# Loop Equations Characterize Random Matrix Statistics

Paul Bourgade
*Courant Institute, NYU*
*bourgade@cims.nyu.edu*

Jiaoyang Huang
*University of Pennsylvania*
*huangjy@wharton.upenn.edu*

**Abstract**

We prove that the universal local point processes of random matrix theory are characterized by their loop equation hierarchies. More precisely, for every rational $\beta > 0$, the $\mathrm{Sine}_\beta$ point process is the unique solution of the bulk loop equation hierarchy, and the $\mathrm{Airy}_\beta$ point process is the unique solution of the edge loop equation hierarchy.

These uniqueness results provide a direct route to universality: it suffices to verify the corresponding approximate loop equations for the ensemble. In many models, these equations follow from local laws and integration by parts.

# Contents

# 1 Introduction

A central theme in random matrix theory is *universality*: after a microscopic scaling, local eigenvalue statistics in the bulk and at regular edges depend only on the symmetry class of the model, or more generally on the parameter $\beta$, and not on the microscopic details of the underlying distribution or potential.

This prediction, originally formulated by Wigner, was extended to non-spectral correlated systems, as demonstrated by several paradigmatic integrable models. Examples include the Tracy–Widom fluctuations of the longest increasing subsequence of a uniform random permutation [5] and integrable models of directed last-passage percolation and random growth [58], which belong to the broader one-dimensional KPZ universality class [25, 76, 77].

Random-matrix predictions also play a central role in two major conjectural settings. The high-lying zeros of the Riemann zeta function are expected to exhibit GUE local statistics, beginning with Montgomery's pair-correlation conjecture and supported by Odlyzko's numerical computations [70, 73]. In quantum chaos, the Bohigas–Giannoni–Schmit conjecture predicts that quantum systems whose classical dynamics are sufficiently chaotic have local spectral statistics described by the appropriate Wigner–Dyson ensemble, determined by the symmetries of the system [12]. In contrast with the preceding random-matrix and integrable-probability examples, these conjectural directions remain open.

We recall two major approaches to proving the universality of random matrix statistics. The first is based on exact algebraic formulas. For unitary invariant ensembles, the eigenvalues form determinantal point processes, and the asymptotic analysis of the associated Christoffel–Darboux kernels leads to the sine-kernel process in the bulk and the Airy-kernel process at a regular soft edge. For orthogonal and symplectic ensembles, analogous Pfaffian, or quaternion determinantal, formulas play the same role. These exact structures, together with orthogonal polynomial and Riemann–Hilbert steepest-descent methods, have provided one of the most powerful routes to local random matrix statistics; see, for example, [2, 26, 27, 68].

A second major route is comparison with integrable ensembles and their perturbations, such as Gaussian and Gaussian-divisible ensembles. In the setting of Wigner matrices, this approach led to the dynamical three-step scheme developed by Erdős, Schlein, Yau, Yin, and collaborators [37, 38, 40]:

(i) establish a local semicircle law, together with eigenvalue rigidity and eigenvector delocalization estimates [36, 42];

(ii) prove universality after a short Gaussian evolution, equivalently for Gaussian divisible ensembles. In the complex Hermitian case, this can be done using exact determinantal formulas for GUE-divisible ensembles [35,59]. For general universality classes, it follows from the short-time relaxation of the Dyson Brownian motion [37, 38, 63], identifying local statistics by comparison with the integrable ensembles;

(iii) remove the Gaussian component by a perturbative, density argument. For sufficiently regular entry distributions this can be done by the reverse heat flow argument [35]. More robust comparison methods include the four moment theorem of Tao and Vu [88, 89], the Green function comparison theorem [41] and the continuity along matrix dynamics [21].

This comparison approach through dynamics applies beyond Wigner matrices, to prove random matrix universality for random graphs, band matrices, $\beta$-ensembles, non-centered models with variance profiles, convolution models appearing in free probability among others, see e.g. [91]. A different comparison mechanism, based on approximate transport maps, provides alternative proofs of universality for $\beta$-ensembles [11,86] and also applies to perturbative multi-matrix models [44].

The purpose of the present paper is to develop a different point of view. Rather than proving universality by exact algebraic formulas for correlation functions, or by reducing a model to an exactly solvable reference ensemble, we seek an intrinsic characterization of the limiting point processes themselves, through a BBGKY-type hierarchy. Our characterization is expressed through *loop equations*. Loop equations, also known as Dyson–Schwinger equations, master equations, or Ward identities, originate from linear responses, invariance principles or integration by parts identities. They appeared in the theoretical physics literature in the context of gauge theory and matrix models, for instance in the work of 't Hooft, Makeenko–Migdal and Migdal [66, 69, 87]. They were later introduced into the mathematical analysis of random matrices by Johansson [57], who used them to derive macroscopic central limit theorems for $\beta$-ensembles. Loop-equation methods have since become a standard tool in the study of global fluctuations and large-$N$ expansions of log-gases and matrix models; see, for example, [14, 15, 43, 85] and the monograph [48].

For a $\beta$-ensemble with $N$ particles and potential $V$, the loop equation hierarchy is a system of identities for resolvent observables, or equivalently for cumulants of Stieltjes transforms of the eigenvalue process. This hierarchy encodes the interplay between the logarithmic repulsion, the confining potential $V$, and the finite-size correction terms.

Although the finite-$N$ loop equations are model-dependent, their microscopic limits are universal. When one zooms into the bulk at the scale of the mean eigenvalue spacing, or into a regular spectral edge at the $N^{-2/3}$ scale, the rescaled loop equations converge to universal limiting hierarchies. We call these limiting identities the *bulk loop equations* (1.2) and the *edge loop equations* (1.5). Unlike the finite-$N$ loop equations, these microscopic equations no longer contain the original potential. They are universal equations for the limiting local point processes.

Our main results show that the universal microscopic loop equations do not merely hold for the $\beta$-random matrix limits, but characterize them. More precisely, for every rational $\beta > 0$, we prove that the $\mathrm{Sine}_\beta$ point process is the unique solution of the bulk loop equation hierarchy, and that the $\mathrm{Airy}_\beta$ point process is the unique solution of the edge loop equation hierarchy. These processes arise, respectively, as the universal bulk and edge limits of Gaussian $\beta$-ensembles [78, 90].

This result is analogous in spirit to Stein's method. The standard Gaussian distribution is characterized by the integration by parts identity

$$\mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)]$$

for a suitable class of test functions $f$. In our setting, the role of Stein's identity is played by the loop equation hierarchy. These identities are again derived by integration by parts, or equivalently by infinitesimal changes of variables, but now in the setting of the one-dimensional logarithmic gas and for a specific family of test functions, namely moments of Stieltjes transforms.

The key point is the passage from identities to characterization, which turns the approximate loop equations into a convergence criterion for universal local statistics of random matrices and related models. In the bulk, verifying the approximate loop equations yields convergence to the $\mathrm{Sine}_\beta$ point process. At the edge, the analogous criterion yields convergence to the $\mathrm{Airy}_\beta$ point process, and hence to the Tracy–$\mathrm{Widom}_\beta$ law for the top eigenvalue. These equations can often be verified by integration by parts, or by a discrete analogue thereof, with the local law as the main input.

To demonstrate the scope of this approach beyond invariant ensembles, we apply our characterization to random $d$-regular graphs. We prove bulk universality for random $d$-regular graphs on $N$ vertices with degree $d$ growing polylogarithmically in $N$, and we show that the same characterization principle significantly simplifies the proof of edge universality for fixed-degree random $d$-regular graphs in [52].

## 1.1 Main results

Throughout the paper, we fix $\beta > 0$. The $\mathrm{Sine}_\beta$ point process is the universal bulk scaling limit of the Gaussian $\beta$-ensemble, characterized by Valkó and Virág through the Brownian carousel construction [90]. We use the normalization in which its intensity is $1/\pi$, or equivalently, the local mean spacing is $\pi$.

The $\mathrm{Airy}_\beta$ point process is the universal soft-edge scaling limit of the Gaussian $\beta$-ensemble, characterized by Ramírez, Rider, and Virág through the spectrum of the stochastic Airy operator [78]. We use the reflected convention in which the Airy points are ordered decreasingly and extend to $-\infty$, matching the usual ordering of the negative zeros of the Airy function. In this normalization, the one-point density has the standard square-root left-tail asymptotics with leading constant $1/\pi$.

We regard both $\mathrm{Sine}_\beta$ and $\mathrm{Airy}_\beta$ as locally finite random point configurations on $\mathbb{R}$. We write a configuration as $\boldsymbol{x} = \{x_i\}_{i\in\mathbb{I}}$, where $\mathbb{I} = \mathbb{Z}$ in the bulk case and $\mathbb{I} = \mathbb{N}$ in the edge case. In the bulk case the particles are indexed in increasing order, while in the edge case they are indexed in decreasing order, so that $x_1 \geqslant x_2 \geqslant \cdots$ and $x_i \to -\infty$. Finally, we denote by $\mathbb{C}_+$ and $\mathbb{C}_-$ the open upper and lower half-planes, respectively.

**Definition 1.1.** *A holomorphic function $s : \mathbb{C}_+ \to \mathbb{C}_+ \cup \mathbb{R}$ is called a* Nevanlinna function*. Such functions are also called* Herglotz *or* Pick *functions in the literature. Any Nevanlinna function $s$ admits the integral representation*

$$s(w) = b + cw + \int_{\mathbb{R}} \left( \frac{1}{x-w} - \frac{x}{1+x^2} \right) \mathrm{d}\mu(x), \tag{1.1}$$

*where $b, c \in \mathbb{R}$, $c \geqslant 0$, and $\mu$ is a Borel measure satisfying*

$$\int_{\mathbb{R}} \frac{\mathrm{d}\mu(x)}{1+x^2} < \infty.$$

Nevanlinna functions form a subclass of $\mathcal{O}(\mathbb{C}_+) := \{f : \mathbb{C}_+ \to \mathbb{C} \text{ holomorphic}\}$. Whenever appropriate, we extend $f \in \mathcal{O}(\mathbb{C}_+)$ to the lower half-plane by Schwarz reflection, $f(\overline{w}) = \overline{f(w)}$, and interpret boundary values on $\mathbb{R}$ as non-tangential limits.

**Definition 1.2** (Configuration)**.** *A* configuration *on $\mathbb{R}$ is a Radon measure of the form*

$$\mu = \sum_{x\in P} \delta_x,$$

*where $P$ is a finite or countable multiset in $\mathbb{R}$, such that*

$$\mu(K) < \infty \qquad \text{for every compact set } K \subset \mathbb{R}.$$

*Equivalently, $\mu$ is a purely atomic Radon measure whose atoms have positive integer masses and such that every bounded set contains only finitely many atoms.*

**Definition 1.3.** *We say that a Nevanlinna function $s$ is* particle-generated *if the measure $\mu$ in the representation (1.1) is a configuration. In this case, $s$ extends to a meromorphic function on $\mathbb{C}$, satisfies $s(\overline{w}) = \overline{s(w)}$, and its poles are precisely the atoms of $\mu$, each with positive integer residue.*[1]

In particular, when $P$ is finite, the Stieltjes transform

$$\sum_{x\in P} \frac{1}{x-w}$$

is a particle-generated Nevanlinna function, corresponding to the choice $c = 0$ and $b = \int_{\mathbb{R}} x/(1+x^2)\,\mathrm{d}\mu(x)$.

When $P$ is infinite, well-balanced, in the bulk case one could define the Stieltjes transform as a principal value. We work with the broader class of particle-generated Nevanlinna functions instead, for three reasons. First, it allows a more unified treatment of loop equations in the bulk and at the edge. Second, Nevanlinna functions are stable under locally uniform limits. Third, the Stieltjes transform is a basic observable in the random matrix setting which may not be natural in other contexts. Our characterization criterion applies to general holomorphic observables $s : \mathbb{C}_+ \to \mathbb{C}_+ \cup \mathbb{R}$.

**Assumption 1.4** (Bulk loop equations)**.** *We assume that $s(w)$ is a particle-generated Nevanlinna function, in $\mathrm{L}^q$ for any $q > 0$. For every integer $p \geqslant 1$ and every $w_1, \ldots, w_p \in \mathbb{C}_+ \cup \mathbb{C}_-$, $s$ satisfies the loop equation*

$$\mathbb{E}\left[ \left( \frac{2-\beta}{\beta}\, \partial_{w_1} s(w_1) + s(w_1)^2 + 1 \right) \prod_{j=2}^{p} s(w_j) \right] + \frac{2}{\beta}\, \mathbb{E}\left[ \sum_{j=2}^{p} \partial_{w_j} \frac{s(w_1) - s(w_j)}{w_1 - w_j} \prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s(w_\ell) \right] = 0. \tag{1.2}$$

[1]More precisely, if $x \in \mathbb{R}$ appears in $P$ with multiplicity $m$, then the residue of $s$ at $x$ is $m$.

*When $p=1$, the empty product is interpreted as 1, and the second term is absent. When $w_j = w_1$, the corresponding difference quotient is interpreted by its holomorphic extension to the diagonal; in particular,*

$$\partial_{w_j} \frac{s(w_1)-s(w_j)}{w_1-w_j}\bigg|_{w_j=w_1} = \frac{1}{2}\,\partial_{w_1}^2 s(w_1).$$

**Theorem 1.5.** *Assume that $\beta>0$ is rational. Under Assumption 1.4, the poles of $s$ define a point process on $\mathbb{R}$, and this point process has the same law as the $\mathrm{Sine}_\beta$ point process.*

**Remark 1.6** (Non-uniqueness for $\beta=0$)**.** *In the degenerate limit $\beta\to 0$, the bulk loop-equation hierarchy (1.2) loses the interaction term and reduces to*

$$\mathbb{E}\left[\partial_{w_1} s(w_1)\prod_{j=2}^{p} s(w_j)\right] + \mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s(w_1)-s(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s(w_\ell)\right] = 0. \tag{1.3}$$

*The resulting hierarchy is highly non-unique. In fact, it is satisfied by any homogeneous Poisson point process of intensity $\lambda>0$. Moreover, since the identities are linear in the law of the point process, arbitrary convex mixtures of such solutions are again solutions.*

**Remark 1.7** (Uniqueness for general hierarchies)**.** *Constraints on the correlations of particle systems at thermal equilibrium, often referred to as sum rules, are a basic tool for the study of charged fluids, see e.g. [67]. Various families of such constraints, or hierarchies thereof, such as Dobrushin-Lanford-Ruelle, Kirkwood–Salzburg, Kubo-Martin-Schwinger, and Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY), have been instrumental to describe particle systems on mesoscopic and macroscopic scales, often by truncations of the hierarchy.*

*How full hierarchies may characterize the microscopic statistics is less understood. In the case of regular, short-range interaction, existence and uniqueness of solutions holds at low density [46, 71], and for integrable interaction [47]. Appendix F proves that the bulk loop equations from Assumption 1.4 are equivalent to the BBGKY hierarchy at equilibrium, for the logarithmic interaction. Theorem 1.5 therefore provides one example of sum rules characterizing a Gibbs state in the case of singular, long-range interactions.*

**Remark 1.8** (The distribution of $s$ on the real line)**.** *Herglotz–Pick (Nevanlinna) functions also play an important role in [1], where it is proved that Stieltjes transforms of a wide class configurations, evaluated at random points on the real line, have a Cauchy distribution.*

The above theorem allows to prove the following convergence criterion in the bulk.

**Theorem 1.9.** *Assume that $\beta>0$ is rational. Let $s_N$ be a sequence of random particle-generated Nevanlinna functions, and let $\mu_N$ denote the associated configurations. Assume that there exists a sequence of events $\Omega_N$ such that*

$$\mathbb{P}(\Omega_N) = 1 - \mathrm{o}_N(1)$$

*and $s_N(w)\mathbf{1}(\Omega_N)$ is in $\mathrm{L}^q$ for any $q>0$. Assume moreover that for every $p\geqslant 1$ and every compact set $K\subset\mathbb{C}_+\cup\mathbb{C}_-$, the following holds uniformly for all $w_1,w_2,\dots,w_p\in K$:*

$$\begin{aligned}\Bigg|&\mathbb{E}\left[\mathbf{1}(\Omega_N)\left(\frac{2-\beta}{\beta}\partial_{w_1}s_N(w_1)+s_N(w_1)^2+1\right)\prod_{j=2}^{p}s_N(w_j)\right]\\ &+\frac{2}{\beta}\mathbb{E}\left[\mathbf{1}(\Omega_N)\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w_1)-s_N(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p}s_N(w_\ell)\right]\Bigg| = \mathrm{o}_N(1)\cdot\left(1+\mathbb{E}\left[\mathbf{1}(\Omega_N)\mathcal{R}_N^{p+1}\right]\right),\end{aligned} \tag{1.4}$$

*where $\mathcal{R}_N = \max\{|s_N(w_1)|,|s_N(w_2)|,\cdots,|s_N(w_p)|\}$. Then $\mu_N$ converges in distribution, with respect to the vague topology, to the $\mathrm{Sine}_\beta$ point process.*

We now turn to analogous statements at the edge. Note that the only difference in the loop equations below, when compared to Assumption 1.4, is the constant $+1$ transformed into $-w_1$, which is eventually responsible for breaking translation invariance.

**Assumption 1.10** (Edge loop equations). *We assume that $s(w)$ is a particle-generated Nevanlinna function, in $\mathrm{L}^q$ for any $q>0$. For every integer $p \geqslant 1$ and every $w_1, w_2, \ldots, w_p \in \mathbb{C}_+ \cup \mathbb{C}_-$, $s$ satisfies the loop equations*

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\,\partial_{w_1} s(w_1) + s(w_1)^2 - w_1\right)\prod_{j=2}^{p} s(w_j)\right] + \frac{2}{\beta}\,\mathbb{E}\left[\sum_{j=2}^{p} \partial_{w_j} \frac{s(w_1)-s(w_j)}{w_1-w_j} \prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s(w_\ell)\right] = 0. \tag{1.5}$$

*When $p=1$, the empty product is interpreted as 1, and the second term is absent.*

**Theorem 1.11.** *Assume that $\beta>0$ is rational. Under Assumption 1.10, the poles of $s$ define a point process on $\mathbb{R}$, and this point process has the same law as the* $\mathrm{Airy}_\beta$ *point process.*

**Theorem 1.12.** *Assume that $\beta>0$ is rational. Let $s_N$ be a sequence of random particle-generated Nevanlinna functions, and let $\mu_N$ denote the associated configurations. Assume that there exists a sequence of events $\Omega_N$ such that*

$$\mathbb{P}(\Omega_N) = 1 - \mathrm{o}_N(1)$$

*and $s_N(w)\mathbf{1}(\Omega_N)$ is in $\mathrm{L}^q$ for any $q>0$. Assume moreover that for every $p \geqslant 1$ and every compact set $K \subset \mathbb{C}_+ \cup \mathbb{C}_-$, the following holds uniformly for all $w_1, w_2, \ldots, w_p \in K$:*

$$\begin{aligned}\Bigg|\mathbb{E}&\left[\mathbf{1}(\Omega_N)\left(\frac{2-\beta}{\beta}\partial_{w_1} s_N(w_1) + s_N(w_1)^2 - w_1\right)\prod_{j=2}^{p} s_N(w_j)\right]\\ &+\frac{2}{\beta}\mathbb{E}\left[\mathbf{1}(\Omega_N)\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w_1)-s_N(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s_N(w_\ell)\right]\Bigg| = \mathrm{o}_N(1)\cdot\left(1+\mathbb{E}\left[\mathbf{1}(\Omega_N)\mathcal{R}_N^{p+1}\right]\right),\end{aligned} \tag{1.6}$$

*where $\mathcal{R}_N = \max\{|s_N(w_1)|, |s_N(w_2)|, \cdots, |s_N(w_p)|\}$. Then $\mu_N$ converges in distribution, with respect to the vague topology, to the* $\mathrm{Airy}_\beta$ *point process.*

In the preceding theorems, we assume that $\beta>0$ is rational. This assumption enters through the proof, rather than through the expected characterization itself. More precisely, our argument uses the loop equations to identify certain exponential observables, which in turn determine the correlation functions of the point process. At present, the construction of these exponential observables requires $\beta$ to be rational.

This rationality phenomenon is consistent with the integrable-structure literature surrounding $\beta$-ensembles. Exact formulas for rational $\beta$ have appeared in [49, 65, 74, 80]; these formulas are often naturally organized in terms of finitely many quasiparticle and quasihole variables. We believe, however, that the uniqueness statements themselves should remain valid for every $\beta>0$.

**Conjecture 1.13.** *For any $\beta>0$, not necessarily rational, the following statements hold. Under Assumption 1.4, the poles of $s$ define a point process on $\mathbb{R}$, and this point process has the same law as the* $\mathrm{Sine}_\beta$ *point process. Under Assumption 1.10, the poles of $s$ define a point process on $\mathbb{R}$, and this point process has the same law as the* $\mathrm{Airy}_\beta$ *point process.*

Finally, $\mathrm{Sine}_\beta$ has also been studied from the perspective of Gibbs measures: it was shown to satisfy the Dobrushin–Lanford–Ruelle equations [28]. Our approach instead starts from the BBGKY/loop equation hierarchy and proves that it uniquely determines the process.

Another approach to uniqueness of the $\mathrm{Sine}_\beta$ process is based on an infinite-volume variational principle. Leblé and Serfaty [64] established a large deviation principle for microscopic empirical fields of log gases, whose rate function is an infinite-volume free energy consisting of a specific relative entropy term and a renormalized logarithmic energy term. Erbar, Huesmann, and Leblé [33] subsequently proved that, in one dimension, this free energy has a unique minimizer, namely the $\mathrm{Sine}_\beta$ point process.

The relation between this variational result and the loop-equation approach is not direct. In particular, it is not known whether solving the full hierarchy of loop equations forces a point process to minimize the infinite-volume log-gas free energy. Moreover, the variational problem in [33] is formulated for stationary point processes, it is intrinsically a bulk theory after local averaging. It therefore does not cover non-stationary edge limits, such as the $\mathrm{Airy}_\beta$ process.

## 1.2 Proof ideas

We explain the proof in the bulk case; the proof of the edge theorem is parallel, with the bulk loop equations replaced by the edge loop equations and $\mathrm{Sine}_\beta$ replaced by $\mathrm{Airy}_\beta$. The argument has three main steps.

*Step 1: the loop equations imply concentration of the Stieltjes transform.* The first role of the loop-equation hierarchy is quantitative: it forces the microscopic Stieltjes transform to concentrate around its deterministic values, i in the upper half-plane and $-\mathrm{i}$ in the lower half-plane, with sub-Gaussian tails. This phenomenon is already known in the setting of general $\beta$-ensembles, where the loop equations have been shown to imply an optimal local law, together with rigidity and sharp fluctuation estimates [19]. Thus, in our argument, the local law is not an additional input but a consequence of the loop equations. We use this concentration estimate as the asymptotic input at infinity for the auxiliary observables introduced in the next step.

*Step 2: the nonlinear loop hierarchy can be linearized.* The central observation is that the loop equation hierarchy admits a linear reformulation after passing to the observables

$$F(t_1,\ldots,t_n;s_1,\ldots,s_m) := \mathbb{E}\left[\mathrm{P.V.}\prod_{j\in\mathbb{Z}}\frac{\prod_{r=1}^n(t_r-x_j)}{\prod_{a=1}^m(s_a-x_j)^{\beta/2}}\right], \tag{1.7}$$

where the expectation is taken with respect to the random point process and $t_1,\ldots,t_n,s_1,\ldots,s_m\in\mathbb{C}_+\cup\mathbb{C}_-$. These observables are exponential linear statistics of the Nevanlinna function.

Under the half-plane balance condition (5.2), which in particular forces $\beta>0$ to be rational, the function $F$ satisfies the following system of linear differential equations:

$$\left(\partial_{t_i}^2+1+\frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i}-\partial_{t_j}}{t_i-t_j}-\sum_{a=1}^m\frac{\partial_{t_i}+(2/\beta)\partial_{s_a}}{t_i-s_a}\right)F=0,\qquad 1\leqslant i\leqslant n, \tag{1.8}$$

$$\left(\partial_{s_a}^2+\frac{\beta^2}{4}+\frac{\beta}{2}\sum_{b\neq a}\frac{\partial_{s_a}-\partial_{s_b}}{s_a-s_b}-\sum_{i=1}^n\frac{\partial_{s_a}+(\beta/2)\partial_{t_i}}{s_a-t_i}\right)F=0,\qquad 1\leqslant a\leqslant m. \tag{1.9}$$

The differential system (1.8)–(1.9) is closely related to deformed Calogero–Moser–Sutherland (CMS) systems with two species of variables. These are integrable systems of particles with long-range pair interactions, where the interaction strengths depend on whether the two particles belong to the same species or to different species. Such systems arise naturally from the representation theory of Lie superalgebras, which are variants of Lie algebras with two types of generators. Deformed CMS systems were introduced and studied, in the case $m=1$, by Chalykh–Feigin–Veselov [22], and were later systematically developed by Sergeev and Veselov in the framework of generalized root systems and Lie superalgebras [81–84]. In the trigonometric case, the integrability of the model is encoded by a family of mutually commuting differential operators, often called commuting quantum integrals. Their joint eigenfunctions are the super-Jack polynomials, introduced in [60]. Recent work [4, 50] further develops the orthogonality theory of super-Jack polynomials and its relation to rational and rational-harmonic deformed CMS systems. In our setting, the same two-species structure appears in a non-symmetric holonomic form, (1.8)–(1.9). Rather than working only with the commuting Hamiltonians of the integrable system, we obtain one second-order differential equation for each external variable. This form is naturally adapted to the loop-equation hierarchy.

Closely related holonomic systems also appear in Selberg-type integral formulas for $\beta$-ensembles. In particular, Desrosiers and Liu [31] studied a Selberg-type integral with $n+m$ external variables and showed that it satisfies a holonomic system of $n+m$ non-symmetric linear PDEs related to deformed Calogero–Sutherland systems. They further identified a distinguished normalized solution in terms of a hypergeometric function built from super-Jack polynomials, and applied the result to ratios of characteristic polynomials in classical $\beta$-ensembles. Our differential system (1.8)–(1.9) can be viewed as the bulk scaling limit of this type of holonomic PDE system. In contrast, the limiting system by itself does not determine a unique local solution: as we show below, its local solution space has dimension $2^{n+m}$, corresponding to the $2^{n+m}$ possible asymptotic branches.

In the special case $\beta = 2$, the observable (1.7) becomes an average of ratios of characteristic polynomials. This connects the system to the classical theory of ratios of characteristic-polynomial averages in unitary ensembles, including the works [3, 6, 13, 24].

The coupled equations (1.8)–(1.9) have a simple duality. In the $\boldsymbol{t}$-equations, the coefficient of the interaction between two $t$-variables is $2/\beta$. In the $\boldsymbol{s}$-equations, the corresponding coefficient between two $s$-variables is $\beta/2$. These two coefficients are reciprocals. More precisely, if we exchange the two sets of variables and their sizes, $\boldsymbol{t} \leftrightarrow \boldsymbol{s}, n \leftrightarrow m$, replace

$$\beta \longmapsto \frac{4}{\beta},$$

and make the corresponding rescaling of the variables by $\beta/2$, then the two systems of differential equations are transformed into each other, up to an overall nonzero factor. Since the equations are set equal to zero, this factor does not change the solution set. In this sense, the $\boldsymbol{t}$-variables and the $\boldsymbol{s}$-variables play dual roles.

This is the same type of $\beta \leftrightarrow 4/\beta$ duality that appears in $\beta$-ensembles. A common example comes from averages of characteristic polynomials. In many such formulas, one can exchange the number of eigenvalues with the number of inserted characteristic polynomials, provided one also replaces $\beta$ by $4/\beta$. The equations above can be viewed as a differential-equation form of this same symmetry. We refer to [29, 45] for this duality and for a recent survey of related dualities in random matrix theory.

*Step 3: solve the deformed CMS system and identify the physical branch.* Once the observables have been shown to satisfy (1.8)–(1.9), the remaining task is to solve a linear holonomic system. Concretely, this system consists of $n+m$ second-order linear partial differential equations in $n+m$ variables. Using standard $D$-module arguments, we show that the corresponding differential operators form a Gröbner basis. We refer the reader to [79] for further background. Hence the quotient module admits a basis given by square-free derivatives, and the local solution space has dimension $2^{n+m}$. Equivalently, the system can be prolonged to a flat first-order system, whose solutions are determined by parallel transport.

The $2^{n+m}$ asymptotic branches can moreover be described explicitly. For each sign pattern

$$(\epsilon_1, \ldots, \epsilon_n, \eta_1, \ldots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m,$$

there is a solution with leading exponential behavior

$$\exp\left(\mathrm{i}\sum_{i=1}^{n} \epsilon_i t_i + \mathrm{i}\frac{\beta}{2}\sum_{a=1}^{m} \eta_a s_a\right). \tag{1.10}$$

Thus the deformed CMS system has exactly $2^{n+m}$ linearly independent formal branches at infinity.

Once $F$ has been uniquely identified, the correlation functions of the point process can be recovered from it; these correlation functions, in turn, identify the limiting process. In the bulk case, this limiting process is $\mathrm{Sine}_\beta$.

## Outline of the paper.

In Section 2, we develop the convergence framework used throughout the paper and prove Theorems 1.9 and 1.12, assuming Theorems 1.5 and 1.11. As an application, we apply our characterization results to Wigner matrices in Section 2.3 and to random $d$-regular graphs in Section 2.4.

In Section 3.1, we show that the $\text{Sine}_\beta$ point process satisfies the bulk loop equation, while the $\text{Airy}_\beta$ point process satisfies the edge loop equation. These results are obtained by taking the bulk and edge scaling limits of the corresponding loop equations for $\beta$-ensembles; the detailed proofs are deferred to Appendix B. In Section 3.2, we show that the bulk loop equation implies concentration of the Stieltjes transform. The analogous proof in the edge case is deferred to Appendix C.

In Section 4, we introduce the bulk and edge exponential observables. Using the concentration estimates as input, we derive their asymptotic behaviors as the spectral parameters tend to infinity. In Section 5, we prove that these observables satisfy systems of linear differential equations related to deformed CMS operators.

In Section 6, we prove our main characterization results, assuming the classification of solutions to the deformed CMS systems. This classification is carried out in Sections 7 and 8, where we solve the bulk and edge deformed CMS systems, respectively. We show that each system has exactly $2^{n+m}$ linearly independent solutions, corresponding to the $2^{n+m}$ possible leading exponential behaviors at infinity. For $\beta = 2$, these asymptotic solutions become exact; this is verified in Appendix D in the bulk and in Appendix E at the edge. Finally, we collect some basic estimates on Airy functions in Appendix A.

## Notation.

We write $\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ and $\mathbb{C}_- := \{z \in \mathbb{C} : \operatorname{Im} z < 0\}$ for the open upper and lower half-planes. We denote by $\mathbb{Q}_{>0}$ the set of positive rational numbers and by $\mathbb{R}_{<0}$ the set of negative real numbers. For a positive integer, we write $[\![N]\!] := \{1, 2, \ldots, N\}$. For two numbers $a, b$, we use

$$a \wedge b := \min\{a, b\}, \qquad a \vee b := \max\{a, b\}.$$

For a random variable $X$ and $\ell > 0$, we denote

$$\|X\|_{L^\ell} := \mathbb{E}[|X|^\ell]^{1/\ell}.$$

For a matrix $X$, we write $\|X\|$ for its spectral norm. For a vector $\boldsymbol{v}$, we write $\|\boldsymbol{v}\|_2$ for its Euclidean norm. We denote the identity matrix by $\mathbb{I}$, and write $\mathbb{I}_n$ when we wish to emphasize that it is the $n \times n$ identity matrix.

For two quantities $X_N$ and $Y_N$ depending on $N$, we write that $X_N = \mathrm{O}(Y_N)$ or $X_N \lesssim Y_N$ if there exists some universal constant such that $|X_N| \leqslant CY_N$ . We write $X_N = \mathrm{o}(Y_N)$, or $X_N \ll Y_N$ if the ratio $|X_N|/Y_N \to 0$ as $N$ goes to infinity. We write $X_N \asymp Y_N$ if there exists a universal constant $C > 0$ such that $Y_N/C \leqslant |X_N| \leqslant CY_N$.

In Definition 1.3, the Nevanlinna functions $s$ correspond to configuration with interparticle distance $\asymp 1$, in particular it differs from the usual Stieltjes transform of a configuration of $N$ particles $\lambda_i$'s on a compact set, denoted in this paper $m_N(z) = (1/N)\sum_{i=1}^N 1/(\lambda_i - z)$.

Throughout the paper, fractional powers such as $\sqrt{w}$, $w^{3/2}$, $w^{-1/4}$, and $w^{\beta/2}$ are taken with respect to the principal branch of the logarithm. That is, for $w \in \mathbb{C} \setminus \mathbb{R}_{<0}$,

$$w^a := |w|^a e^{\mathrm{i}a \arg w}, \qquad \arg w \in (-\pi, \pi). \tag{1.11}$$

## Acknowledgements.

The research of P.B. was supported by the NSF Grant DMS-2348202. The research of J.H. was supported in part by NSF Grant DMS-2246664, DMS-2331096 and a Sloan Research Fellowship. We thank Alice Guionnet, Michel Pain, Xin Sun, Shengjing Xu, Horng-Tzer Yau, Ofer Zeitouni and Lingfu Zhang for enlightening discussions on loop equations. P.B. is particularly thankful to Alice Guionnet for early discussions about the central question addressed in this article, namely, the characterization of local statistics as the unique solution of loop equations, which began at the IAS program *Non-equilibrium Dynamics and Random Matrices* in 2013-2014. We thank Amol Aggarwal and Alexei Borodin for useful references related to the algebraic approach developed in this paper.

# 2 Applications

In this section we prove Theorem 1.9 and Theorem 1.12. These results state that, for rational $\beta > 0$, any sequence of random point processes on $\mathbb{R}$ satisfying the approximate bulk or edge loop equations with vanishing error converges, respectively, to the $\mathrm{Sine}_\beta$ point process or to the $\mathrm{Airy}_\beta$ point process.

To illustrate the scope of our results for random matrix models, we apply our characterization to Wigner matrices and to adjacency matrices of random $d$-regular graphs. First, we give a short proof of bulk universality for Wigner matrices. We then prove bulk universality for random $d$-regular graphs on $N$ vertices in the growing-degree regime $d \gg (\log N)^{24}$. Finally, we explain how the proof of edge universality for fixed-degree random $d$-regular graphs in [52] can be significantly simplified using our characterization.

## 2.1 Properties of Nevanlinna functions

We start with some estimates on the particle-generated Nevanlinna function $s$. Recall the integral representation from (1.1). Writing

$$s(w) = b + cw + \sum_{x \in P} \left( \frac{1}{x-w} - \frac{x}{1+x^2} \right), \qquad c \geqslant 0,$$

we have, for $w = E + \mathrm{i}\eta$,

$$\operatorname{Im}[s(w)] = c\eta + \sum_{x \in P} \frac{\eta}{|x-w|^2}. \tag{2.1}$$

Hence

$$|s'(w)| \leqslant c + \sum_{x \in P} \frac{1}{|x-w|^2} \leqslant \frac{\operatorname{Im}[s(w)]}{\eta}. \tag{2.2}$$

Also, for $z, w \in \mathbb{C} \setminus \mathbb{R}$,

$$\frac{s(z)-s(w)}{z-w} = c + \sum_{x \in P} \frac{1}{(x-z)(x-w)}, \qquad \partial_w \left( \frac{s(z)-s(w)}{z-w} \right) = \sum_{x \in P} \frac{1}{(x-z)(x-w)^2}. \tag{2.3}$$

Finally,

$$\begin{aligned} \left| \frac{s(z)-s(w)}{z-w} \right| &\leqslant c + \frac{1}{2} \sum_{x \in P} \frac{1}{|z-x|^2} + \frac{1}{|w-x|^2} \leqslant \frac{\operatorname{Im}[s(z)]}{2\operatorname{Im}[z]} + \frac{\operatorname{Im}[s(w)]}{2\operatorname{Im}[w]}, \\ \left| \partial_w \left( \frac{s(z)-s(w)}{z-w} \right) \right| &\leqslant \sum_{x \in P} \frac{1}{\operatorname{Im}[z]} \frac{1}{|x-w|^2} \leqslant \frac{\operatorname{Im}[s(w)]}{\operatorname{Im}[z]\ \operatorname{Im}[w]}. \end{aligned} \tag{2.4}$$

We regard Nevanlinna functions as elements of

$$\mathcal{O}(\mathbb{C}_+) := \{ f : \mathbb{C}_+ \to \mathbb{C} \text{ holomorphic} \}.$$

We endow $\mathcal{O}(\mathbb{C}_+)$ with the topology of locally uniform convergence: a sequence $\{f_n\}_{n \geqslant 1} \subset \mathcal{O}(\mathbb{C}_+)$ converges to $f \in \mathcal{O}(\mathbb{C}_+)$ if and only if $f_n \to f$ uniformly on every compact subset of $\mathbb{C}_+$. With this topology, $\mathcal{O}(\mathbb{C}_+)$ is a Fréchet space. Throughout this section, convergence of Nevanlinna functions is understood in this sense.

The following lemma shows that locally uniform convergence of Nevanlinna functions implies vague convergence of the associated Borel measures. We postpone the proof of the lemma to the end of this section.

**Lemma 2.1.** *Let $s$ and $\{s_N\}_{N\geqslant 1}$ be Nevanlinna functions, and let $\mu$ and $\{\mu_N\}_{N\geqslant 1}$ be the corresponding Borel measures in the representation (1.1). If $s_N$ converges to $s$ uniformly on compact subsets of $\mathbb{C}_+$, then $\mu_N$ converges to $\mu$ in the vague topology.*

*Moreover, if each $\mu_N$ is a configuration, then $\mu$ is also a configuration. Equivalently, if each $s_N$ is particle-generated, then so is $s$.*

*Proof of Lemma 2.1.* From the representation (1.1) of Nevanlinna functions, we have

$$s(w) = b + cw + \int_{\mathbb{R}} \left(\frac{1}{\lambda - w} - \frac{\lambda}{1+\lambda^2}\right) \mathrm{d}\mu(\lambda), \quad s_N(w) = b_N + c_N w + \int_{\mathbb{R}} \left(\frac{1}{\lambda - w} - \frac{\lambda}{1+\lambda^2}\right) \mathrm{d}\mu_N(\lambda), \tag{2.5}$$

where $b, b_N, c, c_N \in \mathbb{R}$, $c, c_N \geqslant 0$. By taking imaginary parts, and using that $s_N(z)$ converges to $s(z)$, we have

$$\left(c_N + \int \frac{\mathrm{d}\mu_N(\lambda)}{(\lambda - x)^2 + y^2}\right) y = \mathrm{Im}[s_N(x+\mathrm{i}y)] \to \mathrm{Im}[s(x+\mathrm{i}y)] = \left(c + \int \frac{\mathrm{d}\mu(\lambda)}{(\lambda - x)^2 + y^2}\right) y. \tag{2.6}$$

By taking $x + \mathrm{i}y = \mathrm{i}$, we conclude that

$$\sup_{N\geqslant 1} \left(c_N + \int_{\mathbb{R}} \frac{\mathrm{d}\mu_N(\lambda)}{1+\lambda^2}\right) < \infty, \quad c + \int_{\mathbb{R}} \frac{\mathrm{d}\mu(\lambda)}{1+\lambda^2} < \infty. \tag{2.7}$$

We now prove vague convergence. Let $f \in C_c^\infty(\mathbb{R})$, and choose $M \geqslant 1$ such that $\operatorname{supp} f \subset [-M, M]$. Fix $y \in (0,1]$. We write

$$\int_{\mathbb{R}} f(x)\, \mathrm{Im}[s_N(x+\mathrm{i}y)]\, \mathrm{d}x - \int_{\mathbb{R}} f(x)\, \mathrm{Im}[s(x+\mathrm{i}y)]\, \mathrm{d}x = \int_{\mathbb{R}} f(x)\, \mathrm{Im}\big[s_N(x+\mathrm{i}y) - s(x+\mathrm{i}y)\big]\, \mathrm{d}x,$$

which tends to 0 as $N \to \infty$ by local uniform convergence. Thus it suffices to show that

$$\int_{\mathbb{R}} f(x)\, \mathrm{Im}[s_N(x+\mathrm{i}y)]\, \mathrm{d}x \longrightarrow \pi \int_{\mathbb{R}} f(\lambda)\, \mathrm{d}\mu_N(\lambda) \qquad \text{as } y \downarrow 0, \tag{2.8}$$

with an error bound uniform in $N$, and similarly for $s, \mu$. Once this is established, we obtain

$$\int_{\mathbb{R}} f\, \mathrm{d}\mu_N \longrightarrow \int_{\mathbb{R}} f\, \mathrm{d}\mu,$$

hence $\mu_N \to \mu$ vaguely.

The statement (2.8) follows from showing the following estimate and its analogue obtained by replacing $\mathrm{Im}[s_N(x+\mathrm{i}y)]$ and $\mu_N$ with $\mathrm{Im}[s(x+\mathrm{i}y)]$ and $\mu$:

$$\begin{aligned} &\left|\int_{\mathbb{R}} \mathrm{Im}[s_N(x+\mathrm{i}y)] f(x) \mathrm{d}x - \pi \int_{\mathbb{R}} f(x) \mathrm{d}\mu_N(x)\right| \\ &\leqslant C\big(My\|f\|_\infty + M^2 y \ln(M/y)\|f'\|_\infty\big) \int_{\mathbb{R}} \frac{\mathrm{d}\mu_N(\lambda)}{1+\lambda^2} + 2c_N M \|f\|_\infty y. \end{aligned} \tag{2.9}$$

In fact, by (2.7), the right-hand side of (2.9) tends to 0 as $y \to 0+$, uniformly in $N$.

Using (2.6), we can rewrite the left-hand side of (2.9) as

$$\int_{\mathbb{R}} \mathrm{Im}[s_N(x+\mathrm{i}y)] f(x) \mathrm{d}x - \pi \int_{\mathbb{R}} f(x) \mathrm{d}\mu_N(x) = \int \left(\int_{\mathbb{R}} \frac{y f(x) \mathrm{d}x}{(x-\lambda)^2 + y^2} - \pi f(\lambda)\right) \mathrm{d}\mu_N(\lambda) + c_N y \int_{\mathbb{R}} f(x) \mathrm{d}x, \tag{2.10}$$

where the last term is bounded by $2c_N M \|f\|_\infty y$.

To bound the first term on the right-hand side of (2.10), we consider two cases. For $|\lambda| \geqslant 2M$, we have

$$\int_{\mathbb{R}} \frac{yf(x)\mathrm{d}x}{(x-\lambda)^2+y^2} \leqslant \int_{-M}^{M} \frac{y\|f\|_\infty \mathrm{d}x}{(x-\lambda)^2+y^2} \leqslant \frac{10My\|f\|_\infty}{1+\lambda^2}. \tag{2.11}$$

For $|\lambda| \leqslant 2M$,

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{yf(x)\mathrm{d}x}{(x-\lambda)^2+y^2} - \pi f(\lambda) = \int_{-4M}^{4M} \frac{y\big(f(x)-f(\lambda)\big)\mathrm{d}x}{(x-\lambda)^2+y^2} - \int_{\mathbb{R}\setminus[-4M,4M]} \frac{yf(\lambda)\mathrm{d}x}{(x-\lambda)^2+y^2} \\
&\lesssim y\|f'\|_\infty \int_{-4M}^{4M} \frac{|x-\lambda|\mathrm{d}x}{(x-\lambda)^2+y^2} + \|f\|_\infty \int_{x\in\mathbb{R}\setminus[-4M,4M]} \frac{y}{(x-\lambda)^2+y^2}\mathrm{d}x \\
&\lesssim y\ln(M/y)\|f'\|_\infty + \frac{y\|f\|_\infty}{M} \leqslant \frac{8M^2}{1+\lambda^2}\left(y\ln(M/y)\|f'\|_\infty + \frac{y\|f\|_\infty}{M}\right).
\end{aligned} \tag{2.12}$$

By plugging (2.11) and (2.12) into (2.10), we get

$$\left|\int\left(\int_{\mathbb{R}} \frac{yf(x)\mathrm{d}x}{(x-\lambda)^2+y^2} - \pi f(\lambda)\right)\mathrm{d}\mu_N(\lambda)\right| \leqslant C\big(My\|f\|_\infty + M^2 y\ln(M/y)\|f'\|_\infty\big)\int_{\mathbb{R}} \frac{\mathrm{d}\mu_N(\lambda)}{1+\lambda^2}, \tag{2.13}$$

and the claim (2.9) follows.

It remains to show that $\mu$ is also a configuration. From (2.7) we know that $\mu$ is locally finite. Now let $I=(a,b)$ be a bounded interval such that $\mu(\{a\})=\mu(\{b\})=0$. By vague convergence,

$$\mu(I) = \lim_{N\to\infty} \mu_N(I).$$

Since each $\mu_N$ is a configuration, $\mu_N(I)\in\mathbb{Z}_{\geqslant 0}$. So its limit $\mu(I)$ must also belong to $\mathbb{Z}_{\geqslant 0}$.

Thus $\mu$ assigns a nonnegative integer mass to every bounded interval whose endpoints are not atoms of $\mu$. It follows that $\mu$ is purely atomic, and each atom has positive integer mass. Since $\mu$ is locally finite, it has only finitely many atoms in each bounded interval. Therefore $\mu$ is a configuration. □

## 2.2 Proof of Theorem 1.9 and Theorem 1.12

*Proof of Theorem 1.9.* The proof consists of four steps.
*Step 1. Pointwise moment bounds.* Let $K\subset\mathbb{C}_+$ be compact and set

$$\eta_K := \inf_{z\in K} \operatorname{Im} z > 0. \tag{2.14}$$

Fix an integer $q\geqslant 2$. Taking $p=2q-1$, and choosing $w_1=w$, and $w_2,w_3,\dots,w_p$ so that $q-2$ of them are equal to $w$ and the remaining $q$ are equal to $\bar{w}$, in the bulk loop equation (1.4), we get, since $s_N(\bar{w})=\overline{s_N(w)}$,

$$s_N(w)^2\prod_{j=2}^{2q-1} s_N(w_j) = s_N(w)^2 s_N(w)^{q-2} s_N(\bar{w})^q = |s_N(w)|^{2q}.$$

Moreover, $\mathcal{R}_N = \max\{|s_N(w_1)|,\dots,|s_N(w_p)|\} = |s_N(w)|$. Thus, after isolating the leading quadratic term and taking absolute values, (1.4) gives

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}(\Omega_N)|s_N(w)|^{2q}\right] \leqslant C\mathbb{E}\left[\mathbf{1}(\Omega_N)\big(|\partial_w s_N(w)|+1\big)|s_N(w)|^{2q-2}\right] \\
&\quad + C\mathbb{E}\left[\mathbf{1}(\Omega_N)\sum_{j=1}^{2q-2}\left|\partial_{w_j}\frac{s_N(w)-s_N(w_j)}{w-w_j}\right||s_N(w)|^{2q-3}\right] + \mathrm{o}_N(1)\left(1+\mathbb{E}\left[\mathbf{1}(\Omega_N)|s_N(w)|^{2q}\right]\right),
\end{aligned} \tag{2.15}$$

where $C$ depends only on $\beta$ and the $\mathrm{o}_N(1)$ error is uniform for $w\in K$.

For a particle-generated Nevanlinna function $s_N$, the estimates (2.2) and (2.4) imply that, uniformly for $w \in K$ and $z \in \{w, \bar{w}\}$,

$$|\partial_w s_N(w)| \leqslant \frac{|s_N(w)|}{\operatorname{Im} w} \leqslant C_K |s_N(w)|, \qquad \left| \partial_z \frac{s_N(w) - s_N(z)}{w - z} \right| \leqslant C_K |s_N(w)|. \tag{2.16}$$

Plugging (2.16) into (2.15), we obtain

$$\mathbb{E}\left[\mathbf{1}(\Omega_N)|s_N(w)|^{2q}\right] \leqslant C_K \mathbb{E}\left[\mathbf{1}(\Omega_N)\big(|s_N(w)|^{2q-1} + q|s_N(w)|^{2q-2}\big)\right] + \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[\mathbf{1}(\Omega_N)|s_N(w)|^{2q}\right]\right). \tag{2.17}$$

We now apply Young's inequality in the forms

$$C_K b^{2q-1} \leqslant \frac{1}{4} b^{2q} + C_{q,K}, \qquad C_K q b^{2q-2} \leqslant \frac{1}{4} b^{2q} + C_{q,K}.$$

Using these inequalities in (2.17) and absorbing the $\mathrm{o}_N(1)$-term for all sufficiently large $N$, we conclude that there exists a constant $C_{q,K} > 0$ such that

$$\sup_{z \in K} \mathbb{E}\big[\mathbf{1}(\Omega_N)|s_N(z)|^{2q}\big] \leqslant C_{q,K} \tag{2.18}$$

for all sufficiently large $N$.
*Step 2. Compact-uniform moment bounds.* We next upgrade the pointwise estimate (2.18) to a compact-uniform estimate. Let $\delta > 0$, and choose a finite $\delta$-net $\mathcal{N}_\delta \subset K$. Thus, for every $z \in K$, there exists $w \in \mathcal{N}_\delta$ with $|z - w| \leqslant \delta$. For such $z$ and $w$, define the line segment

$$\gamma(t) := w + t(z - w), \qquad 0 \leqslant t \leqslant 1.$$

Then $\operatorname{Im} \gamma(t) \geqslant \eta_K$ (recall from (2.14)). Using the derivative estimate (2.16)

$$|s_N'(\gamma(t))| \leqslant \frac{|s_N(\gamma(t))|}{\operatorname{Im} \gamma(t)} \leqslant \frac{|s_N(\gamma(t))|}{\eta_K},$$

we get, for $0 \leqslant t \leqslant 1$,

$$|s_N(\gamma(t))| \leqslant |s_N(w)| + |z - w| \int_0^t |s_N'(\gamma(r))| \,\mathrm{d}r \leqslant |s_N(w)| + \frac{\delta}{\eta_K} \int_0^t |s_N(\gamma(r))| \,\mathrm{d}r.$$

By Gronwall's inequality,

$$|s_N(z)| \leqslant \exp\left(\frac{\delta}{\eta_K}\right) |s_N(w)|.$$

Consequently,

$$\sup_{z \in K} |s_N(z)|^{2q} \leqslant \exp\left(\frac{2q\delta}{\eta_K}\right) \sum_{w \in \mathcal{N}_\delta} |s_N(w)|^{2q}. \tag{2.19}$$

Multiplying by $\mathbf{1}(\Omega_N)$, taking expectations, and using (2.18), we obtain

$$\mathbb{E}\left[\mathbf{1}(\Omega_N) \sup_{z \in K} |s_N(z)|^{2q}\right] \leqslant C_{q,K}. \tag{2.20}$$

By the symmetry $s_N(\bar{z}) = \overline{s_N(z)}$, the same estimate also holds for compact sets $K \subset \mathbb{C} \setminus \mathbb{R}$.
*Step 3. Tightness and subsequential limits.* By Markov's inequality, (2.20) implies

$$\mathbb{P}\left(\Omega_N \cap \left\{\sup_{z \in K} |s_N(z)| > M\right\}\right) \leqslant \frac{C_{q,K}}{M^{2q}}.$$

Given $\varepsilon > 0$, choosing $M$ sufficiently large gives

$$\limsup_{N\to\infty} \mathbb{P}\left(\Omega_N \cap \left\{\sup_{z\in K} |s_N(z)| > M\right\}\right) \leqslant \varepsilon. \tag{2.21}$$

Since $\mathbb{P}(\Omega_N^{\complement}) = \mathrm{o}_N(1)$, it follows that

$$\limsup_{N\to\infty} \mathbb{P}\left(\sup_{z\in K} |s_N(z)| > M\right) \leqslant \varepsilon. \tag{2.22}$$

Thus the sequence $\{s_N\}$ is tight in the topology of locally uniform convergence on compact subsets of $\mathbb{C}\setminus\mathbb{R}$.

