Cauchy matrix approach to the SU(2) self-dual Yang–Mills equation

Shangshuai Li1 | Changzheng Qu2 | Xiangxuan Yi3 | Da-jun Zhang1

1 Department of Mathematics, Shanghai University, Shanghai, China
2 School of Mathematics and Statistics, Ningbo University, Ningbo, China
3 Qianweichang College, Shanghai University, Shanghai, China

Correspondence
Da-jun Zhang, Department of Mathematics, Shanghai University, Shanghai 200444, China.
Email: djzhang@staff.shu.edu.cn

Funding information
Science and Technology Commission of Shanghai Municipality; National Natural Science Foundation of China

Abstract
The Cauchy matrix approach is developed to solve the SU(2) self-dual Yang–Mills (SDYM) equation. Starting from a Sylvester matrix equation coupled with certain dispersion relations for an infinite number of coordinates, we derive some new relations that give rise to the SDYM equation under Yang’s formulation. By imposing further constraints on complex independent variables, a broad class of explicit solutions of the equation under Yang’s formulation are obtained.

KEYWORDS
Cauchy matrix approach, integrable system, self-dual Yang–Mills equation, solution

1 INTRODUCTION

The Yang–Mills theory is the most important development in physics in the second half of last century (see Ref. 1). The pioneer work is that of Yang and Mills,2 which laid a foundation of non-abelian gauge theories that explain the electromagnetic, the strong, and weak nuclear interactions. The theory has also geometric interpretations in its nature, where the concept of gauge fields is found to be identical to fiber bundles, e.g., Refs. 3–6 (also see Refs. 7, 8 and the references therein).

The Yang–Mills equation of motion is9 (please refer Section 2 for notations)

$$\partial_\mu F_{\mu\nu} + [B_\mu, F_{\mu\nu}] = [D_\mu, F_{\mu\nu}] = 0,$$

(1)
which corresponds to the case where action functional $S$ and energy functional $E$ take local minima in the sense of semiclassical approximation. It is hard to solve this equation exactly. Even if one could find solutions to (1), it is still difficult to verify $S$ or $E$ reaches local minima. However, it can be shown that (see Refs. 9, 10) when the gauge field strength $F_{\mu \nu}$ is self-dual, for a given integer topological charge $q$, the gauge fields are absolute minima of the action functional $S$. The corresponding solutions are called instantons and monopoles (for static gauge fields). Ref. 9 collected comprehensively the approaches earlier than 1980 to solutions of the self-dual Yang–Mills (SDYM) equation, such as the approach based on the so-called Corrigan–Fairlie–’t Hooft–Wilczek ansatz,11,12 the Atiyah–Hitchin–Drinfeld–Manin construction13 on instantons with integer topological charge $q$ and depending on (for $SU(N)$ case) at least $8q − 3$ parameters,14 and the Bäcklund transformation approach15 based on the Atiyah–Ward ansatz.16 Both Refs. 13 and 16 follow Ward’s observation in 197717 on the connection between self-dual gauge fields and twistor theory. One can also refer to Ref. 9 and the references therein for more details.

The SDYM equation is an integrable system. It has a Lax pair and has Painlevé property for any gauge group,18,19 therefore some methods based on integrability have been employed to solve the SDYM equation, such as a direct transform approach using Lax pair,20 Bäcklund transformation based on Riemann-Hilbert problem,21 an approach inspired from Sato’s theory,22 bilinear method,23,24 and Darboux transformation.25

In this paper, we aim to solve the $SU(2)$ SDYM equation via a direct method, the Cauchy matrix approach. The Cauchy matrix approach has been used to construct and study integrable equations based on the Sylvester-type equations. In this approach, integrable equations are presented as closed forms of some recurrence relations involving derivatives (or shifts). It was first used systematically in Ref. 26 to investigate integrable quadrilateral equations and later developed in Refs. 27, 28 to more general cases. We will construct explicit solutions to the $SU(2)$ SDYM equation.

The paper is organized as follows. In Section 2 we recall Yang’s formulation for the self-duality of the Yang–Mills equation. Then, in Section 3 we make use of the Cauchy matrix approach to construct solutions to a SDYM equation. These solutions will be elaborated in Section 4 by imposing constraints so that they meet self-duality of $SU(2)$ gauge group. Finally, concluding remarks are given in Section 5. There are three appendices, which are devoted to proving Lemma 2, presenting solutions of the Sylvester equation (16), and presenting examples of solutions of the SDYM equation.

## 2 | THE SDYM EQUATION: YANG’S FORMULATION

Let $B_\mu$ be matrix valued gauge potentials defined on $\mathbb{R}^4$ and $F_{\mu \nu}$ the gauge field strength defined by

$$ F_{\mu \nu} \equiv \partial_\nu B_\mu - \partial_\mu B_\nu - [B_\mu, B_\nu], $$

(2)

where $[\cdot, \cdot]$ is the Lie bracket defined as $[G, H] = GH - HG$, and $\partial_\mu$ stands for the differential operator $\partial/\partial x^\mu, (x^0, x^1, x^2, x^3) \in \mathbb{R}^4$. For a given $SU(N)$ gauge field defined on $\mathbb{R}^4$, the self-duality gives rise to

$$ F_{\mu \nu} \# F_{\mu \nu} \equiv \frac{1}{2} \epsilon_{\mu \nu \alpha \beta} F_{\alpha \beta}, $$

(3)
where \( \epsilon_{\mu\nu\alpha\beta} \) is the Levi–Civita tensor, \( *F \) is the dual of field strength \( F \), and \( \mu, \nu, \alpha, \beta \) run over \{0, 1, 2, 3\}. In this case, the gauge potentials \( B_\mu \) are also self-dual\(^9\) and belong to \( \text{su}(N) \). Yang extended the gauge potentials into complex space \( \mathbb{C}^4 \) where \( (x^0, x^1, x^2, x^3) \) are complex. He introduced coordinates (transformation)\(^{29}\) (cf. Ref. 17)

\[
\mathbf{Y} = \sqrt{2} \begin{pmatrix} y - \bar{z} \\ z - \bar{y} \end{pmatrix} = x^0 - i \mathbf{x} \cdot \mathbf{\sigma},
\]

i.e.,

\[
y = \frac{\sqrt{2}}{2} (x^0 - ix^3), \quad \bar{y} = \frac{\sqrt{2}}{2} (x^0 + ix^3), \quad z = \frac{\sqrt{2}}{2} (x^2 - ix^1), \quad \bar{z} = \frac{\sqrt{2}}{2} (x^2 + ix^1),
\]

where \( i^2 = -1, \mathbf{x} = (x^1, x^2, x^3), \mathbf{\sigma} = (\sigma_1, \sigma_2, \sigma_3) \), and \( \sigma_i \) are Pauli matrices. Note that here \( \bar{y} \) and \( \bar{z} \) denote variables independent of the complex conjugates, \( y^*, z^* \), of \( y \) and \( z \), so that real Euclidean space is specified by \( \bar{y} = y^* \) and \( \bar{z} = z^* \). The self-dual condition (3) reduces to

\[
F_{yz} = F_{\bar{y}\bar{z}} = 0, \quad (6a)
\]

\[
F_{y\bar{y}} + F_{z\bar{z}} = 0. \quad (6b)
\]

Next, introducing covariant derivative \( D_\chi := \partial_\chi + B_\chi, \chi \in \{y, \bar{y}, z, \bar{z}\} \), where (cf. Ref. 30)

\[
B_y = B_0 + iB_3, \quad B_{\bar{y}} = B_0 - iB_3, \quad B_z = B_2 + iB_1, \quad B_{\bar{z}} = B_2 - iB_1,
\]

one obtains a representation of (6a) with respect to \( D_\chi \),

\[
[D_y, D_z] = 0, \quad [D_{\bar{y}}, D_{\bar{z}}] = 0. \quad (8)
\]

