11 to special cases of the parameters A and h in $\mathbf{H}_{d,q}^{\mathrm{Potts}}$. The central tool in the proofs of the theorems concerning the sparsity pattern of eigenvectors in Chapter 2 was the result of Lemma 2.3, a statement about the relation between subspaces invariant with respect to a matrix and eigenvectors of this matrix.

Chapter 3 was devoted to an overview of the terms in the context of tensors, tensor networks, and tensor formats as an approach to overcome the exponential scaling of the eigenvalue problem with increasing number of particles. Section 3.1 mentioned basic con-

cepts for tensors like the vector space structure, inner product and norm, and reshaping operations that relate tensors of different order, vectors, and matrices to each other. Section 3.2 introduced the notion of tensor networks which turned out to be a powerful instrument to express tensors and their interplay via tensor contractions. In Section 3.3 respectively 3.4, we explained the necessary details with regard to the hierarchical Tucker format respectively the tensor train format as the particular tensor formats we dealt with in the thesis. We pointed to the growth of rank which occurs by arithmetical operations with representatives in these two tensor formats and the resulting need of truncating the ranks regularly when working algorithmically with such representatives. Furthermore, we stated how to represent the tensorized versions of the Hamilton operators of the XYZ model and the Potts model in these tensor formats with an amount of parameters that scales only linearly in d, thereby satisfying an assumption that was imposed in Section 1.1.

In Chapter 4 we described two numerical methods for the computation of an eigenvector associated with the minimal eigenvalue of a symmetric matrix which we assumed throughout this thesis to be simple. Since the Hamilton operator, regarded as a matrix, is symmetric, the approach was to minimize the Rayleigh quotient. One method was the well-known locally optimal preconditioned conjugate gradient method, but we argued not to consider preconditioning and hence abbreviated the method by LOCG, see Algorithm 4.1. Instead of computing with matrices and vectors, all constituents of the method like the Hamilton operator, the current iterate, the gradient, and as a third search direction a combination of gradients from previous iteration steps, were replaced by their respective representatives in the HT format, due to the truncation of ranks in general only approximately. This method was summarized in Algorithm 4.2.

The other method was the modified alternating linear scheme (MALS), also known in the physics community as density matrix renormalization group (DMRG), usually formulated in the tensor train format. We explained, and recapitulated in Algorithm 4.3, that this method operates by updating repeatedly, in a so-called sweeping procedure, one after another the single cores of the TT representative of the iterate of the tensorized eigenvector. Adaptivity of the TT ranks was incorporated by optimizing in each so-called micro iteration step the contraction of two neighboring cores via solving a local eigenvalue problem and decomposing the fourth-order tensorization of the resulting eigenvector by a truncated SVD.

Chapter 5 contributed the construction of an initial guess for an iterative numerical method that computes an eigenvector associated with the minimal eigenvalue. This was formulated in Section 1.1 as a main goal of this thesis and we named three properties (P1), (P2), (P3) to be essential, now relating our findings to them. We proposed the construction of an initial guess based on the relation of eigenvectors $\mathbf{v}_{\min}^{(d_1)} \in \mathbb{R}^{q^{d_1}}$ and $\mathbf{v}_{\min}^{(d_2)} \in \mathbb{R}^{q^{d_2}}$ associated with the minimal eigenvalues of $\mathbf{H}_{d_1,q}$ respectively $\mathbf{H}_{d_2,q}$ defined by the same A, B, Δ, h or A, h like $\mathbf{H}_{d,q}$, where $d_1 < d_2 < d$. By means of matricizations of $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$, a linear mapping was defined which we called a prolongation operator, with reference to a certain analogy of our concept to the idea of multigrid methods. Via applying this prolongation operator in turn to $\mathbf{v}_{\min}^{(d_2)}$, or possibly to an eigenvector associated with a larger eigenvalue related to d_2, we were able to define a vector corresponding to problem size d, hence being an element of $\mathbb{R}^{q^d}$, if $d = d_2 + (d_2 - d_1)i$ for an $i \in \mathbb{N}$.

The feasibility of this construction procedure, whose result was then regarded as the constructed initial guess, depended on the invertibility of the quadratic matricization of $\mathbf{v}_{\min}^{(d_1)}$ in case of even d_1 respectively of q blocks thereof for d_1 odd. In the situation of the q-XYZ model, it turned out that, based on the results of Section 2.1 about the sparsity pattern

of eigenvectors, the prolongation strategy is feasible only if it holds $A \neq B$ for the coupling parameters in the Hamilton operator. If $q = 2$ and $d_1 = 2$, Lemma C.1, to which we referred in Subsection 5.1.1, showed the invertibility of the matricization of $\mathbf{v}_{\min}^{(d_1)}$ in case $A \neq B$. For the 3-XYZ model and $d_1 = 2$, we stated in Subsection 5.1.2 for $A \neq 0 \neq h$, $B = \Delta = 0$ the invertibility of the matricization and conducted for other coupling parameters promising numerical tests. An analogous result was obtained in Subsection 5.1.3 for the 3-Potts model with $d_1 = 2$ in the general case of $A \neq 0 \neq h$ not already addressed by Proposition 2.19.

Theorem 5.3 clarified that the prolongation procedure indeed yields a vector matching the sparsity patterns that are admissible for an eigenvector related to problem size d and stated a formula how to determine the actual sparsity of this prolongated vector. Together with the comparison of the Rayleigh quotient of the prolongated vector with the exact minimal eigenvalue in different test cases, this indicates that Property (P3) referring to the approximation quality of the prolongated initial guess is satisfied at least in the situations of an XYZ model with $A \neq B$ and a Potts model summarized so far.

We continued in Section 5.2 with translating the construction of an initial guess by prolongation to the HT format based on a linear dimension tree. We pointed out that the construction in linear HT format is possible in such a way that it yields a representative of the prolongated initial guess with a maximal rank independent of the final problem size d and collected the explicit ranks for some relevant cases in Remark 5.5. The conversion of the representative of the initial guess from the HT format with linear dimension tree to that with balanced tree was explained in Section 5.3 by means of straightforward manipulations of the tensor network. We did not elaborate on the direct construction of the prolongated initial guess in balanced HT format and the resulting behavior of the HT ranks. Since the linear HT format is closely related to the TT format, the representation of the prolongated initial guess in the latter was mentioned in Section 5.4 quite briefly.

Section 5.5 picked up the problem that for an XYZ model with $A = B$, the prolongation strategy is not feasible in general as the eigenvectors contain too many zero entries. This in fact finer classification of the structure of the eigenvectors gave rise to propose an alternatively constructed initial guess, or rather a collection of $(q-1)d+1$ initial guesses, one for each admissible sparsity pattern defined by $\mathcal{E}_{d,q}^{(k)}$, $0 \leq k \leq (q-1)d$. Since all the nonzero entries were chosen to equal a constant value and to yield a normalized vector, we called this the constant initial guess. The rationale for this strategy was that in a computing environment suitable for the parallel execution of several instances of the eigensolver with different initial guesses, it might be advantageous to start at least one instance with the same sparsity as an exact solution, and pick in the course of the iteration that instance with the best convergence behavior. We described the construction of a representative of the constant initial guess in HT and TT format and stated that the maximal rank scales linearly in d.

From the expositions made so far, we infer that the prolongated just as the constant initial guess matches very well a low-rank tensor format and hence satisfies Property (P2). As demonstrated for the linear HT and the TT format, once the exact eigenvectors $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ are determined, the actual prolongation procedure up to problem size d does not require any computational work at all, since the representatives in the tensor formats only contain information directly obtained from $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$. Combined with the fact that the computation of $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ up to machine precision is not expensive for small d_1 and d_2, respectively recognizing that the construction of the constant initial guess does not rely on any computable quantity, Property (P1) concerning the comparably cheap accessibility of an initial guess is also fulfilled.

As a last part of this summary, we comment on the results of the numerical tests in Chapter 6. Since the numerical methods require parameters which control the truncation of HT and TT ranks, we conducted some tests regarding this issue in Section 6.1. We decided to choose the upper bound of the ranks as well as a specific relative error bound to be the same for all subsequent tests and especially independent of d and q.