Now take an arbitrary subsequence $N_1 < N_2 < N_3 < \cdots$. By tightness, after passing to a further subsequence, still denoted by $N_j$, we may assume that $s_{N_j} \Longrightarrow s$ in the topology of locally uniform convergence on compact subsets of $\mathbb{C}_+$. We also keep track of the indicators $\mathbf{1}(\Omega_{N_j})$. Since $\mathbb{P}(\Omega_{N_j}) = 1 - \mathrm{o}_{N_j}(1)$, the pair $(s_{N_j}, \mathbf{1}(\Omega_{N_j}))$ converges subsequentially in distribution to $(s, 1)$.

By the Skorokhod representation theorem, we may realize $s_{N_j}$, $s$, and $\mathbf{1}(\Omega_{N_j})$ on the same probability space so that $s_{N_j} \longrightarrow s$ almost surely uniformly on compact subsets of $\mathbb{C}_+$, and $\mathbf{1}(\Omega_{N_j}) \longrightarrow 1$ almost surely.

Since each $s_{N_j}$ is a Nevanlinna function, the locally uniform limit $s$ is again a Nevanlinna function. Since each $s_{N_j}$ is particle-generated, Lemma 2.1 implies that $s$ is also particle-generated. If $\mu$ denotes the Borel measure corresponding to $s$ in the representation (1.1), then, along this subsequence, $\mu_{N_j} \longrightarrow \mu$ almost surely in the vague topology.

*Step 4. Passing to the exact loop equations.* It remains to pass to the limit in the approximate loop equations. Fix $p \geqslant 1$ and $w_1, \ldots, w_p \in \mathbb{C}\setminus\mathbb{R}$. We first record the needed uniform integrability. Combining (2.16), (2.18), and Hölder's inequality, we obtain that, for every fixed $p$ and every fixed $w_1, \ldots, w_p \in \mathbb{C}\setminus\mathbb{R}$, the random variables

$$\mathbf{1}(\Omega_N)\left(\frac{2-\beta}{\beta}\partial_{w_1} s_N(w_1) + s_N(w_1)^2 + 1\right)\prod_{j=2}^{p} s_N(w_j), \quad \mathbf{1}(\Omega_N)\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w_1) - s_N(w_j)}{w_1 - w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s_N(w_\ell) \tag{2.23}$$

are uniformly integrable.

On the Skorokhod probability space, the almost sure locally uniform convergence $s_{N_j} \to s$ implies, by Cauchy's integral formula, that the derivatives also converge locally uniformly. Therefore the integrands in (1.4), multiplied by $\mathbf{1}(\Omega_{N_j})$, converge almost surely to the corresponding expressions with $s$ in place of $s_N$, since $\mathbf{1}(\Omega_{N_j}) \to 1$ almost surely. By uniform integrability (2.23), we may pass to the limit in expectation. Thus, using the approximate loop equation (1.4), this gives

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\partial_{w_1} s(w_1) + s(w_1)^2 + 1\right)\prod_{j=2}^{p} s(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s(w_1) - s(w_j)}{w_1 - w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s(w_\ell)\right] = 0. \tag{2.24}$$

Thus $s$ satisfies the exact bulk loop equations.

By the characterization Theorem 1.5 for the exact bulk loop equations, the configuration $\mu$ has the same law as the $\mathrm{Sine}_\beta$ point process. We have shown that every subsequence of $\{\mu_N\}_{N\geqslant 1}$ has a further subsequence converging in distribution to the $\mathrm{Sine}_\beta$ point process. Therefore the full sequence $\mu_N$ converges in distribution to the $\mathrm{Sine}_\beta$ point process. □

*Proof of Theorem 1.12.* Using Theorem 1.11 as input, the proof of Theorem 1.12 is identical to the proof of Theorem 1.9. We therefore omit the details. □

## 2.3 Wigner matrices

In this section, we illustrate how our characterization can be applied to Wigner matrices. Bulk universality for Wigner matrices was previously established in several forms: averaged-energy universality and gap universality were proved in [35,37–39,41,42,89], while fixed-energy universality was proved later in [18,63]. We show that the local semicircle law, together with a cumulant expansion, implies the microscopic approximate loop equation. Consequently, by Theorem 1.9, this yields a short proof of bulk universality at fixed energy.

**Definition 2.2** (Real Wigner matrix). *Let $H = (h_{ij})_{i,j=1}^N$ be an $N \times N$ real symmetric random matrix. We assume that the upper-triangular entries $\{h_{ij} : 1 \leqslant i \leqslant j \leqslant N\}$ are independent, and that $h_{ji} = h_{ij}$ for $i \leqslant j$. We assume*

$$\mathbb{E}[h_{ij}] = 0, \qquad \mathbb{E}[(\sqrt{N} h_{ij})^2] = 1, \qquad 1 \leqslant i, j \leqslant N.$$

*Moreover, for every $q \in \mathbb{N}$, there exists a constant $C_q > 0$, independent of $N, i, j$, such that $\mathbb{E}[|\sqrt{N} h_{ij}|^q] \leqslant C_q$.*

**Proposition 2.3** (Microscopic approximate loop equation for Wigner matrices). *Consider real Wigner matrices $H$ as in Definition 2.2, with eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$. Fix $E \in (-2, 2)$, and set*

$$\varrho_E := \varrho_{\mathrm{sc}}(E) := \frac{\sqrt{4 - E^2}}{2\pi}. \tag{2.25}$$

*For $w \in \mathbb{C} \setminus \mathbb{R}$, define the rescaled Stieltjes transform*

$$s_N(w) := \sum_{i=1}^N \frac{1}{N\pi\varrho_E(\lambda_i - E) - w} + \frac{E}{2\pi\varrho_E}. \tag{2.26}$$

*Then, for every $p \geqslant 1$ and every compact set $K \subset \mathbb{C} \setminus \mathbb{R}$, uniformly for $w, w_2, \ldots, w_p \in K$, we have*

$$\begin{aligned} &\mathbb{E}\left[\left(\partial_w s_N(w) + s_N(w)^2 + 1\right) \prod_{j=2}^p s_N(w_j)\right] \\ &\qquad + 2\mathbb{E}\left[\sum_{j=2}^p \partial_{w_j} \frac{s_N(w) - s_N(w_j)}{w - w_j} \prod_{\substack{i=2 \\ i \neq j}}^p s_N(w_i)\right] = \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[\mathcal{R}_N^{p+1}\right]\right), \end{aligned} \tag{2.27}$$

*where*

$$\mathcal{R}_N := \max\left\{|s_N(w)|, |s_N(w_2)|, \ldots, |s_N(w_p)|\right\}.$$

*As a consequence, by Theorem 1.9 with $\beta = 1$, the rescaled bulk eigenvalue statistics converge to the Sine$_1$ point process.*

#### 2.3.1 Preparatory estimates.

We recall the following local semicircle law for Wigner matrices from [34, Theorem 2.3].

**Theorem 2.4** ( [34, Theorem 2.3]). *Consider a real Wigner matrix as in Definition 2.2. Fix $\gamma > 0$. Then, for any small $\varepsilon > 0$ and large $D > 0$, the following holds with probability at least $1 - N^{-D}$, provided $N$ is sufficiently large. Uniformly for spectral parameters $z = u + \mathrm{i}\eta$ satisfying $|u| \leqslant 10$ and $N^{-1+\gamma} \leqslant \eta \leqslant 10$,*

$$\max_{i,j} |G_{ij}(z) - \delta_{ij} m_{\mathrm{sc}}(z)| \leqslant N^\varepsilon \left(\sqrt{\frac{\mathrm{Im}[m_{\mathrm{sc}}(z)]}{N\eta}} + \frac{1}{N\eta}\right). \tag{2.28}$$

Fix, once and for all, a small exponent

$$0 < \gamma < \frac{1}{12}.$$

Choosing $\varepsilon < \gamma/2$ in Theorem 2.4, using complex conjugation to pass from the upper half-plane to the lower half-plane, and using the moment assumption in Definition 2.2, we obtain the following good event. For every fixed $D > 0$, after increasing the moment order in the entry estimate if necessary,

$$\mathbb{P}(\Omega_N^{\complement}) \leqslant N^{-D},$$

where

$$\Omega_N := \left\{ \begin{array}{l} H : \max_{i,j} |h_{ij}| \leqslant N^{-1/2+\gamma}, \\ \max_{i,j} |G_{ij}(z)| \leqslant 2 \quad \text{for all } z = u + \mathrm{i}\eta,\ |u| \leqslant 10,\ N^{-1+\gamma} \leqslant |\eta| \leqslant 10 \end{array} \right\}. \tag{2.29}$$

**Lemma 2.5.** *On the event $\Omega_N$, for every $z = u + \mathrm{i}\eta \in \mathbb{C} \setminus \mathbb{R}$ with $|u| \leqslant 10$,*

$$\max_{i,j} |G_{ij}(z)| \leqslant \Lambda(z), \qquad \Lambda(z) := 2 \cdot \max\left\{1, \frac{N^\gamma}{N|\eta|}\right\}. \tag{2.30}$$

*In particular, if $z = E + w/(N\pi\varrho_E)$ with $w$ in a compact subset $K \subset \mathbb{C} \setminus \mathbb{R}$, then*

$$\Lambda(z) \leqslant C_K N^\gamma.$$

*Proof.* It is enough to prove the estimate in the upper half-plane; the lower half-plane follows by complex conjugation. Fix $|u| \leqslant 10$ and define

$$M(s) := \max_{a,b} |G_{ab}(u + \mathrm{i}s)|, \qquad s > 0.$$

If $N^{-1+\gamma} \leqslant s \leqslant 10$, then the bound

$$M(s) \leqslant 2 \tag{2.31}$$

holds directly from the definition of $\Omega_N$. If $s > 10$, then the trivial resolvent bound gives $M(s) \leqslant s^{-1} \leqslant 1$.

For $s > 0$, the identity $\partial_s G(u + \mathrm{i}s) = \mathrm{i}G(u + \mathrm{i}s)^2$ and the Ward identity imply

$$\begin{aligned} |\partial_s G_{ab}(u + \mathrm{i}s)| &\leqslant \sum_{k=1}^N |G_{ak}(u + \mathrm{i}s)|\,|G_{kb}(u + \mathrm{i}s)| \\ &\leqslant \left(\sum_k |G_{ak}(u + \mathrm{i}s)|^2\right)^{1/2} \left(\sum_k |G_{kb}(u + \mathrm{i}s)|^2\right)^{1/2} \\ &= \frac{\sqrt{\operatorname{Im}[G_{aa}(u + \mathrm{i}s)]\ \operatorname{Im}[G_{bb}(u + \mathrm{i}s)]}}{s} \leqslant \frac{M(s)}{s}. \end{aligned} \tag{2.32}$$

Now suppose $0 < s \leqslant N^{-1+\gamma}$. For any indices $a, b$,

$$|G_{ab}(u + \mathrm{i}s)| \leqslant |G_{ab}(u + \mathrm{i}N^{-1+\gamma})| + \int_s^{N^{-1+\gamma}} |\partial_r G_{ab}(u + \mathrm{i}r)|\,\mathrm{d}r.$$

Using (2.32), we get

$$M(s) \leqslant 2 + \int_s^{N^{-1+\gamma}} \frac{M(r)}{r}\,\mathrm{d}r.$$

By Gronwall's inequality,

$$M(s) \leqslant 2\exp\left(\int_s^{N^{-1+\gamma}} \frac{\mathrm{d}r}{r}\right) = \frac{2N^\gamma}{Ns}. \tag{2.33}$$

Combining (2.31) and (2.33) proves (2.30). ☐

We recall the following cumulant expansion from [51, Lemma 2.4].

**Lemma 2.6** ( [51, Lemma 2.4]). *Let $h$ be a centered real random variable with*

$$\mathbb{E}h = 0, \qquad \mathbb{E}h^2 = \frac{1}{N},$$

*and finite moments of all orders. Let $f : \mathbb{R} \to \mathbb{C}$ be smooth. Then*

$$\mathbb{E}[hf(h)] = \frac{1}{N}\mathbb{E}[f'(h)] + \mathcal{R},$$

*where*

$$|\mathcal{R}| \leqslant \mathbb{E}\left[\left(N^{-1}|h| + |h|^3\right) \sup_{|x|\leqslant|h|} |f''(x)|\right].$$

For $i \leqslant j$ and a real parameter $x$, define

$$(\Delta_{ij})_{ab} := \delta_{ia}\delta_{jb} + \mathbf{1}(i \neq j)\delta_{ib}\delta_{ja}, \qquad 1 \leqslant a, b \leqslant N, \tag{2.34}$$

and

$$H^{(ij,x)} := H + (x - h_{ij})\Delta_{ij}, \quad G^{(ij,x)}(z) := (H^{(ij,x)} - z)^{-1}, \quad m_N^{(ij,x)}(z) := \frac{1}{N}\operatorname{Tr} G^{(ij,x)}(z). \tag{2.35}$$

For a differentiable function $F$ of the matrix entries, set

$$\partial_{ij}F(H) := \frac{\mathrm{d}}{\mathrm{d}x}F\big(H^{(ij,x)}\big)\Big|_{x=h_{ij}}, \qquad i \leqslant j,$$

and use the convention $\partial_{ji} := \partial_{ij}$ for $i \leqslant j$.

**Lemma 2.7** (Single-entry resolvent stability). *Assume that, for a spectral parameter $z \in \mathbb{C} \setminus \mathbb{R}$,*

$$\max_{a,b}|G_{ab}(z)| \leqslant \Lambda, \qquad \max_{i,j}|h_{ij}| \leqslant \Lambda N^{-1/2}, \qquad \Lambda \leqslant CN^{\gamma},$$

*where $\gamma > 0$ is sufficiently small. Then, uniformly in all indices $a, b, i, j \in [\![N]\!]$,*

$$\sup_{|x|\leqslant\Lambda N^{-1/2}} \left|G_{ab}^{(ij,x)}(z) - G_{ab}(z)\right| \leqslant CN^{-1/2}\Lambda^3, \tag{2.36}$$

*and*

$$\sup_{|x|\leqslant\Lambda N^{-1/2}} \left|m_N^{(ij,x)}(z) - m_N(z)\right| \leqslant CN^{-1/2}\Lambda^3. \tag{2.37}$$

*Proof.* We use the finite-rank form of the resolvent identity. For $i \neq j$, write

$$\Delta_{ij} = UV^\top, \qquad U = (e_i, e_j), \qquad V = (e_j, e_i).$$

For $i = j$, the same argument applies with $U = V = e_i$. By the Woodbury formula,

$$G^{(ij,x)} = G - (x - h_{ij})GU\big(\mathbb{I} + (x - h_{ij})V^\top GU\big)^{-1}V^\top G.$$

Every entry of $GU$, $V^\top G$, and $V^\top GU$ is a linear combination of at most two entries of $G$. Hence

$$\max_{\alpha,\beta}|(V^\top GU)_{\alpha\beta}| \leqslant C\Lambda.$$

Since by our assumption,$|x| + |h_{ij}| \leqslant 2\Lambda N^{-1/2}$, we have

$$|x - h_{ij}|\Lambda \leqslant 2\Lambda^2 N^{-1/2} \leqslant 2N^{-1/2+2\gamma},$$

the matrix $\mathbb{I} + (x - h_{ij})V^\top GU$ is invertible for $N$ large, provided $\gamma < 1/4$, and its inverse has norm at most $C$. Therefore, for any $a, b$,

$$|(G^{(ij,x)} - G)_{ab}| \leqslant C|x - h_{ij}| \max_{\alpha}|(GU)_{a\alpha}| \max_{\beta}|(V^\top G)_{\beta b}| \leqslant C|x - h_{ij}|\Lambda^2 \leqslant CN^{-1/2}\Lambda^3.$$

This proves (2.36), and (2.37) follows by taking the normalized trace. $\square$

**2.3.2 Proof of Proposition 2.3.** Proposition 2.3 follows from cumulant expansion, and the following cumulant error estimate. We postpone its proof to the end of this section.

**Lemma 2.8** (Cumulant error estimate)**.** *Fix $p \geqslant 1$ and a compact set $K \subset \mathbb{C}\setminus\mathbb{R}$. Let $w, w_2, w_3, \ldots, w_p \in K$, and set*

$$z := E + \frac{w}{N\pi\varrho_E}, \quad z_a := E + \frac{w_a}{N\pi\varrho_E}, \qquad 2 \leqslant a \leqslant p. \tag{2.38}$$

*Let*

$$M_p := \prod_{a=2}^{p} m_N(z_a), \qquad R_N := \max\{|m_N(z)|, |m_N(z_2)|, \cdots, |m_N(z_p)|\}.$$

*For $i, j \in [\![N]\!]$, define*

$$f_{ij}(x) := G_{ij}^{(ij,x)}(z) \prod_{a=2}^{p} m_N^{(ij,x)}(z_a).$$

*Then*

$$\frac{1}{N}\sum_{i,j=1}^{N} \mathbb{E}\left[\left(N^{-1}|h_{ij}| + |h_{ij}|^3\right) \sup_{|x|\leqslant|h_{ij}|} |\partial_x^2 f_{ij}(x)|\right] + \frac{1}{N^2}\sum_{i=1}^{N} |\mathbb{E}\left[\partial_{ii}\left(G_{ii}(z)M_p\right)\right]| = \mathrm{o}_N(1)\left(1 + \mathbb{E}[R_N^{p+1}]\right) \tag{2.39}$$

*uniformly for $w, w_2, w_3, \ldots, w_p \in K$.*

*Proof of Proposition 2.3.* Fix $w, w_2, w_3, \cdots, w_p \in K$, and define $z, z_2, z_3 \cdots, z_p$ as in (2.38). Also set

$$M_p := \prod_{a=2}^{p} m_N(z_a), \qquad M_p^{(a)} := \prod_{\substack{b=2\\ b\neq a}}^{p} m_N(z_b), \qquad R_N := \max\{|m_N(z)|, |m_N(z_2)|, \cdots, |m_N(z_p)|\}.$$

From $(H - z)G(z) = \mathbb{I}$, taking the normalized trace gives

$$1 + zm_N(z) = \frac{1}{N}\sum_{i,j=1}^{N} h_{ij}G_{ji}(z).$$

Therefore

$$\mathbb{E}\left[\left(1 + zm_N(z) + m_N^2(z)\right) M_p\right] = \frac{1}{N}\sum_{i,j=1}^{N} \mathbb{E}\left[h_{ij}G_{ji}(z)M_p\right] + \mathbb{E}\left[m_N^2(z)M_p\right]. \tag{2.40}$$

We apply Lemma 2.6 to the independent upper-triangular entries. Rewriting the result as an ordered double sum, and adding one harmless diagonal copy in order to use uniform derivative identities, gives

$$\frac{1}{N}\sum_{i,j=1}^{N} \mathbb{E}\left[h_{ij}G_{ij}(z)M_p\right] = \frac{1}{N^2}\sum_{i,j=1}^{N} (1 + \delta_{ij})\mathbb{E}\left[\partial_{ij}\left(G_{ij}(z)M_p\right)\right] + \mathcal{E}_N, \tag{2.41}$$

where the additional diagonal contribution is included in $\mathcal{E}_N$, so it is given by (2.39). By Lemma 2.8,

$$|\mathcal{E}_N| \leqslant \mathrm{o}_N(1)\left(1 + \mathbb{E}[R_N^{p+1}]\right), \tag{2.42}$$

uniformly for $w, w_2, w_3, \cdots, w_p \in K$.

We now compute the main term in (2.41). The resolvent derivative identity gives

$$\partial_{ij}G_{kl}(z) = -G_{ki}(z)G_{jl}(z) - \mathbf{1}(i \neq j)G_{kj}(z)G_{il}(z).$$

Consequently,

$$(1+\delta_{ij})\partial_{ij}G_{ij}(z) = -G_{ii}(z)G_{jj}(z) - G_{ij}^2(z),$$

and, for every $2 \leqslant a \leqslant p$,

$$(1+\delta_{ij})\partial_{ij}m_N(z_a) = -\frac{2}{N}G_{ij}^2(z_a).$$

Therefore

$$(1+\delta_{ij})\partial_{ij}\left(G_{ij}(z)M_p\right) = \left(-G_{ii}(z)G_{jj}(z) - G_{ij}^2(z)\right)M_p - \frac{2G_{ij}(z)}{N}\sum_{a=2}^{p} G_{ij}^2(z_a)M_p^{(a)}. \tag{2.43}$$

Summing over $i, j$, we obtain

$$\frac{1}{N^2}\sum_{i,j=1}^{N}(1+\delta_{ij})\partial_{ij}\left(G_{ij}(z)M_p\right) = -m_N^2(z)M_p - \frac{1}{N^2}\operatorname{Tr}G^2(z)\,M_p - \frac{2}{N^3}\sum_{a=2}^{p}\operatorname{Tr}(G(z)G^2(z_a))M_p^{(a)}. \tag{2.44}$$

Using

$$\partial_z m_N(z) = \frac{1}{N}\operatorname{Tr}G^2(z), \quad \partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a} = \frac{1}{N}\operatorname{Tr}(G(z)G^2(z_a)), \qquad 2 \leqslant a \leqslant p,$$

we get

$$\begin{aligned}\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbb{E}\left[(1+\delta_{ij})\partial_{ij}\left(G_{ij}(z)M_p\right)\right] &= -\mathbb{E}[m_N^2(z)M_p] - \frac{1}{N}\mathbb{E}\left[(\partial_z m_N(z))M_p\right] \\ &\quad - \frac{2}{N^2}\sum_{a=2}^{p}\mathbb{E}\left[\partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a}M_p^{(a)}\right].\end{aligned} \tag{2.45}$$

Substituting this into (2.40) and (2.41), the term $-\mathbb{E}[m_N^2(z)M_p]$ cancels the term $+\mathbb{E}[m_N^2(z)M_p]$. Hence

$$\begin{aligned}\mathbb{E}\left[\left(1+zm_N(z)+m_N(z)^2\right)M_p\right] &= -\frac{1}{N}\mathbb{E}\left[\left(\partial_z m_N(z)\right)M_p\right] \\ &\quad - \frac{2}{N^2}\sum_{a=2}^{p}\mathbb{E}\left[\partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a}M_p^{(a)}\right] + \mathrm{o}_N(1)\left(1+\mathbb{E}\left[R_N^{p+1}\right]\right).\end{aligned} \tag{2.46}$$

Starting from the loop equation in the form (2.46), by linearity of the loop equation, we may replace $M_p$ by $\prod_{a=2}^{p}(m_N(z_a)+E/2)$, and get

$$\begin{aligned}&\mathbb{E}\left[\left(1+zm_N(z)+m_N(z)^2+\frac{1}{N}\partial_z m_N(z)\right)\prod_{a=2}^{p}\left(m_N(z_a)+\frac{E}{2}\right)\right] \\ &\quad + \frac{2}{N^2}\sum_{a=2}^{p}\mathbb{E}\left[\partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a}\prod_{\substack{b=2\\ b\neq a}}^{p}\left(m_N(z_b)+\frac{E}{2}\right)\right] = \mathrm{o}_N(1)\left(1+\mathbb{E}\left[R_N^{p+1}\right]\right).\end{aligned} \tag{2.47}$$

We now pass to microscopic variables. Recall from (2.26) and (2.38) that

$$m_N(z) = \pi\varrho_E s_N(w) - \frac{E}{2}, \qquad m_N(z_a) = \pi\varrho_E s_N(w_a) - \frac{E}{2}, \quad 2 \leqslant a \leqslant p,$$

and the chain rule gives

$$\frac{1}{N}\partial_z m_N(z) = (\pi\varrho_E)^2\partial_w s_N(w).$$

Moreover, by (2.25),

$$(\pi\varrho_E)^2 = 1 - \frac{E^2}{4}.$$

Therefore

$$\begin{aligned}
&1 + zm_N(z) + m_N(z)^2 + \frac{1}{N}\partial_z m_N(z) \\
&= 1 + \left(E + \frac{w}{N\pi\varrho_E}\right)\left(\pi\varrho_E s_N(w) - \frac{E}{2}\right) + \left(\pi\varrho_E s_N(w) - \frac{E}{2}\right)^2 + (\pi\varrho_E)^2\partial_w s_N(w) \\
&= (\pi\varrho_E)^2\left(1 + s_N(w)^2 + \partial_w s_N(w)\right) + \frac{w}{N}\left(s_N(w) - \frac{E}{2\pi\varrho_E}\right).
\end{aligned} \tag{2.48}$$

Similarly, for $a \geqslant 2$,

$$\partial_{z_a}\frac{m_N(z) - m_N(z_a)}{z - z_a} = N^2(\pi\varrho_E)^3\partial_{w_a}\frac{s_N(w) - s_N(w_a)}{w - w_a}. \tag{2.49}$$

Substituting (2.48) and (2.49) into (2.47), using $R_N \lesssim 1 + \mathcal{R}_N$, and dividing by $(\pi\varrho_E)^{p+1}$, we get

$$\begin{aligned}
&\mathbb{E}\left[\left(\partial_w s_N(w) + s_N(w)^2 + 1\right)\prod_{a=2}^{p} s_N(w_a)\right] + 2\mathbb{E}\left[\sum_{a=2}^{p}\left(\partial_{w_a}\frac{s_N(w) - s_N(w_a)}{w - w_a}\right)\prod_{\substack{b=2 \\ b\neq a}}^{p} s_N(w_b)\right] \\
&= \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[\mathcal{R}_N^{p+1}\right]\right) - \frac{w}{N(\pi\varrho_E)^2}\mathbb{E}\left[\left(s_N(w) - \frac{E}{2\pi\varrho_E}\right)\prod_{a=2}^{p} s_N(w_a)\right].
\end{aligned} \tag{2.50}$$

It remains to absorb the last term. Since $K \subset \mathbb{C}\backslash\mathbb{R}$ is compact, $w$ is uniformly bounded on $K$. Moreover,

$$\left|\left(s_N(w) - \frac{E}{2\pi\varrho_E}\right)\prod_{a=2}^{p} s_N(w_a)\right| \leqslant C\left(1 + \mathcal{R}_N^{p-1} + \mathcal{R}_N^{p}\right) \leqslant C'\left(1 + \mathcal{R}_N^{p+1}\right).$$

Thus

$$\frac{|w|}{N(\pi\varrho_E)^2}\mathbb{E}\left[\left|\left(s_N(w) - \frac{E}{2\pi\varrho_E}\right)\prod_{a=2}^{p} s_N(w_a)\right|\right] = \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[\mathcal{R}_N^{p+1}\right]\right).$$

The final term on the right-hand side of (2.50) can therefore be absorbed into the error, yielding (2.27). $\square$

*Proof of Lemma 2.8.* Since $K \subset \mathbb{C}\setminus\mathbb{R}$ is compact, there are constants $0 < c_K < C_K < \infty$ such that

$$c_K N^{-1} \leqslant |\operatorname{Im}[z]|, |\operatorname{Im}[z_a]| \leqslant C_K N^{-1}, \qquad 2 \leqslant a \leqslant p.$$

On $\Omega_N$, Lemma 2.5 and Lemma 2.7 imply that, for $|x| \leqslant |h_{ij}| \leqslant N^{-1/2+\gamma}$ and all $z$

$$\max_{k,l}|G_{kl}^{(ij,x)}(z)| \leqslant C_K N^{\gamma} =: \Lambda.$$

The resolvent derivative identities are

$$\begin{aligned}
\partial_x G^{(ij,x)}(z) &= -G^{(ij,x)}(z)\Delta_{ij}G^{(ij,x)}(z), \\
\partial_x^2 G^{(ij,x)}(z) &= 2G^{(ij,x)}(z)\Delta_{ij}G^{(ij,x)}(z)\Delta_{ij}G^{(ij,x)}(z).
\end{aligned} \tag{2.51}$$

Therefore

$$|\partial_x^r G_{kl}^{(ij,x)}(z)| \leqslant C\Lambda^{r+1}, \qquad r = 0, 1, 2.$$

For the normalized trace we use the Ward identity. Since $N|\operatorname{Im} z_a| \asymp_K 1$,

$$|\partial_x m_N^{(ij,x)}(z_a)| \leqslant C_K \Lambda, \qquad |\partial_x^2 m_N^{(ij,x)}(z_a)| \leqslant C_K \Lambda^2.$$

Consequently, by the Leibniz rule,

$$\mathbf{1}(\Omega_N) \sup_{|x| \leqslant |h_{ij}|} |\partial_x^2 f_{ij}(x)| \leqslant C_{p,K} \Lambda^3 (1+R_N)^{p-1}. \tag{2.52}$$

On $\Omega_N$, we also have

$$N^{-1}|h_{ij}| + |h_{ij}|^3 \leqslant C N^{-3/2+3\gamma}.$$

Thus

$$\begin{aligned} &\frac{1}{N} \sum_{i,j} \mathbb{E}\left[\mathbf{1}(\Omega_N)\left(N^{-1}|h_{ij}| + |h_{ij}|^3\right) \sup_{|x| \leqslant |h_{ij}|} |\partial_x^2 f_{ij}(x)|\right] \\ &\qquad \leqslant C_{p,K} N^{-1/2+6\gamma} \mathbb{E}\left[(1+R_N)^{p-1}\right] = \mathrm{o}_N(1) \mathbb{E}\left[(1+R_N)^{p-1}\right], \end{aligned} \tag{2.53}$$

because $\gamma < 1/12$.

On $\Omega_N^{\complement}$, we use the trivial resolvent bound

$$\|G^{(ij,x)}(z_a)\| \leqslant |\operatorname{Im} z_a|^{-1} \leqslant C_K N.$$

Hence

$$\sup_{|x| \leqslant |h_{ij}|} |\partial_x^2 f_{ij}(x)| \leqslant C_{p,K} N^{p+2}.$$

By Hölder's inequality, the moment assumptions on the entries, and the bound $\mathbb{P}(\Omega_N^{\complement}) \leqslant N^{-D}$ with $D$ chosen sufficiently large, we get

$$\frac{1}{N} \sum_{i,j} \mathbb{E}\left[\mathbf{1}(\Omega_N^{\complement})\left(N^{-1}|h_{ij}| + |h_{ij}|^3\right) \sup_{|x| \leqslant |h_{ij}|} |\partial_x^2 f_{ij}(x)|\right] = \mathrm{o}_N(1). \tag{2.54}$$

It remains to estimate the diagonal correction. On $\Omega_N$, the preceding derivative bounds give

$$|\partial_{ii}(G_{ii}(z)M_p)| \leqslant C_{p,K} \Lambda^2 (1+R_N)^{p-1}.$$

Therefore

$$\frac{1}{N^2} \sum_i \mathbb{E}\left[\mathbf{1}(\Omega_N) |\partial_{ii}(G_{ii}(z)M_p)|\right] \leqslant C_{p,K} N^{-1+2\gamma} \mathbb{E}\left[(1+R_N)^{p-1}\right]. \tag{2.55}$$

The contribution of $\Omega_N^{\complement}$ is again $\mathrm{o}_N(1)$ by the trivial resolvent bound and Hölder's inequality. Combining (2.53), (2.54), and (2.55), and using $(1+R_N)^{p-1} \leqslant C_p(1+R_N^{p+1})$, proves (2.39). □

**Remark 2.9** (The role of Quantum Unique Ergodicity)**.** *For symmetric random matrices with independent entries and a general expectation or variance profile, the approximate loop equations can also be established provided one more a priori estimate is available, namely Quantum Unique Ergodicity (QUE) [21].*

*To illustrate this, consider centered random matrices and denote $s_{ij} = \mathbb{E}[|h_{ij}|^2]$, normalized so that $\sum_j s_{ij} = 1$ for all $i$. Then the main term on the right-hand side of (2.41) becomes*

$$\frac{1}{N} \sum_{i,j=1}^{N} s_{ij} \mathbb{E}\left[\partial_{ij}\left(G_{ij}(z) M_p\right)\right].$$

*Summing over $j$ the $G_{ii}G_{jj}$ term in the analogue of (2.43) yields the following multiple of $M_p$, where $v_\alpha$ denotes the normalized eigenvector associated with the eigenvalue $\lambda_\alpha$:*

$$G_{ii}\sum_{j=1}^N s_{ij}G_{jj} = G_{ii}\sum_\alpha\sum_{j=1}^N \frac{s_{ij}v_\alpha(j)^2}{\lambda_\alpha - z} = G_{ii}m_N + \sum_\alpha \frac{1}{\lambda_\alpha - z}\sum_j s_{ij}\Big(v_\alpha(j)^2 - \frac{1}{N}\Big).$$

*QUE implies that the latter centered sum exhibits cancellations, namely it is $\mathrm{O}(N^{-1-\varepsilon})$ for some $\varepsilon > 0$, with overwhelming probability. Applying this estimate to all the terms arising from the summation in (2.43), one obtains the approximate loop equations.*

*For models with spatial dependence, including random band matrices, QUE is known to be essential in the third step of the three-step strategy; see, for example, [91] and the references therein. QUE itself follows from multi-point resolvent estimates on mesoscopic scales [23].*

## 2.4 Random $d$-regular graphs

In this section, we apply our characterization to random $d$-regular graphs. We first prove bulk universality for random $d$-regular graphs on $N$ vertices in the growing-degree regime $d \gg (\log N)^{24}$ in Section 2.4.1. We then explain how the proof of edge universality for fixed-degree random $d$-regular graphs in [52] can be simplified using our new characterization in Section 2.4.6.

**2.4.1 Bulk loop equation.** Let $A$ be the adjacency matrix of the uniform simple $d$-regular graph on $[\![N]\!]$, and introduce the normalized adjacency matrix

$$H := \frac{A}{\sqrt{d-1}}. \tag{2.56}$$

We give a short proof of bulk universality for random $d$-regular graphs on $N$ vertices in the regime $d \gg (\log N)^{24}$. This extends the previous bulk universality results [7, 63], which treated the range $N^{\mathfrak{c}} \leqslant d \leqslant N^{2/3-\mathfrak{c}}$, to degrees growing as slowly as a polylogarithm of $N$. By Theorem 1.9, bulk universality follows from the following approximate local bulk loop equation. Our main focus is on the sparse regime, for simplicity we assume that $d \leqslant N^{1/2}$.

**Proposition 2.10** (Microscopic approximate loop equation for random $d$-regular graphs)**.** *Consider the uniform simple random $d$-regular graph in the regime*

$$(\log N)^{24} \ll d \leqslant N^{1/2}.$$

*We denote the eigenvalues of the rescaled adjacency matrix $H$ (recall from (2.56)) by $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$.*

*Fix $E \in (-2,2)$, and set*

$$\varrho_E := \varrho_{\mathrm{sc}}(E) := \frac{\sqrt{4-E^2}}{2\pi}, \tag{2.57}$$

*For $w \in \mathbb{C}\setminus\mathbb{R}$, define the rescaled Stieltjes transform*

$$s_N(w) := \sum_{i=1}^N \frac{1}{N\pi\varrho_E(\lambda_i - E) - w} + \frac{E}{2\pi\varrho_E} \tag{2.58}$$

*Then, for every $p \geqslant 1$ and every compact set $K \subset \mathbb{C}\setminus\mathbb{R}$, uniformly for $w, w_2, w_3, \dots, w_p \in K$, we have*

$$\begin{aligned}\mathbb{E}&\left[\left(\partial_w s_N(w) + s_N(w)^2 + 1\right)\prod_{j=2}^p s_N(w_j)\right] \\ &+ 2\mathbb{E}\left[\sum_{j=2}^p \partial_{w_j}\frac{s_N(w) - s_N(w_j)}{w - w_j}\prod_{\substack{i=2\\ i\neq j}}^p s_N(w_i)\right] = \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[\mathcal{R}_N^{p+1}\right]\right),\end{aligned} \tag{2.59}$$

*where* $\mathcal{R}_N := \max\{|s_N(w)|, |s_N(w_2)|, \ldots, |s_N(w_p)|\}$. *As a consequence, by Theorem 1.9 with* $\beta = 1$*, the rescaled bulk eigenvalue statistics converge to the* $\mathit{Sine}_1$ *point process.*

The exponent in the condition $d \gg (\log N)^{24}$ is not optimal, and we do not attempt to optimize it. In fact, bulk universality is expected to hold for every fixed degree $d \geqslant 3$ [56].

**Conjecture 2.11.** *For every fixed integer* $d \geqslant 3$*, the local statistics of the bulk eigenvalues of random* $d$*-regular graphs on* $N$ *vertices converge, as* $N \to \infty$*, to those of the* GOE.

### 2.4.2 Local law.

In this section, we recall the local law for random $d$-regular graphs with polylogarithmic degree growth from [9]. We recall the normalized adjacency matrix $H$ from (2.56), and define the Green's function and Stieltjes transform

$$G(z) := (H - z)^{-1}, \qquad m_N(z) := \frac{1}{N} \operatorname{Tr} G(z). \tag{2.60}$$

In the rest of the proof, we will repeatedly use the following relations

$$\partial_z G(z) = G(z)^2, \quad G(z)\mathbf{1} = \frac{1}{d/\sqrt{d-1} - z}\mathbf{1}, \quad AG(z) = \sqrt{d-1}\big(\mathbb{I} + zG(z)\big), \tag{2.61}$$

where $\mathbf{1}$ denotes the trivial all-ones eigenvector of $H$, and the Ward identities

$$\sum_{j=1}^N |G_{ij}(z)|^2 = \sum_{j=1}^N |G_{ji}(z)|^2 = \frac{\operatorname{Im}[G_{ii}(z)]}{\operatorname{Im}[z]}, \quad \frac{1}{N} \sum_{i,j=1}^N |G_{ij}(z)|^2 = \frac{\operatorname{Im}[m_N(z)]}{\operatorname{Im}[z]}. \tag{2.62}$$

For $z \in \mathbb{C}_+$, let

$$\varrho_{\mathrm{sc}}(x) := \frac{1}{2\pi}\sqrt{(4 - x^2)_+}, \qquad m_{\mathrm{sc}}(z) := \int_{\mathbb{R}} \frac{\varrho_{\mathrm{sc}}(x)}{x - z}\,\mathrm{d}x = \frac{-z + \sqrt{z^2 - 4}}{2},$$

where the square root is chosen so that $\sqrt{z^2 - 4} \sim z$ as $|z| \to \infty$.

Define

$$D := d \wedge \frac{N^2}{d^3}, \qquad \Phi(z) := \frac{1}{\sqrt{N\eta}} + \frac{1}{\sqrt{D}},$$

and, for $r \in [0, 1]$,

$$F_z(r) := \left[\left(1 + \frac{1}{\sqrt{|z^2 - 4|}}\right) r\right] \wedge \sqrt{r}.$$

We recall the following local semicircle law for the uniform random regular graph from [9, Theorem 1.1].

**Theorem 2.12** ( [9, Theorem 1.1])**.** *Let* $G(z)$ *be the Green's function* (2.60) *of the uniform simple random* $d$*-regular graph. Let* $\xi = \xi_N$ *satisfy* $\xi \log \xi \gg (\log N)^2$*, and assume that* $\xi^2 \ll d \ll (N/\xi)^{2/3}$*. Then, with probability at least* $1 - \exp(-\xi \log \xi)$*,*

$$\max_i |G_{ii}(z) - m_{\mathrm{sc}}(z)| = \mathrm{O}\big(F_z(\xi\Phi(z))\big), \qquad \max_{i \neq j} |G_{ij}(z)| = \mathrm{O}\big(\xi\Phi(z)\big),$$

*simultaneously for all* $z = E + \mathrm{i}\eta \in \mathbb{C}_+$ *satisfying* $\eta \gg \xi^2/N$*.*

Choosing

$$\xi := \frac{(\log N)^2}{\sqrt{\log\log N}},$$

the assumptions of the local law Theorem 2.12 are satisfied whenever

$$d \geqslant (\log N)^4, \qquad d \leqslant \frac{N^{2/3}}{(\log N)^{4/3}}, \qquad \eta \geqslant \frac{(\log N)^4}{N}.$$

Therefore, with probability at least $1-\exp(-\xi\log\xi) \geqslant 1-e^{-(\log N)^2}$,

$$\max_i |G_{ii}(z)-m_{\rm sc}(z)| = \mathrm{O}\big(F_z(\xi\Phi(z))\big), \qquad \max_{i\neq j}|G_{ij}(z)| = \mathrm{O}\big(\xi\Phi(z)\big),$$

simultaneously for all $z=E+\mathrm{i}\eta\in\mathbb{C}_+$ with $\eta\geqslant(\log N)^4/N$. In particular,

$$\max_{i,j}|G_{ij}(z)|\leqslant 2$$

throughout this spectral domain.

We introduce the following event $\Omega_N$ of $d$-regular graphs on $N$ vertices:

$$\Omega_N := \left\{\mathcal{G} : \max_{i,j}|G_{ij}(z)|\leqslant 2 \text{ for all } z=E+\mathrm{i}\eta\in\mathbb{C}_+,\ \eta\geqslant\frac{(\log N)^4}{N}\right\}. \tag{2.63}$$

**Lemma 2.13.** *On the event $\Omega_N$, for any $z=E+\mathrm{i}\eta\in\mathbb{C}_+$,*

$$\max_{i,j}|G_{ij}(z)|\leqslant\Lambda(z), \qquad \Lambda(z):=2\cdot\max\left\{1,\frac{(\log N)^4}{N\eta}\right\}, \qquad \eta>0. \tag{2.64}$$

*Thus, if $z=E+w/(\pi\varrho_E N)$ with $w$ in a compact subset $K$ of $\mathbb{C}_+$, then*

$$\Lambda(z)\leqslant C_K(\log N)^4.$$

*Proof.* The proof is exactly the same as that of Lemma 2.5 after replacing $N^\gamma$ by $(\log N)^4$, so we omit it. □

**2.4.3 Switching calculus.** We define the signed adjacency matrices

$$(\Delta_{ij})_{ab}:=\delta_{ia}\delta_{jb}+\delta_{ib}\delta_{ja}, \qquad \xi_{ij}^{kl}:=\Delta_{ij}+\Delta_{kl}-\Delta_{ik}-\Delta_{jl}. \tag{2.65}$$

Thus $\Delta_{ij}$ is the adjacency matrix of the edge $ij$, while $\xi_{ij}^{kl}$ is the signed matrix associated with the switching of edges. For any function $F$ of the adjacency matrix $A$, define the discrete difference operator and the continuous differential operator by

$$(D_{ij}^{kl}F)(A):=F(A+\xi_{ij}^{kl})-F(A), \qquad (\partial_{ij}^{kl}F)(A):=\sqrt{d-1}\,\frac{\mathrm{d}}{\mathrm{d}t}F(A+t\xi_{ij}^{kl})\Big|_{t=0}. \tag{2.66}$$

We recall the following discrete integration-by-parts formula from [8, Corollary 3.2].

**Lemma 2.14** ( [8, Corollary 3.2])**.** *Fix indices $i,j\in[\![N]\!]$. Let $F=F_{ij}$ be a random variable, possibly depending on $i,j$. Define*

$$\mathcal{C}_{ij}(F,A):=|F(A)|+\max_{k,l}\left|F\big(A+\xi_{ij}^{kl}\big)\right|.$$

*Then for $i\neq j$ we have*

$$\mathbb{E}\left[A_{ij}F(A)\right]=\frac{d}{N}\,\mathbb{E}\left[F(A)\right]+\frac{1}{Nd}\sum_{k,l}\mathbb{E}\left[A_{ik}A_{jl}\left(F\big(A+\xi_{ij}^{kl}\big)-F(A)\right)\right]+\mathrm{O}\left(\frac{d}{N}\mathbb{E}\left[A_{ij}\mathcal{C}_{ij}(F,A)\right]\right). \tag{2.67}$$

The next lemma replaces the discrete switching derivative by its linearized continuous version.

**Lemma 2.15** (Resolvent expansion under one switching)**.** *Assume*

$$\max_{a,b}|G_{ab}(z)|\leqslant\Lambda, \qquad \frac{\Lambda}{\sqrt{d}}\ll 1.$$

*Then, uniformly in all indices, $a,b,i,j,k,l \in [\![N]\!]$, we have*

$$D_{ij}^{kl} G_{ab}(z) = \frac{1}{\sqrt{d-1}} \partial_{ij}^{kl} G_{ab}(z) + \mathrm{O}\left(\frac{\Lambda^3}{d}\right), \tag{2.68}$$

*and similarly*

$$D_{ij}^{kl} m_N(z) = \frac{1}{\sqrt{d-1}} \partial_{ij}^{kl} m_N(z) + \mathrm{O}\left(\frac{\Lambda^3}{d}\right). \tag{2.69}$$

*Moreover, the derivatives are given by*

$$\begin{aligned} \partial_{ij}^{kl} G(z) = -G(z)\xi_{ij}^{kl} G(z) = &-G_{ji}(z)^2 - G_{jj}(z)G_{ii}(z) - G_{jk}(z)G_{li}(z) - G_{jl}(z)G_{ki}(z) \\ &+ G_{ji}(z)G_{ki}(z) + G_{jl}(z)G_{ji}(z) + G_{jk}(z)G_{ii}(z) + G_{jj}(z)G_{li}(z), \end{aligned} \tag{2.70}$$

*and*

$$\partial_{ij}^{kl} m_N(w) = -\frac{1}{N} \operatorname{Tr}\left(G(w)\xi_{ij}^{kl} G(w)\right) = -\frac{2}{N}\left[(G(w)^2)_{ij} + (G(w)^2)_{kl} - (G(w)^2)_{ik} - (G(w)^2)_{jl}\right]. \tag{2.71}$$

*Proof.* We recall from (2.65) that $\xi_{ij}^{kl} = \Delta_{ij} + \Delta_{kl} - \Delta_{ik} - \Delta_{jl}$. Write

$$\xi_{ij}^{kl} = UV^\top, \qquad U = (e_i - e_l,\ e_j - e_k), \qquad V = (e_j - e_k,\ e_i - e_l).$$

By the Woodbury formula,

$$G(A + \xi_{ij}^{kl}) = G - \frac{1}{\sqrt{d-1}} GU \left(\mathbb{I}_2 + \frac{1}{\sqrt{d-1}} V^\top GU\right)^{-1} V^\top G. \tag{2.72}$$

Every entry of $V^\top GU$, $GU$, and $V^\top G$ is a linear combination of at most four entries of $G$. Hence, using $\Lambda/\sqrt{d} \ll 1$, the matrix

$$\mathbb{I}_2 + \frac{1}{\sqrt{d-1}} V^\top GU$$

is invertible, and the Neumann series gives

$$\left\| \left(\mathbb{I}_2 + \frac{1}{\sqrt{d-1}} V^\top GU\right)^{-1} - \mathbb{I}_2 \right\| \leqslant C \frac{\Lambda}{\sqrt{d}}.$$

Subtracting the linear term

$$-\frac{1}{\sqrt{d-1}} GUV^\top G = -\frac{1}{\sqrt{d-1}} G\xi_{ij}^{kl} G$$

the expression (2.72) gives

$$\left| D_{ij}^{kl} G_{ab} + \frac{1}{\sqrt{d-1}} (G\xi_{ij}^{kl} G)_{ab} \right| \leqslant C \frac{\Lambda^3}{d}.$$

Since

$$\partial_{ij}^{kl} G(z) = -G(z)\xi_{ij}^{kl} G(z),$$

this proves (2.68). Taking the normalized trace gives the claim for $m_N$. Finally, (2.70) and (2.71) follow from (2.61) by straightforward computations. □

**2.4.4 Proof of Proposition 2.10.** In this section we prove Proposition 2.10, using the local law Theorem 2.12 as input, the bulk loop equation follows from the discrete integration by part formula Lemma 2.14.