This implies that there exist two \( \mathcal{N} \times \mathcal{N} \) generating matrices \( D \) and \( \bar{D} \) such that

\[
D_y(D) = D_z(D) = 0, \quad D_{\bar{y}}(\bar{D}) = D_{\bar{z}}(\bar{D}) = 0, \quad (9)
\]

which leads to

\[
B_y = D\partial_y(D^{-1}), \quad B_z = D\partial_z(D^{-1}), \quad B_{\bar{y}} = \bar{D}\partial_{\bar{y}}(\bar{D}^{-1}), \quad B_{\bar{z}} = \bar{D}\partial_{\bar{z}}(\bar{D}^{-1}). \quad (10)
\]

Then, for the field strength \( F_{\mu\nu} \) defined by (2) with the above \( B_\chi \), equation (6a) holds automatically and (6b) reduces to Ref. 31

\[
(J_{\bar{y}}J^{-1})_y + (J_zJ^{-1})_{\bar{z}} = 0, \quad (11)
\]

or its alternative form

\[
(J^{-1}J_{\bar{y}})_y + (J^{-1}J_z)_{\bar{z}} = 0, \quad (12)
\]
where
\[ J = D \tilde{D}^{-1}. \] (13)

When the gauge group is SU(\(N\)), it turns out that \(D, \tilde{D} \in \text{SL}(N)\) and \(\tilde{D} = (D^\dagger)^{-1}\), where \(D^\dagger = (D^*)^T\), and hence \(J = DD^\dagger\). In other words, \(J\) is a positive-definite Hermitian matrix with \(|J| = 1\).

Equation (11) (or (12)) is the usual form that was studied using methods of integrable systems, see Refs. 23–25, 32. In SU(2) case, once \(J\) is obtained, \(D\) can be recovered by Yang’s R-gauge\(29\) as the following. Writing \(J\) in the form
\[ J = \frac{1}{f} \begin{pmatrix} 1 & -g \\ e & f^2 - eg \end{pmatrix}, \] (14)
where \(f\) is real and \(e = -g^*\). Then \(D\) takes the form
\[ D = \frac{1}{\sqrt{f}} \begin{pmatrix} 1 & 0 \\ e & f \end{pmatrix} U \] (15)
where \(U \in \text{SU}(2)\). Hence \(B_X\) are recovered from (10) and so are \(F_{\mu\nu}\) from (2).

In this paper, we will develop the Cauchy matrix approach to construct explicit solutions of Equation (11).

3 THE CAUCHY MATRIX APPROACH TO THE SDYM EQUATION

To construct the SU(2) SDYM equation (11), we start from the Sylvester equation
\[ KM - MK = rs^T, \] (16)
equipped with dispersion relations
\[ r_{x_n} = AK^n r, \quad s_{x_n} = A(K^T)^n s, \quad (n \in \mathbb{Z}), \] (17)
where \(\{x_n\}\) are infinitely many complex independent variables, \(K, M, A, r, s\) are block matrices in the form of
\[ K = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix}, \quad M = \begin{pmatrix} 0 & M_1 \\ M_2 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} I_{N_1} & 0 \\ 0 & -I_{N_2} \end{pmatrix}, \quad r = \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix}, \quad s = \begin{pmatrix} 0 & s_1 \\ s_2 & 0 \end{pmatrix}, \] (18)
with \(K_i \in \mathbb{C}_{N_i \times N_i}, M_i \in \mathbb{C}_{N_1 \times N_2}[x], M_2 \in \mathbb{C}_{N_2 \times N_1}[x], r_i, s_i \in \mathbb{C}_{N \times 1}[x], I_{N_i}\) being the \(N_i\)-th order identity matrix for \(i = 1, 2, x = (\ldots, x_{-1}, x_0, x_1, \ldots),\) and \(N_1 + N_2 = 2N\). In addition, we assume \(K_1\) and \(K_2\) are invertible and do not share any eigenvalues so that the Sylvester equation (16) has a unique solution \(M\) for given \(K, r, s\).\(33\) Define an infinite matrix \(S = (S^{(i,j)})_{\infty \times \infty}\) where each \(S^{(i,j)}\)
is a $2 \times 2$ matrix defined as

\[
S^{(i,j)} = s^T K^i (I + M)^{-1} K^j r = \begin{pmatrix} s_1^{(i,j)} & s_2^{(i,j)} \\ s_3^{(i,j)} & s_4^{(i,j)} \end{pmatrix}, \quad (i, j \in \mathbb{Z}),
\]

i.e.,

\[
s_1^{(i,j)} = -s_2^T K_2^j (I_N - M_2 M_1)^{-1} M_2 K_1^i r_1,
\]

\[
s_2^{(i,j)} = s_2^T K_2^j (I_N - M_2 M_1)^{-1} K_2^i r_2,
\]

\[
s_3^{(i,j)} = s_1^T K_1^j (I_N - M_1 M_2)^{-1} K_1^i r_1,
\]

\[
s_4^{(i,j)} = -s_1^T K_1^j (I_N - M_1 M_2)^{-1} M_1 K_1^i r_2,
\]

where $I$ denotes the $2N$-th order identity matrix. The above settings are the same as those for deriving the Ablowitz–Kaup–Newell–Segur (AKNS) system,\(^{34}\) except in (17) we have introduced variable $x_0$ and the corresponding dispersion relation. Thus we may make use of the results already obtained in Ref. 34. However, the variable $x_0$ does play a useful role in our procedure of deriving the SDYM equation (11).

For $S^{(i,j)}$ defined in (19) where $K, M, r, s$ are governed by the Sylvester equation (16), there exists a recursive relation independent of dispersion relations, see Proposition 2 in Ref. 34 and cf. Ref. 27.

**Lemma 1.** \{$S^{(i,j)}$\} defined in (19) satisfy

\[
S^{(i,j+s)} = S^{(i+s,j)} - \sum_{l=0}^{s-1} S^{(s-1-l,j)} S^{(i,l)}, \quad (s = 1, 2, ...),
\]

where $K, M, r, s$ obey the Sylvester equation (16). In particular, when $s = 1$, it reads

\[
S^{(0,j)} S^{(1,0)} = S^{(1,1)} - S^{(1,2)}.
\]

With respect to the dispersion relation (17), $S^{(i,j)}$ evolves as the following.

**Lemma 2.** $S^{(i,j)}$ defined by (19) obeys evolutions

\[
S_{x}^{(i,j)} = S^{(i+n,j)} a - a S^{(i,j+n)} - \sum_{l=0}^{n-1} S^{(n-1-l,j)} a S^{(i,l)}, \quad (n \in \mathbb{Z}^+),
\]

\[
S_{x_0}^{(i,j)} = S^{(i,j)} a - a S^{(i,j)} = [S^{(i,j)}, a],
\]

(23a, 23b)
\[ S_{x_n}^{(i,j)} = S^{(i+n,j)} + aS^{(i,j+n)} + \sum_{l=-1}^{n} S^{(n-1-l,j)}aS^{(l,i)}, \quad (n \in \mathbb{Z}^-), \quad (23c) \]

where \( K, M, r, s \) satisfy the Sylvester equation (16) and dispersion relation (17), and \( a = \sigma_3 = \text{diag}(1, -1) \).

Formulae (23a) and (23c) have been derived in Ref. 34 and (23b) can be obtained similarly. For the completeness of the paper, a proof of Lemma 2 is given in Appendix A.

Now we come to the first main result of this paper.

