Bispectral duality and separation of variables from surface defect transition

Saebyeok Jeong [ORCID]^a **and Norton Lee** [ORCID]^b

^{a}*Department of Theoretical Physics, CERN,*
1211 Geneva 23, Switzerland

^{b}*Center for Geometry and Physics, Institute for Basic Science (IBS),*
Pohang 37673, Republic of Korea

E-mail: saebyeok.jeong@cern.ch, nortonxs@gmail.com

ABSTRACT: We study two types of surface observables − the **Q**-observables and the **H**-observables − of the 4d $\mathcal{N} = 2$ A_1-quiver U(N) gauge theory obtained by coupling a 2d $\mathcal{N} = (2,2)$ gauged linear sigma model. We demonstrate that the transition between the two surface defects manifests as a Fourier transformation between the surface observables. Utilizing the results from our previous works, which establish that the **Q**-observables and the **H**-observables give rise, respectively, to the Q-operators on the evaluation module over the Yangian $Y(\mathfrak{gl}(2))$ and the Hecke operators on the twisted $\widehat{\mathfrak{sl}}(N)$-coinvariants, we derive an exact duality between the spectral problems of the $\mathfrak{gl}(2)$ XXX spin chain with N sites and the $\mathfrak{sl}(N)$ Gaudin model with 4 sites, both of which are defined on bi-infinite modules. Moreover, we present a dual description of the monodromy surface defect as coupling a 2d $\mathcal{N} = (2,2)$ gauged linear sigma model. Employing this dual perspective, we demonstrate how the monodromy surface defect undergoes a transition to multiple **Q**-observables or **H**-observables, implemented through integral transformations between their surface observables. These transformations provide, respectively, $\hbar$-deformation and a higher-rank generalization of the KZ/BPZ correspondence. In the limit $\varepsilon_2 \to 0$, they give rise to the quantum separation of variables for the $\mathfrak{gl}(2)$ XXX spin chain and the $\mathfrak{sl}(N)$ Gaudin model, respectively.

Contents

1 Introduction

In this paper, we explain how transitions between surface defects in the 4d $\mathcal{N} = 2$ gauge theory reveal new features of its correspondence with a quantum integrable system. We consider the $\mathcal{N} = 2$ theory of class $\mathcal{S}$ [1] of type A_{N-1} associated with the cylinder $\mathbb{C}^\times = \mathbb{P}^1 \setminus \{0, \infty\}$, with n marked points $S = \{p_1, p_2, \cdots, p_n\} \subset \mathbb{C}^\times$. Here, we choose the marked point data such that the $\mathcal{N} = 2$ theory has a Lagrangian description as the linear A_{n-1}-quiver $\mathrm{SU}(N)$ gauge theory in an appropriate S-duality frame. The associated integrable system is defined on the moduli space $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$ of Hitchin's equations on $\mathbb{P}^1$ with (regular) ramification

data at $S \cup \{0, \infty\}$ [2]. We primarily restrict to the case $n = 2$ (keeping N arbitrary), but we set n generic in the introduction to provide a comprehensive overview.

We study this $\mathcal{N} = 2$ gauge theory subject to the Donaldson-Witten twist [3] and further to the $\Omega_{\varepsilon_1, \varepsilon_2}$-deformation associated with the $U(1)^2$ isometry of spacetime $\mathbb{C}^2$ [4]. Compactifying along the torus fiber of $\mathbb{C}^2 \simeq \Sigma \tilde{\times} T^2$, we arrive at a two-dimensional $\mathcal{N} = (2, 2)$ sigma model of maps from the worldsheet Σ to the Hitchin moduli space target $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$ [5, 6]. In the cohomology of the preserved supercharge, the sigma model is either an A-model associated with a symplectic structure or a B-model associated with a complex structure, determined by the ratio $\kappa = -\frac{\varepsilon_2}{\varepsilon_1}$ of the Ω-background parameters [7, 8].

In the context of the topological sigma model, the quantization of the moduli space $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$ is realized by open strings stretched between A-branes [9]. The $(\mathcal{B}_{cc}, \mathcal{B}_{cc})$-strings, where $\mathcal{B}_{cc}$ is the space-filling canonical coisotropic brane [10], form the twisted differential operators on the moduli space $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$ of parabolic $\mathrm{PGL}(N)$-bundles. They provide a non-commutative deformation of the algebra of holomorphic functions on $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$. Then, any A-brane defines a twisted $\mathcal{D}$-module on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$, acted on by this non-commutative algebra through the action of joining the open strings.

In our previous work [8], we studied how half-BPS surface defects in the $\mathcal{N} = 2$ gauge theory are associated with certain A-branes and functors acting on them. In particular, the regular monodromy surface defect descends to the brane of λ-connections, while the canonical surface defect, which we will refer to as the **H**-observable in this work, leads to the Hecke operator. The *magnetic* eigenbrane on which the Hecke operators act *diagonally* was revealed to have its origin in the boundary condition at infinity.[1] From the perspective of the geometric Langlands correspondence, these eigenbranes are associated with special $\mathrm{SL}(N)$-local systems called opers [12, 13], and we showed how these opers are parametrized by the Coulomb moduli of the $\mathcal{N} = 2$ theory [8, 14]. We showed that the vacuum expectation value of the regular monodromy surface defect in the limit $\varepsilon_2 \to 0$ provides a distinguished basis for the sections of the twisted $\widehat{\mathfrak{sl}}(N)$-coinvariants on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$, on which the Hecke operators act diagonally. We also proved that the same vev of the regular monodromy surface defect is a common eigenfunction of the quantum Hamiltonians of the associated $\mathfrak{sl}(N)$ Gaudin model. We suggested a string duality that relates our $\mathcal{N} = 2$ gauge theoretical formulation of the geometric Langlands correspondence (with ramifications) [12, 13, 15, 16] to the more conventional approach from the topologically twisted (GL-twisted) $\mathcal{N} = 4$ gauge theory [11, 17–19]. See also [20–23] for the vertex algebra approach to the geometric Langlands correspondence.

Meanwhile, the same Hitchin integrable system is approached from a dual perspective in the companion paper [24]. The moduli space $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$ is known to be isomorphic, as a hyper-Kähler space, to the moduli space of periodic $U(n)$-monopoles on $\mathbb{P}^1 \times S^1$ with a framing at $\infty \in \mathbb{P}^1$ and Dirac singularities at $D \times \{0\}$, where $D = \{m_0^+, m_1^+, \cdots, m_{N-1}^+\} \subset \mathbb{P}^1 \setminus \{\infty\} = \mathbb{C}$, related to each other by a Nahm transform [25–29]. In one of the complex structures, it is viewed as the moduli space $\mathcal{M}_{\mathrm{mHiggs}}(\mathrm{GL}(n), \mathbb{P}^1; D)$ of *multiplicative* $\mathrm{GL}(n)$-Higgs bundles on $\mathbb{P}^1$ with a framing at $\infty \in \mathbb{P}^1$ and regular singularities at D [25, 29, 30]. The

---

[1]It is called the magnetic eigenbrane since the Hecke operators are realized by the 't Hooft line defects in the GL-twisted $\mathcal{N} = 4$ theory [11].

quantization of $\mathcal{M}_{\mathrm{mHiggs}}(\mathrm{GL}(n), \mathbb{P}^1; D)$ is known to lead to a representation of the Yangian of $\mathfrak{gl}(n)$, which we denote by $Y(\mathfrak{gl}(n))$ [27–29].

In the companion paper [24], we found that the vacuum expectation value of the regular monodromy surface defect gives a distinguished basis for the evaluation module over the Yangian $Y(\mathfrak{gl}(n))$. We showed that the $\mathbf{Q}$-observable, defined as a surface defect in the $\mathcal{N} = 2$ gauge theory, acts on this Yangian module as the Q-operator. In particular, we derived the universal $\hbar$-oper difference equation (equivalently, Baxter's TQ equation) satisfied by the $\mathbf{Q}$-observables using the constraints of the surface defect correlation functions. The action of the Q-operators on the distinguished basis was shown to be diagonal due to the cluster decomposition of the surface defects. Combined with the universal $\hbar$-oper equation, it followed that the Q-eigenstates, constructed as the vevs of the regular monodromy surface defect, are also the common eigenstates of the quantum Hamiltonians of the $\mathfrak{gl}(n)$ XXX spin chain. This statement was regarded as the $\hbar$-deformation of the geometric Langlands correspondence (restricted to the locus of ($\hbar$-)opers) [12, 13], called the $\hbar$-*Langlands correspondence*.[2]

1.1 Q-operator/Hecke operator transition and bispectral duality

Let us revisit the isomorphism between the hyper-Kähler target spaces at the level of the topological sigma model. We note that any A-brane $\mathcal{B}'$, which provides a twisted $\mathcal{D}$-module on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$, is expected to give a $\hbar$-*difference module* on $\mathrm{Bun}_{\mathrm{GL}(n)}(\mathbb{P}^1; D)$, since they are two dual interpretations of the same $(\mathcal{B}_{cc}, \mathcal{B}')$-strings acted on by the $(\mathcal{B}_{cc}, \mathcal{B}_{cc})$-strings, which quantize the algebra of holomorphic functions on the same moduli space.

Our findings in the companion paper [24] give an example of such a duality, realized at specific A-branes descending from the regular monodromy surface defect and the boundary condition at infinity of the $\mathcal{N} = 2$ gauge theory. In the context of the ordinary Langlands correspondence, the regular monodromy surface defect (at fixed monodromy parameters) gives rise to the E-twisted coinvariants of $\widehat{\mathfrak{sl}}(N)$-modules, where $E \in \mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$ is determined by the monodromy parameters [8]. In the context of the $\hbar$-Langlands correspondence, we explained in the companion paper [24] that it leads to elements in an evaluation module $\mathcal{H}$ over the Yangian $Y(\mathfrak{gl}(n))$. Moreover, for the magnetic eigenbrane $\mathbf{F_a}$ descending from the boundary condition $\mathbf{a}$ at infinity, the section of $(\mathcal{B}_{cc}, \mathbf{F_a})$-strings is obtained as the vacuum expectation value of the regular monodromy surface defect at the vacuum $\mathbf{a}$ in the limit $\varepsilon_2 \to 0$, with varying monodromy parameters. It is a section of the Hecke eigensheaf and a Q-eigenstate at the same time, corresponding to a $\mathrm{SL}(N)$ oper and a $\mathrm{GL}(n)$ $\hbar$-oper, respectively, determined by the same Coulomb moduli $\mathbf{a}$.

Therefore, it is natural to investigate the relation between the Q-operator and the Hecke operator, since their eigen-object is realized by the same A-brane. This is the first goal of the present work. We describe a transition between the Q-operator and the Hecke operator using their realization as surface defects in the $\mathcal{N} = 2$ gauge theory, and study its implications for the associated integrable models.

---

[2]The $\hbar$-Langlands correspondence (additive) is closely related to but different from the q-Langlands correspondence (multiplicative), since the latter is relevant to the moduli space of multiplicative Higgs bundles on $\mathbb{P}^1$ with framing at $0, \infty \in \mathbb{P}^1$. See [31–35] for previous works on the q-Langlands correspondence.

Note that the Q-operator, which is the $\hbar$-analogue of the Hecke operator, is realized as a surface defect in the $\mathcal{N} = 2$ gauge theory created by coupling a two-dimensional $\mathcal{N} = (2, 2)$ gauged linear sigma model with the vev of the complex adjoint scalar turned on [24]. Meanwhile, recall that in our $\mathcal{N} = 2$ gauge theoretical formulation of the ordinary Langlands correspondence [8], the Hecke operator follows from the **H**-observable, which is defined by coupling the two-dimensional gauged linear sigma model with the same field contents but in the non-linear sigma model phase instead. Thus, we explain the relation between the Q-operator and the Hecke operator through a transition in the surface defect theory, where the two observables are distinguished by turning on either the vev of the complex adjoint scalar or the complexified FI parameter. This surface defect transition manifests as a Fourier transform between their observable expressions [8, 14], which schematically reads

$$\mathbf{H}(y) = \oint_{\mathcal{C}} dy\, y^{-\frac{x}{\varepsilon_1}} \mathbf{Q}(x), \qquad y \in \mathbb{C}^\times \setminus S. \tag{1.1}$$

This equation directly connects the Q-operator and the Hecke operator in their gauge theoretical realization. We emphasize that the contour $\mathcal{C}$ is explicitly given to pick up certain semi-infinite ε_1-lattice of simple poles in the integrand. There are N choices for the contour, corresponding to the N choices of the vacuum at infinity of the two-dimensional non-linear sigma model, and also to the N fundamental coweights of $\mathrm{PGL}(N)$, each of which defines a Hecke operator. Additionally, the **Q**-observable in the integrand is selected from n inequivalent **Q**-observables according to the domain of $y \in \mathbb{C}^\times \setminus S$. In particular, the Fourier transform is not just a formal one but is evaluated as a concrete convergent series.

The transformation explains the existence of the common eigen-objects in an explicit manner. Moreover, the (universal) $\mathrm{GL}(n)$ $\hbar$-oper equation and the (universal) $\mathrm{SL}(N)$ oper equation satisfied by them are shown to be exchanged with each other precisely under the transformation. In terms of the associated integrable models − the $\mathfrak{gl}(n)$ XXX spin chain defined on N bi-infinite modules and the $\mathfrak{sl}(N)$ Gaudin model defined on two (co-)Verma modules and n bi-infinite modules − the transition induces an exact duality between their spectral problems, giving a one-to-one matching of their quantum Hamiltonians as well as their quantum spectra in the $\mathcal{N} = 2$ gauge theoretical terms.

1.2 Surface defect transition and separation of variables

Next, we move on to another point of view on the moduli space $\mathcal{M}_H(\mathrm{SU}(N), \mathbb{P}^1; S)$ of Hitchin's equations. When viewed as the moduli space of parabolic Higgs bundles $\mathcal{M}_{\mathrm{Higgs}}(\mathrm{PGL}(N), \mathbb{P}^1; S)$, it is birational to the affine deformation of the cotangent bundle over $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$. We observe that there is also another birational map [36]

$$\mathcal{M}_{\mathrm{Higgs}}(\mathrm{PGL}(N), \mathbb{P}^1; S) \longrightarrow (T^*\mathbb{C}^\times)^{[(n-1)(N-1)]}, \tag{1.2}$$

which is holomorphic symplectic in open dense subsets. Here, $[(n-1)(N-1)]$ denotes the Hilbert scheme of points of length $(n-1)(N-1)$. Importantly, the map provides the classical separation of variables for the associated integrable model.

When we regard $(T^*\mathbb{C}^\times)^{[(n-1)(N-1)]}$ as the target space of the topological sigma model compatible with the birational map ($\varepsilon_2 = 0$ and $\kappa = 0$; the A-model in the symplectic structure

ω_K in the convention of [11]), the A-branes now give rise to $\mathcal{D}$-modules on $\mathrm{Sym}^{(n-1)(N-1)}\mathbb{C}^\times$. In particular, the counterpart of the Hecke eigensheaf Δ_ρ on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$ associated with an oper ρ is identified as the $\mathcal{D}$-module on $\mathrm{Sym}^{(n-1)(N-1)}\mathbb{C}^\times$ defined by the oper differential equation for ρ [37]. The sections of the two $\mathcal{D}$-modules are related by a formal integral transform [37]. In other words, the section of the Hecke eigensheaf associated with the oper ρ is expressed as an integral transform of the product of $(n-1)(N-1)$ independent oper solutions. Recalling that this Hecke eigensheaf is also the twisted $\mathcal{D}$-module representing the spectral equations for the quantum Gaudin Hamiltonians, we observe that the integral transform precisely induces the quantum separation of variables for the Gaudin model.[3]

In fact, there is a dual perspective here as well. The image of the birational map can simply be viewed as $(\mathbb{C} \times \mathbb{C}^\times)^{[(n-1)(N-1)]}$, since the cotangent bundle on $\mathbb{C}^\times$ is trivial. By exchanging the base and the fiber, the A-branes should lead to $\hbar$-difference modules on $\mathrm{Sym}^{(n-1)(N-1)}\mathbb{C}$. By analogy with the dual case, we expect that the A-brane for the Q-eigenstate (magnetic eigenbrane) would be related to the $\hbar$-difference module on $\mathrm{Sym}^{(n-1)(N-1)}\mathbb{C}$ representing the $\hbar$-oper difference equation. The Q-eigenstate would be expressed as an integral transform of the product of $(n-1)(N-1)$ independent $\hbar$-oper solutions, i.e., Q-functions. Recalling that a Q-eigenstate is also a common eigenstate of the Yangian transfer matrices represented on the relevant evaluation module, such a transformation would lead to the quantum separation of variables for the associated integrable model in the bispectral dual perspective; namely, the XXX spin chain.

The second goal of the present work is to elucidate these separation of variable transformations as transitions of the surface defect in the $\mathcal{N} = 2$ gauge theory; specifically, from a regular monodromy surface defect to multiple **Q**-observables or **H**-observables. In fact, the $\mathcal{N} = 2$ gauge theoretical formulation enhances the previous analysis in two aspects. Firstly, we establish the surface defect transition at generic values of ε_2 (namely, $\kappa \neq 0$), and the integral transformations between the correlation functions are valid even away from $\varepsilon_2 = 0$. Secondly, the integral transformations are not just formal ones but are obtained as definite contour integrals with explicit contour prescriptions.

We begin by giving a dual description of the monodromy surface defect as coupling to a two-dimensional $\mathcal{N} = (2, 2)$ sigma model, by which the monodromy of the fields can be thought of as the remnant of integrating out two-dimensional degrees of freedom [38]. The target space of this sigma model is the total space of a vector bundle over a flag variety, and the complexified Kähler parameters are identified with the monodromy parameters. The duality exhibits the surface defect observable as an equivariant integral over a hand-saw quiver variety [39, 40] modified by the effect of the coupling to the four-dimensional theory. We then derive a crossing formula relating the equivariant integrals performed in two different stability chambers of the modified hand-saw quiver variety. The observable expression in the other (negative) stability chamber is shown to admit two contour integral representations. One has the product of **Q**-observables in the integrand with a contour of Barnes type. The other has the product of **H**-observables in the integrand with a contour of Pochhammer type.

---

[3]The term *quantum* sometimes indicates a $\varepsilon_2 \neq 0$ deviation, but the *quantum* separation of variables here is relevant to $\varepsilon_1 = \hbar \neq 0$ and $\varepsilon_2 = 0$, as opposed to the classical separation of variables at $\varepsilon_1 = \varepsilon_2 = 0$.

Namely, upon taking the vacuum expectation values of the two relations, we arrive at

$$\Big\langle \Psi(\mathbf{u}) \Big\rangle_{\mathbf{a}} = \int_{\mathcal{C}} d\mathbf{x}\, \mathcal{K}(\mathbf{u}; \mathbf{x}) \left\langle \prod_{i=1}^{(n-1)(N-1)} \mathbf{Q}(x_i) \right\rangle_{\mathbf{a}} = \int_{\mathrm{PC}} d\mathbf{y}\, \mathcal{K}(\mathbf{u}; \mathbf{y}) \left\langle \prod_{i=1}^{(n-1)(N-1)} \mathbf{H}(y_i) \right\rangle_{\mathbf{a}} .$$
(1.3)

We emphasize again that the integral transformations are derived at generic values of ε_1 and ε_2. The correlation functions of the surface observables in the integrands are not factorized in this general case. In the second integral transformation, the vacuum expectation value of the regular monodromy surface defect is identified with a twisted coinvariant of two (co-)Verma modules and n bi-infinite $\widehat{\mathfrak{sl}}(N)$-modules [41–43], while the correlation function of the canonical surface defects on the right-hand side is a $\mathcal{W}_N$-algebra conformal block with two generic vertex operators, n semi-degenerate vertex operators, and $(n-1)(N-1)$ fully degenerate vertex operators at $(y_i)_{i=1}^{(n-1)(N-1)}$ [14, 41, 44, 45]. In this sense, it provides higher-rank generalizations of the KZ/BPZ correspondence studied in [37, 46, 47].[4] See [49, 50] for a vertex algebra approach to the higher-rank generalization of the KZ/BPZ correspondence, and [51] for a vertex algebra approach to the surface defect transition and the separation of variables.

In the limit $\varepsilon_2 \to 0$, the correlation functions in the integrand factorize as the product of individual vacuum expectation values due to the cluster decomposition of the surface defects, since they are local on the topological $\mathbb{C}_2$-plane. Recall that in the limit $\varepsilon_2 \to 0$, the vacuum expectation value of the $\mathbf{Q}$-observable is a $\mathrm{GL}(n)$ $\hbar$-oper solution [24], while the vacuum expectation value of the $\mathbf{H}$-observable is an $\mathrm{SL}(N)$ oper solution [8, 14]. Therefore, we precisely obtain the quantum separation of variables for the XXX spin chain and the Gaudin model, respectively, from the two integral transformations. Note that the procedure provides the separation of variables not just at a formal level but is equipped with definite contour prescriptions for the integral transformations.

1.3 Twisted M-theory perspective

The $\mathcal{N} = 2$ gauge theory with its half-BPS surface defect can be constructed in the IIB string theory for the gauge origami [41, 52, 53]. While our primary methodology involves the use of supersymmetric localization in the IIB gauge origami setup, it is instructive to consider an alternative duality frame to enrich our understanding.

It was suggested in [8] that the IIB gauge origami is related to the twisted M-theory [54] by a string duality, after unrefining the Ω-background so that the number of independent parameters is reduced from three to two. We are led to the twisted M-theory defined on the 11-dimensional worldvolume composed of the topological part $\mathbb{C}_{\varepsilon_1} \times \mathbb{C}_{\varepsilon_2} \times \mathbb{C}_{\varepsilon_4} \times \mathbb{R}$ and the holomorphic part $\mathbb{C} \times \mathbb{C}^\times$. The Ω-background associated with the $\mathrm{U}(1)^2$ isometry of $\mathbb{C}_{\varepsilon_1} \times \mathbb{C}_{\varepsilon_2} \times \mathbb{C}_{\varepsilon_4}$ is turned on, with two independent parameters ($\varepsilon_1 + \varepsilon_2 + \varepsilon_4 = 0$) to keep the supersymmetry.[5]

---

[4]The first integral transformation is expected to be interpreted as the *$\hbar$-KZ/$\hbar$-BPZ correspondence*, relating the $\hbar$-conformal blocks of the quantum vertex algebras for the Yangian double and the $\hbar$-deformed $\mathcal{W}$-algebra. See section 5 for a discussion. See [48] for a $\mathcal{N} = 2$ theory realization of the $\hbar$-deformed $\mathcal{W}$-algebra.

[5]We used the convention that the unrefinement $\varepsilon_3 = 0$ was made in the IIB gauge origami. See [8].

IIA/M branes	0	1	2	3	4	5	6	7	8	9	10
D4/M5	x	x	x	x	x						x
NS5/M5	x	x	x	x					x	x	
KK5$_l$/KK6$_l$	x	x			x	x			x	x	x
D2/M2	x	x				x					

Table 1. Twisted M-theory with M-branes and A_{l-1} singularity.

| $\mathbb{C}_{\varepsilon_1}$	$\mathbb{C}_{\varepsilon_2}$	$\mathbb{C}_{\varepsilon_4}$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{C}^\times$
x^0, x^1	x^2, x^3	x^6, x^7	x^5	x^8, x^9	x^4, x^{10}

Table 2. Spacetime of the twisted M-theory.

To begin with, we consider two stacks of M5-branes supported on $\mathbb{C}_{\varepsilon_1} \times \mathbb{C}_{\varepsilon_2}$ in the topological directions: a stack of N M5-branes wrapping $\mathbb{C}^\times$ and another stack of n M5-branes wrapping $\mathbb{C}$ in the holomorphic directions. This is precisely the M-theory uplift of the IIA brane setup where the four-dimensional $\mathcal{N} = 2$ linear A_{n-1}-quiver $SU(N)$ gauge theory is engineered on the worldvolume of N D4-branes stretched between n NS5-branes and infinity [55], augmented by the Ω-background (the first two rows in table 1). Putting differently, we build the $\mathcal{N} = 2$ theory of class $\mathcal{S}$ from the six-dimensional $\mathcal{N} = (0,2)$ theory of type A_{N-1} on the worldvolume of the N M5-branes, viewing the holomorphic cylinder $\mathbb{C}^\times$ as the associated Riemann surface. The other n M5-branes are transverse to $\mathbb{C}^\times$, located at $S = \{p_1, p_2, \cdots, p_n\} \subset \mathbb{C}^\times$. They provide codimension-two defects in the six-dimensional $\mathcal{N} = (0,2)$ theory, thereby determining the contents of the $\mathcal{N} = 2$ gauge theory. In the associated Hitchin moduli space, these defects are responsible for the ramifications at the marked points S [17].

The moduli space of multiplicative Higgs bundles emerges in a T-dual IIB frame [25, 26]. We may compactify x^3 and T-dualize along this direction (at least in the absence of the Ω-background). In turn, there are N D3-branes suspended between n NS5-branes and infinity. The D3-branes stretched between NS5-branes precisely give periodic monopoles in the effective theory on the transverse part $\mathbb{C} \times S^1$ of the worldvolume of the NS5-branes. The N D3-branes stretched between the NS5-brane and infinity give Dirac singularities at their positions $D = \{m_0^+, m_1^+, \cdots, m_{N-1}^+\} \subset \mathbb{C}$. This duality, which exchanges the underlying effective theory between the one on the N M5-branes wrapping $\mathbb{C}^\times$ and the other on the n M5-branes wrapping $\mathbb{C}$, conceptually explains the isomorphism between the moduli space of Higgs bundles on $\mathbb{C}^\times$ and the moduli space of multiplicative Higgs bundles on $\mathbb{C}$ with the rank of the bundle and the number of regular singularities swapped between N and n.

Next, let us turn to the surface defects in the class $\mathcal{S}$ theory. There are two different origins of the surface defects in the twisted M-theory setup: the A_{l-1}-singularity and the M2-branes. In this work, we consider M2-branes wrapping $\mathbb{C}_{\varepsilon_1} \times \mathbb{R}$, which are local on the holomorphic planes $\mathbb{C} \times \mathbb{C}^\times$ (the fourth row of table 1). They can end on either stack of the M5-branes, the one supported on $\mathbb{C}^\times$ or the one supported on $\mathbb{C}$. We refer to the surface defect constructed from the former (resp. the latter) as the **H**-observable (resp. the **Q**-observable). They respectively give rise to the Hecke operator and the Q-operator in our $\mathcal{N} = 2$ gauge
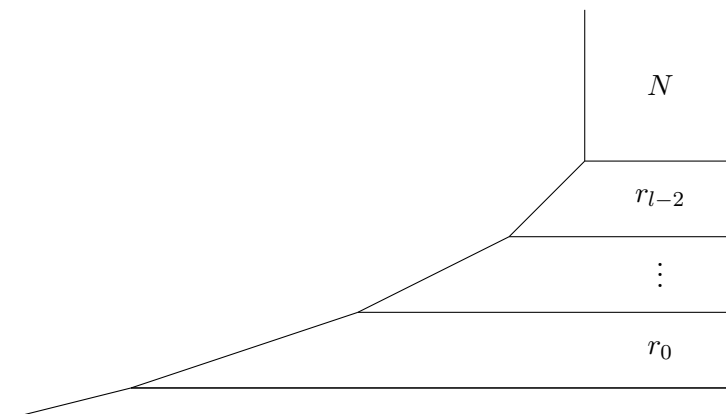
Figure 1. Resolution of $\mathbb{C}_{\varepsilon_1} \times (\mathbb{C}_{\varepsilon_2} \times \mathbb{C}_{\varepsilon_4})/\mathbb{Z}_l$ and M5-branes wrapping $\mathbb{C}^{\times}$.

theoretical implementation of the ($\hbar$-)Langlands correspondence, as we extensively study in [8, 24]. We provide more details of the **Q**-observable in section 2.

The A_{l-1}-singularity is introduced in the twisted M-theory by replacing the topological part of the 11-dimensional worldvolume by $\mathbb{C}_{\varepsilon_1} \times (\mathbb{C}_{\varepsilon_2} \times \mathbb{C}_{\varepsilon_4})/\mathbb{Z}_l \times \mathbb{R}$ (the third row in table 1) [54]. It amounts to add Kaluza-Klein 6-monopoles supported on the transverse worldvolume, intersecting the stack of N M5-branes along $\mathbb{C}_{\varepsilon_1} \times \mathbb{C}^{\times}$. Hence it leads to a codimension-two defect in the six-dimensional $\mathcal{N} = (0,2)$ theory, which reduces to a monodromy surface defect lying on $\mathbb{C}_{\varepsilon_1}$ in the four-dimensional $\mathcal{N} = 2$ theory [53, 56]. We will provide more details of the implementation of the monodromy surface defect by the $\mathbb{Z}_l$-orbifold in section 2.

Here, we discuss dual description and transition of the monodromy surface defect that lead to the separation of variables. First, we make a continuous deformation of the metric to bring the A_{l-1}-singularity to the l-centered Taub-NUT space. The singularity is resolved by $l - 1$ copies of $\mathbb{P}^1$, and the M5-branes wrap these compact cycles according to the $\mathbb{Z}_l$-charges that they were assigned before resolution. We assign $\mathbb{Z}_l$-charge 0 for all the n M5-branes wrapping $\mathbb{C}$, while having $\mathbb{Z}_l$-charges of the N M5-branes wrapping $\mathbb{C}^{\times}$ at our disposal. Let r_ω denote the number of these M5-branes carrying $\mathbb{Z}_l$-charge equal to or less than ω for $\omega = 0, 1, \cdots, l - 1$ ($r_{l-1} = N$). Then, we may visualize the M5-brane configuration after the resolution by the toric diagram in figure 1.

Note that the M5-branes wrapping the compact cycles intersect the worldvolume of the four-dimensional $\mathcal{N} = 2$ gauge theory along $\mathbb{C}_{\varepsilon_1}$. Thus, it is expected that they create a surface defect by coupling an effective two-dimensional $\mathcal{N} = (2,2)$ linear A_{l-1}-quiver gauge theory on their worldvolume, providing a dual description for the monodromy surface defect (see figure 7, for our main case of $n = 2$). We will confirm this explicitly by matching their observable expressions in section 4.

Now, let us restrict to the case of the regular monodromy surface defect where we set $l = N$ and $r_\omega = \omega + 1$, $\omega = 0, 1, \cdots, N - 1$. In the IIA frame, we can initiate consecutive D2-brane creations by pulling out the D4-branes wrapping the compact cycles along the topological $\mathbb{R}$-direction (x^5), starting from the bottom of the toric diagram. The D4-branes wrapping adjacent compact cycles can be joined across their junctions, so that they can

be pulled out together while creating $n-1$ D2-brane ending on $n-1$ out of n NS5-branes supported on the non-compact worldvolume $\mathbb{C}_{\varepsilon_1} \times \mathbb{C}_{\varepsilon_2} \times \mathbb{C}$. The net result of the consecutive transitions is $(n-1)(N-1)$ D2-branes stretched between $n-1$ of the NS5-branes and distinct D4-branes, on top of the usual IIA brane setup for the underlying $\mathcal{N} = 2$ gauge theory. This is precisely the IIA brane engineering of the **Q**-observable surface defect [8] (see section 2.1). Thus, the brane transition exhibits the surface defect transition from the regular monodromy surface defect to $(n-1)(N-1)$ **Q**-observables. Using the transition from the **Q**-observable to the **H**-observable (see section 2), we also achieve the surface defect transition from the regular monodromy surface defect to multiple **H**-observables. We will establish these surface defect transitions as concrete integral transformations between their correlation functions in section 4.

The M-theory on a toric Calabi-Yau threefold is dual to the (p,q)-web of fivebranes in IIB theory, where the (p,q)-web is given by the toric diagram [57]. Applying the duality to our setting, the two stacks of M5-branes become two stacks of D3-branes filling the face between the NS5-brane and the D5-brane(s), wrapping $\mathbb{C}^\times$ and $\mathbb{C}$ in the holomorphic direction respectively. In the presence of the A_{l-1}-singularity, the number of D5-branes increases from 1 to l. The D3-branes end on these D5-branes according to the $\mathbb{Z}_l$-charge that the M5-branes were assigned. After resolving the singularity, the l D5-branes are separated as in the (p,q)-web of figure 1, each of which forms a junction with a $(1,\omega)$-brane and a $(1, \omega+1)$-brane, $\omega = 0, 1, \cdots, l-1$. Meanwhile, the M2-branes wrapping $\mathbb{C}_{\varepsilon_1}$ become D1-branes attached to the D5-brane(s). In the effective GL-twisted $\mathcal{N} = 4$ theory on the worldvolume of the D3-branes wrapping $\mathbb{C}^\times$, the NS5-brane gives a (deformed) Neumann boundary, while the D5-brane(s) gives a (deformed) Nahm pole boundary [58]. The D1-branes attached to the D5-branes become the 't Hooft line defects superposed on the Nahm pole boundary. The transition from the A_{N-1}-singularity to multiple M2-branes, which leads to the separation of variables, would be translated to a transition of the Dirichlet boundary into multiple 't Hooft line defects attached to the regular Nahm pole boundary.

Another version of M-brane transition, leading to the surface defect transition and the separation of variables, was also presented in [51]. We anticipate that these two perspectives are connected through a string duality. In particular, the monodromy surface defect is realized by M5-branes there, and they are expected to arise from the M-theory uplift of another IIA theory related to the one that we described earlier by two T-dualities, mapping the KK5-monopoles to D4-branes. See also [59] where the resolution yields gluing of corner vertex algebras [60].

1.4 Outline

In section 2, we explain that the Q-operator and the Hecke operator are realized by the same two-dimensional $\mathcal{N} = (2,2)$ gauged linear sigma model coupled to the four-dimensional $\mathcal{N} = 2$ theory, where the vev of the complex adjoint scalar is turned on in the former while the complexified FI parameter is turned on in the latter. We achieve the transition between the two as the Fourier transform between their surface defect observables. In section 3, we recap how the **Q**-observables and the **H**-observables give rise to the Q-operators and the Hecke operators, respectively. Then, using the Fourier transformation between them, we

derive the bispectral duality between the $\mathfrak{gl}(2)$ XXX spin chain and the $\mathfrak{sl}(N)$ Gaudin model. In section 4, we elucidate the monodromy surface defect has a dual description as coupling a two-dimensional $\mathcal{N} = (2,2)$ sigma model. We establish the crossing formula between the observable expressions in two different stability chambers of the modified hand-saw quiver variety. Using the expression in the negative chamber, we re-express the monodromy defect observable into two different contour integrals, exhibiting its transition into multiple **Q**-observables or **H**-observables. In particular, we obtain definite expressions for the integral kernels and the contour prescriptions. We show the integral transformation leads to the quantum separation of variables for the XXX spin chain and the Gaudin model in the limit $\varepsilon_2 \to 0$. We also confirm that, when applied to some simple cases, the separation of variable transformations we found reproduce the known prescription of Sklyanin. Finally, we conclude in section 5 with discussions. The appendices contain a brief review on the gauge origami constructions of the $\mathcal{N} = 2$ theory and its surface defects and some computational details.

Acknowledgments

2 Q- and H-observables as surface defects

The contents of this section is mainly a review of the **Q**-observables, the **H**-observables, and the monodromy surface defect in the $\mathcal{N} = 2$ supersymmetric A_1-quiver U(N) gauge theory. Nevertheless, we reorganize the known facts from the perspective of the bispectral duality. In particular, we give a description of the first two surface defects as coupling the same two-dimensional $\mathcal{N} = (2,2)$ gauged linear sigma model with either the vev of the complex adjoint scalar or the complexified FI parameter turned on. The variation of the respective parameters is visualized in the IIA brane engineering of the two surface defects. We also provide higgsing constructions of the two surface defects which is in accordance with the suggested IIA brane picture.

2.1 Q-observables

In the IIB gauge origami setup, the **Q**-observables are built by inserting an additional D3-brane that intersects the worldvolume of the original stack of D3-branes for the gauge theory along a surface. To avoid duplication with the companion paper [24], we will present the gauge origami construction of the $\mathbf{Q}/\tilde{\mathbf{Q}}$-observables in the appendix A.2 and A.3.

Here, we explain that they are these surface defects created by coupling a $\mathcal{N} = (2,2)$ two-dimensional gauged linear sigma model whose complexified FI parameter is turned off and the vev of the complex adjoint scalar is turned on. The coupling is implemented by gauging a

part of the flavor group by the bulk gauge field restricted to the surface. We will also give a higgsing construction of $\mathbf{Q}/\tilde{\mathbf{Q}}$-observables which is in accordance with the IIA brane picture.

2.1.1 Q-observables from coupling two-dimensional $\mathcal{N} = (2,2)$ theory

Due to the supersymmetric localization, the path integral of the $\mathcal{N} = 2$ A_1-quiver $\mathrm{U}(N)$ gauge theory localizes into a finite-dimensional equivariant integral over the moduli space of instantons. Further by the equivariant localization, any gauge theory observable produces an observable on the ensemble over the fixed points $\{\boldsymbol{\lambda}\} = \{(\lambda^{(0)}, \cdots, \lambda^{(N-1)})\}$ of the moduli space of instantons enumerated by N partitions.