For $w, w_2, w_3, \dots, w_p \in K$, set

$$z = E + \frac{w}{N\pi\varrho_E}, \quad z_a = E + \frac{w_a}{N\pi\varrho_E}, \quad 2 \leqslant a \leqslant p, \tag{2.73}$$

and

$$M_p := \prod_{a=2}^{p} m_N(z_a), \qquad M_p^{(a)} := \prod_{\substack{b=2\\ b\neq a}}^{p} m_N(z_b), \qquad R_N := \max\{|m_N(z)|, |m_N(z_2)|, \dots, |m_N(z_p)|\}.$$

Taking the trace in the identity

$$HG(z) = \mathbb{I} + zG(z)$$

gives

$$1 + zm_N(z) = \frac{1}{N}\sum_{i,j} H_{ij}G_{ji}(z). \tag{2.74}$$

Multiplying (2.74) by $M_p$ and adding $m_N(z)^2 M_p$, we obtain

$$\mathbb{E}\left[\left(1 + zm_N(z) + m_N(z)^2\right) M_p\right] = \mathbb{E}\left[m_N(z)^2 M_p\right] + \frac{1}{N\sqrt{d-1}}\sum_{i,j}\mathbb{E}\left[A_{ij}G_{ji}(z)M_p\right]. \tag{2.75}$$

We apply the switching integration-by-parts formula Lemma 2.14 to the observable

$$F_{ij}(A) := G_{ji}(z)M_p.$$

This gives

$$\begin{aligned} \frac{1}{N\sqrt{d-1}}\sum_{i,j}\mathbb{E}\left[A_{ij}F_{ij}(A)\right] &= \frac{d}{N^2\sqrt{d-1}}\sum_{i\neq j}\mathbb{E}\left[F_{ij}(A)\right] \\ &+ \frac{1}{N^2 d\sqrt{d-1}}\sum_{i\neq j,k,l}\mathbb{E}\left[A_{ik}A_{jl}D_{ij}^{kl}F_{ij}(A)\right] + \mathrm{O}\left(\frac{\sqrt{d}}{N^2}\sum_{i\neq j}\mathbb{E}\left[A_{ij}\mathcal{C}_{ij}(F_{ij}, A)\right]\right). \end{aligned} \tag{2.76}$$

For the first two terms on the right-hand side of (2.76), the summations are over $i \neq j$, we can replace them by the corresponding summations with all $i, j$, together with an extra error

$$\frac{d}{N^2\sqrt{d-1}}\sum_{i}\mathbb{E}\left[F_{ii}(A)\right] + \frac{1}{N^2 d\sqrt{d-1}}\sum_{i,k,l}\mathbb{E}\left[A_{ik}A_{il}D_{ii}^{kl}F_{ii}(A)\right] = \mathrm{O}\left(\frac{\sqrt{d}}{N^2}\sum_{i}\mathbb{E}[\mathcal{C}_{ii}(F_{ii}, A)]\right) \tag{2.77}$$

Thus we can rewrite (2.76) as

$$\begin{aligned} \frac{1}{N\sqrt{d-1}}\sum_{i,j}\mathbb{E}\left[A_{ij}F_{ij}(A)\right] &= \frac{d}{N^2\sqrt{d-1}}\sum_{i,j}\mathbb{E}\left[F_{ij}(A)\right] \\ + \frac{1}{N^2 d\sqrt{d-1}}\sum_{i,j,k,l}\mathbb{E}\left[A_{ik}A_{jl}D_{ij}^{kl}F_{ij}(A)\right] &+ \mathrm{O}\left(\frac{\sqrt{d}}{N^2}\left(\sum_{i\neq j}\mathbb{E}[A_{ij}\mathcal{C}_{ij}(F_{ij}, A)] + \sum_{i}\mathbb{E}[\mathcal{C}_{ii}(F_{ii}, A)]\right)\right). \end{aligned} \tag{2.78}$$

The first term on the right-hand side of (2.78) is negligible. Indeed, by (2.61),

$$\sum_{i} G_{ji}(z) = \frac{1}{d/\sqrt{d-1} - z} = \mathrm{O}(d^{-1/2}), \qquad \sum_{i,j} G_{ji}(z) = \mathrm{O}\left(\frac{N}{\sqrt{d}}\right).$$

Therefore

$$\frac{d}{N^2\sqrt{d-1}}\sum_{i,j}\mathbb{E}\left[F_{ij}(A)\right] = \mathrm{O}\left(\frac{1}{N}\mathbb{E}\left[|M_p|\right]\right) = \mathrm{o}_N(1)\left(1+\mathbb{E}[R_N^{p+1}]\right). \tag{2.79}$$

We next control the integration-by-parts error in (2.78). Recall the event $\Omega_N$ from (2.63) and $\Lambda(z)$ from (2.64). For $z = E + w/(\pi\varrho_E N)$ with $w \in K$, we have

$$\Lambda(z) \leqslant C_K(\log N)^4, \qquad \frac{1}{N\,\mathrm{Im}[z]} \leqslant C_K, \qquad (\log N)^{24} \ll d \leqslant N^{1/2}. \tag{2.80}$$

On $\Omega_N$, by Lemma 2.15 and (2.64),

$$|G_{ji}(z)| + \max_{k,l}|G_{ji}(z; A+\xi_{ij}^{kl})| \leqslant C_K\Lambda(z), \quad |m_N(z_a)| + \max_{k,l}|m_N(z_a; A+\xi_{ij}^{kl})| \leqslant C_K\big(1+|m_N(z_a)|\big),$$

for $a = 2,\dots,p$. Thus

$$\mathbf{1}(\Omega_N)\mathcal{C}_{ij}(F,A) \leqslant C_K\Lambda(z)\prod_{a=2}^{p}\big(1+|m_N(z_a)|\big) \leqslant C_K\Lambda(z)\prod_{a=2}^{p}(1+R_N)^{p-1}. \tag{2.81}$$

On the complement $\Omega_N^{\complement}$, we use the trivial resolvent bound. Since $w, w_2, \dots, w_p \in K \subset \mathbb{C}\setminus\mathbb{R}$, all imaginary parts satisfy $|\mathrm{Im}[z_a]| \asymp_K N^{-1}$. Hence

$$|G_{uv}(z_a)| \leqslant \frac{C_K}{|\mathrm{Im}[z_a]|} \leqslant C_K N, \qquad |m_N(z_a)| \leqslant C_K N, \tag{2.82}$$

and the same bound holds for the switched matrix $A+\xi_{ij}^{kl}$. Therefore

$$\mathbf{1}(\Omega_N^{\complement})\mathcal{C}_{ij}(F,A) \leqslant \mathbf{1}(\Omega_N^{\complement})(C_K N)^p.$$

Taking expectations and using $\mathbb{P}(\Omega_N^{\complement}) \leqslant e^{-(\log N)^2}$, we obtain

$$\mathbb{E}\left[\mathbf{1}(\Omega_N^{\complement})\mathcal{C}_{ij}(F,A)\right] \leqslant (C_K N)^p e^{-(\log N)^2} = \mathrm{o}_N(1), \tag{2.83}$$

since $p$ is fixed.

Combining the estimates (2.81) and (2.83) on $\Omega_N$ and $\Omega_N^{\complement}$, and using $\sum_{i,j} A_{ij} = Nd$, we get

$$\frac{\sqrt{d}}{N^2}\sum_{i,j}\mathbb{E}\left[A_{ij}\mathcal{C}_{ij}(F,A)\right] \leqslant C_K\frac{d^{3/2}\Lambda(z)}{N}\mathbb{E}\left[(1+R_N)^{p-1}\right] + \mathrm{o}_N(1) = \mathrm{o}_N(1)\left(1+\mathbb{E}\left[R_N^{p+1}\right]\right), \tag{2.84}$$

where we used $d \leqslant N^{1/2}$, $\Lambda(z) \leqslant C_K(\log N)^4$, and $(1+R_N)^{p-1} \leqslant C_p(1+R_N^{p+1})$.

It remains to expand the discrete derivative in the main switching term. We write

$$D_{ij}^{kl}F_{ij}(A) = \frac{1}{\sqrt{d-1}}\left[\left(\partial_{ij}^{kl}G_{ji}(z)\right)M_p + G_{ji}(z)\sum_{a=2}^{p}\left(\partial_{ij}^{kl}m_N(z_a)\right)M_p^{(a)}\right] + \mathcal{E}_{ij}^{kl}. \tag{2.85}$$

All nonlinear terms in the discrete product rule and all Taylor remainders are included in $\mathcal{E}_{ij}^{kl}$.

We shall use the following deterministic estimates for the linearized switching terms in (2.85). The first estimate controls the Taylor and product-rule errors, while the last two identities identify the contributions of the derivatives falling on $G_{ji}(z)$ and on the external Stieltjes transforms. We postpone the proof to Section 2.4.5.

**Lemma 2.16.** *Adopt the notation and assumptions of Proposition 2.10, and let $\Lambda = \Lambda(z)$ be as in (2.64). For $z = E + w/(N\pi\varrho_E)$, $z_a = E + w_a/(N\pi\varrho_E)$, and $w, w_2, \cdots, w_p \in K \subset \mathbb{C}\setminus\mathbb{R}$, we have*

$$\frac{\mathbf{1}(\Omega_N)}{N^2 d\sqrt{d-1}} \sum_{i,j,k,l} A_{ik}A_{jl}|\mathcal{E}_{ij}^{kl}| \leqslant \frac{C_K\Lambda^3}{\sqrt{d}}(1+R_N)^{p-1}. \tag{2.86}$$

*Moreover, for $z \in \mathbb{C}_+$,*

$$\frac{1}{N^2 d(d-1)} \sum_{i,j,k,l} A_{ik}A_{jl}\partial_{ij}^{kl} G_{ji}(z) = -m_N(z)^2 - \frac{1}{N}\partial_z m_N(z) + \mathrm{O}\left(\frac{1}{d}\left(1 + |m_N(z)|^2 + \frac{|m_N(z)|}{N\,\mathrm{Im}[z]}\right)\right). \tag{2.87}$$

*Finally, for $z, w \in \mathbb{C}_+$, we have*

$$\begin{aligned}&\frac{1}{N^2 d(d-1)} \sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\,\partial_{ij}^{kl} m_N(w)\\ &\qquad = -\frac{2d}{N^2(d-1)}\partial_w \frac{m_N(z)-m_N(w)}{z-w} + \mathrm{O}\left(\frac{|m_N(z)|+|m_N(w)|}{N^2 d\ \mathrm{Im}[z]\ \mathrm{Im}[w]}\right).\end{aligned} \tag{2.88}$$

*If $w = z$, the right-hand side is understood by taking the limit $w \to z$.*

*Proof of Proposition 2.10.* Splitting the error term according to $\Omega_N$ and $\Omega_N^{\complement}$, using (2.86) and (2.80) such that $\Lambda(z) \leqslant C_K(\log N)^4 \ll d^{1/6}$ on $\Omega_N$, and using the trivial resolvent bound (2.82) on $\Omega_N^{\complement}$, we have

$$\frac{1}{N^2 d\sqrt{d-1}} \sum_{i,j,k,l} \mathbb{E}\left[A_{ik}A_{jl}\mathcal{E}_{ij}^{kl}\right] = \mathrm{o}_N(1)\left(1 + \mathbb{E}\left[R_N^{p+1}\right]\right).$$

Substituting (2.85) into (2.78), the main switching term becomes

$$\begin{aligned}&\frac{1}{N^2 d(d-1)} \sum_{i,j,k,l} \mathbb{E}\left[A_{ik}A_{jl}\left(\partial_{ij}^{kl}G_{ji}(z)\right)M_p\right]\\ &\quad + \frac{1}{N^2 d(d-1)} \sum_{a=2}^{p}\sum_{i,j,k,l} \mathbb{E}\left[A_{ik}A_{jl}G_{ji}(z)\left(\partial_{ij}^{kl} m_N(z_a)\right)M_p^{(a)}\right] + \mathrm{o}_N(1)\left(1+\mathbb{E}\left[R_N^{p+1}\right]\right).\end{aligned} \tag{2.89}$$

We now apply the two main estimates in Lemma 2.16. The estimate (2.87) gives

$$\frac{1}{N^2 d(d-1)} \sum_{i,j,k,l} A_{ik}A_{jl}\partial_{ij}^{kl}G_{ji}(z) = -m_N(z)^2 - \frac{1}{N}\partial_z m_N(z) + \mathrm{o}_N(1)\big(1+R_N^2\big), \tag{2.90}$$

after using (2.80). Multiplying by $M_p$ and taking expectations, the error contributes $\mathrm{o}_N(1)(1+\mathbb{E}[R_N^{p+1}])$.

Similarly, for each $a = 2, \ldots, p$, the estimate (2.88) gives

$$\frac{1}{N^2 d(d-1)} \sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\partial_{ij}^{kl} m_N(z_a) = -\frac{2}{N^2}\partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a} + \mathrm{o}_N(1)\big(1+R_N\big), \tag{2.91}$$

where the factor $d/(d-1) = 1 + \mathrm{O}(d^{-1})$ has been absorbed into the error. Again, after multiplying by $M_p^{(a)}$ and taking expectations, the error is $\mathrm{o}_N(1)(1+\mathbb{E}[R_N^{p+1}])$.

Combining (2.75), (2.79), (2.84), (2.89), (2.90), and (2.91), we get

$$\begin{aligned}\mathbb{E}\left[\left(1 + zm_N(z) + m_N(z)^2\right)M_p\right] &= \mathbb{E}\left[\left(-\frac{1}{N}\partial_z m_N(z)\right)M_p\right]\\ &\quad - \frac{2}{N^2}\sum_{a=2}^{p}\mathbb{E}\left[\partial_{z_a}\frac{m_N(z)-m_N(z_a)}{z-z_a}M_p^{(a)}\right] + \mathrm{o}_N(1)\left(1+\mathbb{E}\left[R_N^{p+1}\right]\right).\end{aligned} \tag{2.92}$$

It remains only to pass from (2.92) to microscopic variables. This is identical to the final step in the proof of Proposition 2.3, and hence yields (2.59). We omit the details. □

### 2.4.5 Proof of Lemma 2.16.

*Proof of Lemma 2.16.* We prove the three estimates separately.

*Step 1: The replacement error.*

On the event $\Omega_N$, we recall from Lemma 2.15 with $\Lambda = \Lambda(z) \leqslant C_K(\log N)^4 \ll \sqrt{d}$ as defined in (2.64), that

$$\begin{aligned} D_{ij}^{kl} G_{ji}(z) &= (d-1)^{-1/2}\partial_{ij}^{kl} G_{ji}(z) + \mathrm{O}\left(\frac{\Lambda^3}{d}\right), \\ D_{ij}^{kl} m_N(z_a) &= (d-1)^{-1/2}\partial_{ij}^{kl} m_N(z_a) + \mathrm{O}\left(\frac{\Lambda^3}{d}\right). \end{aligned} \tag{2.93}$$

Moreover using the explicit formulas (2.70) and (2.71),

$$|\partial_{ij}^{kl} G_{ji}(z)| \lesssim \Lambda^2, \qquad |\partial_{ij}^{kl} m_N(z_a)| \lesssim \frac{\Lambda}{N\,\mathrm{Im}[z_a]} \leqslant C_K\Lambda.$$

By definition,

$$\mathcal{E}_{ij}^{kl} = D_{ij}^{kl}\big(G_{ji}(z)M_p\big) - \frac{1}{\sqrt{d-1}}\partial_{ij}^{kl}\big(G_{ji}(z)M_p\big).$$

Using the discrete product rule and (2.93), we get

$$|\mathcal{E}_{ij}^{kl}| \lesssim |G_{ji}(z)|\left|D_{ij}^{kl}M_p - \frac{1}{\sqrt{d-1}}\partial_{ij}^{kl}M_p\right| + \frac{1}{\sqrt{d}}|\partial_{ij}^{kl}G_{ji}(z)|\,|D_{ij}^{kl}M_p| + \frac{\Lambda^3}{d}\prod_{a=2}^{p}\left(|m_N(z_a)| + \frac{C_K\Lambda}{\sqrt{d}}\right). \tag{2.94}$$

Again by the product rule and (2.93),

$$\left|D_{ij}^{kl}M_p - \frac{1}{\sqrt{d-1}}\partial_{ij}^{kl}M_p\right| \lesssim \frac{\Lambda^3}{d}(1+R_N)^{p-2}, \quad |D_{ij}^{kl}M_p| \lesssim \frac{\Lambda}{\sqrt{d}}(1+R_N)^{p-2}. \tag{2.95}$$

Substituting (2.95) into (2.94) gives

$$|\mathcal{E}_{ij}^{kl}| \lesssim |G_{ji}(z)|\frac{\Lambda^3}{d}(1+R_N)^{p-2} + \frac{\Lambda^3}{d}(1+R_N)^{p-1}. \tag{2.96}$$

We now sum this estimate. By the Ward identity (2.62),

$$\sum_{i,j}|G_{ij}(z)| \leqslant N\left(\sum_{i,j}|G_{ij}(z)|^2\right)^{1/2} = N\left(\frac{N\,\mathrm{Im}[m_N(z)]}{\eta}\right)^{1/2} \leqslant N\left(\frac{N|m_N(z)|}{\eta}\right)^{1/2}.$$

Since

$$\sum_{k,l} A_{ik}A_{jl} = d^2,$$

the first term in (2.96) gives

$$\frac{1}{N^2 d\sqrt{d-1}}\sum_{i,j,k,l} A_{ik}A_{jl}|G_{ji}(z)|\frac{\Lambda^3}{d}(1+R_N)^{p-2} \lesssim \frac{\Lambda^3}{\sqrt{d}}\frac{1}{\sqrt{N\eta}}(1+R_N)^{p-1}.$$

The second term in (2.96) gives

$$\frac{1}{N^2 d\sqrt{d-1}}\sum_{i,j,k,l} A_{ik}A_{jl}\frac{\Lambda^3}{d}(1+R_N)^{p-1} \lesssim \frac{\Lambda^3}{\sqrt{d}}(1+R_N)^{p-1}.$$

Therefore

$$\frac{\mathbf{1}(\Omega_N)}{N^2 d\sqrt{d-1}} \sum_{i,j,k,l} A_{ik}A_{jl}|\mathcal{E}_{ij}^{kl}| \lesssim \frac{\Lambda^3}{\sqrt{d}} \left( \frac{1}{\sqrt{N\eta}}(1+R_N)^{p-1} + (1+R_N)^{p-1} \right) \lesssim \frac{C_K \Lambda^3}{\sqrt{d}}(1+R_N)^{p-1}. \quad (2.97)$$

*Step 2: The weighted sum with $\partial_{ij}^{kl} G_{ji}(z)$.*

We will use the following identities from (2.61) repeatedly in this proof

$$AG(z) = \sqrt{d-1}\big(\mathbb{I} + zG(z)\big), \quad \sum_i G_{ij}(z) = \sum_j G_{ij}(z) = \frac{1}{d/\sqrt{d-1} - z}. \quad (2.98)$$

We also recall from (2.70)

$$\begin{aligned} \partial_{ij}^{kl} G_{ji}(z) = &-G_{ji}(z)^2 - G_{jj}(z)G_{ii}(z) - G_{jk}(z)G_{li}(z) - G_{jl}(z)G_{ki}(z) \\ &+ G_{ji}(z)G_{ki}(z) + G_{jl}(z)G_{ji}(z) + G_{jk}(z)G_{ii}(z) + G_{jj}(z)G_{li}(z). \end{aligned}$$

We compute the weighted sum term by term.

The first two terms are

$$\begin{aligned} -\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}^2 &= -d^2 \operatorname{Tr} G^2 = -d^2 N\, \partial_z m_N, \\ -\sum_{i,j,k,l} A_{ik}A_{jl}G_{jj}G_{ii} &= -d^2 (\operatorname{Tr} G)^2 = -d^2 N^2 m_N^2. \end{aligned}$$

For the third term, we use (2.98)

$$\begin{aligned} -\sum_{i,j,k,l} A_{ik}A_{jl}G_{jk}G_{li} &= -\sum_{i,j}\left(\sum_k A_{ik}G_{jk}\right)\left(\sum_l A_{jl}G_{li}\right) = -\sum_{i,j}(AG)_{ij}(AG)_{ji} \\ &= -(d-1)\left(N + 2z \operatorname{Tr} G + z^2 \operatorname{Tr} G^2\right) = -(d-1)N\left(1 + 2zm_N + z^2 \partial_z m_N\right). \end{aligned}$$

For the fourth term,

$$-\sum_{i,j,k,l} A_{ik}A_{jl}G_{jl}G_{ki} = -\left(\operatorname{Tr} AG\right)^2 = -(d-1)N^2(1+zm_N)^2.$$

For the fifth and sixth terms, using the row-sum identity (2.98),

$$\begin{aligned} \sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}G_{ki} &= d\sqrt{d-1}\,N \frac{1+zm_N}{d/\sqrt{d-1} - z}, \\ \sum_{i,j,k,l} A_{ik}A_{jl}G_{jl}G_{ji} &= d\sqrt{d-1}\,N \frac{1+zm_N}{d/\sqrt{d-1} - z}. \end{aligned}$$

For the seventh and eighth terms,

$$\begin{aligned} \sum_{i,j,k,l} A_{ik}A_{jl}G_{jk}G_{ii} &= d^2 N \frac{m_N}{d/\sqrt{d-1} - z}, \\ \sum_{i,j,k,l} A_{ik}A_{jl}G_{jj}G_{li} &= d^2 N \frac{m_N}{d/\sqrt{d-1} - z}. \end{aligned}$$

Adding the eight contributions, we obtain

$$\begin{aligned}\sum_{i,j,k,l} A_{ik}A_{jl}\partial_{ij}^{kl}G_{ji}(z) &= -d^2N^2m_N^2 - d^2N\partial_z m_N - (d-1)N\left(1+2zm_N+z^2\partial_z m_N\right)\\ &\quad -(d-1)N^2(1+zm_N)^2 + 2d\sqrt{d-1}\,N\frac{1+zm_N}{d/\sqrt{d-1}-z} + 2d^2N\frac{m_N}{d/\sqrt{d-1}-z}.\end{aligned}$$

Dividing by $N^2d(d-1)$, and using $|\partial_z m_N(z)| \leqslant |m_N(z)|/\mathrm{Im}[z]$, we get

$$\frac{1}{N^2d(d-1)}\sum_{i,j,k,l} A_{ik}A_{jl}\partial_{ij}^{kl}G_{ji}(z) = -m_N(z)^2 - \frac{1}{N}\partial_z m_N(z) + \mathrm{O}\left(\frac{1+|m_N(z)|^2+\frac{|m_N(z)|}{N\,\mathrm{Im}[z]}}{d}\right).$$

*Step 3: The weighted sum with $\partial_{ij}^{kl}m_N(w)$.*

We recall from (2.71)

$$\partial_{ij}^{kl}m_N(w) = -\frac{2}{N}\left[(G(w)^2)_{ij} + (G(w)^2)_{kl} - (G(w)^2)_{ik} - (G(w)^2)_{jl}\right]. \tag{2.99}$$

Substituting (2.99) into (2.88), we get

$$\begin{aligned}&\frac{1}{N^2d(d-1)}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\,\partial_{ij}^{kl}m_N(w)\\ &= -\frac{2}{d(d-1)N^3}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\left[(G(w)^2)_{ij} + (G(w)^2)_{kl} - (G(w)^2)_{ik} - (G(w)^2)_{jl}\right].\end{aligned} \tag{2.100}$$

We compute the four sums separately. First,

$$\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)(G(w)^2)_{ij} = d^2\,\mathrm{Tr}\big(G(z)G(w)^2\big).$$

Second,

$$\begin{aligned}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)(G(w)^2)_{kl} &= \mathrm{Tr}\big(G(z)AG(w)^2A\big)\\ &= (d-1)\left[\mathrm{Tr}\,G(z) + 2w\,\mathrm{Tr}(G(z)G(w)) + w^2\,\mathrm{Tr}(G(z)G(w)^2)\right].\end{aligned}$$

Third,

$$\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)(G(w)^2)_{ik} = d\sqrt{d-1}\,\frac{\mathrm{Tr}\,G(w) + w\,\mathrm{Tr}\,G(w)^2}{d/\sqrt{d-1}-z}.$$

The fourth term is the same:

$$\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)(G(w)^2)_{jl} = d\sqrt{d-1}\,\frac{\mathrm{Tr}\,G(w) + w\,\mathrm{Tr}\,G(w)^2}{d/\sqrt{d-1}-z}.$$

Therefore

$$\begin{aligned}\frac{1}{N^2d(d-1)}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\,\partial_{ij}^{kl}m_N(w) &= -\frac{2}{d(d-1)N^3}\Bigg[d^2\,\mathrm{Tr}\big(G(z)G(w)^2\big)\\ + (d-1)\big(\mathrm{Tr}\,G(z) + 2w\,\mathrm{Tr}(G(z)G(w)) + w^2\,\mathrm{Tr}(G(z)G(w)^2)\big) &- 2d\sqrt{d-1}\frac{\mathrm{Tr}\,G(w) + w\,\mathrm{Tr}\,G(w)^2}{d/\sqrt{d-1}-z}\Bigg].\end{aligned} \tag{2.101}$$

Finally, by the resolvent identity,

$$\frac{1}{N}\operatorname{Tr} G(w)^2 = \partial_w m_N(w), \quad \frac{1}{N}\operatorname{Tr}(G(z)G(w)) = \frac{m_N(z)-m_N(w)}{z-w},$$

and differentiating in $w$ gives

$$\frac{1}{N}\operatorname{Tr}(G(z)G(w)^2) = \partial_w \frac{m_N(z)-m_N(w)}{z-w}.$$

Substituting these identities into (2.101), we obtain

$$\frac{1}{N^2 d(d-1)}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\,\partial_{ij}^{kl} m_N(w) = -\frac{2d}{N^2(d-1)}\partial_w \frac{m_N(z)-m_N(w)}{z-w} \tag{2.102}$$

$$-\frac{2}{dN^2}\left[m_N(z) + 2w\frac{m_N(z)-m_N(w)}{z-w} + w^2\partial_w \frac{m_N(z)-m_N(w)}{z-w}\right] + \frac{4(m_N(w)+w\partial_w m_N(w))}{\sqrt{d-1}N^2(d/\sqrt{d-1}-z)}. \tag{2.103}$$

We recall from (2.2) and (2.4)

$$|\partial_w m_N(w)| \lesssim \frac{|m_N(w)|}{\operatorname{Im}[w]}, \quad \left|\frac{m_N(z)-m_N(w)}{z-w}\right| \lesssim \frac{|m_N(z)|}{\operatorname{Im}[z]} + \frac{|m_N(w)|}{\operatorname{Im}[w]},$$
$$\left|\partial_w \frac{m_N(z)-m_N(w)}{z-w}\right| \lesssim \frac{|m_N(z)|+|m_N(w)|}{\operatorname{Im}[z]\operatorname{Im}[w]},$$

the last line of (2.103) are bounded by

$$\mathrm{O}\left(\frac{|m_N(z)|+|m_N(w)|}{N^2 d\operatorname{Im}[z]\operatorname{Im}[w]}\right).$$

Thus

$$\frac{1}{N^2 d(d-1)}\sum_{i,j,k,l} A_{ik}A_{jl}G_{ji}(z)\,\partial_{ij}^{kl} m_N(w) = -\frac{2d}{N^2(d-1)}\partial_w \frac{m_N(z)-m_N(w)}{z-w} + \mathrm{O}\left(\frac{|m_N(z)|+|m_N(w)|}{N^2 d\operatorname{Im}[z]\operatorname{Im}[w]}\right).$$

This completes the proof of Lemma 2.16. □

**2.4.6 Edge loop equations for random $d$-regular graphs.** In this section, we explain how the proof of edge universality for random $d$-regular graphs with fixed degree in [52] can be significantly simplified using the new characterization. In that work, edge universality is proved by first establishing the edge loop equations for the Gaussian divisible ensemble $H + \sqrt{t}\,\mathrm{GOE}$, where $H$ is the normalized adjacency matrix of a random $d$-regular graph and $t > 0$ is small. The rest of the argument then follows the standard three-step scheme for random matrix universality [40]: a local law, universality for a short Gaussian divisible flow, and a comparison step back to the original ensemble.

For random $d$-regular graphs, especially when $d$ is fixed or small, the usual Green function comparison methods for Wigner-type matrices do not apply directly. The microscopic loop equations provide an alternative comparison mechanism: they show that the edge statistics are unchanged along the Gaussian divisible flow $H + \sqrt{t}\,\mathrm{GOE}$, and hence allow one to transfer edge universality from the Gaussian divisible ensemble back to $H$.

Our new characterization, Theorem 1.12, simplifies this argument. Instead of deriving and analyzing the edge loop equations for the entire Gaussian divisible family $H + \sqrt{t}\,\mathrm{GOE}$, it is enough to verify the approximate edge loop equations directly for the original ensemble $H$. In other words, the characterization turns the comparison problem into a direct criterion: once the edge loop equations for $H$ are established with sufficiently small errors, edge universality follows.

We recall the Kesten–McKay distribution

$$\varrho_d(x) := \mathbf{1}_{x\in[-2,2]} \left(1+\frac{1}{d-1}-\frac{x^2}{d}\right)^{-1} \frac{\sqrt{4-x^2}}{2\pi}. \tag{2.104}$$

Near the spectral edges $\pm 2$, the Kesten–McKay distribution has square-root behavior:

$$x \to \pm 2, \qquad \varrho_d(x) = \frac{\mathcal{A}\sqrt{2\mp x}}{\pi} + \mathrm{O}(|2\mp x|), \qquad \mathcal{A} := \frac{d(d-1)}{(d-2)^2}. \tag{2.105}$$

We denote by $m_d(z)$ the Stieltjes transform of $\varrho_d$:

$$m_d(z) = \int_{\mathbb{R}} \frac{\varrho_d(x)\,\mathrm{d}x}{x-z}, \qquad z \in \mathbb{C}^+.$$

**Definition 2.17.** *Fix $d \geqslant 3$ and a sufficiently small constant $0 < \mathfrak{c} < 1$. Let $\mathfrak{R} := (\mathfrak{c}/4)\log_{d-1} N$. We define $\Omega_N$ to be the event on which the following two conditions hold:*

1. *The number of vertices whose radius-$\mathfrak{R}$ neighborhood is not a tree is at most $N^{\mathfrak{c}}$.*
2. *The radius-$\mathfrak{R}$ neighborhood of every vertex has excess at most one. Here the excess means the number of independent cycles.*

The event $\Omega_N$ is typical. More precisely, by [53, Proposition 2.1], applied with $\omega_d = 1$, we have

$$\mathbb{P}(\Omega_N) \geqslant 1 - \mathrm{O}\big(N^{-(1-\mathfrak{c})}\big).$$

For random $d$-regular graphs, the required approximate loop equations are given in the following proposition. These equations can be viewed as the direct analogue, for the original random regular graph ensemble, of the edge loop equations proved in [52, Theorem 3.8] for the Gaussian divisible family $H + \sqrt{t}\,\mathrm{GOE}$. For an expository discussion of the first loop equation, corresponding to $p = 1$, we refer to [54].

**Proposition 2.18.** *Fix a degree $d \geqslant 3$, and set the right spectral edge $E = 2$. Fix a small $\mathfrak{c} > 0$. Introduce the microscopic window*

$$\mathbf{M} := \left\{ w \in \mathbb{C} : N^{-2/3-\mathfrak{c}} \leqslant |\operatorname{Im}[w]| \leqslant N^{-2/3+\mathfrak{c}}, \quad -N^{-2/3+\mathfrak{c}} \leqslant \operatorname{Re}[w] \leqslant N^{-2/3+\mathfrak{c}} \right\}. \tag{2.106}$$

*Let $\Delta(z) := m_N(z) - m_d(z)$. Fix $p \geqslant 1$ and $w, w_2, \dots, w_p \in \mathbf{M}$, and set*

$$z := E + w, \qquad z_a := E + w_a, \qquad 2 \leqslant a \leqslant p.$$

*Then, for $N$ sufficiently large,*

$$\begin{aligned} &\mathbb{E}\left[\mathbf{1}(\Omega_N)\left(\Delta(z)^2 + 2\mathcal{A}\sqrt{z-E}\,\Delta(z) + \frac{1}{N}\partial_z m_N(z)\right)\prod_{a=2}^{p}\Delta(z_a)\right] \\ &\quad + \frac{2}{N^2}\sum_{a=2}^{p}\mathbb{E}\left[\mathbf{1}(\Omega_N)\partial_{z_a}\left(\frac{m_N(z)-m_N(z_a)}{z-z_a}\right)\prod_{\substack{b=2\\ b\neq a}}^{p}\Delta(z_b)\right] = \mathrm{O}\left(N^{-(p+1)/3-10\mathfrak{c}}\right). \end{aligned} \tag{2.107}$$

*Here, when $p = 1$, the empty product is interpreted as $1$, and the sum over $a = 2, \dots, p$ is absent.*

The equations (2.107) can be viewed as a centered version of the edge loop equations (1.5). Namely, under the edge scaling,

$$\frac{N^{1/3}}{\mathcal{A}^{2/3}}\,\Delta\left(E + \frac{w}{(\mathcal{A}N)^{2/3}}\right)$$

converges to $s(w) - \sqrt{w}$ in the edge loop equations (1.5). The full set of edge loop equations then follows by taking suitable linear combinations.

## 3 Bulk and edge loop equations

The purpose of this section is twofold. First, we identify the loop equations satisfied by the limiting point processes in the bulk and edge regimes. Namely, the $\mathrm{Sine}_\beta$ point process satisfies the bulk loop equation, while the $\mathrm{Airy}_\beta$ point process satisfies the edge loop equation. Second, we show that these loop equations yield concentration estimates for the corresponding normalized Stieltjes transforms. More precisely, after the appropriate normalization, these Stieltjes transforms concentrate around the deterministic quantities $\pm\mathrm{i}$ in the bulk and $\sqrt{z}$ at the edge, with sub-Gaussian tails. All results in this section hold for every $\beta > 0$.

### 3.1 $\mathrm{Sine}_\beta$ and $\mathrm{Airy}_\beta$ point process

For the $\mathrm{Sine}_\beta$ point process, we use the normalization in which its intensity is $1/\pi$, or equivalently, the local mean spacing is $\pi$. For the $\mathrm{Airy}_\beta$ point process, we use the convention in which the Airy points are ordered decreasingly and extend to $-\infty$. In this normalization, the one-point density has the standard square-root left-tail asymptotics with leading constant $1/\pi$. The following theorems state that the $\mathrm{Sine}_\beta$ point process satisfies the bulk loop equation, and the $\mathrm{Airy}_\beta$ point process satisfies the edge loop equation.

**Theorem 3.1.** *Fix $\beta > 0$, and let $(x_j)_{j\in\mathbb{Z}}$ denote the $\mathrm{Sine}_\beta$ point process with intensity $1/\pi$, indexed so that*

$$\cdots \leqslant x_{-2} \leqslant x_{-1} \leqslant x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots . \tag{3.1}$$

*Then its principal-value Stieltjes transform*

$$s(w) = \mathrm{P.V.} \sum_{j\in\mathbb{Z}} \frac{1}{x_j - w} \tag{3.2}$$

*satisfies the bulk loop equation (1.2).*

**Theorem 3.2.** *Fix $\beta > 0$, and let $(x_j)_{j\geqslant 1}$ denote the $\mathrm{Airy}_\beta$ point process, ordered decreasingly so that*

$$x_1 \geqslant x_2 \geqslant x_3 \geqslant \cdots, \qquad x_j \to -\infty \quad \text{as } j \to \infty. \tag{3.3}$$

*Then its normalized Stieltjes transform*

$$s(w) = \sum_{j=1}^{\infty} \left( \frac{1}{x_j - w} - \frac{1}{\mathfrak{a}_j} \right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}. \tag{3.4}$$

*satisfies the edge loop equation (1.5).*

Both Theorem 3.1 and Theorem 3.2 are proved by taking the bulk and edge scaling limits, respectively, of the Gaussian $\beta$-ensemble. Their proofs are given in Section 3.3.

### 3.2 Local Law

In this section, we collect some consequences for point processes satisfying the bulk/edge loop equations. Proposition 3.3 states that if a Nevanlinna function $s$ associated to a point process satisfies the bulk loop equation (1.2), then $s(w)$ has a sub-Gaussian tail and is the Stieltjes transform of its empirical density. Proposition 3.4 states that if a point process satisfies the edge loop equation (1.5), then $s(w)$ has a sub-Gaussian tail and is the *normalized* Stieltjes transform of its empirical density.

The proofs follow a method introduced in [19], which was recently used in [20] to compute the excess of the $\mathrm{Sine}_\beta$ point process. The local law below is formulated for general Nevanlinna functions, whereas [20] works with the principal-value Stieltjes transform, and obtains concentration estimates with explicit dependence on the inverse temperature $\beta$.

**Proposition 3.3.** *Under Assumption 1.4, for each fixed $\beta > 0$, there exists a constant $C = C(\beta) > 0$ such that, for every $w = E + \mathrm{i}\eta \in \mathbb{C}_+$ with $\eta > 0$ and every integer $q \geqslant 1$,*

$$\mathbb{E}\big[|s(w) - \mathrm{i}|^{2q}\big] \leqslant \frac{(Cq)^q}{\eta^{2q}}. \tag{3.5}$$

*Moreover, the poles of $s(w)$ can be ordered as*

$$\cdots \leqslant x_{-2} \leqslant x_{-1} \leqslant x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots .$$

*There exists a random variable $M > 0$ such that, almost surely,*

$$|x_j - \pi j| \leqslant M\big(\log(2 + |j|)\big)^2. \tag{3.6}$$

*Furthermore, for any $w \in \mathbb{C}_+ \cup \mathbb{C}_-$, almost surely,*

$$s(w) = \mathrm{P.V.} \sum_{j=-\infty}^{\infty} \frac{1}{x_j - w}. \tag{3.7}$$

We next recall the Airy function, denoted by Ai, which solves the Airy equation: $\mathrm{Ai}''(w) - w\,\mathrm{Ai}(w) = 0$, with $\mathrm{Ai}(w) \to 0$ as $w \to \infty$ along $\mathbb{R}_+$. All the zeros of Ai are on the real line, and are all negative, and we denote them as $0 > \mathfrak{a}_1 > \mathfrak{a}_2 > \mathfrak{a}_3 > \cdots$. We refer to Appendix A for further background on Airy functions.

**Proposition 3.4.** *Under Assumption 1.10, for each fixed $\beta > 0$, there exists a constant $C = C(\beta) > 0$ such that the following holds. Let $w = E + \mathrm{i}\eta \in \mathbb{C}_+$, with $\eta > 0$, and let $q \geqslant 1$ be an integer, then*

$$\mathbb{E}\left[\big|s(w) - \sqrt{w}\big|^{2q}\right] \leqslant \frac{(Cq)^q}{\eta^{2q}} \frac{|w|^q}{(\mathrm{Im}[\sqrt{w}])^{2q}}. \tag{3.8}$$

*If we further assume that $E \geqslant \eta$, then*

$$\mathbb{E}\left[\big|\mathrm{Im}\left[s(w) - \sqrt{w}\right]\big|^{2q}\right] \leqslant \frac{(Cq)^q}{\eta^q E^q} + \frac{(Cq)^{2q}}{\eta^{4q} E^q}. \tag{3.9}$$

*Moreover, the poles of $s(w)$ can be ordered as*

$$x_1 \geqslant x_2 \geqslant \cdots .$$

*There exists a random variable $M > 0$ such that, almost surely,*

$$\left| x_j + \left(\frac{3\pi j}{2}\right)^{2/3} \right| \leqslant M \frac{\big(\log(2 + j)\big)^2}{j^{1/3}}. \tag{3.10}$$

*Furthermore, for any $w \in \mathbb{C}_+ \cup \mathbb{C}_-$, almost surely,*

$$s(w) = \sum_{i=1}^{\infty} \left(\frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i}\right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}. \tag{3.11}$$

Section 3.4 is devoted to the proof of Proposition 3.3. The proof of Proposition 3.4 is analogous and is postponed to Appendix C.

## 3.3 Bulk and edge loop equations for $\beta$-ensemble

In this section, we show that, under the one-cut condition, the $\beta$-ensemble with a general potential satisfies the approximate bulk and edge loop equations, namely (1.4) and (1.6). Both results follow from straightforward computations, with the local laws for $\beta$-ensembles up to microscopic scale [19] as an input. We defer the proofs to Appendix B. We then prove Theorem 3.1 and Theorem 3.2 by taking the bulk and edge scaling limits of these approximate loop equations.

Consider the $\beta$-ensemble

$$\mathrm{d}\mu_N(\lambda_1,\dots,\lambda_N)=\frac{1}{Z_N}\prod_{1\leqslant i<j\leqslant N}|\lambda_i-\lambda_j|^\beta e^{-\frac{\beta N}{2}\sum_{k=1}^N V(\lambda_k)}\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_N. \tag{3.12}$$

**Assumption 3.5.** *We assume the following.*

*(A1) The potential $V$ is analytic on $\mathbb{C}$ and real-valued on $\mathbb{R}$.*

*(A2) There exist constants $M_0,C,c>0$ such that*

$$V'(x)\geqslant c,\qquad \sup_{y\in[M_0,x]}\frac{V'(y)}{y}\leqslant CV(x),\qquad x\geqslant M_0,$$

*and the analogous estimates hold as $x\to-\infty$, equivalently for $\widetilde{V}(x):=V(-x)$.*

*(A3) The equilibrium measure $\mu_V$ is one-cut regular. More precisely, its density $\varrho_V$ is positive on a single interval $[A,B]$ and has square-root behavior at both endpoints.*

*(A4) The effective potential*

$$x\longmapsto\frac{V(x)}{2}-\int\log|x-t|\,\mathrm{d}\mu_V(t)$$

*achieves its minimum only on $[A,B]$.*

Write

$$\varrho_V(t)=\frac{1}{\pi}r(t)\sqrt{(t-A)(B-t)},\qquad t\in[A,B], \tag{3.13}$$

and denote by $m_V$ the Stieltjes transform of $\mu_V$,

$$m_V(z):=\int_A^B\frac{\mathrm{d}\mu_V(t)}{t-z}.$$

We also write

$$m_B:=m_V(B), \tag{3.14}$$

where $m_V(B)$ denotes the continuous extension of $m_V$ to the right edge.

**Proposition 3.6** (Approximate bulk loop equations for $\beta$-ensembles)**.** *Assume that the potential $V$ satisfies Assumption 3.5. Fix a bulk energy $E\in(A,B)$, and write $\varrho_E:=\varrho_V(E)$. For $w\in\mathbb{C}\setminus\mathbb{R}$, define*

$$s_N(w)=\sum_{k=1}^N\frac{1}{N\pi\varrho_E(\lambda_k-E)-w}+\frac{V'(E)}{2\pi\varrho_E}. \tag{3.15}$$

*Thus $s_N$ is a particle-generated Nevanlinna function associated with the configuration*

$$\sum_{k=1}^N\delta_{N\pi\varrho_E(\lambda_k-E)}.$$

*Then, for every $p \geqslant 1$ and every compact set $K \subset \mathbb{C} \setminus \mathbb{R}$, uniformly for $w_1, w_2 \dots, w_p \in K$,*

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\partial_{w_1} s_N(w_1) + s_N(w_1)^2 + 1\right)\prod_{j=2}^{p} s_N(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w_1)-s_N(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s_N(w_\ell)\right] = \mathrm{o}_N(1). \tag{3.16}$$

**Proposition 3.7** (Approximate edge loop equations for $\beta$-ensembles)**.** *Assume that $V$ satisfies Assumption 3.5. Let*

$$\gamma := \big(r(B)^2(B-A)\big)^{-1/3}. \tag{3.17}$$

*For $w \in \mathbb{C} \setminus \mathbb{R}$, define the edge-scaled empirical Stieltjes transform*

$$s_N(w) = \sum_{k=1}^{N}\frac{1}{\gamma^{-1}N^{2/3}(\lambda_k - B) - w} - \gamma N^{1/3} m_B. \tag{3.18}$$

*Thus $s_N$ is a particle-generated Nevanlinna function associated with the configuration*

$$\sum_{k=1}^{N}\delta_{\gamma^{-1}N^{2/3}(\lambda_k - B)}.$$

*Then, for every $p \geqslant 1$ and every compact set $K \subset \mathbb{C} \setminus \mathbb{R}$, uniformly for $w_1, \dots, w_p \in K$,*

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\partial_{w_1} s_N(w_1) + s_N(w_1)^2 - w_1\right)\prod_{j=2}^{p} s_N(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w_1)-s_N(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p} s_N(w_\ell)\right] = \mathrm{o}_N(1). \tag{3.19}$$

**Remark 3.8.** *Propositions 3.6 and 3.7 together with Theorems 1.9 and 1.12 imply universality for $\beta$-ensembles, in the bulk and at the edge. Previous proofs proceed by comparison with the Gaussian $\beta$-ensemble [11, 16, 17, 63, 86].*

*Proof of Theorem 3.1.* We apply Proposition 3.6 to the Gaussian $\beta$-ensemble, that is, to $V(x) = x^2$, at the bulk point $E = 0$. Then $V'(0) = 0$. Writing $\varrho_0 := \varrho_V(0)$, the renormalized Stieltjes transform in (3.15) becomes

$$s_N(w) = \sum_{k=1}^{N}\frac{1}{N\pi\varrho_0\lambda_k - w}.$$

The corresponding microscopic point process

$$\mu_N := \sum_{k=1}^{N}\delta_{N\pi\varrho_0\lambda_k}$$

converges in distribution to the $\mathrm{Sine}_\beta$ point process with intensity $1/\pi$, as constructed in [90].

By the same tightness and uniform-integrability argument as in Theorem 1.9, after passing to a subsequence we may assume that $s_N \Longrightarrow s$ in the topology of locally uniform convergence on compact subsets of $\mathbb{C} \setminus \mathbb{R}$, and the limit $s$ satisfies the exact bulk loop equation (1.2). Moreover, the associated limiting particle configuration $\mu$ has the $\mathrm{Sine}_\beta$ law.

Ordering the atoms of $\mu$ as

$$\cdots \leqslant x_{-2} \leqslant x_{-1} \leqslant x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots,$$

the representation (3.7) gives

$$s(w) = \text{P.V.} \int \frac{\mathrm{d}\mu(x)}{x-w} = \text{P.V.} \sum_{j=-\infty}^{\infty} \frac{1}{x_j - w}.$$

Therefore the Stieltjes transform of the $\mathrm{Sine}_\beta$ point process satisfies the bulk loop equation (1.2). ☐

*Proof of Theorem 3.2.* The result follows by the same argument as in the proof of Theorem 3.1, now applying Proposition 3.7 to the Gaussian $\beta$-ensemble and using the fact that its edge scaling limit is the $\mathrm{Airy}_\beta$ point process [78]. We therefore omit the details. ☐

## 3.4 Proof of bulk local law

**3.4.1 Proof of Equation (3.5).** The high-moment estimate (3.5) follows essentially from the analysis of loop equations for $\beta$-ensembles in [19, Theorem 1.1]. The proof is based on a moment bootstrap from the bulk loop equation. We first derive deterministic bounds on $s'(w)$ and on the difference quotient $\partial_w((s(z)-s(w))/(z-w))$ in terms of $\mathrm{Im}[s(w)]$. Applying the loop equation to suitable polynomial test functions then yields a self-consistent estimate for the high moments of $Q(w) := s(w)^2+1$. Since $Q(w) = (s(w)-\mathrm{i})(s(w)+\mathrm{i})$ and $s$ is a Nevanlinna function, moments of $Q(w)$ control moments of $s(w)-\mathrm{i}$. A Young inequality argument closes the estimate and gives the desired sub-Gaussian-type bound for $s(w)-\mathrm{i}$.

*Proof of* (3.5). Let $w = E + \mathrm{i}\eta$, with $\eta > 0$. Define

$$Q(w) := s(w)^2 + 1, \qquad f(z,w) := \partial_w \left( \frac{s(z)-s(w)}{z-w} \right), \qquad f(w,w) := \frac{1}{2} s''(w).$$

For a monomial $F(u,v) = u^a v^b$, the loop equation (1.2), applied with $w_1 = w$, $a$ copies of $w$ and $b$ copies of $\bar{w}$, together with $s(\bar{w}) = \overline{s(w)}$, gives

$$\begin{aligned}\mathbb{E}[Q(w)\, F(s(w), s(\bar{w}))] = &\left(1 - \frac{2}{\beta}\right) \mathbb{E}[s'(w)\, F(s(w), s(\bar{w}))] \\ &- \frac{2}{\beta}\, \mathbb{E}[f(w,w)\, \partial_1 F(s(w), s(\bar{w})) + f(w,\bar{w})\, \partial_2 F(s(w), s(\bar{w}))] \,.\end{aligned}$$

By linearity, the same identity holds for every polynomial $F$.