**Theorem 1.** Let

\[ u \equiv S^{(0,0)} = s^T(I + M)^{-1}r, \quad v \equiv I_2 - S^{(-1,0)} = I_2 - s^T(I + M)^{-1}K^{-1}r. \quad (24) \]

Then \( u \) and \( v \) satisfy the following differential recurrence relation:

\[ v_{x_{n+1}} - u_{x_n} = -u_{x_n}, \quad (n \in \mathbb{Z}). \quad (25) \]

**Proof.** First, the recursive formula (22) with \( i = -1 \) yields

\[ S^{(-1,j+1)} = S^{(0,j)}v, \quad (26) \]

which looks simple but will play a crucial role in the following proof. It gives rise to (with \( j = 0 \))

\[ uv = S^{(-1,1)} \quad (27) \]

and (with \( j = -1 \))

\[ v^{-1} = I_2 + S^{(0,-1)}. \quad (28) \]

In addition, the evolution relation (23b) indicates (considering \( (i, j) = (0, 0) \) and \( (-1, 0) \))

\[ u_{x_0} = [u, a] \quad (29) \]

and

\[ v_{x_0} = [v, a]. \quad (30) \]

Next, looking at (23a) with \( n = 1 \) and \( (i, j) = (-1, 0) \), and making use of (27) and (29), we have

\[ -v_{x_1} = ua - aS^{(-1,1)} - ua(I_2 - v) = uav - auv = [u, a]v = u_{x_0}v. \quad (31) \]

Similarly, with \( n = -1 \) and \( (i, j) = (0, 0) \) formula (23c) gives rise to

\[ u_{x_{-1}} = a - va(I_2 + S^{(0,-1)}) = a - vav^{-1} = (av - va)v^{-1} = -v_{x_0}v^{-1}. \quad (32) \]
where (28) and (30) have been used. These two equations cover the cases \( n = 0 \) and \(-1\) of (25).

Next, we prove (25) for positive \( n \). Formula (23a) gives rise to (with \( i = j = 0 \))

\[
\mathbf{u}_{x_n} = S^{(n,0)}a - aS^{(0,n)} - \sum_{l=0}^{n-1} S^{(n-1-l,0)}aS^{(0,l)}
\]

(33)

and (with \( i = -1, j = 0 \))

\[
-v_{x_{n+1}} = S^{(n,0)}a - aS^{(-1,n+1)} - \sum_{l=0}^{n} S^{(n-l,0)}aS^{(-1,l)}
\]

\[
= S^{(n,0)}a - aS^{(0,n)}v - S^{(n,0)}aS^{(-1,0)} - \sum_{l=1}^{n} S^{(n-l,0)}aS^{(-1,l)},
\]

(34)

where we have made use of (26) and separated the first term from the summation. Then, in light of definition of \( v \), replacing the index \( l \) with \( l + 1 \) and using (26) once again, we arrive at

\[
-v_{x_{n+1}} = S^{(n,0)}av - aS^{(0,n)}v - \sum_{l=0}^{n-1} S^{(n-1-l,0)}aS^{(-1,l+1)}
\]

\[
= S^{(n,0)}av - aS^{(0,n)}v - \sum_{l=0}^{n-1} S^{(n-1-l,0)}aS^{(-1,l)}v = \mathbf{u}_{x_n}v,
\]

(35)

which is (25) with \( n \geq 1 \).

The case of \( n \) less than \(-1\) can be proved similarly from (23c). In detail, we have

\[
\mathbf{u}_{x_n} = S^{(n,0)}a - aS^{(0,n)} + \sum_{l=-1}^{n} S^{(n-1-l,0)}aS^{(0,l)}
\]

(36)

and

\[
-v_{x_{n+1}} = S^{(n,0)}a - aS^{(-1,n+1)} + \sum_{l=-1}^{n+1} S^{(n-l,0)}aS^{(-1,l)}
\]

\[
= S^{(n,0)}av + S^{(n,0)}aS^{(-1,0)} - aS^{(0,n)}v + \sum_{l=-1}^{n+1} S^{(n-l,0)}aS^{(-1,l)}
\]

\[
= S^{(n,0)}av - aS^{(0,n)}v + \sum_{l=0}^{n+1} S^{(n-l,0)}aS^{(-1,l)}
\]

\[
= S^{(n,0)}av - aS^{(0,n)}v + \sum_{l=0}^{n} S^{(n-l,0)}aS^{(-1,l+1)}
\]

\[
= S^{(n,0)}av - aS^{(0,n)}v + \sum_{l=0}^{n} S^{(n-l,0)}aS^{(0,l)}v = \mathbf{u}_{x_n}v.
\]

(37)
Thus, we have proved (25) for all \( n \in \mathbb{Z} \).

Considering the compatibility \((u_{x_n})_{x_m} = (u_{x_m})_{x_n}\), we immediately arrive at the following.

**Theorem 2.** For \( v \equiv I_2 - S^{(-1,0)} = I_2 - s^T (I + M)^{-1} K^{-1} r \) where \( K, M, r, s \) satisfy the Sylvester equation (16) and dispersion relation (17), the following relation holds,

\[
(v_{x_{n+1}} v^{-1})_{x_m} - (v_{x_{m+1}} v^{-1})_{x_n} = 0,
\]

where \( n, m \in \mathbb{Z} \).

As a by-product of (38), \( u \) satisfies a potential SDYM equation (cf. Ref. 35)

\[
(u_{x_{n+1}} u^{-1})_{x_m} - (u_{x_{m+1}} u^{-1})_{x_n} - [u_{x_m}, u_{x_n}] = 0.
\]

Note that Equation (38) differs from the SDYM equation (11) by a sign “−”. In the next section, we will recover (11) from (38) by imposing certain reductions. Besides, apart from \( \{x_n\} \) with dispersion relation (17), we may introduce for \( \{y_m\} \) such that

\[
r_{y_m} = A(-K)^m r, \quad s_{y_m} = A(-K^T)^m s, \quad (m \in \mathbb{Z}).
\]

The resulting equation is

\[
(v_{x_{n+1}} v^{-1})_{y_m} + (v_{y_{m+1}} v^{-1})_{x_n} = 0,
\]

which is in a same form as (11).

### 4 | SOLUTIONS TO THE SU(2) SDYM EQUATION

When the gauge group is SU(2), \( J \) in the SDYM equation (11) should be a positive-definite Hermitian matrix with \( |J| = 1 \).29,30 In the following we will look for a Hermitian matrix \( v \) with \( |v| = 1 \). This will be able to be achieved by imposing some constraints on \( K \) and \( \{x_n\} \).

The Sylvester equation (16) can be written as a more explicit form

\[
K_1 M_1 - M_1 K_2 = r_1 s_2^T,
\]

\[
K_2 M_2 - M_2 K_1 = r_2 s_1^T.
\]

For given invertible \( K_1 \) and \( K_2 \) that do not share any eigenvalues, \( M_1 \) and \( M_2 \) can be determined uniquely. In practice, because \( S^{(i,j)} \) is invariant with respect to \( K_1, K_2 \) and any matrices similar to them34 (cf. Refs. 27, 28), we can consider \( K_1 \) and \( K_2 \) to be their canonical forms, say \( \Gamma \) and \( \Lambda \). Solutions \( M_i \) together with \( r_i \) and \( s_i \) can be explicitly presented. One may refer to Ref. 34 or Appendix B of the present paper. It turns out that these solutions can be presented via the following form:

\[
M_1 = F_1 G_1 H_2, \quad M_2 = F_2 G_2 H_1, \quad r_1 = F_1 E_1, \quad r_2 = F_2 E_2, \quad s_1 = H_1 E_1, \quad s_2 = H_2 E_2,
\]
and these elements satisfy (symmetric or commutative) the relations

\[
\begin{align*}
G_1 &= -G_2^T, \quad F_1 \Gamma = \Gamma F_1, \quad F_2 \Lambda = \Lambda F_2, \quad \Gamma H_1 = H_1 \Gamma^T, \quad \Lambda H_2 = H_2 \Lambda^T, \\
H_i^T &= H_i, \quad (H_i F_i)^T = F_i^T H_i = H_i F_i, \quad i = 1, 2,
\end{align*}
\]

(44)

where \( F_i \) and \( H_i \) are \( N_i \times N_i \) matrices, \( G_1 \) is a \( N_1 \times N_2 \) matrix, \( G_2 \) is a \( N_2 \times N_1 \) matrix, and \( E_i \) is a \( N_i \)-th order column vector, \( i = 1, 2 \).

With the above notations, we are able to investigate symmetric property of the infinite matrix \( S \), i.e., the relations between \( S^{(i,j)} \) and \( S^{(j,i)} \).