On this ensemble over the fixed points $\{\boldsymbol{\lambda}\}$, the $\mathbf{Q}$-observable is expressed as (A.11),

$$\mathcal{Q}(x)[\boldsymbol{\lambda}] = \mathbb{E}\left[\frac{e^x(M^+ - S[\boldsymbol{\lambda}])^*}{P_1^*}\right], \tag{2.1}$$

where $x \in \mathbb{C}$ is a parameter of the $\mathbf{Q}$-observable, M^+ is a flavor bundle, S is the universal sheaf of the instantons, and $\mathbb{E}[\cdots]$ is the symbol that converts the equivariant Chern character into the product of equivariant weights. To avoid duplication, we refer to the companion paper [24] and appendix A for the details of the notations and conventions. We sometimes discard the Γ-functions involving the flavor bundle, writing

$$\mathbf{Q}(x) = \mathbb{E}\left[-\frac{e^x S[\boldsymbol{\lambda}]^*}{P_1^*}\right]. \tag{2.2}$$

Note that the $\mathbf{Q}$-observable (2.1) reduces to a product of Γ-functions in the decoupling limit $\mathfrak{q} \to 0$, where the universal sheaf S becomes a flavor bundle N. In this limit, the $\mathbf{Q}$-observable (2.1) can be viewed as the 1-loop fluctuation of two sets of N 2d $\mathcal{N} = (2,2)$ chiral multiplets supported on the $\mathbb{C}_1$-plane, with the twisted masses assigned as $m_\alpha^+ - x$ and $a_\alpha - x$, $\alpha = 0, 1, \cdots, N-1$. Namely, the x-parameter is introduced as an overall shift of the twisted masses. Equivalently, we may instead consider the two-dimensional $\mathcal{N} = (2,2)$ supersymmetric gauged linear sigma model, whose field contents are

- U(1) gauge multiplet

- N chiral multiplets of charge $+1$ (flavor group M^+)

- N chiral multiplets of charge -1 (flavor group N)

with the complexified FI parameter turned off and the vev of the complex adjoint scalar in the twisted chiral multiplet turned on to be $x \in \mathbb{C}$ [61]. There is no non-perturbative sector of the field configuration since the complexified FI parameter is turned off, and the partition function is precisely given by the product of Γ-functions from the 1-loop fluctuations of the chiral multiplets (see [62] for instance).

When coupled to the four-dimensional theory, the flavor symmetry group N is gauged by the four-dimensional gauge field so that the associated flavor bundle gets uplifted to the universal sheaf S for the four-dimensional instantons, yielding the surface observable (2.1). Therefore, the $\mathbf{Q}$-observable (2.1) can indeed be viewed as the surface defect engineered by
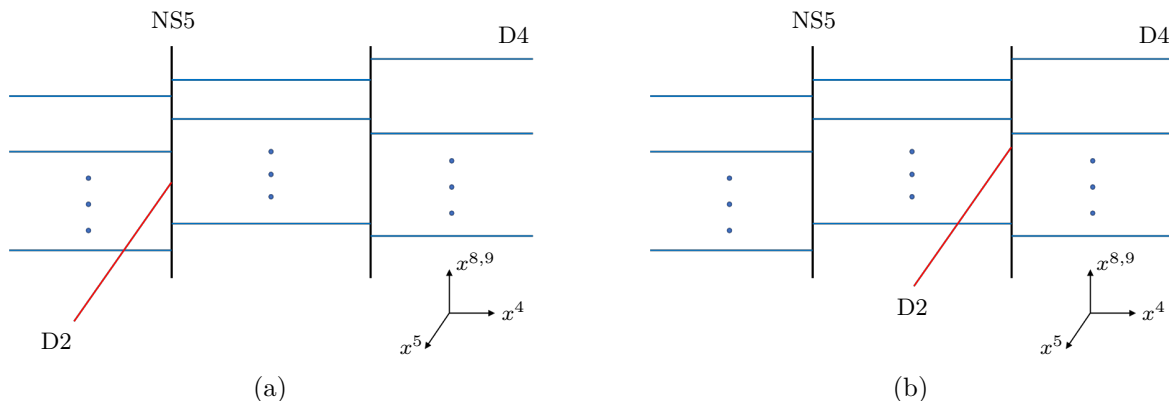
Figure 2. IIA brane engineering of the **Q**-observables, in our case $n = 2$; (a) **Q**-observable and (b) $\tilde{\mathbf{Q}}$-observable (see table 1 and 2).

coupling the two-dimensional $\mathcal{N} = (2, 2)$ sigma model supported on the $\mathbb{C}_1$-plane, where the complexified FI parameter turned off and the vev of the complex adjoint scalar turned on.

Recall that the IIB gauge origami setup (see appendix A) is dualized to IIA theory, where the D3-brane supported on $\mathbb{C}_{13}^2$ becomes a D2-brane (see figure 2, and table 1 and 2). The D2-brane ends on one of the n NS5-branes according to the $\mathbb{Z}_{n+1}$-charge that the D3-brane carried. We will restrict our focus to cases where the D2-branes end on the NS5-branes from the positive x^5-direction, as the opposite direction can be treated symmetrically. In our work, we mainly consider the case $n = 2$, so that there are **Q**- and $\tilde{\mathbf{Q}}$-observables which are inequivalent to each other.[6] We choose our convention that the D3-brane carrying $\mathbb{Z}_3$-charge 1 is dualized to the D2-brane ending on the NS5-brane on the left, while the D3-brane carrying $\mathbb{Z}_3$-charge 0 is dualized to the D2-brane ending on the NS5-brane on the right. In the point of view of the low-energy effective theory, the D2-brane engineers a half-BPS surface defect by coupling its two-dimensional worldvolume theory to the $\mathcal{N} = 2$ theory on the worldvolume of the D4-branes. The effective field theory on worldvolume of the D2-brane can easily be read off from the IIA brane picture to be the already described two-dimensional $\mathcal{N} = (2, 2)$ gauged linear sigma model, where the adjoint scalar vev $x \in \mathbb{C}$ is identified with the position of the D2-brane on the (x^8, x^9)-plane [61]. We call these surface defects the **Q**-observable and the $\tilde{\mathbf{Q}}$-observable, respectively. We recall that the number matches the order ($n = 2$) of the universal $\hbar$-oper equation for $Y(\mathfrak{gl}(2))$, which is solved by two inequivalent Q-operators [24].

When we take the limit $\varepsilon_2 \to 0$ of turning off one of the Ω-background parameters, the four-dimensional $\mathcal{N} = 2$ gauge theory is effectively described by a $\mathcal{N} = (2, 2)$ gauge theory on the topological plane $\mathbb{C}_2$. The **Q**-observable, being a surface defect supported on the $\mathbb{C}_1$-plane, becomes a local observable in this effective two-dimensional theory. Thus, its vacuum expectation value gives

$$\lim_{\varepsilon_2 \to 0} \left\langle \mathbf{Q}(x) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a}; \mathfrak{q})}{\varepsilon_2}} Q(\mathbf{a}; x), \tag{2.3}$$

---

[6]If we had $\mathbb{Z}_{n+1}$-orbifold in the IIB gauge origami, the resulting $\mathcal{N} = 2$ gauge theory is the linear A_{n-1}-quiver gauge theory whose IIA brane realization is the N D4-branes stretched between n NS5-branes and infinity. Therefore, there are n inequivalent **Q**-observables in general case.

where we abuse the notation a bit and denote the normalized vacuum expectation value of the **Q**-observable in the limit $\varepsilon_2 \to 0$ by $Q(\mathbf{a}; x)$. The crucial point here is that the effective twisted superpotential $\widetilde{\mathcal{W}}(\mathbf{a}; \mathfrak{q})$ is not affected by the insertion of the **Q**-observable, since the **Q**-observable is local in the effective two-dimensional theory.

The IIA brane picture (see figure 2) suggests that n inequivalent **Q**-observables exist as surface defects of A_{n-1}-quiver gauge theory. Therefore, in our main case of $n = 2$, we expect to have another $\tilde{\mathbf{Q}}$-observable which is inequivalent to (2.1). We present the gauge origami configuration for the **Q**-observable in the appendix A.3. The $\tilde{\mathbf{Q}}$-observable is given by

$$\tilde{\mathbf{Q}}(x) := \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \prod_{i=1}^{d} \frac{i\varepsilon_1 + \varepsilon_2}{i\varepsilon_1} \frac{Q(x)M(x + d\varepsilon_1)}{Q(x + d\varepsilon_1)Q(x + \varepsilon_2 + (d+1)\varepsilon_1)}. \tag{2.4}$$

Since the $\tilde{\mathbf{Q}}$-observable is a surface defect of the four-dimensional $\mathcal{N} = 2$ theory supported on the $\mathbb{C}_1$-plane, it becomes a local observable in the effective two-dimensional theory in the limit $\varepsilon_2 \to 0$ on the $\mathbb{C}_2$-plane. Thus, the vacuum expectation value in this limit gives

$$\left\langle \tilde{\mathbf{Q}}(x) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})}{\varepsilon_2}} \tilde{Q}(\mathbf{a}; x). \tag{2.5}$$

where we abuse the notation a bit and denote the normalized vacuum expectation value of the $\tilde{\mathbf{Q}}$-observable in the limit $\varepsilon_2 \to 0$ by $\tilde{Q}(\mathbf{a}; x)$. More explicitly, from (2.4) we have

$$\tilde{Q}(\mathbf{a}; x) = \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \frac{Q(\mathbf{a}; x)M(x + d\varepsilon_1)}{Q(\mathbf{a}; x + d\varepsilon_1)Q(\mathbf{a}; x + (d+1)\varepsilon_1)}. \tag{2.6}$$

Just like the **Q**-observable, the insertion of $\tilde{\mathbf{Q}}$-observable does not alter the effective twisted superpotential $\widetilde{\mathcal{W}}(\mathbf{a}; \mathfrak{q})$ in the limit $\varepsilon_2 \to 0$, since the $\tilde{\mathbf{Q}}$-observable is local in the effective two-dimensional theory.

2.1.2 Q-observables from higgsing

We have constructed the **Q**/$\tilde{\mathbf{Q}}$-observables as surface defects in the four-dimensional $\mathcal{N} = 2$ gauge theory created by coupling a two-dimensional $\mathcal{N} = (2, 2)$ theory. Here, we show that these surface defects can be engineered through partial Higgsing of a higher-rank gauge theory. We will first provide a qualitative description of this higgsing construction within the IIA brane setup, and then implement it at the level of exact partition functions. It should be emphasized that the IIA brane picture is ignorant of the Ω-background and should be regarded only as a qualitative description once it is turned on.

Let us consider having $N + 1$ D4-branes stretched between two NS5-branes and infinity. The low energy effective theory is the $SU(N + 1)$ gauge theory with $N + 1$ fundamental hypermultiplets and $N + 1$ anti-fundamental hypermultiplets (see figure 3). Now, consider aligning the position (on the (x^8, x^9)-plane) of one of the D4-branes across the two NS5-branes. This alignment enables the D4-brane to move along the x^5-direction (positive or negative), creating (possibly multiple) D2-branes, each of which is stretched between the D4-brane and one of the NS5-branes. Thus, after sending the D4-brane to the infinity in the x^5-direction, we are left with the original stack of N D4-branes across two NS5-branes augmented by D2-branes ending on the NS5-branes. This is precisely the IIA brane engineering of the **Q**/$\tilde{\mathbf{Q}}$-observables that we explained earlier in section 2.1.1.
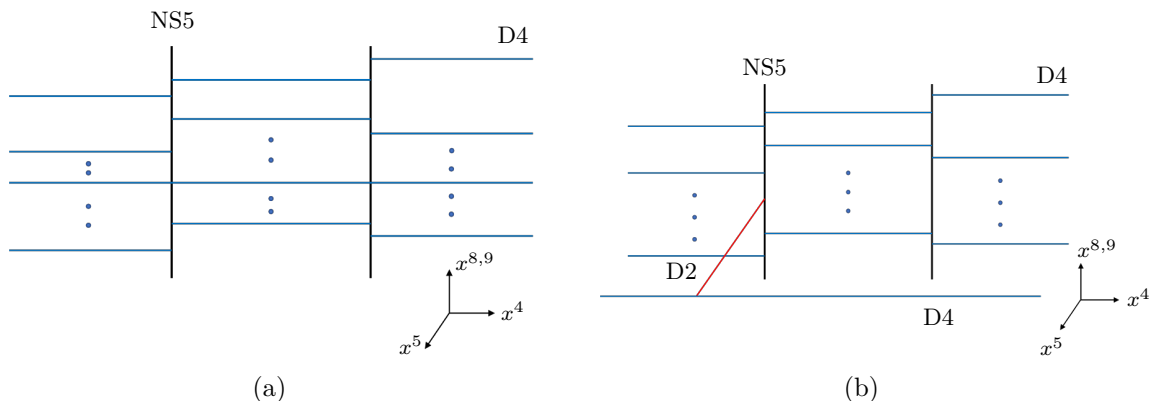
Figure 3. D2-brane creation and higgsing construction of **Q**-observable. (a) One of $N + 1$ D4-branes is aligned across the two NS5-branes, enabling the D4-brane to move along the x^5-direction. (b) A D2-brane ending on one of the NS5-branes is created as a result of the transition.

In terms of the $\mathcal{N} = 2$ gauge theory in the low-energy, the alignment corresponds to tune the masses of one of fundamental hypermultiplets and one of anti-fundamental hypermultiplets to zero. Then we can partially higgs the gauge group $\mathrm{SU}(N+1)$ down to $\mathrm{SU}(N)$ by giving vacuum expectation values for these hypermultiplets, proportional to the deviation of the D4-brane in the x^5-direction. The relevant $\mathrm{U}(1)$ part of the gauge field is squeezed into a two-dimensional plane, giving a two dimensional $\mathcal{N} = (2, 2)$ gauged linear sigma model with the vev of the adjoint scalar coupled to the remnant four-dimensional $\mathcal{N} = 2$ gauge theory with the gauge group $\mathrm{SU}(N)$. Note that the number of the created D2-branes ending on the respective NS5-branes, as well as their orientations, are unspecified within the brane picture; however, in the presence of the Ω-background the implementation at the level of the partition function will reveal that the higgsing constraints single out those numbers and orientations, lifting all the other possible configurations.

Now, in the presence of the Ω-background, let us confirm the higgsing construction by implementing it in exact computation of the partition function, both for the **Q**-observable and the $\tilde{\mathbf{Q}}$-observable. Let us begin with the partition function of the four-dimensional $\mathcal{N} = 2$ gauge theory with the gauge group $\mathrm{U}(N+1)$ and $N+1$ fundamental and $N+1$ anti-fundamental hypermultiplets. The equivariant Chern characters of the framing and flavor bundles are set to be

$$N = \sum_{\alpha=0}^{N-1} e^{a_\alpha} + e^{x+\varepsilon_1+\varepsilon_2}, \quad M^+ = \sum_{\alpha=0}^{N-1} e^{m_\alpha^+} + e^{x+\varepsilon_1+\varepsilon_2}, \quad M^- = \sum_{\alpha=0}^{N-1} e^{m_\alpha^-} + e^{x-\varepsilon_2}, \quad (2.7)$$

where we identified the $(N+1)$-th Coulomb moduli and the masses up to $\varepsilon_{1,2}$-deviation. Note that the constraint $a_N = m_N^+ = m_N^- + \varepsilon_1 + 2\varepsilon_2$ indeed makes one of the fundamentals and one of the antifundamentals massless up to an $\varepsilon_{1,2}$-deviation. It turns out that this constraint engineers a single D2-brane supported on the $\mathbb{C}_1$-plane ending on the left NS5-brane; namely, a single **Q**-observable.[7]

---

[7]More precisely, the number of the created D2-branes is determined by the integer coefficients of the Ω-background parameters in the constraint. The orientation is determined by whether it is ε_1 or ε_2 that is used to deviate from the exact massless condition.

The partition function of the U($N + 1$) gauge theory under the higgsing constraint gives

$$\mathcal{Z}_{\mathrm{U}(N+1)}(\mathbf{a}, \mathbf{m}^{\pm}, a_N - \varepsilon_1 - \varepsilon_2 = m_N^+ - \varepsilon_1 - \varepsilon_2 = m_N^- - \varepsilon_1 = x; \mathfrak{q})$$

$$= \sum_{\lambda} \mathfrak{q}^{|\lambda|} \mathbb{E}\left[\frac{-(S + e^{a_N})(S^* + e^{-a_N}) + (M^+ + e^{m_N^+})(S^* + e^{-a_N}) + q_{12}^{-1}(S + e^{a_N})((M^-)^* + e^{-m_N^-})}{P_{12}^*} \right]$$

$$= \sum_{\lambda} \mathfrak{q}^{|\lambda|} \mathbb{E}\left[\frac{-SS^* + M^+ S^* + q_{12}^{-1} S(M^-)^*}{P_{12}^*} \right] \times \mathbb{E}\left[-\frac{Se^{-x}}{P_1} \right] \mathbb{E}\left[\frac{M^+ e^{-x}}{P_{12}} + \frac{e^x(M^-)^* + q_2}{P_{12}^*} \right]$$

$$= \mathbb{E}\left[\frac{M^+ e^{-x}}{P_{12}} + \frac{e^x(M^-)^* + q_2}{P_{12}^*} \right] \times \left\langle Q(x) \right\rangle_{\mathbf{a}}.$$

(2.8)

The additional 1-loop contribution in front is from the open strings stretching between the distinguished D4-brane and the rest of N D4-branes extended to infinity. It decouples from the $\mathcal{N} = 2$ gauge theory and thus has to be discarded. Hence, we confirm that the partition function of the partially higgsed U($N + 1$) gauge theory reproduces the vacuum expectation value of the $\mathbf{Q}$-observable defined in the U(N) gauge theory, as expected.

Next, we construct the $\tilde{\mathbf{Q}}$-observable by imposing a different higgsing condition to the U($N + 1$) gauge theory. The equivariant Chern character of the framing bundle is now set as

$$N = \sum_{\alpha=0}^{N-1} e^{a_\alpha} + e^{x+\varepsilon_1}, \quad M^+ = \sum_{\alpha=0}^{N-1} e^{m_\alpha^+} + e^{x+\varepsilon_1+\varepsilon_2}, \quad M^- = \sum_{\alpha=0}^{N-1} e^{m_\alpha^-} + e^{x-\varepsilon_2}.$$ (2.9)

The higgsing condition makes the relevant instanton configuration squeezed into the $\mathbb{C}_1$-plane, constraining the corresponding universal sheaf as

$$K_N = e^{a_N} \sum_{i=1}^{d} q_1^{i-1}, \qquad S_N = e^{a_N}(q_2 + P_2 q_1^d),$$ (2.10)

where we set $|K_N| = d$. The partition function of the U($N + 1$) gauge theory subject to the higgsing constraint is computed as

$$\mathcal{Z}_{\mathrm{U}(N+1)}(\mathbf{a}, \mathbf{m}^{\pm}, a_N - \varepsilon_1 = m_N^+ - \varepsilon_1 - \varepsilon_2 = m_N^- + \varepsilon_2 = x; \mathfrak{q})$$

$$= \sum_{\lambda} \sum_{d=0}^{\infty} \mathfrak{q}^{|\lambda|+d} \mathbb{E}\left[-\frac{(S + e^{a_N}(q_2 + P_2 q_1^d))(S + e^{a_N}(q_2 + P_2 q_1^d))^*}{P_{12}^*} \right.$$

$$\left. + \frac{(M^+ + e^{m_N^+})(S + e^{a_N}(q_2 + P_2 q_1^d))^* + q_{12}^{-1}(S + e^{a_N}(q_2 + P_2 q_1^d))(M^- + e^{m_N^-})^*}{P_{12}^*} \right]$$

$$= \sum_{\lambda} \mathfrak{q}^{|\lambda|} \mathbb{E}\left[\frac{-SS^* + M^+ S^* + q_{12}^{-1} S(M^-)^*}{P_{12}^*} \right]$$

$$\times \sum_{d=0}^{\infty} \mathfrak{q}^d \prod_{i=1}^{d} \frac{i\varepsilon_1 + \varepsilon_2}{i\varepsilon_1} \mathbb{E}\left[\frac{e^x q_1^{d+1} q_2 S^*}{P_1^*} + \frac{e^{-x} S(1 - q_1^{-d})}{P_1} \right] \mathbb{E}\left[-\frac{e^{-x} q_1^{-d} M^+}{P_1} - \frac{e^x q_1^d (M^-)^*}{P_1^*} \right]$$

$$\times \mathbb{E}\left[\frac{M^+ e^{-x}}{P_{12}} + \frac{e^x(M^-)^* + q_2}{P_{12}^*} \right]$$

$$= \mathbb{E}\left[\frac{M^+ e^{-x}}{P_{12}} + \frac{e^x(M^-)^* + q_2}{P_{12}^*} \right] \times \left\langle \tilde{Q}(x) \right\rangle_{\mathbf{a}}.$$

(2.11)

By the same reason as above, the extra 1-loop part in front has to be discarded. Thus, the partition function of the partially higgsed $U(N + 1)$ gauge theory gives the vacuum expectation value of the $\tilde{\mathbf{Q}}$-observable, as suggested.

Remark. We stress that we have chosen the higgsing constraints at which the positive vevs of hypermultiplets can be turned on (namely, in figure 3, deviation of the D4-brane to the positive x^5-direction). Alternatively, we could have imposed different Higgsing constraints, resulting in the D4-brane deviating in the negative x^5-direction. Indeed, consider imposing

$$a_N = x + \varepsilon_1, \qquad\qquad m_N^+ = x - \varepsilon_2, \qquad m_N^- = x + \varepsilon_1 + \varepsilon_2, \qquad (2.12a)$$

$$a_N = m_N^- = x + \varepsilon_1 + \varepsilon_2, \qquad m_N^+ = x - \varepsilon_2. \qquad (2.12b)$$

Under the constraint, the partition function of the higher-rank gauge theory produces a single $\tilde{\mathbf{Q}}$-observable and a single $\mathbf{Q}$-observable, respectively. The former should be thought of as a D2-brane ending on the left NS5-brane from the negative x^5-direction, while the latter should be thought of as a D2-brane ending on the right NS5-brane from the negative x^5-direction. The confirmation of these higgsing constructions at the level of partition functions are straightforward as the cases we discussed in detail.

2.1.3 Q-observables across stability chambers

Even though the $\mathbf{Q}$-observable and the $\tilde{\mathbf{Q}}$-observable have similar origin in the IIA brane picture, their observable expressions look fairly distinct. We illuminate the two expressions are in fact connected to each other across different stability chambers of the four-dimensional instanton moduli space.

In the other (negative) stability chamber of the instanton moduli space, the fixed points enumerated by partitions $\boldsymbol{\lambda}$ carry different equivariant weights under the symmetry group action,

$$K = \sum_{\alpha} \sum_{(i,j)\in\lambda^{(\alpha)}} e^{a_\alpha - i\varepsilon_1 - j\varepsilon_2}. \qquad (2.13)$$

Let us proceed with the partial higgsing (2.9) of the $U(N + 1)$ gauge theory that led to the $\tilde{\mathbf{Q}}$-observable in the original chamber. The higgsing condition now enforces the relevant instanton configuration to be trivial, reducing the associated universal sheaf to framing bundle of the resultant two-dimensional theory. The partition function of the partially higgsed $U(N + 1)$ gauge theory is computed as

$$\mathcal{Z}_{U(N+1)}(\mathbf{a}, \mathbf{m}^{\pm}, a_N - \varepsilon_1 = m_N^+ - \varepsilon_1 - \varepsilon_2 = m_N^- + \varepsilon_2 = x; \mathfrak{q})^{<0}$$

$$= \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathcal{Z}_{U(N)}[\boldsymbol{\lambda}]|_{<0} \mathbb{E}\left[\frac{e^x q_{12} S^*}{P_1^*}\right] \mathbb{E}\left[\frac{M^+ e^{-x} q_2}{P_{12}} + \frac{e^x q_2^* (M^-)^*}{P_{12}^*}\right] \mathbb{E}\left[\frac{q_2}{P_{12}^*}\right] \qquad (2.14)$$

$$= \mathbb{E}\left[\frac{M^+ e^{-x} q_2}{P_{12}} + \frac{e^x q_2^* (M^-)^*}{P_{12}^*}\right] \mathbb{E}\left[\frac{q_2}{P_{12}^*}\right] \times \left\langle Q(x + \varepsilon_1 + \varepsilon_2) \right\rangle_{\mathbf{a}}^{<0},$$

where the notation < 0 indicates the other stability chamber. We discard the extra 1-loop contribution in front that decouples from the $\mathcal{N} = 2$ gauge theory. Thus, we confirm that the $\tilde{\mathbf{Q}}$-observable in the original chamber passes to the $\mathbf{Q}$-observable in the other chamber.

Similarly, we can start from the partial higgsing (2.7) which gave the **Q**-observable in the original chamber. The instanton configuration for the $(N + 1)$-th U(1) gauge field is no longer trivial, but is squeezed into the two-dimensional plane $\mathbb{C}_1$. The associated universal sheaf is constrained as

$$K_N = \sum_{i=1}^{k} e^{a_N - i\varepsilon_1 - \varepsilon_2} = e^x \frac{1 - q_1^{-k}}{1 - q_1^{-1}}, \qquad S_N = e^{a_N}(q_2^{-1} - P_2^* q_1^{-k}). \tag{2.15}$$

As a result, the partition function of the partially higgsed U(N+1) gauge theory is computed as

$$
\begin{aligned}
&\mathcal{Z}_{\text{U}(N+1)}(\mathbf{a}, \mathbf{m}^{\pm}, a_N - \varepsilon_1 - \varepsilon_2 = m_N^+ - \varepsilon_1 - \varepsilon_2 = m_N^- - \varepsilon_1 = x; \mathfrak{q})^{<0} \\
&= \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathcal{Z}_{\text{U}(N)}[\boldsymbol{\lambda}]|_{<0} \sum_{k=0}^{\infty} \mathfrak{q}^k \prod_{j=1}^{k} \frac{Q(x + \varepsilon_1 + \varepsilon_2) M^-(x - k\varepsilon_1)^{-1} M^+(x - k\varepsilon_1)^{-1}}{Q(x - k\varepsilon_1) Q(x + \varepsilon_1 + \varepsilon_2 - k\varepsilon_1)} \prod_{j=1}^{k} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_1} \\
&\quad \times \mathbb{E}\left[\frac{M^+ e^{-a_{N+1}} q_2}{P_{12}^*} + \frac{e^{a_{N+1}} q_2^* (M^-)^*}{P_{12}^*} \right] \mathbb{E}\left[\frac{q_2}{P_{12}^*} \right] \\
&= \mathbb{E}\left[\frac{M^+ e^{-x} q_2}{P_{12}} + \frac{e^x q_2^* (M^-)^*}{P_{12}^*} \right] \mathbb{E}\left[\frac{q_2}{P_{12}^*} \right] \times \mathfrak{q}^{\frac{x}{\varepsilon_1}} \left\langle \tilde{Q}(x + \varepsilon_1 + \varepsilon_2) \right\rangle_{\mathbf{a}}^{<0},
\end{aligned}
\tag{2.16}
$$

where the extra 1-loop part in front has to be discarded by the same reason above. Thus, we also verify that the **Q**-observable in the original chamber translates into the **Q̃**-observable in the other chamber.
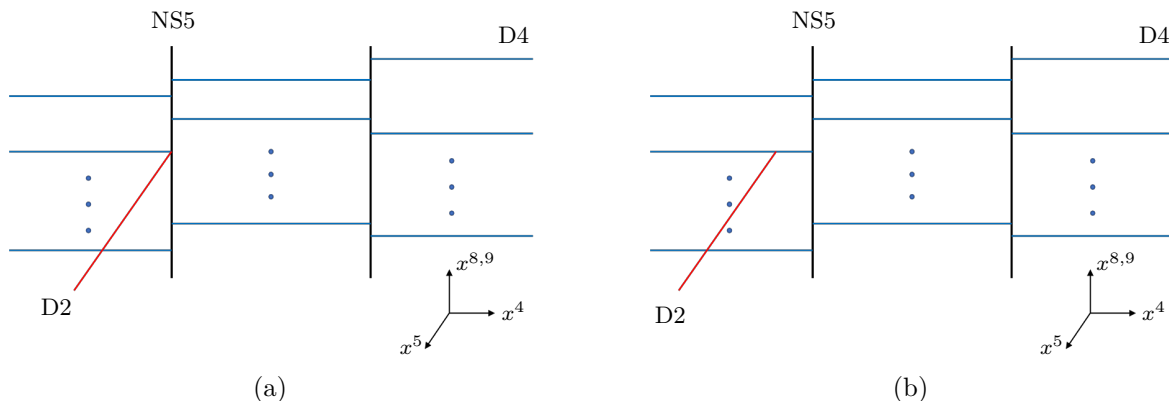
2.2 H-observables

We have seen that the **Q**/**Q̃**-observables are the surface defects created by coupling a two-dimensional $\mathcal{N} = (2, 2)$ theory with the vev of the complex adjoint scalar. Here, we consider the surface defect constructed by coupling the same gauged linear sigma model with the complexified FI parameter $\log y = r - i\theta$ turned on. It is called the canonical surface defect in [8, 14, 63], which we use interchangeably with the **H**-observable in this work.

The real FI parameter r can be either positive or negative (namely, $|y| > 1$ or $|y| < 1$). In either case, the gauged linear sigma model flows to the non-linear sigma model with the target being the total space of

$$\mathcal{O}(-1) \otimes \mathbb{C}^N \longrightarrow \mathbb{P}^{N-1}, \tag{2.17}$$

where the complexified FI parameter becomes the complexified Kähler parameter. It has N discrete vacua. When coupled to the four-dimensional $\mathcal{N} = 2$ theory, the U(N) flavor symmetry of rotating the fiber gets gauged by the bulk gauge field restricted to the surface.

The transition can be intuitively understood in the IIA brane picture, at least in the absence of the Ω-background (see figure 4). The D2-brane ending on one of the NS5-branes can be brought to the position (on the (x^8, x^9)-plane) of one of the $2N$ D4-branes ending on the same NS5-brane. There are N D4-branes on the left and N D4-branes on the right. Depending on which D4-brane it meets, the D2-brane can be moved toward left or right, deviating the real FI parameter from the locus $y = 1$. In either case, there are N choices

Figure 4. Transition between **Q**-observable and **H**-observable in IIA brane picture. (a) D2-brane meets one of the D4-branes on the left (resp. right) by tuning vev of the complex adjoint scalar (b) D2-brane moves toward left (resp. right) along the D4-brane, turning on positive (resp. negative) real FI parameter.

of D4-brane that the D2-brane can end on, which correspond to the choice of vacuum at infinity among N discrete vacua of the non-linear sigma model. Away from the locus $y = 1$, the two non-linear sigma models are smoothly connected to each other, undergoing a flop transition of the target space across the zero Kähler parameter.

Implementing the Ω-background, the **H**-observable gives rise to an observable defined on the ensemble of fixed points of the instanton moduli space enumerated by the partitions $\{\boldsymbol{\lambda}\}$. The transition between the **Q**- and **H**-observables manifests at the level of their observable expressions, in the contour integral presentation of the latter [14] which generalizes the *Coulomb branch presentation* [62] of the vortex partition function of two-dimensional $\mathcal{N} = (2, 2)$ gauged linear sigma model. When the real FI parameter is positive ($|y| > 1$), the contour integral presentation is given by

$$\mathbf{H}_{\text{left}}^{(\alpha)}(y)[\boldsymbol{\lambda}] = \oint_{\mathcal{C}_\alpha} dx \, y^{-\frac{x}{\varepsilon_1}} \mathcal{Q}(x)[\boldsymbol{\lambda}], \qquad \begin{array}{l} \alpha = 0, 1, \cdots, N-1, \\ |y| > 1 \end{array} \tag{2.18}$$

where the integrand is the 1-loop fluctuation of the chiral multiplets at the vev of the complex adjoint scalar $x \in \mathbb{C}$, which is essentially the **Q**-observable (with the Γ-functions from the 1-loop part of the **Q**-observable). The integral kernel $y^{-\frac{x}{\varepsilon_1}} = e^{-\frac{x \log y}{\varepsilon_1}}$ is the classical contribution, with the complexified FI parameter $\log y$. There are N semi-infinite ε_1-lattices of simple poles due to the Γ-functions, and the contour $\mathcal{C}_\alpha$ is chosen to enclose only the poles of single lattice, $x = m_\alpha^+ + k\varepsilon_1$, $k \geq 0$. The resulting expressions are series in y^{-1}, converging in the domain $0 < |\mathfrak{q}| < 1 < |y|$ of $y \in \mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$. The N choices for the contour exactly correspond to the choice of a vacuum at infinity among the N discrete vacua of the non-linear sigma model.

In the limit $\varepsilon_2 \to 0$, the **H**-observable gives a local observable in the effective two-dimensional theory on the $\mathbb{C}_2$-plane. Accordingly, the vacuum expectation value in the limit becomes

$$\left\langle \mathbf{H}^{(\alpha)}(y) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})}{\varepsilon_2}} \chi_\alpha(\mathbf{a}; y), \qquad \alpha = 0, 1, \cdots, N-1, \tag{2.19}$$

where we denoted the normalized vacuum expectation value of the **H**-observable in the limit $\varepsilon_2 \to 0$ by $\chi_\alpha(\mathbf{a}; y)$. Just as the **H**-observables themselves, the vev $\chi_\alpha(\mathbf{a}; y)$, which is given as a series which is a convergent in a specific domain in $\mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$, analytically continues to the other domains of $\mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$ [14].

2.2.1 H-observables across domains of convergence

Due to the coupling to the four-dimensional theory, the parameter space (y-space) for the **H**-observable extends to the Riemann surface associated to the class $\mathcal{S}$ theory [64]. In the present case of the A_1-quiver theory, it was explicitly shown in [14] that the vacuum expectation value of the **H**-observables (2.18) analytically continues across different domains of convergence in the y-space $\mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$. We will not repeat the derivation of the analytic continuation here, but only gather some necessary facts to present the complete map of transitions between the **Q**-observables and the **H**-observables.

As noted earlier, the non-linear sigma model experiences a flop transition of the target across (but away from) $y = 1$. At the level of the **H**-observable, the flop transition manifests as the analytic continuation from one domain (say, $0 < |\mathfrak{q}| < 1 < |y|$) to the other ($0 < |\mathfrak{q}| < |y| < 1$). Once analytically continued, each of N **H**-observables (2.18) can be re-expressed as a linear combination of

$$\mathbf{H}^{(\alpha)}_{\text{middle}}(y) = \sum_{x \in L_\alpha} y^{-\frac{x}{\varepsilon_1}} \mathcal{Q}(x), \qquad \begin{array}{l} \alpha = 0, 1, \cdots, N-1, \\ 0 < |\mathfrak{q}| < |y| < 1 \end{array} \tag{2.20}$$

where $L_\alpha = a_\alpha + \varepsilon_1 \mathbb{Z}$ is the ε_1-lattice centered at the Coulomb parameter a_α. Due to the zeros of the inverse Γ-functions in $\mathcal{Q}(x)$, the non-zero contribution to the summation truncates from above and $\frac{x}{\varepsilon_1}$ grows only to the negative infinity. It follows that the series converges in the domain $0 < |\mathfrak{q}| < |y| < 1$ [14]. See [14] for the exact connection formulas across the two domains.

It is apparent from the IIA brane picture (figure 4) that a D2-brane ending on a D4-brane in the middle can originate from a D2-brane ending on either the left NS5-brane or the right NS5-brane. In the gauge theory perspective, this implies the **H**-observables in the domain $0 < |\mathfrak{q}| < |y| < 1$ can also be obtained by a transition from the $\tilde{\mathbf{Q}}$-observable. In the field theory terms, it suggests that there is an alternative weakly coupled description of the 2d/4d system where the 2d flavor group is identified with 4d flavor group in a different way (relevant to the masses $(m_\alpha^-)_{\alpha=0}^{N-1}$ instead of $(m_\alpha^+)_{\alpha=0}^{N-1}$) and the complexified Kähler parameter is given by $\frac{\mathfrak{q}}{y}$ instead of y. The corresponding **H**-observables are obtained as

$$\tilde{\mathbf{H}}^{(\alpha)}_{\text{middle}}(y) = \oint_{\tilde{\mathcal{C}}_\alpha} dx \, y^{-\frac{x}{\varepsilon_1}} \tilde{\mathcal{Q}}(x), \qquad \begin{array}{l} \alpha = 0, 1, \cdots, N-1, \\ 0 < |\mathfrak{q}| < |y| < 1 \end{array}, \tag{2.21}$$

where

$$\tilde{\mathcal{Q}}(x) = \frac{\tilde{\mathbf{Q}}(x)}{M^+(x)} = \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \left(\prod_{j=1}^{d} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_1} \frac{P^+(x + j\varepsilon_1)}{\mathcal{Y}(x + j\varepsilon_1)} \right) \frac{M^-(x + d\varepsilon_1)}{Q(x + \varepsilon_2 + (d+1)\varepsilon_1)}, \tag{2.22}$$
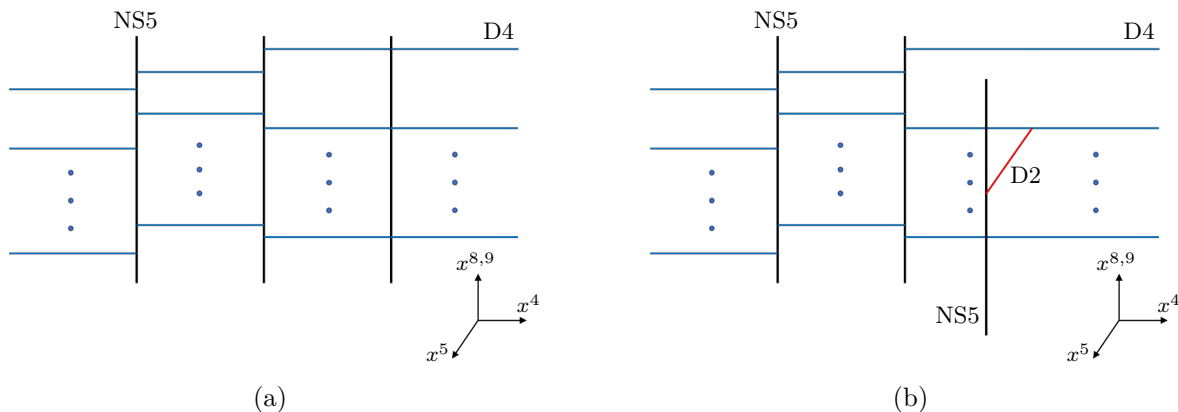
Figure 5. D2-brane creation and higgsing construction of the **H**-observable. (a) All N D4-branes are aligned across one of the NS5-branes, enabling the NS5-brane to move along the x^5-direction. (b) A D2-brane ending on one of the D4-branes is created as a result of the higgsing.

and the contour $\tilde{\mathcal{C}}_\alpha$ encloses the simple poles at $x = a_\alpha + (l(\lambda^{(\alpha)}) + l - d - 1)\varepsilon_1 - \varepsilon_2$, $l \geq 0$. The vevs of the two **H**-observables, (2.20) and (2.21), in the same domain $0 < |\mathfrak{q}| < |y| < 1$ are related to each other by a ε_2-shift of a Coulomb parameter [14].