Now take

$$F(u,v) = (u^2+1)^{q-1}(v^2+1)^q.$$

Then

$$\partial_1 F(u,v) = 2(q-1)u(u^2+1)^{q-2}(v^2+1)^q, \qquad \partial_2 F(u,v) = 2qv(u^2+1)^{q-1}(v^2+1)^{q-1}.$$

Therefore,

$$\mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant C\, \mathbb{E}\big[|s'(w)|\, |Q(w)|^{2q-1}\big] + Cq\, \mathbb{E}\big[\big(|f(w,w)| + |f(w,\bar{w})|\big)|s(w)|\, |Q(w)|^{2q-2}\big] \,. \tag{3.20}$$

Substituting the bounds (2.2) and (2.4) into (3.20) yields

$$\mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant C\, \mathbb{E}\left[\frac{\mathrm{Im}[s(w)]}{\eta}\, |Q(w)|^{2q-1}\right] + Cq\, \mathbb{E}\left[\frac{\mathrm{Im}[s(w)]\, |s(w)|}{\eta^2}\, |Q(w)|^{2q-2}\right].$$

We now apply Young's inequality in the forms

$$Cab^{2q-1} \leqslant \frac{1}{4} b^{2q} + (C'a)^{2q}, \qquad Cqab^{2q-2} \leqslant \frac{1}{4} b^{2q} + (C'qa)^q. \tag{3.21}$$

After absorbing the resulting $\frac{1}{2}\mathbb{E}[|Q(w)|^{2q}]$ contribution into the left-hand side, we obtain

$$\mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q \left(\mathbb{E}\big[(\operatorname{Im}[s(w)])^{2q}\big] + \mathbb{E}\big[(\operatorname{Im}[s(w)])^q |s(w)|^q\big]\right).$$

In particular, since $|\operatorname{Im}[s(w)]| \leqslant |s(w)|$, we have

$$\mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q \mathbb{E}\big[(\operatorname{Im}[s(w)])^q |s(w)|^q\big]. \tag{3.22}$$

Since $s$ is a Nevanlinna function, $\operatorname{Im}[s(w)] \geqslant 0$. Moreover,

$$Q(w) = (s(w) - \mathrm{i})(s(w) + \mathrm{i}),$$

and for $\operatorname{Im}[s(w)] \geqslant 0$ we have

$$|s(w) - \mathrm{i}| \leqslant |s(w) + \mathrm{i}|, \qquad |s(w) + \mathrm{i}| \geqslant 1.$$

Hence

$$|s(w) - \mathrm{i}|^2 \leqslant |Q(w)|, \qquad |s(w) - \mathrm{i}| \leqslant |Q(w)|. \tag{3.23}$$

As a consequence of (3.23), we have

$$\operatorname{Im}[s(w)]^q |s(w)|^q \leqslant |s(w)|^{2q} \leqslant (1 + |s(w) - \mathrm{i}|)^{2q} \leqslant C^q (1 + |Q(w)|^q).$$

By plugging the above estimate into (3.22), we conclude

$$\mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q \left(1 + \mathbb{E}\big[|Q(w)|^{2q}\big]^{1/2}\right). \tag{3.24}$$

By (3.23),

$$\mathbb{E}\big[|s(w) - \mathrm{i}|^{2q}\big] \leqslant \mathbb{E}\big[|Q(w)|^q\big] \leqslant \mathbb{E}\big[|Q(w)|^{2q}\big]^{1/2}, \quad \mathbb{E}\big[|s(w) - \mathrm{i}|^{2q}\big] \leqslant \mathbb{E}\big[|Q(w)|^{2q}\big]. \tag{3.25}$$

There are two cases. If $\mathbb{E}[|Q(w)|^{2q}] \leqslant 1$, then (3.24) and (3.25) imply

$$\mathbb{E}\big[|s(w) - \mathrm{i}|^{2q}\big] \leqslant \mathbb{E}\big[|Q(w)|^{2q}\big] \leqslant 2\left(\frac{Cq}{\eta^2}\right)^q.$$

If $\mathbb{E}[|Q(w)|^{2q}] \geqslant 1$, then dividing (3.24) by $\mathbb{E}[|Q(w)|^{2q}]^{1/2}$ gives

$$\mathbb{E}\big[|s(w) - \mathrm{i}|^{2q}\big] \leqslant \mathbb{E}\big[|Q(w)|^{2q}\big]^{1/2} \leqslant 2\left(\frac{Cq}{\eta^2}\right)^q.$$

Hence (3.5) holds in all cases. □

The sub-Gaussian tail estimate (3.5), together with a union bound, implies the following almost-sure uniform bound.

**Lemma 3.9** (Almost-sure uniform bound on dyadic rectangles)**.** *Adopt the assumptions of Proposition 3.3. For $n \geqslant 1$, define*

$$\mathcal{R}_n := \Big\{w = E + \mathrm{i}\eta : \ |E| \leqslant 2^n, \ 2^{-n} \leqslant \eta \leqslant 2^n\Big\}, \qquad \mathcal{R}_n' := \Big\{w = E + \mathrm{i}\eta : \ |E| \leqslant 2^n, \ 0 < \eta \leqslant 2^n\Big\}.$$

*Then there exists an almost surely finite random variable $M$ such that, almost surely,*

$$\sup_{w = E + \mathrm{i}\eta \in \mathcal{R}_n} \eta \, |s(w) - \mathrm{i}| \leqslant Mn, \qquad n \geqslant 1, \tag{3.26}$$

*and*

$$\sup_{w = E + \mathrm{i}\eta \in \mathcal{R}_n'} \eta \, |\operatorname{Im}[s(w) - \mathrm{i}]| \leqslant Mn, \qquad n \geqslant 1. \tag{3.27}$$

*Proof.* By (3.5) and Markov's inequality,

$$\mathbb{P}(\eta|s(E+\mathrm{i}\eta)-\mathrm{i}|\geqslant t)\leqslant \inf_{q\geqslant 1}\frac{(Cq)^q}{t^{2q}}.$$

Optimizing in $q$ gives constants $c, C>0$ such that

$$\mathbb{P}(\eta|s(E+\mathrm{i}\eta)-\mathrm{i}|\geqslant t)\leqslant Ce^{-ct^2},\qquad E\in\mathbb{R},\ \eta>0,\ t\geqslant 1. \tag{3.28}$$

Fix $n\geqslant 1$, and define the lattice net

$$\mathcal{N}_n := \big\{2^{-n}(a+\mathrm{i}b):\ a\in\mathbb{Z},\ |a|\leqslant 2^{2n},\ b\in\mathbb{Z},\ 1\leqslant b\leqslant 2^{2n}\big\}.$$

Its cardinality satisfies

$$|\mathcal{N}_n|\leqslant C2^{4n}$$

By (3.28) and a union bound,

$$\mathbb{P}\Big(\exists\, w=E+\mathrm{i}\eta\in\mathcal{N}_n : \eta|s(w)-\mathrm{i}|\geqslant n\Big)\leqslant C2^{4n}e^{-cn^2}.$$

The right-hand side is summable in $n$. Hence, by Borel–Cantelli, almost surely there exists $n_0(\omega)$ such that, for all $n\geqslant n_0(\omega)$,

$$\eta|s(w)-\mathrm{i}|\leqslant n\qquad \text{for every } w=E+\mathrm{i}\eta\in\mathcal{N}_n. \tag{3.29}$$

We next pass from the net to the whole rectangle $\mathcal{R}_n$. From (2.2) and $\mathrm{Im}[s(w)]\leqslant|s(w)-\mathrm{i}|+1$, we have

$$|s'(w)|\leqslant\frac{|s(w)-\mathrm{i}|+1}{\mathrm{Im}[w]}. \tag{3.30}$$

Let $w_0=E_0+\mathrm{i}\eta_0\in\mathcal{N}_n$, and let $w=E+\mathrm{i}\eta\in\mathcal{R}_n$ satisfy

$$|\operatorname{Re}[w]-\operatorname{Re}[w_0]|\leqslant 2^{-n-1},\qquad |\operatorname{Im}[w]-\operatorname{Im}[w_0]|\leqslant 2^{-n-1}.$$

Since $\eta_0\geqslant 2^{-n}$, the straight-line path from $w_0$ to $w$ stays in the region $\mathrm{Im}[\zeta]\geqslant\eta_0/2$. Applying (3.30) along this path and using Gronwall's inequality gives

$$|s(w)-s(w_0)|\leqslant C\frac{2^{-n}}{\eta_0}\big(|s(w_0)-\mathrm{i}|+1\big). \tag{3.31}$$

Moreover, $\eta/\eta_0\leqslant C$. Therefore, using (3.29),

$$\eta|s(w)-\mathrm{i}|\leqslant C\eta_0|s(w_0)-\mathrm{i}|+C2^{-n}\big(|s(w_0)-\mathrm{i}|+1\big).$$

Since $2^{-n}\leqslant\eta_0$, $2^{-n}\leqslant 1\leqslant n$, and $\eta_0|s(w_0)-\mathrm{i}|\leqslant n$, we get

$$\eta|s(w)-\mathrm{i}|\leqslant Cn.$$

Thus, almost surely,

$$\sup_{w\in\mathcal{R}_n}\eta|s(w)-\mathrm{i}|\leqslant Cn,\qquad n\geqslant n_0(\omega).$$

Increasing the random constant to handle the finitely many indices $n<n_0(\omega)$, we obtain (3.26).

It remains to control the imaginary part for $0<\eta\leqslant 2^{-n}$. By (2.1), for each fixed $E\in\mathbb{R}$, the function

$$\eta\longmapsto\eta\ \mathrm{Im}[s(E+\mathrm{i}\eta)]=c\eta^2+\sum_{x\in P}\frac{\eta^2}{(x-E)^2+\eta^2}$$

is nondecreasing. Hence, with $\eta_0:=2^{-n}$,

$$\eta\ \mathrm{Im}[s(E+\mathrm{i}\eta)]\leqslant\eta_0\ \mathrm{Im}[s(E+\mathrm{i}\eta_0)].$$

Using (3.26) at $E+\mathrm{i}\eta_0$, we get

$$\eta\,|\operatorname{Im}[s(E+\mathrm{i}\eta)-\mathrm{i}]|\leqslant\eta\ \mathrm{Im}[s(E+\mathrm{i}\eta)]+\eta\leqslant\eta_0\ \mathrm{Im}[s(E+\mathrm{i}\eta_0)]+1\leqslant Mn+2.$$

After increasing $M$, this gives (3.27). $\square$

**3.4.2 Helffer–Sjöstrand formula.** In this section, we recall the Helffer–Sjöstrand formula, Lemma 3.11, which allows us to control the difference between two measures through their associated Nevanlinna functions. We will use Lemma 3.11 to derive particle concentration estimates in Lemma 3.12.

Let $f \in C_c^\infty(\mathbb{R};\mathbb{R})$, and let $\chi \in C_c^\infty(\mathbb{R};\mathbb{R})$ be an even cutoff function such that $\chi \equiv 1$ in a neighborhood of 0. Define the almost-analytic extension

$$\widetilde{f}(x+\mathrm{i}y) := (f(x)+\mathrm{i}yf'(x))\chi(y), \qquad w = x+\mathrm{i}y. \tag{3.32}$$

Then, for every $\lambda \in \mathbb{R}$,

$$f(\lambda) = \frac{1}{\pi}\int_{\mathbb{R}^2} \frac{\partial_{\bar{w}}\widetilde{f}(x+\mathrm{i}y)}{\lambda - x - \mathrm{i}y}\,\mathrm{d}x\,\mathrm{d}y, \qquad \partial_{\bar{w}} := \frac{1}{2}(\partial_x + \mathrm{i}\partial_y).$$

A direct computation gives

$$\partial_{\bar{w}}\widetilde{f}(x+\mathrm{i}y) = \frac{\mathrm{i}}{2}\,y\chi(y)f''(x) + \frac{\mathrm{i}}{2}\,\big(f(x)+\mathrm{i}yf'(x)\big)\chi'(y). \tag{3.33}$$

The scalar formula immediately implies the following standard consequence.

**Lemma 3.10.** *Let $s$ be a Nevanlinna function associated with the Borel measure $\mu$. Then, for every test function $f \in C_c^\infty(\mathbb{R};\mathbb{R})$,*

$$\int_{\mathbb{R}} f(\lambda)\,\mathrm{d}\mu(\lambda) = \frac{2}{\pi}\int_{x+\mathrm{i}y\in\mathbb{C}_+} \mathrm{Re}\Big[\partial_{\bar{w}}\widetilde{f}(x+\mathrm{i}y)\,s(x+\mathrm{i}y)\Big]\,\mathrm{d}x\,\mathrm{d}y.$$

**Lemma 3.11.** *Let $s$ and $s_\varrho$ be two Nevanlinna functions associated with the Borel measures*

$$\sum_{x\in P}\delta_x, \qquad \frac{1}{\pi}\varrho(x)\,\mathrm{d}x,$$

*respectively. Then, for any $\eta > 0$ such that $\chi(\eta) = 1$,*

$$\left|\sum_{x\in P} f(x) - \frac{1}{\pi}\int_{\mathbb{R}} f(x)\varrho(x)\,\mathrm{d}x\right| \leqslant C\big(\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}\big), \tag{3.34}$$

*where*

$$\begin{aligned}
\mathrm{I} &:= \iint_{y\geqslant 0} \big(|f(x)| + y|f'(x)|\big)\,|\chi'(y)|\,|s(x+\mathrm{i}y) - s_\varrho(x+\mathrm{i}y)|\,\mathrm{d}x\,\mathrm{d}y,\\
\mathrm{II} &:= \iint_{0\leqslant y\leqslant\eta} |f''(x)|\,y|\chi(y)|\,\big|\mathrm{Im}\big[s(x+\mathrm{i}y) - s_\varrho(x+\mathrm{i}y)\big]\big|\,\mathrm{d}x\,\mathrm{d}y,\\
\mathrm{III} &:= \iint_{y\geqslant\eta} |f'(x)|\,\big|\partial_y\big(y\chi(y)\big)\big|\,|s(x+\mathrm{i}y) - s_\varrho(x+\mathrm{i}y)|\,\mathrm{d}x\,\mathrm{d}y,\\
\mathrm{IV} &:= \eta\int_{\mathbb{R}} |f'(x)|\,|s(x+\mathrm{i}\eta) - s_\varrho(x+\mathrm{i}\eta)|\,\mathrm{d}x.
\end{aligned}$$

*Here $C$ depends only on the cutoff function $\chi$.*

These identities and estimates are classical; see, for example, [40, Section 11.2]. We omit the proofs.

**3.4.3 Proof of Equation (3.6).** The concentration of the particle locations follows from the following counting estimate. We prove it using the Helffer–Sjöstrand formula, Lemma 3.11, together with the uniform bound on the Nevanlinna function from Lemma 3.9.

**Lemma 3.12** (Bulk counting estimate)**.** *Adopt the assumptions of Proposition 3.3. Then there exists an almost surely finite random variable $M$ such that, almost surely, for every $a \geqslant 2$,*

$$\left|\#(P \cap [0,a]) - \frac{a}{\pi}\right| \leqslant M(\log a)^2, \qquad \left|\#(P \cap [-a,0]) - \frac{a}{\pi}\right| \leqslant M(\log a)^2. \tag{3.35}$$

*Proof of* (3.6). Fix an outcome $\omega$ on which Lemma 3.12 holds. Enumerate the poles of $s$ increasingly as

$$\cdots \leqslant x_{-1} \leqslant x_0 \leqslant x_1 \leqslant \cdots,$$

with only a finite ambiguity near the origin. Then there exists an almost surely finite random variable $M$ such that

$$|x_j - \pi j| \leqslant M\big(\log(2+|j|)\big)^2, \qquad j \in \mathbb{Z}. \tag{3.36}$$

Indeed, for $j \geqslant 1$, Lemma 3.12 gives

$$j = \#(P \cap (0,x_j]) + \mathrm{O}(1) = \frac{x_j}{\pi} + \mathrm{O}\big((\log x_j)^2\big),$$

and hence

$$|x_j - \pi j| \leqslant M\big(\log(2+j)\big)^2.$$

The argument for $j \leqslant -1$ is identical, using the estimate on $[-a,0]$. Enlarging $M$ handles the finitely many indices near the origin. □

*Proof of Lemma 3.12.* We prove the estimate on $[0,a]$; the estimate on $[-a,0]$ is identical.

Fix an outcome $\omega$ on which Lemma 3.9 holds, and let $M = M(\omega)$ be the corresponding random constant. Given $a \geqslant 2$, choose $n \geqslant 1$ such that

$$2^{n-1} \leqslant a+1 \leqslant 2^n.$$

Then $n \asymp \log a$.

Choose smooth functions $f_+, f_- \in C_c^\infty(\mathbb{R})$ such that

$$0 \leqslant f_+ \leqslant 1, \qquad f_+ \equiv 1 \text{ on } [0,a], \qquad \operatorname{supp} f_+ \subset [-1, a+1],$$

and

$$0 \leqslant f_- \leqslant 1, \qquad f_- \equiv 1 \text{ on } [1,a-1], \qquad \operatorname{supp} f_- \subset [0,a],$$

with

$$|f_\pm'| + |f_\pm''| \leqslant C.$$

We claim that, for either $f = f_+$ or $f = f_-$,

$$\left|\sum_{x \in P} f(x) - \frac{1}{\pi}\int_{\mathbb{R}} f(x)\,\mathrm{d}x\right| \leqslant CMn^2. \tag{3.37}$$

To prove this claim, choose the cutoff in the almost-analytic extension (3.32) so that

$$\chi \equiv 1 \text{ on } [-a,a], \qquad \operatorname{supp}\chi \subset [-a-1, a+1], \qquad |\chi'| \leqslant C.$$

We apply Lemma 3.11 with $s_\varrho(w) \equiv \mathrm{i}$, corresponding to $\varrho \equiv 1$, and with $\eta = 1$. Since

$$\operatorname{supp} f \subset [-1, a+1] \subset [-2^n, 2^n], \qquad \operatorname{supp}\chi \subset [-a-1, a+1] \subset [-2^n, 2^n],$$

the estimates from Lemma 3.9 apply throughout the region of integration.

We now bound the four terms in Lemma 3.11. For the term II, we use

$$y\,|\operatorname{Im}[s(x+\mathrm{i}y)-\mathrm{i}]| \leqslant Mn, \qquad 0<y\leqslant 1,$$

and the fact that $f''$ is supported in two intervals of O(1)-length. Thus

$$\mathrm{II} \leqslant CMn.$$

For IV, using $|s(x+\mathrm{i})-\mathrm{i}| \leqslant Mn$ and the fact that $f'$ is supported in two intervals of O(1)-length gives

$$\mathrm{IV} \leqslant CMn.$$

For III, note that $\partial_y(y\chi(y)) = 1$ for $1 \leqslant y \leqslant a$, while the transition region has O(1)-length. Since

$$|s(x+\mathrm{i}y)-\mathrm{i}| \leqslant \frac{Mn}{y}, \qquad 1 \leqslant y \leqslant a+1,$$

and $f'$ is supported in a set of O(1)-measure, we obtain

$$\mathrm{III} \leqslant CMn \int_1^{a+1} \frac{\mathrm{d}y}{y} \leqslant CMn^2.$$

Finally, I is supported where $\chi' \neq 0$, hence where $y \in [a, a+1]$. Using again $|s(x+\mathrm{i}y)-\mathrm{i}| \leqslant Mn/y$, together with $|\operatorname{supp} f| \leqslant Ca$ and $|\operatorname{supp} f'| \leqslant C$, yields

$$\mathrm{I} \leqslant CMn.$$

Combining these bounds gives (3.37).

We now apply (3.37) to $f_+$. Since $f_+ \geqslant \mathbf{1}_{[0,a]}$,

$$\#(P\cap[0,a]) \leqslant \sum_{x\in P} f_+(x) \leqslant \frac{1}{\pi}\int_{\mathbb{R}} f_+(x)\,\mathrm{d}x + CMn^2 \leqslant \frac{a}{\pi} + C + CMn^2.$$

Similarly, applying (3.37) to $f_-$ and using $f_- \leqslant \mathbf{1}_{[0,a]}$, we get

$$\#(P\cap[0,a]) \geqslant \sum_{x\in P} f_-(x) \geqslant \frac{1}{\pi}\int_{\mathbb{R}} f_-(x)\,\mathrm{d}x - CMn^2 \geqslant \frac{a}{\pi} - C - CMn^2.$$

Combining the upper and lower bounds, and using $n \asymp \log a$, gives

$$\left|\#(P\cap[0,a]) - \frac{a}{\pi}\right| \leqslant M(\omega)(\log a)^2,$$

after increasing $M(\omega)$. This proves the first estimate in (3.35). □

**3.4.4 Proof of Equation (3.7).** The proof compares $s$ with an explicitly constructed principal-value sum (3.38) having the same poles and the same normalization at infinity; the Nevanlinna representation then shows that their imaginary parts agree, so their difference is a real constant, which must vanish by the normalization.

*Proof of* (3.7). Define

$$\widetilde{s}(w) := -\cot w + \sum_{j\in\mathbb{Z}} \left(\frac{1}{x_j - w} - \frac{1}{\pi j - w}\right), \qquad w \in \mathbb{C}\setminus\mathbb{R}. \tag{3.38}$$

By (3.6), the series in (3.38) converges absolutely and locally uniformly on $\mathbb{C}\setminus\mathbb{R}$. Indeed, for every compact $K\subset\mathbb{C}\setminus\mathbb{R}$,

$$\left|\frac{1}{x_j-w}-\frac{1}{\pi j-w}\right|=\frac{|x_j-\pi j|}{|x_j-w|\,|\pi j-w|}\leqslant C_K\frac{(\log(2+|j|))^2}{1+j^2},\qquad w\in K. \tag{3.39}$$

Hence $\widetilde{s}$ is holomorphic on $\mathbb{C}\setminus\mathbb{R}$. Using the partial fraction expansion

$$-\cot w=\mathrm{P.V.}\sum_{j\in\mathbb{Z}}\frac{1}{\pi j-w},$$

we obtain

$$\widetilde{s}(w)=\mathrm{P.V.}\sum_{j\in\mathbb{Z}}\frac{1}{x_j-w}.$$

We next check the normalization at infinity. Since

$$-\cot(\mathrm{i}y)\to\mathrm{i},\qquad y\to+\infty,$$

it remains to show that the summation term in (3.38) tends to zero at $w=\mathrm{i}y$. Split the sum into $|j|\leqslant y$ and $|j|>y$. For $|j|\leqslant y$, (3.6) implies

$$|x_j-\mathrm{i}y|+|\pi j-\mathrm{i}y|\gtrsim y,$$

and therefore

$$\sum_{|j|\leqslant y}\left|\frac{1}{x_j-\mathrm{i}y}-\frac{1}{\pi j-\mathrm{i}y}\right|\leqslant\sum_{|j|\leqslant y}\frac{C(\log(2+|j|))^2}{y^2}\leqslant\frac{C(\log y)^2}{y}.$$

For $|j|>y$, (3.6) gives

$$|x_j-\mathrm{i}y|+|\pi j-\mathrm{i}y|\gtrsim|j|,$$

and hence

$$\sum_{|j|>y}\left|\frac{1}{x_j-\mathrm{i}y}-\frac{1}{\pi j-\mathrm{i}y}\right|\leqslant\sum_{|j|>y}\frac{C(\log(2+|j|))^2}{j^2}\leqslant\frac{C(\log y)^2}{y}.$$

Thus the summation term in (3.38) tends to zero, and therefore

$$\widetilde{s}(\mathrm{i}y)\longrightarrow\mathrm{i},\qquad y\to+\infty.$$

On the other hand, Lemma 3.9 implies

$$|s(\mathrm{i}y)-\mathrm{i}|\leqslant\frac{M\log(2+y)}{y}\longrightarrow 0,\qquad y\to+\infty.$$

Thus

$$s(\mathrm{i}y)\longrightarrow\mathrm{i}.$$

For $w=E+\mathrm{i}\eta\in\mathbb{C}_+$, the definition of $\widetilde{s}$ gives

$$\mathrm{Im}[\widetilde{s}(w)]=\sum_{j\in\mathbb{Z}}\frac{\eta}{(x_j-E)^2+\eta^2}.$$

Since $s$ is a Nevanlinna function whose representing measure is $\sum_{j\in\mathbb{Z}}\delta_{x_j}$, its imaginary part has the form

$$\mathrm{Im}[s(w)]=c\eta+\sum_{j\in\mathbb{Z}}\frac{\eta}{(x_j-E)^2+\eta^2}$$

for some $c \geqslant 0$. The convergence $s(\mathrm{i}y) \to \mathrm{i}$ as $y \to +\infty$ forces $c = 0$. Therefore,

$$\operatorname{Im}[s(w)] = \operatorname{Im}[\widetilde{s}(w)], \qquad w \in \mathbb{C}_+.$$

It follows that

$$g(w) := s(w) - \widetilde{s}(w)$$

is holomorphic on $\mathbb{C}_+$ and has identically vanishing imaginary part. Hence $g$ is a real constant on $\mathbb{C}_+$. Since both $s(\mathrm{i}y)$ and $\widetilde{s}(\mathrm{i}y)$ converge to i as $y \to +\infty$, this constant is zero. Therefore

$$s(w) = \widetilde{s}(w) = \mathrm{P.V.} \sum_{j \in \mathbb{Z}} \frac{1}{x_j - w}, \qquad w \in \mathbb{C}_+.$$

By Schwarz reflection, the same identity holds for all $w \in \mathbb{C} \setminus \mathbb{R}$. This proves (3.7). $\square$

# 4 Exponential observables

In this section, we introduce certain observables, which are exponential linear statistics of the Nevanlinna function $s$. Equivalently, when the summation representation of $s$ is available, namely (3.7) and (3.11), they can be viewed as products of ratios of linear factors over the random point configuration. These observables are the analogues of ratios of characteristic polynomials. Using the concentration estimates Proposition 3.3 and Proposition 3.4 as input, we derive the asymptotic behavior of these observables at infinity.

## 4.1 Bulk exponential observables

**Proposition 4.1** (Bulk exponential observables). *Adopt Assumption 1.4, and assume*

$$\frac{\beta}{2} = \frac{p}{q} \in \mathbb{Q}_{>0}.$$

*Fix $k \geqslant 1$, signs $\tau_\alpha \in \{\pm 1\}$, and points*

$$t_r^{(\alpha)},\, s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k.$$

*For each triple $(\alpha, r, a)$, choose a piecewise $C^1$ path*

$$\gamma_{r,a}^{(\alpha)} \subset \mathbb{C}_{\tau_\alpha}$$

*from $s_a^{(\alpha)}$ to $t_r^{(\alpha)}$, and define*

$$G := -\frac{1}{q} \sum_{\alpha=1}^{k} \sum_{r=1}^{p} \sum_{a=1}^{q} \int_{\gamma_{r,a}^{(\alpha)}} s(u)\, \mathrm{d}u. \tag{4.1}$$

*Since $s$ is holomorphic on each half-plane, $G$ is independent of the choice of paths.*

*From the representation (3.7), the exponential $\mathrm{e}^G$ coincides with the principal-value product*

$$e^G = \mathrm{P.V.} \prod_{j \in \mathbb{Z}} \prod_{\alpha=1}^{k} \frac{\prod_{r=1}^{p} (t_r^{(\alpha)} - x_j)}{\prod_{a=1}^{q} (s_a^{(\alpha)} - x_j)^{\beta/2}},$$

*where the powers $(s_a^{(\alpha)} - x_j)^{\beta/2}$ are defined using the principal branch on $\mathbb{C} \setminus \mathbb{R}_{<0}$, as specified in (1.11). Moreover, this random variable is integrable. We set*

$$F := \mathbb{E}[e^G]. \tag{4.2}$$

*Define*

$$M := -\mathrm{i}\sum_{\alpha=1}^{k}\tau_\alpha\left(\sum_{r=1}^{p}t_r^{(\alpha)}-\frac{\beta}{2}\sum_{a=1}^{q}s_a^{(\alpha)}\right), \tag{4.3}$$

$$L := \sum_{\alpha=1}^{k}\sum_{r=1}^{p}\sum_{a=1}^{q}\log\left(1+\frac{|t_r^{(\alpha)}-s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}}\right). \tag{4.4}$$

*Then the following estimates hold.*

(i) *There exists a constant $C>0$, depending only on $p,q,k$, such that*

$$|F| \leqslant e^{\operatorname{Re}[M]+C(L+L^2)}. \tag{4.5}$$

(ii) *There exists a deterministic constant $L_0>0$, depending only on $p,q,k$, such that, whenever $L\leqslant L_0$,*

$$F = e^{M+\mathrm{O}(L)} \tag{4.6}$$

*where the implicit constant depends only on $p,q,k$.*

**Corollary 4.2.** *Adopt Assumption 1.4, and assume*

$$\frac{\beta}{2}=\frac{p}{q}\in\mathbb{Q}_{>0}, \qquad n=kp, \qquad m=kq.$$

*Fix*

$$t_r,s_a\in\mathbb{C}_+\cup\mathbb{C}_-, \qquad 1\leqslant r\leqslant n, \quad 1\leqslant a\leqslant m.$$

*Define*

$$M := \sum_{r=1}^{n}|\operatorname{Im}[t_r]|-\frac{\beta}{2}\sum_{a=1}^{m}|\operatorname{Im}[s_a]|, \tag{4.7}$$

$$L := \sum_{r=1}^{n}\sum_{a=1}^{m}\log\left(1+\frac{|(\operatorname{Re}[t_r]+\mathrm{i}|\operatorname{Im}[t_r]|)-(\operatorname{Re}[s_a]+\mathrm{i}|\operatorname{Im}[s_a]|)|}{\sqrt{|\operatorname{Im}[t_r]|\,|\operatorname{Im}[s_a]|}}\right). \tag{4.8}$$

*Then there exists a constant $C>0$, depending only on $p,q,k$, such that,*

$$\mathbb{E}\left[\left|\mathrm{P.V.}\prod_{j\in\mathbb{Z}}\frac{\prod_{r=1}^{n}(t_r-x_j)}{\prod_{a=1}^{m}(s_a-x_j)^{\beta/2}}\right|\right]\leqslant\mathbb{E}\left[\left|\mathrm{P.V.}\prod_{j\in\mathbb{Z}}\frac{\prod_{r=1}^{n}(t_r-x_j)}{\prod_{a=1}^{m}(s_a-x_j)^{\beta/2}}\right|^2\right]^{1/2}\leqslant e^{M+C(L+L^2)}. \tag{4.9}$$

*Moreover, the exponential observable*

$$F(t_1,\dots,t_n;s_1,\dots,s_m) := \mathbb{E}\left[\mathrm{P.V.}\prod_{j\in\mathbb{Z}}\frac{\prod_{r=1}^{n}(t_r-x_j)}{\prod_{a=1}^{m}(s_a-x_j)^{\beta/2}}\right] \tag{4.10}$$

*is holomorphic on $\mathbb{C}^n\times(\mathbb{C}_+\cup\mathbb{C}_-)^m$.*

*Proof of Proposition 4.1.* We first recall the concentration estimates in both half-planes. By (3.5) and Schwarz reflection, for every $z\in\mathbb{C}_\tau$, $\tau\in\{\pm1\}$, and every integer $\ell\geqslant 1$,

$$\mathbb{E}\left[|s(z)-\mathrm{i}\tau|^\ell\right]\leqslant\mathbb{E}\left[|s(z)-\mathrm{i}\tau|^{2\ell}\right]^{1/2}\leqslant\frac{(C\sqrt{\ell})^\ell}{|\operatorname{Im}[z]|^\ell}\Longrightarrow\|s(z)-\mathrm{i}\tau\|_{L^\ell}\leqslant\frac{C\ell}{|\operatorname{Im}[z]|} \tag{4.11}$$

For $w, z \in \mathbb{C}_\tau$, define

$$Y_\tau(w,z) := -\int_\gamma \big(s(u) - \mathrm{i}\tau\big)\,\mathrm{d}u,$$

where $\gamma \subset \mathbb{C}_\tau$ is any piecewise $C^1$ path from $z$ to $w$. Again, $Y_\tau(w,z)$ is path-independent because $s$ is holomorphic on $\mathbb{C}_\tau$.

By Minkowski's inequality and (4.11),

$$\|Y_\tau(w,z)\|_{L^\ell} \leqslant \int_\gamma \|s(u) - \mathrm{i}\tau\|_{L^\ell}\,|\mathrm{d}u| \leqslant C\sqrt{\ell}\int_\gamma \frac{|\mathrm{d}u|}{|\operatorname{Im}[u]|}.$$

Choose $\gamma$ to be the hyperbolic geodesic (see e.g. [10, Section 7]) from $z$ to $w$ in the half-plane $\mathbb{C}_\tau$. Then the last integral is the hyperbolic distance, so

$$\|Y_\tau(w,z)\|_{L^\ell} \leqslant C\sqrt{\ell}\cdot 2\operatorname{arsinh}\left(\frac{|w-z|}{2\sqrt{\operatorname{Im}[w]\operatorname{Im}[z]}}\right) \leqslant 2C\sqrt{\ell}\,\log\left(1 + \frac{|w-z|}{\sqrt{|\operatorname{Im}[w]|\,|\operatorname{Im}[z]|}}\right). \tag{4.12}$$

Now decompose the integral in (4.1) as

$$-\int_{\gamma^{(\alpha)}_{r,a}} s(u)\,\mathrm{d}u = -\mathrm{i}\tau_\alpha\big(t_r^{(\alpha)} - s_a^{(\alpha)}\big) + Y_{\tau_\alpha}\big(t_r^{(\alpha)}, s_a^{(\alpha)}\big).$$

Summing over $(\alpha, r, a)$ and using $p/q = \beta/2$, we obtain

$$G = -\mathrm{i}\sum_{\alpha=1}^k \tau_\alpha\left(\sum_{r=1}^p t_r^{(\alpha)} - \frac{\beta}{2}\sum_{a=1}^q s_a^{(\alpha)}\right) + \Xi = M + \Xi,$$

where

$$\Xi := \frac{1}{q}\sum_{\alpha=1}^k\sum_{r=1}^p\sum_{a=1}^q Y_{\tau_\alpha}\big(t_r^{(\alpha)}, s_a^{(\alpha)}\big). \tag{4.13}$$

Applying Minkowski again and using (4.12), we get

$$\|\Xi\|_{L^\ell} \leqslant \frac{1}{q}\sum_{\alpha=1}^k\sum_{r=1}^p\sum_{a=1}^q \Big\|Y_{\tau_\alpha}\big(t_r^{(\alpha)}, s_a^{(\alpha)}\big)\Big\|_{L^\ell} \leqslant C\sqrt{\ell}\,L. \tag{4.14}$$

In particular,

$$|\mathbb{E}[\Xi]| \leqslant \|\Xi\|_{L^1} \leqslant CL. \tag{4.15}$$

To prove (4.5), note that $\operatorname{Re}[\Xi]$ is a real sub-Gaussian random variable with scale $CL$ and mean bounded by $CL$, by (4.14) and (4.15). Hence

$$\mathbb{E}[e^{\operatorname{Re}[\Xi]}] \leqslant e^{C(L+L^2)}.$$

Therefore

$$|F| \leqslant e^{\operatorname{Re}[M]}\,\mathbb{E}[e^{\operatorname{Re}[\Xi]}] \leqslant e^{\operatorname{Re}[M] + C(L+L^2)},$$

which proves (4.5).

We now turn to the small-$L$ regime. By Taylor's expansion

$$\mathbb{E}[e^{\Xi}] - 1 = \sum_{\ell=1}^\infty \frac{\mathbb{E}[\Xi^\ell]}{\ell!}.$$

Using (4.14),

$$\left|\frac{\mathbb{E}[\Xi^\ell]}{\ell!}\right| \leqslant \frac{\|\Xi\|_{L^\ell}^\ell}{\ell!} \leqslant \frac{(C\sqrt{\ell}\,L)^\ell}{\ell!}.$$

By Stirling's bound $\ell! \geqslant (\ell/e)^\ell$,

$$\frac{(C\sqrt{\ell}\,L)^\ell}{\ell!} \leqslant \left(\frac{CeL}{\sqrt{\ell}}\right)^\ell \leqslant (C_1 L)^\ell.$$

Hence there exists $L_0 > 0$ such that for all $L \leqslant L_0$,

$$\left|\mathbb{E}[e^{\Xi}] - 1\right| \leqslant \sum_{\ell=1}^{\infty} (C_1 L)^\ell \leqslant C_2 L.$$

Consequently,

$$|e^{-M} F - 1| \leqslant C_2 L, \tag{4.16}$$

and (4.6) follows. □

*Proof of Corollary 4.2.* The first inequality in (4.9) is the Cauchy–Schwarz inequality.

For the second inequality, note first that both sides of (4.9) are unchanged if any individual $t_r$ or $s_a$ is replaced by its complex conjugate. Indeed, on the left this is immediate since each $x_j \in \mathbb{R}$, so the modulus of the principal-value product depends only on $\mathrm{Re}[t_r], |\operatorname{Im}[t_r]|, \mathrm{Re}[s_a], |\operatorname{Im}[s_a]|$. On the right, the quantities $M$ and $L$ are defined precisely in terms of these data. Therefore, without loss of generality, we may assume that $t_1, \ldots, t_n,\ s_1, \ldots, s_m \in \mathbb{C}_+$.

Set

$$X := \mathrm{P.V.} \prod_{j \in \mathbb{Z}} \frac{\prod_{r=1}^n (t_r - x_j)}{\prod_{a=1}^m (s_a - x_j)^{\beta/2}}.$$

Since $x_j \in \mathbb{R}$, we then have

$$|X|^2 = \mathrm{P.V.} \prod_{j \in \mathbb{Z}} \frac{\prod_{r=1}^n (t_r - x_j)(\overline{t_r} - x_j)}{\prod_{a=1}^m (s_a - x_j)^{\beta/2} (\overline{s_a} - x_j)^{\beta/2}}. \tag{4.17}$$

Next choose arbitrary partitions

$$\{1, \ldots, n\} = I_1 \sqcup \cdots \sqcup I_k, \qquad \{1, \ldots, m\} = A_1 \sqcup \cdots \sqcup A_k,$$

with $|I_\alpha| = p$ and $|A_\alpha| = q$ for every $1 \leqslant \alpha \leqslant k$. Using these partitions, define $2k$ blocks by

$$\begin{aligned}
&t_r^{(\alpha)} := t_{I_\alpha(r)}, \qquad s_a^{(\alpha)} := s_{A_\alpha(a)}, \qquad \tau_\alpha := +1, \qquad 1 \leqslant \alpha \leqslant k, \quad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q,\\
&t_r^{(k+\alpha)} := \overline{t_{I_\alpha(r)}}, \qquad s_a^{(k+\alpha)} := \overline{s_{A_\alpha(a)}}, \qquad \tau_{k+\alpha} := -1, \qquad 1 \leqslant \alpha \leqslant k, \quad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q,
\end{aligned}$$

where $I_\alpha(r)$ is the $r$-th element of $I_\alpha$, and $A_\alpha$ is the $a$-th element of $A_\alpha$. Then (4.17) is exactly of the form covered by Proposition 4.1. Hence (4.5) yields

$$\mathbb{E}[|X|^2] \leqslant \exp\{\mathrm{Re}[M_*] + C(L_* + L_*^2)\}, \tag{4.18}$$

where

$$M_* := -\mathrm{i} \sum_{\alpha=1}^{2k} \tau_\alpha \left( \sum_{r=1}^{p} t_r^{(\alpha)} - \frac{\beta}{2} \sum_{a=1}^{q} s_a^{(\alpha)} \right),$$

and

$$L_* := \sum_{\alpha=1}^{2k} \sum_{r=1}^{p} \sum_{a=1}^{q} \log\left(1 + \frac{|t_r^{(\alpha)} - s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}}\right).$$

Recall $M$ and $L$ from (4.7) and (4.8). Since all $t_r, s_a \in \mathbb{C}_+$, we have

$$\mathrm{Re}[M_*] = 2\sum_{r=1}^{n} \operatorname{Im}[t_r] - \beta \sum_{a=1}^{m} \operatorname{Im}[s_a] = 2M.$$

Moreover,

$$L_* = 2\sum_{\alpha=1}^{k}\sum_{r=1}^{p}\sum_{a=1}^{q}\log\left(1+\frac{|t_r^{(\alpha)}-s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}}\right) = 2\sum_{r=1}^{n}\sum_{a=1}^{m}\log\left(1+\frac{|t_r-s_a|}{\sqrt{\operatorname{Im}[t_r]\operatorname{Im}[s_a]}}\right) = 2L.$$

Substituting these relations into (4.18), taking square root, and enlarging the constant $C$ if necessary, we obtain

$$\mathbb{E}[|X|^2]^{1/2} \leqslant e^{M+C(L+L^2)},$$

which is exactly the second inequality in (4.9).

In the following we prove that $F$ as in (4.10) is holomorphic on

$$\mathbb{C}^n \times (\mathbb{C}_+ \cup \mathbb{C}_-)^m.$$

Since $(\mathbb{C}_+ \cup \mathbb{C}_-)^m$ is a disjoint union of connected components, it suffices to fix a sign vector

$$\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m) \in \{\pm 1\}^m$$

and prove holomorphicity on

$$\Omega_{\boldsymbol{\eta}} := \mathbb{C}^n \times \prod_{a=1}^{m} \mathbb{C}_{\eta_a}.$$

Fix $s_1, \ldots, s_m \in \prod_{a=1}^m \mathbb{C}_{\eta_a}$. For each realization of the point process, the principal-value product

$$\mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m) := \mathrm{P.V.}\prod_{j\in\mathbb{Z}} \frac{\prod_{r=1}^n (t_r - x_j)}{\prod_{a=1}^m (s_a - x_j)^{\beta/2}} \tag{4.19}$$

is entire in $t_1, \ldots, t_n$. Moreover, after fixing the branches of $(s_a - x_j)^{\beta/2}$ on the half-plane $\mathbb{C}_{\eta_a}$, it is holomorphic in each $s_a \in \mathbb{C}_{\eta_a}$.

We first prove a local $L^1$ bound. By Cauchy–Schwarz,

$$\mathbb{E}\big[|\mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)|\big] \leqslant \mathbb{E}\big[|\mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)|^2\big]^{1/2}. \tag{4.20}$$

Let $K \subset \Omega_{\boldsymbol{\eta}}$ be compact. Choose $R > 0$ such that

$$R \geqslant \max_{1\leqslant r\leqslant n}\sup_{(t,s)\in K} |\operatorname{Im}[t_r]| + 1.$$

For $(t_1, \cdots, t_n, s_1, \cdots, s_m) \in K$, define

$$\widehat{t}_r := \operatorname{Re}[t_r] + \mathrm{i}R, \qquad 1 \leqslant r \leqslant n.$$

Since $x_j \in \mathbb{R}$, we have $|t_r - x_j| \leqslant |\widehat{t}_r - x_j|$, for $j \in \mathbb{Z}$. Therefore, first for finite symmetric truncations of the principal-value product and then by passing to the limit,

$$|\mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)| \leqslant |\mathcal{G}(\widehat{t}_1, \ldots, \widehat{t}_n; s_1, \ldots, s_m)|.$$

The variables $\widehat{t}_1, \ldots, \widehat{t}_n$ lie in the upper half-plane and range over a compact subset of $(\mathbb{C}_+)^n$, while $(s_1, \ldots, s_m)$ ranges over a compact subset of $\mathbb{C}_{\eta_1} \times \cdots \times \mathbb{C}_{\eta_m}$. Hence the absolute bound (4.9) implies

$$\sup_{(t,s)\in K} \mathbb{E}\big[|\mathcal{G}(\widehat{t}_1, \ldots, \widehat{t}_n; s_1, \ldots, s_m)|^2\big]^{1/2} < \infty.$$

Consequently,

$$\sup_{(t,s)\in K} \mathbb{E}\big[|\mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)|\big] < \infty. \tag{4.21}$$

We now prove separate holomorphicity. Fix all variables except $t_r$, and let $\omega_r$ be any closed contour in the $t_r$-plane. The contour together with the fixed values of the remaining variables is contained in some compact subset of $\Omega_{\boldsymbol{\eta}}$. By (4.21), Fubini's theorem allows us to interchange expectation and contour integration:

$$\int_{\omega_r} F(t_1, \ldots, t_n; s_1, \ldots, s_m)\, \mathrm{d}t_r = \mathbb{E}\left[\int_{\omega_r} \mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)\, \mathrm{d}t_r\right].$$

For each realization, $t_r \mapsto \mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)$ is entire, so the inner contour integral vanishes. Hence

$$\int_{\omega_r} F(t_1, \ldots, t_n; s_1, \ldots, s_m)\, \mathrm{d}t_r = 0.$$

By Morera's theorem, $F$ is holomorphic in each $t_r$.

The same argument applies to each $s_a$, with contours lying inside the fixed half-plane $\mathbb{C}_{\eta_a}$. Indeed, for each realization, $s_a \mapsto \mathcal{G}(t_1, \ldots, t_n; s_1, \ldots, s_m)$ is holomorphic on $\mathbb{C}_{\eta_a}$, and the local bound (4.21) justifies the same interchange of expectation and contour integration. Thus $F$ is separately holomorphic in all variables on $\Omega_{\boldsymbol{\eta}}$.

Finally, (4.21) shows that $F$ is locally bounded on $\Omega_{\boldsymbol{\eta}}$. By the Hartogs–Osgood theorem (see [62, Theorem 1.2.5]), separate holomorphicity together with local boundedness implies joint holomorphicity. Therefore $F$ is jointly holomorphic on $\Omega_{\boldsymbol{\eta}}$. Since $\boldsymbol{\eta}$ was arbitrary, $F$ is holomorphic on $\mathbb{C}^n \times (\mathbb{C}_+ \cup \mathbb{C}_-)^m$. □

## 4.2 Edge exponential observables

**Proposition 4.3** (Edge exponential observables)**.** *Adopt Assumption 1.10, and assume*

$$\frac{\beta}{2} = \frac{p}{q} \in \mathbb{Q}_{>0}.$$

*Fix $k \geqslant 1$, signs $\tau_\alpha \in \{\pm 1\}$, and points*

$$t_r^{(\alpha)},\ s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k. \tag{4.22}$$

*For each triple $(\alpha, r, a)$, choose a piecewise $C^1$ path*

$$\gamma_{r,a}^{(\alpha)} \subset \mathbb{C}_{\tau_\alpha}$$

*from $s_a^{(\alpha)}$ to $t_r^{(\alpha)}$, and define*

$$G := -\frac{1}{q} \sum_{\alpha=1}^{k} \sum_{r=1}^{p} \sum_{a=1}^{q} \int_{\gamma_{r,a}^{(\alpha)}} s(u)\, \mathrm{d}u. \tag{4.23}$$

*Since $s$ is holomorphic on each half-plane, $G$ is independent of the choice of paths. From the representation (3.11), $e^G$ admits the renormalized product representation*

$$e^G = \prod_{\alpha=1}^{k} e^{c_0\left(\sum_{r=1}^{p} t_r^{(\alpha)} - \frac{\beta}{2}\sum_{a=1}^{q} s_a^{(\alpha)}\right)} \times \prod_{j \geqslant 1} \prod_{\alpha=1}^{k} \frac{\prod_{r=1}^{p} (t_r^{(\alpha)} - x_j) e^{t_r^{(\alpha)}/\mathfrak{a}_j}}{\prod_{a=1}^{q} \left((s_a^{(\alpha)} - x_j) e^{s_a^{(\alpha)}/\mathfrak{a}_j}\right)^{\beta/2}}, \tag{4.24}$$

*where $c_0 := \mathrm{Ai}'(0)/\mathrm{Ai}(0)$, and the powers $(s_a^{(\alpha)} - x_j)^{\beta/2}$ are defined using the principal branch on $\mathbb{C} \setminus \mathbb{R}_{<0}$, as specified in (1.11). Moreover, this random variable is integrable. We set*

$$F := \mathbb{E}[e^G]. \tag{4.25}$$

*Then $F$ is holomorphic on the domain specified in (4.22).*

*Define*

$$M := -\frac{2}{3}\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\left(t_r^{(\alpha)}\right)^{3/2} - \frac{\beta}{2}\sum_{a=1}^{q}\left(s_a^{(\alpha)}\right)^{3/2}\right), \tag{4.26}$$

$$L := \sum_{\alpha=1}^{k}\sum_{r=1}^{p}\sum_{a=1}^{q}\log\left(1 + \frac{|t_r^{(\alpha)} - s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}}\right). \tag{4.27}$$

*Here the powers $w^{3/2}$ are taken with the principal branch.*

*If, in addition,*

$$|\arg t_r^{(\alpha)}|,\ |\arg s_a^{(\alpha)}| \in (\pi/4, \pi), \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k, \tag{4.28}$$

*then there exists a deterministic constant $L_0 > 0$, depending only on $p, q, k$, such that, whenever $L \leqslant L_0$,*

$$F = \exp\{M + \mathrm{O}(L)\}, \tag{4.29}$$

*where the implicit constant depends only on $p, q, k$.*

*Proof.* The proof is parallel to that of the bulk estimate, so we only indicate the changes.

Write

$$G = M + \Xi,$$

where $M$ is from (4.26) and

$$\Xi := -\frac{1}{q}\sum_{\alpha=1}^{k}\sum_{r=1}^{p}\sum_{a=1}^{q}\int_{\gamma_{r,a}^{(\alpha)}}\left(s(u) - \sqrt{u}\right)\mathrm{d}u.$$

We recall the edge concentration estimate (3.8): for any $w = E + \mathrm{i}\eta$ with $\eta > 0$, and for every $q \geqslant 1$,

$$\mathbb{E}\left[\left|s(w) - \sqrt{w}\right|^{2q}\right] \leqslant \frac{(Cq)^q}{\eta^{2q}}\frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}. \tag{4.30}$$

We also note that $s(\overline{w}) = \overline{s(w)}$ and $\sqrt{\overline{w}} = \overline{\sqrt{w}}$. Therefore, $s(w) - \sqrt{w}$ has a uniform sub-Gaussian tail on every compact subset of $\mathbb{C} \setminus \mathbb{R}$. In particular, $e^G$ is integrable. And $F$, as defined in (4.25), is holomorphic on the domain specified in (4.22).

When $\arg w \in (\pi/4, \pi)$, we have

$$\operatorname{Im}[\sqrt{w}] \asymp |w|^{1/2},$$

and hence (4.30) reduces to

$$\mathbb{E}\left[\left|s(w) - \sqrt{w}\right|^{2q}\right] \leqslant \frac{(Cq)^q}{\eta^{2q}}. \tag{4.31}$$

This is the same form as the bulk concentration estimate (3.5).

Thus, under the assumption (4.28), using (4.31) as input, noticing that the region $\arg w \in (\pi/4, \pi)$ is hyperbolically geodesically convex, exactly as in the proof of Proposition 4.1, this gives, for every integer $\ell \geqslant 1$,

$$\|\Xi\|_{L^\ell} \leqslant C\sqrt{\ell}\, L.$$

Therefore

$$F = \mathbb{E}[e^G] = e^M\mathbb{E}[e^\Xi].$$

As in the bulk case, the sub-Gaussian bound implies

$$\left|\mathbb{E}[e^\Xi] - 1\right| \leqslant CL$$

provided $L \leqslant L_0$, for some sufficiently small deterministic constant $L_0 > 0$. Hence

$$F = e^{M + \mathrm{O}(L)},$$

which is exactly (4.29). □

# 5 Linear differential equations

In this section, we prove that the exponential observables introduced in Section 4 satisfy linear systems of second-order differential equations. These systems are the bulk and edge deformed CMS equations corresponding to the two sets of variables $t$ and $s$. They are obtained from the loop equation hierarchy for the Nevanlinna function. While this hierarchy is nonlinear in the Nevanlinna function itself, its action on exponential observables linearizes, yielding the deformed CMS equations.