**Lemma 3.** \{\( S^{(i,j)} \)\} defined by (19) with \( K_1 = \Gamma, \ K_2 = \Lambda, \ M, \ r, \ s \) satisfy the Sylvester equation (16), furthermore, the elements in \( S^{(i,j)} \) and \( S^{(j,i)} \) are related as the following,

\[
\begin{align*}
S^{(i,j)}_1 &= -S^{(j,i)}_4, \quad S^{(i,j)}_2 = S^{(j,i)}_2, \quad S^{(i,j)}_3 = S^{(j,i)}_3, \quad i, j \in \mathbb{Z},
\end{align*}
\]

(45)

i.e., \( S^{(i,j)} \) and \( S^{(j,i)} \) are related as

\[
(46)
\]

Proof. Making use of expressions (43) and relations (44), from (20) we find

\[
\begin{align*}
S^{(i,j)}_1 &= -S^{(j,i)}_2 \Lambda \Gamma^{-1} (I_{N_2} - M_2 M_1)^{-1} M_2 \Gamma^i r_1 \\
&= \left( -E_2^T (\Lambda^T)^{(H_2 F_2)^{-1}} - G_2 H_1 F_1 G_1 \right) (I_{N_2} - M_2 M_1)^{-1} M_2 \Gamma^i r_1 \\
&= S^{(i,j)}_3 (I_{N_2} - M_2 M_1)^{-1} \Lambda^i r_2 = S^{(j,i)}_4,
\end{align*}
\]

and

\[
\begin{align*}
S^{(i,j)}_2 &= S^{(j,i)}_2 \Lambda \Gamma^{-1} (I_{N_2} - M_2 M_1)^{-1} \Lambda^i r_2 \\
&= E_2^T (\Lambda^T)^{((H_2 F_2)^{-1} - G_2 H_1 F_1 G_1)^{-1} \Lambda^i F_2 E_2} \\
&= (E_2^T (\Lambda^T)^{(H_2 F_2)^{-1}} - G_2 H_1 F_1 G_1)^{-1} \Lambda^i E_2 \\
&= S^{(i,j)}_2 \Lambda \Gamma^{-1} (I_{N_2} - M_2 M_1)^{-1} \Lambda^i r_2 = S^{(j,i)}_2.
\end{align*}
\]

(47)

The third relation \( S^{(i,j)}_3 = S^{(j,i)}_3 \) can be proved in a similar way.

With this lemma we are able to prove \(|v| = 1\). In fact, the relation (28) yields

\[
\left( 1 + s^{(0,-1)}_1 \right) \left( 1 - s^{(-1,0)}_1 \right) - s^{(0,-1)}_2 s^{(-1,0)}_3 = 1.
\]

(48)
Then, by Lemma 3 we may replace $s_1^{(0,-1)}$ and $s_2^{(0,-1)}$ using (45) and the resulting equation gives rise to
\[
\begin{bmatrix}
1 - s_1^{(-1,0)} & -s_2^{(-1,0)} \\
-s_3^{(-1,0)} & 1 - s_4^{(-1,0)}
\end{bmatrix} = |J_2 - S^{(-1,0)}| = |\mathbf{v}| = 1. \tag{49}
\]

Next, we make $\mathbf{v}$ to be a Hermitian matrix by imposing constraints. First, we introduce
\[
z_n \doteq x_n = \xi_n + i\eta_n, \quad \bar{z}_n \doteq (-1)^{n+1} x_{-n} = \xi_n - i\eta_n, \quad n = 1, 2, \ldots, \tag{50}
\]
where $\xi_n, \eta_n \in \mathbb{R}$. This indicates $\bar{z}_n = z_n^\ast$. Then we take $m = -n - 1$. The resulting Equation (38) reads
\[
(\mathbf{v}_{zn+1} \mathbf{v}^{-1})_{zn+1} + (\mathbf{v}_{zn} \mathbf{v}^{-1})_{zn} = 0, \quad n = 1, 2, \ldots, \tag{51}
\]
which coincides with the form (11). We next introduce further constraints by $N_2 = N_1$ and
\[
K_2 = -(K_1^\ast)^{-1} \tag{52}
\]
to the Cauchy matrix scheme (16, 17, 18). Equation (50) implies
\[
\partial_{\xi_n} = \partial_{x_n} + \partial_{\bar{x}_n}, \quad \partial_{\eta_n} = i(\partial_{x_n} - \partial_{\bar{x}_n}), \tag{53}
\]
and the dispersion relation (17) is equivalently written in terms of $\xi_n$ and $\eta_n$ as
\[
\begin{align*}
\partial_{\xi_n} \mathbf{r}_1 &= (K_1^n + (-1)^{n+1}K_1^{-n}) \mathbf{r}_1, \quad \partial_{\eta_n} \mathbf{r}_1 = i(K_1^n - (-1)^{n+1}K_1^{-n}) \mathbf{r}_1, \\
\partial_{\xi_n} \mathbf{r}_2 &= ((K_1^\ast)^n + (-1)^{n+1}(K_1^\ast)^{-n}) \mathbf{r}_2, \quad \partial_{\eta_n} \mathbf{r}_2 = -i((K_1^\ast)^n - (-1)^{n+1}(K_1^\ast)^{-n}) \mathbf{r}_2,
\end{align*} \tag{54}
\]
and
\[
\begin{align*}
\partial_{\xi_n} \mathbf{s}_1 &= \left((K_1^T)^n + (-1)^{n+1}(K_1^T)^{-n}\right) \mathbf{s}_1, \quad \partial_{\eta_n} \mathbf{s}_1 = i\left((K_1^T)^n - (-1)^{n+1}(K_1^T)^{-n}\right) \mathbf{s}_1, \\
\partial_{\xi_n} \mathbf{s}_2 &= \left((K_1^\dagger)^n + (-1)^{n+1}(K_1^\dagger)^{-n}\right) \mathbf{s}_2, \quad \partial_{\eta_n} \mathbf{s}_2 = -i\left((K_1^\dagger)^n - (-1)^{n+1}(K_1^\dagger)^{-n}\right) \mathbf{s}_2. \tag{55}
\end{align*}
\]
We are able to take
\[
\mathbf{r}_2 = (K_1^\ast)^{-1}\mathbf{r}_1^\ast, \quad \mathbf{s}_2 = -\delta(K_1^\dagger)^{-1}\mathbf{s}_1^\ast, \tag{56}
\]
where $\delta = \pm 1$, such that
\[
M_2 = \delta M_1^\ast \tag{57}
\]
in light of the uniqueness of solutions of the Sylvester equations (42). This also means the reduction (50) is allowed. Denote $\mathbf{v} = \begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix}$. Then, by direct calculation we find

$$
v^*_1 = 1 - \left( s_1^{(-1,0)} \right)^* = 1 + (s_2^{T})^* \mathbf{M}_2^* (\mathbf{I}_{N_2} - \mathbf{M}_1^* \mathbf{M}_2^*)^{-1} (\mathbf{K}_2^*)^{-1} \mathbf{r}_1^* = 1 - \mathbf{s}_1^T \mathbf{K}_1^{-1} \mathbf{M}_1 \mathbf{r}_1 = 1 + s_4^{(0,-1)} = 1 - s_1^{(-1,0)} = v_1,
$$

(58)

and $v^*_4 = v_4$ in a similar way, and

$$
v^*_2 = -\left( s_2^{(-1,0)} \right)^* = - (s_2^{T})^* \mathbf{M}_2^* (\mathbf{I}_{N_2} - \mathbf{M}_1^* \mathbf{M}_2^*)^{-1} \mathbf{r}_2^* = - \mathbf{s}_2^T \mathbf{K}_1^{-1} \mathbf{M}_2^{-1} \mathbf{r}_1 = - \mathbf{s}_2^{(0,-1)} = - \mathbf{s}_3^{(-1,0)} = \delta v_3.
$$

(59)

All these together indicate $\mathbf{v} = \mathbf{v}^\dagger$ when we take $\delta = 1$, i.e., $\mathbf{v}$ is a Hermitian matrix when $\delta = 1$.