The LOCG method, as discussed in Section 6.2, profited in each situation clearly from the prolongated initial guess compared to a random initial guess in the sense that the number of iteration steps until reaching a small error level is at least halved or even reduced down to one tenth. We observed that the iterates stay in the same set of sparsity pattern as the initial guess and hence possibly approach an eigenvector associated with a non-minimal eigenvalue. However, by applying the prolongation operator not to $\mathbf{v}_{\min}^{(d_2)}$ but instead to another eigenvector related to problem size d_2 with a different sparsity pattern, the sparsity of the initial guess may be altered. Of course, the correct sparsity of $\mathbf{v}_{\min}^{(d)}$ is not known a priori, but even executing the method two times for both an initial guess located in $\mathcal{E}_{d,q}^{\text{even}}$ and $\mathcal{E}_{d,q}^{\text{odd}}$ is competitive to the performance of a random initial guess. We observed in some situations that for d_1 and d_2 getting larger, convergence to the minimal eigenvalue gets faster. The geometry of the underlying dimension tree, whether linear or balanced, turned out to have practically no effect on the convergence behavior. However, in some cases the numerical method consumes more CPU time for the balanced tree.

For the MALS, see Section 6.3, which was generally much faster than the LOCG method with respect to the CPU time, the influence of the prolongated initial guess on the necessary number of half sweeps was less striking. An advantage occured in some situations, mainly when the approximation error compared to the random initial guess differed rather much at the beginning of the iteration. In one of four tests with comparably large three-digit d, the prolongated initial guess was superior. For a setting where the prolongated initial guess has the wrong sparsity pattern, altering this pattern by prolongating a vector different to $\mathbf{v}_{\min}^{(d_2)}$ did not yield an improvement. Except this single event, the prolongated initial guess showed an equivalent or slightly better performance than the random one.

The tests in Section 6.4 for the situation of coupling parameters $A = B$ in an XYZ model implied that for LOCG, the constant initial guess should be preferred instead of a prolongated initial guess constructed via the pseudoinverse of a non-invertible matricization of $\mathbf{v}_{\min}^{(d_1)}$. The constant initial guess with the correct sparsity pattern performs much better than the random initial guess. Even if there is a priori no knowledge about that correct pattern, starting in parallel many instances of the method for constant initial guesses with the different admissible sparsity patterns appears as an acceptable approach.

For the MALS in case $A = B$, see Section 6.5, we received a reversed impression. Even with the constant initial guess having the correct sparsity pattern, it took more half sweeps to converge as with the random initial guess, while the prolongated initial guess constructed via the pseudoinverse of the matricization of $\mathbf{v}_{\min}^{(d_1)}$ performed like the random initial guess.

Summing up, we have introduced specific notions of sparsity patterns of vectors and have proven the existence of eigenvectors with these patterns for certain Hamilton operators. Based on these findings, we have proposed strategies to construct an initial guess for numerical methods that compute an eigenvector associated with the minimal eigenvalue. We have demonstrated that these construction strategies match very well a low-rank tensor format. As a result of numerical tests, we have observed that the constructed initial guesses substantially accelerate one particular type of numerical method, while for another type of algorithm they have shown a performance at least equivalent to a random initial guess.

7.2. Outlook

To recommend in which directions further research should be done, we distinguish two different objectives. On the one hand, there are topics which are summarized under the term of "generalization and variation" while on the other hand, especially based on the results of this thesis, issues exist where one should aim for "improvement".

With regard to the former, we mention the possibility to consider other Hamilton operators, namely those stemming from the interactions of more than only directly neighbored particles, cf. [BFSB14, Section II] or [BBF17, Section II], or the presence of more than one external magnetic field, see [BBF19, Section II]. The construction of these Hamilton operators proceeds like for the models considered so far in this thesis. Each of the interactions of several specific particles with each other enters the Hamilton operator as a summand which is a Kronecker product containing the respective spin matrix as a factor exactly at those positions which correspond to the location of the interacting particles. All remaining Kronecker factors in each summand are identity matrices. Just as well, each interaction of a particle with an external magnetic fields amounts to an additional summand with only one factor being a spin matrix and not an identity.

Moreover, Hamilton operators modeling so-called long-range interactions, see [Can95] or [SGL$^+$20, Section IV], may be taken into account. This means that also distant particles are regarded as interacting with each other. The strength of the interaction typically decreases with increasing distance. In order to apply the developed ideas also to such scenarios, one has to check whether the eigenvectors satisfy a similar type of sparsity pattern upon which the prolongation scheme is built. Depending on these findings, the procedure of constructing an initial guess has to be adapted to the particular structure of the eigenvectors.

Another type of variation arises from formulating the DMRG/MALS not only based on the tensor train format, but instead to allow more general geometries of the tensor network as done in [GSR$^+$14] or [STG$^+$19, Section 5]. The sweeping procedure is in principle not limited to the linear arrangement of core tensors in the tensor train format. If one employs a tree structure like in the hierarchical Tucker format as an ansatz for a representative of a tensorized eigenvector, there is some freedom to choose the order in which the nodes of the tensor network are optimized. Depending on this order, some nodes are visited more frequently than other ones. Besides, a representative in the HT format contains more individual tensors than a representative in the TT format for fixed d. This may lead to larger computational cost until having updated all of these tensors at least once which would be seen as a half sweep in that case. Hence one has to discuss which order is the most efficient one and whether there is a benefit compared to the TT format.

Concerning the objective of improving the strategies proposed in this thesis, we name the situation of $A = B$ in the XYZ model where the construction of an initial guess by the developed type of prolongation does not work properly. One has to think about alternative ways to relate information available for different small problem sizes. To this end, it might be advantageous also to include information about the representation of the relevant eigenvectors $\mathbf{v}_{\min}^{(d_1)}$ and $\mathbf{v}_{\min}^{(d_2)}$ in a tensor format instead of only relying on their structure in full vector format when setting up the prolongation procedure. It has to be investigated how the undoubtedly helpful knowledge of the various sparsity patterns in full vector format may be combined with the properties of the eigenvectors or their tensorizations revealed by a tensor decomposition. With respect to an improvement in efficiency for the constant initial guess, a priori statements about the actual sparsity pattern of $\mathbf{v}_{\min}^{(d)}$ for different domains of A, B, Δ, h are desirable.

A further issue to address is the observation that the MALS profits only partially from employing a constructed initial guess instead of a random one. While the LOCG method seems to retain relevant properties of the initial guess like the sparsity pattern since it treats the iterates as a whole and utilizes tensorization and representation in a tensor format rather for auxiliary means, the procedure of optimizing only single nodes of the tensor network in MALS has a different spirit. An exploration to what extent the construction of an initial guess may be tailored better to the idea of MALS is very advisable.

In addition, a topic which was not discussed so far and might be examined is the consideration of possible symmetries in the eigenvectors. If some vector entries can be inferred from other ones, this reduces the number of quantities to compute. However, one has to clarify how a symmetry in full vector format translates to the values of the parameters constituting a representative in a tensor format. It has also to be argued in which particular way a numerical method based on low-rank tensor representations can take symmetries into account. Since for one-dimensional spin systems there is no specification at which end of the chain particle 1 respectively particle d is located, one expects that the coefficients of $|j_1\ j_2\ \dots\ j_{d-1}\ j_d\rangle$ and $|j_d\ j_{d-1}\ \dots\ j_2\ j_1\rangle$ stored in an eigenvector have at least the same absolute value, cf. [Bal15, Section 2.4]. This is related to the concept of reverse symmetry from [HWSH13b, Section 3.2.3] and, despite the difficulties to incorporate knowledge about symmetries especially into the DMRG/MALS as mentioned in [HWSH13b, Section 3.2.8], further investigations are recommendable.

As it was pursued by our proposed strategy of constructing an initial guess via prolongation, we referred to the idea of utilizing cheaply obtained information to gain an insight into situations being more complicated. Stepping back from the concrete subject matter of quantum mechanical spin systems, it remains a quite general question to which other types of problems this principle of concluding from the small to the large may be applied.

A. Dirac bra-ket notation

In this thesis we employ the *Dirac bra-ket notation*, dating back to [Dir39], only in a quite specific and limited way. The present section is intended to summarize the features most important for our purposes. A more general introduction and overview may be found in [Gri14, Chap. 3].

We denote by the *ket* $|j\rangle$, $j \in \{0, \dots, q-1\}$, the i-th standard basis vector

$$\mathbf{e}_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^q,$$

where the element 1 is at position $i = j + 1$, beginning counting that position with 1. This offset in counting by one is mainly due to the traditional notational convention for the standard basis like in [HJ13, Sect. 0.1.7] on the one hand, and practical reasons when working with the Dirac notation, see below, on the other hand. The value of q is not referred to in the notation and will be clear from the context.