## 5.1 Bulk deformed Calogero–Moser–Sutherland operators

**Proposition 5.1.** *Adopt Assumption 1.4, assume that $\beta/2 \in \mathbb{Q}_{>0}$, and let*

$$n = \frac{\beta m}{2}.$$

*Define the exponential observable*

$$F(t_1, \dots, t_n; s_1, \dots, s_m) := \mathbb{E}\left[\mathrm{P.V.} \prod_{j\in\mathbb{Z}} \frac{\prod_{r=1}^{n}(t_r - x_j)}{\prod_{a=1}^{m}(s_a - x_j)^{\beta/2}}\right], \tag{5.1}$$

*where the expectation is taken with respect to the underlying random point process, and the powers $(s_a - x_j)^{\beta/2}$ are defined using the principal branch on $\mathbb{C} \setminus \mathbb{R}_{<0}$, as specified in (1.11).*

*Assume that $t_1, \dots, t_n, s_1, \dots, s_m \in \mathbb{C}_+ \cup \mathbb{C}_-$, satisfy the half-plane balance conditions*

$$\sum_{r=1}^{n} \mathbf{1}(t_r \in \mathbb{C}_+) - \frac{\beta}{2}\sum_{a=1}^{m} \mathbf{1}(s_a \in \mathbb{C}_+) = 0, \quad \sum_{r=1}^{n} \mathbf{1}(t_r \in \mathbb{C}_-) - \frac{\beta}{2}\sum_{a=1}^{m} \mathbf{1}(s_a \in \mathbb{C}_-) = 0. \tag{5.2}$$

*Then $F$ satisfies the bulk deformed CMS equations*

$$\left(\partial_{t_i}^2 + 1 + \frac{2}{\beta}\sum_{j\neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^{m} \frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a}\right) F = 0, \qquad 1 \leqslant i \leqslant n, \tag{5.3}$$

*and*

$$\left(\partial_{s_a}^2 + \frac{\beta^2}{4} + \frac{\beta}{2}\sum_{b\neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} - \sum_{i=1}^{n} \frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i}\right) F = 0, \qquad 1 \leqslant a \leqslant m. \tag{5.4}$$

*Moreover, by Corollary 4.2, the observable $F$ defined in (5.1) is holomorphic in $t_1, \dots, t_n$ on $\mathbb{C}^n$, and holomorphic in each $s_a$ on its chosen half-plane. The equations (5.3)–(5.4) continue to hold under this analytic continuation, as meromorphic identities in the variables.*

**Remark 5.2.** *When $\beta = 2$, the two systems (5.3) and (5.4) coincide and reduce to*

$$\begin{aligned}
&\left(\partial_{t_i}^2 + 1 + \sum_{j\neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{j=1}^{m} \frac{\partial_{t_i} + \partial_{s_j}}{t_i - s_j}\right) F = 0, \qquad 1 \leqslant i \leqslant n,\\
&\left(\partial_{s_i}^2 + 1 + \sum_{j\neq i} \frac{\partial_{s_i} - \partial_{s_j}}{s_i - s_j} - \sum_{j=1}^{m} \frac{\partial_{s_i} + \partial_{t_j}}{s_i - t_j}\right) F = 0, \qquad 1 \leqslant i \leqslant n.
\end{aligned} \tag{5.5}$$

*Proof of Proposition 5.1.* We divide the proof into several steps.

*Step 1. Integrated loop equation.* For $u \in \mathbb{C}_+ \cup \mathbb{C}_-$, define

$$\varepsilon(u) = \begin{cases} +1, & u \in \mathbb{C}_+, \\ -1, & u \in \mathbb{C}_-, \end{cases} \qquad S(u) := \int_{\varepsilon(u)\mathrm{i}}^{u} s(\zeta)\,\mathrm{d}\zeta.$$

Then $S$ is holomorphic on each half-plane and satisfies $S'(u) = s(u)$.

Throughout the proof, all difference quotients are interpreted by their holomorphic extensions whenever their two arguments coincide. In particular,

$$\left.\frac{S'(w_1) - S'(w_j)}{w_1 - w_j}\right|_{w_j = w_1} = S''(w_1), \qquad \partial_{w_j} \left.\frac{S'(w_1) - S'(w_j)}{w_1 - w_j}\right|_{w_j = w_1} = \frac{1}{2} S'''(w_1).$$

We start from the bulk loop equation (1.2). For $w_1, \dots, w_p \in \mathbb{C}_+ \cup \mathbb{C}_-$, it gives

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right) \prod_{j=2}^{p} S'(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p} \partial_{w_j} \frac{S'(w_1) - S'(w_j)}{w_1 - w_j} \prod_{\substack{\ell=2 \\ \ell \neq j}}^{p} S'(w_\ell)\right] = 0. \tag{5.6}$$

Because the integrand in the second term is holomorphic in $w_j$, including at $w_j = w_1$ under the convention above, we may integrate (5.6) in each variable $w_j$, $2 \leqslant j \leqslant p$, from $\varepsilon(w_j)\mathrm{i}$ to $w_j$. We obtain

$$\begin{aligned} &\mathbb{E}\left[\left(\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right) \prod_{j=2}^{p} S(w_j)\right] \\ &\quad + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p} \left(\frac{S'(w_1) - S'(w_j)}{w_1 - w_j} - \frac{S'(w_1) - S'(\varepsilon(w_j)\mathrm{i})}{w_1 - \varepsilon(w_j)\mathrm{i}}\right) \prod_{\substack{\ell=2 \\ \ell \neq j}}^{p} S(w_\ell)\right] = 0. \end{aligned} \tag{5.7}$$

*Step 2. Exponential test functions.* Choose

$$O_\ell(u) = e^{-a_\ell u}, \qquad a_\ell \in \{1, -\beta/2\}, \qquad 1 \leqslant \ell \leqslant p,$$

and set

$$\widetilde{O}_\ell(u) := e^{|a_\ell| u}, \qquad u \geqslant 0.$$

Then

$$|O_\ell(S(w_\ell))| + |O'_\ell(S(w_\ell))| \leqslant C\, \widetilde{O}_\ell(|S(w_\ell)|),$$

where $C$ depends only on $\beta$.

Using (2.2), we have

$$\left|\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right| \leqslant C\left(\frac{\operatorname{Im}[s(w_1)]}{\operatorname{Im}[w_1]} + |s(w_1)|^2 + 1\right).$$

Hence

$$\begin{aligned} &\mathbb{E}\left[\left|\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right| \prod_{\ell=1}^{p} \widetilde{O}_\ell(|S(w_\ell)|)\right] \\ &\qquad \leqslant C\,\mathbb{E}\left[\left(\frac{\operatorname{Im}[s(w_1)]}{\operatorname{Im}[w_1]} + |s(w_1)|^2 + 1\right) \prod_{\ell=1}^{p} \widetilde{O}_\ell(|S(w_\ell)|)\right] < \infty. \end{aligned} \tag{5.8}$$

Here the last finiteness follows from (3.5), which gives sub-Gaussian tails for $s(w_1)$, together with the corresponding exponential integrability of the path integrals $S(w_\ell)$.

Similarly, (2.2) and (2.4) give

$$\begin{aligned} &\mathbb{E}\left[\left|S''(w_1) - \frac{S'(w_1) - S'(\varepsilon(w_1)\mathrm{i})}{w_1 - \varepsilon(w_1)\mathrm{i}}\right| \prod_{\ell=1}^{p} \widetilde{O}_\ell(|S(w_\ell)|)\right] < \infty, \\ &\mathbb{E}\left[\left|\frac{S'(w_1) - S'(w_j)}{w_1 - w_j} - \frac{S'(w_1) - S'(\varepsilon(w_j)\mathrm{i})}{w_1 - \varepsilon(w_j)\mathrm{i}}\right| \prod_{\ell=1}^{p} \widetilde{O}_\ell(|S(w_\ell)|)\right] < \infty, \end{aligned} \tag{5.9}$$

for every $2 \leqslant j \leqslant p$. The second estimate remains valid when $w_j = w_1$, in which case it reduces to the first estimate.

To pass from (5.7) to general test functions, we first consider polynomials in $S(w_1), \dots, S(w_p)$. We apply (5.7) at higher levels of the loop hierarchy, allowing additional variables to coincide with any of $w_1, \dots, w_p$, and then group the equal variables. This is legitimate because the difference quotients have been extended holomorphically to the diagonals. A repeated variable at $w_1$ produces the factor

$$S''(w_1) - \frac{S'(w_1) - S'(\varepsilon(w_1)\mathrm{i})}{w_1 - \varepsilon(w_1)\mathrm{i}}$$

multiplying the derivative of the polynomial with respect to its first argument. Similarly, a repeated variable at $w_j$, $j \geqslant 2$, produces

$$\frac{S'(w_1) - S'(w_j)}{w_1 - w_j} - \frac{S'(w_1) - S'(\varepsilon(w_j)\mathrm{i})}{w_1 - \varepsilon(w_j)\mathrm{i}}$$

multiplying the derivative with respect to its $j$-th argument.

Thus (5.7) extends to polynomial test functions of $S(w_1), \dots, S(w_p)$. Using the power-series expansions of the exponentials and the integrability estimates (5.8)–(5.9), we may then pass to the limit by dominated convergence. This gives

$$\begin{aligned} 0 = {} & \mathbb{E}\left[\left(\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right) \prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right] \\ & + \frac{2}{\beta}\mathbb{E}\left[\left(S''(w_1) - \frac{S'(w_1) - S'(\varepsilon(w_1)\mathrm{i})}{w_1 - \varepsilon(w_1)\mathrm{i}}\right) O_1'(S(w_1)) \prod_{\ell=2}^{p} O_\ell(S(w_\ell))\right] \\ & + \frac{2}{\beta}\sum_{j=2}^{p} \mathbb{E}\left[\left(\frac{S'(w_1) - S'(w_j)}{w_1 - w_j} - \frac{S'(w_1) - S'(\varepsilon(w_j)\mathrm{i})}{w_1 - \varepsilon(w_j)\mathrm{i}}\right) O_1(S(w_1)) O_j'(S(w_j)) \prod_{\ell \neq 1, j} O_\ell(S(w_\ell))\right]. \end{aligned} \tag{5.10}$$

Since $O_\ell'(u) = -a_\ell O_\ell(u)$, this becomes

$$\begin{aligned} 0 = {} & \mathbb{E}\left[\left(\frac{2-\beta}{\beta} S''(w_1) + (S'(w_1))^2 + 1\right) \prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right] \\ & - \frac{2a_1}{\beta}\mathbb{E}\left[\left(S''(w_1) - \frac{S'(w_1) - S'(\varepsilon(w_1)\mathrm{i})}{w_1 - \varepsilon(w_1)\mathrm{i}}\right) \prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right] \\ & - \frac{2}{\beta}\sum_{j=2}^{p} a_j \mathbb{E}\left[\left(\frac{S'(w_1) - S'(w_j)}{w_1 - w_j} - \frac{S'(w_1) - S'(\varepsilon(w_j)\mathrm{i})}{w_1 - \varepsilon(w_j)\mathrm{i}}\right) \prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right]. \end{aligned} \tag{5.11}$$

*Step 3. Cancellation of base-point terms.* Assume that

$$\sum_{\ell\colon \varepsilon(w_\ell)=+1} a_\ell = 0, \qquad \sum_{\ell\colon \varepsilon(w_\ell)=-1} a_\ell = 0. \tag{5.12}$$

Then the terms involving $S'(\mathrm{i})$ and $S'(-\mathrm{i})$ cancel in (5.11). Hence (5.11) reduces to

$$\begin{aligned} 0 = \mathbb{E}\left[\left(\frac{2-\beta-2a_1}{\beta}S''(w_1)+(S'(w_1))^2+1\right)\prod_{\ell=1}^{p}O_\ell(S(w_\ell))\right] \\ -\frac{2}{\beta}\sum_{j=2}^{p}a_j\mathbb{E}\left[\frac{S'(w_1)-S'(w_j)}{w_1-w_j}\prod_{\ell=1}^{p}O_\ell(S(w_\ell))\right]. \end{aligned} \tag{5.13}$$

Since $a_1 \in \{1, -\beta/2\}$, we have

$$\frac{2-\beta-2a_1}{\beta} = -\frac{1}{a_1}.$$

Therefore

$$\left(-\frac{1}{a_1}S''(w_1)+(S'(w_1))^2+1\right)e^{-a_1S(w_1)} = \left(\frac{\partial_{w_1}^2}{a_1^2}+1\right)e^{-a_1S(w_1)}.$$

Moreover,

$$-a_j\frac{S'(w_1)-S'(w_j)}{w_1-w_j}\prod_{\ell=1}^{p}O_\ell(S(w_\ell)) = \frac{(a_j/a_1)\partial_{w_1}-\partial_{w_j}}{w_1-w_j}\left(\prod_{\ell=1}^{p}O_\ell(S(w_\ell))\right).$$

Substituting these identities into (5.13), we obtain the master identity

$$\mathbb{E}\left[\left(\frac{\partial_{w_1}^2}{a_1^2}+1\right)\prod_{\ell=1}^{p}O_\ell(S(w_\ell))\right]+\frac{2}{\beta}\sum_{j=2}^{p}\mathbb{E}\left[\frac{(a_j/a_1)\partial_{w_1}-\partial_{w_j}}{w_1-w_j}\left(\prod_{\ell=1}^{p}O_\ell(S(w_\ell))\right)\right]=0. \tag{5.14}$$

*Step 4. Derivation of* (5.3) *and* (5.4) By the representation (3.7)

$$e^{-a_\ell S(w_\ell)} = \prod_{j\in\mathbb{Z}}\frac{(w_\ell-x_j)^{a_\ell}}{(\varepsilon(w_\ell)\mathrm{i}-x_j)^{a_\ell}},$$

the balance condition (5.12) implies that all base-point factors cancel. Therefore

$$\prod_{\ell=1}^{p}O_\ell(S(w_\ell)) = \prod_{j\in\mathbb{Z}}\prod_{\ell=1}^{p}(w_\ell-x_j)^{a_\ell}. \tag{5.15}$$

To derive (5.3), take $w_1 = t_i$ and $a_1 = 1$. Let the remaining variables be the other $t$-variables with exponent 1, and the $s$-variables with exponent $-\beta/2$. Then (5.12) is exactly (5.2), and (5.15) becomes

$$\prod_{j\in\mathbb{Z}}\frac{\prod_{r=1}^{n}(t_r-x_j)}{\prod_{a=1}^{m}(s_a-x_j)^{\beta/2}}.$$

We obtain (5.3) from (5.14).

Similarly, to derive (5.4), take $w_1 = s_i$ and $a_1 = -\beta/2$, with the remaining variables again given by the other $t$- and $s$-variables. Multiplying the resulting (5.14) by $\beta^2/4$, this gives (5.4).

*Step 5. Holomorphic continuation in the t-variables.*

The equations (5.3) and (5.4) hold on the nonempty open domain where the half-plane balance condition (5.2) is satisfied. Since both sides are meromorphic in $t_1, \ldots, t_n$, and by Corollary 4.2, $F$ is entire in these variables, the identities extend by analytic continuation to all $t_1, \ldots, t_n \in \mathbb{C}$, away from the collision hyperplanes where the coefficients have poles. Equivalently, they hold as meromorphic identities in the $t$-variables. This completes the proof. ☐

### 5.2 Edge deformed Calogero–Moser–Sutherland operators

**Proposition 5.3.** *Adopt Assumption 1.10, assume that $\beta/2 \in \mathbb{Q}_{>0}$, and let*

$$n = \frac{\beta m}{2}.$$

*Define the exponential observable*

$$F(t_1, \dots, t_n; s_1, \dots, s_m) = e^{c_0\left(\sum_{r=1}^n t_r - \frac{\beta}{2}\sum_{a=1}^m s_a\right)} \times \mathbb{E}\left[\prod_{j\geqslant 1} \frac{\prod_{r=1}^n (t_r - x_j)e^{t_r/\mathfrak{a}_j}}{\prod_{a=1}^m \left((s_a - x_j)e^{s_a/\mathfrak{a}_j}\right)^{\beta/2}}\right], \tag{5.16}$$

*where $c_0 := \mathrm{Ai}'(0)/\mathrm{Ai}(0)$, and the powers $(s_a^{(\alpha)} - x_j)^{\beta/2}$ are defined using the principal branch on $\mathbb{C} \setminus \mathbb{R}_{<0}$, as specified in (1.11).*

*Assume that $t_1, \dots, t_n, s_1, \dots, s_m \in \mathbb{C}_+ \cup \mathbb{C}_-$ satisfy*

$$\sum_{r=1}^n \mathbf{1}(t_r \in \mathbb{C}_+) - \frac{\beta}{2}\sum_{a=1}^m \mathbf{1}(s_a \in \mathbb{C}_+) = 0, \qquad \sum_{r=1}^n \mathbf{1}(t_r \in \mathbb{C}_-) - \frac{\beta}{2}\sum_{a=1}^m \mathbf{1}(s_a \in \mathbb{C}_-) = 0. \tag{5.17}$$

*Then $F$ satisfies*

$$\left(\partial_{t_i}^2 - t_i + \frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{j=1}^m \frac{\partial_{t_i} + (2/\beta)\partial_{s_j}}{t_i - s_j}\right) F = 0, \qquad 1 \leqslant i \leqslant n, \tag{5.18}$$

*and*

$$\left(\partial_{s_i}^2 - \frac{\beta^2}{4}s_i + \frac{\beta}{2}\sum_{j\neq i}\frac{\partial_{s_i} - \partial_{s_j}}{s_i - s_j} - \sum_{j=1}^n \frac{\partial_{s_i} + (\beta/2)\partial_{t_j}}{s_i - t_j}\right) F = 0, \qquad 1 \leqslant i \leqslant m. \tag{5.19}$$

**Remark 5.4.** *When $\beta = 2$, the two systems (5.18) and (5.19) coincide, and reduce to*

$$\begin{aligned}
&\left(\partial_{t_i}^2 - t_i + \sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{j=1}^m \frac{\partial_{t_i} + \partial_{s_j}}{t_i - s_j}\right) F = 0, \qquad 1 \leqslant i \leqslant n,\\
&\left(\partial_{s_i}^2 - s_i + \sum_{j\neq i}\frac{\partial_{s_i} - \partial_{s_j}}{s_i - s_j} - \sum_{j=1}^m \frac{\partial_{s_i} + \partial_{t_j}}{s_i - t_j}\right) F = 0, \qquad 1 \leqslant i \leqslant n.
\end{aligned} \tag{5.20}$$

*Proof of Proposition 5.3.* For $u \in \mathbb{C}_+ \cup \mathbb{C}_-$, define

$$\varepsilon(u) = \begin{cases} +1, & u \in \mathbb{C}_+,\\ -1, & u \in \mathbb{C}_-,\end{cases} \qquad S(u) := \int_{\varepsilon(u)\mathrm{i}}^{u} s(\zeta)\,\mathrm{d}\zeta.$$

Then $S$ is holomorphic on each half-plane and satisfies $S'(u) = s(u)$.

Assume that

$$\sum_{\ell:\,\varepsilon(w_\ell)=+1} a_\ell = 0, \qquad \sum_{\ell:\,\varepsilon(w_\ell)=-1} a_\ell = 0. \tag{5.21}$$

Take

$$O_\ell(u) = e^{-a_\ell u}, \qquad a_\ell \in \{1, -\beta/2\}.$$

Exactly as in the bulk case, using the edge loop equation in place of the bulk loop equation, this gives the master identity

$$\mathbb{E}\left[\left(\frac{\partial_{w_1}^2}{a_1^2} - w_1\right)\prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right] + \frac{2}{\beta}\sum_{j=2}^{p}\mathbb{E}\left[\frac{(a_j/a_1)\partial_{w_1} - \partial_{w_j}}{w_1 - w_j}\left(\prod_{\ell=1}^{p} O_\ell(S(w_\ell))\right)\right] = 0. \tag{5.22}$$

Next we identify the observable. In the edge case, the natural analogue of the characteristic product is the Airy-regularized product

$$\Gamma_{\boldsymbol{x}}(z) := e^{c_0 z}\prod_{j\geqslant 1}(z - x_j)e^{z/\mathfrak{a}_j}, \qquad c_0 := \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}. \tag{5.23}$$

We emphasize that (5.23) should be understood as a formal regularized product; the ratios appearing in (5.16), however, are well defined.

From the representation (3.11), we formally have

$$\partial_u \log \Gamma_{\boldsymbol{x}}(u) = -s(u) = -S'(u).$$

Therefore,

$$e^{-S(u)} = \frac{\Gamma_{\boldsymbol{x}}(u)}{\Gamma_{\boldsymbol{x}}(\varepsilon(u)\mathrm{i})}.$$

Hence

$$\prod_{\ell=1}^{p} O_\ell(S(w_\ell)) = \prod_{\ell=1}^{p}\left(\frac{\Gamma_{\boldsymbol{x}}(w_\ell)}{\Gamma_{\boldsymbol{x}}(\varepsilon(w_\ell)\mathrm{i})}\right)^{a_\ell}.$$

Under the balance condition (5.21), the base-point factors cancel, and so

$$\prod_{\ell=1}^{p} O_\ell(S(w_\ell)) = \prod_{\ell=1}^{p}\Gamma_{\boldsymbol{x}}(w_\ell)^{a_\ell}. \tag{5.24}$$

To derive (5.18), take $w_1 = t_i$ and $a_1 = 1$. Let the remaining variables be the other $t$-variables with exponent 1, and the $s$-variables with exponent $-\beta/2$. Then (5.21) is exactly (5.17), and (5.24) becomes

$$\exp\left\{c_0\left(\sum_{r=1}^{n} t_r - \frac{\beta}{2}\sum_{a=1}^{m} s_a\right)\right\}\prod_{j\geqslant 1}\frac{\prod_{r=1}^{n}(t_r - x_j)e^{t_r/\mathfrak{a}_j}}{\prod_{a=1}^{m}\left((s_a - x_j)e^{s_a/\mathfrak{a}_j}\right)^{\beta/2}}.$$

We obtain (5.18) from (5.22).

Similarly, to derive (5.19), take $w_1 = s_i$ and $a_1 = -\beta/2$, with the remaining variables again given by the other $t$- and $s$-variables. Multiplying the resulting identity by $\beta^2/4$, this gives (5.19). □

# 6 Unique characterization

In this section, we prove our main results Theorem 1.5 and Theorem 1.11.

## 6.1 Solutions of the bulk deformed CMS operators

In this section we discuss solutions of the bulk deformed CMS system (5.3)–(5.4). These are second-order linear differential equations in the $n$ variables $t_1, \ldots, t_n$ and the $m$ variables $s_1, \ldots, s_m$. We construct sectorial solutions indexed by sign patterns

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = (\epsilon_1, \ldots, \epsilon_n, \eta_1, \ldots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m.$$

The solution associated with $(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ is characterized by the leading exponential behavior

$$\exp\left(\mathrm{i}\sum_{i=1}^{n}\epsilon_i t_i + \mathrm{i}\frac{\beta}{2}\sum_{a=1}^{m}\eta_a s_a\right).$$

We first consider the special case $\beta = 2$, where the bulk deformed CMS operators are given explicitly in (5.5). In this case, one can write down the solutions in closed form. The proof is a direct verification and is postponed to Appendix D.

**Theorem 6.1** (Explicit solutions for $\beta = 2$)**.** *For each sign pattern*

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m,$$

*write*

$$(\boldsymbol{t}, \boldsymbol{s}) = (t_1, \dots, t_n, s_1, \dots, s_m).$$

*Define*

$$\begin{aligned} F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s}) = \exp\left(\mathrm{i}\sum_{i=1}^{n}\epsilon_i t_i + \mathrm{i}\sum_{a=1}^{m}\eta_a s_a\right) &\prod_{1 \leqslant i < j \leqslant n}(t_i - t_j)^{(\epsilon_i\epsilon_j - 1)/2} \\ \times \prod_{1 \leqslant a < b \leqslant m}(s_a - s_b)^{(\eta_a\eta_b - 1)/2} &\prod_{i=1}^{n}\prod_{a=1}^{m}(t_i - s_a)^{(1+\epsilon_i\eta_a)/2}. \end{aligned} \tag{6.1}$$

*Then $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ solves the $\beta = 2$ system (5.5). Moreover, the $2^{n+m}$ solutions obtained in this way are linearly independent.*

We now return to general $\beta > 0$. Fix a sign pattern

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m.$$

For $1 \leqslant i < j \leqslant n$, $1 \leqslant a < b \leqslant m$, and $1 \leqslant i \leqslant n$, $1 \leqslant a \leqslant m$, define

$$\alpha_{ij} := \frac{\epsilon_i\epsilon_j - 1}{\beta}, \qquad \gamma_{ab} := \frac{\beta}{4}(\eta_a\eta_b - 1), \qquad \rho_{ia} := \frac{1 + \epsilon_i\eta_a}{2}. \tag{6.2}$$

Equivalently,

$$\alpha_{ij} = \begin{cases} 0, & \epsilon_i = \epsilon_j, \\ -2/\beta, & \epsilon_i \neq \epsilon_j, \end{cases} \qquad \gamma_{ab} = \begin{cases} 0, & \eta_a = \eta_b, \\ -\beta/2, & \eta_a \neq \eta_b, \end{cases} \qquad \rho_{ia} = \begin{cases} 1, & \epsilon_i = \eta_a, \\ 0, & \epsilon_i \neq \eta_a. \end{cases} \tag{6.3}$$

Let

$$\mathfrak{X} := \Big\{(\boldsymbol{t}, \boldsymbol{s}) \in \mathbb{C}^{n+m} : t_i \neq t_j,\ s_a \neq s_b,\ t_i \neq s_a \text{ for all } i \neq j,\ a \neq b\Big\}. \tag{6.4}$$

Fix a simply connected domain $\mathfrak{C} \subset \mathfrak{X}$. On $\mathfrak{C}$, choose branches of

$$\log(t_i - t_j), \qquad \log(s_a - s_b), \qquad \log(t_i - s_a).$$

We define the branch factor

$$\begin{aligned} \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s}) := \exp\left(\mathrm{i}\sum_{i=1}^{n}\epsilon_i t_i + \mathrm{i}\frac{\beta}{2}\sum_{a=1}^{m}\eta_a s_a\right) &\prod_{1 \leqslant i < j \leqslant n}(t_i - t_j)^{\alpha_{ij}} \\ \times \prod_{1 \leqslant a < b \leqslant m}(s_a - s_b)^{\gamma_{ab}} &\prod_{i=1}^{n}\prod_{a=1}^{m}(t_i - s_a)^{\rho_{ia}}. \end{aligned} \tag{6.5}$$

We remark that when $\beta = 2$, the branch factor $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})$ solves the $\beta = 2$ system (5.5).

We also introduce the separation scale

$$R(\boldsymbol{t},\boldsymbol{s}) := \min\Big( \min_{1\leqslant i<j\leqslant n} |t_i - t_j|, \min_{1\leqslant i\leqslant n,\ 1\leqslant a\leqslant m} |t_i - s_a|, \min_{1\leqslant a<b\leqslant m} |s_a - s_b| \Big). \tag{6.6}$$

The following theorem shows that, for general $\beta > 0$, the same phenomenon persists: each sign pattern gives rise to a solution of the bulk deformed CMS system (5.3)–(5.4), whose leading asymptotic factor is the branch factor (6.5). We postpone its proof to Section 7.

**Theorem 6.2** (Bulk local solution space and ordered sectorial solutions)**.** *The following two statements hold.*

(i) *For every point $(\boldsymbol{t}_0,\boldsymbol{s}_0) \in \mathfrak{X}$, the space of germs at $(\boldsymbol{t}_0,\boldsymbol{s}_0)$ of holomorphic solutions of the bulk deformed CMS system (5.3)–(5.4) has dimension $2^{n+m}$.*

(ii) *There exists a sufficiently large constant $D = D(\beta, n, m)$ such that the following holds. Write*

$$(z_1, \ldots, z_{n+m}) := (t_1, \ldots, t_n, s_1, \ldots, s_m),$$

*and consider the unbounded ordered domain*

$$\mathfrak{C} := \{(\boldsymbol{t},\boldsymbol{s}) \in \mathfrak{X} : \mathrm{Re}[z_{\ell+1}] - \mathrm{Re}[z_\ell] > D, \quad 1 \leqslant \ell < n+m\}\,. \tag{6.7}$$

*On $\mathfrak{C}$, choose branches of the logarithms entering $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$. For each sign pattern $(\boldsymbol{\epsilon},\boldsymbol{\eta}) \in \{\pm1\}^n \times \{\pm1\}^m$, there exists a holomorphic solution $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ of (5.3)–(5.4) on $\mathfrak{C}$ of the form*

$$F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}). \tag{6.8}$$

*The correction factor satisfies*

$$W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = 1 + \mathrm{O}\big(R(\boldsymbol{t},\boldsymbol{s})^{-1}\big), \tag{6.9}$$

*and, for subsets $I \subseteq \{1, \ldots, n\}$ and $A \subseteq \{1, \ldots, m\}$,*

$$\prod_{i\in I} \partial_{t_i} \prod_{a\in A} \partial_{s_a} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = \mathrm{O}\big(R(\boldsymbol{t},\boldsymbol{s})^{-|I|-|A|-1}\big). \tag{6.10}$$

*The estimates are uniform on $\mathfrak{C}$, and the implicit constants depend only on $\beta, n, m$. Moreover, the $2^{n+m}$ solutions obtained in this way are linearly independent. The same statement holds if the inequalities defining the ordered domain are imposed in any other fixed order of the $t$- and $s$-coordinates.*

## 6.2 Proof of Theorem 1.5

Our main result, Theorem 1.5, follows from the proposition below. The proposition states that the classification of the bulk deformed CMS system (5.3)–(5.4) in Theorem 6.2, together with the asymptotics of the exponential observables from Proposition 4.1 and Corollary 4.2, uniquely determines the exponential observable.

In the remainder of this section, we first prove Theorem 1.5, using Proposition 6.3 as an input. We then prove Proposition 6.3 in Section 6.3.

**Proposition 6.3.** *Adopt Assumption 1.4, and assume*

$$\frac{\beta}{2} = \frac{p}{q} \in \mathbb{Q}_{>0}, \qquad n = kp, \qquad m = kq.$$

*Define*

$$F(t_1, \ldots, t_n; s_1, \ldots, s_m) := \mathbb{E}\left[ \mathrm{P.V.} \prod_{j\in\mathbb{Z}} \frac{\prod_{i=1}^n (t_i - x_j)}{\prod_{a=1}^m (s_a - x_j)^{\beta/2}} \right].$$

*Fix $\boldsymbol{\eta} = (\eta_1, \dots, \eta_m) \in \{\pm 1\}^m$ such that*

$$\#\{a : \eta_a = +1\} = k_+ q, \qquad \#\{a : \eta_a = -1\} = k_- q, \qquad k_+ + k_- = k.$$

*Let $D$ be as in Theorem 6.2, write $(z_1, \dots, z_{n+m}) = (t_1, \dots, t_n, s_1, \dots, s_m)$, and consider the unbounded domain*

$$\mathfrak{C} := \{(\boldsymbol{t}, \boldsymbol{s}) \in \mathfrak{X} : \mathrm{Re}[z_{\ell+1}] - \mathrm{Re}[z_\ell] > D,\ 1 \leqslant \ell < n+m, \quad s_a \in \mathbb{C}_{\eta_a},\ 1 \leqslant a \leqslant m\}.$$

*Recall the sectorial solutions from (6.8), then, on $\mathfrak{C}$,*

$$F = \sum_{\boldsymbol{\epsilon} \in \mathcal{E}_{\boldsymbol{\eta}}} F_{\boldsymbol{\epsilon}, \boldsymbol{\eta}}, \qquad \mathcal{E}_{\boldsymbol{\eta}} := \left\{ \boldsymbol{\epsilon} \in \{\pm 1\}^n : \sum_{i=1}^n \epsilon_i + \frac{\beta}{2} \sum_{a=1}^m \eta_a = 0 \right\}.$$

*Equivalently,*

$$\#\{i : \epsilon_i = +1\} = k_- p, \qquad \#\{i : \epsilon_i = -1\} = k_+ p.$$

*Proof of Theorem 1.5.* By Theorem 3.1, we already know that the $\mathrm{Sine}_\beta$ point process satisfies the microscopic bulk loop equations (1.2). It remains to prove uniqueness: any random point process satisfying (1.2) and the assumptions of the Theorem 1.5 must have the same law.

By the representation (3.7),

$$s(w) = \mathrm{P.\,V.} \sum_{j=-\infty}^{\infty} \frac{1}{x_j - w}.$$

Thus the random point process $\{\cdots \leqslant x_{-2} \leqslant x_{-1} \leqslant x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots\}$ is characterized by the resolvent correlation functions

$$\mathbb{E}\big[s(w_1) s(w_2) \cdots s(w_k)\big], \qquad w_1, \dots, w_k \in \mathbb{C}_+ \cup \mathbb{C}_-. \tag{6.11}$$

We next show that the correlation functions in (6.11) can be recovered from the exponential observables. Consequently, once these observables are uniquely determined, the correlation functions are uniquely determined as well. Let $\tau_\alpha \in \{\pm 1\}$ be such that

$$w_\alpha \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant \alpha \leqslant k.$$

Consider

$$F = \mathbb{E}\left[\mathrm{P.\,V.} \prod_{j \in \mathbb{Z}} \prod_{\alpha=1}^k \frac{\prod_{i=1}^p (t_i^{(\alpha)} - x_j)}{\prod_{a=1}^q (s_a^{(\alpha)} - x_j)^{\beta/2}}\right],$$

where

$$t_i^{(\alpha)},\ s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant i \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k.$$

Since $\beta/2 = p/q$, the observable equals 1 after the diagonal specialization

$$w_\alpha = t_1^{(\alpha)} = \cdots = t_p^{(\alpha)} = s_1^{(\alpha)} = \cdots = s_q^{(\alpha)}, \qquad 1 \leqslant \alpha \leqslant k. \tag{6.12}$$

Differentiating before this specialization gives

$$\partial_{t_1^{(\alpha)}} \log \left[\mathrm{P.\,V.} \prod_{j \in \mathbb{Z}} \frac{\prod_{i=1}^p (t_i^{(\alpha)} - x_j)}{\prod_{a=1}^q (s_a^{(\alpha)} - x_j)^{\beta/2}}\right] = \mathrm{P.\,V.} \sum_{j \in \mathbb{Z}} \frac{1}{t_1^{(\alpha)} - x_j} = -s(t_1^{(\alpha)}).$$

Therefore, after applying $\partial_{t_1^{(1)}} \cdots \partial_{t_1^{(k)}}$ and then imposing (6.12), all undifferentiated product factors become 1, and we obtain

$$\mathbb{E}\left[s(w_1) \cdots s(w_k)\right] = (-1)^k\, \partial_{t_1^{(1)}} \cdots \partial_{t_1^{(k)}} F \Big|_{w_\alpha = t_1^{(\alpha)} = \cdots = t_p^{(\alpha)} = s_1^{(\alpha)} = \cdots = s_q^{(\alpha)}, 1 \leqslant \alpha \leqslant k.}.$$

By Proposition 6.3, the observable $F$ is uniquely determined on the nonempty ordered domain $\mathfrak{C}$ appearing there. Moreover, by Corollary 4.2, $F$ is holomorphic in the numerator variables $t_i^{(\alpha)} \in \mathbb{C}$ and holomorphic in the denominator variables $s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}$. Hence, by the identity theorem for holomorphic functions, this uniquely determines $F$ on the whole domain

$$(t_i^{(\alpha)}) \in \mathbb{C}, \qquad (s_a^{(\alpha)}) \in \mathbb{C}_{\tau_\alpha}, \quad 1 \leqslant i \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k.$$

In particular, the derivatives at the specialization (6.12) are unique. Thus all correlation functions (6.11) are uniquely determined. This proves Theorem 1.5. □

### 6.3 Proof of Proposition 6.3

*Proof of Proposition 6.3.* Choose real numbers

$$\xi_1 < \cdots < \xi_n < \zeta_1 < \cdots < \zeta_m$$

whose successive gaps are larger than $2D$. We use these same real parts throughout the proof. Since the constants in (6.9)–(6.10) are universal, we may increase $D$ so that

$$|W_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} - 1| \leqslant \frac{1}{4}$$

on the ordered domain, uniformly over all sign patterns. In particular, all correction factors are bounded above and away from zero.

The observable $F$ is a holomorphic solution of (5.3)–(5.4) on $\mathfrak{C}$. By Theorem 6.2, the sectorial solutions form a basis, and hence

$$F = \sum_{(\boldsymbol{\epsilon},\boldsymbol{\eta}') \in \{\pm 1\}^{n+m}} c_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} F_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}, \qquad c_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} \in \mathbb{C}. \tag{6.13}$$

All configurations below remain in the same unbounded domain $\mathfrak{C}$, so the solutions and the coefficients in this expansion do not change during the three comparisons.

*Step 1: Forcing $\boldsymbol{\eta}' = \boldsymbol{\eta}$.*

For a large parameter $\Lambda > 0$, set

$$t_r := \xi_r + \mathrm{i}\Lambda^r, \quad 1 \leqslant r \leqslant n, \qquad s_a := \zeta_a + \eta_a \mathrm{i}\Lambda^{n+a}, \quad 1 \leqslant a \leqslant m.$$

For all sufficiently large $\Lambda$, these variables are pairwise distinct and $R(\boldsymbol{t}, \boldsymbol{s}) \gtrsim \Lambda$.

By Corollary 4.2, applied to the configuration $(\boldsymbol{t}, \boldsymbol{s})$ above, we have

$$|F(\boldsymbol{t}, \boldsymbol{s})| \leqslant \exp\left( \sum_{r=1}^{n} \Lambda^r - \frac{\beta}{2} \sum_{a=1}^{m} \Lambda^{n+a} + \mathrm{O}\big((\log \Lambda)^2\big) \right). \tag{6.14}$$

On the other hand, by the explicit formula for $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}$ and the estimate

$$W_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}(\boldsymbol{t}, \boldsymbol{s}) = 1 + \mathrm{O}\big(R(\boldsymbol{t}, \boldsymbol{s})^{-1}\big),$$

each branch satisfies

$$F_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}(\boldsymbol{t}, \boldsymbol{s}) = \exp\left( -\sum_{r=1}^{n} \epsilon_r \Lambda^r - \frac{\beta}{2} \sum_{a=1}^{m} \eta_a \eta'_a \Lambda^{n+a} \right) \Lambda^{\nu_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}} C_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} \left(1 + \mathrm{O}(\Lambda^{-1})\right), \tag{6.15}$$

for some exponent $\nu_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} \in \mathbb{R}$ and some nonzero coefficient $C_{\boldsymbol{\epsilon},\boldsymbol{\eta}'}$ depending on $(\boldsymbol{\xi}, \boldsymbol{\zeta})$.

Suppose that $c_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} \neq 0$ for some $\boldsymbol{\eta}' \neq \boldsymbol{\eta}$, and let $a_0$ be the largest index such that $\eta'_{a_0} \neq \eta_{a_0}$. Then

$$-\frac{\beta}{2}\sum_{a=1}^{m}\eta_a\eta'_a\Lambda^{n+a} = -\frac{\beta}{2}\sum_{a=1}^{m}\Lambda^{n+a} + \beta\sum_{\{a:\eta'_a\neq\eta_a\}}\Lambda^{n+a} = -\frac{\beta}{2}\sum_{a=1}^{m}\Lambda^{n+a} + \beta\Lambda^{n+a_0} + \mathrm{O}(\Lambda^{n+a_0-1}). \tag{6.16}$$

The additional term $\beta\Lambda^{n+a_0}$ dominates all contributions involving the $t$-variables, which are of order at most $\Lambda^n$, as well as the error $\mathrm{O}((\log\Lambda)^2)$ in (6.14).

Moreover, because all variables have been placed on separated scales, distinct sign patterns $(\boldsymbol{\epsilon},\boldsymbol{\eta}')$ give distinct exponential weights in $\Lambda$. Thus, among all terms with nonzero coefficients and $\boldsymbol{\eta}' \neq \boldsymbol{\eta}$, there is a unique lexicographically dominant exponential weight, which cannot be cancelled by the remaining terms. This contradicts the upper bound (6.14). Hence

$$c_{\boldsymbol{\epsilon},\boldsymbol{\eta}'} = 0 \qquad \text{whenever } \boldsymbol{\eta}' \neq \boldsymbol{\eta}.$$

Therefore,

$$F = \sum_{\boldsymbol{\epsilon}\in\{\pm1\}^n} c_{\boldsymbol{\epsilon}} F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}, \qquad c_{\boldsymbol{\epsilon}} := c_{\boldsymbol{\epsilon},\boldsymbol{\eta}}. \tag{6.17}$$

*Step 2: only balanced sign patterns survive.*

Let

$$u := \#\{a : \eta_a = +1\} = k_+q, \qquad m-u = \#\{a : \eta_a = -1\} = k_-q.$$

Fix $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_n) \in \{\pm1\}^n$, and let

$$v := \#\{i : \epsilon_i = -1\}.$$

Take large $\Lambda > 0$ and set

$$t_r := \xi_r - \mathrm{i}\epsilon_r\Lambda, \quad 1 \leqslant r \leqslant n, \quad s_a := \zeta_a + \mathrm{i}\eta_a\Lambda, \quad 1 \leqslant a \leqslant m. \tag{6.18}$$

Then by Corollary 4.2, with

$$M = \sum_{r=1}^{n}\Lambda - \frac{\beta}{2}\sum_{a=1}^{m}\Lambda = 0, \quad L = \sum_{r=1}^{n}\sum_{a=1}^{m}\log\left(1 + \frac{|\xi_r - \zeta_a|}{\Lambda}\right) = \mathrm{O}(1),$$

we conclude that

$$|F(\boldsymbol{t},\boldsymbol{s})| \leqslant e^{M+C(L+L^2)} \lesssim 1, \tag{6.19}$$

provided we take $\Lambda$ large enough.

With the choice of parameters in (6.18), (6.5) gives

$$\begin{aligned}|\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})| &= \exp\left(\Lambda\sum_{i=1}^{n}\epsilon_i^2 - \Lambda\frac{\beta}{2}\sum_{a=1}^{m}\eta_a^2\right)\prod_{1\leqslant i<j\leqslant n}|t_i-t_j|^{\alpha_{ij}}\prod_{1\leqslant a<b\leqslant m}|s_a-s_b|^{\gamma_{ab}}\prod_{i=1}^{n}\prod_{a=1}^{m}|t_i-s_a|^{\rho_{ia}}\\ &= \prod_{1\leqslant i<j\leqslant n}|t_i-t_j|^{\alpha_{ij}}\prod_{1\leqslant a<b\leqslant m}|s_a-s_b|^{\gamma_{ab}}\prod_{i=1}^{n}\prod_{a=1}^{m}|t_i-s_a|^{\rho_{ia}}.\end{aligned}$$

We recall that $v = \#\{i : \epsilon_i = -1\}$ and $u = \#\{a : \eta_a = +1\}$, and recall $\alpha_{ij}$, $\gamma_{ab}$ and $\rho_{ia}$ from (6.3). By a direct count of the algebraic powers of $\Lambda$, we can further simplify the above expression as

$$\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = \Lambda^{\kappa}C_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{\xi},\boldsymbol{\zeta})(1+\mathrm{O}(\Lambda^{-1}), \quad F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})(1+\mathrm{O}(R(\boldsymbol{t},\boldsymbol{s})^{-1})), \tag{6.20}$$

where $|C_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{\xi},\boldsymbol{\zeta})| \asymp 1$, and

$$\begin{aligned}\kappa &= -\frac{2}{\beta}v(n-v) - \frac{\beta}{2}u(m-u) + v(m-u) + u(n-v)\\ &= \frac{q}{p}v^2 + \frac{p}{q}u^2 - 2uv = \frac{q}{p}\left(v - \frac{pu}{q}\right)^2 \geqslant 0.\end{aligned}$$

Thus the exponent $\kappa = 0$ if and only if

$$v = \frac{pu}{q} = pk_+ = \frac{\beta u}{2}.$$

Since

$$\sum_{i=1}^{n} \epsilon_i = n - 2v, \qquad \sum_{a=1}^{m} \eta_a = 2u - m,$$

this is equivalent to

$$\sum_{i=1}^{n} \epsilon_i + \frac{\beta}{2}\sum_{a=1}^{m} \eta_a = 0. \tag{6.21}$$

If (6.21) fails, then $\kappa > 0$, so the branch $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})$ in (6.20) grows like a positive power of $\Lambda$. For generic choices of $(\boldsymbol{\xi},\boldsymbol{\zeta})$, the leading coefficients $C_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{\xi},\boldsymbol{\zeta})$ corresponding to distinct $\boldsymbol{\epsilon}$ are linearly independent, since they come from distinct monomials in the explicit factor $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$. Because $F(\boldsymbol{t},\boldsymbol{s})$ in (6.19) remains bounded, this forces

$$c_{\boldsymbol{\epsilon}} = 0 \qquad \text{unless} \qquad \sum_{i=1}^{n} \epsilon_i + \frac{\beta}{2}\sum_{a=1}^{m} \eta_a = 0.$$

Hence

$$F = \sum_{\boldsymbol{\epsilon}\in\mathcal{E}_{\boldsymbol{\eta}}} c_{\boldsymbol{\epsilon}}\, F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}, \qquad \mathcal{E}_{\boldsymbol{\eta}} := \left\{\boldsymbol{\epsilon} \in \{\pm 1\}^n : \sum_{i=1}^{n} \epsilon_i + \frac{\beta}{2}\sum_{a=1}^{m} \eta_a = 0\right\}. \tag{6.22}$$

*Step 3: for every admissible $\boldsymbol{\epsilon}$, one has $c_{\boldsymbol{\epsilon}} = 1$.* Fix $\boldsymbol{\epsilon} \in \mathcal{E}_{\boldsymbol{\eta}}$. Then there exist partitions

$$\{1,\dots,n\} = I_1 \sqcup \cdots \sqcup I_k, \qquad \{1,\dots,m\} = A_1 \sqcup \cdots \sqcup A_k,$$

with $|I_\alpha| = p$ and $|A_\alpha| = q$ for every $1 \leqslant \alpha \leqslant k$, such that for each block there is a sign $\tau_\alpha \in \{\pm 1\}$ satisfying

$$\epsilon_{I_\alpha(r)} = -\tau_\alpha, \qquad \eta_{A_\alpha(a)} = \tau_\alpha, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q.$$

Here $I_\alpha(r)$ denotes the $r$-th element of $I_\alpha$, and $A_\alpha(a)$ denotes the $a$-th element of $A_\alpha$.

After relabelling the variables according to these blocks, write

$$\epsilon_r^{(\alpha)} := \epsilon_{I_\alpha(r)}, \qquad t_r^{(\alpha)} := t_{I_\alpha(r)}, \qquad \eta_a^{(\alpha)} := \eta_{A_\alpha(a)}, \qquad s_a^{(\alpha)} := s_{A_\alpha(a)}.$$

Thus

$$\epsilon_r^{(\alpha)} = -\tau_\alpha, \quad \eta_a^{(\alpha)} = \tau_\alpha, \quad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k. \tag{6.23}$$

We then specialize the variables in each block by setting

$$t_r^{(\alpha)} := \Lambda\xi_{I_\alpha(r)} + \tau_\alpha \mathrm{i}(\Lambda^4 + \Lambda), \qquad s_a^{(\alpha)} := \Lambda\zeta_{A_\alpha(a)} + \tau_\alpha \mathrm{i}(\Lambda^4 - \Lambda). \tag{6.24}$$

For all sufficiently large $\Lambda$, this configuration lies in the same domain $\mathfrak{C}$, and $R(\boldsymbol{t},\boldsymbol{s}) \asymp \Lambda$. Moreover, the variables in each block lie in $\mathbb{C}_{\tau_\alpha}$.

Then (4.6), together with

$$
\begin{aligned}
M &= -\mathrm{i} \sum_{\alpha=1}^{k} \tau_\alpha \left( \sum_{r=1}^{p} t_r^{(\alpha)} - \frac{\beta}{2} \sum_{a=1}^{q} s_a^{(\alpha)} \right) \\
&= -\mathrm{i} \sum_{\alpha=1}^{k} \tau_\alpha \left[ \Lambda \left( \sum_{r=1}^{p} \xi_{I_\alpha(r)} - \frac{\beta}{2} \sum_{a=1}^{q} \zeta_{A_\alpha(a)} \right) + 2p\tau_\alpha \mathrm{i}\Lambda \right] \\
&= 2pk\Lambda - \mathrm{i}\Lambda \sum_{\alpha=1}^{k} \tau_\alpha \left( \sum_{r=1}^{p} \xi_{I_\alpha(r)} - \frac{\beta}{2} \sum_{a=1}^{q} \zeta_{A_\alpha(a)} \right),
\end{aligned}
\tag{6.25}
$$

and

$$
L = \sum_{\alpha=1}^{k} \sum_{r=1}^{p} \sum_{a=1}^{q} \log\left( 1 + \frac{|t_r^{(\alpha)} - s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}} \right) = \mathrm{O}\left( \frac{\Lambda}{\Lambda^4} \right) = \mathrm{O}(\Lambda^{-3}),
$$

yields

$$
F(\boldsymbol{t}, \boldsymbol{s}) = e^{2pk\Lambda} \exp\left( -\mathrm{i}\Lambda \sum_{\alpha=1}^{k} \tau_\alpha \left( \sum_{r=1}^{p} \xi_{I_\alpha(r)} - \frac{\beta}{2} \sum_{a=1}^{q} \zeta_{A_\alpha(a)} \right) \right) \big(1 + \mathrm{O}(\Lambda^{-1})\big).
\tag{6.26}
$$

For the distinguished branch $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$, the formula (6.5) gives

$$
F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} \big(1 + \mathrm{O}(\Lambda^{-1})\big),
\tag{6.27}
$$

because $R(\boldsymbol{t}, \boldsymbol{s}) \asymp \Lambda$ and $W_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = 1 + \mathrm{O}(R(\boldsymbol{t}, \boldsymbol{s})^{-1})$. By plugging in (6.23) and (6.24), the exponential part of $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ is

$$
\exp\left( \mathrm{i} \sum_{\alpha=1}^{k} \sum_{r=1}^{p} \epsilon_r^{(\alpha)} t_r^{(\alpha)} + \mathrm{i}\frac{\beta}{2} \sum_{\alpha=1}^{k} \sum_{a=1}^{q} \eta_a^{(\alpha)} s_a^{(\alpha)} \right) = \exp\left[ -\mathrm{i} \sum_{\alpha=1}^{k} \tau_\alpha \left( \sum_{r=1}^{p} t_r^{(\alpha)} - \frac{\beta}{2} \sum_{a=1}^{q} s_a^{(\alpha)} \right) \right].
$$

By the computation of $M$ in (6.25), this equals

$$
\exp\left[ 2pk\Lambda - \mathrm{i}\Lambda \sum_{\alpha=1}^{k} \tau_\alpha \left( \sum_{r=1}^{p} \xi_{I_\alpha(r)} - \frac{\beta}{2} \sum_{a=1}^{q} \zeta_{A_\alpha(a)} \right) \right].
\tag{6.28}
$$

Next we show that the algebraic part of $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ contributes only a

$$
1 + \mathrm{O}(\Lambda^{-1})
\tag{6.29}
$$

factor. Within a fixed block, or for two different blocks $\alpha \neq \beta$ such that $\tau_\alpha = \tau_\beta$, all $t$-variables have the same $\epsilon$-sign, all $s$-variables have the same $\eta$-sign, and the $\epsilon$-signs and $\eta$-signs are opposite. Therefore

$$
\alpha_{ij} = 0, \qquad \gamma_{ab} = 0, \qquad \rho_{ia} = 0
$$

for all such pairs of variables. Thus these pairs give no algebraic contribution.