We end this section with the following summarization.

**Theorem 3.** The SDYM equation (51) has the following solutions:

$$
\mathbf{v} = \mathbf{I}_2 - \mathbf{S}^{(-1,0)} = \begin{pmatrix} 1 - s_1^{(-1,0)} & \mathbf{s}_2^{(-1,0)} \\ \mathbf{s}_3^{(-1,0)} & 1 - s_4^{(-1,0)} \end{pmatrix},
$$

(60)

where

$$
s_1^{(-1,0)} = s_1^\dagger (\mathbf{K}_1^*)^{-1} (\mathbf{I}_N - \mathbf{M}_1^* \mathbf{M}_1) \mathbf{K}_1^{-1} \mathbf{r}_1,
$$

$$
s_3^{(-1,0)} = s_3^T (\mathbf{I}_N - \mathbf{M}_1^* \mathbf{M}_1)^{-1} \mathbf{K}_1^{-1} \mathbf{r}_1,
$$

$$
s_4^{(-1,0)} = s_4^T \mathbf{M}_1 (\mathbf{I}_N - \mathbf{M}_1^* \mathbf{M}_1)^{-1} \mathbf{r}_1,
$$

(61)

and

$$
z_n = \xi_n + i \eta_n, \quad \bar{z}_n = z_n^* = \xi_n - i \eta_n, \quad \xi_n, \eta_n \in \mathbb{R}, \quad n = 1, 2, \ldots.
$$

(62)

Here, $K_1 \in \mathbb{C}_{N \times N}$, $K_1$ and $-(K_1^*)^{-1}$ do not share any eigenvalues, $M_1, r_1$, and $s_1$ are determined by the system

$$
K_1 M_1 + M_1 (K_1^*)^{-1} = - \mathbf{r}_1 s_1^\dagger (K_1^*)^{-1},
$$

(63a)

$$
\partial_{\xi_n} r_1 = (K_1^n + (-1)^{n+1} K_1^{-n}) r_1, \quad \partial_{\eta_n} r_1 = i (K_1^n - (-1)^{n+1} K_1^{-n}) r_1,
$$

(63b)

$$
\partial_{\xi_n} s_1 = \left( (K_1^T)^n + (-1)^{n+1} (K_1^T)^{-n} \right) s_1, \quad \partial_{\eta_n} s_1 = i \left( (K_1^T)^n - (-1)^{n+1} (K_1^T)^{-n} \right) s_1,
$$

(63c)
for $n = 1, 2, \ldots$. Solution $v$ satisfies $v = v^\dagger$ and $|v| = 1$. $v$ is piecewisely positive-definite or negative-definite, depending on the domains where $v$ is positive or negative.

Explicit solutions of the system (63) can be formulated from Appendix B. Some examples of solution $v$ will be listed in Appendix C.

5 | CONCLUDING REMARKS

In this paper we have constructed explicit solution $v$ for the SDYM equation (51) by using a direct method, namely, the Cauchy matrix approach. We started with the Sylvester equation (16) together with the dispersion relation (17) with respect to infinitely many coordinates $\{x_n\}$, and then, making use of the recursive relation (22) we proved $u$ and $v$ satisfy the key equation (25) that gives rise to Equation (38). After that, we introduced real independent coordinates $\xi_n$ and $\eta_n$, and imposed constraints on $z_n$, $\bar{z}_n$ and $K_1$ and $K_2$. Finally, we obtained solutions of the SDYM equation (51). Solutions have been presented via Theorem 3.

Our Cauchy matrix approach is based on the scheme for the AKNS system, cf. Ref. 34. Compared with Ref. 34, here we introduced the auxiliary variable $x_0$, frequently made use of the recursive relation (22), and finally we were able to construct the key equation (25). We also discussed symmetric relation of $S^{(i,j)}$ and $S^{(j,i)}$. All these elaborations enabled us to finally prove the property $v = v^\dagger$ and $|v| = 1$, and obtain exact solutions to the SDYM equation (51).

There are similar direct approaches to construct Equation (25), e.g., the one based on biddifferential graded algebra, 36,37 where it is assumed there exists a function $v$ to satisfy Equation (25). In our approach, $v$ is clearly defined by $I_2 - S^{(-1,0)}$, which provides an explicit solution to the SDYM equation (51).

In our scheme, after introducing $z_n$ and $\bar{z}_n$ by (50), we took $m = -n - 1$. Instead of doing that, if we take $m = n - 1$ in Equation (38), we have an equation

$$
(v_{x_{n+1}} v^{-1})_{x_{n-1}} - (v_{x_n} v^{-1})_{x_n} = 0, \quad (n \in \mathbb{Z}),
$$

(64)

and by redefining $x_n$ by $i x_n$, it becomes

$$
(v_{x_{n+1}} v^{-1})_{x_{n-1}} + (v_{x_n} v^{-1})_{x_n} = 0, \quad (n \in \mathbb{Z}).
$$

(65)

Manakov and Zakharov constructed its solutions using its Lax pair.32 Because the equation depends only on three independent variables, its solutions may generate monopoles (see Ref. 9). Note that solutions obtained in Ref. 32 are different from ours. In fact, the two constraints on $n$ and $m$ hold simultaneously only when $n = 0$. In this case, our results can be applied and Equation (65) can be written as

$$
(v_{x_1} v^{-1})_{x_{-1}} - [v, a]v^{-1}, a = 0,
$$

(66)

where we have made use of relation (30). Note that this equation is a 2D equation and was derived recently in Ref. 38 by means of a similar direct method.

With regard to the methods of solving the SDYM equation, most of them are direct and constructive. The approach in Ref. 20 can be thought of a Darboux–Bäcklund transformation employing the Lax pair of the SDYM equation. Solutions obtained are algebraic type, i.e., rational solutions
in terms of polynomials of independent variables. The bilinear approach given in Ref. 24 (cf. Ref. 23) is to reformulate the bilinear relations satisfied by the functions $e, f, g$ in (14) to a new set of bilinear equations involving nine $\tau$ functions in Hankelians. In principle, they provide algebraic solutions as well. Our solutions are expressed in terms of exponential functions, which are in expression formally similar to those obtained via Darboux transformation. However, our solutions are different from those in Ref. 25, as $J$ is a solution of the SU(2) SDYM equation but it is not necessary $J \in \text{SU}(2)$ (cf. Ref. 25). In addition, our expression for $v$ is more explicit and includes solutions generated by possible canonical forms of $K_\perp$ (not only diagonal form or a Jordan form, but also any combinations of them).

There are several further investigations related to the approach and results of the present paper. For example, extend the approach to the SU($\mathcal{N}$) SDYM equation and noncommutative case, e.g., Refs. 39–41. Besides, it is well known that as a 4D integrable system, the SDYM equation allows various reductions to lower dimensional integrable equations. That would be interesting to understand how reductions play roles in generating solutions to lower dimensional integrable equations. In addition, Ward used to discuss discretization of the SDYM equation. Because the Cauchy matrix approach originates from solving discrete integrable systems, it would be interesting to have a discrete analogue of the SDYM equation from this approach.

**ACKNOWLEDGMENTS**

This project is supported by the NSF of China (11631007, 11875040, 11971251) and Science and Technology Innovation Plan of Shanghai (20590742900).