A ket with more than one "entry" denotes the Kronecker product of the respective vectors,

$$|j_1\ j_2\ \dots\ j_d\rangle := |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_d\rangle\,.$$

Such a ket is a unit vector in $\mathbb{R}^{q^d}$. The position $i \in \{1, \dots, q^d\}$ of the 1 in this unit vector $\mathbf{e}_i$ is determined as follows: Identifying each j_n, $n \in \{1, \dots, d\}$, with a digit in the q-ary number system, we are given by $(j_1 j_2 \dots j_d)$ the q-ary representation of a number ι in the decimal system between 0 and $q^d - 1$ to which we add $+1$ to obtain $i = \iota + 1$. As an example, with $q = 3$ and $d = 5$, consider $(01202) \mathrel{\widehat{=}} 01202_3 = 47_{10}$, so $|0\ 1\ 2\ 0\ 2\rangle = \mathbf{e}_{48} \in \mathbb{R}^{243}$.

By the *bra* $\langle j|$, $j \in \{0, \dots, q-1\}$, we denote the (conjugate) transposed unit vector $\mathbf{e}_{j+1}^* = \mathbf{e}_{j+1}^\top$. Here as well

$$\langle j_1\ j_2\ \dots\ j_d| := \langle j_1| \otimes \langle j_2| \otimes \cdots \otimes \langle j_d|\,,$$

with the same rule to determine the position of the 1 in that transposed unit vector.

As it simplifies some situations, we regard the entries in $|\cdot\rangle$ modulo q, thus $|\overline{j_1}\ \overline{j_2}\ \dots\ \overline{j_d}\rangle$ with $\overline{j_n} \notin \{0, \dots, q-1\}$ for some n equals $|j_1\ j_2\ \dots\ j_d\rangle$ with $j_n \in \{0, \dots, q-1\}$ whenever $\overline{j_n} \equiv j_n \bmod q$ for all n. To the entries in $\langle\cdot|$ this applies analogously.

By definition, $|j\rangle\langle l|$ denotes a $q \times q$ matrix whose $(j+1, l+1)$ element equals 1 while all other elements are 0. There is the property

$$\big(|j\rangle\langle l|\big) \otimes \big(|\widetilde{j}\rangle\langle \widetilde{l}|\big) = |j\ \widetilde{j}\rangle\langle l\ \widetilde{l}|,$$

where the additional brackets ( ) are unnecessary since multiplying in the order

$$\underbrace{|j\rangle}_{q\times 1}\underbrace{\left(\langle l|\otimes|\widetilde{j}\rangle\right)}_{q\times q}\underbrace{\langle\widetilde{l}|}_{1\times q}$$

is not possible. Therefore we have

$$|j_1\ \dots\ j_d\rangle\langle l_1\ \dots\ l_d|\otimes|\widetilde{j}_1\ \dots\ \widetilde{j}_{\widetilde{d}}\rangle\langle\widetilde{l}_1\ \dots\ \widetilde{l}_{\widetilde{d}}|=|j_1\ \dots\ j_d\,\widetilde{j}_1\ \dots\ \widetilde{j}_{\widetilde{d}}\rangle\langle l_1\ \dots\ l_d\,\widetilde{l}_1\ \dots\ \widetilde{l}_{\widetilde{d}}|.$$

Yet another feature is

$$\langle j|l\rangle=\delta_{j,l}:=\begin{cases}1, & j=l\\ 0, & j\neq l\end{cases},$$

noticing that a possible second bar | in the middle is typically left out.

B. Sparsity patterns of vectors

Example B.1. The vectors contained in $\mathcal{E}_{d,q}^{\text{even}}$, $\mathcal{E}_{d,q}^{\text{odd}}$, and $\mathcal{E}_{d,q}^{(k)}$, $k \in \{0, \dots, (q-1)d\}$, see Definition 2.6, have the following structure. An entry $*$ may also be 0.

(i) $d = 2$, $q = 2$:

$$\begin{pmatrix} * & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{\text{even}} \qquad \begin{pmatrix} 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{\text{odd}} = \mathcal{E}_{2,2}^{(1)}$$
$$\begin{pmatrix} * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{(0)}$$
$$\begin{pmatrix} 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,2}^{(2)}$$

(ii) $d = 3$, $q = 2$:

$$\begin{pmatrix} * & 0 & 0 & * & 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{\text{even}} \qquad \begin{pmatrix} 0 & * & * & 0 & * & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{\text{odd}}$$
$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{(0)} \qquad \begin{pmatrix} 0 & * & * & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{(1)}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{(2)} \qquad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{3,2}^{(3)}$$

(iii) $d = 4$, $q = 2$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{(0)}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & * & * & 0 & 0 & * & * & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{(2)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{(4)}$$
$$\begin{pmatrix} * & 0 & 0 & * & 0 & * & * & 0 & 0 & * & * & 0 & * & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{\text{even}}$$

$$\begin{pmatrix} 0 & * & * & 0 & * & 0 & 0 & * & * & 0 & 0 & * & 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{\text{odd}}$$
$$\begin{pmatrix} 0 & * & * & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{(1)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & * & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{4,2}^{(3)}$$

(iv) $d = 6$, $q = 2$:

$$\begin{array}{cccccccccccccccccc}
 & \left(\overset{*}{0}\right. & 0 & 0 & \overset{*}{2} & 0 & \overset{*}{2} & \overset{*}{2} & 0 & 0 & \overset{*}{2} & \overset{*}{2} & 0 & \overset{*}{2} & 0 & 0 & \overset{*}{4} & \cdots & \\
\cdots & 0 & \overset{*}{2} & \overset{*}{2} & 0 & \overset{*}{2} & 0 & 0 & \overset{*}{4} & \overset{*}{2} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{4} & \overset{*}{4} & 0 & \cdots & \\
\cdots & 0 & \overset{*}{2} & \overset{*}{2} & 0 & \overset{*}{2} & 0 & 0 & \overset{*}{4} & \overset{*}{2} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{4} & \overset{*}{4} & 0 & \cdots & \\
\cdots & \overset{*}{2} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{4} & \overset{*}{4} & 0 & 0 & \overset{*}{4} & \overset{*}{4} & 0 & \overset{*}{4} & 0 & 0 & \left.\overset{*}{6}\right)^\top & & \in \mathcal{E}_{6,2}^{\text{even}}
\end{array}$$

$$\begin{array}{cccccccccccccccccc}
 & \left(0\right. & \overset{*}{1} & \overset{*}{1} & 0 & \overset{*}{1} & 0 & 0 & \overset{*}{3} & \overset{*}{1} & 0 & 0 & \overset{*}{3} & 0 & \overset{*}{3} & \overset{*}{3} & 0 & \cdots & \\
\cdots & \overset{*}{1} & 0 & 0 & \overset{*}{3} & 0 & \overset{*}{3} & \overset{*}{3} & 0 & 0 & \overset{*}{3} & \overset{*}{3} & 0 & \overset{*}{3} & 0 & 0 & \overset{*}{5} & \cdots & \\
\cdots & \overset{*}{1} & 0 & 0 & \overset{*}{3} & 0 & \overset{*}{3} & \overset{*}{3} & 0 & 0 & \overset{*}{3} & \overset{*}{3} & 0 & \overset{*}{3} & 0 & 0 & \overset{*}{5} & \cdots & \\
\cdots & 0 & \overset{*}{3} & \overset{*}{3} & 0 & \overset{*}{3} & 0 & 0 & \overset{*}{5} & \overset{*}{3} & 0 & 0 & \overset{*}{5} & 0 & \overset{*}{5} & \overset{*}{5} & \left.0\right)^\top & & \in \mathcal{E}_{6,2}^{\text{odd}}
\end{array}$$

The vectors with $\overset{*}{a} = *$ if $a = k$ and $\overset{*}{a} = 0$ else constitute $\mathcal{E}_{6,2}^{(k)}$.