For two different blocks $\alpha \neq \beta$ with $\tau_\alpha \neq \tau_\beta$, all cross-block differences are of order $\Lambda^4$, while their lower-order corrections are of size $\mathrm{O}(\Lambda)$. Hence, after factoring out the corresponding powers of $\Lambda^4$, their product is a fixed nonzero constant times $1 + \mathrm{O}(\Lambda^{-3})$. Moreover, the total power of $\Lambda^4$ cancels. Indeed, the cross-block exponents are

$$
-\frac{2}{\beta} \cdot p^2, \qquad -\frac{\beta}{2} \cdot q^2, \qquad 1 \cdot 2pq,
$$

and their sum is

$$
-\frac{2p^2}{\beta} - \frac{\beta q^2}{2} + 2pq = -pq - pq + 2pq = 0,
$$

where we used $\beta/2 = p/q$. Thus the algebraic part is a fixed nonzero constant times $1 + \mathrm{O}(\Lambda^{-3})$. By the multiplicative normalization of $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ fixed above, this constant is equal to 1. Therefore the algebraic part is $1 + \mathrm{O}(\Lambda^{-1})$.

Combining (6.27), the exponential part (6.28), and the algebraic estimate (6.29), we obtain

$$F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = e^{2pk\Lambda} \exp\left(-\mathrm{i}\Lambda \sum_{\alpha=1}^{k} \tau_\alpha \left(\sum_{r=1}^{p} \xi_{I_\alpha(r)} - \frac{\beta}{2} \sum_{a=1}^{q} \zeta_{A_\alpha(a)}\right)\right) \left(1 + \mathrm{O}(\Lambda^{-1})\right). \tag{6.30}$$

Now let $\boldsymbol{\epsilon}' \in \mathcal{E}_{\boldsymbol{\eta}}$ with $\boldsymbol{\epsilon}' \neq \boldsymbol{\epsilon}$. In the following, we estimate $F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})$. For each block $\alpha$, define the number of mismatches

$$m_\alpha(\boldsymbol{\epsilon}') := \#\left\{1 \leqslant r \leqslant p : \epsilon'_{I_\alpha(r)} \neq -\tau_\alpha\right\}.$$

Since $\boldsymbol{\epsilon}' \neq \boldsymbol{\epsilon}$, we have

$$m(\boldsymbol{\epsilon}') := \sum_{\alpha=1}^{k} m_\alpha(\boldsymbol{\epsilon}') \geqslant 1.$$

We compare the exponential part of $\Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}$ with that of the distinguished branch. Using (6.24), the modulus of the exponential contribution from the $\alpha$-th block is

$$\begin{aligned}
&\left|\exp\left(\mathrm{i}\sum_{r=1}^{p} \epsilon'_{I_\alpha(r)} t_r^{(\alpha)} + \mathrm{i}\frac{\beta}{2}\sum_{a=1}^{q} \eta_a^{(\alpha)} s_a^{(\alpha)}\right)\right| \\
&\qquad = \exp\left(-\tau_\alpha(\Lambda^4 + \Lambda)\sum_{r=1}^{p} \epsilon'_{I_\alpha(r)} - \frac{\beta}{2} q(\Lambda^4 - \Lambda)\right) \\
&\qquad = \exp\left((p - 2m_\alpha(\boldsymbol{\epsilon}'))(\Lambda^4 + \Lambda) - p(\Lambda^4 - \Lambda)\right) \\
&\qquad = \exp\left(2p\Lambda - 2m_\alpha(\boldsymbol{\epsilon}')(\Lambda^4 + \Lambda)\right),
\end{aligned}$$

where we used $\beta q/2 = p$. Multiplying over $\alpha = 1, \ldots, k$, this gives

$$|\text{exponential part of } \Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}| = e^{2pk\Lambda} e^{-2m(\boldsymbol{\epsilon}')(\Lambda^4+\Lambda)} \leqslant e^{2pk\Lambda} e^{-2(\Lambda^4+\Lambda)}. \tag{6.31}$$

It remains to bound the algebraic part. Between different blocks, the separations are at most of order $\Lambda^4$, while all nonzero separations are bounded below by a constant multiple of $\Lambda$. Hence all algebraic factors in $\Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}$ are bounded by $C\Lambda^4$, and there are at most $(n+m)^2$ such factors. Therefore

$$|\text{algebraic part of } \Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}| \leqslant (C\Lambda)^{4(n+m)^2}. \tag{6.32}$$

Since $W_{\boldsymbol{\epsilon}',\boldsymbol{\eta}} = 1 + \mathrm{O}(\Lambda^{-1})$, combining (6.31) and (6.32) gives

$$|F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s})| \leqslant e^{2pk\Lambda} e^{-2(\Lambda^4+\Lambda)} (C\Lambda)^{4(n+m)^2} = \mathrm{o}(e^{2pk\Lambda}), \tag{6.33}$$

provided $\Lambda$ is sufficiently large.

Substituting the special configuration above into (6.22), and using (6.26), (6.30), and (6.33), we obtain

$$1 + \mathrm{o}(1) = c_{\boldsymbol{\epsilon}}(1 + \mathrm{o}(1)) + \mathrm{o}(1).$$

Hence $c_{\boldsymbol{\epsilon}} = 1$. Since $\boldsymbol{\epsilon} \in \mathcal{E}_{\boldsymbol{\eta}}$ was arbitrary, we conclude that

$$c_{\boldsymbol{\epsilon}} = 1 \qquad \text{for every } \boldsymbol{\epsilon} \in \mathcal{E}_{\boldsymbol{\eta}}.$$

Combining this with (6.22), we have proved on $\mathfrak{C}$ that

$$F = \sum_{\boldsymbol{\epsilon} \in \mathcal{E}_{\boldsymbol{\eta}}} F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}.$$

□

### 6.4 Solutions of the edge deformed CMS operators

In this section, we study solutions of the edge deformed CMS system (5.18)–(5.19). As in the bulk case, we construct sectorial solutions indexed by sign patterns

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m.$$

For $z \in \mathbb{C}_\pm$, set

$$\phi_-(z) := \mathrm{Ai}(z), \qquad \phi_+(z) := \begin{cases} \mathrm{i}\omega \, \mathrm{Ai}(\omega z), & z \in \mathbb{C}_+, \\ -\mathrm{i}\omega^2 \, \mathrm{Ai}(\omega^2 z) & z \in \mathbb{C}_-. \end{cases} \tag{6.34}$$

Both functions solve the Airy equation

$$\phi''(z) - z\phi(z) = 0.$$

Moreover, by (A.5), (A.6), and (A.7), they satisfy the asymptotics

$$\phi_\pm(z) \sim \frac{1}{2\sqrt{\pi}} z^{-1/4} \exp\left( \pm \frac{2}{3} z^{3/2} \right), \qquad z \to \infty, \quad |\arg z| < \pi - \delta. \tag{6.35}$$

We refer to Appendix A for further background on Airy functions. We remark that $\phi_+$ in (6.34) is constructed such that the asymptotics (6.35) hold.

For the $s$-variables, we introduce the rescaled Airy functions

$$\psi_\pm(s) := \phi_\pm\left( \left( \frac{\beta}{2} \right)^{2/3} s \right). \tag{6.36}$$

Then $\psi_\pm$ solve

$$\psi_\pm''(s) - \frac{\beta^2}{4} s \, \psi_\pm(s) = 0.$$

The solution associated with a sign pattern $(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ is characterized by the leading Airy behavior

$$\left( \prod_{i=1}^n \phi_{\epsilon_i}(t_i) \right) \left( \prod_{a=1}^m \psi_{\eta_a}(s_a) \right).$$

We first consider the special case $\beta = 2$, where the edge deformed CMS operators are given explicitly in (5.20). In this case, one can write down the solutions in closed form. The proof is a direct verification and is postponed to Appendix E.

**Theorem 6.4** (Explicit solutions for $\beta = 2$)**.** *For each sign pattern*

$$\boldsymbol{\sigma} = (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^{n+m},$$

*write*

$$\boldsymbol{x} = (x_1, \dots, x_{n+m}) := (\boldsymbol{t}, \boldsymbol{s}) = (t_1, \dots, t_n, s_1, \dots, s_m).$$

*Define*

$$D_{\boldsymbol{\sigma}}(\boldsymbol{x}) := \det\left[ \partial_x^{j-1} \phi_{\sigma_r}(x_r) \right]_{1 \leqslant r, j \leqslant n+m}.$$

*Let*

$$\Delta(\boldsymbol{t}) := \prod_{1 \leqslant i < j \leqslant n} (t_i - t_j), \qquad \Delta(\boldsymbol{s}) := \prod_{1 \leqslant a < b \leqslant m} (s_a - s_b),$$

*and set*

$$F_{\boldsymbol{\sigma}}(\boldsymbol{t}, \boldsymbol{s}) := \frac{D_{\boldsymbol{\sigma}}(\boldsymbol{x})}{\Delta(\boldsymbol{t}) \Delta(\boldsymbol{s})}. \tag{6.37}$$

*Then $F_{\boldsymbol{\sigma}}$ solves the $\beta = 2$ edge system (5.20). Moreover, the $2^{n+m}$ solutions obtained in this way are linearly independent.*

We now return to general $\beta > 0$. Fix a sign pattern

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^n \times \{\pm 1\}^m.$$

Define the configuration space

$$\mathsf{X} := \Big\{(\boldsymbol{t}, \boldsymbol{s}) \in (\mathbb{C}_+ \cup \mathbb{C}_-)^{n+m} : t_i, s_a \neq 0,\ t_i \neq t_j,\ s_a \neq s_b,\ t_i \neq s_a,\ |\arg t_i|, |\arg s_a| < \pi\Big\}. \tag{6.38}$$

Fix a simply connected domain $\mathsf{C} \subset \mathsf{X}$. Since $|\arg t_i|, |\arg s_a| < \pi - \delta$, the principal square-root branches $\sqrt{t_i}, \sqrt{s_a}$ are well defined on $\mathsf{C}$. Then, for all choices of signs, the functions

$$\pm\sqrt{t_i} \pm \sqrt{t_j}, \qquad \pm\sqrt{s_a} \pm \sqrt{s_b}, \qquad \pm\sqrt{t_i} \pm \sqrt{s_a}$$

are holomorphic and nonvanishing on $\mathsf{C}$. Hence, we may also choose branches of

$$\log(\pm\sqrt{t_i} \pm \sqrt{t_j}), \qquad \log(\pm\sqrt{s_a} \pm \sqrt{s_b}), \qquad \log(\pm\sqrt{t_i} \pm \sqrt{s_a}).$$

We then define the branch factor

$$\begin{aligned}\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s}) := \left(\prod_{i=1}^{n} \phi_{\epsilon_i}(t_i)\right)\left(\prod_{a=1}^{m} \psi_{\eta_a}(s_a)\right) \prod_{1 \leqslant i<j \leqslant n} (\epsilon_i\sqrt{t_i} + \epsilon_j\sqrt{t_j})^{-2/\beta} \\ \times \prod_{1 \leqslant a<b \leqslant m} (\eta_a\sqrt{s_a} + \eta_b\sqrt{s_b})^{-\beta/2} \prod_{i=1}^{n}\prod_{a=1}^{m} (\epsilon_i\sqrt{t_i} - \eta_a\sqrt{s_a}).\end{aligned} \tag{6.39}$$

We also introduce the separation scale

$$\mathsf{R}(\boldsymbol{t}, \boldsymbol{s}) := \min\Big(\min_{1\leqslant i \leqslant n} |t_i|, \min_{1\leqslant a \leqslant m} |s_a|, \min_{1\leqslant i<j \leqslant n} |t_i - t_j|, \min_{1\leqslant i \leqslant n,\ 1\leqslant a \leqslant m} |t_i - s_a|, \min_{1\leqslant a<b \leqslant m} |s_a - s_b|\Big). \tag{6.40}$$

The following theorem shows that, for general $\beta > 0$, the same phenomenon persists: each sign pattern gives rise to a solution of the edge deformed CMS system (5.18)–(5.19), whose leading asymptotic factor is the branch factor (6.39). We postpone the proof to Section 8.

**Theorem 6.5** (Edge local solution space and sectorial solutions)**.** *Fix $0 < \delta < \pi/2$. The following two statements hold.*

(i) *For every point $(\boldsymbol{t}_0, \boldsymbol{s}_0) \in \mathsf{X}$, the space of germs at $(\boldsymbol{t}_0, \boldsymbol{s}_0)$ of holomorphic solutions of the edge deformed CMS system (5.18)–(5.19) has dimension $2^{n+m}$.*

(ii) *There exists a sufficiently large constant $D = D(\beta, n, m)$ such that the following holds. Write*

$$(z_1, \dots, z_{n+m}) := (t_1, \dots, t_n, s_1, \dots, s_m).$$

*For any $\boldsymbol{\tau} \in \{\pm\}^{n+m}$, consider the unbounded ordered domain*

$$\begin{aligned}\mathsf{C} = \mathsf{C}(\boldsymbol{\tau}, D) := \Big\{\boldsymbol{z} \in \mathbb{C}^{n+m} : \Big|\arg z_j - \frac{\pi}{3}\tau_j\Big| < \frac{\pi}{12}, \quad 1 \leqslant j \leqslant n+m, \\ \mathrm{Re}[z_N] > D, \qquad \mathrm{Re}[z_r] - \mathrm{Re}[z_{r+1}] > D, \quad 1 \leqslant r < n+m\Big\}.\end{aligned} \tag{6.41}$$

*On $\mathsf{C}$, choose branches of the logarithms entering $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$. For each sign pattern $(\boldsymbol{\epsilon}, \boldsymbol{\eta}) \in \{\pm 1\}^n \times \{\pm 1\}^m$, there exists a holomorphic solution $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ of (5.18)–(5.19) on $\mathsf{C}$ of the form*

$$F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}. \tag{6.42}$$

*The correction factor satisfies*

$$W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = 1 + \mathrm{O}\Big(\mathsf{R}(\boldsymbol{t},\boldsymbol{s})^{-3/2}\Big)\,, \tag{6.43}$$

*and, for subsets $I \subseteq \{1,\dots,n\}$ and $A \subseteq \{1,\dots,m\}$,*

$$\partial_{\boldsymbol{t}_I}\partial_{\boldsymbol{s}_A} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t},\boldsymbol{s}) = \mathrm{O}\Big(\mathsf{R}(\boldsymbol{t},\boldsymbol{s})^{-|I|-|A|-3/2}\Big)\,. \tag{6.44}$$

*The estimates are uniform on $\mathsf{C}$, and their implicit constants depend only on $\beta, n, m$. Moreover, the $2^{n+m}$ solutions obtained in this way are linearly independent.*

### 6.5 Proof of Theorem 1.11

Our characterization of Airy point process Theorem 1.11, follows from the proposition below. The proposition states that the classification of the edge deformed CMS system (5.18)–(5.19) in Theorem 6.5, together with the asymptotics of the exponential observables from Proposition 4.3, uniquely determines the exponential observable.

In the remainder of this section, we first prove Theorem 1.11, using Proposition 6.6 as an input. We then prove Proposition 6.6 in Section 6.6.

**Proposition 6.6.** *Adopt Assumption 1.10, and assume*

$$\frac{\beta}{2} = \frac{p}{q} \in \mathbb{Q}_{>0}, \qquad n = kp, \qquad m = kq.$$

*Define*

$$F(t_1,\dots,t_n;s_1,\dots,s_m) := \mathbb{E}\left[\frac{\prod_{r=1}^n e^{c_0 t_r}}{\prod_{a=1}^m e^{(\beta/2)c_0 s_a}} \prod_{j\geqslant 1} \frac{\prod_{r=1}^n (t_r - x_j) e^{t_r/\mathfrak{a}_j}}{\prod_{a=1}^m \big((s_a - x_j)e^{s_a/\mathfrak{a}_j}\big)^{\beta/2}}\right], \tag{6.45}$$

*where $c_0 = \mathrm{Ai}'(0)/\mathrm{Ai}(0)$.*

*Fix sign patterns*

$$\boldsymbol{\epsilon} = (\epsilon_1,\dots,\epsilon_n) \in \{\pm 1\}^n, \qquad \boldsymbol{\eta} = (\eta_1,\dots,\eta_m) \in \{\pm 1\}^m,$$

*satisfying*

$$\#\{r : \epsilon_r = +1\} = k_+ p, \quad \#\{r : \epsilon_r = -1\} = k_- p, \quad \#\{a : \eta_a = +1\} = k_+ q, \quad \#\{a : \eta_a = -1\} = k_- q, \tag{6.46}$$

*where $k_+ + k_- = k$, and set*

$$\boldsymbol{\tau} := (\epsilon_1,\dots,\epsilon_n,\eta_1,\dots,\eta_m).$$

*For $D = D(\beta,n,m) > 0$ sufficiently large, and let $\mathsf{C} = \mathsf{C}(\boldsymbol{\tau}, D)$ be as in (6.41).*

*Recall the sectorial solutions $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ from (6.42). Choose the logarithmic branches entering $\Phi_{-,+}$ according to the conventions used in the proof below. Then, on $\mathsf{C}$,*

$$F = 2^{k(p+q)/2}\left(\frac{\beta}{2}\right)^{kq/6} \pi^{k(p+q)/2} \exp(-\mathrm{i}\pi\vartheta_{p,q}(k_+,k_-))\, F_{-,+}, \tag{6.47}$$

*where*

$$\vartheta_{p,q}(k_+,k_-) := \frac{q}{2}\left[p\big(k_-^2 + 2k_+k_- - k_+^2\big) + k_- - k_+\right]. \tag{6.48}$$

*Here $+$ and $-$ denote the all-plus and all-minus sign patterns, respectively.*

*Proof of Theorem 1.11.* By Theorem 3.2, we already know that the $\mathrm{Airy}_\beta$ point process satisfies the microscopic edge loop equations (1.5). It remains to prove uniqueness: any random point process satisfying (1.5) and the assumptions of Theorem 1.11 must have the same law.

By the representation (3.11),

$$s(w) = \sum_{i=1}^{\infty} \left( \frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i} \right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}.$$

Hence to show the uniqueness of the random point process, it is enough to show that the resolvent correlation functions

$$\mathbb{E}\big[s(w_1)s(w_2)\cdots s(w_k)\big], \qquad w_1, \dots, w_k \in \mathbb{C}_+ \cup \mathbb{C}_-, \tag{6.49}$$

are uniquely determined.

We next show that the correlation functions in (6.49) can be recovered from the edge exponential observables. Consequently, once these observables are uniquely determined, the correlation functions are uniquely determined as well.

Let $\tau_\alpha \in \{\pm 1\}$ be such that

$$w_\alpha \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant \alpha \leqslant k.$$

Consider the edge exponential observable

$$F = \prod_{\alpha=1}^{k} e^{c_0\left(\sum_{r=1}^{p} t_r^{(\alpha)} - \frac{\beta}{2}\sum_{a=1}^{q} s_a^{(\alpha)}\right)} \times \mathbb{E}\left[ \prod_{j\geqslant 1}\prod_{\alpha=1}^{k} \frac{\prod_{r=1}^{p}(t_r^{(\alpha)} - x_j)e^{t_r^{(\alpha)}/\mathfrak{a}_j}}{\prod_{a=1}^{q}\left((s_a^{(\alpha)} - x_j)e^{s_a^{(\alpha)}/\mathfrak{a}_j}\right)^{\beta/2}} \right], \tag{6.50}$$

where

$$t_r^{(\alpha)},\, s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k, \quad c_0 := \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}.$$

As in the bulk case, differentiating the observable before taking the diagonal specialization recovers the resolvent correlations. Therefore,

$$\mathbb{E}\big[s(w_1)s(w_2)\cdots s(w_k)\big] = (-1)^k\, \partial_{t_1^{(1)}} \cdots \partial_{t_1^{(k)}} F\Big|_{w_\alpha = t_1^{(\alpha)} = \cdots = t_p^{(\alpha)} = s_1^{(\alpha)} = \cdots = s_q^{(\alpha)},\ 1 \leqslant \alpha \leqslant k}. \tag{6.51}$$

By Proposition 6.6, the observable $F$ is uniquely determined on the nonempty open domain $\mathsf{C}$ introduced there. Moreover, by Proposition 4.3, $F$ is holomorphic in the variables

$$t_r^{(\alpha)},\, s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \quad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k.$$

Hence, by the identity theorem for holomorphic functions, $F$ is uniquely determined on this domain. In particular, the derivatives at the specialization in (6.51) are unique. Thus all correlation functions (6.49) are uniquely determined. This proves Theorem 1.11. $\square$

## 6.6 Proof of Proposition 6.6

*Proof of Proposition 6.6.* By Theorem 6.5, the sectorial solutions form a basis on $\mathsf{C}$. Hence

$$F = \sum_{(\boldsymbol{\epsilon}', \boldsymbol{\eta}') \in \{\pm 1\}^{n+m}} c_{\boldsymbol{\epsilon}', \boldsymbol{\eta}'} F_{\boldsymbol{\epsilon}', \boldsymbol{\eta}'}, \qquad c_{\boldsymbol{\epsilon}', \boldsymbol{\eta}'} \in \mathbb{C}. \tag{6.52}$$

We will show that all coefficients vanish except possibly the one corresponding to the $F_{-,+}$ branch.

*Step 1: forcing $\boldsymbol{\epsilon}' = -$ and $\boldsymbol{\eta}' = +$.*

By the balance condition (6.46), there exist partitions

$$\{1,\dots,n\} = I_1 \sqcup \cdots \sqcup I_k, \qquad \{1,\dots,m\} = A_1 \sqcup \cdots \sqcup A_k,$$

with $|I_\alpha| = p$ and $|A_\alpha| = q$ for every $1 \leqslant \alpha \leqslant k$, such that each block has a sign $\tau_\alpha \in \{\pm 1\}$ satisfying

$$\epsilon_{I_\alpha(r)} = \tau_\alpha, \qquad \eta_{A_\alpha(a)} = \tau_\alpha, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q.$$

Here $I_\alpha(r)$ denotes the $r$-th element of $I_\alpha$, and $A_\alpha(a)$ denotes the $a$-th element of $A_\alpha$.

After relabelling the variables according to these blocks, write

$$\epsilon_r^{(\alpha)} := \epsilon_{I_\alpha(r)}, \qquad t_r^{(\alpha)} := t_{I_\alpha(r)}, \qquad \eta_a^{(\alpha)} := \eta_{A_\alpha(a)}, \qquad s_a^{(\alpha)} := s_{A_\alpha(a)}.$$

Then

$$\epsilon_r^{(\alpha)} = \eta_a^{(\alpha)} = \tau_\alpha, \qquad t_r^{(\alpha)}, s_a^{(\alpha)} \in \mathbb{C}_{\tau_\alpha}, \qquad 1 \leqslant r \leqslant p, \quad 1 \leqslant a \leqslant q, \quad 1 \leqslant \alpha \leqslant k. \tag{6.53}$$

Choose positive parameters

$$\xi_1 > \xi_2 > \cdots > \xi_n > 0, \qquad 0 < \zeta_1 < \zeta_2 < \cdots < \zeta_m,$$

outside the finitely many proper real hyperplanes on which two distinct sign patterns have the same real exponential rate below. For $\Lambda > 0$, set

$$t_r^{(\alpha)} := \Lambda^2 \left(1 - \frac{\mathrm{i}\tau_\alpha \xi_{I_\alpha(r)}}{\Lambda}\right)^{2/3} e^{\mathrm{i}\tau_\alpha \pi/3}, \qquad s_a^{(\alpha)} := \Lambda^2 \left(1 + \frac{\mathrm{i}\tau_\alpha \zeta_{A_\alpha(a)}}{\Lambda}\right)^{2/3} e^{\mathrm{i}\tau_\alpha \pi/3}. \tag{6.54}$$

All powers are principal powers. Uniformly in the indices,

$$\begin{aligned} t_r^{(\alpha)} &= \Lambda^2 e^{\mathrm{i}\tau_\alpha \pi/3} + \frac{\Lambda}{3}(\sqrt{3} - \mathrm{i}\tau_\alpha)\xi_{I_\alpha(r)} + \mathrm{O}(1), \\ s_a^{(\alpha)} &= \Lambda^2 e^{\mathrm{i}\tau_\alpha \pi/3} + \frac{\Lambda}{3}(-\sqrt{3} + \mathrm{i}\tau_\alpha)\zeta_{A_\alpha(a)} + \mathrm{O}(1). \end{aligned} \tag{6.55}$$

In particular,

$$\mathrm{Re}[t_i] = \frac{\Lambda^2}{2} + \frac{\sqrt{3}}{3}\Lambda\xi_i + \mathrm{O}(1), \qquad \mathrm{Re}[s_a] = \frac{\Lambda^2}{2} - \frac{\sqrt{3}}{3}\Lambda\zeta_a + \mathrm{O}(1).$$

Consequently,

$$\mathrm{Re}[t_1] > \cdots > \mathrm{Re}[t_n] > \mathrm{Re}[s_1] > \cdots > \mathrm{Re}[s_m] > D$$

for every sufficiently large $\Lambda$. Moreover,

$$\arg t_r^{(\alpha)} \longrightarrow \frac{\pi}{3}\tau_\alpha, \qquad \arg s_a^{(\alpha)} \longrightarrow \frac{\pi}{3}\tau_\alpha.$$

Thus $(\boldsymbol{t}(\Lambda), \boldsymbol{s}(\Lambda)) \in \mathsf{C}$ for every sufficiently large $\Lambda$. The strict inequalities among the parameters also imply

$$\mathsf{R}(\boldsymbol{t}(\Lambda), \boldsymbol{s}(\Lambda)) \asymp \Lambda. \tag{6.56}$$

Indeed, two variables in opposite half-planes are separated on the scale $\Lambda^2$, while two variables in the same half-plane are separated on the scale $\Lambda$.

Using the principal branches, we have

$$\big(t_r^{(\alpha)}\big)^{3/2} = \mathrm{i}\tau_\alpha \Lambda^3 + \xi_{I_\alpha(r)}\Lambda^2, \qquad \big(s_a^{(\alpha)}\big)^{3/2} = \mathrm{i}\tau_\alpha \Lambda^3 - \zeta_{A_\alpha(a)}\Lambda^2. \tag{6.57}$$

The quantity $L$ in Proposition 4.3 satisfies

$$L = \sum_{\alpha=1}^{k} \sum_{r=1}^{p} \sum_{a=1}^{q} \log\left(1 + \frac{|t_r^{(\alpha)} - s_a^{(\alpha)}|}{\sqrt{|\operatorname{Im}[t_r^{(\alpha)}]|\,|\operatorname{Im}[s_a^{(\alpha)}]|}}\right) = \mathrm{O}(\Lambda^{-1}),$$

Furthermore, using $p=(\beta/2)q$, the $\mathrm{i}\tau_\alpha\Lambda^3$-terms cancel inside each block, and therefore

$$M=-\frac{2}{3}\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\big(t_r^{(\alpha)}\big)^{3/2}-\frac{\beta}{2}\sum_{a=1}^{q}\big(s_a^{(\alpha)}\big)^{3/2}\right)=-\frac{2}{3}\Lambda^2\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\xi_{I_\alpha(r)}+\frac{\beta}{2}\sum_{a=1}^{q}\zeta_{A_\alpha(a)}\right). \tag{6.58}$$

Hence (4.29) gives

$$F(\boldsymbol{t},\boldsymbol{s})=\exp\left[-\frac{2}{3}\Lambda^2\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\xi_{I_\alpha(r)}+\frac{\beta}{2}\sum_{a=1}^{q}\zeta_{A_\alpha(a)}\right)\right]\big(1+\mathrm{O}(\Lambda^{-1})\big). \tag{6.59}$$

Next we estimate the sectorial branches. Recall that

$$F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}=\Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}W_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'},$$

where the explicit branch factor is given by (6.39). Namely,

$$\begin{aligned}\Phi_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}(t,s)=\left(\prod_{r=1}^{n}\phi_{\epsilon_r'}(t_r)\right)\left(\prod_{a=1}^{m}\psi_{\eta_a'}(s_a)\right)&\prod_{1\leqslant i<j\leqslant n}(\epsilon_i'\sqrt{t_i}+\epsilon_j'\sqrt{t_j})^{-2/\beta}\\ \times\prod_{1\leqslant a<b\leqslant m}(\eta_a'\sqrt{s_a}+\eta_b'\sqrt{s_b})^{-\beta/2}&\prod_{i=1}^{n}\prod_{a=1}^{m}(\epsilon_i'\sqrt{t_i}-\eta_a'\sqrt{s_a}).\end{aligned} \tag{6.60}$$

For $0<|\arg z|<\pi-\delta$, the Airy asymptotics (6.35) imply

$$\phi_+(z)=\frac{1}{2\sqrt{\pi}\,z^{1/4}}e^{\frac{2}{3}z^{3/2}}\left(1+\mathrm{O}_\delta(|z|^{-3/2})\right),\quad \phi_-(z)=\frac{1}{2\sqrt{\pi}\,z^{1/4}}e^{-\frac{2}{3}z^{3/2}}\left(1+\mathrm{O}_\delta(|z|^{-3/2})\right). \tag{6.61}$$

Consequently, for the normalized Airy functions $\psi_\pm$ defined in (6.36), we have

$$\begin{aligned}\psi_+(z)&=\frac{1}{2^{5/6}\beta^{1/6}\sqrt{\pi}\,z^{1/4}}e^{\frac{\beta}{3}z^{3/2}}\left(1+\mathrm{O}_\delta(|z|^{-3/2})\right),\\ \psi_-(z)&=\frac{1}{2^{5/6}\beta^{1/6}\sqrt{\pi}\,z^{1/4}}e^{-\frac{\beta}{3}z^{3/2}}\left(1+\mathrm{O}_\delta(|z|^{-3/2})\right).\end{aligned} \tag{6.62}$$

For the special configuration (6.54) above, (6.57) implies that

$$\begin{aligned}\phi_+\big(t_r^{(\alpha)}\big)&=e^{\frac{2}{3}\mathrm{i}\tau_\alpha\Lambda^3+\frac{2}{3}\xi_{I_\alpha(r)}\Lambda^2+\mathrm{O}(\log\Lambda)},\quad &\phi_-\big(t_r^{(\alpha)}\big)&=e^{-\frac{2}{3}\mathrm{i}\tau_\alpha\Lambda^3-\frac{2}{3}\xi_{I_\alpha(r)}\Lambda^2+\mathrm{O}(\log\Lambda)},\\ \psi_+\big(s_a^{(\alpha)}\big)&=e^{\frac{\beta}{3}\mathrm{i}\tau_\alpha\Lambda^3-\frac{\beta}{3}\zeta_{A_\alpha(a)}\Lambda^2+\mathrm{O}(\log\Lambda)},\quad &\psi_-\big(s_a^{(\alpha)}\big)&=e^{-\frac{\beta}{3}\mathrm{i}\tau_\alpha\Lambda^3+\frac{\beta}{3}\zeta_{A_\alpha(a)}\Lambda^2+\mathrm{O}(\log\Lambda)}.\end{aligned} \tag{6.63}$$

Moreover, every algebraic factor in (6.60) contributes at most $\exp(\mathrm{O}(\log\Lambda))$. Since $\mathsf{R}(\boldsymbol{t},\boldsymbol{s})\asymp\Lambda$, we also have

$$W_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}(\boldsymbol{t},\boldsymbol{s})=1+\mathrm{O}(\mathsf{R}(\boldsymbol{t},\boldsymbol{s})^{-3/2})=1+\mathrm{O}(\Lambda^{-3/2}). \tag{6.64}$$

Combining these estimates, we obtain

$$F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}(\boldsymbol{t},\boldsymbol{s})=\exp\left[\frac{2}{3}\Lambda^2\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\epsilon_r'^{(\alpha)}\xi_{I_\alpha(r)}-\frac{\beta}{2}\sum_{a=1}^{q}\eta_a'^{(\alpha)}\zeta_{A_\alpha(a)}\right)+\mathrm{i}\Lambda^3\Theta_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}+\mathrm{O}(\log\Lambda)\right], \tag{6.65}$$

where

$$\Theta_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}=\frac{2}{3}\sum_{\alpha=1}^{k}\tau_\alpha\left(\sum_{r=1}^{p}\epsilon_r'^{(\alpha)}+\frac{\beta}{2}\sum_{a=1}^{q}\eta_a'^{(\alpha)}\right)\in\mathbb{R}. \tag{6.66}$$

In particular, the term $\mathrm{i}\Lambda^3\Theta_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}$ is purely oscillatory and does not affect the magnitude of $F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}(\boldsymbol{t},\boldsymbol{s})$, and we conclude

$$|F_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'}(\boldsymbol{t},\boldsymbol{s})| = \exp\!\left[\frac{2}{3}\Lambda^2\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\epsilon'_{I_\alpha(r)}\xi_{I_\alpha(r)} - \frac{\beta}{2}\sum_{a=1}^{q}\eta'_{A_\alpha(a)}\zeta_{A_\alpha(a)}\right) + \mathrm{O}(\log\Lambda)\right]. \tag{6.67}$$

The real rate of the $(-,+)$-branch is

$$-\frac{2}{3}\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\xi_{I_\alpha(r)} + \frac{\beta}{2}\sum_{a=1}^{q}\zeta_{A_\alpha(a)}\right),$$

which is exactly the rate in (6.59). Since all the parameters are positive, every other sign pattern has a strictly larger real rate. By the generic choice of the parameters, all these rates are distinct.

If a coefficient of a branch other than $(-,+)$ in (6.52) were nonzero, the unique branch of maximal real rate among the nonzero terms would dominate the right-hand side exponentially, contradicting (6.59). Therefore

$$c_{\boldsymbol{\epsilon}',\boldsymbol{\eta}'} = 0 \qquad \text{unless} \qquad (\boldsymbol{\epsilon}',\boldsymbol{\eta}') = (-,+),$$

and hence

$$F = c_{-,+}F_{-,+}.$$

*Step 2: determine $c_{-,+}$.*

It remains to determine the coefficient $c_{-,+}$. Using the Airy asymptotics (6.62), we can rewrite this distinguished branch as

$$\begin{aligned}\Phi_{-,+} &= \left(\prod_{r=1}^{n}\phi_-(t_r)\right)\left(\prod_{a=1}^{m}\psi_+(s_a)\right)\prod_{1\leqslant i<j\leqslant n}(-\sqrt{t_i}-\sqrt{t_j})^{-2/\beta}\\ &\quad\times\prod_{1\leqslant a<b\leqslant m}(\sqrt{s_a}+\sqrt{s_b})^{-\beta/2}\prod_{i=1}^{n}\prod_{a=1}^{m}(-\sqrt{t_i}-\sqrt{s_a})\\ &= (1+\mathrm{O}(\Lambda^{-3})\exp\!\left[-\frac{2}{3}\Lambda^2\sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p}\xi_{I_\alpha(r)} + \frac{\beta}{2}\sum_{a=1}^{q}\zeta_{A_\alpha(a)}\right)\right]\times I\end{aligned} \tag{6.68}$$

where the algebraic factors $I$ in (6.68) is given by

$$I := \frac{\prod_{1\leqslant i<j\leqslant n}(-\sqrt{t_i}-\sqrt{t_j})^{-2/\beta}\prod_{1\leqslant a<b\leqslant m}(\sqrt{s_a}+\sqrt{s_b})^{-\beta/2}\prod_{i=1}^{n}\prod_{a=1}^{m}(-\sqrt{t_i}-\sqrt{s_a})}{\prod_{i=1}^{n}2\sqrt{\pi}t_r^{1/4}\prod_{a=1}^{m}2^{5/6}\beta^{1/6}\sqrt{\pi}s_a^{1/4}} \tag{6.69}$$

For any $t_i$ or $s_a$ lying in $\mathbb{C}_\sigma$, $\sigma\in\{\pm1\}$, the test configuration (6.54) gives

$$\sqrt{t_i},\sqrt{s_a} = \Lambda e^{\mathrm{i}\sigma\pi/6}\big(1+\mathrm{O}(\Lambda^{-1})\big), \qquad t_i^{1/4}, s_a^{1/4} = \Lambda^{1/2}e^{\mathrm{i}\sigma\pi/12}\big(1+\mathrm{O}(\Lambda^{-1})\big).$$

Hence, if two variables lie in the same half-plane $\mathbb{C}_\sigma$, then

$$\sqrt{u}+\sqrt{v} = 2\Lambda e^{\mathrm{i}\sigma\pi/6}\big(1+\mathrm{O}(\Lambda^{-1})\big), \quad -\sqrt{u}-\sqrt{v} = 2\Lambda e^{-\mathrm{i}5\sigma\pi/6}\big(1+\mathrm{O}(\Lambda^{-1})\big),$$

whereas if they lie in opposite half-planes, then

$$\sqrt{u}+\sqrt{v} = \sqrt{3}\,\Lambda\big(1+\mathrm{O}(\Lambda^{-1})\big), \quad -\sqrt{u}-\sqrt{v} = \sqrt{3}e^{\mathrm{i}\pi}\,\Lambda\big(1+\mathrm{O}(\Lambda^{-1})\big).$$

In the above, we chose the logarithmic branches so that, same-half-plane negative sums have arguments tending to $-5\pi/6$ in the upper half-plane and $5\pi/6$ in the lower half-plane, while cross-half-plane negative sums have arguments tending to $\pi$, Using $\beta = 2p/q$, $n = kp$ and $m = kq$, the total power of $\Lambda$ in the algebraic prefactor $I$ is

$$-\frac{kp(kp-1)}{\beta} - \frac{\beta kq(kq-1)}{4} + k^2pq - \frac{k(p+q)}{2} = 0.$$

The absolute value of the remaining constant is

$$2^{-k(p+q)/2} \left(\frac{\beta}{2}\right)^{-kq/6} \pi^{-k(p+q)/2}.$$

We recall the balance sign condition (6.46). Then the total phase of the algebraic prefactor $I$ divided by $\pi$ is

$$-\frac{p+q}{12}(k_+ - k_-) + \frac{5}{3\beta}\left[\binom{k_+p}{2} - \binom{k_-p}{2}\right] - \frac{2p^2}{\beta}k_+k_- - \frac{\beta}{12}\left[\binom{k_+q}{2} - \binom{k_-q}{2}\right] - \frac{5pq}{6}(k_+^2 - k_-^2) + 2pqk_+k_-.$$

Using $\beta/2 = p/q$, this equals

$$\frac{q}{2}\left[p\big(k_-^2 + 2k_+k_- - k_+^2\big) + k_- - k_+\right].$$

Since $W_{-,+} = 1 + \mathrm{O}(\Lambda^{-3/2})$, we obtain

$$\begin{aligned} F_{-,+}(\boldsymbol{t}, \boldsymbol{s}) = {} & 2^{-k(p+q)/2} \left(\frac{\beta}{2}\right)^{-kq/6} \pi^{-k(p+q)/2} \\ & \times \exp\left\{\frac{\mathrm{i}\pi q}{2}\left[p\big(k_-^2 + 2k_+k_- - k_+^2\big) + k_- - k_+\right]\right\} \\ & \times \exp\left[-\frac{2}{3}\Lambda^2 \sum_{\alpha=1}^{k}\left(\sum_{r=1}^{p} \xi_{I_\alpha(r)} + \frac{\beta}{2}\sum_{a=1}^{q} \zeta_{A_\alpha(a)}\right)\right]\big(1 + \mathrm{O}(\Lambda^{-1})\big). \end{aligned} \tag{6.70}$$

Comparing (6.70) with (6.59), we conclude that

$$c_{-,+} = 2^{k(p+q)/2} \left(\frac{\beta}{2}\right)^{kq/6} \pi^{k(p+q)/2} \exp\left\{-\frac{\mathrm{i}\pi q}{2}\left[p\big(k_-^2 + 2k_+k_- - k_+^2\big) + k_- - k_+\right]\right\}.$$

This proves (6.47). $\square$

# 7 Solutions of the bulk deformed CMS operators

We prove the first statement of Theorem 6.2, namely that the solution space has dimension $2^{n+m}$, in Section 7.1. We derive the conjugated equations in Section 7.2. We then prove the second statement, namely the construction of solutions with prescribed leading behavior, in Section 7.3.

## 7.1 Local holonomic rank

For $1 \leqslant i \leqslant n$ and $1 \leqslant a \leqslant m$, define

$$\mathcal{T}_i := \partial_{t_i}^2 + 1 + \frac{2}{\beta}\sum_{j \neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^{m} \frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a}, \tag{7.1}$$

$$\mathcal{S}_a := \partial_{s_a}^2 + \frac{\beta^2}{4} + \frac{\beta}{2}\sum_{b \neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} - \sum_{i=1}^{n} \frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i}. \tag{7.2}$$

Then the bulk deformed CMS system (5.3)–(5.4) can be written as

$$\mathcal{T}_i F = 0, \qquad 1 \leqslant i \leqslant n, \qquad \mathcal{S}_a F = 0, \qquad 1 \leqslant a \leqslant m. \tag{7.3}$$

We first record the following commutator identities for the operators $\mathcal{T}_i$ and $\mathcal{S}_a$.

**Proposition 7.1** (Bulk commutator identities)**.** *For the operators $\mathcal{T}_i, \mathcal{S}_a$ defined in* (7.1)*–*(7.2)*, the following commutator identities hold:*

$$[\mathcal{T}_i, \mathcal{T}_j] = \frac{4}{\beta(t_i - t_j)^2}\,(\mathcal{T}_j - \mathcal{T}_i), \qquad 1 \leqslant i \neq j \leqslant n, \tag{7.4}$$

$$[\mathcal{S}_a, \mathcal{S}_b] = \frac{\beta}{(s_a - s_b)^2}\,(\mathcal{S}_b - \mathcal{S}_a), \qquad 1 \leqslant a \neq b \leqslant m, \tag{7.5}$$

$$[\mathcal{T}_i, \mathcal{S}_a] = \frac{4}{\beta(t_i - s_a)^2}\,\mathcal{S}_a - \frac{\beta}{(t_i - s_a)^2}\,\mathcal{T}_i, \qquad 1 \leqslant i \leqslant n,\ 1 \leqslant a \leqslant m. \tag{7.6}$$

*Proof of Proposition 7.1.* Introduce the building blocks

$$\mathcal{A}_i := \partial_{t_i}^2 + 1, \qquad \mathcal{B}_a := \partial_{s_a}^2 + \frac{\beta^2}{4},$$

$$\mathcal{U}_{ij} := \frac{2}{\beta}\frac{1}{t_i - t_j}\big(\partial_{t_i} - \partial_{t_j}\big) = \mathcal{U}_{ji}, \qquad \mathcal{W}_{ab} := \frac{\beta}{2}\frac{1}{s_a - s_b}\big(\partial_{s_a} - \partial_{s_b}\big) = \mathcal{W}_{ba},$$

$$\mathcal{X}_{ia} := \frac{1}{t_i - s_a}\left(\frac{\beta}{2}\partial_{t_i} + \partial_{s_a}\right).$$

Then the definitions of $\mathcal{T}_i$ and $\mathcal{S}_a$ become

$$\mathcal{T}_i = \mathcal{A}_i + \sum_{j \neq i} \mathcal{U}_{ij} - \frac{2}{\beta}\sum_{a=1}^{m} \mathcal{X}_{ia},$$

$$\mathcal{S}_a = \mathcal{B}_a + \sum_{b \neq a} \mathcal{W}_{ab} + \sum_{i=1}^{n} \mathcal{X}_{ia}.$$

We shall repeatedly use the following elementary consequences of the Leibniz rule. If $f, g$ are scalar functions and $\mathcal{P}, \mathcal{Q}$ are commuting constant-coefficient first-order differential operators, then

$$[\partial_x^2, f\mathcal{P}] = (\partial_x^2 f)\mathcal{P} + 2(\partial_x f)\partial_x \mathcal{P}, \tag{7.7}$$

$$[f\mathcal{P}, g\mathcal{Q}] = f(\mathcal{P}g)\mathcal{Q} - g(\mathcal{Q}f)\mathcal{P}. \tag{7.8}$$

First fix distinct $i, j$, and set

$$\lambda_{ij} := \frac{4}{\beta(t_i - t_j)^2}.$$

Applying (7.7)–(7.8) gives the following local identities:

$$[\mathcal{A}_i + \mathcal{U}_{ij},\ \mathcal{A}_j + \mathcal{U}_{ij}] = \lambda_{ij}(\mathcal{A}_j - \mathcal{A}_i),$$

$$[\mathcal{A}_i + \mathcal{U}_{ij} + \mathcal{U}_{ik},\ \mathcal{U}_{jk}] + [\mathcal{U}_{ik},\ \mathcal{A}_j + \mathcal{U}_{ij}] = \lambda_{ij}(\mathcal{U}_{jk} - \mathcal{U}_{ik}), \qquad k \neq i, j,$$

$$\Big[\mathcal{A}_i + \mathcal{U}_{ij} - \frac{2}{\beta}\mathcal{X}_{ia},\ -\frac{2}{\beta}\mathcal{X}_{ja}\Big] + \Big[-\frac{2}{\beta}\mathcal{X}_{ia},\ \mathcal{A}_j + \mathcal{U}_{ij}\Big] = \lambda_{ij}\Big(-\frac{2}{\beta}\mathcal{X}_{ja} + \frac{2}{\beta}\mathcal{X}_{ia}\Big).$$

In the expansion of $[\mathcal{T}_i, \mathcal{T}_j]$, all terms involving disjoint sets of variables commute. Thus the only nonzero contributions are precisely the local pieces above. Hence

$$[\mathcal{T}_i, \mathcal{T}_j] = \lambda_{ij}(\mathcal{A}_j - \mathcal{A}_i) + \lambda_{ij}\sum_{k \neq i,j}(\mathcal{U}_{jk} - \mathcal{U}_{ik}) + \lambda_{ij}\sum_{a=1}^{m}\Big(-\frac{2}{\beta}\mathcal{X}_{ja} + \frac{2}{\beta}\mathcal{X}_{ia}\Big) = \lambda_{ij}(\mathcal{T}_j - \mathcal{T}_i).$$

This proves (7.4).

Next fix distinct $a, b$, and set

$$\mu_{ab} := \frac{\beta}{(s_a - s_b)^2}.$$

The corresponding local identities are

$$\begin{aligned}
[\mathcal{B}_a + \mathcal{W}_{ab},\, \mathcal{B}_b + \mathcal{W}_{ab}] &= \mu_{ab}(\mathcal{B}_b - \mathcal{B}_a), \\
[\mathcal{B}_a + \mathcal{W}_{ab} + \mathcal{W}_{ac},\, \mathcal{W}_{bc}] + [\mathcal{W}_{ac},\, \mathcal{B}_b + \mathcal{W}_{ab}] &= \mu_{ab}(\mathcal{W}_{bc} - \mathcal{W}_{ac}), \qquad c \neq a, b, \\
[\mathcal{B}_a + \mathcal{W}_{ab} + \mathcal{X}_{ia},\, \mathcal{X}_{ib}] + [\mathcal{X}_{ia},\, \mathcal{B}_b + \mathcal{W}_{ab}] &= \mu_{ab}(\mathcal{X}_{ib} - \mathcal{X}_{ia}).
\end{aligned}$$

Again, after expanding $[\mathcal{S}_a, \mathcal{S}_b]$, all disjoint-variable commutators vanish. Therefore

$$[\mathcal{S}_a, \mathcal{S}_b] = \mu_{ab}(\mathcal{B}_b - \mathcal{B}_a) + \mu_{ab} \sum_{c \neq a,b} (\mathcal{W}_{bc} - \mathcal{W}_{ac}) + \mu_{ab} \sum_{i=1}^{n} (\mathcal{X}_{ib} - \mathcal{X}_{ia}) = \mu_{ab}(\mathcal{S}_b - \mathcal{S}_a).$$

This proves (7.5).

Finally fix $i, a$, and set

$$\rho_{ia} := \frac{4}{\beta(t_i - s_a)^2}, \qquad \sigma_{ia} := \frac{\beta}{(t_i - s_a)^2}.$$

The mixed local identities are

$$\begin{aligned}
\Big[\mathcal{A}_i - \frac{2}{\beta}\mathcal{X}_{ia},\, \mathcal{B}_a + \mathcal{X}_{ia}\Big] &= \rho_{ia}(\mathcal{B}_a + \mathcal{X}_{ia}) - \sigma_{ia}\Big(\mathcal{A}_i - \frac{2}{\beta}\mathcal{X}_{ia}\Big), \\
[\mathcal{A}_i - \tfrac{2}{\beta}\mathcal{X}_{ia} + \mathcal{U}_{ij},\, \mathcal{X}_{ja}] + [\mathcal{U}_{ij},\, \mathcal{B}_a + \mathcal{X}_{ia}] &= \rho_{ia}\mathcal{X}_{ja} - \sigma_{ia}\mathcal{U}_{ij}, \qquad j \neq i, \\
\Big[\mathcal{A}_i - \frac{2}{\beta}\mathcal{X}_{ia} - \frac{2}{\beta}\mathcal{X}_{ib},\, \mathcal{W}_{ab}\Big] + \Big[-\frac{2}{\beta}\mathcal{X}_{ib},\, \mathcal{B}_a + \mathcal{X}_{ia}\Big] &= \rho_{ia}\mathcal{W}_{ab} - \sigma_{ia}\Big(-\frac{2}{\beta}\mathcal{X}_{ib}\Big), \qquad b \neq a.
\end{aligned}$$

Expanding $[\mathcal{T}_i, \mathcal{S}_a]$ and keeping only the nonzero local contributions gives

$$\begin{aligned}
[\mathcal{T}_i, \mathcal{S}_a] &= \rho_{ia}(\mathcal{B}_a + \mathcal{X}_{ia}) - \sigma_{ia}\Big(\mathcal{A}_i - \frac{2}{\beta}\mathcal{X}_{ia}\Big) \\
&\quad + \sum_{j \neq i} (\rho_{ia}\mathcal{X}_{ja} - \sigma_{ia}\mathcal{U}_{ij}) + \sum_{b \neq a} \Big(\rho_{ia}\mathcal{W}_{ab} - \sigma_{ia}\Big(-\frac{2}{\beta}\mathcal{X}_{ib}\Big)\Big) = \rho_{ia}\mathcal{S}_a - \sigma_{ia}\mathcal{T}_i.
\end{aligned}$$

This proves (7.6). $\square$

Let

$$K := \mathbb{C}(t_1, \ldots, t_n, s_1, \ldots, s_m), \tag{7.9}$$

and let

$$D_K := K\langle \partial_{t_1}, \ldots, \partial_{t_n}, \partial_{s_1}, \ldots, \partial_{s_m} \rangle \tag{7.10}$$

be the algebra of differential operators with coefficients in $K$.