Finally, the **H**-observables (2.21) can be analytically continued to the last domain $0 < |y| < |\mathfrak{q}| < 1$ to be linear combinations of

$$\tilde{\mathbf{H}}^{(\alpha)}_{\text{right}}(y) = \sum_{x \in \tilde{L}_\alpha} y^{-\frac{x}{\varepsilon_1}} \tilde{\mathcal{Q}}(x), \qquad \begin{matrix} \alpha = 0, 1, \cdots, N-1, \\ 0 < |y| < |\mathfrak{q}| < 1 \end{matrix}, \qquad (2.23)$$

where $\tilde{L}_\alpha = m_\alpha^- + \varepsilon_1\mathbb{Z}$ is the ε_1-lattice centered at the mass parameter m_α^-. Due to the inverse Γ-functions in $\tilde{\mathcal{Q}}(x)$, the non-zero contribution to the summation truncates from above, guaranteeing the convergence of the series in the domain $0 < |y| < |\mathfrak{q}| < 1$. See [14] for the exact connection formulas across the two domains.

2.2.2 H-observables from higgsing

The **H**-observable built by coupling the non-linear sigma model with the target (2.17) can also be realized by a partial higgsing [14, 53, 65, 66]. In this construction, N hypermultiplets are brought to be nearly massless, admitting a non-zero vacuum expectation value. After the vev is turned on, the BPS field configuration is squeezed into the $\mathbb{C}_1$-plane, forming a vortex string coupled to the 4d theory. See [14, 53], for instance, for the implementation of the partial higgsing construction of the **H**-observable at the level of the partition function.

The higgsing construction of the **H**-observables can easily visualized in the IIA brane setup (see figure 5) [65, 66]. We begin with a higher-rank quiver gauge theory with one more gauge node, which is realized by the IIA branes with an additional NS5-brane. A D2-brane creation can be initiated by aligning all the D4-branes across one of the NS5-branes, by pulling out the relevant NS5-brane along the x^5-direction. The worldvolume theory on the D2-brane is precisely the non-linear sigma model with the target (2.17), which couples to the remaining 4d theory by gauging its flavor symmetry.

We refer to [14, 53, 65, 66] more details on the higgsing construction of the **H**-observable surface defect. Here, we only remind that the higgsing construction of the **H**-observables exactly parallels the higgsing construction of the **Q**-observables in section 2.1.2. Here, *all* the D4-branes are aligned across *one* of the NS5-brane, enabling the NS5-brane to move along the x^5-direction. A D2-brane ending on one of the D4-branes is created, engineering a **H**-observable in the remnant $\mathcal{N} = 2$ gauge theory. In higgsing construction of the **Q**-observables, instead, only *one* of the D4-branes is aligned across *all* the NS5-branes, enabling the D4-brane to move along the x^5-direction. The newly created D2-brane ends on one of the NS5-branes, giving a **Q**-observable.

2.3 Monodromy surface defect

The monodromy surface defect in the four-dimensional $\mathcal{N} = 2$ gauge theory is engineered by assigning singular behavior of the gauge field along a surface [67]. The singular behavior of the gauge field can be modelled by placing the $\mathcal{N} = 2$ gauge theory on an orbifold [56], $\mathbb{C}_1 \times (\mathbb{C}_2/\mathbb{Z}_l)$. As we review in appendix A.4, the construction can be embedded into the IIB origami setup by replacing a part of the worldvolume to the orbifold.

2.3.1 Construction by orbifold

To model the singularity of the gauge field, we consider the $\mathcal{N} = 2$ gauge theory supported on the orbifold $\mathbb{C}_1 \times (\mathbb{C}_2/\mathbb{Z}_l)$, where the $\mathbb{Z}_l$-action is given by [56],

$$(z_1, z_2) \mapsto (z_1, \zeta z_2), \qquad \zeta = \exp\frac{2\pi i}{l}. \tag{2.24}$$

The map $z_2 \mapsto z_2^l$ sends the $\mathcal{N} = 2$ gauge theory supported on the $\mathbb{Z}_l$-orbifold to the one supported on the ordinary $\mathbb{C}_1 \times \mathbb{C}_2$, compensating the presence of the orbifold by singularity of fields along the locus $\{z_2 = 0\}$. Replacing the support of the worldvolume of the $\mathcal{N} = 2$ gauge theory to the $\mathbb{Z}_l$-orbifold engineers a monodromy surface defect in this sense.

The monodromy surface defects are classified by the Levi subgroups of the gauge group and the flavor group it preserves [67], which we encode in *coloring functions*. In our case of the A_1-quiver U(N) gauge theory, the coloring functions (c, c_f, c_{af}) assign $\mathbb{Z}_l$-charges to the framing and the flavor bundles by

$$\begin{aligned} c : N &\longrightarrow \{0, 1, \cdots, l-1\} \\ c_f : M^+ &\longrightarrow \{0, 1, \cdots, l-1\} \\ c_{af} : M^- &\longrightarrow \{0, 1, \cdots, l-1\}. \end{aligned} \tag{2.25}$$

The preserved Levi subgroups are determined to be $S\left(\times_{\omega=0}^{l-1} \mathrm{U}(|c^{-1}(\omega)|)\right)$, etc.

In the special case where $l = N$ and the coloring functions (c, c_f, c_{af}) are one-to-one functions, we may set $c(\alpha) = c_f(\alpha) = c_{af}(\alpha) = \alpha)$ without loss of generality. We call the corresponding monodromy surface defect to be regular. In the present work, we will mainly restrict to the regular monodromy surface defect, even though we will give a dual description of monodromy surface defect defined in the most general case.

2.3.2 Monodromy surface observable

To avoid duplication with the companion paper [24], we do not present the derivation of the observable expression for the monodromy surface defect here. See [8, 24] and appendix A.4 for the detail of the definitions and the conventions. The observable expression of the regular monodromy surface defect labelled by the coloring function **c** is given by (cf. (A.25))

$$\Psi_{\mathbf{c}}(\mathbf{u})[\boldsymbol{\lambda}] = \sum_{\hat{\boldsymbol{\lambda}} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{k_\omega - k_{l-1}} \mathbb{E}\left[\frac{S\Sigma_0^*}{P_1^*} + \sum_{\omega=0}^{l-2} \frac{-\Sigma_\omega(\Sigma_\omega - \Sigma_{\omega+1})^* - M_\omega^+ \Sigma_\omega^* + q_1^{-1}(M_\omega^-)^*\Sigma_\omega}{P_1^*}\right].$$
(2.26)

There are $l-1$ defect parameters $(\mathfrak{q}_\omega)_{\omega=0}^{l-2}$ encoding the singularity of the gauge field and the magnetic flux along the support $\mathbb{C}_1$.

We will re-parameterize these defect parameters using $\mathbf{u} = (u_\omega)_{\omega=0}^{l-1}$,

$$\mathfrak{q}_\omega = \frac{u_{\omega+1}}{u_\omega}, \qquad \omega = 0, 1, \cdots, l-2,$$
(2.27)

with a redundancy of overall scaling. Also, we will incorporate the classical contribution to the observable (2.26), given by

$$\Psi_{\mathbf{c}}(\mathbf{u})^{\text{classical}} = \prod_{\omega=0}^{l-1} u_\omega^{\sum_{\alpha \in c^{-1}(\omega)} \frac{m_\alpha^+ - a_\alpha}{\varepsilon_1}}.$$
(2.28)

2.3.3 Space of regular monodromy surface defects

For our discussion on the ($\hbar$-)Langlands correspondence and the associated bispectral duality, we will restrict to the case of the regular monodromy surface defect.

It is immediate that the vacuum expectation value of the monodromy surface defect in the space of degree-zero Laurent polynomials in the monodromy defect parameters $\mathbf{u} = (u_\omega)_{\omega=0}^{N-1}$; namely,

$$\left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} \in \mathcal{H}^{\mathbb{C}^\times}, \qquad \mathcal{H} = \bigotimes_{\omega=0}^{N-1} \mathcal{H}_\omega,$$
(2.29)

where $\mathcal{H}_\omega = u_\omega^{\frac{m_\omega^+ - a_\omega}{\varepsilon_1}} \mathbb{C}(u_\omega)$ is the space of Laurent polynomials in u_ω. Here, the degree-zero condition $\mathcal{H}^{\mathbb{C}^\times}$ really means being degree-zero except the factor in front given by (2.28); namely, $\mathcal{H}^{\mathbb{C}^\times} = \prod_{\omega=0}^{N-1} u_\omega^{\frac{m_\omega^+ - a_\omega}{\varepsilon_1}} \mathbb{C}(u_0, u_1, \cdots, u_{l-1})^{\mathbb{C}^\times}$. In the companion paper [24], we viewed $\mathcal{H}^{\mathbb{C}^\times}$ as the degree-zero subspace of the evaluation module over the Yangian $Y(\mathfrak{gl}(2))$, defined upon N bi-infinite $\mathfrak{gl}(2)$-modules.[8] We refer to the companion paper [24] for the precise definition of the $\mathfrak{gl}(2)$-module structure on $\mathcal{H}_\omega$, and the mapping between the module parameters and the $\mathcal{N} = 2$ gauge theory parameters. In particular, the associated $\mathfrak{gl}(2)$ XXX spin chain was formulated on this bi-infinite Yangian module.

Meanwhile, let us briefly recap from [8, 43, 68] that the same vev of the regular monodromy surface defect is valued in the space of E-twisted coinvariants of four $\widehat{\mathfrak{sl}}(N)$-modules, where

---

[8]These modules are said to be of bi-infinite type, since they are not of highest-weight (nor lowest-weight) so that the weights grow toward both infinities. See [8] for precise definitions.

$E \in \mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$. Here, the four $\widehat{\mathfrak{sl}}(N)$-modules are induced from two (co-)Verma $\mathfrak{sl}(N)$-modules (V_ζ and $\tilde{V}_{\tilde{\zeta}}$) and two bi-infinite $\mathfrak{sl}(N)$-modules ($H_{\boldsymbol{\tau}-\zeta,\sigma}$ and $H_{\boldsymbol{\tau}-\tilde{\zeta},\tilde{\sigma}}$). The space of twisted $\widehat{\mathfrak{sl}}(N)$-coinvariants is isomorphic to the space of $\mathfrak{sl}(N)$-invariants [69], so that

$$\left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} \in M^{\mathfrak{sl}(N)}, \qquad M = V_\zeta \otimes H_{\boldsymbol{\tau}-\zeta,\sigma} \otimes H_{\boldsymbol{\tau}-\tilde{\zeta},\tilde{\sigma}} \otimes \tilde{V}_{\tilde{\zeta}}, \tag{2.30}$$

where the monodromy parameters $\mathbf{u}$ give holomorphic coordinates on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$. We refer to our previous work [8] for the precise definitions of the (co-)Verma $\mathfrak{sl}(N)$-modules (V_ζ and $\tilde{V}_{\tilde{\zeta}}$) and the bi-infinite $\mathfrak{sl}(N)$-modules ($H_{\boldsymbol{\tau}-\zeta,\sigma}$ and $H_{\boldsymbol{\tau}-\tilde{\zeta},\tilde{\sigma}}$), and the mapping between the module parameters and the $\mathcal{N}=2$ gauge theory parameters. In particular, the associated $\mathfrak{sl}(N)$ Gaudin model was defined on these (bi-)infinite $\mathfrak{sl}(N)$-modules.

At this point, let us split the Coulomb moduli $(a_\alpha)_{\alpha=0}^{N-1}$ into two parts: the reference values $\frac{a_\alpha}{\varepsilon_1} \bmod \mathbb{Z}$ and the ε_1-integral shifts. We regard the reference values $\frac{a_\alpha}{\varepsilon_1} \bmod \mathbb{Z}$ to be fixed, so that the vacuum expectation value (2.29) and (2.30) always gives an element of the same spaces $\mathcal{H}^{\mathbb{C}^\times}$ and $M^{\mathfrak{sl}(N)}$. Namely,

$$\left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}+\varepsilon_1\mathbf{n}} \in \mathcal{H}^{\mathbb{C}^\times}, \qquad \left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}+\varepsilon_1\mathbf{n}} \in M^{\mathfrak{sl}(N)}. \tag{2.31}$$

There are $\mathbb{Z}^{N-1}$ amount of remaining degrees of freedom, $\mathbf{n}$, for the ε_1-shifts of N Coulomb moduli up to the vanishing trace constraint $\sum_{\alpha=0}^{N-1} a_\alpha = 0$. The vacuum expectation value of the monodromy surface defect provides a basis of the spaces $\mathcal{H}^{\mathbb{C}^\times}$ and $M^{\mathfrak{sl}(N)}$ enumerated by the ε_1-integral shifts in the Coulomb moduli $\mathbf{a}$.

In the limit $\varepsilon_2 \to 0$ where the four-dimensional theory is described by the effective two-dimensional theory on the $\mathbb{C}_2$-plane, the monodromy surface defect becomes a local defect since it is supported on the $\mathbb{C}_1$-plane. Its vacuum expectation value becomes

$$\lim_{\varepsilon_2 \to 0} \left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})}{\varepsilon_2}} \psi(\mathbf{a};\mathbf{u};\mathfrak{q}), \tag{2.32}$$

where we denote the normalized vacuum expectation value of the monodromy surface defect in the limit $\varepsilon_2 \to 0$ as $\psi(\mathbf{a};\mathbf{u};\mathfrak{q})$. Note the effective twisted superpotential $\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})$ does not depend on the defect parameters $\mathbf{u}$, and the normalized vacuum expectation value $\psi(\mathbf{a};\mathbf{u};\mathfrak{q})$, with ε_1-integral shifts in the Coulomb moduli $\mathbf{a}$, still provides a distinguished basis of the spaces $\mathcal{H}^{\mathbb{C}^\times}$ and $M^{\mathfrak{sl}(N)}$.

3 Bispectral duality from Q/H surface defect transition

In this section, we explain that the bispectral duality between the associated XXX spin chain and the Gaudin model follows from the transition between the **Q**-observables and the **H**-observables. First, we recap how **Q**- and **H**-observables give rise to the Q-operators and the Hecke operators. The transition between the two observables immediately yields the Fourier transform between the universal $\hbar$-oper equation and the universal oper equation satisfied by the respective operators. The one-to-one mapping between the spectral equations of the $\mathfrak{gl}(2)$ XXX spin chain with N sites and the $\mathfrak{sl}(N)$ Gaudin model with 4 sites is obtained as a direct consequence.

3.1 Surface defects as operators

We consider the configuration where the **Q**-observable or the **H**-observable is inserted on the same $\mathbb{C}_1$-plane on top of the regular monodromy surface defect. In the gauge origami setup, such a configuration descends from adding a stack of N D3-branes to the stack of D3-branes engineering the $\mathcal{N} = 2$ gauge theory with the monodromy surface defect, each of which carries a $\mathbb{Z}_N$-charge $\omega = 0, 1, \cdots, N-1$ (see [8, 24]).

3.1.1 $\mathcal{N} = 2$ gauge theory formulation of ($\hbar$-)Langlands correspondence

Let us first briefly recap the role of such parallel surface defect configurations in the $\mathcal{N} = 2$ gauge theoretical account for the ordinary [8] and the $\hbar$-deformed Langlands correspondence [24], at our main example of the A_1-quiver U(N) theory.

We studied the geometric Langlands correspondence on the Riemann sphere $\mathcal{C} = \mathbb{P}^1$ with ramifications at four marked points $S = \{0, \mathfrak{q}, 1, \infty\} \subset \mathbb{P}^1$.[9] On the automorphic side, we had coinvariants of four $\widehat{\mathfrak{sl}}(N)$-modules twisted by parabolic PGL(N)-bundles $E \in \text{Bun}_{\text{PGL}(N)}(\mathbb{P}^1; S)$, which form sections of a twisted $\mathcal{D}$-module on $\text{Bun}_{\text{PGL}(N)}(\mathbb{P}^1; S)$ by the Ward identities. The relevant $\widehat{\mathfrak{sl}}(N)$-modules were the ones induced from $\mathfrak{sl}(N)$-modules by evaluation homomorphism, two of which are (co-)Verma modules (attached to 0 and ∞) and the other two were of bi-infinite type (attached to $\mathfrak{q}$ and 1). We constructed a distinguished basis $\psi(\mathbf{a}; \mathbf{u})$, enumerated by the Coulomb moduli $\mathbf{a}$, of these sections as the normalized vacuum expectation value of the regular monodromy surface defect in the limit $\varepsilon_2 \to 0$. In particular, the monodromy defect parameters $\mathbf{u}$ were identified with holomorphic coordinates on an open patch in $\text{Bun}_{\text{PGL}(N)}(\mathbb{P}^1; S)$.

On the Galois side, we considered the oper submanifold in the moduli space of parabolic SL(N) local systems on $\mathbb{P}^1$ with four regular singularities at $S = \{0, \mathfrak{q}, 1, \infty\}$, $\text{Op}_{\text{SL}(N)}(\mathbb{P}^1; S) \subset \mathcal{M}_{\text{loc}}(\text{SL}(N), \mathbb{P}^1; S)$. We realized an oper $\rho_{\mathbf{a}} \in \text{Op}_{\text{SL}(N)}(\mathbb{P}^1; S)$, enumerated by the Coulomb moduli $\mathbf{a}$, as a N-th order differential equation of the quantized chiral ring for the $\mathcal{N} = 2$ gauge theory coupled to the surface defect theory for the **H**-observable. In particular, N independent solutions to the oper differential equation were obtained as the normalized vacuum expectation values $\chi_\alpha(\mathbf{a}; y)$ of the **H**-observable, where $y \in \mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$ and $\alpha = 0, 1, \cdots, N-1$ labels the choice of the vacuum at infinity.

To establish the geometric Langlands correspondence in the $\mathcal{N} = 2$ theory context, we showed that the action of inserting a **H**-observable on top of the regular monodromy surface defect defines a Hecke operator on the twisted $\mathcal{D}$-module. The action was proven to be diagonal on the distinguished basis element $\psi(\mathbf{a}; \mathbf{u})$ due to cluster decomposition of the two surface defects in the limit $\varepsilon_2 \to 0$. Thus, the normalized vacuum expectation value $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy defect was shown to be a section of the Hecke eigensheaf corresponding to the oper $\rho_{\mathbf{a}}$, with the corresponding *eigenvalue* being precisely the oper solution $\chi_\alpha(\mathbf{a}; y)$, the normalized vacuum expectation value of the **H**-observable.

The $\mathcal{N} = 2$ gauge theoretical account of the $\hbar$-Langlands correspondence that we have established in the companion paper [24] completely parallels the prior study [8]. In particular, we showed that the action of inserting a **Q**-observable on top of the same regular monodromy

---

[9]The marked points at 0 and ∞ were taken to be *maximal* while the ones at $\mathfrak{q}$ and 1 were *minimal*. For their precise definitions, see [8].

surface defect defines a Q-operator on the module over the Yangian. The action is shown to be diagonal on the distinguished basis element $\psi(\mathbf{a}; \mathbf{u})$ by the cluster decomposition of the two surface defects in the limit $\varepsilon_2 \to 0$. Accordingly, the normalized vev $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy surface defect is shown to be a Q-eigenstate corresponding to a $\hbar$-oper, with the eigenvalue being the Q-function, i.e., the normalized vev of the $\mathbf{Q}$-observable.

3.1.2 Q-observables as Q-operators

In the companion paper [24], we viewed the $\mathbf{Q}$-observables as Q-operators acting on the space $\mathcal{H}^{\mathbb{C}^\times}$ as follows. Even with the additional insertion of the $\mathbf{Q}$-observable, the correlation function of the $\mathbf{Q}$-observable and the monodromy surface defect is valued in the space of degree-zero Laurent polynomials in the monodromy defect parameters $\mathbf{u} = (u_\omega)_{\omega=0}^{N-1}$,

$$\mathbf{Q}(x) \cdot \left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} := \left\langle \mathbf{Q}(x)\Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} \in \mathcal{H}^{\mathbb{C}^\times} = \left(\bigotimes_{\omega=0}^{N-1} \mathcal{H}_\omega \right)^{\mathbb{C}^\times}, \quad \text{also for } \tilde{\mathbf{Q}}(x), \quad (3.1)$$

just as the vacuum expectation value of the regular monodromy surface defect itself. In other words, we can view the $\mathbf{Q}$-observable also as an operator

$$\mathbf{Q}(x), \tilde{\mathbf{Q}}(x) \in \mathrm{End}\left(\mathcal{H}^{\mathbb{C}^\times}\right), \quad (3.2)$$

whose action on the basis element $\langle \Psi(\mathbf{u}) \rangle_{\mathbf{a}}$ is given by the above correlation function. Under this context, we refer to $\mathbf{Q}(x)$ and $\tilde{\mathbf{Q}}(x)$ as the *Q-operators*. Since the contribution at each partition $\boldsymbol{\lambda}$ is modified by the insertion of the $\mathbf{Q}$-observable, the action of Q-operator is highly non-trivial.

One of the main results of the companion paper [24] is that the so-defined Q-operators satisfy the universal $\hbar$-oper difference equation in the limit $\varepsilon_2 \to 0$,

$$\begin{aligned} 0 &= \left[1 - \hat{t}(x)e^{-\varepsilon_1 \partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1 \partial_x} \right] \mathbf{Q}(x) \\ 0 &= \left[1 - \hat{t}(x)e^{-\varepsilon_1 \partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1 \partial_x} \right] \tilde{\mathbf{Q}}(x) \end{aligned}, \quad (3.3)$$

where $\hat{t}(x) \in \mathrm{End}(\mathcal{H}^{\mathbb{C}^\times})$ is the transfer matrix of the Yangian $Y(\mathfrak{gl}(2))$, i.e., the trace of the monodromy matrix whose coefficients generate the Bethe subalgebra of $Y(\mathfrak{gl}(2))$, represented on $\mathcal{H}^{\mathbb{C}^\times}$. The universal $\hbar$-oper equation is nothing but the Baxter's TQ equation for the Q-operators and the transfer matrix for the $\mathfrak{gl}(2)$ XXX spin chain. This explains our terminology of referring to $\mathbf{Q}$-observables as the Q-operators.

3.1.3 H-observables as Hecke operators

Meanwhile, in our previous work [8], we defined the action of the $\mathbf{H}$-observables on the twisted coinvariants $\langle \Psi(\mathbf{u}) \rangle_{\mathbf{a}} \in M^{\mathfrak{sl}(N)}$ by its insertion into the correlation function in the limit $\varepsilon_2 \to 0$. Namely,

$$\mathbf{H}^{(\alpha)}(y) \cdot \left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = \left\langle \mathbf{H}^{(\alpha)}(y)\Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} \in M^{\mathfrak{sl}(N)}, \quad \alpha = 0, 1, \cdots, N-1. \quad (3.4)$$

More precisely, the action of the Hecke modifications on the twisted $\widehat{\mathfrak{sl}}(N)$-coinvariants is realized by the insertion of the $\mathbf{H}$-observable and a generating function of additional defect

0-observables, with the extra fugacities. The extra fugacities precisely parametrize the space $\mathbb{P}^{N-1}$ of Hecke modifications. Then, the Hecke operator was defined by a contour integral on this $\mathbb{P}^{N-1}$ in the limit $\varepsilon_2 \to 0$, which picks up the insertion of the **H**-observable without any further insertion of the defect 0-observable.

Using the constraints of the correlation function of the parallel surface defects, it was shown that the Hecke operator satisfies the universal oper equation

$$0 = \left[\partial_y^N + \hat{\mathfrak{t}}_2(y)\partial_y^{N-2} + \cdots + \hat{\mathfrak{t}}_{N-1}(y)\partial_y + \hat{\mathfrak{t}}_N(y) \right] \hat{\mathbf{H}}^{(\alpha)}(y), \quad \alpha = 0, 1, \cdots, N-1, \quad (3.5)$$

where we incorporated a classical part and decoupled an abelian U(1) factor,

$$\hat{\mathbf{H}}(y) = y^{\frac{\bar{m}^-}{\varepsilon_1} + \frac{N-3}{2}} (y - \mathfrak{q})^{-\frac{\bar{m}^- - \bar{a}}{\varepsilon_1}} (y-1)^{1 + \frac{\bar{m}^+ - \bar{a}}{\varepsilon_1}} \mathbf{H}(y), \quad (3.6)$$

so that the universal oper has vanishing coefficient of ∂_y^{N-1}. The universal oper equation is indeed the characteristic feature of the Hecke operator [70–72]. Here, the operator-valued meromorphic functions $\hat{\mathfrak{t}}_k(y)$ have poles of order $\mathrm{ord}_p(\hat{\mathfrak{t}}_k(y)) = -k$ at the marked points $S = \{0, \mathfrak{q}, 1, \infty\}$. Their Laurent coefficients are precisely the Gaudin Hamiltonians (valued in $\mathrm{End}(M^{\mathfrak{sl}(N)})$), given by twisted differential operators on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$. See our previous work [8] for more detail on the $\mathcal{N} = 2$ gauge theoretical construction of the universal oper and its application to the geometric Langlands correspondence.

3.2 Eigenvalue properties and bispectral duality

In our previous work [8] and the companion paper [24], it was shown that the vevs of the regular monodromy surface defects in the limit $\varepsilon_2 \to 0$, on which the Hecke operators and Q-operators act diagonally, give common eigenfunctions of the quantum Hamiltonians of the Gaudin model and the XXX spin chain, respectively. Here, we show that the two spectral problems are equivalent precisely through the transformation between the **Q**-observables and the **H**-observables.

3.2.1 Eigenvalue properties for Q-operators and Hecke operators

In [8], we showed that insertion of a **H**-observable in the limit $\varepsilon_2 \to 0$ gives the action of a Hecke operator on the twisted covariants of $\widehat{\mathfrak{sl}}(N)$-modules. On the distinguished basis $\psi(\mathbf{a}; \mathbf{u})$ constructed as the normalized vacuum expectation value of the regular monodromy surface defect, the action of Hecke operator was shown to be diagonal,

$$\mathbf{H}^{(\alpha)}(y) \cdot \psi(\mathbf{a}; \mathbf{u}) = \lim_{\varepsilon_2 \to 0} \left\langle \mathbf{H}^{(\alpha)}(y)\Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = \chi_\alpha(\mathbf{a}; y)\psi(\mathbf{a}; \mathbf{u}), \quad \alpha = 0, 1, \cdots, N-1, \quad (3.7)$$

due to the cluster decomposition of the **H**-observable and the regular monodromy surface defect in the limit $\varepsilon_2 \to 0$. Here, the eigenvalue $\chi_\alpha(\mathbf{a}; y)$ is precisely the α-th solution to the oper differential equation, constructed as the normalized vacuum expectation value (2.19) of the **H**-observable with the α-th vacuum at infinity [8, 14].

In the point of view of the transition (2.18) between the **Q**-observable and the **H**-observable, this Hecke eigensheaf property is equivalent to the *Q-eigenstate property* of the normalized vacuum expectation value $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy surface defect, now

viewed as an element in the Yangian module $\psi(\mathbf{a}; \mathbf{u}) \in \mathcal{H}^{\mathbb{C}^\times}$. Namely, upon the action of the Q-operators in the limit $\varepsilon_2 \to 0$, we have

$$\mathbf{Q}(x) \cdot \psi(\mathbf{a}; \mathbf{u}) = \lim_{\varepsilon_2 \to 0} \left\langle \mathbf{Q}(x)\Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = Q(\mathbf{a}; x)\psi(\mathbf{a}; \mathbf{u}), \quad \text{same for } \tilde{\mathbf{Q}}(x), \qquad (3.8)$$

again by the cluster decomposition. Thus, just as the Hecke operators, the Q-operators also act diagonally on the basis $\psi(\mathbf{a}; \mathbf{u})$ of the space $\mathcal{H}^{\mathbb{C}^\times}$ enumerated by the ε_1-integral shifts of the Coulomb moduli $\mathbf{a}$, with the eigenvalue being the Q-functions $Q(\mathbf{a}; x)$ (and $\tilde{Q}(\mathbf{a}; x)$) solving the $\hbar$-oper difference equation [24].

In this sense, the Q-operator and the Hecke operator are related to each other by the transition in the surface defect theory, turning on either the vev of the complex adjoint scalar or the complexified FI parameter. In particular, their eigenvalues $- Q(\mathbf{a}; x)$ (or $\tilde{Q}(\mathbf{a}; x)$) and $\chi_\alpha(\mathbf{a}; y) -$ are related precisely by the Fourier transformation (2.18) and its analogues in other convergence domains of $\mathbb{P}^1 \setminus \{0, \mathfrak{q}, 1, \infty\}$. We emphasize again that the surface defect realization of the Q-operator and the Hecke operator in our $\mathcal{N} = 2$ theory setup gives a simple intuitive picture for these eigenvalue properties: the cluster decomposition of the correlation function of surface defects supported on the $\mathbb{C}_1$-plane in the limit $\varepsilon_2 \to 0$.

3.2.2 Bispectral duality of XXX spin chain and Gaudin model

Finally, let us investigate the implication of the transition between the $\mathbf{Q}$-observables and the $\mathbf{H}$-observables on the associated integrable models. Due to the eigenvalue property (3.8), the universal $\hbar$-oper equation (3.3) satisfied by the $\mathbf{Q}$-observables gives

$$\begin{aligned}
0 &= \left[1 - \hat{t}(x)e^{-\varepsilon_1 \partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1 \partial_x}\right] Q(\mathbf{a}; x + \varepsilon_1)\psi(\mathbf{a}; \mathbf{u}; \mathfrak{q}), \\
0 &= \left[1 - \hat{t}(x)e^{-\varepsilon_1 \partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1 \partial_x}\right] \tilde{Q}(\mathbf{a}; x + \varepsilon_1)\psi(\mathbf{a}; \mathbf{u}; \mathfrak{q}),
\end{aligned} \qquad (3.9)$$

at each basis element $\psi(\mathbf{a}; \mathbf{u}; \mathfrak{q})$. Recall that the coefficients of the degree N polynomials $\hat{t}(x) \in \text{End}(\mathcal{H}^{\mathbb{C}^\times})$ are quantum Hamiltonians of the $\mathfrak{gl}(2)$ XXX spin chain with N sites. By the transition (2.18), we can convert the Q-function $Q(\mathbf{a}; x)$ into the oper solution $\chi_\alpha(\mathbf{a}; y)$ by a Fourier transformation. Since the kernel of the Fourier transformation is $y^{-\frac{x}{\varepsilon_1}}$, the shift operator $e^{-\varepsilon_1 \partial_x}$ gets compensated by a multiplication by y while the multiplication by x is converted into the differential operator $-\varepsilon_1 y\partial_y$. Note that $y\partial_y$ obviously commutes with itself so that polynomials in $y\partial_y$ are well-defined. All in all, the above equation is brought into an operator-valued N-th order differential equation [8, 14, 43]

$$0 = \left[\partial_y^N + \hat{t}_2(y)\partial_y^{N-2} + \cdots + \hat{t}_{N-1}(y)\partial_y + \hat{t}_N(y)\right] \hat{\chi}_\alpha(\mathbf{a}; y)\psi(\mathbf{a}; \mathbf{u}; \mathfrak{q}), \quad \alpha = 0, 1, \cdots, N - 1, \qquad (3.10)$$

at each basis element $\psi(\mathbf{a}; \mathbf{u}; \mathfrak{q})$. This is nothing but the universal oper equation (3.5) with the eigenvalue property (3.4) applied. Here, the coefficient of the next-to-highest order differential ∂_y^{N-1} is absent due to the redefinition (3.6) of the $\mathbf{H}$-observable.

By construction above, the coefficients of the Laurent series $\hat{t}_k(y)$ are merely rearrangements of the coefficients of the transfer matrix $\hat{t}(x)$ in the universal $\hbar$-oper (3.9). Hence, we conclude that the quantum Hamiltonians of the XXX spin chain (defined on N bi-infinite

modules over $\mathfrak{gl}(2)$) and the Gaudin model (defined on two (co-)Verma modules and two bi-infinite modules over $\mathfrak{sl}(N)$) are exactly identical as differential operators in the monodromy defect parameters.[10]

We can proceed in the same way for the quantum spectra of the integrable models. Consider the $\hbar$-oper equation

$$
\begin{aligned}
0 &= \left[1 - t(\mathbf{a};x)e^{-\varepsilon_1\partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1\partial_x}\right] Q(\mathbf{a};x+\varepsilon_1), \\
0 &= \left[1 - t(\mathbf{a};x)e^{-\varepsilon_1\partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1\partial_x}\right] \tilde{Q}(\mathbf{a};x+\varepsilon_1),
\end{aligned}
\tag{3.11}
$$

where the coefficients of the degree N polynomial $t(\mathbf{a};x)$ are now normalized vacuum expectation values of the local observables in the limit $\varepsilon_2 \to 0$. Under the transition (2.18), the equation becomes a scalar-valued N-th order differential equation,

$$
0 = \left[\partial_y^N + \mathrm{t}_2(\mathbf{a};y)\partial_y^{N-2} + \cdots + \mathrm{t}_{N-1}(\mathbf{a};y)\partial_y + \mathrm{t}_N(\mathbf{a};y)\right] \hat{\chi}_\alpha(\mathbf{a};y), \quad \alpha = 0, 1, \cdots, N-1.
\tag{3.12}
$$

This is precisely the oper equation derived in [8, 14]. By construction, the Laurent coefficients of $\mathrm{t}_k(\mathbf{a};y)$, $k = 2, 3, \cdots, N$, are simple rearrangements of the coefficients of $t(\mathbf{a};x)$. Thus, we conclude that the quantum spectra of the XXX spin chain and the Gaudin model are also rearrangements of the same set of normalized vacuum expectation values of the local observables, $\lim_{\varepsilon_2\to 0} \left\langle \mathrm{Tr}\,\phi^k \right\rangle_{\mathbf{a}}$, $k = 2, 3, \cdots, N$.

All in all, we prove the equivalence between the spectral problems of the XXX spin chain (defined on N bi-infinite modules over $\mathfrak{gl}(2)$) and the Gaudin model (defined on two (co-)Verma modules and two bi-infinite modules over $\mathfrak{sl}(N)$), associated to the same A_1-quiver $\mathcal{N} = 2$ gauge theory. In particular, the vacuum expectation values $\psi(\mathbf{a};\mathbf{u})$ of the regular monodromy surface defect provide a basis of bispectral wavefunctions.

We stress that in both sides of the duality, the XXX spin chain and the Gaudin model are constructed on bi-infinite modules. In this sense, our $\mathcal{N} = 2$ gauge theoretical approach generalizes the bispectral duality between the two models only involving highest-weight (or lowest-weight) modules. See [73–76] for previous studies of the bispectral duality between the two integrable models. When the gauge theory parameters are specialized to trigger higgsing to a two-dimensional $\mathcal{N} = (2,2)$ theory, the bispectral duality that we established above will reduce to the one for the integrable models on lowest-weight (and further finite-dimensional) modules. For instance, it would be possible to match the Bethe roots of the two models.

4 Surface defect transition and separation of variables

In the previous section, we proved that a common eigenfunction of the Hitchin integrable system — whether in the form of the Gaudin model or of the XXX spin chain — is given by the vacuum expectation value of the regular monodromy surface defect,

$$
\langle \Psi(\mathbf{u}) \rangle_{\mathbf{a}} = \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mu_{\boldsymbol{\lambda}} \Psi(\mathbf{u})[\boldsymbol{\lambda}],
\tag{4.1}
$$

---

[10]Recall that, in the companion paper [24], we presented the quantum Hamiltonians of the XXX spin chain both as differential operators in the monodromy defect parameters $(u_\omega)_{\omega=0}^{N-1}$ and as ε_1-difference operators in the Fourier dual variables $(w_\omega)_{\omega=0}^{N-1}$.

in the limit $\varepsilon_2 \to 0$. Recall that the observable expression is given by (for general coloring) (2.26),

$$\Psi_{\mathbf{c}}(\mathbf{u})[\boldsymbol{\lambda}] = \sum_{\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{k_\omega - k_{l-1}} \mathbb{E}\left[\frac{S\Sigma_0^*}{P_1^*} + \sum_{\omega=0}^{l-2} \frac{-\Sigma_\omega(\Sigma_\omega - \Sigma_{\omega+1})^* - M_\omega^+ \Sigma_\omega^* + q_1^{-1}(M_\omega^-)^* \Sigma_\omega}{P_1^*}\right].$$
(4.2)

In this section, we will re-express this as a surface observable created by coupling a $\mathcal{N} = (2,2)$ gauged linear sigma model and show that such a dual description leads to the separation of variables for the XXX spin chain and the Gaudin model.