(v) $d = 2$, $q = 3$:

$$\begin{array}{ll}
\left(* \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; *\right)^\top \in \mathcal{E}_{2,3}^{\text{even}} & \left(0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0\right)^\top \in \mathcal{E}_{2,3}^{\text{odd}} \\
\left(* \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{2,3}^{(0)} & \left(0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{2,3}^{(1)} \\
\left(0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{2,3}^{(2)} & \left(0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0\right)^\top \in \mathcal{E}_{2,3}^{(3)} \\
\left(0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; *\right)^\top \in \mathcal{E}_{2,3}^{(4)} &
\end{array}$$

(vi) $d = 3$, $q = 3$:

$$\left(* \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(0)}$$
$$\left(0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(2)}$$
$$\left(0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(4)}$$
$$\left(0 \;\; *\right)^\top \in \mathcal{E}_{3,3}^{(6)}$$
$$\left(* \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; *\right)^\top \in \mathcal{E}_{3,3}^{\text{even}}$$

$$\left(0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{\text{odd}}$$
$$\left(0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(1)}$$
$$\left(0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(3)}$$
$$\left(0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; * \;\; 0 \;\; * \;\; 0\right)^\top \in \mathcal{E}_{3,3}^{(5)}$$

(vii) $d = 2$, $q = 4$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(0)}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(2)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(4)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(6)}$$
$$\begin{pmatrix} * & 0 & * & 0 & 0 & * & 0 & * & * & 0 & * & 0 & 0 & * & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{\mathrm{even}}$$

$$\begin{pmatrix} 0 & * & 0 & * & * & 0 & * & 0 & 0 & * & 0 & * & * & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{\mathrm{odd}}$$
$$\begin{pmatrix} 0 & * & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(1)}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(3)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{(5)}$$

(viii) $d = 3$, $q = 4$:

$$\begin{pmatrix} \overset{*}{0} & 0 & \overset{*}{2} & 0 & 0 & \overset{*}{2} & 0 & \overset{*}{4} & \overset{*}{2} & 0 & \overset{*}{4} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{6} & \cdots \\ \cdots & 0 & \overset{*}{2} & 0 & \overset{*}{4} & \overset{*}{2} & 0 & \overset{*}{4} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{6} & \overset{*}{4} & 0 & \overset{*}{6} & 0 & \cdots \\ \cdots & \overset{*}{2} & 0 & \overset{*}{4} & 0 & 0 & \overset{*}{4} & 0 & \overset{*}{6} & \overset{*}{4} & 0 & \overset{*}{6} & 0 & 0 & \overset{*}{6} & 0 & \overset{*}{8} & \cdots \\ \cdots & 0 & \overset{*}{4} & 0 & \overset{*}{6} & \overset{*}{4} & 0 & \overset{*}{6} & 0 & 0 & \overset{*}{6} & 0 & \overset{*}{8} & \overset{*}{6} & 0 & \overset{*}{8} & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,4}^{\mathrm{even}}$$

$$\begin{pmatrix} 0 & \overset{*}{1} & 0 & \overset{*}{3} & \overset{*}{1} & 0 & \overset{*}{3} & 0 & 0 & \overset{*}{3} & 0 & \overset{*}{5} & \overset{*}{3} & 0 & \overset{*}{5} & 0 & \cdots \\ \cdots & \overset{*}{1} & 0 & \overset{*}{3} & 0 & 0 & \overset{*}{3} & 0 & \overset{*}{5} & \overset{*}{3} & 0 & \overset{*}{5} & 0 & 0 & \overset{*}{5} & 0 & \overset{*}{7} & \cdots \\ \cdots & 0 & \overset{*}{3} & 0 & \overset{*}{5} & \overset{*}{3} & 0 & \overset{*}{5} & 0 & 0 & \overset{*}{5} & 0 & \overset{*}{7} & \overset{*}{5} & 0 & \overset{*}{7} & 0 & \cdots \\ \cdots & \overset{*}{3} & 0 & \overset{*}{5} & 0 & 0 & \overset{*}{5} & 0 & \overset{*}{7} & \overset{*}{5} & 0 & \overset{*}{7} & 0 & 0 & \overset{*}{7} & 0 & \overset{*}{9} \end{pmatrix}^\top \in \mathcal{E}_{3,4}^{\mathrm{odd}}$$

The vectors with $\overset{*}{a} = *$ if $a = k$ and $\overset{*}{a} = 0$ else constitute $\mathcal{E}_{3,4}^{(k)}$.

(ix) $d = 2$, $q = 5$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(0)}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(2)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(4)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(6)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(8)}$$
$$\begin{pmatrix} * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{\mathrm{even}}$$

$$\begin{pmatrix} 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{\mathrm{odd}}$$
$$\begin{pmatrix} 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(1)}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(3)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(5)}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{(7)}$$

Example B.2. The vectors contained in $\mathcal{E}_{d,q}^{[k]}$, $k \in \{0, \dots, q-1\}$, see Definition 2.15, have the following structure. An entry $*$ may also be 0.

(i) $d = 2$, $q = 3$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & * & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,3}^{[0]}$$
$$\begin{pmatrix} 0 & * & 0 & * & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,3}^{[1]}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & * & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,3}^{[2]}$$

(ii) $d = 3$, $q = 3$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & * & 0 & * & 0 & 0 & 0 & * & 0 & * & 0 & * & 0 & 0 & 0 & * & 0 & * & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{3,3}^{[0]}$$
$$\begin{pmatrix} 0 & * & 0 & * & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & * & 0 & * & 0 & 0 & 0 & * & 0 & * & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,3}^{[1]}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & * & 0 & * & 0 & 0 & 0 & * & 0 & * & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & * & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{3,3}^{[2]}$$

(iii) $d = 2$, $q = 4$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{[0]}$$
$$\begin{pmatrix} 0 & * & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{[1]}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{[2]}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,4}^{[3]}$$

(iv) $d = 2$, $q = 5$:

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{[0]}$$
$$\begin{pmatrix} 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{[1]}$$
$$\begin{pmatrix} 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{[2]}$$
$$\begin{pmatrix} 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{[3]}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \end{pmatrix}^\top \in \mathcal{E}_{2,5}^{[4]}$$

C. Invertibility of reshaped eigenvectors for $d = 2$

This appendix contains some statements referred to in Section 5.1.

Lemma C.1. *Consider*

$$\begin{aligned}\mathbf{H}_{2,2}^{\mathrm{XYZ}} &= A\mathbf{S}_{x,2}\otimes\mathbf{S}_{x,2} + B\mathbf{S}_{y,2}\otimes\mathbf{S}_{y,2} + \Delta\mathbf{S}_{z,2}\otimes\mathbf{S}_{z,2} + h(\mathbf{S}_{z,2}\otimes\mathbf{I}_2 + \mathbf{I}_2\otimes\mathbf{S}_{z,2}) \\ &= \begin{pmatrix} \frac{\Delta}{4}+h & 0 & 0 & \frac{A-B}{4} \\ 0 & -\frac{\Delta}{4} & \frac{A+B}{4} & 0 \\ 0 & \frac{A+B}{4} & -\frac{\Delta}{4} & 0 \\ \frac{A-B}{4} & 0 & 0 & \frac{\Delta}{4}-h \end{pmatrix},\end{aligned}$$

assume $A \neq B$, and let λ be an eigenvalue of $\mathbf{H}_{2,2}^{\mathrm{XYZ}}$.

(i) If the eigenspace of λ is contained in $\mathcal{E}_{2,2}^{\mathrm{even}}$, then an associated eigenvector is given by

$$\mathbf{v} = \begin{pmatrix} w_1 & 0 & 0 & w_2 \end{pmatrix}^\top$$

with $w_1 = 1$ and $w_2 = \frac{-\Delta-4h+4\lambda}{A-B} \neq 0$.

(ii) If the eigenspace of λ is contained in $\mathcal{E}_{2,2}^{\mathrm{odd}}$, then an associated eigenvector is given by

$$\mathbf{v} = \begin{pmatrix} 0 & w_1 & w_2 & 0 \end{pmatrix}^\top \quad \text{with} \quad \begin{cases} w_1 = 1,\ w_2 = \frac{\Delta+4\lambda}{A+B} \neq 0, & A \neq -B \\ w_1 = w_2 = 1, & A = -B \end{cases}.$$

Proof. Theorem 2.9 ensures the existence of eigenvectors with the sparsity pattern stated in (i) resp. (ii). So it suffices to consider the submatrices of $\mathbf{H}_2 := \mathbf{H}_{2,2}^{\mathrm{XYZ}}$ consisting of rows and columns 1 and 4 resp. 2 and 3.