For $1 \leqslant i \leqslant n$ and $1 \leqslant a \leqslant m$, write

$$\mathcal{T}_i = \partial_{t_i}^2 + \mathcal{P}_i, \qquad \mathcal{S}_a = \partial_{s_a}^2 + \mathcal{Q}_a,$$

where $\mathcal{P}_i, \mathcal{Q}_a \in D_K$ have derivative order at most 1. Let

$$\mathcal{I} := D_K\mathcal{T}_1 + \cdots + D_K\mathcal{T}_n + D_K\mathcal{S}_1 + \cdots + D_K\mathcal{S}_m \subset D_K \tag{7.11}$$

be the left ideal generated by $\mathcal{T}_1, \dots, \mathcal{T}_n, \mathcal{S}_1, \dots, \mathcal{S}_m$.

For subsets

$$I \subseteq \{1, \dots, n\}, \qquad A \subseteq \{1, \dots, m\},$$

write

$$\partial_{t_I} := \prod_{i \in I} \partial_{t_i}, \qquad \partial_{s_A} := \prod_{a \in A} \partial_{s_a},$$

where the factors are arranged in increasing order. We call

$$\partial_{t_I} \partial_{s_A} F$$

a *square-free derivative* of $F$. The word "square-free" means that no variable is differentiated more than once. There are exactly $2^{n+m}$ such derivatives.

We shall prove that

$$G := \{\mathcal{T}_1, \dots, \mathcal{T}_n, \mathcal{S}_1, \dots, \mathcal{S}_m\}$$

is a left Gröbner basis of $\mathcal{I}$. In the present setting, this means that every operator in $D_K$ can be reduced, modulo the generators in $G$, to a unique normal form whose derivative monomials are not divisible by any of the leading monomials

$$\operatorname{lm}(\mathcal{T}_i) = \partial_{t_i}^2, \qquad \operatorname{lm}(\mathcal{S}_a) = \partial_{s_a}^2.$$

Consequently, the reduced monomials are precisely the square-free derivative monomials, namely those in which no derivative appears with exponent 2 or larger. For further background, see [79].

**Proposition 7.2** (Square-free normal forms)**.** *Assume that the commutator identities in Proposition 7.1 hold. Then the residue classes of the square-free derivative monomials*

$$\partial_{t_I} \partial_{s_A}, \qquad I \subseteq \{1, \dots, n\}, \quad A \subseteq \{1, \dots, m\},$$

*form a $K$-basis of $D_K/\mathcal{I}$. Equivalently, every class in $D_K/\mathcal{I}$ has a unique representative of the form*

$$\sum_{I,A} c_{I,A}\, \partial_{t_I} \partial_{s_A}, \qquad c_{I,A} \in K.$$

*In particular,*

$$\dim_K(D_K/\mathcal{I}) = 2^{n+m}.$$

*Proof.* The equations $\mathcal{T}_i = 0$ and $\mathcal{S}_a = 0$ should be thought of as reduction rules:

$$\partial_{t_i}^2 \equiv -\mathcal{P}_i, \qquad \partial_{s_a}^2 \equiv -\mathcal{Q}_a \qquad \bmod \mathcal{I}. \tag{7.12}$$

Since $\mathcal{P}_i$ and $\mathcal{Q}_a$ have derivative order at most 1, each reduction replaces a repeated derivative by lower-order terms. Thus the square-free derivative monomials are the natural candidates for a basis.

The point is to prove that these reductions are consistent. For example, suppose we want to reduce $\partial_{t_i}^2 \partial_{t_j}^2$. There are two possible first steps: we may first replace $\partial_{t_i}^2$ by $-\mathcal{P}_i$, or we may first replace $\partial_{t_j}^2$ by $-\mathcal{P}_j$. This gives $-\partial_{t_j}^2 \mathcal{P}_i$ or $-\partial_{t_i}^2 \mathcal{P}_j$. Thus the obstruction to consistency is, up to sign, the so-called $S$-pair

$$\operatorname{Sp}(\mathcal{T}_i, \mathcal{T}_j) := \partial_{t_j}^2 \mathcal{T}_i - \partial_{t_i}^2 \mathcal{T}_j = \partial_{t_j}^2 \mathcal{P}_i - \partial_{t_i}^2 \mathcal{P}_j.$$

Buchberger's criterion is a systematic way to check exactly this. It says that, to prove that a set of generators is a left Gröbner basis, it is enough to check the differences coming from all possible pairwise overlaps of leading monomials. These differences are called $S$-pairs. It is enough to check that the $S$-pair is trivial modulo the generators, in the sense that it can be written as a left linear combination of the generators whose terms all have leading monomial strictly smaller than the original overlap. Equivalently, its remainder after division by the generators is zero.

In our case,
$$\mathcal{T}_i = \partial_{t_i}^2 + \mathcal{P}_i, \qquad \mathcal{S}_a = \partial_{s_a}^2 + \mathcal{Q}_a,$$
with $\mathcal{P}_i, \mathcal{Q}_a$ of order at most 1. Hence the only possible overlaps are
$$\partial_{t_i}^2 \partial_{t_j}^2, \qquad \partial_{s_a}^2 \partial_{s_b}^2, \qquad \partial_{t_i}^2 \partial_{s_a}^2.$$
Equivalently, we only need to check the three types of pairs
$$(\mathcal{T}_i, \mathcal{T}_j), \qquad (\mathcal{S}_a, \mathcal{S}_b), \qquad (\mathcal{T}_i, \mathcal{S}_a).$$

1. *The pair $(\mathcal{T}_i, \mathcal{T}_j)$, with $i \neq j$.*
The corresponding $S$-pair is
$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{T}_j) := \partial_{t_j}^2 \mathcal{T}_i - \partial_{t_i}^2 \mathcal{T}_j.$$
Using
$$\partial_{t_j}^2 = \mathcal{T}_j - \mathcal{P}_j, \qquad \partial_{t_i}^2 = \mathcal{T}_i - \mathcal{P}_i,$$
we get
$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{T}_j) = (\mathcal{T}_j - \mathcal{P}_j)\mathcal{T}_i - (\mathcal{T}_i - \mathcal{P}_i)\mathcal{T}_j = [\mathcal{T}_j, \mathcal{T}_i] - \mathcal{P}_j \mathcal{T}_i + \mathcal{P}_i \mathcal{T}_j.$$
By the commutator identity (7.4),
$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{T}_j) = \left( \frac{4}{\beta(t_i - t_j)^2} - \mathcal{P}_j \right) \mathcal{T}_i + \left( \mathcal{P}_i - \frac{4}{\beta(t_i - t_j)^2} \right) \mathcal{T}_j.$$
This is a left linear combination of $\mathcal{T}_i$ and $\mathcal{T}_j$. Moreover, the terms involving $\mathcal{P}_i$ and $\mathcal{P}_j$ have derivative order at most 3, while the overlap $\partial_{t_i}^2 \partial_{t_j}^2$ has derivative order 4. Hence this $S$-pair reduces to 0 modulo the generators.

2. *The pair $(\mathcal{S}_a, \mathcal{S}_b)$, with $a \neq b$.*
The corresponding $S$-pair is
$$\mathrm{Sp}(\mathcal{S}_a, \mathcal{S}_b) := \partial_{s_b}^2 \mathcal{S}_a - \partial_{s_a}^2 \mathcal{S}_b.$$
Using
$$\partial_{s_b}^2 = \mathcal{S}_b - \mathcal{Q}_b, \qquad \partial_{s_a}^2 = \mathcal{S}_a - \mathcal{Q}_a,$$
we compute
$$\mathrm{Sp}(\mathcal{S}_a, \mathcal{S}_b) = (\mathcal{S}_b - \mathcal{Q}_b)\mathcal{S}_a - (\mathcal{S}_a - \mathcal{Q}_a)\mathcal{S}_b = [\mathcal{S}_b, \mathcal{S}_a] - Q_b \mathcal{S}_a + \mathcal{Q}_a \mathcal{S}_b.$$
By the commutator identity (7.5), we have
$$\mathrm{Sp}(\mathcal{S}_a, \mathcal{S}_b) = \left( \frac{\beta}{(s_a - s_b)^2} - \mathcal{Q}_b \right) \mathcal{S}_a + \left( Q_a - \frac{\beta}{(s_a - s_b)^2} \right) \mathcal{S}_b.$$
Again this is a left linear combination of the generators. The terms involving $\mathcal{Q}_a$ and $\mathcal{Q}_b$ have derivative order at most 3, while the overlap $\partial_{s_a}^2 \partial_{s_b}^2$ has derivative order 4. Hence this $S$-pair reduces to 0.

3. *The mixed pair $(\mathcal{T}_i, \mathcal{S}_a)$.*
The corresponding $S$-pair is
$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{S}_a) := \partial_{s_a}^2 \mathcal{T}_i - \partial_{t_i}^2 \mathcal{S}_a.$$
Using
$$\partial_{s_a}^2 = \mathcal{S}_a - \mathcal{Q}_a, \qquad \partial_{t_i}^2 = \mathcal{T}_i - \mathcal{P}_i,$$
we get
$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{S}_a) = (\mathcal{S}_a - \mathcal{Q}_a)\mathcal{T}_i - (\mathcal{T}_i - \mathcal{P}_i)\mathcal{S}_a = [\mathcal{S}_a, \mathcal{T}_i] - \mathcal{Q}_a \mathcal{T}_i + \mathcal{P}_i \mathcal{S}_a.$$

By the mixed commutator identity (7.6), we have

$$\mathrm{Sp}(\mathcal{T}_i, \mathcal{S}_a) = \left(\frac{\beta}{(t_i - s_a)^2} - \mathcal{Q}_a\right)\mathcal{T}_i + \left(\mathcal{P}_i - \frac{4}{\beta(t_i - s_a)^2}\right)\mathcal{S}_a.$$

This is a left linear combination of $\mathcal{T}_i$ and $\mathcal{S}_a$. The terms involving $\mathcal{P}_i$ and $\mathcal{Q}_a$ have derivative order at most 3, while the overlap $\partial_{t_i}^2\partial_{s_a}^2$ has derivative order 4. Hence this $S$-pair also reduces to 0.

We have checked all possible overlaps. By Buchberger's criterion, $G := \{\mathcal{T}_1, \dots, \mathcal{T}_n, \mathcal{S}_1, \dots, \mathcal{S}_m\}$ is a left Gröbner basis of $\mathcal{I}$. Therefore the initial ideal is generated by the leading monomials of these operators:

$$\left(\partial_{t_1}^2, \dots, \partial_{t_n}^2, \partial_{s_1}^2, \dots, \partial_{s_m}^2\right).$$

A derivative monomial is standard if and only if it is not divisible by any of these leading monomials. Equivalently, no derivative $\partial_{t_i}$ or $\partial_{s_a}$ appears with exponent 2 or larger. Thus the standard monomials are precisely

$$\partial_{t_I}\partial_{s_A}, \qquad I \subseteq \{1, \dots, n\}, \quad A \subseteq \{1, \dots, m\}.$$

These are exactly the square-free derivative monomials. The normal-form theorem for Gröbner bases now implies that the residue classes of these standard monomials form a $K$-basis of $D_K/\mathcal{I}$. Since there are $2^{n+m}$ choices of $(I, A)$, we obtain $\dim_K(D_K/\mathcal{I}) = 2^{n+m}$. □

*Proof of the first statement of Theorem 6.2.* We show that the second-order system (5.3)–(5.4) is equivalent, near every point of $\mathfrak{X}$, to a flat first-order system of size $2^{n+m}$.

*Step 1: repeated derivatives can be eliminated.*

We can rewrite the equations (5.3)–(5.4) as follows:

$$\partial_{t_i}^2 F = -F - \frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j}F + \sum_{a=1}^{m}\frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a}F, \tag{7.13}$$

$$\partial_{s_a}^2 F = -\frac{\beta^2}{4}F - \frac{\beta}{2}\sum_{b\neq a}\frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b}F + \sum_{i=1}^{n}\frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i}F. \tag{7.14}$$

Thus, whenever a derivative of $F$ contains $\partial_{t_i}^2$ or $\partial_{s_a}^2$, we may replace that repeated derivative by lower-order derivatives using (7.13) or (7.14).

Proposition 7.2, says exactly that all such reduction choices are compatible. Equivalently, every derivative of a solution can be written uniquely as a rational linear combination of the $2^{n+m}$ square-free derivatives $\partial_{t_I}\partial_{s_A}F$.

*Step 2: introduce the finite set of unknowns.* We recall the notation from (7.9), (7.10), and (7.11): $K$ is the field of rational functions, $D_K$ is the algebra of differential operators with coefficients in $K$, and $\mathcal{I}$ is the left ideal generated by the differential equations.

By Proposition 7.2, the residue classes

$$e_{I,A} := \left[\partial_{t_I}\partial_{s_A}\right], \qquad I \subseteq \{1, \dots, n\}, \quad A \subseteq \{1, \dots, m\},$$

form a $K$-basis of $D_K/\mathcal{I}$. In concrete terms, this means that every derivative can be reduced uniquely to a $K$-linear combination of the square-free derivatives $\partial_{t_I}\partial_{s_A}F$.

*Step 3: the square-free derivatives satisfy a first-order system.*

Since the classes $e_{I,A}$ form a basis of $D_K/\mathcal{I}$, multiplication by any coordinate derivative can be expressed in this basis. Thus there are matrices $M_{t_i}$ and $M_{s_a}$, with entries in $K$, such that

$$\begin{aligned} \partial_{t_i} e_{I,A} &= \sum_{J,B} (M_{t_i})_{(I,A),(J,B)}\, e_{J,B}, \\ \partial_{s_a} e_{I,A} &= \sum_{J,B} (M_{s_a})_{(I,A),(J,B)}\, e_{J,B}. \end{aligned} \tag{7.15}$$

The entries of these matrices are rational functions on $\mathfrak{X}$, because they are obtained by repeatedly using the reduction rules (7.13) and (7.14).

Now let $F$ be a local holomorphic solution of (7.3). Define the vector of square-free derivatives

$$U(F) := \big(\partial_{t_I}\partial_{s_A}F\big)_{I,A}.$$

Then (7.15) implies

$$\begin{aligned} \partial_{t_i}U(F) &= M_{t_i}\,U(F), \qquad 1 \leqslant i \leqslant n, \\ \partial_{s_a}U(F) &= M_{s_a}\,U(F), \qquad 1 \leqslant a \leqslant m. \end{aligned} \tag{7.16}$$

Thus the original scalar second-order system gives a first-order linear system for the $2^{n+m}$ square-free derivatives of $F$.

*Step 4: the first-order system is flat.*

The coordinate derivatives commute:

$$[\partial_{t_i}, \partial_{t_j}] = 0, \qquad [\partial_{s_a}, \partial_{s_b}] = 0, \qquad [\partial_{t_i}, \partial_{s_a}] = 0.$$

Therefore their induced actions on $D_K/\mathcal{I}$ commute. Written in the basis $\{e_{I,A}\}$, this gives

$$\begin{aligned} \partial_{t_i}M_{t_j} - \partial_{t_j}M_{t_i} &= [M_{t_i}, M_{t_j}], \\ \partial_{s_a}M_{s_b} - \partial_{s_b}M_{s_a} &= [M_{s_a}, M_{s_b}], \\ \partial_{t_i}M_{s_a} - \partial_{s_a}M_{t_i} &= [M_{t_i}, M_{s_a}], \end{aligned} \tag{7.17}$$

where $[A,B] = AB - BA$. These are exactly the compatibility equations for the first-order system (7.16). Hence the system is flat.

*Step 5: local existence and dimension of the solution space.*

Fix $(\boldsymbol{t}_0, \boldsymbol{s}_0) \in \mathfrak{X}$, and choose a sufficiently small simply connected neighborhood $\Omega \subset \mathfrak{X}$ of $(\boldsymbol{t}_0, \boldsymbol{s}_0)$. On $\Omega$, the matrices $M_{t_i}, M_{s_a}$ are holomorphic. Since $\Omega$ is simply connected and the system (7.16) is flat, parallel transport is independent of the path. Equivalently, for each initial value

$$c \in \mathbb{C}^{2^{n+m}},$$

there is a unique holomorphic solution of (7.16) with $U(p) = c$.

Indeed, along any smooth path $\gamma : [0,1] \to \Omega$, the restriction of $U$ satisfies the linear ordinary differential equation

$$\frac{\mathrm{d}V}{\mathrm{d}\tau} = \sum_{i=1}^{n} \dot\gamma_i(\tau)\, M_{t_i}(\gamma(\tau))\, V + \sum_{a=1}^{m} \dot\gamma_{n+a}(\tau)\, M_{s_a}(\gamma(\tau))\, V, \qquad V(0) = c,$$

and flatness means that the resulting transport depends only on the endpoints. Therefore the space of local holomorphic solutions of the first-order system (7.16) has dimension $2^{n+m}$.

It remains to check that this first-order system (7.16) is equivalent to the original scalar system (7.3). We have already seen that every scalar solution $F$ of (7.3) produces a solution $U(F)$ of the first-order system (7.16). Conversely, suppose $U = (U_{I,A})$ solves the first-order system (7.16). Define

$$F := U_{\emptyset,\emptyset}.$$

Since

$$\partial_{t_i}e_{\emptyset,\emptyset} = e_{\{i\},\emptyset}, \qquad \partial_{s_a}e_{\emptyset,\emptyset} = e_{\emptyset,\{a\}},$$

the first-order system gives

$$U_{\{i\},\emptyset} = \partial_{t_i}F, \qquad U_{\emptyset,\{a\}} = \partial_{s_a}F.$$

Repeating the same argument for larger subsets $I, A$, we obtain

$$U_{I,A} = \partial_{t_I}\partial_{s_A} F$$

for every square-free pair $(I, A)$.

Finally, the reduction rules encoded in the matrices $M_{t_i}$ and $M_{s_a}$ are precisely the equations $\mathcal{T}_i = 0$ and $\mathcal{S}_a = 0$. Therefore the identities

$$\partial_{t_i}^2 \equiv -\mathcal{P}_i, \qquad \partial_{s_a}^2 \equiv -\mathcal{Q}_a \qquad \mod \mathcal{I}$$

imply

$$\mathcal{T}_i F = 0, \qquad \mathcal{S}_a F = 0.$$

Hence $F$ is a holomorphic solution of the original system.

Therefore the space of germs at $(\boldsymbol{t}_0, \boldsymbol{s}_0)$ of holomorphic solutions of (7.3) is naturally identified with the space of germs of solutions of the flat first-order system. This space has dimension $2^{n+m}$, as claimed. $\square$

## 7.2 The conjugated system

For each sign pattern $(\boldsymbol{\epsilon}, \boldsymbol{\eta})$, we write $F = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W$, where $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ contains the prescribed leading exponential and algebraic factors. This separates the expected leading asymptotic behavior from the unknown correction term. The correction $W$ is expected to satisfy $W \to 1$ at infinity. In the following proposition, we conjugate the original operators by $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ and derive the resulting equations for $W$.

We recall $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ from (6.5)

$$\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s}) := \exp\left(\mathrm{i}\sum_{i=1}^{n} \epsilon_i t_i + \mathrm{i}\frac{\beta}{2}\sum_{a=1}^{m} \eta_a s_a\right) \prod_{1\leqslant i<j\leqslant n} (t_i - t_j)^{\alpha_{ij}} \prod_{1\leqslant a<b\leqslant m} (s_a - s_b)^{\gamma_{ab}} \prod_{i=1}^{n}\prod_{a=1}^{m}(t_i - s_a)^{\rho_{ia}}, \tag{7.18}$$

and define

$$\begin{aligned} \Omega_i &:= \sum_{j\neq i} \frac{\alpha_{ij}}{t_i - t_j} + \sum_{a=1}^{m} \frac{\rho_{ia}}{t_i - s_a}, \\ \Sigma_a &:= \sum_{b\neq a} \frac{\gamma_{ab}}{s_a - s_b} + \sum_{i=1}^{n} \frac{\rho_{ia}}{s_a - t_i}. \end{aligned} \tag{7.19}$$

**Proposition 7.3** (Conjugated operator family)**.** *Let $F = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W$. Then the original system (7.3)*

$$\mathcal{T}_i F = 0, \qquad 1 \leqslant i \leqslant n, \qquad \mathcal{S}_a F = 0, \qquad 1 \leqslant a \leqslant m, \tag{7.20}$$

*is equivalent to*

$$(\mathcal{L}_i^{(t)} + \mathcal{R}_i^{(t)})W = 0, \qquad i = 1, \dots, n, \tag{7.21}$$

$$(\mathcal{L}_a^{(s)} + \mathcal{R}_a^{(s)})W = 0, \qquad a = 1, \dots, m, \tag{7.22}$$

*where*

$$\mathcal{L}_i^{(t)} = \partial_{t_i}^2 + 2\mathrm{i}\epsilon_i\,\partial_{t_i}, \tag{7.23}$$

$$\mathcal{R}_i^{(t)} = 2\Omega_i\,\partial_{t_i} + \frac{2}{\beta}\sum_{j\neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^{m} \frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a} + V_i^{(t)}, \tag{7.24}$$

$$\mathcal{L}_a^{(s)} = \partial_{s_a}^2 + \mathrm{i}\beta\eta_a\,\partial_{s_a}, \tag{7.25}$$

$$\mathcal{R}_a^{(s)} = 2\Sigma_a\,\partial_{s_a} + \frac{\beta}{2}\sum_{b\neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} - \sum_{i=1}^{n} \frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i} + V_a^{(s)}, \tag{7.26}$$

*and*

$$V_i^{(t)} = \partial_{t_i}\Omega_i + \Omega_i^2 + \frac{2}{\beta}\sum_{j\neq i}\frac{\Omega_i - \Omega_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{\Omega_i + (2/\beta)\Sigma_a}{t_i - s_a}, \tag{7.27}$$

$$V_a^{(s)} = \partial_{s_a}\Sigma_a + \Sigma_a^2 + \frac{\beta}{2}\sum_{b\neq a}\frac{\Sigma_a - \Sigma_b}{s_a - s_b} - \sum_{i=1}^{n}\frac{\Sigma_a + (\beta/2)\Omega_i}{s_a - t_i}. \tag{7.28}$$

*Proof.* Let $\Phi := \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ be as in (7.18). Since $\Phi$ is nowhere vanishing on $\mathfrak{X}$, the equations

$$\mathcal{T}_i F = 0, \qquad \mathcal{S}_a F = 0, \qquad F = \Phi W,$$

are equivalent to

$$\widetilde{\mathcal{T}}_i W = 0, \qquad \widetilde{\mathcal{S}}_a W = 0,$$

where

$$\widetilde{\mathcal{T}}_i := \Phi^{-1}\mathcal{T}_i\Phi, \qquad \widetilde{\mathcal{S}}_a := \Phi^{-1}\mathcal{S}_a\Phi.$$

We recall $\Omega_i$ and $\Sigma_a$ from (7.19) and set

$$\mu_i := \partial_{t_i}\log\Phi = \mathrm{i}\epsilon_i + \Omega_i, \qquad \nu_a := \partial_{s_a}\log\Phi = \mathrm{i}\frac{\beta}{2}\eta_a + \Sigma_a. \tag{7.29}$$

Then

$$\Phi^{-1}\partial_{t_i}\Phi = \partial_{t_i} + \mu_i, \qquad \Phi^{-1}\partial_{s_a}\Phi = \partial_{s_a} + \nu_a, \tag{7.30}$$

and

$$\begin{aligned}\Phi^{-1}\partial_{t_i}^2\Phi &= (\partial_{t_i} + \mu_i)^2 = \partial_{t_i}^2 + 2\mu_i\partial_{t_i} + \partial_{t_i}\mu_i + \mu_i^2,\\ \Phi^{-1}\partial_{s_a}^2\Phi &= (\partial_{s_a} + \nu_a)^2 = \partial_{s_a}^2 + 2\nu_a\partial_{s_a} + \partial_{s_a}\nu_a + \nu_a^2.\end{aligned} \tag{7.31}$$

We first compute $\widetilde{\mathcal{T}}_i$. A substitution of (7.30) and (7.31) into $\mathcal{T}_i$, recalled in (7.1), gives

$$\widetilde{\mathcal{T}}_i = (\partial_{t_i} + \mu_i)^2 + 1 + \frac{2}{\beta}\sum_{j\neq i}\frac{(\partial_{t_i} + \mu_i) - (\partial_{t_j} + \mu_j)}{t_i - t_j} - \sum_{a=1}^{m}\frac{(\partial_{t_i} + \mu_i) + (2/\beta)(\partial_{s_a} + \nu_a)}{t_i - s_a}. \tag{7.32}$$

Using $\mu_i = \mathrm{i}\epsilon_i + \Omega_i$ from (7.29), the derivative terms in (7.32) are

$$\partial_{t_i}^2 + 2\mathrm{i}\epsilon_i\partial_{t_i} + 2\Omega_i\partial_{t_i} + \frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^{m}\frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a}. \tag{7.33}$$

The scalar part is

$$\begin{aligned}&\partial_{t_i}\Omega_i + (\mathrm{i}\epsilon_i + \Omega_i)^2 + 1 + \frac{2}{\beta}\sum_{j\neq i}\frac{\mathrm{i}(\epsilon_i - \epsilon_j) + \Omega_i - \Omega_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{\mathrm{i}(\epsilon_i + \eta_a) + \Omega_i + (2/\beta)\Sigma_a}{t_i - s_a}\\ &= (1 - \epsilon_i^2) + \mathrm{i}\left(2\epsilon_i\Omega_i + \frac{2}{\beta}\sum_{j\neq i}\frac{\epsilon_i - \epsilon_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{\epsilon_i + \eta_a}{t_i - s_a}\right) + V_i^{(t)},\end{aligned} \tag{7.34}$$

where we have separated the terms involving only $\Omega_i, \Sigma_a$ from the terms involving $\epsilon_i, \eta_a$.

Since $\epsilon_i^2 = 1$, the constant term $1 - \epsilon_i^2$ in (7.34) vanishes. For the remaining term in parentheses, substitute the definition of $\Omega_i$ from (7.19):

$$2\epsilon_i\Omega_i + \frac{2}{\beta}\sum_{j\neq i}\frac{\epsilon_i - \epsilon_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{\epsilon_i + \eta_a}{t_i - s_a} = \sum_{j\neq i}\frac{2\epsilon_i\alpha_{ij} + \frac{2}{\beta}(\epsilon_i - \epsilon_j)}{t_i - t_j} + \sum_{a=1}^{m}\frac{2\epsilon_i\rho_{ia} - (\epsilon_i + \eta_a)}{t_i - s_a} = 0.$$

This vanishes by the defining relations (6.2) for the exponents $\alpha_{ij}$ and $\rho_{ia}$. Hence the scalar part (7.34) is $V_i^{(t)}$, and (7.21) follows by combining (7.32) and (7.33).

The computation for $\widetilde{\mathcal{S}}_a$ is analogous. Substituting (7.30) and (7.31) into $\mathcal{S}_a$, recalled in (7.2), we obtain

$$\widetilde{\mathcal{S}}_a = (\partial_{s_a} + \nu_a)^2 + \frac{\beta^2}{4} + \frac{\beta}{2}\sum_{b\neq a}\frac{(\partial_{s_a}+\nu_a) - (\partial_{s_b}+\nu_b)}{s_a - s_b} - \sum_{i=1}^{n}\frac{(\partial_{s_a}+\nu_a) + (\beta/2)(\partial_{t_i}+\mu_i)}{s_a - t_i}. \tag{7.35}$$

Using $\nu_a = \mathrm{i}\beta\eta_a/2 + \Sigma_a$ from (7.29), the derivative terms in (7.35) are

$$\partial_{s_a}^2 + \mathrm{i}\beta\eta_a\,\partial_{s_a} + 2\Sigma_a\partial_{s_a} + \frac{\beta}{2}\sum_{b\neq a}\frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} - \sum_{i=1}^{n}\frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i}. \tag{7.36}$$

The scalar part is

$$\begin{aligned}&\partial_{s_a}\Sigma_a + \left(\mathrm{i}\frac{\beta}{2}\eta_a + \Sigma_a\right)^2 + \frac{\beta^2}{4} + \frac{\beta}{2}\sum_{b\neq a}\frac{\mathrm{i}\frac{\beta}{2}(\eta_a-\eta_b) + \Sigma_a - \Sigma_b}{s_a - s_b} - \sum_{i=1}^{n}\frac{\mathrm{i}\frac{\beta}{2}(\eta_a+\epsilon_i) + \Sigma_a + (\beta/2)\Omega_i}{s_a - t_i}\\ &= \frac{\beta^2}{4}(1-\eta_a^2) + \mathrm{i}\left(\beta\eta_a\Sigma_a + \frac{\beta^2}{4}\sum_{b\neq a}\frac{\eta_a-\eta_b}{s_a-s_b} - \frac{\beta}{2}\sum_{i=1}^{n}\frac{\eta_a+\epsilon_i}{s_a-t_i}\right) + V_a^{(s)}.\end{aligned} \tag{7.37}$$

Since $\eta_a^2 = 1$, the constant term vanishes. For the remaining term in parentheses, substitute the definition of $\Sigma_a$ from (7.19):

$$\beta\eta_a\Sigma_a + \frac{\beta^2}{4}\sum_{b\neq a}\frac{\eta_a-\eta_b}{s_a-s_b} - \frac{\beta}{2}\sum_{i=1}^{n}\frac{\eta_a+\epsilon_i}{s_a-t_i} = \sum_{b\neq a}\frac{\beta\eta_a\gamma_{ab} + \frac{\beta^2}{4}(\eta_a-\eta_b)}{s_a-s_b} + \sum_{i=1}^{n}\frac{\beta\eta_a\rho_{ia} - \frac{\beta}{2}(\eta_a+\epsilon_i)}{s_a-t_i} = 0.$$

Again, this vanishes by the defining relations (6.2) for the exponents $\gamma_{ab}$ and $\rho_{ia}$. Hence the scalar part (7.37) is $V_a^{(s)}$, and (7.22) follows by combining (7.35) and (7.36). This proves the proposition. $\square$

## 7.3 Solutions attached to the branch factors

In this section, we prove the second statement of Theorem 6.2. The proof reduces the conjugated system (7.21)–(7.22) to a one-dimensional problem along a carefully chosen shifted ray, constructed in Proposition 7.4. In Proposition 7.5, we show that along this ray the scaled square-free derivatives satisfy a block first-order system in the radial variable $\lambda$. The non-empty derivative block has a positive spectral gap, which allows us to impose the condition that this block decays at infinity by solving a Volterra integral equation. The Volterra operator is a contraction, and hence gives a unique solution of the first-order system with the desired asymptotic estimates. Translating these estimates back from the scaled derivative system gives the desired estimates for $W$ and its square-free derivatives.

**Proposition 7.4.** *Let $\mathfrak{C}$ be the domain (6.7). Fix*

$$\boldsymbol{\sigma} = (\sigma_1,\dots,\sigma_{n+m}) \in \{\pm 1\}^{n+m},$$

*and define*

$$\boldsymbol{\omega} = (\omega_1,\dots,\omega_{n+m}), \qquad \omega_j = j + \mathrm{i}\sigma_j, \qquad 1 \leqslant j \leqslant n+m.$$

*Then, for every*

$$\boldsymbol{z} = (z_1,\dots,z_{n+m}) \in \mathfrak{C}$$

*and every $r \geqslant 0$, we have $\boldsymbol{z} + r\boldsymbol{\omega} \in \mathfrak{C}$ and*

$$R\big(\boldsymbol{z} + r\boldsymbol{\omega}\big) \geqslant \frac{1}{4}\big(R(\boldsymbol{z}) + r\big), \tag{7.38}$$

*where $R(\boldsymbol{z})$ is defined in (6.6).*

*Proof.* For $1 \leqslant \ell < n+m$, we have

$$\operatorname{Re}\left[(z_{\ell+1} + r\omega_{\ell+1}) - (z_\ell + r\omega_\ell)\right] = \operatorname{Re}[z_{\ell+1} - z_\ell] + r > D.$$

Thus $\boldsymbol{z} + r\boldsymbol{\omega} \in \mathfrak{C}$.

Fix $1 \leqslant i < j \leqslant n+m$. Since $j - i \geqslant 1$,

$$\left|(z_j + r\omega_j) - (z_i + r\omega_i)\right| \geqslant \operatorname{Re}[z_j - z_i] + r(j-i) \geqslant r.$$

On the other hand,

$$\left|(z_j + r\omega_j) - (z_i + r\omega_i)\right| = |z_j - z_i + r(j-i) + \mathrm{i}r(\sigma_j - \sigma_i)| \geqslant |z_j - z_i + r(j-i)| - r|\sigma_j - \sigma_i|.$$

Since $\operatorname{Re}[z_j - z_i] > 0$ and $r(j-i) \geqslant 0$, adding $r(j-i)$ to the real part cannot decrease the modulus. Moreover, $|\sigma_j - \sigma_i| \leqslant 2$. Therefore

$$\left|(z_j + r\omega_j) - (z_i + r\omega_i)\right| \geqslant |z_j - z_i| - 2r \geqslant R(\boldsymbol{z}) - 2r.$$

Combining the two lower bounds gives

$$\left|(z_j + r\omega_j) - (z_i + r\omega_i)\right| \geqslant \max\{r,\ R(\boldsymbol{z}) - 2r\} \geqslant \frac{3}{4}r + \frac{1}{4}\big(R(\boldsymbol{z}) - 2r\big) = \frac{1}{4}\big(R(\boldsymbol{z}) + r\big).$$

Taking the minimum over $1 \leqslant i < j \leqslant n+m$ proves (7.38). $\square$

**Proposition 7.5** (Radial block system along shifted rays)**.** *Let $\mathfrak{C}$ be the domain (6.7), fix a sign pattern*

$$(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = \boldsymbol{\sigma} = (\epsilon_1, \ldots, \epsilon_n, \eta_1, \ldots, \eta_m) = (\sigma_1, \ldots, \sigma_{n+m}) \in \{\pm 1\}^{n+m},$$

*and let*

$$\boldsymbol{z} = (\boldsymbol{t}, \boldsymbol{s}) \in \mathfrak{C}, \qquad R_0 \leqslant R(\boldsymbol{z}) \leqslant 2R_0.$$

*Let $\boldsymbol{\omega}$ be as in Proposition 7.4, and define*

$$\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda) := \boldsymbol{z} + (\lambda - R_0)\boldsymbol{\omega}, \qquad \lambda \in [R_0, \infty).$$

*Let $\mathcal{J}$ be the set of square-free index pairs*

$$(I, A), \qquad I \subseteq \{1, \ldots, n\}, \quad A \subseteq \{1, \ldots, m\},$$

*and write*

$$\mathcal{J}^* := \mathcal{J} \setminus \{(\varnothing, \varnothing)\}.$$

*Order the components indexed by $\mathcal{J}^*$ by increasing $|I| + |A|$.*

*Let $W$ be a holomorphic solution of the conjugated bulk system (7.21)–(7.22). Define*

$$Y_{I,A}(\lambda; \boldsymbol{z}) := \lambda^{|I|+|A|} \partial_{t_I} \partial_{s_A} W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big), \qquad (I, A) \in \mathcal{J},$$

*and write*

$$Y = (Y_0, Y_*), \qquad Y_0(\lambda; \boldsymbol{z}) := W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big),$$

*where $Y_*$ collects the components corresponding to $(I, A) \in \mathcal{J}^*$.*

*Then there exist matrix-valued functions*

$$C_{01}(\lambda; \boldsymbol{z}), \qquad C_{10}(\lambda; \boldsymbol{z}), \qquad C_{11}(\lambda; \boldsymbol{z}),$$

*and a matrix $\Lambda(\lambda;\boldsymbol{z})$ on the non-empty derivative block, whose entries are uniformly bounded for $\lambda \geqslant R_0$, such that*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_0 = \frac{1}{\lambda}C_{01}(\lambda;\boldsymbol{z})Y_*, \tag{7.39}$$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_* = \Lambda(\lambda;\boldsymbol{z})Y_* + \frac{1}{\lambda}C_{10}(\lambda;\boldsymbol{z})Y_0 + \frac{1}{\lambda}C_{11}(\lambda;\boldsymbol{z})Y_*. \tag{7.40}$$

*The matrix $\Lambda(\lambda;\boldsymbol{z})$ is block lower triangular with respect to the ordering by $|I|+|A|$. More precisely, for $(I,A),(J,B)\in\mathcal{J}^*$ with $(I,A)\neq(J,B)$,*

$$\Lambda_{(I,A),(J,B)}(\lambda;\boldsymbol{z}) = 0 \qquad \textit{unless} \qquad |J|+|B| < |I|+|A|. \tag{7.41}$$

*Its diagonal entries are independent of $\lambda$ and $\boldsymbol{z}$, and are given by*

$$\Lambda_{(I,A),(I,A)}(\lambda;\boldsymbol{z}) = \Lambda_{I,A}(\boldsymbol{\omega}) := -2\mathrm{i}\sum_{i\in I}\epsilon_i\omega_i - \beta\mathrm{i}\sum_{a\in A}\eta_a\omega_{n+a}. \tag{7.42}$$

*In particular,*

$$\operatorname{Re}\Lambda_{I,A}(\boldsymbol{\omega}) = 2|I| + \beta|A|, \qquad (I,A)\in\mathcal{J}^*. \tag{7.43}$$

*Proof.* By Proposition 7.4, along the shifted ray $\gamma_{\boldsymbol{z},\boldsymbol{\omega}}$ we have

$$R\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) \gtrsim \lambda, \quad \lambda \geqslant R_0.$$

*Step 1: The equation for $Y_0$.* First we notice that

$$Y_0(\lambda;\boldsymbol{z}) = W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big).$$

By the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_0 = \sum_{i=1}^{n}\omega_i\,\partial_{t_i}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) + \sum_{a=1}^{m}\omega_{n+a}\,\partial_{s_a}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big).$$

Since

$$Y_{\{i\},\varnothing} = \lambda\partial_{t_i}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big), \qquad Y_{\varnothing,\{a\}} = \lambda\partial_{s_a}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big),$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_0 = \frac{1}{\lambda}\left(\sum_{i=1}^{n}\omega_i\,Y_{\{i\},\varnothing} + \sum_{a=1}^{m}\omega_{n+a}\,Y_{\varnothing,\{a\}}\right).$$

This is (7.39). The entries of $C_{01}$ are either 0 or one of the components of $\boldsymbol{\omega}$, hence are uniformly bounded.

*Step 2: The equations for the non-empty square-free derivatives.*
Fix a non-empty square-free pair

$$(I,A)\in\mathcal{J}^*.$$

By the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_{I,A} = \frac{|I|+|A|}{\lambda}Y_{I,A} + \lambda^{|I|+|A|}\left(\sum_{i=1}^{n}\omega_i\,\partial_{t_i}\partial_{t_I}\partial_{s_A}W + \sum_{a=1}^{m}\omega_{n+a}\,\partial_{s_a}\partial_{t_I}\partial_{s_A}W\right)\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big). \tag{7.44}$$

If $i\notin I$, then

$$\partial_{t_i}\partial_{t_I}\partial_{s_A}W = \partial_{t_{I\cup\{i\}}}\partial_{s_A}W$$

is still square-free. Therefore

$$\lambda^{|I|+|A|}\partial_{t_i}\partial_{t_I}\partial_{s_A}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) = \frac{1}{\lambda}Y_{I\cup\{i\},A}(\lambda;\boldsymbol{z}).$$

Similarly, if $a \notin A$, then

$$\lambda^{|I|+|A|}\partial_{s_a}\partial_{t_I}\partial_{s_A}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) = \frac{1}{\lambda}Y_{I,A\cup\{a\}}(\lambda;\boldsymbol{z}).$$

These terms contribute only to the $\lambda^{-1}$-part of (7.40).

Next suppose $i \in I$. Write $I' = I \setminus \{i\}$. Then the relevant derivative is

$$\lambda^{|I|+|A|}\partial_{t_i}^2\partial_{t_{I'}}\partial_{s_A}W. \tag{7.45}$$

We recall the conjugated equations from Proposition 7.3:

$$\partial_{t_i}^2 W = -2\mathrm{i}\epsilon_i\,\partial_{t_i}W - \mathcal{R}_i^{(t)}W, \tag{7.46}$$

$$\partial_{s_a}^2 W = -\mathrm{i}\beta\eta_a\,\partial_{s_a}W - \mathcal{R}_a^{(s)}W. \tag{7.47}$$

Using (7.46), we rewrite (7.45) as

$$\begin{aligned}\lambda^{|I|+|A|}\partial_{t_i}^2\partial_{t_{I'}}\partial_{s_A}W &= \lambda^{|I|+|A|}\partial_{t_{I'}}\partial_{s_A}\left(-2\mathrm{i}\epsilon_i\,\partial_{t_i} - \mathcal{R}_i^{(t)}\right)W\\ &= -2\mathrm{i}\epsilon_i\,\lambda^{|I|+|A|}\partial_{t_I}\partial_{s_A}W - \lambda^{|I|+|A|}\partial_{t_{I'}}\partial_{s_A}\big(\mathcal{R}_i^{(t)}W\big).\end{aligned} \tag{7.48}$$

Multiplying by $\omega_i$, the leading term in (7.48) contributes

$$-2\mathrm{i}\epsilon_i\omega_i\,Y_{I,A}. \tag{7.49}$$

We now estimate the second term on the right-hand side of (7.48). The coefficients

$$\frac{1}{t_i - t_j},\qquad \frac{1}{s_a - s_b},\qquad \frac{1}{t_i - s_a} \tag{7.50}$$

have degree $-1$. From (7.19), the functions $\Omega_i$ and $\Sigma_a$ also have degree $-1$. Hence, in (7.24) and (7.26), the coefficients of first-order derivatives have degree $-1$. Finally, the formulas (7.27) and (7.28) show directly that $V_i^{(t)}$ and $V_a^{(s)}$ are sums of terms of degree $-2$. Therefore, along the shifted ray, for square-free pairs $(J,B)$,

$$\partial_{t_J}\partial_{s_B}\big(\text{a first-order coefficient}\big) = \mathrm{O}\big(\lambda^{-1-|J|-|B|}\big),\quad \partial_{t_J}\partial_{s_B}V_i^{(t)},\quad \partial_{t_J}\partial_{s_B}V_a^{(s)} = \mathrm{O}\big(\lambda^{-2-|J|-|B|}\big). \tag{7.51}$$

Here we used Proposition 7.4, which gives $R(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)) \geqslant \lambda/4$.

Consider first a term obtained from a first-order part of $\mathcal{R}_i^{(t)}$. Suppose that, after applying Leibniz' rule, exactly $r$ of the derivatives in $\partial_{t_{I'}}\partial_{s_A}$ fall on its coefficient. The coefficient is then $\mathrm{O}(\lambda^{-1-r})$, while the differential monomial acting on $W$ has total order $|I|+|A|-r$. If this monomial is square-free, then, after passing to the scaled variables, its coefficient is

$$\lambda^{|I|+|A|}\lambda^{-1-r}\lambda^{-(|I|+|A|-r)} = \lambda^{-1}.$$

Hence it belongs to the $\lambda^{-1}C_{11}Y_*$ part of (7.40).

If the resulting monomial is not square-free, it contains exactly one repeated variable. Suppose first that the repeated variable is $t_j$. We use

$$\partial_{t_j}^2 W = -2\mathrm{i}\epsilon_j\partial_{t_j}W - \mathcal{R}_j^{(t)}W.$$

The principal term $-2\mathrm{i}\epsilon_j\partial_{t_j}W$ removes the repetition without supplying an additional factor $\lambda^{-1}$. The resulting square-free derivative has total order $|I|+|A|-r-1$, and its scaled coefficient is

$$\lambda^{|I|+|A|}\lambda^{-1-r}\lambda^{-(|I|+|A|-r-1)} = 1.$$

Since

$$|I| + |A| - r - 1 < |I| + |A|,$$

this gives a uniformly bounded coupling from a strictly lower-order square-free derivative into the equation for $Y_{I,A}$. The case in which the repeated variable is $s_b$ is identical, using (7.47).

If instead we use the corresponding remainder $\mathcal{R}_j^{(t)} W$, or $\mathcal{R}_b^{(s)} W$ when the repeated variable is $s_b$, every first-order coefficient supplies one additional factor $\lambda^{-1}$, while the total order of the derivative on $W$ decreases by one. Derivatives falling on that new coefficient increase its decay and decrease the remaining derivative order by the same amount. A scalar coefficient supplies two powers of $\lambda^{-1}$, and the resulting differential monomial is square-free. Thus the sum

$$\text{decay order of the coefficient} + \text{total order of the derivative on } W$$

remains equal to $|I| + |A| + 1$. The new differential monomial is either square-free, in which case its scaled coefficient is $\mathrm{O}(\lambda^{-1})$, or it contains again exactly one repeated variable. In the latter case we repeat the same reduction. Each use of a remainder decreases the total order of the derivative on $W$, so the procedure terminates after finitely many steps.

There are only two possible terminal cases. If no principal term is used at the last step, the final coefficient decay order and the final square-free derivative order add up to $|I| + |A| + 1$, and hence the scaled coefficient is $\mathrm{O}(\lambda^{-1})$. If a principal term is used, it removes the unique repetition and lowers the derivative order by one without changing the coefficient decay. The final coefficient decay order and the final derivative order then add up to $|I| + |A|$; the scaled coefficient is $\mathrm{O}(1)$, and the final square-free derivative has strictly smaller total order than $|I| + |A|$.

A term obtained directly from the scalar part $V_i^{(t)}$ is simpler. If exactly $r$ derivatives in $\partial_{t_{I'}} \partial_{s_A}$ fall on the scalar coefficient, then the coefficient is $\mathrm{O}(\lambda^{-2-r})$, while the derivative acting on $W$ has total order $|I| + |A| - 1 - r$. Hence its scaled coefficient is

$$\lambda^{|I|+|A|} \lambda^{-2-r} \lambda^{-(|I|+|A|-1-r)} = \lambda^{-1}.$$

Thus all scalar terms contribute to the $\lambda^{-1} C_{10} Y_0$ or $\lambda^{-1} C_{11} Y_*$ part of (7.40).

Similarly, if $a \in A$ and

$$A' = A \setminus \{a\},$$

then

$$\lambda^{|I|+|A|} \partial_{s_a}^2 \partial_{t_I} \partial_{s_{A'}} W$$

is reduced using (7.47). Its leading contribution, before multiplication by $\omega_{n+a}$, is

$$-\mathrm{i}\beta\eta_a \, \lambda^{|I|+|A|} \partial_{t_I} \partial_{s_A} W.$$

After multiplying by $\omega_{n+a}$, this contributes

$$-\beta \mathrm{i} \eta_a \omega_{n+a} \, Y_{I,A}. \tag{7.52}$$

Again, all remaining terms come from $\mathcal{R}_a^{(s)}$, and the same conclusion as above holds.

We have therefore proved the following dichotomy for every term produced by $\mathcal{R}_i^{(t)}$ or $\mathcal{R}_a^{(s)}$:

- it is a uniformly bounded multiple of a square-free derivative of strictly smaller total order; or
- it carries a factor $\lambda^{-1}$, after which its coefficient is uniformly bounded.

The first class gives the strict block-lower-triangular part of $\Lambda(\lambda; \boldsymbol{z})$, while the second class gives $C_{10}$ and $C_{11}$. Notice also that an $\mathrm{O}(1)$ term cannot involve $Y_0$, because a principal reduction always leaves at least one derivative of $W$.

Combining the leading contributions (7.49) and (7.52) for all $i \in I$ and $a \in A$, the diagonal coefficient of $Y_{I,A}$ is

$$\Lambda_{I,A}(\boldsymbol{\omega}) := -2\mathrm{i}\sum_{i\in I}\epsilon_i\omega_i - \beta\mathrm{i}\sum_{a\in A}\eta_a\omega_{n+a}.$$

This gives (7.40), (7.41), and (7.42). Uniform boundedness follows from (7.51), the fact that there are only finitely many square-free derivatives, and the uniform boundedness of the components of $\boldsymbol{\omega}$.