4.1 Dual description of monodromy surface defect as coupling to sigma model

First, let us remind that the equivariant Chern characters of the framing bundle and the tautological bundle can be written as

$$\begin{aligned}
N_\omega &= \sum_{\omega=0}^{N-1} e^{a_\omega} \\
K_\omega &= \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{\substack{(i,j) \in \Lambda^{(\alpha)} \\ c(\alpha)+j-1 = \omega + l(J-1)}} q_1^{i-1} q_2^{J-1} \\
&= \frac{1}{P_1} \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=J_{\min,\alpha,\omega}}^{J_{\max,\alpha,\omega}} \left(1 - q_1^{\Lambda^{(\alpha)t}_{\omega+1-c(\alpha)+l(J-1)}}\right) q_2^{J-1},
\end{aligned}$$
(4.3)

where the lower and the upper bound of the J-summation are given by

$$J_{\min,\alpha,\omega} = \begin{cases} 1 & \text{if } 0 \le c(\alpha) \le \omega \\ 2 & \text{if } \omega < c(\alpha) \le l-1 \end{cases},$$

$$J_{\max,\alpha,\omega} = 1 + \left\lfloor \frac{c(\alpha) + \Lambda_1^{(\alpha)} - 1 - \omega}{l} \right\rfloor, \qquad J_{\max,\alpha} := J_{\max,\alpha,l-1} = \left\lfloor \frac{c(\alpha) + \Lambda_1^{(\alpha)}}{l} \right\rfloor.$$
(4.4)

At a fixed point in the moduli space of instantons labelled by the partitions $\boldsymbol{\lambda} = (\lambda^{(0)}, \lambda^{(1)}, \cdots, \lambda^{(N-1)})$, the $\mathcal{Y}$-observable can be reorganized into

$$\begin{aligned}
\mathcal{Y}(x)[\boldsymbol{\lambda}] &= \prod_{\alpha=0}^{N-1} (x - a_\alpha) \prod_{i=1}^{l(\lambda^{(\alpha)})} \prod_{j=1}^{\lambda_i^{(\alpha)}} \frac{(x - c_\square - \varepsilon_1)(x - c_\square - \varepsilon_2)}{(x - c_\square)(x - c_\square - \varepsilon)} \\
&= \prod_{\alpha=0}^{N-1} (x - A_\alpha) \prod_{i=1}^{l(\lambda^{(\alpha)})} \frac{x - (i-1)\varepsilon_1 - \lambda_i^{(\alpha)} \varepsilon_2}{x - i\varepsilon_1 - \lambda_i^{(\alpha)} \varepsilon_2} \\
&= \prod_{\alpha=0}^{N-1} (x - A_\alpha) \prod_{J=1}^{\lambda_1^{(\alpha)}} \frac{x - a_\alpha - \lambda_{J+1}^{(\alpha)t} \varepsilon_1 - J\varepsilon_2}{x - a_\alpha - \lambda_J^{(\alpha)t} \varepsilon_1 - J\varepsilon_2},
\end{aligned}$$
(4.5)

where we defined $A_\alpha := a_\alpha + l(\lambda^{(\alpha)})\varepsilon_1$. Equivalently, the equivariant Chern character of the universal sheaf can be expressed as

$$S[\boldsymbol{\lambda}] = \sum_{\alpha=0}^{N-1} e^{A_\alpha} + \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=1}^{\lambda_1^{(\alpha)}} \left(q_1^{\lambda_{J+1}^{(\alpha)t}} - q_1^{\lambda_J^{(\alpha)t}} \right) q_2^J$$

$$:= F_{l-1}[\boldsymbol{\lambda}] + P_1 W[\boldsymbol{\lambda}],$$

(4.6)

where we defined

$$F_{l-1}[\boldsymbol{\lambda}] := \sum_{\alpha=0}^{N-1} e^{A_\alpha}$$

$$W[\boldsymbol{\lambda}] := \frac{1}{P_1} \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=1}^{\lambda_1^{(\alpha)}} \left(q_1^{\lambda_{J+1}^{(\alpha)t}} - q_1^{\lambda_J^{(\alpha)t}} \right) q_2^J.$$

(4.7)

Note that $W[\boldsymbol{\lambda}]$ is in fact a pure character, since $\lambda_J^{(\alpha)t} \geq \lambda_{J+1}^{(\alpha)t}$ for all $1 \leq J \leq \lambda_1^{(\alpha)}$ so that each summand is divisible by $P_1 = 1 - q_1$. Thus, $W[\boldsymbol{\lambda}]$ is in fact the equivariant Chern character of a vector space, which we denote by the same letter $W[\boldsymbol{\lambda}]$ unless the notation incurs any confusion. It is straightforward to see that $w[\boldsymbol{\lambda}] := \dim W[\boldsymbol{\lambda}] = \sum_{\alpha=0}^{N-1} l(\lambda^{(\alpha)})$.

Now we reorganize the monodromy defect observable (2.26) in such a way that the dual description as coupling two-dimensional sigma model becomes more manifest. Recall that we defined

$$\Sigma_\omega[\boldsymbol{\Lambda}] := S_{\omega+1}[\boldsymbol{\Lambda}] + S_{\omega+2}[\boldsymbol{\Lambda}] + \cdots + S_{l-1}[\boldsymbol{\Lambda}]$$
$$= N_{\omega+1} + \cdots + N_{l-1} + P_1 K_\omega - P_1 K_{l-1}, \qquad \omega = 0, 1, \cdots, l-2.$$
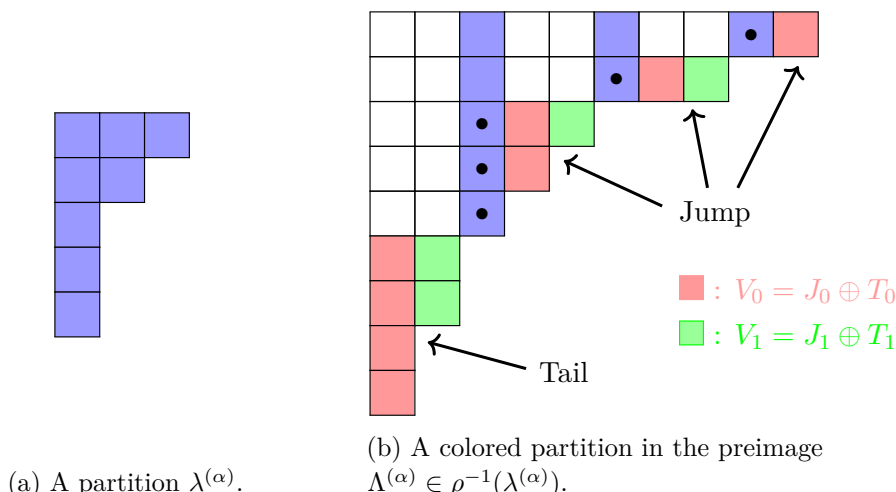
(4.8)

This can be reorganized as

$$\Sigma_\omega = \sum_{\omega < \omega' \leq l-1} L_{\omega'} + P_1 V_\omega, \qquad \omega = 0, 1, \cdots, l-2,$$

(4.9)

where we defined

$$L_\omega[\boldsymbol{\lambda}] = \sum_{c(\alpha)=\omega} e^{A_\alpha}$$

$$V_\omega[\boldsymbol{\Lambda}] = J_\omega[\boldsymbol{\Lambda}] + T_\omega[\boldsymbol{\Lambda}],$$

$$J_\omega[\boldsymbol{\Lambda}] = \frac{1}{P_1} \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=1}^{\lambda_1^{(\alpha)}} \left(q_1^{\lambda_{J+1}^{(\alpha)t}} - q_1^{\Lambda_{\omega+1-c(\alpha)+lJ}^{(\alpha)t}} \right) q_2^J,$$

$$T_\omega[\boldsymbol{\Lambda}] = \frac{1}{P_1} \sum_{0 \leq c(\alpha) \leq \omega} e^{A_\alpha} \left(1 - q_1^{\Lambda_{\omega-c(\alpha)+1}^{(\alpha)t} - \lambda_1^{(\alpha)t}} \right).$$

(4.10)

Note that $L_\omega[\boldsymbol{\lambda}]$ only depends on the four-dimensional instanton configuration represented by $\boldsymbol{\lambda}$. Also, it is crucial that $J_\omega[\boldsymbol{\Lambda}]$ and $T_\omega[\boldsymbol{\Lambda}]$ are in fact pure characters, and therefore they are $\mathsf{T_H}$-equivariant Chern characters of vector spaces which we denote by the same letters. Indeed, there is a sandwich relation constraining the allowed colored partitions $\boldsymbol{\Lambda}$ for given $\boldsymbol{\lambda}$,

$$\lambda_J^{(\alpha)t} = \Lambda_{lJ-c(\alpha)}^{(\alpha)t} \geq \Lambda_{\omega+lJ-c(\alpha)}^{(\alpha)t} \geq \Lambda_{l+lJ-c(\alpha)}^{(\alpha)t} = \lambda_{J+1}^{(\alpha)t}, \qquad \omega = 0, 1, \cdots, l-1,$$

(4.11)

(a) A partition $\lambda^{(\alpha)}$.

(b) A colored partition in the preimage $\Lambda^{(\alpha)} \in \rho^{-1}(\lambda^{(\alpha)})$.

Figure 6. Monodromy defect observable encoded in tails and jumps of colored partitions, exemplified in the case of $l = 3$. The colored partition $\Lambda^{(\alpha)}$ (with $c(\alpha) = 0$ chosen in the example) is represented by a colored Young diagram depicted in (b). The projection $\rho : \Lambda^{(\alpha)} \mapsto \lambda^{(a)}$ maps $\Lambda^{(a)}$ to the subdiagram $\lambda^{(a)}$ composed of the columns carrying $\mathbb{Z}_l$-charge $l - 1$ (the blue boxes carrying $\mathbb{Z}_3$-charge 2 in the example, $W[\boldsymbol{\lambda}]$ is the subspace with the bullets). Conversely, a colored partition $\Lambda^{(a)}$ in the preimage of a given $\lambda^{(\alpha)}$ can be reconstructed from V_ω, $\omega = 0, 1, \cdots, l - 2$ (V_0 and V_1 depicted by the red boxes and green boxes in the example). Each V_ω is the direct sum of the jump J_ω, which is constructed based on $W[\boldsymbol{\lambda}]$, and the tail T_ω, which is independent of $\lambda^{(\alpha)}$. Note that the maximal number of boxes in each jump is constrained by the shape of $\lambda^{(\alpha)}$, while the tail can grow indefinitely toward downward.

which implies the summands of J_ω and T_ω are divisible by P_1. In turn, V_ω is actually a direct sum of two vector spaces,

$$V_\omega = J_\omega \oplus T_\omega, \qquad \omega = 0, 1, \cdots, l - 2, \tag{4.12}$$

with dimensions

$$v_\omega := \dim V_\omega = k_\omega - k_{l-1} + \sum_{\omega < c(\alpha) \leq l-1} l(\lambda^{(\alpha)}) \geq 0. \tag{4.13}$$

We refer to the subspace J_ω as the *jump*, and to the subspace T_ω as the *tail*. The jump and the tail encodes how the colored partition $\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})$ is reconstructed for a given $\boldsymbol{\lambda}$. The jump describes the restricted Young diagrams that are put in between the columns of $\boldsymbol{\lambda}$, while the tail describes the restricted Young diagrams attached to the left of the left-most column of $\boldsymbol{\lambda}$. See figure 6 for an example.

Note that if we decouple the two-dimensional degrees of freedom by turning off the bulk gauge coupling $\mathfrak{q} \to 0$, then there is no four-dimensional instanton so that we only have $\boldsymbol{\lambda} = \varnothing$, which leads to $W[\boldsymbol{\lambda}] = \varnothing$ and therefore all the jumps become empty $J_\omega = \varnothing$, $\omega = 0, 1, \cdots, l - 2$. The resulting two-dimensional partition function of the defect becomes a sum over restricted Young diagrams represented by the tail, which manifests the dual description of the defect as coupling to certain two-dimensional sigma model, as studied in [14]. We will now extend such a dual description of the monodromy defect to non-zero bulk coupling $\mathfrak{q} \neq 0$.
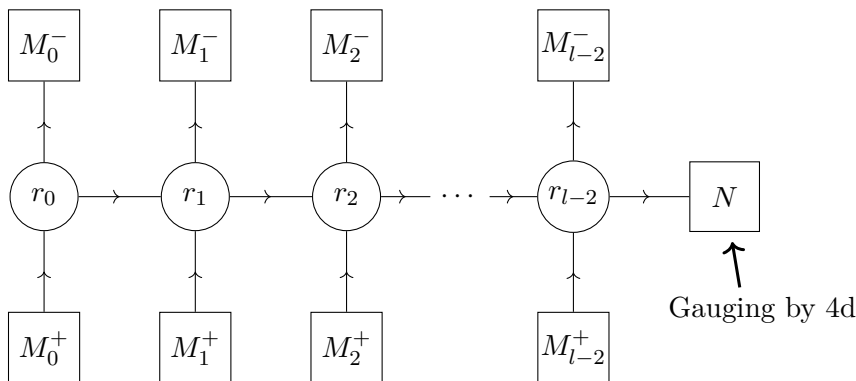
Figure 7. The $\mathcal{N} = (2,2)$ quiver for the two-dimensional gauged linear sigma model coupled to the four-dimensional $\mathcal{N} = 2$ gauge theory, giving dual description of the monodromy surface defect.

For this, let us re-express the monodromy defect observable (2.26) using Σ_ω (4.8) as

$$
\begin{aligned}
\Psi(\mathbf{u})[\boldsymbol{\lambda}] &= \sum_{\hat{\boldsymbol{\lambda}} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{k_\omega - k_{l-1}} \mathbb{E}\left[\frac{S\Sigma_0^*}{P_1^*} + \sum_{\omega=0}^{l-2} \frac{-\Sigma_\omega(\Sigma_\omega - \Sigma_{\omega+1})^* - M_\omega^+ \Sigma_\omega^* + q_1^{-1}(M_\omega^-)^* \Sigma_\omega}{P_1^*} \right] \\
&= \sum_{\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{k_\omega - k_{l-1}} \mathbb{E}\left[\sum_{\omega=0}^{l-2} \left(S - \sum_{\omega < \omega' \leq l-1} L_{\omega'} \right) \frac{L_{\omega+1}^*}{P_1^*} \right. \\
&\quad \left. - \frac{\sum_{\omega=0}^{l-2} \sum_{\omega < \omega' \leq l-1} M_\omega^+ L_{\omega'}^* - q_1^{-1} L_{\omega'}(M_\omega^-)^*}{P_1^*} \right] \\
&\quad \times \mathbb{E}\left[P_1 W V_0^* + \sum_{\omega=0}^{l-2} \left(L_\omega V_\omega^* + q_1 V_\omega L_{\omega+1}^* - P_1 V_\omega(V_\omega - V_{\omega+1})^* - M_\omega^+ V_\omega^* - (M_\omega^-)^* V_\omega \right) \right].
\end{aligned}
\tag{4.14}
$$

We will explain that this expression of the monodromy surface defect observable manifests the dual description of the surface defect, coupling of the two-dimensional $\mathcal{N} = (2,2)$ gauged linear sigma model which is represented as a $\mathcal{N} = (2,2)$ quiver gauge theory in figure 7. Here, the coupling is initiated by gauging the $\mathrm{GL}(N)$ flavor symmetry for the right-most flavor node by the four-dimensional gauge field restricted to the surface. The Higgs branch of this gauged linear sigma model is the total space of a vector bundle,

$$
\bigoplus_{\omega=0}^{l-2} \mathcal{E}_\omega \otimes \mathbb{C}^{|c_f^{-1}(\omega)|} \oplus \bigoplus_{\omega=0}^{l-2} \mathcal{E}_\omega^\vee \otimes \mathbb{C}^{|c_{af}^{-1}(\omega)|} \longrightarrow F(r_0, r_1, \cdots, r_{l-1}),
\tag{4.15}
$$

over the partial flag variety $F(r_0, r_1, \cdots, r_{l-1})$. Here, $r_\omega = |\{\alpha \,|\, 0 \leq c(\alpha) \leq \omega\}|$ and a partial flag in $F(r_0, r_1, \cdots, r_{l-1})$ is a filtration of the complex N-dimensional vector space $\mathbb{C}^N$ by its subspaces of dimensions $0 \leq r_0 \leq r_1 \leq \cdots \leq r_{l-1} = N$. Finally, $\mathcal{E}_\omega$ is the ω-th tautological vector bundle over the flag variety $F(r_0, r_1, \cdots, r_{l-1})$ with the rank $\mathrm{rk}\,\mathcal{E}_\omega = r_\omega$, and $\mathcal{E}_\omega^\vee = \mathrm{Hom}(\mathcal{E}_\omega, \mathbb{C})$ is its dual vector bundle.

It has been well-known that the partition function of such a gauged linear sigma model, without coupling to the four-dimensional theory, localizes to an integral of equivariant Euler class of a vector bundle over (smooth resolution of) the moduli space of quasi-maps from

$\mathbb{P}^1$ to $F(r_0, r_1, \cdots, r_{l-1})$. The moduli space of these quasi-maps admits a construction from the hand-saw quiver [39, 40]. Now, we show that the surface defect observable (4.14) can be obtained by an integral of equivariant Euler class of a vector bundle over this hand-saw quiver variety modified by the effect of the coupling to the bulk four-dimensional theory. In this sense, we can regard the observable expression (4.14) as the manifestation of this dual description of the monodromy surface defect.

4.1.1 Modified hand-saw quiver variety

Let us now introduce our modified hand-saw quiver variety. Consider the following collection of vector spaces,

- $l-1$ vector spaces $V_\omega = \mathbb{C}^{v_\omega}$, $\omega = 0, 1, \cdots, l-2$.

- l vector spaces $L_\omega[\boldsymbol{\lambda}] = \mathbb{C}^{r_\omega - r_{\omega-1}}$, $\omega = 0, 1, \cdots, l-1$.

- A vector space $W[\boldsymbol{\lambda}] = \mathbb{C}^{w[\boldsymbol{\lambda}]}$.

The vector spaces $(L_\omega[\boldsymbol{\lambda}])_{\omega=0}^{l-1}$ and $W[\boldsymbol{\lambda}]$ are defined upon a given instanton configuration represented by $\boldsymbol{\lambda}$, in the following way. Firstly, their dimensions are determined by $\boldsymbol{\lambda}$ as $\dim L_\omega = r_\omega - r_{\omega-1} = |\{\alpha \,|\, c(\alpha) = \omega\}|$ and $\dim W = w = \sum_{\alpha=0}^{N-1} l(\lambda^{(\alpha)})$. Moreover, the equivariant weights of these vector spaces under the Cartan of the global symmetry group $\mathbb{C}_{\varepsilon_1}^\times \times \mathrm{GL}(W) \times \bigtimes_{\omega=0}^{l-1} \mathrm{GL}(L_\omega)$ are imposed as (4.10) and (4.7), namely,

$$
\begin{aligned}
L_\omega[\boldsymbol{\lambda}] &= \sum_{c(\alpha)=\omega} e^{A_\alpha} = \sum_{c(\alpha)=\omega} e^{a_\alpha} q_1^{l(\lambda^{(\alpha)})} \\
W[\boldsymbol{\lambda}] &= \frac{1}{P_1} \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=1}^{\lambda_1^{(\alpha)}} \left(q_1^{\lambda_{J+1}^{(\alpha)t}} - q_1^{\lambda_J^{(\alpha)t}} \right) q_2^J,
\end{aligned}
\tag{4.16}
$$

where we recall that the character of $W[\boldsymbol{\lambda}]$ above is in fact a pure character. Note that W is decomposed into column segments, $W[\boldsymbol{\lambda}] = \bigoplus_{\alpha=0}^{N-1} \bigoplus_{J=1}^{\lambda_1^{(\alpha)}} \left(\bigoplus_{n \geq 0} b^n w_{\alpha,J} \right)$, where $w_{\alpha,J} \in W$ is the basis vector carrying the character $e^{a_\alpha} q_1^{\lambda_{J+1}^{(\alpha)t}} q_2^J$ and we defined the map $b \in \mathrm{End}(W)$ sending each basis vector to another basis vector carrying one more charge of $q_1 \in \mathbb{C}_{\varepsilon_1}^\times$.[11]

The modified hand-saw quiver is represented by figure 8.

Here, the arrows between the vector spaces are the linear maps,

$$
\begin{aligned}
B_1 &\in \bigoplus_{\omega=0}^{l-2} \mathrm{Hom}(V_\omega, V_\omega), \quad B_2 \in \bigoplus_{\omega=0}^{l-3} \mathrm{Hom}(V_\omega, V_{\omega+1}) \\
i &\in \bigoplus_{\omega=0}^{l-2} \mathrm{Hom}(L_\omega, V_\omega), \quad\; j \in \bigoplus_{\omega=0}^{l-2} \mathrm{Hom}(V_\omega, L_{\omega+1}), \\
I &\in \mathrm{Hom}(W, V_0).
\end{aligned}
\tag{4.17}
$$

---

[11]$b \in \mathrm{End}(W)$ is the restriction of $B_1 : K \to K$, which is a part of the ADHM data for the instanton moduli space, to the subspace W.
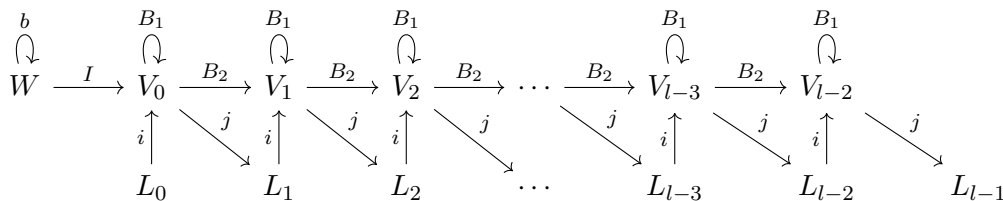
Figure 8. The modified hand-saw quiver provides the dual description of the monodromy defect observable: the $\mathcal{N} = (2, 2)$ gauged linear sigma model coupled to the four-dimensional $\mathcal{N} = 2$ gauge theory. Note that our modified hand-saw quiver includes a *handle* from the newly introduced vector space W and the map $I : W \to V_0$, in contrast to the original hand-saw quiver of [39, 40].

We define the complex moment maps by

$$\mu_\omega := (B_1)_{\omega+1}(B_2)_\omega - (B_2)_\omega(B_1)_\omega + i_{\omega+1}j_\omega : V_\omega \longrightarrow V_{\omega+1}, \qquad \omega = 0, 1, \cdots, l-3, \quad (4.18)$$

and

$$\mu_{-1} := (B_1)_0 I - Ib : W \longrightarrow V_0. \tag{4.19}$$

We also define the real moment maps by

$$\mu_\omega^{\mathbb{R}} := \left[(B_1)_\omega, (B_1)_\omega^\dagger\right] + (B_2)_\omega(B_2)_{\omega-1}^\dagger - (B_2)_{\omega+1}^\dagger(B_2)_\omega + i_\omega i_\omega^\dagger - j_\omega^\dagger j_\omega : V_\omega \longrightarrow V_\omega, \quad (4.20)$$

for $\omega = 0, 1, \cdots, l-2$, where we used the notation $(B_2)_{-1} = I$.

The locus of the moment map equations $\bigcap_{\omega=-1}^{l-3} \mu_\omega^{-1}(0) \cap \bigcap_{\omega=0}^{l-2}(\mu_\omega^{\mathbb{R}})^{-1}(\zeta \operatorname{id}_{V_\omega})$, where we have chosen a stability chamber by $\zeta \in \mathbb{R}$, is acted on by the gauge group

$$G = \bigtimes_{\omega=0}^{l-2} \mathrm{U}(V_\omega), \tag{4.21}$$

where the action is given by

$$((B_1)_\omega, (B_2)_\omega, i_\omega, j_\omega, I) \mapsto (g_\omega(B_1)_\omega g_\omega^{-1}, g_{\omega+1}(B_2)_\omega g_\omega^{-1}, g_\omega i_\omega, j_\omega g_\omega^{-1}, g_0 I), \quad g_\omega \in \mathrm{U}(V_\omega). \tag{4.22}$$

The modified hand-saw quiver variety, at the stability chamber given by $\zeta \in \mathbb{R}$, is defined as the quotient by this gauge group action,

$$\mathcal{M}_{\text{hand-saw}}^\zeta[\boldsymbol{\lambda}] = \bigcap_{\omega=-1}^{l-3} \mu_\omega^{-1}(0) \cap \bigcap_{\omega=0}^{l-2}(\mu_\omega^{\mathbb{R}})^{-1}(\zeta \operatorname{id}_{V_\omega}) \bigg/ G. \tag{4.23}$$

Note that imposing the moment map equation $\mu_{-1} = 0$ implies

$$(B_1)_0 I(W) \subset I(W). \tag{4.24}$$

The modified hand-saw quiver variety is acted on by the global symmetry group $\mathsf{H} = \mathbb{C}_{\varepsilon_1}^\times \times \mathrm{GL}(W) \times \bigtimes_{\omega=0}^{l-1} \mathrm{GL}(L_\omega)$. The $\mathbb{C}_{\varepsilon_1}^\times$-action descends from the two-dimensional spacetime isometry, so that it acts as

$$((B_1)_\omega, (B_2)_\omega, i_\omega, j_\omega, I) \mapsto (q_1(B_1)_\omega, (B_2)_\omega, i_\omega, q_1 j_\omega, I), \qquad q_1 \in \mathbb{C}_{\varepsilon_1}^\times, \tag{4.25}$$

and the action of $\mathrm{GL}(W) \times \times_{\omega=0}^{l-1} \mathrm{GL}(L_\omega)$ framing is given by

$$((B_1)_\omega, (B_2)_\omega, i_\omega, j_\omega, I) \mapsto \left((B_1)_\omega, (B_2)_\omega, i_\omega h_\omega^{-1}, h_{\omega+1} j_\omega, I h_W^{-1}\right), \quad h_\omega \in \mathrm{GL}(L_\omega), h_W \in \mathrm{GL}(W). \tag{4.26}$$

Recall that we want the modified hand-saw quiver variety to describe the supersymmetric localization locus of the two-dimensional gauged linear sigma model coupled to the four-dimensional gauge theory, so that this global symmetry group is embedded in the global symmetry group of the instanton moduli space,

$$\mathsf{H} \hookrightarrow \mathbb{C}_{\varepsilon_1}^\times \times \mathbb{C}_{\varepsilon_2}^\times \times \mathrm{GL}(N), \tag{4.27}$$

where the $\mathbb{C}_{\varepsilon_1}^\times \subset \mathsf{H}$ part is simply identified with $\mathbb{C}_{\varepsilon_1}^\times$ in the image. The framing part of H embeds $\mathrm{GL}(W) \times \times_{\omega=0}^{l-1} \mathrm{GL}(L_\omega) \hookrightarrow \mathbb{C}_{\varepsilon_1}^\times \times \mathbb{C}_{\varepsilon_2}^\times \times \mathrm{GL}(N)$ in such a way that the equivariant Chern characters of W and $(L\omega)\omega = 0^{l-1}$ are given by (4.16). In particular, the embedding breaks the global gauge symmetry group of the four-dimensional $\mathcal{N} = 2$ theory to $\times_{\omega=0}^{l-1} \mathrm{GL}(|c^{-1}(\omega)|) \subset \mathrm{GL}(N)$.

Since the Higgs branch of our gauged linear sigma model is the vector bundle (4.15) over the partial flag variety, we need to consider the corresponding matter bundle over the modified hand-saw quiver variety,

$$\mathcal{E}_{\mathrm{matter}} = \bigoplus_{\omega=0}^{l-2} \mathrm{Hom}(M_\omega^+, V_\omega) \oplus \bigoplus_{\omega=0}^{l-2} \mathrm{Hom}(V_\omega, M_\omega^-). \tag{4.28}$$

The flavor symmetry group $\mathsf{F} = \times_{\omega=0}^{l-1} \left(\mathrm{GL}(M_\omega^+) \times \mathrm{GL}(M_\omega^-)\right)$ acts in an obvious way on the fiber of the bundle. Note that this flavor symmetry group embeds into the flavor symmetry group for the matter bundle over the bulk instanton moduli space. The embedding is determined by the flavor coloring functions, $\mathsf{F} \simeq \times_{\omega=0}^{l-1} \left(\mathrm{GL}(|c_f^{-1}(\omega)|) \times \mathrm{GL}(|c_{af}^{-1}(\omega)|)\right) \subset \mathrm{GL}(M^+) \times \mathrm{GL}(M^-)$.

At this point, we stress again that the modified hand-saw quiver variety is defined based on a given instanton configuration in the four-dimensional bulk. If we turn off the four-dimensional gauge coupling ($\mathfrak{q} \to 0$), the instanton moduli space becomes empty, so we can only take $\boldsymbol{\lambda} = \varnothing$. This implies that $W[\boldsymbol{\lambda}] = \varnothing$ and that $I : W \to V_0$ is a trivial map. The equivariant Chern characters of the vector spaces L_ω also simplify to $L_\omega = \sum_{\alpha \in c^{-1}(\omega)} e^{a_\alpha}$. In this limit, the *handle* (the vector space W and the map I) disappears, and we recover the original hand-saw quiver variety from [39, 40]. In this sense, our modification of the hand-saw quiver variety reflects the effect of coupling the two-dimensional sigma model to the four-dimensional bulk theory.

4.1.2 Recovering monodromy defect observable in positive chamber

We now show that the monodromy defect observable (4.14) can be recovered by an equivariant integral of the Euler class of a vector bundle over the modified hand-saw quiver variety (4.23), in the positive stability chamber $\zeta > 0$.

In this chamber, we can compensate the real moment map equations $\mu_\omega^{\mathbb{R}} = \zeta \, \mathrm{id}_{V_\omega}$ by the stability condition

$$
\begin{aligned}
V_0 &= \mathbb{C}[B_1]I(W) \oplus \mathbb{C}[B_1]i(L_0), \\
V_\omega &= \mathbb{C}[B_1]B_2(V_{\omega-1}) \oplus \mathbb{C}[B_1]i(L_\omega), \qquad \omega = 1, 2, \cdots, l-2.
\end{aligned}
\tag{4.29}
$$

and complexified gauge group action $G_\mathbb{C} = \bigtimes_{\omega=0}^{l-2} \mathrm{GL}(V_\omega)$. Namely,

$$
\mathcal{M}_{\text{hand-saw}}^{\zeta > 0}[\boldsymbol{\lambda}] = \left(\bigcap_{\omega=-1}^{l-3} \mu_\omega^{-1}(0) \right)^{\mathrm{stab}_{\zeta > 0}} \Big/ G_\mathbb{C}.
\tag{4.30}
$$

To evaluate equivariant integral over this moduli space, we need to classify the fixed points under the action of the Cartan of the global symmetry group $\mathsf{T_H} \subset \mathsf{H}$. At a fixed point, the $\mathsf{T_H}$-action is compensated by a gauge transformation of $G_\mathbb{C} = \bigtimes_{\omega=0}^{l-2} \mathrm{GL}(V_\omega)$. This implies that each V_ω is a representation of $\mathsf{T_H}$. Due to the stability condition (4.29), V_ω is the direct sum of a subspace spanned by the vectors of the form $B_1^i B_2^\omega I(W)$ and a subspace spanned by the vectors of the form $B_1^i B_2^{\omega-\omega'} i(L_{\omega'})$ where $i \geq 0$ and $\omega' \leq \omega$. Each of them comprises a Young diagram with varying ω, due to the moment map equations. Consequently, the former subspace gives a jump J_ω, while the latter subspace produces a tail T_ω. To sum up, the fixed points are precisely classified by the colored partitions for the given $\boldsymbol{\lambda}$, i.e., the preimage $\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})$. At each fixed point $\boldsymbol{\Lambda}$ we reproduce V_ω as the representation of $\mathsf{T_H}$ carried by the direct sum of the jump and the tail (4.12), $V_\omega = J_\omega \oplus T_\omega$ (recall figure 6). Their equivariant Chern characters are therefore given by (4.10),

$$
\begin{aligned}
V_\omega[\boldsymbol{\Lambda}] &= J_\omega[\boldsymbol{\Lambda}] + T_\omega[\boldsymbol{\Lambda}] \\
J_\omega[\boldsymbol{\Lambda}] &= \frac{1}{P_1} \sum_{\alpha=0}^{N-1} e^{a_\alpha} \sum_{J=1}^{\lambda_1^{(\alpha)}} \left(q_1^{\lambda_{J+1}^{(\alpha)t}} - q_1^{\Lambda_{\omega+1-c(\alpha)+lJ}^{(\alpha)t}} \right) q_2^J, \qquad \omega = 0, 1, \cdots, l-2. \\
T_\omega[\boldsymbol{\Lambda}] &= \frac{1}{P_1} \sum_{0 \leq c(\alpha) \leq \omega} e^{A_\alpha} \left(1 - q_1^{\Lambda_{\omega-c(\alpha)+1}^{(\alpha)t} - \lambda_1^{(\alpha)t}} \right)
\end{aligned}
\tag{4.31}
$$

Also, we need the equivariant Euler classes of the vector bundles. It can be computed once the equivariant Chern character is obtained by taking the product of all the equivariant weights. For the tangent bundle, it is straightforward to compute

$$
T\mathcal{M}_{\text{hand-saw}}[\boldsymbol{\Lambda}] = P_1 W V_0^* + \sum_{\omega=0}^{l-2} \left(L_\omega V_\omega^* + q_1 V_\omega L_{\omega+1}^* - P_1 V_\omega (V_\omega - V_{\omega+1})^* \right),
\tag{4.32}
$$

from the equivariant Chern characters of the linear maps ((4.25) and (4.26)), the linearized complex moment map equations ((4.18) and (4.19)), and the linearized complexified gauge transformations ((4.22)). For the matter bundle, we also simply get

$$
\mathcal{E}_{\text{matter}}[\boldsymbol{\Lambda}] = \sum_{\omega=0}^{l-2} (M_\omega^+ V_\omega^* + (M_\omega^-)^* V_\omega).
\tag{4.33}
$$

All in all, the integral of the equivariant Euler class of the matter bundle over the modified hand-saw quiver variety is computed by equivariant localization as

$$
\begin{aligned}
\Psi[\boldsymbol{\lambda}]^{\text{non-pert}}_{\zeta>0} &= \sum_{v_0,v_1,\cdots,v_{l-2}=0}^{\infty} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{v_\omega} \int_{\mathcal{M}^{\zeta>0}_{\text{hand-saw}}[\boldsymbol{\lambda}]} e_{\mathsf{T}_{\mathsf{H}}}(\mathcal{E}_{\text{matter}}) \\
&= \sum_{\boldsymbol{\Lambda}\in\rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{v_\omega} \, \mathbb{E}\Bigg[P_1 W V_0^* + \sum_{\omega=0}^{l-2}\Big(L_\omega V_\omega^* + q_1 V_\omega L_{\omega+1}^* - P_1 V_\omega (V_\omega - V_{\omega+1})^* \\
&\qquad\qquad\qquad\qquad\qquad - M_\omega^+ V_\omega^* - (M_\omega^-)^* V_\omega \Big) \Bigg],
\end{aligned}
\tag{4.34}
$$

where we wrote the complexified FI parameters of the $\mathcal{N}=(2,2)$ gauged linear sigma model (which are the complexified Kähler parameters in the non-linear sigma model description) as $(\mathfrak{q}_\omega)_{\omega=0}^{l-2}$. Thus, with the identification between the monodromy parameters and the complexified FI parameters $(\mathfrak{q}_\omega)_{\omega=0}^{l-2}$, the above expression recovers the non-perturbative part of the monodromy defect observable (4.14), after augmented by the following classical contribution of the two-dimensional sigma model,

$$
\mathbf{l_u} := \prod_{\omega=1}^{l-1} \left(\frac{u_\omega}{u_0}\right)^{-\sum_{\alpha\in c^{-1}(\omega)} l(\lambda^{(\alpha)})} = \prod_{\omega=1}^{l-1} \left(\frac{u_\omega}{u_0}\right)^{\sum_{\alpha\in c^{-1}(\omega)} \frac{a_\alpha - A_\alpha}{\varepsilon_1}}.
\tag{4.35}
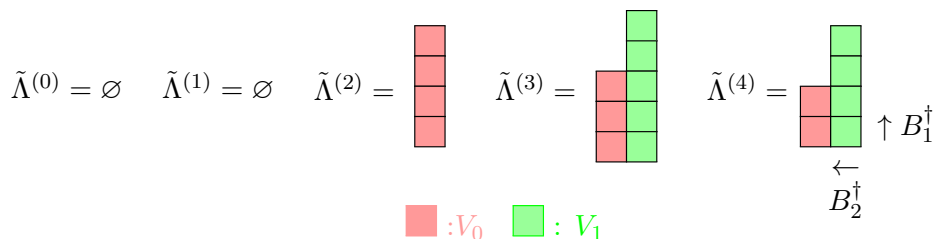$$

4.1.3 Negative chamber and surface defect transition

Now, let us consider the modified hand-saw quiver variety in the negative stability chamber $\zeta < 0$. In this chamber, we can compensate the real moment map equations $\mu_\omega^{\mathbb{R}} = \zeta \, \text{id}_{V_\omega}$ by another stability condition

$$
\begin{aligned}
V_{l-2} &= \mathbb{C}[B_1^\dagger] j^\dagger(L_{l-1}) \\
V_\omega &= \mathbb{C}[B_1^\dagger] B_2^\dagger(V_{\omega+1}) \oplus \mathbb{C}[B_1^\dagger] j^\dagger(L_{\omega+1}), \qquad \omega = 0,1,\cdots,l-3.
\end{aligned}
\tag{4.36}
$$

The modified hand-saw quiver variety in this chamber is thus given by

$$
\mathcal{M}^{\zeta<0}_{\text{hand-saw}} = \left(\bigcap_{\omega=-1}^{l-3} \mu_\omega^{-1}(0)\right)^{\text{stab}_{\zeta<0}} \Big/ G_\mathbb{C}.
\tag{4.37}
$$

Let us classify the T_{H}-fixed points. By the same argument as above, the vector spaces $(V_\omega)_{\omega=0}^{l-2}$ become representations of T_{H} at the fixed points. It should be noted that the jump disappears in this chamber, because there is no equivalent of I map participating in the construction of the vector spaces V_ω. Thus the whole vector space V_ω is spanned by the vectors of the form $(B_1^\dagger)^i (B_2^\dagger)^{\omega'-\omega-1} j^\dagger(L_{\omega'})$ for $\omega' > \omega$ and $i \geq 0$, which can be represented by columns growing in the upward direction. Due to the moment map equations, there is an hierarchy between the heights of the columns with different ω so that they form Young diagrams. These Young diagrams have restricted horizontal lengths, because of the vanishing condition $B_2^\dagger(V_0) = \varnothing$. Hence, the fixed points in the negative chamber are fully classified by N tuples of such restricted colored Young diagrams, which we collectively denote by

$$\tilde{\Lambda}^{(0)} = \varnothing \quad \tilde{\Lambda}^{(1)} = \varnothing \quad \tilde{\Lambda}^{(2)} = \quad \tilde{\Lambda}^{(3)} = \quad \tilde{\Lambda}^{(4)} = \qquad \uparrow B_1^\dagger$$

$$\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \leftarrow \atop B_2^\dagger$$

$$\blacksquare : V_0 \qquad \blacksquare : V_1$$

Figure 9. A fixed point $\tilde{\Lambda}$ in the negative chamber represented by restricted colored Young diagrams, exemplified in the case of $N = 5$ and $l = 3$ with the coloring $c(0, 1, 2, 3, 4) = (0, 0, 1, 2, 2)$. Each Young diagrams $\tilde{\Lambda}^{(\alpha)}$ is built starting from the box in the lower-right corner, which carries the $\mathbb{Z}_l$-charge $c(\alpha) - 1 \bmod l$. It grows toward the upward by the action of $B_1^\dagger$ and toward the left by the action of $B_2^\dagger$. Note the Young diagrams are growing in the opposite directions compared to the positive chamber. The action of $B_2^\dagger$ decreases the $\mathbb{Z}_l$-charge by 1, and the Young diagrams are restricted in the way that there is no box carrying $\mathbb{Z}_l$-charge $l - 1$. In particular, note that $\tilde{\Lambda}^{(\alpha)} = \varnothing$ if $c(\alpha) = 0$.