(i) Setting

$$\mathbf{H}_2^{(1,4)} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}^\top \mathbf{H}_2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\Delta}{4}+h & \frac{A-B}{4} \\ \frac{A-B}{4} & \frac{\Delta}{4}-h \end{pmatrix},$$

it is

$$\begin{aligned}\det\left(\lambda\mathbf{I}_2 - \mathbf{H}_2^{(1,4)}\right) &= \det\begin{pmatrix} \lambda - (\frac{\Delta}{4}+h) & -\frac{A-B}{4} \\ -\frac{A-B}{4} & \lambda - (\frac{\Delta}{4}-h) \end{pmatrix} \\ &= \lambda^2 - \frac{\Delta}{2}\lambda + \frac{\Delta^2}{16} - h^2 - \frac{(A-B)^2}{16}\end{aligned}$$

with zeros

$$\lambda_{1,2}(\mathbf{H}_2^{(1,4)}) = \frac{\Delta}{4} \pm \sqrt{h^2 + \frac{(A-B)^2}{16}}.$$

For $\lambda = \lambda_{1,2}(\mathbf{H}_2^{(1,4)})$ we obtain that

$$\left(\mathbf{H}_2^{(1,4)} - \lambda \mathbf{I}_2\right) \mathbf{w} = \begin{pmatrix} \frac{\Delta}{4} + h - \lambda & \frac{A-B}{4} \\ \frac{A-B}{4} & \frac{\Delta}{4} - h - \lambda \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a solution

$$\mathbf{w} = \begin{pmatrix} 1 \\ \frac{-\Delta - 4h + 4\lambda}{A-B} \end{pmatrix}.$$

Furthermore, $-\Delta - 4h + 4\lambda = -4h \pm 4\sqrt{h^2 + \frac{(A-B)^2}{16}}$ and

$$-4h - 4\sqrt{h^2 + \frac{(A-B)^2}{16}} < -4h - 4\sqrt{h^2} \leq 0 \leq -4h + 4\sqrt{h^2} < -4h + 4\sqrt{h^2 + \frac{(A-B)^2}{16}},$$

hence $w_2 \neq 0$.

(ii) Setting

$$\mathbf{H}_2^{(2,3)} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}^\top \mathbf{H}_2 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{\Delta}{4} & \frac{A+B}{4} \\ \frac{A+B}{4} & -\frac{\Delta}{4} \end{pmatrix},$$

it is

$$\det\left(\lambda \mathbf{I}_2 - \mathbf{H}_2^{(2,3)}\right) = \det \begin{pmatrix} \lambda + \frac{\Delta}{4} & -\frac{A+B}{4} \\ -\frac{A+B}{4} & \lambda + \frac{\Delta}{4} \end{pmatrix} = \lambda^2 + \frac{\Delta}{2}\lambda + \frac{\Delta^2}{16} - \frac{(A+B)^2}{16}$$

with zeros

$$\lambda_{1,2}(\mathbf{H}_2^{(2,3)}) = -\frac{\Delta}{4} \pm \frac{A+B}{4}.$$

For $\lambda = \lambda_{1,2}(\mathbf{H}_2^{(2,3)})$ we obtain that

$$\left(\mathbf{H}_2^{(2,3)} - \lambda \mathbf{I}_2\right) \mathbf{w} = \begin{pmatrix} -\frac{\Delta}{4} - \lambda & \frac{A+B}{4} \\ \frac{A+B}{4} & -\frac{\Delta}{4} - \lambda \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has, if $A \neq -B$, a solution

$$\mathbf{w} = \begin{pmatrix} 1 \\ \frac{\Delta + 4\lambda}{A+B} \end{pmatrix}.$$

Furthermore, $\Delta + 4\lambda = \pm(A+B) \neq 0$, hence $w_2 \neq 0$. If instead $A = -B$, then each vector in $\mathbb{R}^2$, in particular $\mathbf{w} = (1,1)^\top$, is a solution since $\lambda = -\frac{\Delta}{4}$.

□

Lemma C.2. *The equations* (5.21) *and* (5.22) *hold.*

Proof. The characteristic polynomial of $\mathbf{H}_2^{(1,3,5,7,9)}$ is

$$\begin{aligned}\det\left(\lambda\mathbf{I}_5-\mathbf{H}_2^{(1,3,5,7,9)}\right)&=\det\begin{pmatrix}\lambda-2h & 0 & -\frac{A}{2} & 0 & 0\\ 0 & \lambda & -\frac{A}{2} & 0 & 0\\ -\frac{A}{2} & -\frac{A}{2} & \lambda & -\frac{A}{2} & -\frac{A}{2}\\ 0 & 0 & -\frac{A}{2} & \lambda & 0\\ 0 & 0 & -\frac{A}{2} & 0 & \lambda+2h\end{pmatrix}\\&=\lambda\left(\lambda^4-\left(A^2+4h^2\right)\lambda^2+2A^2h^2\right)\end{aligned}$$

with zeros

$$\pm\sqrt{\frac{A^2}{2}+2h^2\pm\sqrt{\left(\frac{A^2}{2}+2h^2\right)^2-2A^2h^2}}=\pm\sqrt{\frac{A^2}{2}+2h^2\pm\sqrt{\frac{A^4}{4}+4h^4}}$$

and 0. Ordering these zeros in an increasing manner yields (5.21). Furthermore,

$$\begin{aligned}\det\left(\lambda\mathbf{I}_4-\mathbf{H}_2^{(2,4,6,8)}\right)&=\det\begin{pmatrix}\lambda-h & -\frac{A}{2} & -\frac{A}{2} & 0\\ -\frac{A}{2} & \lambda-h & 0 & -\frac{A}{2}\\ -\frac{A}{2} & 0 & \lambda+h & -\frac{A}{2}\\ 0 & -\frac{A}{2} & -\frac{A}{2} & \lambda+h\end{pmatrix}\\&=\lambda^4-\left(A^2+2h^2\right)\lambda^2+h^4\end{aligned}$$

with zeros

$$\pm\sqrt{\frac{A^2}{2}+h^2\pm\sqrt{\left(\frac{A^2}{2}+h^2\right)^2-h^4}}=\pm\sqrt{\frac{A^2}{2}+h^2\pm\sqrt{\frac{A^4}{4}+A^2h^2}}$$

and ordering these zeros in an increasing manner yields (5.22). □

Lemma C.3. *Let* $\lambda=\lambda_i(\mathbf{H}_2^{(1,3,5,7,9)})$, $i\in\{1,\dots,5\}$, *see* (5.21), *be an eigenvalue of* $\mathbf{H}_2:=\mathbf{H}_{2,3}^{\mathrm{XYZ}}\in\mathbb{R}^{9\times9}$ *in case* $A\neq0\neq h$, $B=\Delta=0$. *Then an associated eigenvector is given by*

$$\mathbf{v}=\begin{pmatrix}w_1 & 0 & w_2 & 0 & w_3 & 0 & w_4 & 0 & w_5\end{pmatrix}^\top$$

with $w_2=1,w_4=-1,w_1=w_3=w_5=0$ *for* $i=3$ *and*

$$w_1=\frac{-A}{4h-2\lambda},\quad w_2=\frac{A}{2\lambda}=w_4,\quad w_3=1,\quad w_5=\frac{A}{4h+2\lambda}$$

for $i\in\{1,2,4,5\}$.

Proof. Due to the construction (5.20) of $\mathbf{H}_2^{(1,3,5,7,9)}$ motivated by Theorem 2.9, an eigenvector $\mathbf{v}\in\mathbb{R}^9$ of $\mathbf{H}_{2,3}^{\mathrm{XYZ}}$ associated with $\lambda=\lambda_i(\mathbf{H}_2^{(1,3,5,7,9)})$, $i\in\{1,\dots,5\}$, is an element of $\mathcal{E}_{2,3}^{\mathrm{even}}$, i.e. has the sparsity structure

$$\mathbf{v}=\begin{pmatrix}*&0&*&0&*&0&*&0&*\end{pmatrix}^\top.$$

C. Invertibility of reshaped eigenvectors for $d = 2$

Consider

$$\mathbf{H}_2^{(1,3,5,7,9)}\mathbf{u} = \mathbf{P}_{(1,3,5,7,9)}^{\top}\mathbf{H}_2\mathbf{P}_{(1,3,5,7,9)}\mathbf{u} = \begin{pmatrix} 2h & 0 & \frac{A}{2} & 0 & 0 \\ 0 & 0 & \frac{A}{2} & 0 & 0 \\ \frac{A}{2} & \frac{A}{2} & 0 & \frac{A}{2} & \frac{A}{2} \\ 0 & 0 & \frac{A}{2} & 0 & 0 \\ 0 & 0 & \frac{A}{2} & 0 & -2h \end{pmatrix} \mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where $\mathbf{P}_{(1,3,5,7,9)} \in \mathbb{R}^{9\times 5}$ contains a one in the entries $(1,1)$, $(3,2)$, $(5,3)$, $(7,4)$, $(9,5)$ and zeros elsewhere. A solution is given by $\mathbf{u} = (0,1,0,-1,0)^{\top}$, so $\operatorname{span}\{(0,1,0,-1,0)^{\top}\}$ is the eigenspace associated with the eigenvalue $\lambda_3(\mathbf{H}_2^{(1,3,5,7,9)}) = 0$. This yields already $w_2 = w_4$ for all $\mathbf{w} \perp \mathbf{u}$.