**ORCID**

Changzheng Qu https://orcid.org/0000-0003-4859-367X

Da-jun Zhang https://orcid.org/0000-0003-3691-4165

**REFERENCES**

1. Wu TT. Remarks on Yang–Mills theory. In: Goldhaber A, Shrock R, Smith J, et al. eds. *Symmetry and Modern Physics: Yang Retirement Symposium*. World Scientific; 2003:199-204.
2. Yang CN, Mills RL. Conservation of isotopic spin and isotopic gauge invariance. Phys Rev. 1954;96:191-195.
3. Atiyah MF. Classical Geometry of Yang–Mills Fields. Lez. Fermi. Ac. Naz. dei Lincei, Sc. Norm. Sup. Pisa; 1979.
4. Bourguignon JP, Lawson HBJ. Yang–Mills theory: its physical origin and differential geometric aspects. In: Yau ST ed. *Seminar on Differential Geometry*, Ann. Math. Studies, No.102. Princeton University Press; 1982:395-421.
5. Gu CH. On classical Yang–Mills fields. Phys Rep. 1981;80:257-331.
6. Wu TT, Yang CN. Concept of nonintegrable phase factors and global formulation of gauge fields. Phys Rev D. 1975;12:3845-3857.
7. Bourguignon JP. Yang–Mills theory: the differential geometric side. In: Hansen VL ed. *Differential Geometry: Proceedings of the Nordic Summer School Held in Lyngby, Denmark Jul. 29–Aug. 9, 1985*. Springer-Verlag; 1987:13-54.
8. Weatherall JO. Fiber bundles, Yang–Mills theory, and general relativity. Synthese. 2016;193:2389-2425.
9. Prasad MK. Instantons and monopoles in Yang–Mills gauge field theories. Physica. 1980;1D:167-191.
10. Belavin AA, Polyakov AM, Schwartz AS, Tyupkin YS. Pseudoparticle solutions of the Yang–Mills equations. Phys Lett. 1975;59B:85-87.
11. Corrigan EF, Fairlie DB. Scalar field theory and exact solutions to a classical SU(2) gauge theory. Phys Lett. 1977;76B:69-71.
12. Wilczek F. Geometry and interactions of instantons. In: Stump DR, Weingarten DH eds. *Quark Confinement and Field Theory*. John Wiley and Sons; 1977:211-219.
13. Atiyah MF, Hitchin NJ, Drinfeld VG, Manin YI. Construction of instantons. Phys Lett. 1978;76A:185-187.
14. Atiyah MF, Hitchin NJ, Singer IM. Deformations of instantons. Proc Natl Acad Sci USA. 1977;74:2662-2663.
15. Corrigan EF, Fairlie DB, Yates RG, Goddard P. The construction of self-dual solutions to SU(2) gauge theory. *Commun Math Phys.* 1978;58:223-240.

16. Atiyah MF, Ward RS. Instantons and algebraic geometry. *Commun Math Phys.* 1977;55:117-124.

17. Ward RS. On self-dual gauge fields. *Phys Lett.* 1977;61A:81-82.

18. Jimbo M, Kruskal MD, Miwa T. Painlevé test for the self-dual Yang–Mills equation. *Phys Lett.* 1982;92A:59-60.

19. Ward RS. The Painlevé property for the self-dual gauge-field equations. *Phys Lett.* 1984;102A:279-282.

20. Belavin AA, Zakharov VE. Yang–Mills equations as inverse scattering problem. *Phys Lett.* 1978;73B:53-57.

21. Ueno K, Nakamura Y. Transformation theory for anti-self-dual equations and the Riemann–Hilbert problem. *Phys Lett.* 1982;109B:273-278.

22. Takasaki K. A new approach to the self-dual Yang–Mills equations. *Commun Math Phys.* 1984;94:35-59.

23. Ohta Y, Nimmo JJC, Gilson CR. A bilinear approach to a Pfaffian self-dual Yang–Mills equation. *Glasgow Math J.* 2001;43:99-108.

24. Sasa N, Ohta Y, Matsukidaira J. Bilinear form approach to the self-dual Yang–Mills equations and integrable systems in (2+1)-dimension. *J Phys Soc Japan.* 1998;67:83-86.

25. Nimmo JJC, Gilson CR, Ohta Y. Applications of Darboux transformations to the self-dual Yang–Mills equations. *Theor Math Phys.* 2000;122:239-246.

26. Nijhoff FW, Atkinson J, Hietarinta J. Soliton solutions for ABS lattice equations: I. Cauchy matrix approach. *J Phys A: Math Theor.* 2009;42:404005(34pp).

27. Xu DD, Zhang DJ, Zhao SL. The Sylvester equation and integrable equations: I. The Korteweg-de Vries system and sine-Gordon equation. *J Nonl Math Phys.* 2014;21:382-406.

28. Zhang DJ, Zhao SL. Solutions to ABS lattice equations via generalized Cauchy matrix approach. *Stud Appl Math.* 2013;131:72-103.

29. Yang CN. Condition of self-duality for SU(2) gauge fields on Euclidean four-dimensional space. *Phys Rev Lett.* 1977;38:1377-1379.

30. Pohlmeyer K. On the Lagrangian theory of anti-self-dual fields in four-dimensional Euclidean space. *Commun Math Phys.* 1980;72:37-47.

31. Brihaye Y, Fairlie DB, Nuyts J, Yates RG. Properties of the self dual equations for an SU(n) gauge theory. *J Math Phys.* 1978;19:2528-2532.

32. Manakov SV, Zakharov VE. Three-dimensional model of relativistic-invariant field theory, integrable by the inverse scattering transform. *Lett Math Phys.* 1981;5:247-253.

33. Sylvester J. Sur l'équation en matrices $px = qx$. *C R Acad Sci Paris.* 1884;99:67-76.

34. Zhao SL. The Sylvester equation and integrable equations: the Ablowitz-Kaup-Newell-Segur system. *Rep Math Phys.* 2018;82:241-263.

35. Leznov AN, Mukhtarov MA. Deformation of algebras and solution of self-duality equation. *J Math Phys.* 1987;28:2574-2578.

36. Dimakis A, Müller-Hoissen F. Dispersionless limit of the noncommutative potential KP hierarchy and solutions of the pseudodual chiral model in 2+1 dimensions. *J Phys A: Math Theor.* 2008;41:265203(33pp).

37. Dimakis A, Müller-Hoissen F. Bidifferential graded algebras and integrable systems. *Discrete Contin Dyn Syst Ser A.* 2009;2009(Suppl.2009):208-219.

38. Vekslerchik VE. Solitons of some nonlinear sigma-like models. *Symmetry Integrability Geom: Meth Appl.* 2020;16:144(13pp).

39. Gilson CR, Hamanaka M, Nimmo JJC. Bäcklund transformations and the Atiyah–Ward ansatz for noncommutative anti-self-dual Yang–Mills equations. *Proc R Soc A.* 2009;465:2613-2632.

40. Gilson CR, Hamanaka M, Huang SC, Nimmo JJC. Soliton solutions of noncommutative anti-self-dual Yang–Mills equations. *J Phys A: Math Theor.* 2020;53:404002(17pp).

41. Takasaki K. Anti-self-dual Yang–Mills equations on noncommutative space-time. *J Geom Phys.* 2001;37:291-306.

42. Ablowitz MJ, Chakravarty S, Takhtajan LA. A self-dual Yang–Mills hierarchy and its reductions to integrable systems in 1+1 and 2+1 dimensions. *Commun Math Phys.* 1993;158:289-314.

43. Chakravarty S. SDYM hierarchies and classical soliton systems. In: Clarkson PA ed. *Applications of Analytic and Geometric Methods to Nonlinear Differential Equations*. Springer; 1993:1-8.