$\tilde{\Lambda} = (\tilde{\Lambda}^{(0)}, \cdots, \tilde{\Lambda}^{(N-1)})$ (in fact, $\tilde{\Lambda}^{(\alpha)} = \varnothing$ if $c(\alpha) = 0$ due to the restriction). The equivariant Chern characters of $(V_\omega)_{\omega=0}^{l-2}$ at the fixed point $\tilde{\Lambda}$ are thus given by

$$V_\omega[\tilde{\Lambda}] = \frac{1}{P_1^*} \sum_{\omega < c(\alpha) \le l-1} e^{A_\alpha - \varepsilon_1} \left(1 - q_1^{-\tilde{\Lambda}_{c(\alpha)-\omega}^{(\alpha)t}} \right), \qquad \omega = 0, 1, \cdots, l-2. \tag{4.38}$$

which is in fact a pure character since the term in the parenthesis is divisible by $P_1^* = 1 - q_1^{-1}$. Note that

$$\Sigma_\omega[\tilde{\Lambda}] = \sum_{\omega < \omega' \le l-1} L_{\omega'} + P_1 V_\omega = \sum_{\omega < c(\alpha) \le l-1} e^{A_\alpha} q_1^{-\tilde{\Lambda}_{c(\alpha)-\omega}^{(\alpha)t}}. \tag{4.39}$$

See figure 9 for an example.

The integral of the equivariant Euler class of the matter bundle over the modified handsaw quiver variety in this chamber is computed by the same formula (4.34), where the fixed point set is replaced by $\{\tilde{\Lambda}\}$. Namely,

$$\Psi[\boldsymbol{\lambda}]_{\zeta<0}^{\text{non-pert}} = \sum_{\tilde{\Lambda}} \prod_{\omega=0}^{l-2} \mathfrak{q}_\omega^{v_\omega} \mathbb{E}\left[P_1 W V_0^* + \sum_{\omega=0}^{l-2} \left(L_\omega V_\omega^* + q_1 V_\omega L_{\omega+1}^* - P_1 V_\omega (V_\omega - V_{\omega+1})^* \right. \right.$$
$$\left. \left. - M_\omega^+ V_\omega^* - (M_\omega^-)^* V_\omega \right) \right]. \tag{4.40}$$

This expression is obviously different from (4.34), because the fixed point sets, and therefore the equivariant Chern characters of vector bundles, are distinct. However, we will now show an equality relating the two partition functions up to a simple multiplicative factor.

To establish the equality, let us first introduce the contour integral presentation for the monodromy defect observable. Let us express the monodromy defect observable in the

positive chamber as

$$
\Psi[\boldsymbol{\lambda}]_{\zeta>0}^{\text{non-pert}} = \sum_{v_0,\dots,v_{l-2}=0}^{\infty} \prod_{\omega=0}^{l-2} \frac{\mathfrak{q}_\omega^{v_\omega}}{v_\omega!} \oint_{\mathcal{C}_\omega} \prod_{i=1}^{v_\omega} \frac{d\phi_i^{(\omega)}}{2\pi\mathrm{i}} \frac{1}{\varepsilon_1} \prod_{j<i} \frac{(\phi_i^{(\omega)}-\phi_j^{(\omega)})^2}{(\phi_i^{(\omega)}-\phi_j^{(\omega)})^2-\varepsilon_1^2} \prod_{j=1}^{v_{\omega+1}} \frac{\phi_i^{(\omega)}-\phi_j^{(\omega+1)}+\varepsilon_1}{\phi_i^{(\omega)}-\phi_j^{(\omega+1)}}
$$

$$
\times \frac{\prod_{\alpha\in c_f^{-1}(\omega)}(m_\alpha^+-\phi_i^{(\omega)})\prod_{\alpha\in c_{af}^{-1}(\omega)}(\phi_i^{(\omega)}-m_\alpha^-)}{\prod_{\alpha\in c^{-1}(\omega)}(A_\alpha-\phi_i^{(\omega)})\prod_{\alpha\in c^{-1}(\omega+1)}(\phi_i^{(\omega)}-A_\alpha+\varepsilon_1)} \times \prod_{s=1}^{w} \frac{\phi_i^{(0)}-w_s-\varepsilon_1}{\phi_i^{(0)}-w_s},
$$

(4.41)

where we used the notation $W[\boldsymbol{\lambda}] = \sum_{s=1}^{w} e^{w_s}$. The contours $\mathcal{C}_\omega$ are chosen precisely to reproduce (4.34), as we explain now. The integration variables $\phi_i^{(\omega)}$ pick up poles at

1. A_α where $\alpha \in c^{-1}(\omega)$,

2. $\phi_{i'}^{(\omega)} + \varepsilon_1$ of another i', or

3. $\phi_j^{(\omega-1)}$ for some j.

4. If $\omega = 0$, $\phi_i^{(0)}$ can also pick up poles at w_s.

We start by choosing poles for $\phi_i^{(0)}$, $i = 1, 2, \cdots, v_0$, and move on to $\phi_i^{(\omega+1)}$ only after we exhaust all $\phi_i^{(\omega)}$, $i = 1, 2, \cdots, v_\omega$. The combination of 1 and 2 and the combination of 2 and 3 with $\phi_j^{(\omega-1)} = e^{A_l} q_1^{k'-1}$ generate a tail T_ω. Also, the combination of 3 and 4, where $\phi_i^{(\omega)} = w_s$ for some s, spans a jump J_ω. By taking all possible such combinations, we recover all the fixed points of the modified hand-saw quiver in the positive chamber enumerated by $\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})$ as the poles picked up. Moreover, the contour is chosen not to enclose poles elsewhere, and the rest of the integrand provides the correct factors to reproduce the monodromy defect observable (4.34) by the residue. For instance, $\prod_{\alpha\in c^{-1}(\omega+1)}(\phi_i^{(\omega)} - A_\alpha + \varepsilon_1)^{-1}$ term gives $\mathbb{E}[q_1 V_\omega L_{\omega+1}^*]$, etc.

Now we explain that the contour integral (4.41) can be evaluated in a different way, by picking up the exact complement set of the poles including the one at infinity. Namely,

1. We start from $\omega = l - 2$.

2. We split v_ω integration variables $(\phi_i^{(\omega)})_{i=1}^{v_\omega}$ into two groups, where the first group consists of r_ω and the second group consists of $v_\omega - r_\omega$ of them, for some $0 \le r_\omega \le v_\omega$. The second group takes the pole at infinity, while the first group takes poles elsewhere.

3. The residue at infinity from the integration of the second group is computed as $\frac{(-1)^{v_\omega-r_\omega}}{(v_\omega-r_\omega)!} \prod_{k=1}^{v_\omega-r_\omega} (R_\infty^{(\omega)}+v_{\omega-1}-r_{\omega+1}-k+1)$, where $R_\infty^{(\omega)} = \frac{1}{\varepsilon_1}(m_\omega^+ + m_\omega^- - \sum_{\alpha\in c^{-1}(\omega)} A_\alpha - \sum_{\alpha\in c^{-1}(\omega+1)} A_\alpha + \varepsilon_1)$. Here, we set $v_{-1} = w$ and $r_{l-1} = 0$.

4. Change the order of summation by $\sum_{v_\omega=0}^{\infty} \sum_{r_\omega=0}^{v_\omega} = \sum_{r_\omega=0}^{\infty} \sum_{v_\omega=r_\omega}^{\infty}$. The second summation gives $c_\omega^{R_\infty^{(\omega)}+v_{\omega-1}+r_{\omega+1}}$, where c_ω can be recursively determined by $c_{\omega-1} = 1 + \mathfrak{q}_\omega c_\omega$ and $c_{l-2} = 1 + \mathfrak{q}_{l-2}$. Namely,

$$
c_\omega = 1 + \mathfrak{q}_\omega + \mathfrak{q}_\omega\mathfrak{q}_{\omega+1} + \cdots + \mathfrak{q}_\omega\mathfrak{q}_{\omega+1}\cdots\mathfrak{q}_{l-2}. \tag{4.42}
$$

We leave the integration of the remaining r_ω variables in the first group and the summation over r_ω undone and proceed.

5. Repeat the steps 2 to 4 with $\omega \to \omega - 1$, up to $\omega = 0$. Then, we are left with the integration of r_ω variables for each $\omega = 0, 1, \cdots, l-2$, and the summation over r_ω. As a result of step 4, we get the prefactor $c_0^w \prod_{\omega=0}^{l-2} c_\omega^{R_\infty^{(\omega)}}$, and also the mutation of the complexified FI parameters $\mathfrak{q}_\omega \to \mathfrak{q}_\omega \frac{c_{\omega+1}}{c_{\omega-1}}$, $\omega = 0, 1, \cdots, l-2$ appearing in the remaining integrals, where we set $c_{-1} = c_{l-1} = 1$.

6. The integration of the remaining r_ω variables is performed by taking the residues at $A_\alpha - \varepsilon_1 - (i-1)\varepsilon_1$, $i \geq 0$, where $\omega < c(\alpha) \leq l-1$. These poles gathered from all $\omega = 0, 1, \cdots, l-2$ compose a restricted Young diagram $\tilde{\boldsymbol{\Lambda}}$. The residue is precisely the contribution of $\tilde{\boldsymbol{\Lambda}}$ to the equivariant integral (4.40) in the negative chamber, with the mutated $\mathfrak{q}_\omega$.

7. Sum over all $r_\omega \geq 0$. Then we recover (4.40) with the mutated $\mathfrak{q}_\omega$ and the prefactor explained in step 5.

All in all, we achieve the following crossing formula relating the monodromy observable in the positive and the negative chamber:

$$\Psi(\mathfrak{q}_\omega)[\boldsymbol{\lambda}]_{\zeta>0}^{\text{non-pert}} = c_0^{\frac{1}{\varepsilon_1} \sum_{\alpha=0}^{N-1}(A_\alpha - a_\alpha)} \prod_{\omega=0}^{l-2} c_\omega^{R_\infty^{(\omega)}} \times \Psi\left(\mathfrak{q}_\omega \frac{c_{\omega+1}}{c_{\omega-1}}\right) [\boldsymbol{\lambda}]_{\zeta<0}^{\text{non-pert}}. \qquad (4.43)$$

We remind that the classical contribution (4.35) of the sigma model should further be augmented to reproduce the original monodromy surface observable (4.14).

This crossing formula is useful since the expression of the monodromy defect observable is simplified in the negative chamber. Furthermore, the crossing formula gives a direct exhibition of transition of the monodromy surface defect into (multiple) $\mathbf{Q}$-observables or $\mathbf{H}$-observables, as we investigate from now on.

4.2 Transition between $\mathbb{Z}_2$-monodromy defect and H-observable

Before studying the transition that leads to the separation of variables for the Gaudin model and the XXX spin chain, we will apply the crossing formula (4.43) to show dualities among the monodromy surface defects defined by $\mathbb{Z}_2$-orbifold and the $\mathbf{H}$-observables.

4.2.1 Monodromy surface defect dual to H-observable

Let us consider the monodromy defect defined by the $\mathbb{Z}_2$-orbifold, where the coloring function is given by

$$c(\alpha) = \begin{cases} 1 & \alpha = \beta, \\ 0 & \text{otherwise,} \end{cases} \qquad (4.44)$$

for a chosen $\beta \in \{0, 1, \cdots, N-1\}$. In other words, we have

$$L_0 = \sum_{0 \leq \alpha \neq \beta \leq N-1} e^{A_\alpha}, \qquad L_1 = e^{A_\beta}. \qquad (4.45)$$

We also assign the coloring for the flavor as

$$M_0^+ = \sum_{\alpha=0}^{N-1} e^{m_\alpha^+}, \quad M_1^+ = 0$$

$$M_0^- = 0 \qquad\qquad M_1^- = \sum_{\alpha=0}^{N-1} e^{m_\alpha^-}. \tag{4.46}$$

In this case, the Higgs branch of the two-dimensional $\mathcal{N} = (2,2)$ gauged sigma model is the total space of

$$\mathcal{E} \otimes \mathbb{C}^N \longrightarrow F(N-1, N), \tag{4.47}$$

where $\mathcal{E}$ is the tautological vector bundle over the flag variety $F(N-1, N)$. This is dual to the total space of

$$\mathcal{O}(-1) \otimes \mathbb{C}^N \longrightarrow \mathbb{P}^{N-1}. \tag{4.48}$$

This is precisely the gauged linear sigma model that we discussed in section 2. It is in the non-linear sigma model phase here by turning on the complexified FI parameter, rather than the vev of the adjoint scalar, which produces the **H**-observable by coupling to the $\mathcal{N} = 2$ A_1-quiver gauge theory [14]. Therefore, we expect that the monodromy surface defect defined by the $\mathbb{Z}_2$-orbifold is dual to the **H**-observable, where the prescribed singularity in the former can be thought of as a remnant of integrating out two-dimensional degrees of freedom in the latter. We will now explain that this duality can indeed be confirmed by the equality between their observable expressions, as a consequence of the crossing formula derived in (4.43).

The restricted Young diagrams enumerating the fixed points in the negative chamber are Young diagrams with single column of length, say, $s \geq 0$ (see figure 9). These Young diagrams represent the V_0 space generated by the action of $B_1^\dagger$ on the image of the one-dimensional space L_1 under $j^\dagger$, $V_0 = \mathbb{C}[B_1^\dagger]j^\dagger(L_1)$. Its equivariant Chern character is therefore

$$V_0 = \sum_{i=1}^{s} e^{A_\beta} q_1^{-i}, \tag{4.49}$$

so that

$$\Sigma_0 = L_1 + P_1 V_0 = e^{A_\beta} q_1^{-s}. \tag{4.50}$$

The monodromy defect observable in the negative chamber is computed as

$$\begin{aligned}
\Psi_\beta[\boldsymbol{\lambda}]_{\zeta<0} &= \sum_{s=0}^{\infty} \mathfrak{q}_0^s \, \mathbb{E}\left[\frac{(S - \Sigma_0 - M_0^+)\Sigma_0^*}{P_1^*}\right] \\
&= \frac{\prod_\alpha \sin\left(\frac{a_\beta - m_\alpha^+}{\varepsilon_1}\right)}{\pi \prod_{\alpha\neq\beta} \sin\left(\frac{a_\beta - a_\alpha}{\varepsilon_1}\right)} \times \sum_{s=0}^{\infty} \mathfrak{q}_0^s \prod_{\alpha=0}^{N-1} \Gamma\left(s + 1 - \frac{A_\beta - m_\alpha^+}{\varepsilon_1}\right) Q(A_\beta - (s+1)) \\
&= \frac{\prod_\alpha \sin\left(\frac{a_\beta - m_\alpha^+}{\varepsilon_1}\right)}{\pi \prod_{\alpha\neq\beta} \sin\left(\frac{a_\beta - a_\alpha}{\varepsilon_1}\right)} \mathfrak{q}_0^{\frac{A_\beta}{\varepsilon_1} - 1} \mathbf{H}^{(\beta)}(\mathfrak{q}_0)
\end{aligned} \tag{4.51}$$

Here, $Q(x)[\boldsymbol{\lambda}] = \mathbb{E}\left[-\frac{e^x S^*[\boldsymbol{\lambda}]}{P_1^*}\right]$ is the **Q**-observable and $\mathbf{H}^{(\beta)}(\mathfrak{q}_0)$ is precisely the **H**-observable observable with the complexified FI parameter $\mathfrak{q}_0$, where $\beta \in \{0, 1, \cdots, N-1\}$ enumerates the choice of vacuum at infinity among N discrete vacua. Recall that the monodromy defect observable is convergent in the domain $0 < |\mathfrak{q}_0|, |\mathfrak{q}_1| < 1$, or equivalently, $|\mathfrak{q}| < |\mathfrak{q}_0| < 1$. This is precisely the domain where the series expansion of the emergent **H**-observable converges. The **H**-observable can in fact be analytically continued to other domains, $0 < |\mathfrak{q}_0| < |\mathfrak{q}|$ and $1 < |\mathfrak{q}_0|$, by using its contour integral representation [14]. We will not discuss this analytic continuation here, but remark that the equality (4.51) will still hold across different convergence domains when the monodromy surface defect observable is also properly analytically continued.

The crossing formula (4.43) simplifies in this case ($l = 2$) since the mutation of the complexified FI parameters is absent. It gives

$$\begin{aligned}
\Psi_\beta(\mathfrak{q}_0)[\boldsymbol{\lambda}] &= \mathfrak{q}_0^{-l(\lambda^{(\beta)})}(1 + \mathfrak{q}_0)^{1 + \frac{1}{\varepsilon_1}\sum_{\alpha=0}^{N-1}(m_\alpha^+ - a_\alpha)}\Psi[\boldsymbol{\lambda}]_{\zeta < 0} \\
&= \frac{\prod_\alpha \sin\left(\frac{a_\beta - m_\alpha^+}{\varepsilon_1}\right)}{\pi \prod_{\alpha \neq \beta} \sin\left(\frac{a_\beta - a_\alpha}{\varepsilon_1}\right)}\mathfrak{q}_0^{\frac{a_\beta}{\varepsilon_1} - 1}(1 + \mathfrak{q}_0)^{1 + \frac{1}{\varepsilon_1}\sum_{\alpha=0}^{N-1}(m_\alpha^+ - a_\alpha)}\mathbf{H}^{(\beta)}(\mathfrak{q}_0)[\boldsymbol{\lambda}].
\end{aligned} \tag{4.52}$$

The equality between the observable expressions manifests the duality between the monodromy surface defect defined by the $\mathbb{Z}_2$-orbifold with the coloring functions (4.46) and the **H**-observable. The identity was derived and used for the cases of $N = 2$ and 3 in [14], extended to arbitrary N here.

4.2.2 Transition between $\mathbb{Z}_2$-monodromy defects and H-observable

We emphasize here that the $\mathbb{Z}_2$-monodromy defect defined by the coloring in (4.45) and (4.46) is different from the regular $\mathbb{Z}_N$-monodromy defect, even in the case of $N = 2$ where the two defects can be compared on equal footing. The distinction comes from the flavor coloring; the flavor coloring (4.46) of the $\mathbb{Z}_2$-monodromy defect sets all the fundamental (resp. anti-fundamental) hypermultiplets to carry charge 0 (resp. charge 1), while the flavor coloring of the regular monodromy defect sets exactly one fundamental hypermultiplet and one anti-fundamental hypermultiplet to carry charge ω for each $\omega \in \{0, 1, \cdots, N-1\}$.

Let us recall that the flavor coloring determines the embedding of the flavor group of the two-dimensional sigma model into the flavor group of the four-dimensional bulk theory. In the case of $N = 2$, where the complexified flavor group of the four-dimensional gauge theory is GL(4) by treating all four hypermultiplets as fundamental, the two defects correspond to the same two-dimensional sigma model whose GL(2) flavor group is identified with GL(2) subgroup appearing in two different decompositions of GL(4) into GL(2) × GL(2). Here, we show that such a discrete group action on the flavor group embedding is reflected in their observable expressions as an integral transformation between them.

Let us consider the regular monodromy defect at $N = 2$. The coloring for the Coulomb moduli is chosen as the same as in the case of the $\mathbb{Z}_2$-monodromy defect (4.45) at $N = 2$ with $\beta = 1$,

$$L_0 = e^{A_0}, \qquad L_1 = e^{A_1}, \tag{4.53}$$

while the coloring for the flavor is chosen in a different manner as

$$
\begin{aligned}
M_0^+ = e^{m_0^+}, \quad M_1^+ = e^{m_1^+} \\
M_0^- = e^{m_0^-}, \quad M_1^- = e^{m_1^-}.
\end{aligned}
\tag{4.54}
$$

Compared to the $N = 2$ case of (4.46), the $\mathbb{Z}_2$-charge assignment for the masses got shuffled according to a discrete group action. The vacuum expectation value of the regular monodromy defect in $N = 2$ can again be expressed using the $\mathbf{Q}$-observable as

$$
\begin{aligned}
\Psi[\boldsymbol{\lambda}] = {}& \mathfrak{q}_0^{\frac{a_1-A_1}{\varepsilon_1}} (1 + \mathfrak{q}_0)^{1+\frac{1}{\varepsilon_1}(m_0^+ + m_0^- - a_0 - a_1)} \\
& \times \sum_{s=0}^{\infty} \mathfrak{q}_0^s \frac{(-\varepsilon_1)^{1+2s+\frac{m_0^+ + m_0^- - 2A_1}{\varepsilon_1}} \Gamma\left(1 + s + \frac{m_0^- - A_1}{\varepsilon_1}\right)}{\Gamma\left(-s + \frac{A_1 - m_0^+}{\varepsilon_1}\right)} \frac{\pi}{\sin\pi\frac{A_1 - A_0}{\varepsilon_1}} Q(A_1 - (s+1)\varepsilon_1) \\
= {}& -\frac{(-\varepsilon_1)^{-1+\frac{m_0^- - m_1^+}{\varepsilon_1}} e^{\pi i \frac{a_1 - m_1^+}{\varepsilon_1}} \sin\pi\frac{a_1 - m_0^+}{\varepsilon_1}}{4\sin\pi\frac{a_1 - A_0}{\varepsilon_1}\sin\pi\frac{m_0^- - a_1}{\varepsilon_1}\Gamma\left(\frac{m_1^+ - m_0^-}{\varepsilon_1}\right)} \times \mathfrak{q}_0^{\frac{a_1}{\varepsilon_1}-1}(1+\mathfrak{q}_0)^{1+\frac{1}{\varepsilon_1}(m_0^+ + m_0^- - a_0 - a_1)} \\
& \times \sum_{s=0}^{\infty} \mathfrak{q}_0^{-\frac{A_1}{\varepsilon_1}+s+1} Q(A_1 - (s+1)\varepsilon_1) \prod_{\alpha=0,1} (-\varepsilon_1)^{1+s-\frac{A_1 - m_\alpha^+}{\varepsilon_1}} \Gamma\left(1 + s - \frac{A_1 - m_\alpha^+}{\varepsilon_1}\right) \\
& \times \int_{\mathrm{PC}} dt\, t^{s+\frac{m_0^- - A_1}{\varepsilon_1}} (1-t)^{-1+\frac{m_1^+ - m_0^-}{\varepsilon_1}},
\end{aligned}
\tag{4.55}
$$

where we used the integral representation of the beta function,

$$
B(a, b-a) = \frac{\Gamma(a)\Gamma(b-a)}{\Gamma(b)} = -\frac{e^{-\pi i b}}{4\sin\pi a \sin\pi(b-a)} \int_{\mathrm{PC}} dt\, t^{a-1}(1-t)^{b-a-1}.
\tag{4.56}
$$

Here, PC denotes the Pochhammer contour on the t-plane enclosing $t = 0$ and $t = 1$.

Now, we change the order of summation and integration in the last line. The summation converges in the domain $|\mathfrak{q}| < |\mathfrak{q}_0 t| < 1$ and gives the $\mathbf{H}$-observable. Some part of the Pochhammer contour deviates from this domain, where the integrand can be properly analytically continued. As a result, after absorbing irrelevant functions by redefinition of the observable, we have

$$
\begin{aligned}
\Psi(\mathfrak{q}_0)[\boldsymbol{\lambda}] = {}& \mathfrak{q}_0^{1+\frac{a_1 - m_1^+}{\varepsilon_1}} (1 + \mathfrak{q}_0)^{1+\frac{1}{\varepsilon_1}(m_0^+ + m_0^- - a_0 - a_1)} \\
& \times \int_{\mathrm{PC}} dy\, y^{-\frac{1}{2}+\frac{m_0^- - m_1^-}{2\varepsilon_1}} (y - \mathfrak{q}_0)^{-1+\frac{m_1^+ - m_0^-}{\varepsilon_1}} (y - \mathfrak{q})^{\frac{\bar{m}^- - \bar{a}}{\varepsilon_1}} (y - 1)^{-1-\frac{\bar{m}^+ - \bar{a}}{\varepsilon_1}} \mathbf{H}^{(1)}(y)[\boldsymbol{\lambda}],
\end{aligned}
\tag{4.57}
$$

where PC is now the Pochhammer contour in the y-plane enclosing $y = 0$ and $y = \mathfrak{q}_0$. Note that even though the integrand has monodromies around $y = 0$ and $y = \mathfrak{q}_0$, it is single-valued along the Pochhammer contour and therefore it is well-defined.

Note that the integral transformation has been written for observable expressions at each $\boldsymbol{\lambda}$. Thus, it describes the transition of the regular $\mathbb{Z}_2$-monodromy surface defect into a $\mathbf{H}$-observable. Upon taking the expectation value, we get

$$\left\langle \Psi(\mathfrak{q}_0) \right\rangle_{\mathbf{a}} = \int_{\mathrm{PC}} dy\, \mathcal{K}(\mathfrak{q}_0; y) \left\langle \mathbf{H}^{(1)}(y) \right\rangle_{\mathbf{a}}, \tag{4.58}$$

where the integral kernel $\mathcal{K}(\mathfrak{q}_0; y)$ is given by

$$\begin{aligned}
\mathcal{K}(\mathfrak{q}_0; y) = {}& \mathfrak{q}_0^{1 + \frac{a_1 - m_1^+}{\varepsilon_1}} (1 + \mathfrak{q}_0)^{1 + \frac{1}{\varepsilon_1}(m_0^+ + m_0^- - a_0 - a_1)} \\
& \times y^{-\frac{1}{2} + \frac{m_0^- - m_1^-}{2\varepsilon_1}} (y - \mathfrak{q}_0)^{-1 + \frac{m_1^+ - m_0^-}{\varepsilon_1}} (y - \mathfrak{q})^{\frac{\bar{m}^- - \bar{a}}{\varepsilon_1}} (y - 1)^{-1 - \frac{\bar{m}^+ - \bar{a}}{\varepsilon_1}}.
\end{aligned} \tag{4.59}$$

Note that the relation is valid for any value of $\varepsilon_1, \varepsilon_2$. The results of [8, 43, 68] (at $N = 2$) identifies the left hand side to a twisted coinvariant of four $\widehat{\mathfrak{sl}}(2)$-modules (two (co-)Verma modules and two bi-infinite modules). Also, the results of [14, 53] (at $N = 2$) show that the vacuum expectation value $\langle \mathbf{H}(y) \rangle$ of the $\mathbf{H}$-observable is identified with a Virasoro conformal block, with four generic vertex operators and one simplest degenerate operator (see also [45, 77]). Thus, the above equation can be interpreted as a transformation between conformal blocks of two different vertex algebras. Differently put, it reformulates the simplest example of the KZ/BPZ correspondence [37, 46, 47] as a surface defect transition in the $\mathcal{N} = 2$ gauge theory. Note that the integral kernel (4.59) is guaranteed to match with the ones in [46, 51], after reducing the latter to the invariants under the action of global $\mathfrak{sl}(2)$ on $\widehat{\mathfrak{sl}}(2)$-modules and global $\mathrm{SL}(2)$ on $\mathbb{P}^1$ with four marked points, due to the KZ equation solved by the left hand side and the BPZ equation solved by the integrand on the right hand side.

4.3 Separation of variables for XXX spin chain

Finally, let us apply the crossing formula (4.43) to the regular monodromy surface defect defined by the $\mathbb{Z}_N$-orbifold. We show that the crossing formula manifests a transition to $N - 1$ $\mathbf{Q}$-observables or $N - 1$ $\mathbf{H}$-observables, which leads to the separation of variable transformations for the XXX spin chain and the Gaudin model, respectively.

4.3.1 Separated expression for eigenstates

Let us recall that, in the negative chamber, the fixed points of the modified hand-saw quiver variety is enumerated by the collections $\{\tilde{\boldsymbol{\Lambda}}\}$ of restricted Young diagrams. The vector spaces $(V_\omega)_{\omega=0}^{N-2}$ are generated as in (4.36),

$$\begin{aligned}
V_{N-2} &= \mathbb{C}[B_1^\dagger] j^\dagger(L_{N-1}) \\
V_\omega &= \mathbb{C}[B_1^\dagger] B_2^\dagger(V_{\omega+1}) \oplus \mathbb{C}[B_1^\dagger] j^\dagger(L_{\omega+1}), \qquad \omega = 0, 1, \cdots, N - 3,
\end{aligned} \tag{4.60}$$

where each $(L_\omega)_{\omega=0}^{N-1}$ is one-dimensional in this case. The columns in the restricted Young diagram $\tilde{\boldsymbol{\Lambda}}$ represent the basis vectors in V_ω obtained by repetitive $B_1^\dagger$-action on $j^\dagger(L_{\omega+1})$ or $B_2^\dagger(V_{\omega+1})$. We can parametrize the restricted Young diagrams by introducing $\frac{N(N-1)}{2}$ non-negative integers $\mathbf{s} = (s_{n,j})_{1 \le j \le n \le N-1}$ for the differences of lengths of the neighboring columns (see figure 10).
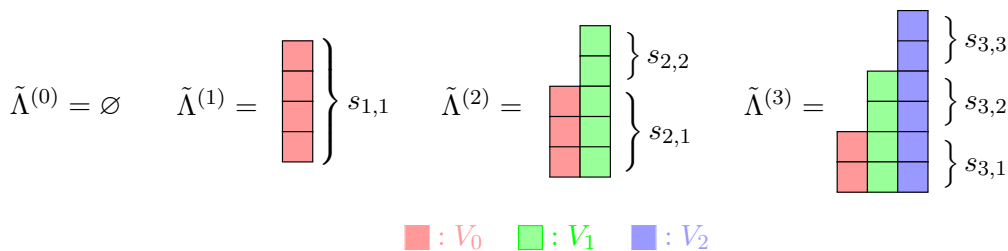
Figure 10. A collection $\tilde{\mathbf{\Lambda}}$ of restricted Young diagrams for the regular monodromy defect observable in the negative chamber, exemplified in the case of $N = 4$. Note that $\tilde{\Lambda}^{(\omega)}$ is composed of ω columns each of which carries different representations of $\mathbb{Z}_N$. The restricted Young diagrams in this case can be parametrized by the differences of lengths of these columns (in the example, $\frac{4 \times 3}{2} = 6$ non-negative integers $(s_{1,1}, s_{2,1}, s_{2,2}, s_{3,1}, s_{3,2}, s_{3,3})$).

Then the equivariant Chern characters of V_ω can be written in terms of $\mathbf{s}$ by

$$V_\omega = \sum_{n=\omega+1}^{N-1} \sum_{i=1}^{\sum_{j=1}^{\omega+1} s_{n,j}} e^{A_n} q_1^{-i}, \qquad \omega = 0, 1, \cdots, N-2. \tag{4.61}$$

This implies

$$
\begin{aligned}
\Sigma_\omega &= \sum_{n=\omega+1}^{N-1} L_n + P_1 V_\omega = \sum_{n=\omega+1}^{N-1} e^{A_n} q_1^{-\sum_{j=1}^{\omega+1} s_{n,j}} \\
&=: \sum_{i=\omega+1}^{N-1} e^{x_{\omega+1,i} + \varepsilon_1}
\end{aligned}
\qquad \omega = 0, 1, \cdots, N-2, \tag{4.62}
$$

where we defined

$$x_{n,i} := A_i - \left(1 + \sum_{j=1}^{n} s_{i,j}\right) \varepsilon_1, \qquad 1 \leq n \leq i \leq N-1. \tag{4.63}$$

Note that we can write this in a recursive way as $x_{n,i} = x_{n-1,i} - s_{i,n} \varepsilon_1$, with $2 \leq n \leq i \leq N-1$, and $x_{1,i} = A_i - (1 + s_{i,1}) \varepsilon_1$ for $i = 1, 2, \cdots, N-1$. The monodromy defect observable in the negative chamber is expressed as

$$
\begin{aligned}
&\Psi(\mathbf{u})[\boldsymbol{\lambda}]_{\zeta<0} \\
&= \sum_{\mathbf{s}=0}^{\infty} \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{\sum_{n=\omega+1}^{N-1} \sum_{j=1}^{\omega+1} s_{n,j}} \mathbb{E}\left[\frac{S\Sigma_0^* + \sum_{\omega=0}^{N-2}\left(-\Sigma_\omega(\Sigma_\omega - \Sigma_{\omega+1})^* - M_\omega^+ \Sigma_\omega^* + q_1^{-1}(M_\omega^-)^* \Sigma_\omega\right)}{P_1^*} \right] \\
&= \sum_{\mathbf{s}=0}^{\infty} \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{\sum_{i=\omega+1}^{N-1}\left(-1+\frac{A_i - x_{\omega+1,i}}{\varepsilon_1}\right)} \frac{\prod_{\substack{\omega+1 \leq i \leq N-1 \\ \omega+2 \leq j \neq i \leq N-1}} (-\varepsilon_1)^{\frac{x_{\omega+2,j} - x_{\omega+1,i}}{\varepsilon_1}} \Gamma\left(\frac{x_{\omega+2,j} - x_{\omega+1,i}}{\varepsilon_1}\right)}{\prod_{\omega+2 \leq i \leq N-1} s_{i,\omega+2}! \varepsilon_1^{s_{i,\omega+2}} \prod_{\omega+1 \leq i \neq j \leq N-1} \Gamma\left(\frac{x_{\omega+1,i} - x_{\omega+1,j}}{\varepsilon_1}\right)} \\
&\quad \times \prod_{1 \leq n \leq i \leq N-1} \frac{(-\varepsilon_1)^{\frac{m_{n-1}^- - x_{n,i}}{\varepsilon_1}} \Gamma\left(\frac{m_{n-1}^- - x_{n,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{n,i} - m_{n-1}^+}{\varepsilon_1}} \Gamma\left(1 + \frac{x_{n,i} - m_{n-1}^+}{\varepsilon_1}\right)} \times \prod_{i=1}^{N-1} (-1)^{s_{i,1}+1} Q(x_{1,i})[\boldsymbol{\lambda}] \prod_{\substack{\alpha=0 \\ \alpha \neq i}}^{N-1} \frac{\pi}{\sin \pi \frac{x_{1,i} - A_\alpha}{\varepsilon_1}}.
\end{aligned}
\tag{4.64}
$$

We can express the summation over non-negative integers $\mathbf{s}$ as a $\frac{N(N-1)}{2}$-dimensional contour integral, which reproduces the summation over $\mathbf{s}$ by picking up the residues at simple poles. For this, we promote $\mathbf{x} = (x_{n,i})_{1 \leq n \leq i \leq N-1}$ to integration variables, and write

$$\Psi[\boldsymbol{\lambda}]_{\zeta<0} = \int_{\mathcal{C}} \prod_{i=1}^{N-1} \frac{dx_{1,i}}{2\pi i \varepsilon_1} \, \mathbf{V}(\mathbf{x}_1) \mu(\mathbf{x}_1) C(\mathbf{x}_1), \tag{4.65}$$

where $\mu(\mathbf{x}_1)$ and $C(\mathbf{x}_1)$ only depends on $\mathbf{x}_1 = (x_{1,i})_{i=1}^{N-1}$, and $\mathbf{V}(\mathbf{x}_1)$ is a $\frac{(N-1)(N-2)}{2}$-dimensional contour integral of remaining variables $(x_{n,i})_{2 \leq n \leq i \leq N-1}$. The contour $\mathcal{C}$ is chosen precisely in the way that $\mathbf{x}$ picks up the simple poles at (4.63), as we now explain.