Let λ be a nonzero eigenvalue of $\mathbf{H}_2^{(1,3,5,7,9)}$. By some steps of Gaussian elimination,

$$\left(\mathbf{H}_2^{(1,3,5,7,9)} - \lambda\mathbf{I}_5\right)\mathbf{w} = \begin{pmatrix} 2h-\lambda & 0 & \frac{A}{2} & 0 & 0 \\ 0 & -\lambda & \frac{A}{2} & 0 & 0 \\ \frac{A}{2} & \frac{A}{2} & -\lambda & \frac{A}{2} & \frac{A}{2} \\ 0 & 0 & \frac{A}{2} & -\lambda & 0 \\ 0 & 0 & \frac{A}{2} & 0 & -2h-\lambda \end{pmatrix} \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

is transformed into

$$\begin{pmatrix} 2h-\lambda & 0 & \frac{A}{2} & 0 & 0 \\ 0 & -\lambda & \frac{A}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda & 2h+\lambda \\ 0 & 0 & \frac{A}{2} & 0 & -2h-\lambda \end{pmatrix} \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Choosing $w_3 = 1$ we infer

$$w_5 = \frac{A}{4h+2\lambda}, \quad w_4 = \frac{A}{2\lambda} = w_2, \quad w_1 = \frac{-A}{4h-2\lambda}.$$

□

The following statement, which we cite from the literature without proof, and the resulting corollary are employed in the discussion of the invertibility of the matricization of $\mathbf{v}_{\min}^{(2)}$ for the 3-Potts model.

Proposition C.4 ([DPTZ22, Theorem 1])**.** *Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be Hermitian with eigenvalues $\lambda_i(\mathbf{A})$, $1 \le i \le n$. Let $\mathbf{M}_j \in \mathbb{C}^{n-1\times n-1}$, $1 \le j \le n$, be that principal submatrix of $\mathbf{A}$ which is formed by deleting the j-th row and column, with eigenvalues $\lambda_i(\mathbf{M}_j)$, $1 \le i \le n-1$. Let $v_{i,j}$ be the j-th component of a normalized eigenvector $\mathbf{v}_i$ associated with $\lambda_i(\mathbf{A})$. Then*

$$|v_{i,j}|^2 \prod_{k=1;\, k\neq i}^{n} (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{A})) = \prod_{k=1}^{n-1} (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{M}_j)).$$

Corollary C.5. *A consequence of Proposition C.4 is that if*

$$\lambda_i(\mathbf{A}) \neq \lambda_k(\mathbf{M}_j) \text{ for all } j,k,$$

then

$$v_{i,j} \neq 0 \text{ for all } j.$$

Now we verify (5.32).

Lemma C.6. *Let $h \in \mathbb{R}$. For*

$$\begin{aligned}\widehat{\chi}^{(1,6,8)}(\lambda) &:= \lambda^2 + (2h+1)\lambda - 2(4h^2 - 2h + 1) \\ &= \left(\lambda - \left(-\left(h+\frac{1}{2}\right) - 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right)\left(\lambda - \left(-\left(h+\frac{1}{2}\right) + 3\sqrt{\left(h-\frac{1}{6}\right)^2 + \frac{2}{9}}\right)\right)\end{aligned}$$

and

$$\lambda_{\min}(\mathbf{H}_2^{(2,4,9)}) := \frac{1}{2}\left(-(-h+1) - 3\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}\right),$$

it is

$$\widehat{\chi}^{(1,6,8)}\left(\lambda_{\min}(\mathbf{H}_2^{(2,4,9)})\right) \begin{cases} < 0 & , \ h \neq 0 \\ = 0 & , \ h = 0 \end{cases}.$$

Proof. If $h = 0$, we have $\widehat{\chi}^{(1,6,8)}(\lambda) = \lambda^2 + \lambda - 2$ and $\lambda_{\min}(\mathbf{H}_2^{(2,4,9)}) = -2$, hence

$$\widehat{\chi}^{(1,6,8)}\left(\lambda_{\min}(\mathbf{H}_2^{(2,4,9)})\right) = 0.$$

So, let in the following $h \neq 0$. A straightforward calculation yields

$$\widehat{\chi}^{(1,6,8)}\left(\frac{1}{2}\left(-(-h+1) - 3\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}\right)\right) = -\frac{9}{2}h^2 + \frac{9}{2}h - \frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}}.$$

We distinguish four cases.

- $\underline{h > 0:}$ It is

  $$\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}} > 1,$$

  hence

  $$-\frac{9}{2}h^2 + \frac{9}{2}h - \frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}} < \frac{9}{2}h - \frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}} < \frac{9}{2}h - \frac{9}{2}h = 0.$$

- $\underline{h < -\frac{3}{2}:}$ In this case we have

  $$\left(h+\frac{1}{3}\right)^2 + \frac{8}{9} = h^2 + \frac{2}{3}h + 1 < h^2,$$

  which is equivalent to

  $$\sqrt{\left(h+\frac{1}{3}\right)^2 + \frac{8}{9}} < |h| = -h$$

and implies

$$\begin{aligned}-\frac{9}{2}h^2+\frac{9}{2}h-\frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}} &< -\frac{9}{2}h^2-\frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\\ &= -\frac{9}{2}h\left(h+\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\right)\\ &= \frac{9}{2}|h|\left(-|h|+\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\right)\\ &< \frac{9}{2}|h|\left(-|h|+|h|\right)=0.\end{aligned}$$

- $\underline{-\frac{3}{2}\le h\le -\frac{2}{3}}$: On the interval $\left[-\frac{3}{2},-\frac{2}{3}\right]$, the function $\varphi_1(h):=-h+1$ is monotonically decreasing with maximum $\varphi_1(-\frac{3}{2})=\frac{5}{2}$ and minimum $\varphi_1(-\frac{2}{3})=\frac{5}{3}$, just as the function $\varphi_2(h):=\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}$ is monotonically decreasing with maximum $\varphi_2(-\frac{3}{2})=\frac{3}{2}<\min\{\varphi_1(h)\}$ and minimum $\varphi_2(-\frac{2}{3})=1$. Hence

$$\varphi_1(h)-\varphi_2(h)=-h+1-\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}>0$$

and thus

$$\begin{aligned}-\frac{9}{2}h^2+\frac{9}{2}h-\frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}} &= \frac{9}{2}h\left(-h+1-\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\right)\\ &= -\frac{9}{2}|h|\left(-h+1-\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\right)<0.\end{aligned}$$

- $\underline{-\frac{2}{3}<h<0}$: Here it is

$$\left(h+\frac{1}{3}\right)^2<\frac{1}{9},$$

which implies

$$\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}<1$$

and so

$$\begin{aligned}-\frac{9}{2}h^2+\frac{9}{2}h-\frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}} &< \frac{9}{2}h-\frac{9}{2}h\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\\ &= \frac{9}{2}|h|\left(-1+\sqrt{\left(h+\frac{1}{3}\right)^2+\frac{8}{9}}\right)<0.\end{aligned}$$

□

List of algorithms

4.1. LOCG in full vector format 54
4.2. LOCG for tensors in HT format 55
4.3. MALS for eigenvalue problems 61

List of figures

3.1. Diagrammatic notation of tensors 34
3.2. Contraction of fourth-order tensor with fifth-order tensor 35

5.1. Approximation quality of prolongated vector for 2-Ising model with varying h 70
5.2. Approximation quality of prolongated vector for 2-Ising model with varying (d_1, d_2) and classical inverse or truncated pseudoinverse 72
5.3. Sparsity pattern of $\mathbf{v}_{\min}^{(2)}$ for 3-XYZ model with 10,000 times randomly chosen $A \neq B$, Δ, h 75
5.4. Approximation quality of prolongated vector for 3-Ising model with varying h 76
5.5. Strategies to construct prolongated vector in case $\mathbf{v}_{\min}^{(2)} \in \mathcal{E}_{2,3}^{\text{odd}}$ 77
5.6. Approximation quality of prolongated vector for 3-Potts model with varying h 81