44. Mason LJ. Nonlinear Schrödinger and Korteweg-de Vries are reductions of the self-dual Yang–Mills. *Phys Lett A.* 1989;137:30-33.
APPENDIX A: PROOF FOR LEMMA 2

First, the following recurrence relations hold (see eq. (2.1) in Ref. 34, cf. Ref. 27),

\[ K^n M - MK^n = \sum_{l=0}^{n-1} K^{n-1-l} rs^T K^l, \quad (n \in \mathbb{Z}^+), \quad \text{(A1a)} \]

\[ K^n M - MK^n = - \sum_{l=-1}^{n} K^{n-1-l} rs^T K^l, \quad (n \in \mathbb{Z}^-). \quad \text{(A1b)} \]

In addition, in light of the dispersion relation (17), we have

\[ KM_{x_n} - M_{x_n} K = r_{x_n} s^T + rs^T x_n \\
\quad = AK^n rs^T + rs^T K^n A \\
\quad = AK^n(KM - MK) - A(KM - MK) K^n \\
\quad = KA(K^n M - MK^n) - A(K^n M - MK^n) K. \quad \text{(A2)} \]

Note that we have assumed \( K_1 \) and \( K_2 \) are invertible and do not share any eigenvalues so that the Sylvester equation (16) has a unique solution \( M \) for given \( K, r, s \). With such a property, it follows that

\[ M_{x_n} = A(K^n M - MK^n), \quad n \in \mathbb{Z}, \]

which, together with (A1), gives rise to

\[ M_{x_n} = \sum_{l=0}^{n-1} K^{n-1-l} ras^T K^l, \quad (n \in \mathbb{Z}^+), \quad \text{(A3a)} \]

\[ M_{x_n} = 0, \quad (n = 0), \quad \text{(A3b)} \]

\[ M_{x_n} = - \sum_{l=-1}^{n} K^{n-1-l} ras^T K^l, \quad (n \in \mathbb{Z}^-). \quad \text{(A3c)} \]
Next, direct calculation yields

\[
S_{x_n}^{(i,j)} = s_{x_n}^T K^j (I + M)^{-1} K^i r + s_{x_n}^T K^j (I + M)^{-1} K^i r_{x_n} + s_{x_n}^T K^j (I + M)^{-1} K^i r_{x_n} K^i r = -a S_{x_n}^{(i,j+n)} + S_{x_n}^{(i+n,j)} a - s_{x_n}^T K^j (I + M)^{-1} M_{x_n} (I + M)^{-1} K^i r,
\]

(A4)

which gives rise to Equations (23) after substituting (A3) into it.

**APPENDIX B: SOLUTIONS TO (42) AND NOTATIONS**

When \( K_1 \) and \( K_2 \) in the Sylvester equations (42) take their canonical forms \( \Gamma \) and \( \Lambda \), solutions to (42) can be represented as (43). The involved notations are the following. Let

\[
\Gamma = \text{diag}(\Gamma_{n_1}(k_1), \Gamma_{n_2}(k_2), \ldots, \Gamma_{n_p}(k_p)), \quad \Lambda = \text{diag}(\Gamma_{m_1}(l_1), \Gamma_{m_2}(l_2), \ldots, \Gamma_{m_q}(l_q)),
\]

where \( \Gamma_{n}(k) \) denotes a \( n \)-th order Jordan block

\[
\Gamma_{n}(k) = \begin{pmatrix}
k & 0 & 0 & \cdots & 0 & 0 \\
1 & k & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & k
\end{pmatrix}_{n \times n},
\]

(B2)

the index \( \{n_i, m_j\} \) are positive integers and satisfy \( \sum_{i=1}^{p} n_i = N_1, \sum_{j=1}^{q} m_j = N_2 \). Note that \( n_i \) and \( m_j \) allow to be 1, which provides diagonal matrix blocks in \( \Gamma \) and \( \Lambda \). In particular, when all \( \{n_i\} \) are one, \( \Gamma = \text{diag}(k_1, k_2, \ldots, k_N) \). Introduce a lower triangular Toeplitz matrix

\[
F_M(\rho(k)) = \begin{pmatrix}
\rho & 0 & 0 & \cdots & 0 \\
\frac{\partial \rho}{\partial k} & \rho & 0 & \cdots & 0 \\
\frac{1}{2!} \frac{\partial^2 \rho}{\partial k^2} & \frac{\partial \rho}{\partial k} & \rho & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
\frac{\partial^{M-1} \rho}{(M-1)!} & \frac{\partial^{M-2} \rho}{(M-2)!} & \frac{\partial^{M-3} \rho}{(M-3)!} & \cdots & \rho
\end{pmatrix},
\]

(B3)

and a symmetric matrix

\[
H_M(\rho(k)) = \begin{pmatrix}
\rho & \frac{\partial \rho}{\partial k} & \frac{\partial^2 \rho}{\partial k^2} & \cdots & \frac{\partial^{M-1} \rho}{(M-1)!} \\
\frac{\partial \rho}{\partial k} & \frac{1}{2!} \frac{\partial^2 \rho}{\partial k^2} & \frac{\partial^3 \rho}{\partial k^3} & \cdots & 0 \\
\frac{1}{2!} \frac{\partial^2 \rho}{\partial k^2} & \frac{\partial \rho}{\partial k} & \frac{\partial^2 \rho}{\partial k^2} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
\frac{\partial^{M-1} \rho}{(M-1)!} & \frac{\partial^{M-2} \rho}{(M-2)!} & \frac{\partial^{M-3} \rho}{(M-3)!} & \cdots & 0
\end{pmatrix},
\]

(B4)
where $\rho = \rho(k)$ is a $C^\infty$ function of $k$. Let

$$e_M = (1, 0, 0, \ldots, 0)^T,$$

and

$$\{G_{n,m}(k, l)\}_{ij} = \frac{(i - 1)(-1)^{i+j}}{i + j - 2} \frac{1}{(k - l)^{i+j-1}}.$$  \hspace{1cm} (B5)

Then the basic elements in (43) are expressed as

$$F_1 = \text{diag}(F_{n_1}(\rho(k_1)), \ldots, F_{n_p}(\rho(k_p))),\quad H_1 = \text{diag}(H_{n_1}(\rho(k_1)), \ldots, H_{n_p}(\rho(k_p))),$$

$$F_2 = \text{diag}(F_{m_1}(\sigma(l_1)), \ldots, F_{m_q}(\sigma(l_q))),\quad H_2 = \text{diag}(H_{m_1}(\sigma(l_1)), \ldots, H_{m_q}(\sigma(l_q))),$$

$$E_1^T = \left( e_{n_1}^T, \ldots, e_{n_p}^T \right),\quad E_2^T = \left( e_{m_1}^T, \ldots, e_{m_q}^T \right),$$

$$(G_1)_{i,j} = G_{n_i,m_j}(k_i, l_j),\quad (G_2)_{j,i} = G_{m_j,n_i}(l_j, k_i),$$  \hspace{1cm} (B6)

where the plane wave factors are given by

$$\rho(k_i) = \exp \left( \sum_{n \in \mathbb{Z}} k^n x_n \right) \rho^{(0)}_i,\quad \sigma(l_j) = \exp \left( - \sum_{n \in \mathbb{Z}} l^n x_n \right) \sigma^{(0)}_j,\quad \rho^{(0)}_i,\quad \sigma^{(0)}_j \in \mathbb{C}.  \hspace{1cm} (B7)$$

**APPENDIX C: EXAMPLES OF SOLUTIONS**

**C.1 One-soliton solution**

When $N = 1$ we have

$$K_1 = k_1,\quad r_1 = \rho_1,\quad s_1 = \sigma_1,\quad M_1 = -(\rho_1 \sigma_1^*)/(|k_1|^2 + 1),$$  \hspace{1cm} (C1)

where (noting that the only difference between $\rho_j$ and $\sigma_j$ is the phase factor $\rho_j^{(0)}$ and $\sigma_j^{(0)}$)

$$\rho_j = \exp \left( \left( k^n_j + (-1)^{n+1} k^{-n}_j \right) \xi_n + i \left( k^n_j - (-1)^{n+1} k^{-n}_j \right) \eta_n \right) \rho_j^{(0)};$$

$$\sigma_j = \exp \left( \left( k^n_j + (-1)^{n+1} k^{-n}_j \right) \xi_n + i \left( k^n_j - (-1)^{n+1} k^{-n}_j \right) \eta_n \right) \sigma_j^{(0)}.  \hspace{1cm} (C2)$$