Firstly, $\mathbf{V}(\mathbf{x}_1)$ is a $\frac{(N-1)(N-2)}{2}$-dimensional contour integral given by

$$\begin{aligned}
\mathbf{V}(\mathbf{x}_1) = \int_{\mathcal{C}} \prod_{\omega=1}^{N-2} \frac{1}{(N-\omega-1)!} \prod_{i=\omega+1}^{N-1} \frac{dx_{\omega+1,i}}{2\pi\varepsilon_1 i} \mathfrak{q}_\omega^{\sum\limits_{i=\omega+1}^{N-1}\left(-1+\frac{A_i - x_{\omega+1,i}}{\varepsilon_1}\right)} \\
\frac{\prod\limits_{\substack{\omega \leq i \leq N-1 \\ \omega+1 \leq j \leq N-1}} (-\varepsilon_1)^{\frac{x_{\omega+1,j}-x_{\omega,i}}{\varepsilon_1}} \Gamma\left(\frac{x_{\omega+1,j}-x_{\omega,i}}{\varepsilon_1}\right)}{\prod_{\omega+1 \leq i \neq j \leq N-1} \Gamma\left(\frac{x_{\omega+1,i}-x_{\omega+1,j}}{\varepsilon_1}\right)} \\
\times \prod_{2 \leq n \leq i \leq N-1} \frac{(-\varepsilon_1)^{\frac{m_{n-1}^{-}-x_{n,i}}{\varepsilon_1}} \Gamma\left(\frac{m_{n-1}^{-}-x_{n,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{n,i}-m_{n-1}^{+}}{\varepsilon_1}} \Gamma\left(1+\frac{x_{n,i}-m_{n-1}^{+}}{\varepsilon_1}\right)},
\end{aligned} \tag{4.66}$$

where the contour $\mathcal{C}$ picks up the simple poles of the Γ-functions in the numerator in the following way.

1. Start with $x_{N-1,N-1}$. Due to the Γ-functions in the numerator, there are two semi-infinite lattices of simple poles in the $\frac{x_{N-1,N-1}}{\varepsilon_1}$-plane starting from $\frac{x_{N-2,N-1}}{\varepsilon_1}$ and $\frac{x_{N-2,N-2}}{\varepsilon_1}$, growing toward the left. We choose the contour that encloses all the poles at $x_{N-1,N-1} = x_{N-2,N-1} - s_{N-1,N-1}\varepsilon_1$ and $x_{N-1,N-1} = x_{N-2,N-2} - s_{N-1,N-1}\varepsilon_1$ with $s_{N-1,N-1} \geq 0$. The residue is given by $\frac{(-1)^{s_{N-1,N-1}}}{s_{N-1,N-1}! \varepsilon_1^{s_{N-1,N-1}}}$.

2. Proceed recursively. For the integral of $x_{\omega+1,j}$, the relevant Γ-functions in the numerator are $\Gamma\left(\frac{x_{\omega+1,j}-x_{\omega,i}}{\varepsilon_1}\right)$ for $1 \leq \omega < i \leq N-1$, $\Gamma\left(\frac{x_{\omega+2,i}-x_{\omega+1,j}}{\varepsilon_1}\right)$ for $3 \leq \omega+2 < i \leq N-1$, and finally $\Gamma\left(\frac{m_{\omega}^{-}-x_{\omega+1,j}}{\varepsilon_1}\right)$ for $3 \leq \omega+1 \leq i \leq N-1$. Correspondingly, there are three groups of semi-infinite lattices of simple poles. We choose the contour that only encloses the poles at $x_{\omega+1,j} = x_{\omega,i} - s_{j,\omega+2}\varepsilon_1$ with $s_{j,\omega+2} \geq 0$. The residue is given by $\frac{(-1)^{s_{j,\omega+2}}}{s_{j,\omega+2}! \varepsilon_1^{s_{j,\omega+2}}}$ and $\omega < i \leq N-1$.

3. The duplication made by rearranging $(x_{\omega+1,i})_{\omega+1 \leq i \leq N-1}$ is cancelled by the $(N-\omega-1)!$ in the denominator.

By taking the residues in the described way, $\mathbf{V}(\mathbf{x}_1)$ precisely reproduce (4.64) except the terms involving $\mathbf{x}_1$. To get the rest, we perform the $(N-1)$-dimensional contour integral

for $\mathbf{x}_1 = (x_{1,i})_{i=1}^{N-1}$. Here, $\mu(\mathbf{x}_1)$ is given by

$$\mu(\mathbf{x}_1) = \frac{\mathfrak{q}_0^{\sum_{i=1}^{N-1}\left(-1+\frac{A_i-x_{1,i}}{\varepsilon_1}\right)}}{(N-1)!} \prod_{i=1}^{N-1} \frac{(-\varepsilon_1)^{\frac{m_0^--x_{1,i}}{\varepsilon_1}} \Gamma\left(\frac{m_0^--x_{1,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{1,i}-m_0^+}{\varepsilon_1}} \Gamma\left(1+\frac{x_{1,i}-m_0^+}{\varepsilon_1}\right)}$$

$$\times \frac{1}{\prod_{1\leq j\neq k\leq N-1}\Gamma\left(\frac{x_{1,j}-x_{1,k}}{\varepsilon_1}\right)} \prod_{i=1}^{N-1}\prod_{\alpha=0}^{N-1} \frac{\pi}{\sin\pi\left(\frac{x_{1,i}-a_\alpha}{\varepsilon_1}\right)} \qquad (4.67)$$

Finally, $C(\mathbf{x}_1)$ consists of $\mathbf{Q}$-observables,

$$C(\mathbf{x}_1)[\boldsymbol{\lambda}] = \prod_{i=1}^{N-1} Q(x_{1,i})[\boldsymbol{\lambda}]. \qquad (4.68)$$

The contour for $x_{1,i}$ is chosen in the following way. There are simple poles in the integrand at $x_{1,i} = A_\alpha - (s_{i,1}+1)\varepsilon_1$ with $s_{i,1} \geq 0$ and $\alpha = 0, 1, \cdots, N-1$ from the sine functions in the denominator of $\mu(\mathbf{x}_1)$ (the semi-infinite lattices of poles toward the other direction are cancelled by the zeros of the inverse Γ-functions in the $\mathbf{Q}$-observables in $C(\mathbf{x}_1)$), and at $x_{1,i} = m_0^- + s_{i,1}\varepsilon_1$ with $s_{i,1} \geq 0$ from the Γ-functions in the numerator of $\mu(\mathbf{x}_1)$. The contour is chosen to enclose the poles at $x_{1,i} = A_\alpha - (s_{i,1}+1)\varepsilon_1$ with $s_{i,1} \geq 0$ and $1 \leq \alpha \leq N-1$. The residue is $(-1)^{s_{i,1}+1}$. The duplication of rearranging $(x_{1,i})_{i=1}^{N-1}$ is cancelled by the $(N-1)!$ in the denominator of $\mu(\mathbf{x}_1)$. Performing the $\mathbf{x}_1$-integral in this way, we precisely reproduce the observable expression (4.64) in the negative chamber at last.

To fully recover the observable expression in the original chamber, we need to take account for the crossing formula (4.43) and the classical contribution (4.35). Thus we get

$$\Psi(\mathbf{u})[\boldsymbol{\lambda}] = c_0^{\frac{\sum_{\alpha=0}^{N-1}(A_\alpha-a_\alpha)}{\varepsilon_1}} \prod_{\omega=0}^{N-2} c_\omega^{\frac{m_\omega^++m_\omega^--A_\omega-A_{\omega+1}+\varepsilon_1}{\varepsilon_1}} \prod_{\omega=1}^{N-1}\left(\frac{u_\omega}{u_0}\right)^{\frac{a_\omega-A_\omega}{\varepsilon_1}}$$

$$\times \int_{\mathcal{C}} \prod_{i=1}^{N-1} \frac{dx_{1,i}}{2\pi\mathrm{i}\varepsilon_1} \left(V(\mathbf{x}_1)\mu(\mathbf{x}_1)\right)_{\mathfrak{q}_\omega\to\mathfrak{q}_\omega\frac{c_{\omega+1}}{c_{\omega-1}}} C(\mathbf{x}_1)[\boldsymbol{\lambda}], \qquad (4.69)$$

The $\mathfrak{q}_\omega$ only appears in elementary functions in the integrand $V(\mathbf{x}_1)\mu(\mathbf{x}_1)$. We can recombine the effect of mutation $\mathfrak{q}_\omega \to \mathfrak{q}_\omega\frac{c_{\omega+1}}{c_{\omega-1}}$ with the terms in the first line. Importantly, all the A_ω's in the exponents cancel with each other, yielding

$$\Psi(\mathbf{u})[\boldsymbol{\lambda}] = \int_{\mathcal{C}} \prod_{i=1}^{N-1} \frac{dx_{1,i}}{2\pi\mathrm{i}\varepsilon_1} \mathcal{K}(\mathbf{u};\mathbf{x}_1) C(\mathbf{x}_1)[\boldsymbol{\lambda}], \qquad (4.70)$$

where the integral kernel $\mathcal{K}(\mathbf{u};\mathbf{x}_1) = \mathbf{c}\mathbf{V}'(\mathbf{x}_1)\mu'(\mathbf{x}_1)$ is given by the product of

$$\mathbf{c} = c_0^{N+\frac{x_{0,0}+x_{0,1}-a_0-a_1}{\varepsilon_1}-\sum_{i=2}^{N-1}\frac{a_i-x_{2,i}}{\varepsilon_1}} \prod_{\omega=0}^{N-2} c_\omega^{-1+\frac{m_\omega^++m_\omega^--x_{\omega,\omega}-x_{\omega,\omega+1}}{\varepsilon_1}}, \qquad (4.71)$$

$$
\mathbf{V}'(\mathbf{x}_1) = \int_{\mathcal{C}} \prod_{\omega=1}^{N-2} \frac{1}{(N-\omega-1)!} \prod_{i=\omega+1}^{N-1} \frac{dx_{\omega+1,i}}{2\pi\varepsilon_1 \mathrm{i}} \mathfrak{q}_{\omega}^{\sum\limits_{i=\omega+1}^{N-1}\left(-1+\frac{a_i-x_{\omega+1,i}}{\varepsilon_1}\right)}
$$

$$
\frac{\prod\limits_{\substack{\omega\leq i\leq N-1 \\ \omega+1\leq j\leq N-1}} (-\varepsilon_1)^{\frac{x_{\omega+1,j}-x_{\omega,i}}{\varepsilon_1}} \Gamma\left(\frac{x_{\omega+1,j}-x_{\omega,i}}{\varepsilon_1}\right)}{\prod_{\omega+1\leq i\neq j\leq N-1} \Gamma\left(\frac{x_{\omega+1,i}-x_{\omega+1,j}}{\varepsilon_1}\right)} \tag{4.72}
$$

$$
\times \prod_{2\leq n\leq i\leq N-1} \frac{(-\varepsilon_1)^{\frac{m_{n-1}^- - x_{n,i}}{\varepsilon_1}} \Gamma\left(\frac{m_{n-1}^- - x_{n,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{n,i}-m_{n-1}^+}{\varepsilon_1}} \Gamma\left(1+\frac{x_{n,i}-m_{n-1}^+}{\varepsilon_1}\right)},
$$

and

$$
\mu'(\mathbf{x}_1) = \frac{\mathfrak{q}_0^{\sum_{i=1}^{N-1}\left(-1+\frac{a_i-x_{1,i}}{\varepsilon_1}\right)}}{(N-1)!} \prod_{i=1}^{N-1} \frac{(-\varepsilon_1)^{\frac{m_0^- - x_{1,i}}{\varepsilon_1}} \Gamma\left(\frac{m_0^- - x_{1,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{1,i}-m_0^+}{\varepsilon_1}} \Gamma\left(1+\frac{x_{1,i}-m_0^+}{\varepsilon_1}\right)} \tag{4.73}
$$

$$
\times \frac{1}{\prod_{1\leq j\neq k\leq N-1} \Gamma\left(\frac{x_{1,j}-x_{1,k}}{\varepsilon_1}\right)} \prod_{i=1}^{N-1} \prod_{\alpha=0}^{N-1} \frac{\pi}{\sin\pi\left(\frac{x_{1,i}-a_\alpha}{\varepsilon_1}\right)},
$$

and the contour $\mathcal{C}$ is chosen exactly as explained above. In particular, all the $A_\omega[\boldsymbol{\lambda}] = a_\omega + l(\lambda^{(\omega)})\varepsilon_1$ in the exponents were cancelled as we emphasized, so that the $\boldsymbol{\lambda}$-dependence of the integrand is solely in the product of $\mathbf{Q}$-observables, $C(\mathbf{x}_1)[\boldsymbol{\lambda}]$. We stress that the equality (4.70) holds at the level of observable expressions at each $\boldsymbol{\lambda}$. Thus, it is a direct manifestation of the transition of the regular monodromy surface defect into the $N-1$ $\mathbf{Q}$-observables.

Upon taking the vacuum expectation value, the integral transformation (4.70) relates the vacuum expectation value of the regular monodromy surface defect and the correlation function of the $\mathbf{Q}$-observables. Namely, we get

$$
\left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = \int_{\mathcal{C}} \prod_{i=1}^{N-1} \frac{dx_{1,i}}{2\pi \mathrm{i}\varepsilon_1} \mathcal{K}(\mathbf{u}; \mathbf{x}_1) \left\langle \prod_{i=1}^{N-1} Q(x_{1,i}) \right\rangle_{\mathbf{a}}. \tag{4.74}
$$

Note that the integral transformation is valid for any value of $\varepsilon_1, \varepsilon_2$.

In particular, the kernel $\mathcal{K}(\mathbf{u}, \mathbf{x}_1)$ does not involve ε_2, and the limit $\varepsilon_2 \to 0$ can be taken without harming the validity of the integral transformation with the same kernel. In the limit $\varepsilon_2 \to 0$, the normalized vacuum expectation value $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy surface defect gives a common eigenstate in $\mathcal{H}^{\mathbb{C}^\times}$ of the Q-operators, as we showed in section 3.1.2. We also showed in the companion paper [24] that this Q-eigenstate $\psi(\mathbf{a}; \mathbf{u})$ is also a common eigenstate of the quantum Hamiltonians of the $\mathfrak{gl}(2)$ XXX spin chain. Meanwhile, the correlation function of the $\mathbf{Q}$-observables in the right hand side of the transition formula (4.74) factorizes into the product of the vacuum expectation values of individual $\tilde{\mathbf{Q}}$-observables,

$$
\lim_{\varepsilon_2\to 0} \left\langle \prod_{i=1}^{N-1} Q(x_{1,i}) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})}{\varepsilon_2}} \prod_{i=1}^{N-1} Q(\mathbf{a}; x_{1,i}), \tag{4.75}
$$

since the **Q**-observables, being local on the topological plane $\mathbb{C}_2$, can be arbitrarily separated from each other in this limit. Therefore, this cluster decomposition of the **Q**-observables yields an integral transformation between the Q-eigenstate and the product of $N-1$ Q-functions,

$$\psi(\mathbf{a}; \mathbf{u}) = \int_{\mathcal{C}} \prod_{i=1}^{N-1} \frac{dx_{1,i}}{2\pi\mathrm{i}\varepsilon_1} \, \mathcal{K}(\mathbf{u}; \mathbf{x}_1) \prod_{i=1}^{N-1} Q(\mathbf{a}; x_{1,i}). \qquad (4.76)$$

Recall that each Q-function $Q(\mathbf{a}; x_{1,i})$ is separately a solution to the same $\hbar$-oper difference equation [24],

$$0 = \left[1 - t(\mathbf{a}; x)e^{-\varepsilon_1\partial_x} + \mathfrak{q}P(x)e^{-2\varepsilon_1\partial_x} \right] Q(\mathbf{a}; x + \varepsilon_1), \qquad (4.77)$$

which is none other than the Baxter TQ equation for the Q-function $Q(\mathbf{a}; x_{1,i})$. Therefore, the integral transformation (4.76) translates the $N-1$ spectral equations for the XXX spin chain (single $(N-1)$-body problem) into $(N-1)$-copies of identical second-order difference equations ($N-1$ single-body problems). Thus, we conclude (4.76), which expresses the transition between the regular monodromy surface defect and the $N-1$ **Q**-observables in the limit $\varepsilon_2 \to 0$, establishes the quantum separation of variables for the XXX spin chain associated to the A_1-quiver $\mathcal{N} = 2$ gauge theory. It was also studied in [78] without making $\varepsilon_2 \neq 0$ refinement.

We stress that our XXX spin chain associated to the four-dimensional A_1-quiver $\mathcal{N} = 2$ gauge theory is defined on bi-infinite $\mathfrak{gl}(2)$-modules, for which the algebraic Bethe ansatz method is not applicable by construction. In this sense, the separation of variables established in (4.76) generalizes the one on highest-weight modules (see [79], for example, for a previous study of the separation of variables for the $\mathfrak{gl}(2)$ spin chain for highest-weight modules). When the gauge theory parameters are specialized to trigger a higgsing to two-dimensional $\mathcal{N} = (2,2)$ theory, we recover the standard Bethe/gauge correspondence where the XXX spin chain is defined on lowest-weight modules, and we will recover from (4.76) the separation of variables for XXX spin chain on lowest-weight modules.

4.3.2 Matching with Sklyanin's prescription

We confirm here that the separation of variables for the XXX spin chain that we achieved above from the surface defect transition precisely reproduces the prescription by Sklyanin [80].

The gauge theoretical construction the R-matrices was established in [24].

$$R_\omega(x) = x - \theta_\omega - \varepsilon_1 \begin{pmatrix} u_\omega\partial_{u_\omega} - s_\omega & 2s_\omega\partial_{u_\omega} - u_\omega\partial_{u_\omega}^2 \\ u_\omega & -u_\omega\partial_{u_\omega} + s_\omega \end{pmatrix} \in \mathrm{End}(\mathbb{C}^2) \otimes \mathrm{End}(\mathcal{H}_\omega), \qquad (4.78)$$

where $\mathcal{H}_\omega = u_\omega^{\frac{m_\omega^+ - a_\omega}{\varepsilon_1}} \mathbb{C}((u_\omega))$ is a space of Laurent polynomials in a formal variable u_ω, realizing an evaluation module over the Yangian of $\mathfrak{gl}(2)$, $Y(\mathfrak{gl}(2))$. The spin s_ω is given in terms of the mass parameters by

$$s_\omega = \frac{m_\omega^+ - m_\omega^- - \varepsilon_1}{2\varepsilon_1}, \qquad \omega = 0, 1, \cdots, N-1. \qquad (4.79)$$

The trace part acts trivially. The evaluation parameter θ_ω is encoded in the rest of the degrees of freedom of the mass parameters by

$$\theta_\omega = \frac{m_\omega^+ + m_\omega^- + \varepsilon_1}{2}, \qquad \omega = 0, 1, \cdots, N-1. \tag{4.80}$$

Note that the $\mathfrak{gl}(2)$-module $\mathcal{H}_\omega$ is irreducible for generic values of gauge theory parameters. Moreover, it is a bi-infinite module, possessing neither a highest-weight state nor a lowest-weight state.

The monodromy matrix is obtained by concatenating these R-matrices,

$$T(x)|_{\mathcal{H}} = R_{N-1}(x)R_{N-2}(x)\cdots R_0(x) =: \begin{pmatrix} A(x) & B(x) \\ C(x) & D(x) \end{pmatrix}, \tag{4.81}$$

and we denoted the entries of the monodromy matrix by $A(x), B(x), C(x), D(x) \in \text{End}(\mathcal{H})$. We verify that the separated variables $\mathbf{x}_1 = (x_{1,i})_{i=1}^{N-1}$ are given by the zeros of the lower-left entry $C(x)$ of the monodromy matrix, deriving the prescription of Sklyanin in our $\mathcal{N} = 2$ gauge theoretical setup.

In fact, we will show that a stronger version of the statement is true. Let us define *partial* monodromy matrices by

$$T(x)|_{\otimes_{\omega'=\omega}^{N-1}\mathcal{H}_{\omega'}} = R_{N-1}(x)\cdots R_\omega(x) =: \begin{pmatrix} A_\omega(x) & B_\omega(x) \\ C_\omega(x) & D_\omega(x) \end{pmatrix}, \qquad \omega = 0, 1, \cdots, N-1, \tag{4.82}$$

where we denoted the entries by $A_\omega(x), B_\omega(x), C_\omega(x), D_\omega(x) \in \text{End}\left(\otimes_{\omega'=\omega}^{N-1}\mathcal{H}_{\omega'}\right)$. In particular, we recover the monodromy matrix at $\omega = 0$, $T_0(x) = T(x)|_{\mathcal{H}}$. It will turn out that the variables $\mathbf{x}_{\omega+1} = (x_{\omega+1,i})_{i=\omega+1}^{N-1}$ are provided by the zeros of $C_\omega(x)$ for each $\omega = 0, 1, \cdots, N-2$ ($C_{N-1}(x)$ does not depend on x by construction, so that there is no zero).

In the convention of the section 5.2 of [24], we undo the transformation of $g_\omega(x)$ and consider

$$g_N(x)^{-1}\left(T(x)|_{\otimes_{\omega'=\omega}^{N-1}\mathcal{H}_{\omega'}}\right)g_\omega(x) = L_{N-1}(x)L_{N-2}(x)\cdots L_\omega(x) =: \begin{pmatrix} \alpha(x) & \beta(x) \\ \gamma(x) & \delta(x) \end{pmatrix}, \tag{4.83}$$

since $C_\omega(x)$ and $\gamma(x)$ are related by a simple relation:

$$\begin{aligned} T(x)|_{\otimes_{\omega'=\omega}^{N-1}\mathcal{H}_{\omega'}} &= g_N(x)L_{N-1}(x)\cdots L_\omega(x)g_\omega(x)^{-1} \\ &= \frac{1}{u_0}\begin{pmatrix} 1 & b_N(x) \\ 0 & a_N(x) \end{pmatrix}\begin{pmatrix} \alpha_\omega(x) & \beta_\omega(x) \\ \gamma_\omega(x) & \delta_\omega(x) \end{pmatrix}\begin{pmatrix} a_\omega(x) & -b_\omega(x) \\ 0 & 1 \end{pmatrix}\frac{u_\omega}{a_\omega} \\ &= \begin{pmatrix} * & * \\ \frac{a_N(x)}{u_0}\gamma_\omega(x)u_\omega & * \end{pmatrix}, \end{aligned} \tag{4.84}$$

namely, $C_\omega(x) = a_N(x)u_0^{-1}\gamma_\omega(x)u_\omega = \frac{1}{P_0^+(x+\varepsilon_1)}u_0^{-1}\gamma_\omega(x)u_\omega$, where we denoted $g(x) = \frac{1}{u_\varpi}\begin{pmatrix} 1 & b_\omega(x) \\ 0 & a_\omega(x) \end{pmatrix}$. An important feature of the fractional TQ equation written in the matrix

form (cf. (5.11) in [24]) is that the lower-right entry of $L_\omega(x)$ is zero, yielding

$$
\begin{pmatrix} \alpha_\omega(x) & \beta_\omega(x) \\ \gamma_\omega(x) & \delta_n(x) \end{pmatrix} = \begin{pmatrix} \alpha_{\omega+1}(x) & \beta_{\omega+1}(x) \\ \gamma_{\omega+1}(x) & \delta_{\omega+1}(x) \end{pmatrix} \begin{pmatrix} T_\omega(x) & -\mathfrak{q}_\omega P_\omega^-(x) \\ P_{\omega+1}^+(x) & 0 \end{pmatrix}
$$
$$
= \begin{pmatrix} * & * \\ \gamma_{\omega+1}(x)T_\omega(x) + P_{\omega+1}^+(x)\delta_{\omega+1}(x) & -\gamma_{\omega+1}(x)\mathfrak{q}_\omega P_\omega^-(x) \end{pmatrix},
$$
(4.85)

which implies that $\gamma_\omega(x)$ satisfies a second-order recursion relation,

$$
\gamma_{\omega-1}(x) = \gamma_\omega(x)T_{\omega-1}(x) - \gamma_{\omega+1}(x)\mathfrak{q}_\omega P_\omega(x).
$$
(4.86)

Thus, having $\gamma_\omega(x) = (x - m_0^+ + \varepsilon_1)u_0 C_\omega(x)u_\omega^{-1}$, we get a recursive formula for the degree $N - \omega - 1$ polynomial $C_\omega(x)$ as

$$
C_{\omega-1}(x) + C_{\omega+1}(x)\frac{u_{\omega-1}}{u_\omega}P_\omega(x) = C_\omega(x)u_\omega^{-1}T_{\omega-1}(x)u_{\omega-1}
$$
$$
=: C_\omega(x)\left(\left(1 + \frac{u_{\omega-1}}{u_\omega}\right)x - W_\omega\right),
$$
(4.87)

where the differential operator W_ω is determined from

$$
T_\omega(x) = (1 + \mathfrak{q}_\omega)x + u_{\omega+1}(\partial_{u_{\omega+1}} - \partial_{u_\omega}) - m_{\omega+1}^+ - \mathfrak{q}_\omega m_\omega^-
$$
(4.88)

as

$$
W_\omega = u_\omega^{-1}\left(-\varepsilon_1 u_\omega(\partial_{u_\omega} - \partial_{u_{\omega-1}}) + m_\omega^+ + \mathfrak{q}_{\omega-1}m_{\omega-1}^-\right)u_{\omega-1}
$$
$$
= -\varepsilon_1\left(\mathfrak{q}_{\omega+1}\partial_{\mathfrak{q}_{\omega+1}} - \mathfrak{q}_\omega\partial_{\mathfrak{q}_\omega} - \mathfrak{q}_\omega(\mathfrak{q}_\omega\partial_{\mathfrak{q}_\omega} - \mathfrak{q}_{\omega-1}\partial_{\mathfrak{q}_{\omega-1}})\right) + m_\omega^+\frac{u_{\omega-1}}{u_\omega} + m_{\omega-1}^- + \varepsilon_1.
$$
(4.89)

To show that the zeros of $C_\omega(x)$ give the variables $\mathbf{x}_{\omega+1} = (x_{\omega+1,i})_{i=\omega+1}^{N-1}$ in the separated expression for the eigenstate, let us act the recursion formula (4.87) for the operators $C_\omega(x)$ to the integrand of the separated expression (4.74). The differential operator W_ω acts in a simple way, since the all the $\mathbf{u}$-dependent terms in the integrand are just elementary functions in $\mathbf{u}$. Specifically, the $\mathbf{u}$-dependent terms in the integrand are gathered into

$$
\mathcal{K}_{\mathbf{u}} = \mathbf{c} \times \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{\sum_{i=\omega+1}^{N-1}\left(-1 + \frac{a_i - x_{\omega+1,i}}{\varepsilon_1}\right)} \times \prod_{\omega=0}^{N-1} u_\omega^{\frac{m_\omega^+ - a_\omega}{\varepsilon_1}},
$$
(4.90)

where the first term is from (4.71), the second term is from $\mathbf{V}'(\mathbf{x}_1)$ (4.72) and $\mu'(\mathbf{x}_1)$ (4.73), and the third term is the classical part (2.28) of the regular monodromy defect observable. Thus, the differential operator W_ω is converted into a number upon acting on the separated expression as

$$
w_\omega := \mathcal{K}_{\mathbf{u}}^{-1}W_\omega\mathcal{K}_{\mathbf{u}}
$$
$$
= \mathfrak{q}_\omega(m_\omega^+ + m_\omega^- + \varepsilon_1) + \sum_{i=\omega+2}^{N-1} x_{\omega+2,i} - \sum_{i=\omega+1}^{N-1} x_{\omega+1,i} - \mathfrak{q}_\omega\left(\sum_{i=\omega+1}^{N-1} x_{\omega+1,i} - \sum_{i=\omega}^{N-1} x_{\omega,i}\right).
$$
(4.91)

After acting on the separated expression, the recursion formula (4.87) reads

$$C_{\omega-1}(x) + C_{\omega+1}(x)\frac{u_{\omega-1}}{u_\omega}P_\omega(x) = C_\omega(x)\left(\left(1 + \frac{u_{\omega-1}}{u_\omega}\right)x - w_\omega\right).$$ (4.92)

It can be solved by a straightforward computation, starting from $\omega = N - 2$ all the way to $\omega = 0$. The solution is simply

$$C_\omega(x) = \prod_{i=\omega+1}^{N-1}(x - x_{\omega+1,i}), \qquad \omega = 0, 1, \cdots, N - 2.$$ (4.93)

Therefore, we prove that the zeros of $C_\omega(x)$ provide the variables $\mathbf{x}_{\omega+1} = (x_{\omega+1,i})_{i=\omega+1}^{N-1}$ in the separated expression (4.76) for the XXX spin chain eigenstate. In particular, the zeros of $C(x) = C_0(x)$ give the separated variables $\mathbf{x}_1 = (x_{1,i})_{i=1}^{N-1}$. This confirms the separation of variables that we established from the surface defect transition recovers the Sklyanin's prescription.

4.4 Separation of variables for Gaudin model

In section 2, we discussed that the **Q**-observable and the **H**-observable arise from the same two-dimensional gauged linear sigma model coupled to the four-dimensional gauge theory. The former has non-zero vev of the complex adjoint scalar, while the latter has non-zero the complexified FI parameter. In section 3, we elucidated that the Fourier transformation between the two observables implements a transition between the Q-operator and the Hecke operator.

In the separated expression (4.76) for the spin chain eigenstate, the integrand contains a product of $N - 1$ normalized vacuum expectation values $Q(\mathbf{a}; x_{1,i})$ of the **Q**-observables. The separated variables $\mathbf{x}_1 = (\mathbf{x}_{1,i})_{i=1}^{N-1}$, which obviously act diagonally on $Q(\mathbf{a}; x_{1,i})$, are precisely the Coulomb parameters of the **Q**-observables. Noting that the transition between the **Q**-observable and the **H**-observable manifests as a integral transformation for their observable expressions (2.18), we expect to have another separated expression for the eigenstate where the integrand is given by a product of $N - 1$ normalized vacuum expectation values of the **H**-observable instead.

Each of the normalized vacuum expectation values of the **H**-observable in the product is separately a solution to the same oper differential equation, where the oper specifies a twisted $\mathcal{D}$-module on $\mathrm{Bun}_{\mathrm{PGL}(N)}(\mathbb{P}^1; S)$, whose section is the normalized vacuum expectation value $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy surface defect, representing the spectral equations for the Gaudin model. Thus, the resulting integral transformation for the twisted coinvariants precisely gives a separation of variables for the Gaudin model, where the separated variables are the complexified FI parameters of the surface defect theory for the **H**-observables. In this way, we observe that the separation of variables for the XXX spin chain and the Gaudin model are connected to each other by the surface defect transition between the **Q**-observable and the **H**-observable. This is indeed in accordance with the bispectral duality between the two models that we established in section 3.2.2.

For this, let us revisit the monodromy defect observable in the negative chamber (4.64) expressed as a summation over non-negative integers. Instead of turning the summation into a contour integral of Barnes type as we did for the separation of variables for the XXX

spin chain, we can actually perform the summation to transform the **Q**-observable into the **H**-observable, after properly expressing some combinations of Γ-functions as contour integrals of Pochhammer type. To be explicit, we will exemplify the case of $N = 2$ and 3 first, and then present the prescription for general N.

4.4.1 $\mathfrak{sl}(2)$ case

We have already studied the regular monodromy defect at $N = 2$ in section 4.2.2, where its transition to a **H**-observable was shown by an integral transformation between their observable expressions. For completeness, let us restate here the integral transformation:

$$\left\langle \Psi(\mathfrak{q}_0) \right\rangle_{\mathbf{a}} = \int_{\mathrm{PC}} dy \, \mathcal{K}(\mathfrak{q}_0; y) \left\langle \mathbf{H}^{(1)}(y) \right\rangle_{\mathbf{a}}, \tag{4.94}$$

where the integral kernel is

$$\begin{aligned}
\mathcal{K}(\mathfrak{q}_0; y) = \mathfrak{q}_0^{1 + \frac{a_1 - m_1^+}{\varepsilon_1}} (1 + \mathfrak{q}_0)^{1 + \frac{1}{\varepsilon_1}(m_0^+ + m_0^- - a_0 - a_1)} \\
\times y^{-\frac{1}{2} + \frac{m_0^- - m_1^-}{2\varepsilon_1}} (y - \mathfrak{q}_0)^{-1 + \frac{m_1^+ - m_0^-}{\varepsilon_1}} (y - \mathfrak{q})^{\frac{\bar{m}^- - \bar{a}}{\varepsilon_1}} (y - 1)^{-1 - \frac{\bar{m}^+ - \bar{a}}{\varepsilon_1}}.
\end{aligned} \tag{4.95}$$

We remind that PC is the Pochhammer contour enclosing $y = 0$ and $y = \mathfrak{q}_0$. The integral transformation is valid for any value of $\varepsilon_1, \varepsilon_2$. In particular, ε_2 is not involved in the kernel $\mathcal{K}(\mathfrak{q}_0; y)$, so that the limit $\varepsilon_2 \to 0$ can be safely taken without harming the validity of the transformation with the same kernel. The limit $\varepsilon_2 \to 0$ of the formula gives

$$\psi(\mathbf{a}; \mathfrak{q}_0) = \int_{\mathrm{PC}} dy \, \mathcal{K}(\mathfrak{q}_0; y) \chi_1(\mathbf{a}; y), \tag{4.96}$$

where $\psi(\mathbf{a}; \mathfrak{q}_0)$ is the normalized vacuum expectation value of the regular monodromy surface defect (at $N = 2$) and $\chi_1(\mathbf{a}; y)$ is the normalized vacuum expectation value of the **H**-observable, in the limit $\varepsilon_2 \to 0$.

Recall that $\psi(\mathbf{a}; \mathfrak{q}_0)$ is a section of twisted $\mathcal{D}$-module on $\mathrm{Bun}_{\mathrm{PGL}(2)}(\mathbb{P}^1; S)$ corresponding to the oper $\rho_{\mathbf{a}}$, namely, an eigenfunction of the Gaudin Hamiltonian with the eigenvalue fixed by the chosen oper.[12] Also note that $\chi_1(\mathbf{a}; y)$ is a solution to the second-order oper differential equation for $\rho_{\mathbf{a}}$. Therefore, the integral transformation is precisely the separation of variable transformation for the Gaudin model. We also confirmed our construction recovers Sklyanin's prescription of the separation of variables by using the Gaudin Lax matrix expressed in the $\mathcal{N} = 2$ gauge theoretical terms, obtained in [8]. Namely, a straightforward computation shows that the zero of the lower-left entry of the 2×2 Lax matrix $L(z)$, when acting on the separation of variable kernel $\mathcal{K}(\mathfrak{q}_0; y)$, is precisely at $z = y$.

Calling it separation of variables is somewhat misleading in this particular case $N = 2$, since the Gaudin model is a single-body spectral problem even before the transformation. Nevertheless, it clarifies the distinction between the eigenfunction $\psi(\mathbf{a}; \mathfrak{q}_0)$ of the Gaudin Hamiltonian and the oper solution $\chi(\mathbf{a}; y)$, where the two relevant second-order differential equations are almost identical except a shuffling of monodromy parameters. Such a coincidence will not persist in the separation of variables for higher-rank Gaudin models that we discuss now.

---

[12]There is only one Gaudin Hamiltonian for this case of $N = 2$.