6.1. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree, $r_{\max} = 100$, $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ 112
6.2. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree, $r_{\max} \in \{100, qd\}$, $\varepsilon_{\text{rel}} \in \{10^{-9}, 10^{-12}\}$ 113
6.3. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree, $r_{\max} = 100$, $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ 114
6.4. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree, $r_{\max} \in \{100, qd\}$, $\varepsilon_{\text{rel}} \in \{10^{-9}, 10^{-12}\}$ 115
6.5. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, gradient descent, linear tree, $r_{\max} = 100$, $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ 116
6.6. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, gradient descent, linear tree, $r_{\max} \in \{100, qd\}$, $\varepsilon_{\text{rel}} \in \{10^{-9}, 10^{-12}\}$ 117
6.7. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (1, 0, 0, -0.7)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree, $r_{\max} = 100$, $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ 118
6.8. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (1, 0, 0, -0.7)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree, $r_{\max} \in \{100, qd\}$, $\varepsilon_{\text{rel}} \in \{10^{-9}, 10^{-12}\}$ 119
6.9. 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.6, -0.3)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree, $r_{\max} = 100$, $\varepsilon_{\text{rel}} \in \{10^{-3}, 10^{-6}\}$ 120
6.10. 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.6, -0.3)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree, $r_{\max} \in \{100, qd\}$, $\varepsilon_{\text{rel}} \in \{10^{-9}, 10^{-12}\}$ 121
6.11. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 3)$, LOCG, linear tree 122
6.12. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, LOCG, linear tree 123
6.13. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (3, 4)$, LOCG, linear tree 124
6.14. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 9)$, LOCG, linear tree 125

6.15. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (8, 9)$, LOCG, linear tree . . . 126
6.16. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 3)$, LOCG, balanced tree . . . 127
6.17. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 4)$, LOCG, balanced tree . . . 128
6.18. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (3, 4)$, LOCG, balanced tree . . . 129
6.19. 3-XYZ, $d = 10$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 4)$, LOCG, linear vs. balanced tree . . . 130
6.20. 3-XYZ, $d = 14$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 6)$, LOCG, linear vs. balanced tree . . . 131
6.21. 3-XYZ, $d = 14$, $(A, B, \Delta, h) = (2.6, 0.7, -1.9, -0.3)$, $(d_1, d_2) = (5, 6)$, LOCG, linear vs. balanced tree . . . 132
6.22. 3-Potts, $d = 14$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) \in \{(2, 3), (5, 6)\}$, LOCG, balanced tree . . . 133
6.23. 3-Potts, $d = 10$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (2, 3)$, LOCG vs. gradient descent, linear tree . . . 134
6.24. 3-Potts, $d = 14$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (3, 4)$, LOCG vs. gradient descent, linear tree . . . 135
6.25. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 3)$, MALS . . . 136
6.26. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 9)$, MALS . . . 139
6.27. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (8, 9)$, MALS . . . 140
6.28. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 3)$, MALS . . . 141
6.29. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (2, 4)$, MALS . . . 142
6.30. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (-0.9, 1.6, 1.7, 1.2)$, $(d_1, d_2) = (3, 4)$, MALS . . . 143
6.31. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) \in \{(2, 4), (2, 6)\}$, MALS . . . 144
6.32. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (2.6, 0.7, -1.9, -0.3)$, $(d_1, d_2) = (5, 6)$, MALS . . . 145
6.33. 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (2, 3)$, MALS . . . 146
6.34. 3-Potts, $d \in \{10, 14\}$, $(A, h) = (1.3, 0.9)$, $(d_1, d_2) = (3, 4)$, MALS . . . 147
6.35. 2-XYZ, $d \in \{100, 300\}$, $(A, B, \Delta, h) = (1.9, 0.4, -1.1, 0.2)$, $(d_1, d_2) = (2, 4)$, MALS . . . 148
6.36. 3-XYZ, $d \in \{100, 300\}$, $(A, B, \Delta, h) = (1.6, -0.3, 2.9, -0.8)$, $(d_1, d_2) = (2, 4)$, MALS . . . 149
6.37. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (3, 4)$, LOCG, balanced tree . . . 150
6.38. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (8, 9)$, LOCG, linear tree . . . 153
6.39. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, LOCG, linear tree . . . 154
6.40. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, LOCG, balanced tree . . . 155

6.41. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (2, 2, 1, 1)$, LOCG, linear tree 156
6.42. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (2, 2, 1, 1)$, LOCG, balanced tree 157
6.43. 2-XYZ, $d \in \{16, 22\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, MALS 158
6.44. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (2, 2, 1, 1)$, MALS 159
6.45. 2-XYZ, $d \in \{16, 23\}$, $(A, B, \Delta, h) = (1, 1, -0.2, 0.3)$, $(d_1, d_2) = (8, 9)$, MALS . 160
6.46. 3-XYZ, $d \in \{10, 14\}$, $(A, B, \Delta, h) = (2, 2, 1, 1)$, $(d_1, d_2) = (3, 4)$, MALS 161

Bibliography

[Arb16] Arbenz, Peter: *Lecture notes on solving large scale eigenvalue problems*, 2016. ETH Zürich, available at https://people.inf.ethz.ch/arbenz/ewp/Lnotes/lsevp.pdf.

[AS65] Abramowitz, Milton and Irene A. Stegun (editors): *Handbook of Mathematical Functions*. Dover, New York, 1965.

[Bac23] Bachmayr, Markus: *Low-rank tensor methods for partial differential equations*. Acta Numerica, 32:1–121, 2023.

[Bal15] Ballentine, Leslie E.: *Quantum Mechanics: A Modern Development*. World Scientific, New Jersey, London, Singapore, 2nd edition, 2015.

[Bax07] Baxter, Rodney J.: *Exactly Solved Models in Statistical Mechanics*. Dover, Mineola, New York, 2007.

[BBF17] Bonfim, Oz F. de Alcantara, Beatriz Boechat, and João Florencio: *Quantum fidelity approach to the ground-state properties of the one-dimensional axial next-nearest-neighbor Ising model in a transverse field*. Phys. Rev. E, 96(4):042140, 2017.

[BBF19] Bonfim, Oz F. de Alcantara, Beatriz Boechat, and João Florencio: *Ground-state properties of the one-dimensional transverse Ising model in a longitudinal magnetic field*. Phys. Rev. E, 99(1):012122, 2019.

[BFSB14] Boechat, Beatriz, João Florencio, Andreia Saguia, and Oz F. de Alcantara Bonfim: *Critical behavior of a quantum chain with four-spin interactions in the presence of longitudinal and transverse magnetic fields*. Phys. Rev. E, 89(3):032143, 2014.

[Can95] Cannas, Sergio A.: *One-dimensional Ising model with long-range interactions: A renormalization-group treatment*. Phys. Rev. B, 52(5):3034–3037, 1995.

[Dir39] Dirac, Paul A. M.: *A new notation for quantum mechanics*. Math. Proc. Cambridge Philos. Soc., 35(3):416–418, 1939.

[DLDMV00] De Lathauwer, Lieven, Bart De Moor, and Joos Vandewalle: *A multilinear singular value decomposition*. SIAM J. Matrix Anal. Appl., 21(4):1253–1278, 2000.

[DPTZ22] Denton, Peter B., Stephen J. Parke, Terence Tao, and Xining Zhang: *Eigenvectors from eigenvalues: A survey of a basic identity in linear algebra*. Bull. Amer. Math. Soc. (N.S.), 59(1):31–58, 2022.

[Eck19] Eckle, Hans-Peter: *Models of Quantum Matter: A First Course on Integrability and the Bethe Ansatz*. Oxford University Press, Oxford, 2019.

[EY36] Eckart, Carl and Gale Young: *The approximation of one matrix by another of lower rank*. Psychometrika, 1(3):211–218, 1936.

[GKT13] Grasedyck, Lars, Daniel Kressner, and Christine Tobler: *A literature survey of low-rank tensor approximation techniques*. GAMM-Mitt., 36(1):53–78, 2013.