Hence the one-soliton solution is given by

$$v = \begin{pmatrix} 1 - s_1^{(-1,0)} & -s_3^{(-1,0)} \\ -s_3^{(-1,0)} & 1 - s_4^{(-1,0)} \end{pmatrix},$$  \hspace{1cm} (C3)

where

$$s_1^{(-1,0)} = -\frac{1}{|k_1|^2} \frac{|\rho_1 \sigma_1|^2 (|k_1|^2 + 1)}{(|k_1|^2 + 1)^2 - |\rho_1 \sigma_1|^2},$$
\[ s_3^{(-1,0)} = \frac{1}{k_1} \frac{\rho_1 \sigma_1 (|k_1|^2 + 1)}{(|k_1|^2 + 1)^2 - |\rho_1 \sigma_1|^2}, \]
\[ s_4^{(-1,0)} = -\frac{|\rho_1 \sigma_1|^2 (|k_1|^2 + 1)}{(|k_1|^2 + 1)^2 - |\rho_1 \sigma_1|^2}. \quad (C4) \]

**C.2 | Two-soliton solution**

In the case \( N = 2 \), we assume

\[
K_1 = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}, \quad r_1 = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \quad s_1 = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}, \quad (C5)
\]

and we have

\[
M_1 = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} -(\rho_1 \sigma_1^*)/(|k_1|^2 + 1) & -(\rho_2 \sigma_2^*)/(|k_2|^2 + 1) \\ -(\rho_2 \sigma_1^*)/(k_1^* k_2 + 1) & -(\rho_2 \sigma_2^*)/(|k_2|^2 + 1) \end{pmatrix}, \quad (C6)
\]

where \( \rho_i \) and \( \sigma_i \) are defined by (C2). Introduce

\[
T = (I_2 - M_1^* M_1)^{-1} = \frac{1}{\tau} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}, \quad (C7)
\]

where

\[
\tau = |I_2 - M_1^* M_1| = 1 - \frac{|\rho_1 \sigma_1|^2}{(|k_1|^2 + 1)^2} - \frac{|\rho_2 \sigma_2|^2}{(|k_2|^2 + 1)^2} - \frac{\rho_1 \sigma_1 \rho_2 \sigma_2^*}{(k_1^* k_2 + 1)^2} - \frac{\rho_1 \sigma_1^* \rho_2 \sigma_2}{(k_1^* k_2 + 1)^2} + \frac{[(|k_1|^2 + 1) (|k_2|^2 + 1) - (k_1^* k_2 + 1)(k_1^* k_2 + 1)]^2}{[(|k_1|^2 + 1)(|k_2|^2 + 1)(k_1^* k_2 + 1)(k_1^* k_2 + 1)]^2}, \quad (C8)
\]

and

\[
T_{11} = 1 - \frac{\rho_1 \sigma_1 \rho_2 \sigma_2^*}{(k_1 k_2^* + 1)^2} - \frac{|\rho_2 \sigma_2|^2}{(|k_2|^2 + 1)^2}, \quad T_{12} = \frac{\rho_1 \rho_2^* |\sigma_1|^2}{(|k_1|^2 + 1)(k_1^* k_2^* + 1)} + \frac{|\rho_2|^2 \sigma_2^* \sigma_2}{(|k_2|^2 + 1)(k_1^* k_2 + 1)},
\]

\[
T_{21} = \frac{|\rho_1|^2 \sigma_1 \sigma_2^*}{(|k_1|^2 + 1)(k_1^* k_2^* + 1)} + \frac{\rho_1^* |\rho_2|^2 |\sigma_2|^2}{(k_1^* k_2 + 1)(|k_2|^2 + 1)}, \quad T_{22} = 1 - \frac{\rho_1 \sigma_1 |\sigma_1|^2}{(|k_1|^2 + 1)^2} - \frac{\rho_1^* \sigma_1 \rho_2 \sigma_2}{(k_1^* k_2 + 1)^2}.
\quad (C9)
In addition, let

\[
P = (I_2 - M_1 M_1^*)^{-1} M_1^* = \frac{1}{\tau} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \frac{1}{\tau} \begin{pmatrix} T_{11} m_{11}^* + T_{12} m_{21}^* & T_{11} m_{12}^* + T_{12} m_{22}^* \\ T_{21} m_{11}^* + T_{22} m_{21}^* & T_{21} m_{12}^* + T_{22} m_{22}^* \end{pmatrix},
\]

\[
Q = M_1 (I_2 - M_1 M_1^*)^{-1} = \frac{1}{\tau} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \frac{1}{\tau} \begin{pmatrix} m_{11} T_{11} + m_{12} T_{21} & m_{11} T_{12} + m_{12} T_{22} \\ m_{21} T_{11} + m_{22} T_{21} & m_{21} T_{12} + m_{22} T_{22} \end{pmatrix}.
\]

(C10)

Those elements in \( \mathbf{v} \) turn out to be

\[
s_{1}^{(-1,0)} = \frac{1}{\tau} \left( \frac{1}{|k_1|^2} \rho_1 \sigma_1^* P_{11} + \frac{1}{k_1 k_2} \rho_1 \sigma_2^* P_{21} + \frac{1}{k_1 k_2} \rho_2 \sigma_1^* P_{12} + \frac{1}{|k_2|^2} \rho_2 \sigma_2^* P_{22} \right),
\]

(C11a)

\[
s_{2}^{(-1,0)} = \frac{1}{\tau} \left( \frac{1}{k_1} \rho_1 \sigma_1 T_{11} + \frac{1}{k_1} \rho_1 \sigma_2 T_{21} + \frac{1}{k_2} \rho_2 \sigma_1 T_{12} + \frac{1}{k_2} \rho_2 \sigma_2 T_{22} \right),
\]

(C11b)

\[
s_{3}^{(-1,0)} = \frac{1}{\tau} \left( \rho_1^* \sigma_1 Q_{11} + \rho_1^* \sigma_2 Q_{21} + \rho_2^* \sigma_1 Q_{12} + \rho_2^* \sigma_2 Q_{22} \right).
\]

(C11c)

### C.3 Jordan block solution

When \( K_1 \) is a \( 2 \times 2 \) Jordan matrix, we have

\[
K_1 = \begin{pmatrix} k & 0 \\ 1 & k \end{pmatrix}, \quad r_1 = \begin{pmatrix} \rho \\ \partial_k \rho \end{pmatrix}, \quad s_1 = \begin{pmatrix} \sigma \\ \partial_k \sigma \end{pmatrix},
\]

(C12)

and

\[
M_1 = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix},
\]

\[
T = (I_2 - M_1 M_1^*)^{-1} = \frac{1}{\tau} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \frac{1}{\tau} \begin{pmatrix} 1 - m_{21}^* m_{12} - |m_{22}|^2 & m_{11}^* m_{12} + m_{12}^* m_{22} \\ m_{21}^* m_{11} + m_{22}^* m_{21} & 1 - |m_{11}|^2 - m_{12}^* m_{22} \end{pmatrix},
\]

(C14)

where \( \rho \) and \( \sigma \) are defined as in (C2) and we have dropped off the index \( j \) without making any confusions. In this case,

\[
T = (I_2 - M_1 M_1^*)^{-1} = \frac{1}{\tau} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \frac{1}{\tau} \begin{pmatrix} 1 - m_{21}^* m_{12} - |m_{22}|^2 & m_{11}^* m_{12} + m_{12}^* m_{22} \\ m_{21}^* m_{11} + m_{22}^* m_{21} & 1 - |m_{11}|^2 - m_{12}^* m_{22} \end{pmatrix},
\]

(C14)

where

\[
\tau = |I_2 - M_1^* M_1| = 1 - m_{21}^* m_{12} - m_{12}^* m_{21} - |m_{11}|^2 - |m_{22}|^2
\]

\[
+ |m_{11}|^2 |m_{22}|^2 + |m_{12}|^2 |m_{21}|^2 - m_{11} m_{22} m_{12} m_{21} - m_{12} m_{22} m_{11} m_{21}
\]

\[
= 1 - 2 \text{Re}[m_{12}^* m_{21} + m_{11} m_{22} m_{12}^* m_{21}^*] - |m_{11}|^2 - |m_{22}|^2 + |m_{11}|^2 |m_{22}|^2 + |m_{12}|^2 |m_{21}|^2.
\]

(C15)

Then \( \mathbf{v} \) can be given via formula (C11) with (C10), while all \( m_{ij} \)s and \( T_{ij} \)s should be taken from this subsection.