4.4.2 $\mathfrak{sl}(3)$ case

The regular monodromy defect observable in the negative chamber is written as a summation over $\frac{3 \times 2}{2} = 3$ non-negative integers $(s_{1,1}, s_{2,1}, s_{2,2})$,

$$
\Psi[\boldsymbol{\lambda}]_{\zeta<0} = \sum_{s_{1,1}, s_{2,1}, s_{2,2}=0}^{\infty} \mathfrak{q}_0^{s_{1,1}+s_{2,1}} \mathfrak{q}_1^{s_{2,1}+s_{2,2}}
$$
$$
\times \mathbb{E} \left[\frac{S\Sigma_0^* - \Sigma_0\Sigma_0^* - \Sigma_1\Sigma_1^* + \Sigma_0\Sigma_1^* - M_0^+\Sigma_0^* + q_1^{-1}(M_0^-)^*\Sigma_0 - M_1^+\Sigma_1^* + q_1^{-1}(M_1^-)^*\Sigma_1}{P_1^*} \right],
\tag{4.97}
$$

where

$$
\begin{aligned}
\Sigma_0 &= e^{A_1-s_{1,1}\varepsilon_1} + e^{A_2-s_{2,1}\varepsilon_1} \\
\Sigma_1 &= e^{A_2-(s_{2,1}+s_{2,2})\varepsilon_1}.
\end{aligned}
\tag{4.98}
$$

By substituting (4.98) into (4.97), we get

$$
\Psi[\boldsymbol{\lambda}]_{\zeta<0} = \sum_{s_{1,1}, s_{2,1}, s_{2,2}=0}^{\infty} \mathfrak{q}_0^{s_{1,1}+s_{2,1}} \mathfrak{q}_1^{s_{2,1}+s_{2,2}} \frac{\Gamma\left(s_{1,1}-s_{2,1}-s_{2,2}+\frac{A_2-A_1}{\varepsilon_1}\right)}{\Gamma\left(s_{2,1}-s_{1,1}+\frac{A_1-A_2}{\varepsilon_1}\right)\Gamma\left(-s_{2,1}+s_{1,1}+\frac{A_2-A_1}{\varepsilon_1}\right)}
$$
$$
\times \frac{(-\varepsilon_1)^{\frac{m_1^--m_1^++2m_0^--2m_2^+}{\varepsilon_1}}}{s_{2,2}!}(-1)^{s_{2,2}} \times \frac{\prod_{\alpha=1}^{2}\prod_{i=1}^{2}\Gamma\left(-s_{i,1}+\frac{A_i-m_\alpha^+}{\varepsilon_1}\right)}{\prod_{i=1}^{2}\Gamma\left(-s_{i,1}+\frac{A_i-m_0^-}{\varepsilon_1}\right)\prod_{\pm}\Gamma\left(-s_{2,1}-s_{2,2}+\frac{A_2-m_1^\pm}{\varepsilon_1}\right)}
$$
$$
\times Q(A_1-(s_{1,1}+1)\varepsilon_1)Q(A_2-(s_{2,1}+1)\varepsilon_1) \prod_{i=1,2} \mathbb{E}\left[\frac{e^{A_i-(s_{i,1}+1)\varepsilon_1}(M^+)^*}{P_1^*}\right].
\tag{4.99}
$$

The summations can be converted into contour integrals as we now explain. First, note that the summation over $s_{2,2}$ produces a hypergeometric function. We use the Euler integral representation of the hypergeometric function

$$
{}_2F_1(a,b;c;z) = \frac{e^{-c\pi i}\Gamma(c)\Gamma(1-b)\Gamma(1+b-c)}{4\pi^2} \int_{\mathrm{PC}} dt\, t^{b-1}(1-t)^{c-b-1}(1-zt)^a, \tag{4.100}
$$

where PC denotes the Pochhammer contour enclosing the singularities at $t=0$ and $t=1$. Then, we can express the ratios of remaining Γ-functions having $s_{1,1}$ or $s_{2,1}$ in their arguments as contour integrals using (4.56). In this way, all the $s_{1,1}$ and $s_{2,1}$ only appear in the exponents of simple rational functions of integration variables and in the arguments of the two **Q**-observables. Finally, we change the order of summation and integration. Each of $s_{1,1}$ and $s_{2,1}$ summations yields a **H**-observable, which is analytically continued from the original

domain if needed. The net result is a four-dimensional contour integral given by

$$
\Psi(\mathbf{u})[\boldsymbol{\lambda}]_{\zeta<0} = \mathfrak{q}_0^{-1+\frac{A_1-\bar{a}}{\varepsilon_1}}(\mathfrak{q}_0\mathfrak{q}_1)^{-1+\frac{A_2-\bar{a}}{\varepsilon_1}}
$$

$$
\times \int_{\mathrm{PC}^{\times 4}} \prod_{i=1}^{4} dt_i \, t_1^{-1+\frac{m_1^-+\bar{m}^-}{\varepsilon_1}}(1-t_1)^{-\frac{m_1^-}{\varepsilon_1}}(1-\mathfrak{q}_1 t_1)^{\frac{m_1^++\bar{m}^-}{\varepsilon_1}} t_2^{-\frac{m_1^++\bar{m}^-}{\varepsilon_1}}(1-t_2)^{-1+\frac{m_1^+-m_1^--\bar{m}^-}{\varepsilon_1}}
$$

$$
\times t_3^{-\frac{m_2^++\bar{m}^-}{\varepsilon_1}}(1-t_3)^{-2+\frac{m_2^+-m_0^-}{\varepsilon_1}} t_4^{-\frac{m_2^++\bar{m}^-}{\varepsilon_1}}(1-t_4)^{-2+\frac{m_2^+-m_0^-}{\varepsilon_1}}(t_3-t_4)
$$

$$
\times ((1-t_1)t_2 t_4 - \mathfrak{q}_1\mathfrak{q}_2)^{\frac{\bar{m}^--\bar{a}}{\varepsilon_1}}\left(\frac{\mathfrak{q}_0\mathfrak{q}_1 t_3}{t_1(1-\mathfrak{q}_1 t_1)}-1\right)^{-1-\frac{\bar{m}^+-\bar{a}}{\varepsilon_1}}
$$

$$
\times \left(\frac{t_3}{t_1(1-\mathfrak{q}_1 t_1)}+\mathfrak{q}_2\right)^{\frac{\bar{m}^--\bar{a}}{\varepsilon_1}}\left(\frac{\mathfrak{q}_0\mathfrak{q}_1 t_3}{t_1(1-\mathfrak{q}_1 t_1)}+1\right)^{-1-\frac{\bar{m}^+-\bar{a}}{\varepsilon_1}}
$$

$$
\times \mathbf{H}^{(1)}\left(\mathfrak{q}_0(1-t_1)t_2 t_4\right)[\boldsymbol{\lambda}]\,\mathbf{H}^{(2)}\left(-\frac{\mathfrak{q}_0\mathfrak{q}_1 t_3}{t_1(1-\mathfrak{q}_1 t_1)}\right)[\boldsymbol{\lambda}],
$$

(4.101)

where we absorbed irrelevant prefactor into the observable expression and omitted it. Now, applying the crossing formula (4.43) we recover the original monodromy surface defect observable,

$$
\Psi[\boldsymbol{\lambda}] = \mathfrak{q}_0^{-l(\lambda^{(1)})}(\mathfrak{q}_0\mathfrak{q}_1)^{-l(\lambda^{(2)})} c_0^{R_\infty^{(0)}+\frac{1}{\varepsilon_1}\sum_{\alpha=0}^{2}(A_\alpha-a_\alpha)} c_1^{R_\infty^{(1)}} \times \Psi\left(\mathfrak{q}_\omega \frac{c_{\omega+1}}{c_{\omega-1}}\right)[\boldsymbol{\lambda}]_{\zeta<0}
$$

$$
= \mathfrak{q}_0^{-1+\frac{a_1-\bar{a}}{\varepsilon_1}}(\mathfrak{q}_0\mathfrak{q}_1)^{-1+\frac{a_2-\bar{a}}{\varepsilon_1}}(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1)^{2+\frac{m_0^++m_0^-2\bar{a}}{\varepsilon_1}}(1+\mathfrak{q}_1)^{-1+\frac{m_0^++m_0^--2\bar{a}}{\varepsilon_1}}
$$

$$
\times \int_{\mathrm{PC}^{\times 4}} \prod_{i=1}^{4} dt_i \, t_1^{-1+\frac{m_1^-+\bar{m}^-}{\varepsilon_1}}(1-t_1)^{-\frac{m_1^-}{\varepsilon_1}}\left(1-\frac{\mathfrak{q}_1 t_1}{1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1}\right)^{\frac{m_1^++\bar{m}^-}{\varepsilon_1}}
$$

$$
\times t_2^{-\frac{m_1^++\bar{m}^-}{\varepsilon_1}}(1-t_2)^{-1+\frac{m_1^+-m_1^--\bar{m}^-}{\varepsilon_1}}
$$

$$
\times t_3^{-\frac{m_2^++\bar{m}^-}{\varepsilon_1}}(1-t_3)^{-2+\frac{m_2^+-m_0^-}{\varepsilon_1}} t_4^{-\frac{m_2^++\bar{m}^-}{\varepsilon_1}}(1-t_4)^{-2+\frac{m_2^+-m_0^-}{\varepsilon_1}}(t_3-t_4)
$$

$$
\times \left((1-t_1)t_2 t_4 - \frac{\mathfrak{q}_1\mathfrak{q}_2}{(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1)(1+\mathfrak{q}_1)}\right)^{\frac{\bar{m}^--\bar{a}}{\varepsilon_1}}\left(\frac{\mathfrak{q}_0\mathfrak{q}_1(1+\mathfrak{q}_1)t_3}{t_1(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1-\mathfrak{q}_1 t_1)}-1\right)^{-1-\frac{\bar{m}^+-\bar{a}}{\varepsilon_1}}
$$

$$
\times \left(\frac{t_3}{t_1\left(1-\frac{\mathfrak{q}_1 t_1}{1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1}\right)}+\frac{\mathfrak{q}_2}{1+\mathfrak{q}_1}\right)^{\frac{\bar{m}^--\bar{a}}{\varepsilon_1}}\left(\frac{\mathfrak{q}_0\mathfrak{q}_1(1+\mathfrak{q}_1)t_3}{t_1(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1-\mathfrak{q}_1 t_1)}+1\right)^{-1-\frac{\bar{m}^+-\bar{a}}{\varepsilon_1}}
$$

$$
\times \mathbf{H}^{(1)}\left(\mathfrak{q}_0(1+\mathfrak{q}_1)(1-t_1)t_2 t_4\right)[\boldsymbol{\lambda}]\,\mathbf{H}^{(2)}\left(-\frac{\mathfrak{q}_0\mathfrak{q}_1(1+\mathfrak{q}_1)t_3}{t_1(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1-\mathfrak{q}_1 t_1)}\right)[\boldsymbol{\lambda}],
$$

(4.102)

expressed as an integral transformation of a product of two $\mathbf{H}$-observables. Note that the relation holds at the level of observable expressions at each $\boldsymbol{\lambda}$, and the $\boldsymbol{\lambda}$-dependence is solely at the $\mathbf{H}$-observables in the right hand side. Therefore, upon taking the vacuum

expectation value, we arrive at

$$\left\langle \Psi(\mathbf{u}) \right\rangle_{\mathbf{a}} = \int_{\mathrm{PC}} dy_1 dy_2 \, \mathcal{K}(\mathbf{u}; \mathbf{y}) \left\langle \mathbf{H}^{(1)}(y_1) \mathbf{H}^{(2)}(y_2) \right\rangle_{\mathbf{a}}, \tag{4.103}$$

where we split the four-dimensional contour integral (4.102) into two parts; the two-dimensional contour integral for $y_1 = \mathfrak{q}_0(1+\mathfrak{q}_1)(1-t_1)t_2 t_4$, $y_2 = -\frac{\mathfrak{q}_0 \mathfrak{q}_1(1+\mathfrak{q}_1)t_3}{t_1(1+\mathfrak{q}_0+\mathfrak{q}_0\mathfrak{q}_1-\mathfrak{q}_1 t_1)}$ and the integral of the rest two. The integral of the rest and the elementary functions of $\mathfrak{q}_0$ in front are gathered into the separation of variable kernel $\mathcal{K}(\mathbf{u}; \mathbf{y})$. Note that the relation is valid for any value of $\varepsilon_1, \varepsilon_2$.

The vacuum expectation value $\langle \Psi(\mathbf{u}) \rangle_{\mathbf{a}}$ of the monodromy surface defect is identified with a twisted coinvariant of four $\widehat{\mathfrak{sl}}(3)$-modules (two (co-)Verma modules (at 0 and ∞) and two bi-infinite modules (at $\mathfrak{q}$ and 1)) [43, 68]. Meanwhile, the correlation function of the two $\mathbf{H}$-observables $\langle \mathbf{H}^{(1)}(y_1) \mathbf{H}^{(2)}(y_2) \rangle_{\mathbf{a}}$ gives a $\mathcal{W}_3$-algebra conformal block with two generic vertex operators (at 0 and ∞), two semi-degenerate vertex operators (at $\mathfrak{q}$ and 1), and two simplest fully degenerate vertex operators (at y_1 and y_2) [14, 43]. In this sense, we may interpret this integral transformation as a generalization of the KZ/BPZ correspondence [37, 46, 47] to a higher-rank.

Moreover, the kernel $\mathcal{K}(\mathbf{u}; \mathbf{y})$ does not involve ε_2 so that the limit $\varepsilon_2 \to 0$ can be taken to the integral transformation. In the integrand, the correlation function of two $\mathbf{H}$-observables in the integrand factorizes in the limit $\varepsilon_2 \to 0$,

$$\lim_{\varepsilon_2 \to 0} \left\langle \mathbf{H}^{(1)}(y_1) \mathbf{H}^{(2)}(y_2) \right\rangle_{\mathbf{a}} = e^{\frac{\widetilde{\mathcal{W}}(\mathbf{a};\mathfrak{q})}{\varepsilon_2}} \chi_1(\mathbf{a}; y_1) \chi_2(\mathbf{a}; y_2), \tag{4.104}$$

since the two $\mathbf{H}$-observables are local on the topological $\mathbb{C}_2$-plane, and thus can be arbitrarily separated from each other. Thus, the integral transformation reads

$$\psi(\mathbf{a}; \mathbf{u}) = \int_{\mathrm{PC}} dy_1 dy_2 \, \mathcal{K}(\mathbf{u}; \mathbf{y}) \, \chi_1(\mathbf{a}; y_1) \chi_2(\mathbf{a}; y_2). \tag{4.105}$$

As briefly reviewed in section 3, it was shown in [8] that the normalized vacuum expectation value $\psi(\mathbf{a}; \mathbf{u})$ of the regular monodromy surface defect in the limit $\varepsilon_2 \to 0$ is identified with a section of the twisted $\mathcal{D}$-module corresponding to an oper $\rho_{\mathbf{a}}$. It is also a common eigenfunction of the Gaudin Hamiltonians with the spectra fixed by the chosen oper. Meanwhile, the integrand on the right hand side contains a product of two normalized vacuum expectation values $\chi(\mathbf{a}; y)$ of the $\mathbf{H}$-observable in the limit $\varepsilon_2 \to 0$, which are solutions to the third-order oper differential equation associated to $\rho_{\mathbf{a}}$. Therefore, the integral transformation converts the Gaudin spectral equations (a two-particle system) into two copies of identical third-order differential equations (two independent single-particle systems). Therefore, we conclude that (4.105) gives a separation of variable transformation for the $\mathfrak{sl}(3)$ Gaudin model.

4.4.3 General $\mathfrak{sl}(N)w$ case

For general $\mathfrak{sl}(N)$ Gaudin model, we can achieve the integral transformation for the separation of variables by the same procedure. The only difference is that we have to deal with a higher multi-dimensional contour integral to sort out the kernel $\mathcal{K}(\mathbf{u}; \mathbf{y})$. Here, we will outline how this computation can actually be done.

Just as in the cases of $N = 2$ and 3, the separation of variable kernel $\mathcal{K}$ for general N is obtained by simply rewriting the following expression in (4.64),

$$
\sum_{\mathbf{s}} \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{\sum_{i=\omega+1}^{N-1}\left(-1+\frac{A_i-x_{\omega,i}}{\varepsilon_1}\right)}
$$
$$
\times \mathbb{E}\left[-\frac{\Sigma_0 q_1^{-1}(M^+)^*}{P_1^*} - \sum_{\omega=0}^{N-2} \frac{\Sigma_\omega(\Sigma_\omega - \Sigma_{\omega+1})^* + M_\omega^+ \Sigma_\omega^* - q_1^{-1}(M_\omega^-)^* \Sigma_\omega - (N-1)}{P_1^*}\right].
$$
(4.106)

The second line gives complicated ratios of Γ-functions, given by

$$
\prod_{i=1}^{N-1} \frac{(-\varepsilon_1)^{\frac{m_0^- - x_{1,i}}{\varepsilon_1}} \Gamma\left(\frac{m_0^- - x_{1,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{1,i}-m_0^+}{\varepsilon_1}} \Gamma\left(1 + \frac{x_{1,i}-m_0^+}{\varepsilon_1}\right)} \prod_{\alpha=0}^{N-1} \frac{(-\varepsilon_1)^{\frac{x_{1,i}-m_\alpha^+ - \varepsilon_1}{\varepsilon_1}}}{\Gamma\left(-\frac{x_{1,i}-m_\alpha^+}{\varepsilon_1}\right)} \frac{1}{\prod_{1\leq i \neq j \leq N-1} \Gamma\left(\frac{x_{1,i}-x_{1,j}}{\varepsilon_1}\right)}
$$

$$
\times \prod_{\omega=0}^{N-2} \prod_{i=\omega+1}^{N-1} \frac{\varepsilon_1^{-s_{i,\omega+2}}}{s_{i,\omega+2}!} \frac{(-\varepsilon_1)^{\frac{m_\omega^- - x_{\omega+1,i}}{\varepsilon_1}} \Gamma\left(\frac{m_\omega^- - x_{\omega+1,i}}{\varepsilon_1}\right)}{(-\varepsilon_1)^{1+\frac{x_{\omega+1,i}-m_\omega^+}{\varepsilon_1}} \Gamma\left(1 + \frac{x_{\omega+1,i}-m_\omega^+}{\varepsilon_1}\right)}
$$

$$
\frac{\prod_{\substack{\omega+1\leq i \leq N-1 \\ \omega+2 \leq j \neq i \leq N-1}} (-\varepsilon_1)^{\frac{x_{\omega+2,j}-x_{\omega+1,i}}{\varepsilon_1}} \Gamma\left(\frac{x_{\omega+2,j}-x_{\omega+1,i}}{\varepsilon_1}\right)}{\prod_{\omega+1\leq i \neq j \leq N-1} \Gamma\left(\frac{x_{\omega+1,i}-x_{\omega+1,j}}{\varepsilon_1}\right)}.
$$
(4.107)

Let us start with plugging in $x_{N-1,N-1} = x_{N-2,N-1} - s_{N-1,N-1}\varepsilon_1$. Collecting all the terms involving $x_{N-1,N-1}$, the summation over $s_{N-1,N-1}$ gives a hypergeometric function. Then we use the Euler integral representation (4.100) to rewrite the hypergeometric function. Namely,

$$
\sum_{s_{N-1,N-1}=0}^{\infty} \frac{(-\mathfrak{q}_{N-2})^{s_{N-1,N-1}}}{s_{N-1,N-1}!} \frac{\Gamma\left(\frac{m_{N-2}^- - x_{N-1,N-1}}{\varepsilon_1}\right) \Gamma\left(\frac{x_{N-1,N-1}-x_{N-2,N-2}}{\varepsilon_1}\right)}{\Gamma\left(1 + \frac{x_{N-1,N-1}-m_{N-2}^+}{\varepsilon_1}\right)}
$$

$$
= \frac{\Gamma\left(\frac{m_{N-2}^- - x_{N-2,N-1}}{\varepsilon_1}\right) \Gamma\left(\frac{x_{N-2,N-1}-x_{N-2,N-2}}{\varepsilon_1}\right)}{\Gamma\left(1 + \frac{x_{N-2,N-1}-m_{N-2}^+}{\varepsilon_1}\right)}
$$

$$
\times {}_2F_1\left(\frac{m_{N-2}^\pm - x_{N-2,N-1}}{\varepsilon_1}; 1 + \frac{x_{N-2,N-2}-x_{N-2,N-1}}{\varepsilon_1}; -\mathfrak{q}_{N-2}\right)
$$

$$
= \frac{e^{-\pi \mathrm{i}\left(1 + \frac{x_{N-2,N-2}-x_{N-2,N-1}}{\varepsilon_1}\right)}}{4\pi} \frac{\Gamma\left(\frac{m_{N-2}^- - x_{N-2,N-1}}{\varepsilon_1}\right) \Gamma\left(\frac{m_{N-2}^+ - x_{N-2,N-2}}{\varepsilon_1}\right)}{\sin \pi\left(\frac{x_{N-2,N-1}-x_{N-2,N-2}}{\varepsilon_1}\right)}
$$

$$
\times \int_{\mathrm{PC}} dt_{N-1}\, t_{N-1}^{\frac{m_{N-2}^+ - x_{N-2,N-1}}{\varepsilon_1}-1} (1-t_{N-1})^{\frac{m_{N-2}^+ - x_{N-2,N-2}}{\varepsilon_1}} (1+\mathfrak{q}_{N-2} t_{N-1})^{-\frac{m_{N-2}^- - x_{N-2,N-2}}{\varepsilon_1}}.
$$
(4.108)

The sine function in the denominator can be combined with the cross terms in the next level,

$$\frac{\pi}{\sin \pi \left(\frac{x_{N-2,N-1} - x_{N-2,N-2}}{\varepsilon_1}\right)} \frac{1}{\Gamma\left(\frac{x_{N-2,N-1} - x_{N-2,N-2}}{\varepsilon_1}\right) \Gamma\left(\frac{x_{N-2,N-2} - x_{N-2,N-1}}{\varepsilon_1}\right)}$$

$$= \frac{\Gamma\left(\frac{x_{N-2,N-2} - x_{N-2,N-1}}{\varepsilon_1} + 1\right)}{\Gamma\left(\frac{x_{N-2,N-2} - x_{N-2,N-1}}{\varepsilon_1}\right)} = \frac{x_{N-2,N-2} - x_{N-2,N-1}}{\varepsilon_1}. \tag{4.109}$$

We combine this term with relevant Γ-functions in the next level using

$$\frac{\Gamma(a)\Gamma(b)\Gamma(c-a)\Gamma(d-b)}{\Gamma(c)\Gamma(d)}(c-d) \tag{4.110}$$

$$= \int_{\mathrm{PC}^{\times 2}} dt_1 dt_2 \, t_1^{a-1}(1-t_1)^{c-a-2} t_2^{b-1}(1-t_2)^{d-b-2} \left[(c-a-1)(1-t_2) - (d-b-1)(1-t_1)\right]$$

We can apply this procedure repetitively, leaving only ratios of Γ-functions which can be converted into Pochhammer contour integrals by (4.56). Then the summation over the non-negative integers $s_{\omega,1}$, $\omega = 1, 2, \cdots, N-1$ can be computed inside the integral, yielding a product of $N-1$ **H**-observables. By leaving the $N-1$ Pochhammer contour integrals for the arguments of the **H**-observables undone while performing the rest of the integrals, we obtain the separation of variable kernel $\mathcal{K}(\mathbf{u}; \mathbf{y})$ for the $\mathfrak{sl}(N)$ Gaudin model. See appendix B for more detail.

5 Discussion

We explained that the Q-operator and the Hecke operator originate from the same two-dimensional $\mathcal{N} = (2, 2)$ sigma model coupled to the four-dimensional gauge theory, but in two different parameter regimes. In the former, the vev of the complex adjoint scalar is turned on, while in the later, the complexified FI parameter is turned on. The transition between the two can be initiated by specializing the adjoint scalar vev (resp. the complexified FI parameter) and turning on the complexified FI parameter (resp. the adjoint scalar vev). Such a transition of the surface defect was interpreted as a Fourier transformation between the Q-operator and the Hecke operator, which then led to the bispectral duality between the XXX spin chain and the Gaudin model associated with the same $\mathcal{N} = 2$ gauge theory.

Moreover, we elucidated that there is a dual two-dimensional sigma model description of the monodromy surface defect. Using this description, its observable expression was computed in a different stability chamber, which describes the transition into multiple $(N-1)$ **Q**-observables or **H**-observables. The transition of the surface observables can be expressed as two different integral transformations, which were shown to yield the separation of variables for both the XXX spin chain and the Gaudin model, respectively.

Separation of variables for $\mathfrak{gl}(n)$ XXX spin chain and Gaudin model with $n+2$ sites. A direct generalization of our formulation would involve the higher-rank linear A_{n-1}-quiver gauge theory, with $n > 2$. We expect that our prescription could be further generalized to the XXX spin chain associated with $Y(\mathfrak{gl}(n))$, steeming from the surface defect transition of the monodromy surface defect to $(n-1)(N-1)$ **Q**-observables. It would be interesting to

connect our construction to more recent studies (see [81, 82] for instance) of the quantum separation of variables for XXX spin chain, and thereby to generalize the prescription to bi-infinite modules. In the bispectral dual perspective, we expect that our prescription would generalize the quantum separation of variables for the $\mathfrak{sl}(N)$ Gaudin model to the one defined on $n+2$ sites. Due to technicalities, we postpone this to future work.

Quantum analytic Langlands correspondence and KZ/BPZ correspondence. In a recent study [83], it was suggested that the separation of variables for the common eigenfunction of the quantum Hitchin Hamiltonians provides a unitary map between properly defined L^2-spaces, which is well-suited for the study of analytic Langlands correspondence [70, 71, 84]. Further, in their study of quantum deformation of the analytic Langlands correspondence [83], which in our setting corresponds to the deviation $\varepsilon_2 \neq 0$ along a certain line, the twisted $\widehat{\mathfrak{sl}}(2)$-coinvariants at $\kappa \neq 0$ form a Hilbert space of states, on which the KZ/BPZ correspondence is applied to analyze the wavefunction by using Liouville correlation functions with degenerate fields. Our integral transformations for the vev of the regular monodromy surface defect (i.e., twisted $\widehat{\mathfrak{sl}}(N)$-coinvariants at $\kappa = -\frac{\varepsilon_2}{\varepsilon_1} \neq 0$) provide two generalizations of the KZ/BPZ correspondence: the $\hbar$-deformation and the higher N generalization. It would be interesting to apply these generalized integral transformations to investigate the (quantum) analytic ($\hbar$-)Langlands correspondence in our $\mathcal{N} = 2$ theory setup. This would require placing the $\mathcal{N} = 2$ theory on $\mathbb{P}^1 \times \mathbb{R}^2$ with two (as opposed to one) fixed points of the isometry. See also [85] for the GL-twisted $\mathcal{N} = 4$ theory formulation of the analytic Langlands correspondence.

Separation of variables and 5d Chern-Simons theory. In the twisted M-theory setup [54], the monodromy surface defect in the $\mathcal{N} = 2$ theory corresponds to having an A_{N-1}-singularity in the topological part of the worldvolume, which can be deformed to the N-centered Taub-NUT space. Reducing along the Taub-NUT circle and localizing, we arrive at the 5d non-commutative $\mathfrak{gl}(N)$ Chern-Simons theory [54] with intersecting holomorphic surface defects. Our integral transformation suggests that this is dual to the 5d non-commutative $\mathfrak{gl}(1)$ Chern-Simons theory with intersecting holomorphic surface defects and additional line defects. It would be interesting to re-examine our integral transformations from this perspective.

A   Gauge origami constructions

The partition function of 4d A_1-quiver U(N) gauge theory, the **Q**-observables, and the monodromy surface defect can be constructed using the *gauge origami* [86], which is a configuration of intersecting D3-branes on $X \times \mathbb{C}$ in the IIB theory, where X is a Calabi-Yau four-fold. Here, we briefly recall this construction.

A.1 A_1-quiver gauge theory

We consider the gauge origami defined on $X = \mathbb{C}^2_{12} \times \mathbb{C}^2_{34}/\mathbb{Z}_3$, where the subscripts enumerates the four $\mathbb{C}$-planes. In particular, we take $3N$ D3-branes supported on $\mathbb{C}^2_{12}$, represented by

the character of their Chan-Paton bundle by

$$N_{12} = \sum_{\alpha=1}^{N} e^{a_\alpha} \cdot \mathcal{R}_0 + \sum_{\alpha=1}^{N} e^{m_\alpha^- - \varepsilon_4} \cdot \mathcal{R}_1 + \sum_{\alpha=1}^{N} e^{m_\alpha^+ - \varepsilon_3} \cdot \mathcal{R}_2, \tag{A.1}$$

where $\mathcal{R}_i$ denotes the i-th representation of $\mathbb{Z}_3$ assigned to the D3-branes. Note that there are N D3-branes for each representation $\mathcal{R}_i$, $i = 0, 1, 2$. The spiked instanton partition function reads

$$\mathcal{Z}_S = \sum_{\boldsymbol{\lambda}} \prod_{i=0,1,2} \mathfrak{q}_i^{|\boldsymbol{\lambda}_i|} \mathbb{E}\left[-\frac{P_3 S_{12} S_{12}^*}{P_{12}^*} \right]^{\mathbb{Z}_3}$$
$$= \sum_{\boldsymbol{\lambda}} \prod_{i=0,1,2} \mathfrak{q}_i^{|\boldsymbol{\lambda}_i|} \mathbb{E}\left[\frac{-S_0 S_0^* - S_1 S_1^* - S_2 S_2^* + S_0 S_1^* + q_3^3 S_1 S_2^* + q_{12}^{-1} S_2 S_0^*}{P_{12}^*} \right] \tag{A.2}$$

where we used the notation

$$q_a = e^{\varepsilon_a}, \quad P_a = 1 - q_a, \qquad a, b = 1, 2, 3, 4, \tag{A.3}$$
$$P_{ab} = P_a P_b,$$

and defined the universal sheaf $S_i = N_i - P_{12} K_i$ for the instanton associated to the i-th node, $i = 0, 1, 2$, where K_i is the vector space encoding the fixed points of the moduli space of instantons for each stack of N D3-branes. For instance,

$$K_0 = \sum_{\square_{(i,j)}^{(\alpha)} \in \boldsymbol{\lambda}_0} e^{a_\alpha + (i-1)\varepsilon_1 + (j-1)\varepsilon_2}, \tag{A.4}$$

with each $\boldsymbol{\lambda}_i$, $i = 0, 1, 2$, being an N-tuple of Young diagrams. $\mathfrak{q}_i$, $i = 0, 1, 2$ are gauge couplings for each stack. We used the symbol $\mathbb{E}$ to convert characters into products of equivariant weights,

$$\mathbb{E}\left[\sum_i e^{x_i} - \sum_j e^{y_j} \right] = \frac{\prod_j y_j}{\prod_i x_i}. \tag{A.5}$$

This is precisely the partition function of the $\hat{A}_2$-quiver U(N) gauge theory.

We may now turn off the gauge coupling $\mathfrak{q}_1 = \mathfrak{q}_2 = 0$ while keeping $\mathfrak{q}_0 = \mathfrak{q}$. This is known as the *freezing limit*. Some of the universal bundle becomes the flavor bundle $S_1 = M^+$, $S_2 = M^-$, $S_0 = S$, so that the spike instanton partition function becomes the partition function of U(N) gauge theory with N fundamental and N anti-fundamental hypermultiplets:

$$\mathcal{Z} = \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathbb{E}\left[\frac{-SS^* + M^+ S^* + q_{12}^{-1} S(M^-)^*}{P_{12}^*} \right] \tag{A.6}$$

with $M = M^+ + M^-$. Here, S is the universal sheaf for the instantons, and $M^\pm$ are the flavor bundles. We sometimes use the notation

$$\bar{a} = \frac{1}{N} \sum_{\alpha=0}^{N-1} a_\alpha, \qquad \bar{m}^\pm = \frac{1}{N} \sum_{\alpha=0}^{N-1} m_\alpha^\pm, \tag{A.7}$$

to denote the average of the Coulomb parameters and the masses.

A.2 Q-observable

We consider the gauge origami set up:

$$N_{12} = \sum_{\alpha=1}^{N} e^{a_\alpha} \cdot \mathcal{R}_0 + \sum_{\alpha=1}^{N} e^{m_\alpha^- - \varepsilon_4} \cdot \mathcal{R}_1 + \sum_{\alpha=1}^{N} e^{m_\alpha^+ - \varepsilon_3} \cdot \mathcal{R}_2$$
$$N_{14} = e^{x + \varepsilon_1 + \varepsilon_4} \cdot \mathcal{R}_2,$$

(A.8)

by adding one more D3-brane. The spike instanton partition function is

$$\mathcal{Z} = \sum_{\boldsymbol{\lambda}} \prod_{i=0,1,2} \mathfrak{q}_i^{|\boldsymbol{\lambda}_{12,i}| + |\boldsymbol{\lambda}_{14,i}|} \mathbb{E}\left[-\frac{P_3 S_{12} S_{12}^*}{P_{12}^*} - \frac{P_2 S_{13} S_{13}^*}{P_{13}^*} + q_2 P_3 \frac{S_{14} S_{12}^*}{P_1^*} \right]^{\mathbb{Z}_3}$$

(A.9)

We take the freezing limit $\mathfrak{q}_1 = \mathfrak{q}_2 = 0$, $\mathfrak{q}_0 = \mathfrak{q}$ which yields the A_1-quiver:

$$\mathcal{Z} = \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathbb{E}\left[-\frac{SS^* - M^+ S^* - q_{12}^* S(M^-)^*}{P_{12}^*} \right] \mathbb{E}\left[-\frac{e^x (S - M^+)^*}{P_1^*} \right]$$
$$= \sum_{\boldsymbol{\lambda}} \mathcal{Z}[\boldsymbol{\lambda}] \mathcal{Q}(x)[\boldsymbol{\lambda}]$$

(A.10)

The **Q**-observable associated to an instanton configuration $\boldsymbol{\lambda}$ is given by

$$\mathcal{Q}(x)[\boldsymbol{\lambda}] = \mathbb{E}\left[\frac{e^x (M^+ - S[\boldsymbol{\lambda}])^*}{P_1^*} \right].$$

(A.11)

We often separate the 1-loop part with the Γ-functions,

$$M^+(x) = \mathbb{E}\left[-\frac{e^x M^+}{P_1^*} \right] = \prod_{\alpha=0}^{N-1} \frac{(-\varepsilon_1)^{\frac{x - m_\alpha^+}{\varepsilon_1}}}{\Gamma\left(-\frac{x - m_\alpha^+}{\varepsilon_1} \right)},$$

(A.12)

and write

$$\mathbf{Q}(x) = \mathbb{E}\left[-\frac{e^x S[\boldsymbol{\lambda}]^*}{P_1^*} \right] = \prod_{\alpha=0}^{N-1} \left(\frac{\varepsilon_1^{\frac{x - a_\alpha}{\varepsilon_1}}}{\Gamma\left(-\frac{x - a_\alpha}{\varepsilon_1} \right)} \prod_{i=1}^{\infty} \frac{x - a_\alpha - (i-1)\varepsilon_1 - \lambda_i^{(\alpha)} \varepsilon_2}{x - a_\alpha - (i-1)\varepsilon_1} \right)$$
$$= \prod_{\alpha=0}^{N-1} \left(\frac{\varepsilon_1^{\frac{x - a_\alpha}{\varepsilon_1}}}{\Gamma\left(l(\lambda^{(\alpha)}) - \frac{x - a_\alpha}{\varepsilon_1} \right)} \prod_{i=1}^{l(\lambda^{(\alpha)})} \frac{x - a_\alpha - (i-1)\varepsilon_1 - \lambda_i^{(\alpha)} \varepsilon_2}{\varepsilon_1} \right).$$

(A.13)

Namely, $\mathcal{Q}(x) = \frac{\mathbf{Q}(x)}{M^+(x)}$. Somewhat loosely, we call both $\mathcal{Q}(x)$ and $\mathbf{Q}(x)$ the **Q**-observable.

A.3 $\tilde{\mathbf{Q}}$-observable

We consider the following gauge origami setup where the $\mathbb{Z}_3$ charge assignment for the $\hat{N}_{14}$ is slightly different:

$$\hat{N}_{12} = \sum_{\alpha=1}^{N} e^{a_\alpha} \cdot \mathcal{R}_0 + \sum_{\alpha=1}^{N} e^{m_\alpha^- - \varepsilon_4} \cdot \mathcal{R}_1 + \sum_{\alpha=1}^{N} e^{m_\alpha^+ - \varepsilon_3} \mathcal{R}_2,$$
$$\hat{N}_{14} = e^x q_1 \cdot \mathcal{R}_0.$$

(A.14)

We impose freezing $\mathfrak{q}_1 = \mathfrak{q}_2 = 0$. Then $\hat{K}_{14}$ grows only along the direction of $\mathbb{C}_1$, yielding

$$P_1\hat{K}_{14} = e^x q_1(1 - q_1^d), \tag{A.15}$$

where $d = |\hat{K}_{13}|$. The spiked instanton partition function gives

$$\begin{aligned}
\mathcal{Z}_S &= \sum_{\boldsymbol{\lambda}} \sum_{d=0}^{\infty} \mathfrak{q}^{|\boldsymbol{\lambda}|+d} \mathbb{E}\left[-\frac{\hat{P}_3 \hat{S}_{12} \hat{S}_{12}^*}{\hat{P}_{12}^*} - \frac{\hat{P}_2 \hat{S}_{14} \hat{S}_{14}^*}{\hat{P}_{14}^*} + \hat{q}_2 \hat{P}_3 \frac{\hat{S}_{14} \hat{S}_{12}^*}{\hat{P}_1^*} \right] \\
&= \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathcal{Z}[\boldsymbol{\lambda}] \sum_{d=0}^{\infty} \prod_{j=1}^{k} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_2} \frac{\mathcal{Q}(x)M(x + d\varepsilon_1)}{Q(x + d\varepsilon_1)Q(x + d\varepsilon_1 + \varepsilon_+)} \\
&=: \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\boldsymbol{\lambda}|} \mathcal{Z}[\boldsymbol{\lambda}] \tilde{Q}(x)[\boldsymbol{\lambda}].
\end{aligned} \tag{A.16}$$

The $\tilde{Q}$-observable associated to instanton configuration $\boldsymbol{\lambda}$ is

$$\begin{aligned}
\tilde{\mathcal{Q}}(x) &= \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \prod_{j=1}^{k} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_2} \frac{\mathcal{Q}(x)[\boldsymbol{\lambda}]M(x + d\varepsilon_1)}{Q(x + d\varepsilon_1)[\boldsymbol{\lambda}]Q(x + (d+1)\varepsilon_1 + \varepsilon_2)[\boldsymbol{\lambda}]} \\
&= \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \left(\prod_{j=1}^{d} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_1} \frac{P^+(x + j\varepsilon_1)}{\mathcal{Y}(x + j\varepsilon_1)} \right) \frac{M^-(x + d\varepsilon_1)}{Q(x + \varepsilon_2 + (d+1)\varepsilon_1)}.
\end{aligned} \tag{A.17}$$

We often separate the 1-loop part $M^+(x) = \mathbb{E}\left[-\frac{e^x M^+}{P_1^*} \right]$ and write

$$\tilde{\mathbf{Q}}(x) = \mathfrak{q}^{\frac{x}{\varepsilon_1}} \sum_{d=0}^{\infty} \mathfrak{q}^d \left(\prod_{j=1}^{d} \frac{j\varepsilon_1 + \varepsilon_2}{j\varepsilon_1} \frac{1}{\mathcal{Y}(x + j\varepsilon_1)} \right) \frac{M(x + d\varepsilon_1)}{Q(x + \varepsilon_2 + (d+1)\varepsilon_1)}. \tag{A.18}$$

Namely, $\tilde{\mathcal{Q}}(x) = \frac{\tilde{\mathbf{Q}}(x)}{M^+(x)}$. Somewhat loosely, we call both $\tilde{\mathcal{Q}}(x)$ and $\tilde{\mathbf{Q}}(x)$ the $\tilde{\mathbf{Q}}$-observable.