[Gra10] Grasedyck, Lars: *Hierarchical singular value decomposition of tensors*. SIAM J. Matrix Anal. Appl., 31(4):2029–2054, 2010.

[Gri14] Griffiths, David J.: *Introduction to Quantum Mechanics*. Pearson, Harlow, 2nd edition, 2014.

[GSR+14] Gerster, Matthias, Pietro Silvi, Matteo Rizzi, Rosario Fazio, Tommaso Calarco, and Simone Montangero: *Unconstrained tree tensor network: An adaptive gauge picture for enhanced performance*. Phys. Rev. B, 90(12):125154, 2014.

[GVL13] Golub, Gene H. and Charles F. Van Loan: *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 4th edition, 2013.

[Hac85] Hackbusch, Wolfgang: *Multi-Grid Methods and Applications*. Springer, Berlin, Heidelberg, 1985.

[Hac19] Hackbusch, Wolfgang: *Tensor Spaces and Numerical Tensor Calculus*. Springer Nature Switzerland, Cham, 2nd edition, 2019.

[HB09] Hanke-Bourgeois, Martin: *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Vieweg+Teubner, Wiesbaden, 3rd edition, 2009.

[Hei28] Heisenberg, Werner: *Zur Theorie des Ferromagnetismus*. Z. Physik, 49:619–636, 1928.

[HJ91] Horn, Roger A. and Charles R. Johnson: *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, New York, Melbourne, 1991.

[HJ13] Horn, Roger A. and Charles R. Johnson: *Matrix Analysis*. Cambridge University Press, Cambridge, New York, Melbourne, 2nd edition, 2013.

[HK09] Hackbusch, Wolfgang and Stefan Kühn: *A new scheme for the tensor representation*. J. Fourier Anal. Appl., 15:706–722, 2009.

[HPS83] Henderson, Harold V., Friedrich Pukelsheim, and Shayle R. Searle: *On the history of the Kronecker product*. Linear Multilinear Algebra, 14(2):113–120, 1983.

[HRS12] Holtz, Sebastian, Thorsten Rohwedder, and Reinhold Schneider: *The alternating linear scheme for tensor optimization in the tensor train format*. SIAM J. Sci. Comput., 34(2):A683–A713, 2012.

[HWSH13a] Huckle, Thomas, Konrad Waldherr, and Thomas Schulte-Herbrüggen: *Computations in quantum tensor networks.* Linear Algebra Appl., 438(2):750–781, 2013.

[HWSH13b] Huckle, Thomas, Konrad Waldherr, and Thomas Schulte-Herbrüggen: *Exploiting matrix symmetries and physical symmetries in matrix product states and tensor trains.* Linear Multilinear Algebra, 61(1):91–122, 2013.

[KB09] Kolda, Tamara G. and Brett W. Bader: *Tensor decompositions and applications.* SIAM Rev., 51(3):455–500, 2009.

[Kny01] Knyazev, Andrew V.: *Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method.* SIAM J. Sci. Comput., 23(2):517–541, 2001.

[KT14] Kressner, Daniel and Christine Tobler: *Algorithm 941:* `htucker` *– a* Matlab *toolbox for tensors in hierarchical Tucker format.* ACM Trans. Math. Softw., 40(3):Article 22, 2014.

[LMJ18] Lubasch, Michael, Pierre Moinier, and Dieter Jaksch: *Multigrid renormalization.* J. Comput. Phys., 372:587–602, 2018.

[LRV04] Latorre, José I., Enrique Rico, and Guifré Vidal: *Ground state entanglement in quantum spin chains.* arXiv:quant-ph/0304098v4, 2004.

[Mir60] Mirsky, Leonid: *Symmetric gauge functions and unitarily invariant norms.* Quart. J. Math. Oxford, 11(1):50–59, 1960.

[Nac04] Nachtergaele, Bruno: *Quantum spin systems.* arXiv:math-ph/0409006v1, 2004.

[NO08] Nakahara, Mikio and Tetsuo Ohmi: *Quantum Computing: From Linear Algebra to Physical Realization.* CRC Press, Boca Raton, London, New York, 2008.

[OD12] Oseledets, Ivan and Sergey Dolgov: *Solution of linear systems and matrix inversion in the TT-format.* SIAM J. Sci. Comput., 34(5):A2718–A2739, 2012.

[ODK+12] Oseledets, Ivan, Sergey Dolgov, Vladimir Kazeev, Thomas Mach, Olga Lebedeva, Dmitry Savostyanov, Pavel Zhlobich, and Le Song: *TT-Toolbox*, 2009-2012. GitHub repository, available at https://github.com/oseledets/TT-Toolbox.

[Orú14] Orús, Román: *A practical introduction to tensor networks: Matrix product states and projected entangled pair states.* Ann. Physics, 349:117–158, 2014.

[Orú19] Orús, Román: *Tensor networks for complex quantum systems.* Nat. Rev. Phys., 1:538–550, 2019.

[Ose10] Oseledets, Ivan: *Approximation of* $2^d \times 2^d$ *matrices using tensor decomposition.* SIAM J. Matrix Anal. Appl., 31(4):2130–2145, 2010.

[Ose11] Oseledets, Ivan: *Tensor-train decomposition.* SIAM J. Sci. Comput., 33(5):2295–2317, 2011.

[Par98] Parlett, Beresford N.: *The Symmetric Eigenvalue Problem.* SIAM, Philadelphia, 1998.

[PESV15] Pfeifer, Robert N. C., Glen Evenbly, Sukhwinder Singh, and Guifré Vidal: *NCON: A tensor network contractor for MATLAB.* arXiv:1402.0939v3 [physics.comp-ph], 2015.

[PF10] Parkinson, John B. and Damian J. J. Farnell: *An Introduction to Quantum Spin Systems.* Lect. Notes Phys. 816. Springer, Berlin, Heidelberg, 2010.

[Pot52] Potts, Renfrey B.: *Some generalized order-disorder transformations.* Math. Proc. Cambridge Philos. Soc., 48(1):106–109, 1952.

[Raa17] Raasch, Thorsten: *Numerische multilineare Algebra.* Vorlesungsskript, Universität Mainz, 2017.

[Saa11] Saad, Yousef: *Numerical Methods for Large Eigenvalue Problems.* SIAM, Philadelphia, 2nd edition, 2011.

[Sch07] Schmidt, Erhard: *Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener.* Math. Ann., 63:433–476, 1907.

[Sch11] Schollwöck, Ulrich: *The density-matrix renormalization group in the age of matrix product states.* Ann. Physics, 326:96–192, 2011.

[SGL+20] Secular, Paul, Nikita Gourianov, Michael Lubasch, Sergey Dolgov, Stephen R. Clark, and Dieter Jaksch: *Parallel time-dependent variational principle algorithm for matrix product states.* Phys. Rev. B, 101(23):235123, 2020.

[SP81] Sólyom, Jenő and Pierre Pfeuty: *Renormalization-group study of the Hamiltonian version of the Potts model.* Phys. Rev. B, 24(1):218–229, 1981.

[Ste93] Stewart, Gilbert W.: *On the early history of the singular value decomposition.* SIAM Rev., 35(4):551–566, 1993.

[STG+19] Silvi, Pietro, Ferdinand Tschirsich, Matthias Gerster, Johannes Jünemann, Daniel Jaschke, Matteo Rizzi, and Simone Montangero: *The tensor networks anthology: Simulation techniques for many-body quantum lattice systems.* SciPost Phys. Lect. Notes, 8:1–106, 2019.

[TB97] Trefethen, Lloyd N. and David Bau, III: *Numerical Linear Algebra.* SIAM, Philadelphia, 1997.

[Tob12] Tobler, Christine: *Low-rank Tensor Methods for Linear Systems and Eigenvalue Problems.* PhD thesis, ETH Zürich, 2012.

[Tow00] Townsend, John S.: *A Modern Approach to Quantum Mechanics.* University Science Books, Sausalito, California, 2000.

[VLRK03] Vidal, Guifré, José I. Latorre, Enrique Rico, and Alexei Kitaev: *Entanglement in quantum critical phenomena.* Phys. Rev. Lett., 90(22):227902, 2003.

[VMC08] Verstraete, Frank, Valentin Murg, and J. Ignacio Cirac: *Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems.* Adv. Phys., 57(2):143–224, 2008.

[Whi92] White, Steven R.: *Density matrix formulation for quantum renormalization groups.* Phys. Rev. Lett., 69(19):2863–2866, 1992.