A.4 Monodromy surface defect

Consider the gauge origami (A.1) in the presence of regular monodromy defect:

$$\hat{N}_{12} = \sum_{\omega=0}^{N-1} e^{a_\omega} \hat{q}_2^\omega \mathcal{R}_0 \otimes \mathfrak{R}_\omega + q_{13} M_\omega^- \hat{q}_2^{\omega+1} \mathcal{R}_1 \otimes \mathfrak{R}_{\omega+1} + q_3^{-1} M_\omega^+ \hat{q}_2^\omega \mathcal{R}_2 \otimes \mathfrak{R}_\omega \tag{A.19}$$

The universal sheaf of the instanton on the orbifold, carrying $\mathbb{Z}_N$ representation, can be written as

$$\hat{S}_{12} = \sum_{\omega=0}^{N-1} S_\omega \hat{q}_2^\omega \mathcal{R}_0 \otimes \mathfrak{R}_\omega + q_{13} M_\omega^- \hat{q}_2^{\omega+1} \mathcal{R}_1 \otimes \mathfrak{R}_{\omega+1} + q_3^{-1} M_\omega^+ \hat{q}_2^\omega \mathcal{R}_2 \otimes \mathfrak{R}_\omega \tag{A.20}$$

where we defined $S_\omega = N_\omega - P_1 K_\omega + q_2^{\delta_{\omega,0}} P_1 K_{\omega-1}$, $M_\omega^\pm = e^{m_\omega^\pm}$. We get the universal sheaf for the instantons in the absence of the orbifold and the flavor bundle by

$$S = \sum_{\omega=0}^{N-1} S_\omega = N - P_{12} K_{N-1}, \quad M^\pm = \sum_{\omega=0}^{N-1} M_\omega^\pm. \tag{A.21}$$

The spiked instanton partition function is given by

$$
\begin{aligned}
\hat{\mathcal{Z}} &= \sum_{\hat{\boldsymbol{\lambda}}} \prod_{\omega=0}^{N-1} \mathfrak{q}_\omega^{|K_\omega|} \mathbb{E}\left[-\frac{\hat{P}_3 \hat{S}_{12} \hat{S}_{12}^*}{\hat{P}_{12}^*}\right]^{\mathbb{Z}_3 \times \mathbb{Z}_N} \\
&= \sum_{\boldsymbol{\lambda}} \mathfrak{q}^{|\lambda|} \mathbb{E}\left[\frac{-SS^* + M^+ S^*}{P_{12}^*} + \frac{S(M^-)^*}{P_{12}}\right] \\
&\quad \times \sum_{\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{k_\omega - k_{N-1}} \mathbb{E}\left[\sum_{\omega_1 < \omega_2} \frac{S_{\omega_1} S_{\omega_2}^*}{P_1^*} - \frac{M_{\omega_1}^+ S_{\omega_2}^*}{P_1^*} - \frac{S_{\omega_2}(M_{\omega_1}^-)^*}{P_1}\right] \\
&= \langle \Psi_{\mathbf{c}}(\mathbf{u})\rangle \mathcal{Z}
\end{aligned}
\tag{A.22}
$$

Here, ρ is a projection from colored partitions $\boldsymbol{\Lambda} = (\Lambda^{(0)}, \ldots, \Lambda^{(N-1)})$ to ordinary partitions $\boldsymbol{\lambda} = (\lambda^{(0)}, \ldots, \lambda^{(N-1)})$, $\rho : \{\hat{\boldsymbol{\lambda}}\} \to \{\boldsymbol{\lambda}\}$, defined by

$$
\rho(\Lambda_i^{(\alpha)}) = \left\lfloor \frac{\Lambda_i^{(\alpha)} + c(\alpha)}{N} \right\rfloor,
\tag{A.23}
$$

where $\lfloor \cdots \rfloor$ is the floor function. Define

$$
\Sigma_\omega = S_{\omega+1} + \cdots + S_{N-1}, \quad \Sigma_{\omega-1} - \Sigma_\omega = S_\omega.
\tag{A.24}
$$

The surface observable is obtained as

$$
\Psi_{\mathbf{c}}(\mathbf{u}) = \sum_{\boldsymbol{\Lambda} \in \rho^{-1}(\boldsymbol{\lambda})} \prod_{\omega=0}^{N-2} \mathfrak{q}_\omega^{k_\omega - k_{N-1}} \mathbb{E}\left[\sum_{\omega=0}^{N-2} \frac{(S - \Sigma_\omega)(\Sigma_\omega - \Sigma_{\omega+1})^*}{P_1^*} - \frac{M_\omega^+ \Sigma_\omega^*}{P_1^*} - \frac{\Sigma_\omega(M_\omega^-)^*}{P_1}\right].
\tag{A.25}
$$

B Separation of variable kernel $\mathcal{K}(\mathbf{u}; \mathbf{y})$ for general N

It will be convenient for us to set $x_{N-\omega, N-j}$, $\omega = 1, \ldots, N-1$, $j = 1, \ldots, \omega$, and proceed starting from $\omega = 1$ up to $\omega = N-1$.

We showed explicit computations for $\omega = 1$ in section 4.4.3. Now, there will be $\omega + 1$ Γ-functions in the numerator for $x_{N-\omega, N-j}$, $\omega = 2, \ldots, N-1$ taking the form of

$$
\prod_{i=1}^{\omega+1} \Gamma\left(\frac{x_{N-\omega, N-j} - x_{N-\omega-1, N-k}}{\varepsilon_1}\right)
$$

The assignment can be represented by a $\omega \times (\omega + 1)$ rectangle matrix $M_{j,k}^{(\omega)}$, whose matrix components are either 0 or 1, satisfying the following condition:

- $M_{j,j}^{(\omega)} = 0$ for $j = 1, \ldots, \omega$.

- The number of 1 in each column must be even.

- The number of 1 in each row must be $\omega - 1$.

- The total number of non-zero entries in $M_{j,k}^{(\omega)}$ is $\omega(\omega - 1)$.

- There can not be two columns in $M_{j,k}^{(\omega)}$ that are identical when $\omega > 2$.

This will leave us with the only two zero-entries on each row of $M_{j,k}^{(\omega)}$: $k = j$, $k_j^{(\omega)}$ for some $k_j^{(\omega)}$.

We list the first few $M^{(\omega)}$:

$$M^{(2)} : \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \tag{B.1a}$$

$$M^{(3)} : \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}; \tag{B.1b}$$

$$M^{(4)} : \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}; \tag{B.1c}$$

Notice that when ω is even, all components of $M_{j,\omega+1}^{(\omega)}$ must be 1. The argument is the following: In the sub-square matrix, the maximal number of 1's are $\omega(\omega - 2)$ as the number of 1 in a single column must be even.

Note that a matrix satisfying all the first 4 conditions but not the last one does not give a legitimate pairing. The first of such matrix occurs at $\omega = 4$:

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{B.2}$$

The third/fourth column is the same as the first/second column.

We want to use (4.110) such that all $x_{N-\omega-1,N-k}$ should only show up in the Γ-function and powers of integration variables, which would require $c - a = d - b$. To guarantee it always happens, we multiply

$$1 = \prod_{j=1}^{\omega} \left[\frac{\Gamma\left(\frac{x_{N-\omega,N-j-f^{(\omega)}}}{\varepsilon_1} \right)}{\Gamma\left(\frac{x_{N-\omega,N-j-f^{(\omega)}}}{\varepsilon_1} \right)} \right]^{\omega-1}. \tag{B.3}$$

The pairing is done in the following way:

- Define a set $\mathbb{S}_k^{(\omega)} := \{j | M_{j,k}^{(\omega)} = 1\}$ that collects the rows in the k-th column of the matrix $M_{j,k}^{(\omega)}$ element being 1. The size of the set $\mathbb{S}_k^{(\omega)}$ is always even according to how $M_{j,k}^{(\omega)}$ is defined.

- We pair $(j, j')_k$, $j < j'$, of all elements $j, j' \in \mathbb{S}_k^{(\omega)}$ for all $k = 1, \ldots, \omega + 1$. (Notice that by the construction of $M_{j,k}^{(\omega)}$, $j \neq k \neq j'$)

- One can not make a pair $(j, j')_k$ for $\mathbb{S}_k^{(\omega)}$ if a pairing $(j, j')_{k'}$ already exist for another $\mathbb{S}_{k'}^{(\omega)}$, $k \neq k'$. We obtain a collection of paring $\mathbb{B}^{(\omega)}$ for each ω. $|\mathbb{B}^{(\omega)}| = \frac{\omega(\omega-1)}{2}$.

- For a pair $(j, j')_k$ defined, we will perform the integral transformation (4.110) on

$$
\frac{\Gamma\left(\frac{x_{N-\omega,N-j}-x_{N-\omega-1,N-k}}{\varepsilon_1}\right)\Gamma\left(\frac{x_{N-\omega,N-j'}-x_{N-\omega-1,N-k}}{\varepsilon_1}\right)}{\Gamma\left(\frac{x_{N-\omega,N-j}-f^{(\omega)}}{\varepsilon_1}\right)\Gamma\left(\frac{x_{N-\omega,N-j'}-f^{(\omega)}}{\varepsilon_1}\right)}\frac{x_{N-\omega,N-j}-x_{N-\omega,N-j'}}{\varepsilon_1}
$$

$$
=\frac{1}{\Gamma\left(\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}\right)\Gamma\left(\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}-1\right)}
$$

$$
\times \int_{\mathrm{PC}^{\times 2}} t_{(j,j')_k,1}^{\frac{x_{N-\omega,N-j}-x_{N-\omega-1,N-k}}{\varepsilon_1}-1}\, t_{(j,j')_k,2}^{\frac{x_{N-\omega,N-j'}-x_{N-\omega-1,N-k}}{\varepsilon_1}}
$$

$$
\times\left[(1-t_{(j,j')_k,1})(1-t_{(j,j')_k,2})\right]^{\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}-2}(t_{(j,j')_k,1}-t_{(j,j')_k,2})\,dt_{(j,j')_k,1}\,dt_{(j,j')_k,2}
$$

(B.4)

By performing (4.110) to all the pairing generated by $M_{j,k}^{(\omega)}$ in $\mathbb{B}^{(\omega)}$

$$
\prod_{k=1}^{\omega+1}\left[\Gamma\left(\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}\right)\Gamma\left(\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}-1\right)\right]^{-\frac{1}{2}\sum_{j=1}^{\omega}M_{j,k}^{(\omega)}}
$$

$$
\times \prod_{(j,j')_k\in\mathbb{B}^{(\omega)}}\int_{\mathrm{PC}^{\times 2}} t_{(j,j')_k,1}^{\frac{x_{N-\omega,N-j}-x_{N-\omega-1,N-k}}{\varepsilon_1}-1}\, t_{(j,j')_k,2}^{\frac{x_{N-\omega,N-j'}-x_{N-\omega-1,N-k}}{\varepsilon_1}}
$$

(B.5)

$$
\times\left[(1-t_{(j,j')_k,1})(1-t_{(j,j')_k,2})\right]^{\frac{x_{N-\omega-1,N-k}-f^{(\omega)}}{\varepsilon_1}-2}(t_{(j,j')_k,1}-t_{(j,j')_k,2})\,dt_{(j,j')_k,1}\,dt_{(j,j')_k,2}
$$

The integration degree is $2\times|\mathbb{B}^{(\omega)}|=\omega(\omega-1)$.

Finally we will use the ensemble for $x_{N-\omega,N-j} = x_{N-\omega-1,j} - s_{N-j,N-\omega}\varepsilon_1$ dependence, we denote $F_{\omega,j}(\mathbf{t}_{\omega,j})$ of the function in all previous $t_{(j,j')_k,l}$, $l = 1,2$ that has $x_{N-\omega,N-j}$ as its power:

$$
\prod_{j=1}^{\omega} \sum_{s_{N-j,N-\omega}=0}^{\infty} (\mathfrak{q}_\omega \cdots \mathfrak{q}_j F_{\omega,j}(\mathbf{t}_{\omega,j}))^{-\frac{x_{N-\omega-1,N-j}}{\varepsilon_1} + s_{N-j,N-\omega}}
$$

$$
\frac{\Gamma\left(1 + s_{N-j,N-\omega} - \frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^-}{\varepsilon_1}\right) \Gamma\left(-s_{N-j,N-\omega} + \frac{x_{N-\omega-1,N-j} - x_{N-\omega-1,N-k_j^{(\omega)}}}{\varepsilon_1}\right)}{\Gamma\left(1 + s_{N-j,N-\omega}\right) \Gamma\left(-n_{\omega,j} + \frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^+}{\varepsilon_1}\right)}
$$

$$
= \prod_{j=1}^{\omega} \Gamma\left(1 - \frac{y_{\omega+1,j} - m_\omega^+}{\varepsilon_1}\right) \Gamma\left(1 - \frac{y_{\omega+1,k_j^{(\omega)}} - m_\omega^-}{\varepsilon_1}\right)
$$

$$
\times \int_{\mathrm{PC}} t_{\omega,j,0}^{-\frac{y_{\omega+1,j} - m_\omega^-}{\varepsilon_1}} (1 - t_{\omega,j,0})^{-\frac{y_{\omega+1,k_j^{(\omega)}} - m_j^+}{\varepsilon_1} - 1} (1 - \mathfrak{q}_\omega \ldots \mathfrak{q}_j F_{\omega,j}(\mathbf{t}_{\omega,j}) t_{\omega,j,0})^{\frac{y_{\omega+1,j} - m_\omega^+}{\varepsilon_1} - 1} \, dt_{\omega,j,0}
$$

$$
= \frac{e^{-\pi i \left(1 + \frac{x_{N-2,N-2} - x_{N-2,N-1}}{\varepsilon_1}\right)}}{4\pi} \frac{\Gamma\left(\frac{m_{N-\omega-1}^- - x_{N-\omega-1,N-k_j^{(\omega)}}}{\varepsilon_1}\right) \Gamma\left(\frac{m_{N-\omega-1}^+ - x_{N-\omega-1,N-j}}{\varepsilon_1}\right)}{\sin\pi\left(\frac{x_{N-\omega-1,N-j} - x_{N-\omega-1,N-k_j^{(\omega)}}}{\varepsilon_1}\right)}
$$

$$
\times \int_{\mathrm{PC}} t_{\omega,j,0}^{-\frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^-}{\varepsilon_1} - 1} (1 - t_{\omega,j,0})^{-\frac{x_{N-\omega-1,N-k_j^{(\omega)}} - m_{N-\omega-1}^+}{\varepsilon_1}}
$$

$$
(1 - \mathfrak{q}_\omega \ldots \mathfrak{q}_j F_{\omega,j}(\mathbf{t}_{\omega,j}) t_{\omega,j,0})^{\frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^+}{\varepsilon_1}} \, dt_{\omega,j,0}
$$

(B.6)

The Γ-factor can be organized to

$$
\prod_{j=1}^{\omega} \Gamma\left(-\frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^+}{\varepsilon_1}\right) \Gamma\left(-\frac{x_{N-\omega-1,k_j^{(\omega)}} - m_{N-\omega-1}^-}{\varepsilon_1}\right)
$$

$$
= \prod_{j=1}^{\omega} \Gamma\left(-\frac{x_{N-\omega-1,N-j} - m_{N-\omega-1}^+}{\varepsilon_1}\right) \prod_{i=1}^{\omega+1} \Gamma\left(-\frac{x_{N-\omega-1,N-i} - m_{N-\omega-1}^-}{\varepsilon_1}\right)^{\sum_{j=1, j\neq i}^{\omega} (1 - M_{j,i}^{(\omega)})}
$$

(B.7)

Up to this moment, we still have Γ-factors generated from the last level $\omega - 1$:

$$
\frac{\prod_{j=1}^{\omega-1} \Gamma\left(-\frac{x_{N-\omega,N-j} - m_{N-\omega}^+}{\varepsilon_1}\right) \prod_{k=1}^{\omega} \Gamma\left(-\frac{x_{N-\omega,N-k} - m_{N-\omega}^-}{\varepsilon_1}\right)^{\sum_{j=1, j\neq k}^{\omega-1} (1 - M_{j,k}^{(\omega-1)})}}{\prod_{k=1}^{\omega} \left[\Gamma\left(\frac{x_{N-\omega,N-k} - f^{(\omega-1)}}{\varepsilon_1}\right) \Gamma\left(\frac{x_{N-\omega,N-k} - f^{(\omega-1)}}{\varepsilon_1} - 1\right)\right]^{\frac{1}{2}\sum_{j=1}^{\omega-1} M_{j,k}^{(\omega-1)}}} \prod_{k=1}^{\omega} \Gamma\left(\frac{x_{N-\omega,N-k} - f^{(\omega)}}{\varepsilon_1}\right)^{\omega-1},
$$

(B.8)

which finally can be rewritten in $\omega \times (\omega - 1)$ integrals using the Pochhammer contour integral (4.56). This can be seen by counting the numbers of Γ-function in both the numerator

(one need to use the reflection formula for the Γ-functions with negative arguments in $x_{N-\omega,N-j}$ before proper counting). The total number of Γ-functions in the denominator with positive argument of $x_{N-\omega,N-j}$ and Γ-functions with negative argument of $x_{N-\omega,N-j}$ in the numerator is

$$
\begin{aligned}
&2 \times \sum_{j=1}^{\omega-1} \sum_{k=1}^{\omega} \frac{1}{2} M_{j,k}^{(\omega-1)} + \sum_{k=1}^{\omega} \sum_{j=1,j\neq k}^{\omega-1} (1 - M_{j,k}^{(\omega-1)}) + \omega - 1 \\
&= (\omega-1)(\omega-2) + (\omega-1)^2 - (\omega-1)(\omega-2) + \omega - 1 \\
&= \omega(\omega-1).
\end{aligned}
\tag{B.9}
$$

It equals to the total number of Γ-function in the numerator with positive argument of $x_{N-\omega,N-k}$. The so-generated Γ-functions are not $x_{N-\omega,N-j}$ dependent.

Data Availability Statement. This article has no associated data or the data will not be deposited.

Code Availability Statement. This article has no associated code or the code will not be deposited.

References

[1] D. Gaiotto, *N=2 dualities*, *JHEP* **08** (2012) 034 [arXiv:0904.2715] [INSPIRE].

[2] D. Gaiotto, G.W. Moore and A. Neitzke, *Wall-crossing, Hitchin systems, and the WKB approximation*, *Adv. Math.* **234** (2013) 239 [arXiv:0907.3987] [INSPIRE].

[3] E. Witten, *Topological Quantum Field Theory*, *Commun. Math. Phys.* **117** (1988) 353 [INSPIRE].

[4] N.A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, *Adv. Theor. Math. Phys.* **7** (2003) 831 [hep-th/0206161] [INSPIRE].

[5] M. Bershadsky, C. Vafa and V. Sadov, *D-branes and topological field theories*, *Nucl. Phys. B* **463** (1996) 420 [hep-th/9511222] [INSPIRE].

[6] J.A. Harvey, G.W. Moore and A. Strominger, *Reducing S duality to T duality*, *Phys. Rev. D* **52** (1995) 7161 [hep-th/9501022] [INSPIRE].

[7] N. Nekrasov and E. Witten, *The Omega Deformation, Branes, Integrability, and Liouville Theory*, *JHEP* **09** (2010) 092 [arXiv:1002.0888] [INSPIRE].

[8] S. Jeong, N. Lee and N. Nekrasov, *Parallel surface defects, Hecke operators, and quantum Hitchin system*, arXiv:2304.04656 [INSPIRE].

[9] S. Gukov and E. Witten, *Branes and Quantization*, *Adv. Theor. Math. Phys.* **13** (2009) 1445 [arXiv:0809.0305] [INSPIRE].

[10] A. Kapustin and D. Orlov, *Remarks on A branes, mirror symmetry, and the Fukaya category*, *J. Geom. Phys.* **48** (2003) 84 [hep-th/0109098] [INSPIRE].

[11] A. Kapustin and E. Witten, *Electric-Magnetic Duality And The Geometric Langlands Program*, *Commun. Num. Theor. Phys.* **1** (2007) 1 [`hep-th/0604151`] [ɪɴSPIRE].

[12] A. Beilinson and Y. Drinfeld, *Quantization of Hitchin's Integrable System and Hecke Eigensheaves* [ɪɴSPIRE].

[13] A. Beilinson and V. Drinfeld, *Opers*, `math/0501398`.

[14] S. Jeong and N. Nekrasov, *Opers, surface defects, and Yang-Yang functional*, *Adv. Theor. Math. Phys.* **24** (2020) 1789 [`arXiv:1806.08270`] [ɪɴSPIRE].

[15] E. Frenkel, D. Gaitsgory and K. Vilonen, *On the geometric Langlands conjecture*, `math/0012255`.

[16] E. Frenkel, *Ramifications of the geometric Langlands Program*, `math/0611294` [ɪɴSPIRE].

[17] S. Gukov and E. Witten, *Gauge Theory, Ramification, And The Geometric Langlands Program*, `hep-th/0612073` [ɪɴSPIRE].

[18] D. Gaiotto, *S-duality of boundary conditions and the geometric Langlands program*, *Proc. Symp. Pure Math.* **98** (2018) 139 [`arXiv:1609.09030`] [ɪɴSPIRE].

[19] E. Frenkel and D. Gaiotto, *Quantum Langlands dualities of boundary conditions, D-modules, and conformal blocks*, *Commun. Num. Theor. Phys.* **14** (2020) 199 [`arXiv:1805.00203`] [ɪɴSPIRE].

[20] E. Frenkel, *Lectures on the Langlands program and conformal field theory*, in the proceedings of the *Les Houches School of Physics: Frontiers in Number Theory, Physics and Geometry*, Les Houches, France, March 09–21 (2003) [`DOI:10.1007/978-3-540-30308-4_11`] [`hep-th/0512172`] [ɪɴSPIRE].

[21] J. Teschner, *Quantization of the Hitchin moduli spaces, Liouville theory, and the geometric Langlands correspondence I*, *Adv. Theor. Math. Phys.* **15** (2011) 471 [`arXiv:1005.2846`] [ɪɴSPIRE].

[22] A. Balasubramanian and J. Teschner, *Supersymmetric field theories and geometric Langlands: the other side of the coin*, *Proc. Symp. Pure Math.* **98** (2018) 79 [`arXiv:1702.06499`] [ɪɴSPIRE].

[23] J. Teschner, *Quantisation conditions of the quantum Hitchin system and the real geometric Langlands correspondence*, `arXiv:1707.07873` [ɪɴSPIRE].

[24] S. Jeong, N. Lee and N. Nekrasov, *di-Langlands correspondence and extended observables*, *JHEP* **06** (2024) 105 [`arXiv:2402.13888`] [ɪɴSPIRE].

[25] S.A. Cherkis and A. Kapustin, *Nahm transform for periodic monopoles and N=2 superYang-Mills theory*, *Commun. Math. Phys.* **218** (2001) 333 [`hep-th/0006050`] [ɪɴSPIRE].

[26] S.A. Cherkis and A. Kapustin, *Periodic monopoles with singularities and N=2 super QCD*, *Commun. Math. Phys.* **234** (2003) 1 [`hep-th/0011081`] [ɪɴSPIRE].

[27] N. Nekrasov and V. Pestun, *Seiberg-Witten Geometry of Four-Dimensional $\mathcal{N} = 2$ Quiver Gauge Theories*, *SIGMA* **19** (2023) 047 [`arXiv:1211.2240`] [ɪɴSPIRE].

[28] N. Nekrasov, V. Pestun and S. Shatashvili, *Quantum geometry and quiver gauge theories*, *Commun. Math. Phys.* **357** (2018) 519 [`arXiv:1312.6689`] [ɪɴSPIRE].

[29] C. Elliott and V. Pestun, *Multiplicative Hitchin systems and supersymmetric gauge theory*, *Selecta M* **ath.** (2019) 1 [`arXiv:1812.05516`] [ɪɴSPIRE].

[30] B. Charbonneau and J. Hurtubise, *Singular Hermitian–Einstein Monopoles on the Product of a Circle and a Riemann Surface*, *Int. Math. Res. Not.* **2011** (2010) 175.

[31] E. Frenkel and N. Reshetikhin, *Quantum affine algebras and deformations of the Virasoro and 237-1237-1237-1*, *Commun. Math. Phys.* **178** (1996) 237.

[32] M. Aganagic, E. Frenkel and A. Okounkov, *Quantum q-Langlands Correspondence*, *Trans. Moscow Math. Soc.* **79** (2018) 1 [arXiv:1701.03146] [InSPIRE].

[33] P. Koroteev, D.S. Sage and A.M. Zeitlin, *(SL(N),q)-Opers, the q-Langlands Correspondence, and Quantum/Classical Duality*, *Commun. Math. Phys.* **381** (2021) 641 [arXiv:1811.09937] [InSPIRE].

[34] E. Frenkel, P. Koroteev, D.S. Sage and A.M. Zeitlin, *q-opers, QQ-systems, and Bethe Ansatz*, *J. Eur. Math. Soc.* **26** (2023) 355 [arXiv:2002.07344] [InSPIRE].

[35] N. Haouzi, *A new realization of quantum algebras in gauge theory and Ramification in the Langlands program*, arXiv:2311.04367 [InSPIRE].

[36] A. Gorsky, N. Nekrasov and V. Rubtsov, *Hilbert schemes, separated variables, and D-branes*, *Commun. Math. Phys.* **222** (2001) 299 [hep-th/9901089] [InSPIRE].

[37] E. Frenkel, *Affine algebras, Langlands duality and Bethe ansatz*, in the proceedings of the *11th International Conference on Mathematical Physics (ICMP-11) (Satellite colloquia: New Problems in the General Theory of Fields and Particles*, Paris, France, Jul 25–28 (1995) [q-alg/9506003] [InSPIRE].

[38] S. Gukov, *Surface operators*, in *New Dualities of Supersymmetric Gauge Theories*, J. Teschner ed., (2016), p. 223–259 [DOI:10.1007/978-3-319-18769-3_8] [arXiv:1412.7127] [InSPIRE].

[39] M. Finkelberg and L. Rybnikov, *Quantization of Drinfeld Zastava in type A*, *J. Eur. Math. Soc.* **16** (2014) 235.

[40] H. Nakajima, *Handsaw quiver varieties and finite W-algebras*, *Moscow Math. J.* **12** (2012) 633 [arXiv:1107.5073] [InSPIRE].

[41] N. Nekrasov, *BPS/CFT correspondence V: BPZ and KZ equations from qq-characters*, arXiv:1711.11582 [InSPIRE].

[42] N. Nekrasov, *Blowups in BPS/CFT Correspondence, and Painlevé VI*, *Annales Henri Poincare* **25** (2024) 1123 [arXiv:2007.03646] [InSPIRE].

[43] S. Jeong, N. Lee and N. Nekrasov, *Intersecting defects in gauge theory, quantum spin chains, and Knizhnik-Zamolodchikov equations*, *JHEP* **10** (2021) 120 [arXiv:2103.17186] [InSPIRE].

[44] S. Jeong, *Splitting of surface defect partition functions and integrable systems*, *Nucl. Phys. B* **938** (2019) 775 [arXiv:1709.04926] [InSPIRE].

[45] S. Jeong and N. Nekrasov, *Riemann-Hilbert correspondence and blown up surface defects*, *JHEP* **12** (2020) 006 [arXiv:2007.03660] [InSPIRE].

[46] A.V. Stoyanovsky, *A relation between the knizhnik-zamolodchikov and Belavin-Polyakov-Zamolodchikov systems of partial differential equations*, math-ph/0012013 [InSPIRE].

[47] S. Ribault and J. Teschner, *H+(3)-WZNW correlators from Liouville theory*, *JHEP* **06** (2005) 014 [hep-th/0502048] [InSPIRE].

[48] F. Nieri and Y. Zenkevich, *Quiver $W_{\epsilon_1,\epsilon_2}$ algebras of 4d $\mathcal{N} = 2$ gauge theories*, *J. Phys. A* **53** (2020) 275401 [arXiv:1912.09969] [InSPIRE].

[49] S. Ribault, *On sl(3) Knizhnik-Zamolodchikov equations and W(3) null-vector equations*, *JHEP* **10** (2009) 002 [arXiv:0811.4587] [InSPIRE].

[50] T. Creutzig, Y. Hikida and P.B. Ronne, *Correspondences between WZNW models and CFTs with W-algebra symmetry*, *JHEP* **02** (2016) 048 [arXiv:1509.07516] [InSPIRE].

[51] E. Frenkel, S. Gukov and J. Teschner, *Surface Operators and Separation of Variables*, *JHEP* **01** (2016) 179 [arXiv:1506.07508] [INSPIRE].

[52] N. Nekrasov, *BPS/CFT correspondence: non-perturbative Dyson-Schwinger equations and qq-characters*, *JHEP* **03** (2016) 181 [arXiv:1512.05388] [INSPIRE].

[53] N. Nekrasov, *BPS/CFT correspondence IV: sigma models and defects in gauge theory*, *Lett. Math. Phys.* **109** (2019) 579 [arXiv:1711.11011] [INSPIRE].

[54] K. Costello, *M-theory in the Omega-background and 5-dimensional non-commutative gauge theory*, arXiv:1610.04144 [INSPIRE].

[55] E. Witten, *Solutions of four-dimensional field theories via M-theory*, *Nucl. Phys. B* **500** (1997) 3 [hep-th/9703166] [INSPIRE].

[56] H. Kanno and Y. Tachikawa, *Instanton counting with a surface operator and the chain-saw quiver*, *JHEP* **06** (2011) 119 [arXiv:1105.0357] [INSPIRE].

[57] N.C. Leung and C. Vafa, *Branes and toric geometry*, *Adv. Theor. Math. Phys.* **2** (1998) 91 [hep-th/9711013] [INSPIRE].

[58] D. Gaiotto and E. Witten, *Supersymmetric Boundary Conditions in N=4 Super Yang-Mills Theory*, *J. Statist. Phys.* **135** (2009) 789 [arXiv:0804.2902] [INSPIRE].

[59] T. Procházka and M. Rapčák, *W-algebra modules, free fields, and Gukov-Witten defects*, *JHEP* **05** (2019) 159 [arXiv:1808.08837] [INSPIRE].

[60] D. Gaiotto and M. Rapčák, *Vertex Algebras at the Corner*, *JHEP* **01** (2019) 160 [arXiv:1703.00982] [INSPIRE].

[61] D. Gaiotto, S. Gukov and N. Seiberg, *Surface Defects and Resolvents*, *JHEP* **09** (2013) 070 [arXiv:1307.2578] [INSPIRE].

[62] K. Hori and M. Romo, *Exact Results In Two-Dimensional (2,2) Supersymmetric Gauge Theories With Boundary*, arXiv:1308.2438 [INSPIRE].

[63] D. Gaiotto, G.W. Moore and A. Neitzke, *Wall-Crossing in Coupled 2d-4d Systems*, *JHEP* **12** (2012) 082 [arXiv:1103.2598] [INSPIRE].

[64] D. Gaiotto, *Surface Operators in N=2 4d Gauge Theories*, *JHEP* **11** (2012) 090 [arXiv:0911.1316] [INSPIRE].

[65] A. Hanany and D. Tong, *Vortex strings and four-dimensional gauge dynamics*, *JHEP* **04** (2004) 066 [hep-th/0403158] [INSPIRE].

[66] N. Dorey, S. Lee and T.J. Hollowood, *Quantization of Integrable Systems and a 2d/4d Duality*, *JHEP* **10** (2011) 077 [arXiv:1103.5726] [INSPIRE].

[67] S. Gukov and E. Witten, *Rigid Surface Operators*, *Adv. Theor. Math. Phys.* **14** (2010) 87 [arXiv:0804.1561] [INSPIRE].

[68] N. Nekrasov and A. Tsymbaliuk, *Surface defects in gauge theory and KZ equation*, *Lett. Math. Phys.* **112** (2022) 28 [arXiv:2103.12611] [INSPIRE].

[69] B. Feigin, E. Frenkel and N. Reshetikhin, *Gaudin model, Bethe ansatz and correlation functions at the critical level*, *Commun. Math. Phys.* **166** (1994) 27 [hep-th/9402022] [INSPIRE].

[70] P. Etingof, E. Frenkel and D. Kazhdan, *Analytic Langlands correspondence for PGL_2 on $\mathbb{P}^1$ with parabolic structures over local fields*, *Geom. Funct. Anal.* **32** (2022) 725 [arXiv:2106.05243] [INSPIRE].

[71] P. Etingof, E. Frenkel and D. Kazhdan, *Hecke operators and analytic Langlands correspondence for curves over local fields*, *Duke Math. J.* **172** (2023) 2015 [arXiv:2103.01509] [INSPIRE].

[72] P. Etingof, E. Frenkel and D. Kazhdan, *A general framework and examples of the analytic Langlands correspondence*, *Pure Appl. Math. Quart.* **20** (2024) 307 [arXiv:2311.03743] [INSPIRE].

[73] E.E. Mukhin, V.O. Tarasov and A.N. Varchenko, *Bispectral and* $(\mathfrak{gl}_N, \mathfrak{gl}_M)$ *dualities*, *Functional Analysis and Other Mathematics* **1** (2007) 47.

[74] A. Mironov, A. Morozov, Y. Zenkevich and A. Zotov, *Spectral Duality in Integrable Systems from AGT Conjecture*, *JETP Lett.* **97** (2013) 45 [arXiv:1204.0913] [INSPIRE].

[75] K. Bulycheva, H.-Y. Chen, A. Gorsky and P. Koroteev, *BPS States in Omega Background and Integrability*, *JHEP* **10** (2012) 116 [arXiv:1207.0460] [INSPIRE].

[76] D. Gaiotto and P. Koroteev, *On Three Dimensional Quiver Gauge Theories and Integrability*, *JHEP* **05** (2013) 126 [arXiv:1304.0779] [INSPIRE].

[77] S. Jeong and X. Zhang, *BPZ equations for higher degenerate fields and nonperturbative Dyson-Schwinger equations*, *Phys. Rev. D* **109** (2024) 125006 [arXiv:1710.06970] [INSPIRE].

[78] N. Lee and N. Nekrasov, *Quantum spin systems and supersymmetric gauge theories. Part I*, *JHEP* **03** (2021) 093 [arXiv:2009.11199] [INSPIRE].

[79] S.E. Derkachov, G.P. Korchemsky and A.N. Manashov, *Separation of variables for the quantum SL(2,R) spin chain*, *JHEP* **07** (2003) 047 [hep-th/0210216] [INSPIRE].

[80] E.K. Sklyanin, *Separation of variables - new trends*, *Prog. Theor. Phys. Suppl.* **118** (1995) 35 [solv-int/9504001] [INSPIRE].

[81] S.E. Derkachov and P.A. Valinevich, *Separation of variables for the quantum* $SL(3,\mathbb{C})$ *spin magnet: eigenfunctions of Sklyanin B-operator*, *Zap. Nauchn. Semin.* **473** (2018) 110 [arXiv:1807.00302] [INSPIRE].

[82] P. Ryan and D. Volin, *Separation of Variables for Rational* $\mathfrak{gl}(\mathsf{n})$ *Spin Chains in Any Compact Representation, via Fusion, Embedding Morphism and Bäcklund Flow*, *Commun. Math. Phys.* **383** (2021) 311 [arXiv:2002.12341] [INSPIRE].

[83] D. Gaiotto and J. Teschner, *Quantum Analytic Langlands Correspondence*, arXiv:2402.00494 [INSPIRE].

[84] P. Etingof, E. Frenkel and D. Kazhdan, *An analytic version of the Langlands correspondence for complex curves*, arXiv:1908.09677 [INSPIRE].

[85] D. Gaiotto and E. Witten, *Gauge Theory and the Analytic Form of the Geometric Langlands Program*, *Annales Henri Poincare* **25** (2024) 557 [arXiv:2107.01732] [INSPIRE].

[86] N. Nekrasov, *BPS/CFT Correspondence III: Gauge Origami partition function and qq-characters*, *Commun. Math. Phys.* **358** (2018) 863 [arXiv:1701.00189] [INSPIRE].