*Step 3: The real parts of the diagonal entries.*

It remains to compute the real part of the diagonal entries. By the construction of $\boldsymbol{\omega}$ in (7.4),

$$\mathrm{Im}[\omega_i] = \epsilon_i, \qquad 1 \leqslant i \leqslant n, \qquad \mathrm{Im}[\omega_{n+a}] = \eta_a, \qquad 1 \leqslant a \leqslant m.$$

Hence

$$\mathrm{Re}[-2\mathrm{i}\epsilon_i\omega_i] = 2\epsilon_i\,\mathrm{Im}[\omega_i] = 2,$$

and

$$\mathrm{Re}[-\beta\mathrm{i}\eta_a\omega_{n+a}] = \beta\eta_a\,\mathrm{Im}[\omega_{n+a}] = \beta.$$

Therefore $\mathrm{Re}[\Lambda_{I,A}(\boldsymbol{\omega})] = 2|I| + \beta|A|$, which proves (7.43). □

*Proof of the second statement in Theorem 6.2.* Fix first $\boldsymbol{z} = (\boldsymbol{t}, \boldsymbol{s}) \in \mathfrak{C}$, set $R_0 := R(\boldsymbol{z})$, let $\boldsymbol{\omega}$ be as in (7.4), and consider the shifted ray

$$\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda) := \boldsymbol{z} + (\lambda - R_0)\boldsymbol{\omega}, \qquad \lambda \in [R_0, \infty).$$

Along this ray, Proposition 7.5 gives

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_0 = \frac{1}{\lambda}C_{01}(\lambda;\boldsymbol{z})Y_*, \tag{7.53}$$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}Y_* = \Lambda(\lambda;\boldsymbol{z})Y_* + \frac{1}{\lambda}C_{10}(\lambda;\boldsymbol{z})Y_0 + \frac{1}{\lambda}C_{11}(\lambda;\boldsymbol{z})Y_*. \tag{7.54}$$

The matrices in these equations are uniformly bounded.

*Step 1: Exponential decay of the principal backward evolution.* Set

$$c_0 := \min\{2, \beta\} > 0.$$

Recall from (7.42) that the diagonal entries of $\Lambda(\lambda;\boldsymbol{z})$ are independent of $\lambda$ and are given by

$$\Lambda_{I,A}(\boldsymbol{\omega}) = -2\mathrm{i}\sum_{i\in I}\epsilon_i\omega_i - \beta\mathrm{i}\sum_{a\in A}\eta_a\omega_{n+a}.$$

By (7.43),

$$\mathrm{Re}\,\Lambda_{I,A}(\boldsymbol{\omega}) = 2|I| + \beta|A| \geqslant c_0, \qquad (I,A) \in \mathcal{J}^*.$$

Consequently, the backward evolution generated by the diagonal matrix $\Lambda(\boldsymbol{\omega})$ satisfies

$$\left\| e^{-(\sigma-\lambda)\Lambda(\boldsymbol{\omega})} \right\| \leqslant e^{-c_0(\sigma-\lambda)}, \qquad \sigma \geqslant \lambda \geqslant R_0. \tag{7.55}$$

Let $\mathcal{E}(\lambda, \sigma; \boldsymbol{z})$ denote the backward evolution associated with the full triangular matrix $\Lambda(\lambda;\boldsymbol{z})$, namely

$$\partial_\lambda\mathcal{E}(\lambda,\sigma;\boldsymbol{z}) = \Lambda(\lambda;\boldsymbol{z})\mathcal{E}(\lambda,\sigma;\boldsymbol{z}), \qquad \mathcal{E}(\sigma,\sigma;\boldsymbol{z}) = I. \tag{7.56}$$

The matrix

$$\Lambda(\lambda;\boldsymbol{z}) - \Lambda(\boldsymbol{\omega})$$

has vanishing diagonal entries. Moreover, by the triangularity statement in Proposition 7.5,

$$\big(\Lambda(\lambda;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})\big)_{(I,A),(J,B)}=0 \qquad \text{unless} \qquad |J|+|B|<|I|+|A|. \tag{7.57}$$

Therefore every product containing $n+m$ factors of $\Lambda(\lambda;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})$ vanishes, even when the factors are evaluated at different values of $\lambda$ and are separated by diagonal evolution factors.

Repeated variation of constants around the diagonal evolution therefore gives the finite expansion

$$\begin{aligned}\mathcal{E}(\lambda,\sigma;\boldsymbol{z}) &= e^{-(\sigma-\lambda)\Lambda(\boldsymbol{\omega})}+\sum_{r=1}^{n+m-1}(-1)^r\int_{\lambda\leqslant\tau_1\leqslant\cdots\leqslant\tau_r\leqslant\sigma} e^{-(\tau_1-\lambda)\Lambda(\boldsymbol{\omega})}\big(\Lambda(\tau_1;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})\big)\\ &\times e^{-(\tau_2-\tau_1)\Lambda(\boldsymbol{\omega})}\big(\Lambda(\tau_2;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})\big)\cdots e^{-(\tau_r-\tau_{r-1})\Lambda(\boldsymbol{\omega})}\big(\Lambda(\tau_r;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})\big)e^{-(\sigma-\tau_r)\Lambda(\boldsymbol{\omega})}\,\mathrm{d}\tau_1\cdots\mathrm{d}\tau_r.\end{aligned} \tag{7.58}$$

Here, when $r=1$, the intermediate diagonal factors are absent.

After increasing $M$ if necessary, the uniform boundedness of the off-diagonal entries gives

$$\|\Lambda(\lambda;\boldsymbol{z})-\Lambda(\boldsymbol{\omega})\|\leqslant M, \qquad \lambda\geqslant R_0.$$

For the term in (7.58) containing $r$ off-diagonal factors, (7.55) gives

$$\Big\|e^{-(\tau_1-\lambda)\Lambda(\boldsymbol{\omega})}\Big\|\,\Big\|e^{-(\tau_2-\tau_1)\Lambda(\boldsymbol{\omega})}\Big\|\cdots\Big\|e^{-(\sigma-\tau_r)\Lambda(\boldsymbol{\omega})}\Big\|\leqslant e^{-c_0(\tau_1-\lambda)}e^{-c_0(\tau_2-\tau_1)}\cdots e^{-c_0(\sigma-\tau_r)}=e^{-c_0(\sigma-\lambda)}.$$

The ordered integration region $\lambda\leqslant\tau_1\leqslant\cdots\leqslant\tau_r\leqslant\sigma$ has volume $(\sigma-\lambda)^r/r!$. Taking norms in (7.58), we therefore obtain, after enlarging the constant $C$,

$$\|\mathcal{E}(\lambda,\sigma;\boldsymbol{z})\|\leqslant e^{-c_0(\sigma-\lambda)}\sum_{r=0}^{n+m-1}\frac{\big(M(\sigma-\lambda)\big)^r}{r!}\leqslant Ce^{-\frac{c_0}{2}(\sigma-\lambda)}, \qquad \sigma\geqslant\lambda\geqslant R_0. \tag{7.59}$$

*Step 2: Reduction to a Volterra equation.*

We impose the normalization

$$Y_0(\lambda)\to 1, \qquad Y_*(\lambda)\to 0, \qquad \lambda\to\infty.$$

Integrating (7.53) from $\lambda$ to $\infty$, we obtain

$$Y_0(\lambda)=1+(\mathcal{K}Y_*)(\lambda), \qquad (\mathcal{K}Y_*)(\lambda):=-\int_\lambda^\infty\frac{1}{\sigma}C_{01}(\sigma;\boldsymbol{z})Y_*(\sigma)\,\mathrm{d}\sigma. \tag{7.60}$$

Substituting this into (7.54) and using variation of constants gives

$$Y_*=\mathcal{T}(Y_*), \tag{7.61}$$

where

$$(\mathcal{T}Y_*)(\lambda):=-\int_\lambda^\infty\mathcal{E}(\lambda,\sigma;\boldsymbol{z})\frac{1}{\sigma}\Big[C_{10}(\sigma;\boldsymbol{z})\big(1+\mathcal{K}Y_*(\sigma)\big)+C_{11}(\sigma;\boldsymbol{z})Y_*(\sigma)\Big]\,\mathrm{d}\sigma. \tag{7.62}$$

*Step 3: The fixed-point space and the basic estimates.*

Let $\mathcal{B}_{\boldsymbol{z}}$ be the Banach space of continuous functions

$$Y_*:[R_0,\infty)\to\mathbb{C}^{|\mathcal{J}^*|}$$

with norm

$$\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}:=\sup_{\lambda\geqslant R_0}\lambda\|Y_*(\lambda)\|_2.$$

If $Y_* \in \mathcal{B}_{\boldsymbol{z}}$, then by (7.60),

$$|(\mathcal{K}Y_*)(\lambda)| \leqslant M \int_\lambda^\infty \frac{\|Y_*(\sigma)\|_2}{\sigma}\, \mathrm{d}\sigma \leqslant M\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}} \int_\lambda^\infty \sigma^{-2}\, \mathrm{d}\sigma = \frac{M}{\lambda}\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}.$$

Hence

$$|(\mathcal{K}Y_*)(\lambda)| \leqslant \frac{C}{\lambda}\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{7.63}$$

With (7.59), (7.62) and (7.63), we have

$$\begin{aligned}\lambda\|(\mathcal{T}Y_*)(\lambda)\|_2 &\leqslant M\lambda \int_\lambda^\infty \|\mathcal{E}(\lambda,\sigma;\boldsymbol{z})\|\,\frac{1}{\sigma}\Big(1 + |(\mathcal{K}Y_*)(\sigma)| + \|Y_*(\sigma)\|_2\Big)\,\mathrm{d}\sigma\\ &\leqslant CM\lambda \int_\lambda^\infty e^{-c_0(\sigma-\lambda)/2}\Big(\sigma^{-1} + \|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}\sigma^{-2}\Big)\,\mathrm{d}\sigma \leqslant C_0 + C_1R_0^{-1}\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}.\end{aligned}$$

Therefore

$$\|\mathcal{T}Y_*\|_{\mathcal{B}_{\boldsymbol{z}}} \leqslant C_0 + C_1R_0^{-1}\|Y_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{7.64}$$

Similarly, for $Y_*, Z_* \in \mathcal{B}_{\boldsymbol{z}}$,

$$|(\mathcal{K}Y_*)(\lambda) - (\mathcal{K}Z_*)(\lambda)| \leqslant \frac{C}{\lambda}\|Y_* - Z_*\|_{\mathcal{B}_{\boldsymbol{z}}}.$$

Hence

$$\begin{aligned}\lambda\|(\mathcal{T}Y_* - \mathcal{T}Z_*)(\lambda)\|_2 &\leqslant M\lambda \int_\lambda^\infty \|\mathcal{E}(\lambda,\sigma;\boldsymbol{z})\|\,\frac{1}{\sigma}\Big(|(\mathcal{K}Y_*)(\sigma) - (\mathcal{K}Z_*)(\sigma)| + \|Y_*(\sigma) - Z_*(\sigma)\|_2\Big)\,\mathrm{d}\sigma\\ &\leqslant CM\lambda \int_\lambda^\infty e^{-c_0(\sigma-\lambda)/2}\sigma^{-2}\mathrm{d}\sigma\, \|Y_* - Z_*\|_{\mathcal{B}_{\boldsymbol{z}}} \leqslant C_1R_0^{-1}\|Y_* - Z_*\|_{\mathcal{B}_{\boldsymbol{z}}}.\end{aligned}$$

Thus

$$\|\mathcal{T}Y_* - \mathcal{T}Z_*\|_{\mathcal{B}_{\boldsymbol{z}}} \leqslant C_1R_0^{-1}\|Y_* - Z_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{7.65}$$

*Step 4: Existence, asymptotics, and raywise uniqueness.*

If $R_0$ is sufficiently large, then $C_1R_0^{-1} < 1$, so by (7.65) $\mathcal{T}$ is a contraction on $\mathcal{B}_{\boldsymbol{z}}$. Hence (7.61) admits a unique fixed point $Y_* \in \mathcal{B}_{\boldsymbol{z}}$. Defining $Y_0$ by (7.60), we obtain a unique solution $(Y_0, Y_*)$ of (7.53)–(7.54) satisfying

$$Y_0(\lambda) \to 1, \qquad Y_*(\lambda) \to 0, \qquad \lambda \to \infty.$$

Moreover, (7.63) and (7.64) imply

$$Y_0(\lambda) = 1 + \mathrm{O}(\lambda^{-1}), \qquad Y_*(\lambda) = \mathrm{O}(\lambda^{-1}), \qquad \lambda \geqslant R_0.$$

Now define

$$W(\boldsymbol{z}) := Y_0(R_0).$$

By construction,

$$W(\boldsymbol{z}) = 1 + \mathrm{O}(R_0^{-1}) = 1 + \mathrm{O}(R(\boldsymbol{z})^{-1}).$$

For each non-empty square-free pair $(I, A) \in \mathcal{J}^*$, the corresponding component of $Y_*(R_0)$ is

$$Y_{I,A}(R_0) = R(\boldsymbol{z})^{|I|+|A|}\partial_{t_I}\partial_{s_A}W(\boldsymbol{z}).$$

Thus the bound $Y_*(R_0) = \mathrm{O}(R_0^{-1})$ gives

$$\partial_{t_I}\partial_{s_A}W(\boldsymbol{z}) = \mathrm{O}\big(R(\boldsymbol{z})^{-|I|-|A|-1}\big).$$

The same argument applied to the difference of two normalized solutions has no constant term in (7.60). Therefore the corresponding Volterra map is a strict contraction with zero forcing, and the difference vanishes. This proves uniqueness along every admissible shifted ray.

*Step 5: Holomorphic dependence and gluing on the ordered domain.*

Fix $\boldsymbol{z}_0 \in \mathfrak{C}$. After shrinking to a small neighborhood $\mathfrak{U}$ of $\boldsymbol{z}_0$, we may choose a fixed number $R_0$ such that $R_0 \leqslant R(\boldsymbol{z}) \leqslant 2R_0$ for all $\boldsymbol{z} \in \mathfrak{U}$. Use the same vector $\boldsymbol{\omega}$ and the rays $\boldsymbol{z} + (\lambda - R_0)\boldsymbol{\omega}$, $\lambda \geqslant R_0$, throughout $\mathfrak{U}$. Proposition 7.4 gives the required separation uniformly on these rays. Since $R_0$ and $\boldsymbol{\omega}$ are fixed, the coefficients of the Volterra equation depend holomorphically on $\boldsymbol{z}$, and the contraction estimate is uniform. Therefore its fixed point depends holomorphically on $\boldsymbol{z}$.

We have so far solved the first-order system only in the radial direction. To check the remaining coordinate directions, fix one coordinate and compare two ways of differentiating the constructed square-free jet: differentiation with respect to the base point, and the corresponding coordinate equation of the full square-free first-order system. Their difference is the error in that coordinate equation. By flatness, radial differentiation commutes with every coordinate equation, so this error satisfies the homogeneous radial system. The estimates above hold uniformly for nearby base points, and therefore the error has zero normalization at infinity. Raywise uniqueness forces it to vanish. Thus the constructed jet is a holomorphic horizontal section of the full flat first-order system, and its empty component is a holomorphic solution $W^{(\mathfrak{U})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ of the conjugated system.

The local constructions agree on overlaps. Indeed, all neighborhoods use the same vector $\boldsymbol{\omega}$. At a point $\boldsymbol{z}$ in an overlap, restrict both horizontal sections to the geometric ray $\boldsymbol{z} + r\boldsymbol{\omega}$, $r \geqslant 0$. After using the same radial parameter, their scaled square-free jets solve the same radial system and have the same normalization at infinity. The raywise uniqueness from Step 4 therefore shows that their values at $r = 0$ agree. Hence the local solutions glue to a holomorphic function $W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ on all of $\mathfrak{C}$, and the estimates (6.9)–(6.10) hold there with constants depending only on $\beta, n, m$. Finally, $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} := \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ solves the original system.

*Step 6: Linear independence.*

From the explicit formula for $\Phi_{\epsilon,\eta}$, we have

$$\partial_{t_i} \log \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \mathrm{i}\epsilon_i + \mathrm{O}(D^{-1}), \qquad \partial_{s_a} \log \Phi_{\epsilon,\eta} = \mathrm{i}\frac{\beta}{2}\eta_a + \mathrm{O}(D^{-1}).$$

For a square-free pair $(I, A)$, the explicit formula for the branch factor and the estimates for $W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ give, uniformly on $\mathfrak{C}$,

$$\Phi^{-1}_{\boldsymbol{\epsilon},\boldsymbol{\eta}} \partial_{t_I} \partial_{s_A} F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \mathrm{i}^{|I|+|A|} \left(\frac{\beta}{2}\right)^{|A|} \prod_{i \in I} \epsilon_i \prod_{a \in A} \eta_a + \mathrm{O}(D^{-1}). \tag{7.66}$$

After dividing each row by its nonzero prefactor, the leading matrix is the character table

$$\left(\prod_{i \in I} \epsilon_i \prod_{a \in A} \eta_a\right)_{(I,A),(\boldsymbol{\epsilon},\boldsymbol{\eta})}$$

of $\{\pm 1\}^{n+m}$. In particular, its rows satisfy the orthogonality relations

$$\sum_{\epsilon,\eta} \left(\prod_{i \in I} \epsilon_i \prod_{a \in A} \eta_a\right) \left(\prod_{j \in J} \epsilon_j \prod_{b \in B} \eta_b\right) = 2^{n+m} \mathbf{1}_{\{(I,A)=(J,B)\}},$$

and therefore the leading matrix is invertible. Increasing $D$, if necessary, the matrix in (7.66) is also invertible. Therefore the $2^{n+m}$ solutions $F_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ are linearly independent. □

# 8 Solutions of the edge deformed CMS operators

The proof of Theorem 6.5 is parallel to the bulk case. We prove the first statement of Theorem 6.5, namely that the solution space has dimension $2^{n+m}$, in Section 8.1. We derive the conjugated equations in Section 8.2. We then prove the second statement, namely the construction of solutions with prescribed leading behavior, in Section 8.3.

## 8.1 Local holonomic rank

For $1 \leqslant i \leqslant n$ and $1 \leqslant a \leqslant m$, define

$$\mathsf{T}_i := \partial_{t_i}^2 - t_i + \frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i}-\partial_{t_j}}{t_i-t_j} - \sum_{a=1}^{m}\frac{\partial_{t_i}+(2/\beta)\partial_{s_a}}{t_i-s_a}, \tag{8.1}$$

$$\mathsf{S}_a := \partial_{s_a}^2 - \frac{\beta^2}{4}s_a + \frac{\beta}{2}\sum_{b\neq a}\frac{\partial_{s_a}-\partial_{s_b}}{s_a-s_b} - \sum_{i=1}^{n}\frac{\partial_{s_a}+(\beta/2)\partial_{t_i}}{s_a-t_i}. \tag{8.2}$$

Then the edge deformed CMS operators (5.18)–(5.19) can be written as

$$\mathsf{T}_i F = 0, \qquad 1 \leqslant i \leqslant n, \qquad \mathsf{S}_a F = 0, \qquad 1 \leqslant a \leqslant m. \tag{8.3}$$

Using the following edge commutator identities as input, the first statement of Theorem 6.5 can be proven in the same way as the bulk case Section 7.1. So we omit its proof.

**Proposition 8.1** (Edge commutator identities)**.** *For the operators $\mathsf{T}_i, \mathsf{S}_a$ defined in (8.1)–(8.2), the following commutator identities hold:*

$$[\mathsf{T}_i,\mathsf{T}_j] = \frac{4}{\beta(t_i-t_j)^2}(\mathsf{T}_j-\mathsf{T}_i), \qquad 1 \leqslant i \neq j \leqslant n, \tag{8.4}$$

$$[\mathsf{S}_a,\mathsf{S}_b] = \frac{\beta}{(s_a-s_b)^2}(\mathsf{S}_b-\mathsf{S}_a), \qquad 1 \leqslant a \neq b \leqslant m, \tag{8.5}$$

$$[\mathsf{T}_i,\mathsf{S}_a] = \frac{4}{\beta(t_i-s_a)^2}\mathsf{S}_a - \frac{\beta}{(t_i-s_a)^2}\mathsf{T}_i, \qquad 1 \leqslant i \leqslant n,\ 1 \leqslant a \leqslant m. \tag{8.6}$$

*Proof.* This is proved by the same local calculation as for the bulk operators Proposition 7.1. One decomposes $\mathsf{T}_i$ and $\mathsf{S}_a$ into the basic pairwise pieces

$$\partial_{t_i}^2 - t_i, \qquad \partial_{s_a}^2 - \frac{\beta^2}{4}s_a, \qquad \frac{2}{\beta}\frac{\partial_{t_i}-\partial_{t_j}}{t_i-t_j}, \qquad \frac{\beta}{2}\frac{\partial_{s_a}-\partial_{s_b}}{s_a-s_b}, \qquad \frac{\partial_{t_i}+(2/\beta)\partial_{s_a}}{t_i-s_a},$$

checks the corresponding two-body and three-body commutators, and then sums the local identities. Compared with the bulk case, the linear potentials $-t_i$ and $-\beta^2 s_a/4$ contribute exactly the scalar terms needed to replace the bulk operators by the edge operators on the right-hand side. They do not change the coefficients of the commutator identities. □

## 8.2 The conjugated system

We recall the (rescaled) Airy functions $\phi_\pm, \psi_\pm$ and their asymptotics from (6.34), (6.35) and (6.36). We have

$$\phi_\pm''(z) = z\,\phi_\pm(z), \qquad q_\pm(z) := \frac{\phi_\pm'(z)}{\phi_\pm(z)} = \pm\sqrt{z} - \frac{1}{4z} + \mathrm{O}(z^{-5/2}),$$

and

$$\psi_\pm''(s) = \frac{\beta^2}{4}s\,\psi_\pm(s), \qquad p_\pm(s) := \frac{\psi_\pm'(s)}{\psi_\pm(s)} = \pm\frac{\beta}{2}\sqrt{s} - \frac{1}{4s} + \mathrm{O}(s^{-5/2}).$$

Now fix sign vectors

$$\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n) \in \{\pm 1\}^n, \qquad \boldsymbol{\eta} = (\eta_1, \ldots, \eta_m) \in \{\pm 1\}^m,$$

and write

$$q_i := q_{\epsilon_i}(t_i), \qquad p_a := p_{\eta_a}(s_a). \tag{8.7}$$

Then $q_i$ and $p_a$ satisfy the Riccati identity

$$q_i' + q_i^2 = t_i, \quad p_a' + p_a^2 = \frac{\beta^2}{4} s_a \tag{8.8}$$

It is convenient to separate off the leading square-root terms:

$$r_i := q_i - \epsilon_i \sqrt{t_i} = -\frac{1}{4t_i} + \mathrm{O}(t_i^{-5/2}), \qquad v_a := p_a - \frac{\beta}{2}\eta_a \sqrt{s_a} = -\frac{1}{4s_a} + \mathrm{O}(s_a^{-5/2}). \tag{8.9}$$

We also recall the branch factor $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s})$ from (6.39)

$$\begin{aligned} \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}(\boldsymbol{t}, \boldsymbol{s}) := \left(\prod_{i=1}^n \phi_{\epsilon_i}(t_i)\right)\left(\prod_{a=1}^m \psi_{\eta_a}(s_a)\right) \prod_{1 \leqslant i < j \leqslant n} (\epsilon_i \sqrt{t_i} + \epsilon_j \sqrt{t_j})^{-2/\beta} \\ \times \prod_{1 \leqslant a < b \leqslant m} (\eta_a \sqrt{s_a} + \eta_b \sqrt{s_b})^{-\beta/2} \prod_{i=1}^n \prod_{a=1}^m (\epsilon_i \sqrt{t_i} - \eta_a \sqrt{s_a}). \end{aligned} \tag{8.10}$$

We also introduce

$$\begin{aligned} \Omega_i &:= -\frac{\epsilon_i}{\beta\sqrt{t_i}} \sum_{j \neq i} \frac{1}{\epsilon_i \sqrt{t_i} + \epsilon_j \sqrt{t_j}} + \frac{\epsilon_i}{2\sqrt{t_i}} \sum_{a=1}^m \frac{1}{\epsilon_i \sqrt{t_i} - \eta_a \sqrt{s_a}}, \\ \Sigma_a &:= -\frac{\beta \eta_a}{4\sqrt{s_a}} \sum_{b \neq a} \frac{1}{\eta_a \sqrt{s_a} + \eta_b \sqrt{s_b}} - \frac{\eta_a}{2\sqrt{s_a}} \sum_{i=1}^n \frac{1}{\epsilon_i \sqrt{t_i} - \eta_a \sqrt{s_a}}. \end{aligned} \tag{8.11}$$

**Proposition 8.2** (Conjugated operator family)**.** *Let $F = \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W$. Then the original edge system (8.3),*

$$\mathsf{T}_i F = 0, \qquad 1 \leqslant i \leqslant n, \qquad \mathsf{S}_a F = 0, \qquad 1 \leqslant a \leqslant m,$$

*is equivalent to*

$$(\mathsf{L}_i^{(t)} + \mathsf{R}_i^{(t)}) W = 0, \qquad i = 1, \ldots, n, \tag{8.12}$$

$$(\mathsf{L}_a^{(s)} + \mathsf{R}_a^{(s)}) W = 0, \qquad a = 1, \ldots, m, \tag{8.13}$$

*where*

$$\mathsf{L}_i^{(t)} = \partial_{t_i}^2 + 2q_i\, \partial_{t_i}, \tag{8.14}$$

$$\mathsf{R}_i^{(t)} = 2\Omega_i\, \partial_{t_i} + \frac{2}{\beta} \sum_{j \neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^m \frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a} + V_i^{(t)}, \tag{8.15}$$

$$\mathsf{L}_a^{(s)} = \partial_{s_a}^2 + 2p_a\, \partial_{s_a}, \tag{8.16}$$

$$\mathsf{R}_a^{(s)} = 2\Sigma_a\, \partial_{s_a} + \frac{\beta}{2} \sum_{b \neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} - \sum_{i=1}^n \frac{\partial_{s_a} + (\beta/2)\partial_{t_i}}{s_a - t_i} + V_a^{(s)}. \tag{8.17}$$

*Here $q_i, p_a, r_i, v_a$ are defined in (8.7) and (8.9). The scalar terms are*

$$V_i^{(t)} = \partial_{t_i}\Omega_i + \Omega_i^2 + 2r_i\Omega_i + \frac{2}{\beta}\sum_{j\neq i}\frac{r_i - r_j + \Omega_i - \Omega_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{r_i + (2/\beta)v_a + \Omega_i + (2/\beta)\Sigma_a}{t_i - s_a}, \tag{8.18}$$

$$V_a^{(s)} = \partial_{s_a}\Sigma_a + \Sigma_a^2 + 2v_a\Sigma_a + \frac{\beta}{2}\sum_{b\neq a}\frac{v_a - v_b + \Sigma_a - \Sigma_b}{s_a - s_b} - \sum_{i=1}^{n}\frac{v_a + (\beta/2)r_i + \Sigma_a + (\beta/2)\Omega_i}{s_a - t_i}. \tag{8.19}$$

*Proof.* Let $\Phi := \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ be as in (8.10). Since $\Phi$ is nowhere vanishing on $\mathsf{X}$, the equations

$$\mathsf{T}_i F = 0, \qquad \mathsf{S}_a F = 0, \qquad F = \Phi W,$$

are equivalent to

$$\widetilde{\mathsf{T}}_i W = 0, \qquad \widetilde{\mathsf{S}}_a W = 0,$$

where

$$\widetilde{\mathsf{T}}_i := \Phi^{-1}\mathsf{T}_i\Phi, \qquad \widetilde{\mathsf{S}}_a := \Phi^{-1}\mathsf{S}_a\Phi.$$

We recall $\Omega_i$ and $\Sigma_a$ from (8.11) and set

$$\mu_i := \partial_{t_i}\log\Phi = q_i + \Omega_i, \qquad \nu_a := \partial_{s_a}\log\Phi = p_a + \Sigma_a. \tag{8.20}$$

Then

$$\Phi^{-1}\partial_{t_i}\Phi = \partial_{t_i} + \mu_i, \qquad \Phi^{-1}\partial_{s_a}\Phi = \partial_{s_a} + \nu_a, \tag{8.21}$$

and

$$\begin{aligned} \Phi^{-1}\partial_{t_i}^2\Phi &= (\partial_{t_i} + \mu_i)^2 = \partial_{t_i}^2 + 2\mu_i\partial_{t_i} + \partial_{t_i}\mu_i + \mu_i^2, \\ \Phi^{-1}\partial_{s_a}^2\Phi &= (\partial_{s_a} + \nu_a)^2 = \partial_{s_a}^2 + 2\nu_a\partial_{s_a} + \partial_{s_a}\nu_a + \nu_a^2. \end{aligned} \tag{8.22}$$

We first compute $\widetilde{\mathsf{T}}_i$. Substituting (8.21) and (8.22) into $\mathsf{T}_i$, recalled in (8.1), we obtain

$$\widetilde{\mathsf{T}}_i = (\partial_{t_i} + \mu_i)^2 - t_i + \frac{2}{\beta}\sum_{j\neq i}\frac{(\partial_{t_i} + \mu_i) - (\partial_{t_j} + \mu_j)}{t_i - t_j} - \sum_{a=1}^{m}\frac{(\partial_{t_i} + \mu_i) + (2/\beta)(\partial_{s_a} + \nu_a)}{t_i - s_a}. \tag{8.23}$$

Using $\mu_i = q_i + \Omega_i$ from (8.20), the derivative terms in (8.23) are

$$\partial_{t_i}^2 + 2q_i\partial_{t_i} + 2\Omega_i\partial_{t_i} + \frac{2}{\beta}\sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} - \sum_{a=1}^{m}\frac{\partial_{t_i} + (2/\beta)\partial_{s_a}}{t_i - s_a}. \tag{8.24}$$

Thus the first two terms give $\mathsf{L}_i^{(t)}$, and the remaining first-order terms give the differential part of $\mathsf{R}_i^{(t)}$.

The scalar part is

$$\partial_{t_i}\Omega_i + q_i' + q_i^2 - t_i + \Omega_i^2 + 2q_i\Omega_i + \frac{2}{\beta}\sum_{j\neq i}\frac{q_i - q_j + \Omega_i - \Omega_j}{t_i - t_j} - \sum_{a=1}^{m}\frac{q_i + (2/\beta)p_a + \Omega_i + (2/\beta)\Sigma_a}{t_i - s_a}. \tag{8.25}$$

Using (8.8), the term $q_i' + q_i^2 - t_i$ vanishes. Using the decomposition (8.9), the scalar part (8.25) becomes

$$V_i^{(t)} + 2\epsilon_i\sqrt{t_i}\,\Omega_i + \frac{2}{\beta}\sum_{j\neq i}\frac{\epsilon_i\sqrt{t_i} - \epsilon_j\sqrt{t_j}}{t_i - t_j} - \sum_{a=1}^{m}\frac{\epsilon_i\sqrt{t_i} + \eta_a\sqrt{s_a}}{t_i - s_a}. \tag{8.26}$$

For the remaining square-root terms, use

$$\frac{\epsilon_i\sqrt{t_i}-\epsilon_j\sqrt{t_j}}{t_i-t_j}=\frac{1}{\epsilon_i\sqrt{t_i}+\epsilon_j\sqrt{t_j}},\qquad \frac{\epsilon_i\sqrt{t_i}+\eta_a\sqrt{s_a}}{t_i-s_a}=\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}.$$

Substituting the definition of $\Omega_i$ from (8.11), we get

$$\begin{aligned}&2\epsilon_i\sqrt{t_i}\,\Omega_i+\frac{2}{\beta}\sum_{j\neq i}\frac{\epsilon_i\sqrt{t_i}-\epsilon_j\sqrt{t_j}}{t_i-t_j}-\sum_{a=1}^{m}\frac{\epsilon_i\sqrt{t_i}+\eta_a\sqrt{s_a}}{t_i-s_a}\\&=-\frac{2}{\beta}\sum_{j\neq i}\frac{1}{\epsilon_i\sqrt{t_i}+\epsilon_j\sqrt{t_j}}+\sum_{a=1}^{m}\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}+\frac{2}{\beta}\sum_{j\neq i}\frac{1}{\epsilon_i\sqrt{t_i}+\epsilon_j\sqrt{t_j}}-\sum_{a=1}^{m}\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}=0.\end{aligned}$$

Hence the scalar part is exactly $V_i^{(t)}$, and (8.12) follows by combining (8.23) and (8.24).

The computation for $\widetilde{\mathsf{S}}_a$ is analogous. Substituting (8.21) and (8.22) into $\mathsf{S}_a$, recalled in (8.2), we obtain

$$\widetilde{\mathsf{S}}_a=(\partial_{s_a}+\nu_a)^2-\frac{\beta^2}{4}s_a+\frac{\beta}{2}\sum_{b\neq a}\frac{(\partial_{s_a}+\nu_a)-(\partial_{s_b}+\nu_b)}{s_a-s_b}-\sum_{i=1}^{n}\frac{(\partial_{s_a}+\nu_a)+(\beta/2)(\partial_{t_i}+\mu_i)}{s_a-t_i}. \tag{8.27}$$

Using $\nu_a=p_a+\Sigma_a$ from (8.20), the derivative terms in (8.27) are

$$\partial_{s_a}^2+2p_a\partial_{s_a}+2\Sigma_a\partial_{s_a}+\frac{\beta}{2}\sum_{b\neq a}\frac{\partial_{s_a}-\partial_{s_b}}{s_a-s_b}-\sum_{i=1}^{n}\frac{\partial_{s_a}+(\beta/2)\partial_{t_i}}{s_a-t_i}. \tag{8.28}$$

Thus the first two terms give $\mathsf{L}_a^{(s)}$, and the remaining first-order terms give the differential part of $\mathsf{R}_a^{(s)}$.

The scalar part is

$$\partial_{s_a}\Sigma_a+p_a'+p_a^2-\frac{\beta^2}{4}s_a+\Sigma_a^2+2p_a\Sigma_a+\frac{\beta}{2}\sum_{b\neq a}\frac{p_a-p_b+\Sigma_a-\Sigma_b}{s_a-s_b}-\sum_{i=1}^{n}\frac{p_a+(\beta/2)q_i+\Sigma_a+(\beta/2)\Omega_i}{s_a-t_i}. \tag{8.29}$$

Using (8.8), the term $p_a'+p_a^2-\beta^2 s_a/4$ vanishes. Using the decomposition (8.9), the scalar part (8.29) becomes

$$V_a^{(s)}+\beta\eta_a\sqrt{s_a}\,\Sigma_a+\frac{\beta^2}{4}\sum_{b\neq a}\frac{\eta_a\sqrt{s_a}-\eta_b\sqrt{s_b}}{s_a-s_b}-\frac{\beta}{2}\sum_{i=1}^{n}\frac{\eta_a\sqrt{s_a}+\epsilon_i\sqrt{t_i}}{s_a-t_i}. \tag{8.30}$$

For the remaining square-root terms, use

$$\frac{\eta_a\sqrt{s_a}-\eta_b\sqrt{s_b}}{s_a-s_b}=\frac{1}{\eta_a\sqrt{s_a}+\eta_b\sqrt{s_b}},\qquad \frac{\eta_a\sqrt{s_a}+\epsilon_i\sqrt{t_i}}{s_a-t_i}=-\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}.$$

Substituting the definition of $\Sigma_a$ from (8.11), we get

$$\begin{aligned}&\beta\eta_a\sqrt{s_a}\,\Sigma_a+\frac{\beta^2}{4}\sum_{b\neq a}\frac{\eta_a\sqrt{s_a}-\eta_b\sqrt{s_b}}{s_a-s_b}-\frac{\beta}{2}\sum_{i=1}^{n}\frac{\eta_a\sqrt{s_a}+\epsilon_i\sqrt{t_i}}{s_a-t_i}\\&=-\frac{\beta^2}{4}\sum_{b\neq a}\frac{1}{\eta_a\sqrt{s_a}+\eta_b\sqrt{s_b}}-\frac{\beta}{2}\sum_{i=1}^{n}\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}+\frac{\beta^2}{4}\sum_{b\neq a}\frac{1}{\eta_a\sqrt{s_a}+\eta_b\sqrt{s_b}}+\frac{\beta}{2}\sum_{i=1}^{n}\frac{1}{\epsilon_i\sqrt{t_i}-\eta_a\sqrt{s_a}}=0.\end{aligned}$$

Hence the scalar part is exactly $V_a^{(s)}$, and (8.13) follows by combining (8.27) and (8.28). This proves the proposition. $\square$

**Lemma 8.3** (Size of the conjugated coefficients)**.** *Let $\mathsf{C}$ be a domain of the form (6.41), with $D$ sufficiently large, let $(\boldsymbol{t},\boldsymbol{s}) \in \mathsf{C}$, and set $R = \mathsf{R}(\boldsymbol{t},\boldsymbol{s})$ recall from (6.40). Then, uniformly on $\mathsf{C}$,*

$$\Omega_i = \mathrm{O}(R^{-1}), \qquad \Sigma_a = \mathrm{O}(R^{-1}), \tag{8.31}$$

$$q_i - \epsilon_i\sqrt{t_i} = \mathrm{O}(R^{-1}), \qquad p_a - \eta_a\frac{\beta}{2}\sqrt{s_a} = \mathrm{O}(R^{-1}), \tag{8.32}$$

$$V_i^{(t)} = \mathrm{O}(R^{-2}), \qquad V_a^{(s)} = \mathrm{O}(R^{-2}). \tag{8.33}$$

*Moreover, if $\mathsf{D}$ is any differential monomial of total order $\ell$, then*

$$\mathsf{D}\Omega_i, \quad \mathsf{D}\Sigma_a, \quad \mathsf{D}\big(q_i - \epsilon_i\sqrt{t_i}\big), \quad \mathsf{D}\left(p_a - \eta_a\frac{\beta}{2}\sqrt{s_a}\right) = \mathrm{O}\big(R^{-1-\ell}\big), \tag{8.34}$$

$$\mathsf{D}V_i^{(t)}, \quad \mathsf{D}V_a^{(s)} = \mathrm{O}\big(R^{-2-\ell}\big). \tag{8.35}$$

*The same all-order estimates hold for every complete first-order coefficient in $\mathsf{R}_i^{(t)}$ and $\mathsf{R}_a^{(s)}$, with $R^{-1-\ell}$ on the right-hand side. The implicit constants may depend on $\beta, n, m$ and on the differential monomial $\mathsf{D}$, but are uniform over $\mathsf{C}$.*

*Proof.* We begin with the zeroth-order estimates. We use the elementary inequality

$$|u(u \pm v)| \gtrsim \min\big(|u|^2, |u^2 - v^2|\big), \qquad u, v \in \mathbb{C}. \tag{8.36}$$

Indeed, if $|v| \geqslant 2|u|$, then

$$|u \pm v| \geqslant |v| - |u| \geqslant |u|,$$

so the left-hand side of (8.36) is at least $|u|^2$. If $|v| \leqslant 2|u|$, then

$$|u \mp v| \leqslant |u| + |v| \leqslant 3|u|,$$

and therefore

$$|u^2 - v^2| = |u+v|\,|u-v| \leqslant 3|u|\,|u \pm v|.$$

This proves (8.36).

Applying (8.36) to the terms in (8.11), we obtain, for $j \neq i$,

$$\big|\sqrt{t_i}\big(\epsilon_i\sqrt{t_i} + \epsilon_j\sqrt{t_j}\big)\big| \gtrsim \min\big(|t_i|, |t_i - t_j|\big) \gtrsim R.$$

Similarly, for $b \neq a$,

$$\big|\sqrt{s_a}\big(\eta_a\sqrt{s_a} + \eta_b\sqrt{s_b}\big)\big| \gtrsim \min\big(|s_a|, |s_a - s_b|\big) \gtrsim R.$$

For the mixed terms,

$$\big|\sqrt{t_i}\big(\epsilon_i\sqrt{t_i} - \eta_a\sqrt{s_a}\big)\big| \gtrsim \min\big(|t_i|, |t_i - s_a|\big) \gtrsim R,$$

and

$$\big|\sqrt{s_a}\big(\epsilon_i\sqrt{t_i} - \eta_a\sqrt{s_a}\big)\big| \gtrsim \min\big(|s_a|, |t_i - s_a|\big) \gtrsim R.$$

Thus every complete summand in (8.11) is $\mathrm{O}(R^{-1})$. Since the numbers of summands depend only on $n$ and $m$, this proves (8.31).

The Airy asymptotics (8.9) give

$$q_i - \epsilon_i\sqrt{t_i} = -\frac{1}{4t_i} + \mathrm{O}(t_i^{-5/2}),$$

and

$$p_a - \eta_a\frac{\beta}{2}\sqrt{s_a} = -\frac{1}{4s_a} + \mathrm{O}(s_a^{-5/2}),$$

uniformly in the sectors under consideration. Since $|t_i| \geqslant R$, $|s_a| \geqslant R$, and $R$ is sufficiently large, this proves (8.32).

We next establish the derivative estimates. The angular restrictions in (6.41) place every coordinate in a sector having a fixed positive angular distance from the real axis and from the branch cut of the principal square root. Consequently, there exists a sufficiently small constant $c > 0$ such that the polydisc obtained by moving each coordinate by at most $4cR$ remains in the same half-plane and in a fixed slightly larger sector on which the square-root branches and the Airy asymptotics are valid.

After decreasing $c$ if necessary, every coordinate modulus and every pairwise difference on this polydisc remain bounded below by a constant multiple of $R$. Indeed, the definition of $R$ gives

$$|t_i|,\ |s_a|,\ |t_i - t_j|,\ |t_i - s_a|,\ |s_a - s_b| \geqslant R$$

at the center, while moving each coordinate by at most $4cR$ changes a coordinate modulus by at most $4cR$ and a pairwise difference by at most $8cR$.

It follows that the zeroth-order estimates proved above hold uniformly throughout this polydisc. In particular,

$$\Omega_i, \quad \Sigma_a, \quad q_i - \epsilon_i\sqrt{t_i}, \quad p_a - \eta_a \frac{\beta}{2}\sqrt{s_a} = \mathrm{O}(R^{-1})$$

there. These functions are holomorphic on the polydisc. Applying the multivariable Cauchy estimates on concentric smaller polydiscs therefore gives

$$\mathsf{D}\Omega_i, \quad \mathsf{D}\Sigma_a, \quad \mathsf{D}\big(q_i - \epsilon_i\sqrt{t_i}\big), \quad \mathsf{D}\left(p_a - \eta_a \frac{\beta}{2}\sqrt{s_a}\right) = \mathrm{O}(R^{-1-\ell}),$$

which is (8.34).

We now use (8.18) and (8.19). Recall that

$$r_i = q_i - \epsilon_i\sqrt{t_i}, \qquad v_a = p_a - \eta_a \frac{\beta}{2}\sqrt{s_a}.$$

On a smaller concentric polydisc,

$$r_i,\ v_a,\ \Omega_i,\ \Sigma_a = \mathrm{O}(R^{-1}),$$

and (8.34) gives

$$\partial_{t_i}\Omega_i = \mathrm{O}(R^{-2}), \qquad \partial_{s_a}\Sigma_a = \mathrm{O}(R^{-2}).$$

Hence

$$\partial_{t_i}\Omega_i, \quad \Omega_i^2, \quad r_i\Omega_i = \mathrm{O}(R^{-2}),$$

and, for $j \neq i$,

$$\frac{r_i - r_j + \Omega_i - \Omega_j}{t_i - t_j} = \mathrm{O}(R^{-2}).$$

Likewise,

$$\frac{r_i + (2/\beta)v_a + \Omega_i + (2/\beta)\Sigma_a}{t_i - s_a} = \mathrm{O}(R^{-2}).$$

Substitution into (8.18) gives $V_i^{(t)} = \mathrm{O}(R^{-2})$. The same estimates applied to (8.19) give $V_a^{(s)} = \mathrm{O}(R^{-2})$. This proves (8.33), uniformly on the smaller polydisc.

The functions $V_i^{(t)}$ and $V_a^{(s)}$ are holomorphic on that polydisc. A further application of the multivariable Cauchy estimates therefore yields

$$\mathsf{D}V_i^{(t)}, \quad \mathsf{D}V_a^{(s)} = \mathrm{O}(R^{-2-\ell}),$$

which proves (8.35).

Finally, after collecting the coefficient of each first-order derivative in (8.15) and (8.17), every such coefficient is a finite sum of terms of the form

$$\Omega_i, \qquad \Sigma_a, \qquad \frac{1}{t_i - t_j}, \qquad \frac{1}{t_i - s_a}, \qquad \frac{1}{s_a - s_b},$$

multiplied by constants depending only on $\beta$. Each complete coefficient is therefore $\mathrm{O}(R^{-1})$ on the same polydisc and is holomorphic there. Cauchy's estimates give the asserted $\mathrm{O}(R^{-1-\ell})$ bound for all of its derivatives. ☐

## 8.3 Solutions attached to the branch factors

**Proposition 8.4.** *Let $\mathsf{C} = \mathsf{C}(\boldsymbol{\tau}, D)$ be the domain (6.41), and let*

$$\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{n+m}) \in \{\pm 1\}^{n+m}.$$

*There exist a vector $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{\tau}, \boldsymbol{\sigma})$, independent of the base point $\boldsymbol{z} \in \mathsf{C}$, and a constant $c > 0$, depending only on $n+m$, such that, for every $\boldsymbol{z} \in \mathsf{C}$ and every $r \geqslant 0$, we have $\boldsymbol{z} + r\boldsymbol{\omega} \in \mathsf{C}$ and*

$$\mathsf{R}(\boldsymbol{z} + r\boldsymbol{\omega}) \geqslant c\big(\mathsf{R}(\boldsymbol{z}) + r\big), \tag{8.37}$$

$$\left|\arg\left(-\sigma_j \omega_j \sqrt{z_j + r\omega_j}\right)\right| \leqslant \frac{23\pi}{48}, \qquad 1 \leqslant j \leqslant n+m, \tag{8.38}$$

$$\operatorname{Re}\left[-\sigma_j \omega_j \sqrt{z_j + r\omega_j}\right] \geqslant c\big(\mathsf{R}(\boldsymbol{z}) + r\big)^{1/2}, \qquad 1 \leqslant j \leqslant n+m. \tag{8.39}$$

*Here the square root is the principal square root.*

*Proof.* Choose $K > 1$, depending only on $n+m$, sufficiently large, and define

$$\omega_j := K^{n+m+1-j} \exp\left[\mathrm{i}\tau_j \left(\frac{\pi}{3} + \sigma_j \frac{\pi}{16}\right)\right], \qquad 1 \leqslant j \leqslant n+m. \tag{8.40}$$

For every $j$,

$$\left|\arg \omega_j - \frac{\pi}{3}\tau_j\right| \leqslant \frac{\pi}{16}.$$

Thus $z_j$ and $\omega_j$ belong to the same convex cone appearing in the definition of $\mathsf{C}$, and hence $(z_j + r\omega_j)_{1 \leqslant j \leqslant n+m}$ remains in that cone for every $r \geqslant 0$.

Moreover,

$$\operatorname{Re} \omega_j \geqslant \cos\left(\frac{19\pi}{48}\right) K^{n+m+1-j}, \quad |\omega_j| \leqslant K^{n+m+1-j}.$$

Consequently, after taking $K$ sufficiently large,

$$\operatorname{Re} \omega_j - \operatorname{Re} \omega_{j+1} \geqslant K^{n+m-j} \left[K \cos\left(\frac{19\pi}{48}\right) - 1\right] > 0, \qquad 1 \leqslant j < n+m,$$

and $\operatorname{Re} \omega_{n+m} > 0$. It follows that

$$\operatorname{Re}(z_j + r\omega_j) - \operatorname{Re}(z_{j+1} + r\omega_{j+1}) > D, \quad 1 \leqslant j < n+m, \quad \operatorname{Re}(z_{n+m} + r\omega_{n+m}) > D.$$

Thus all the inequalities defining $\mathsf{C}$ are preserved along the ray, which proves $\boldsymbol{z} + r\boldsymbol{\omega} \in \mathsf{C}$.

We next prove (8.37). Since $z_j$ and $\omega_j$ lie in the same cone of opening $\pi/6$,

$$|z_j + r\omega_j| \geqslant c\big(|z_j| + r|\omega_j|\big) \geqslant c\big(\mathsf{R}(\boldsymbol{z}) + r\big).$$

For $1 \leqslant i < j \leqslant n+m$, the geometric growth in (8.40) gives

$$\omega_i - \omega_j = \omega_i\big(1 + \mathrm{O}(K^{-1})\big),$$

uniformly in $i<j$. Increasing $K$, if necessary, we may therefore assume that

$$|\arg(\omega_i-\omega_j)|<\frac{11\pi}{24},\qquad |\omega_i-\omega_j|\geqslant 1.$$

On the other hand, the ordering inequalities defining $\mathsf{C}$ give $\operatorname{Re}(z_i-z_j)>0$, so the angle between $z_i-z_j$ and $\omega_i-\omega_j$ is at most $23\pi/24$. Hence

$$|(z_i-z_j)+r(\omega_i-\omega_j)|\geqslant \sin\Bigl(\frac{\pi}{48}\Bigr)\,(|z_i-z_j|+r|\omega_i-\omega_j|)\geqslant c\bigl(\mathsf{R}(\boldsymbol{z})+r\bigr).$$

Taking the minimum over all coordinates and all pairs proves (8.37).

It remains to prove the angular estimate. By $\boldsymbol{z}+r\boldsymbol{\omega}\in\mathsf{C}$,

$$\frac{\pi}{4}<\tau_j\arg(z_j+r\omega_j)<\frac{5\pi}{12}.$$

Since the square root is the principal square root, its argument is one-half of the argument of $z_j+r\omega_j$. If $\sigma_j=-1$, then

$$\arg\left(-\sigma_j\omega_j\sqrt{z_j+r\omega_j}\right)=\tau_j\left(\frac{\pi}{3}-\frac{\pi}{16}+\frac{\tau_j}{2}\arg(z_j+r\omega_j)\right).$$

If $\sigma_j=+1$, its principal argument is

$$-\tau_j\left(\pi-\frac{\pi}{3}-\frac{\pi}{16}-\frac{\tau_j}{2}\arg(z_j+r\omega_j)\right).$$

In either case,

$$\frac{19\pi}{48}<\left|\arg\left(-\sigma_j\omega_j\sqrt{z_j+r\omega_j}\right)\right|<\frac{23\pi}{48},$$

which proves (8.38). Consequently,

$$\operatorname{Re}\left[-\sigma_j\omega_j\sqrt{z_j+r\omega_j}\right]\geqslant \sin\Bigl(\frac{\pi}{48}\Bigr)\,|\omega_j|\,|z_j+r\omega_j|^{1/2}\geqslant c\,\mathsf{R}(\boldsymbol{z}+r\boldsymbol{\omega})^{1/2}\geqslant c\bigl(\mathsf{R}(\boldsymbol{z})+r\bigr)^{1/2},$$

where we used $|\omega_j|\geqslant 1$ and (8.37). This proves (8.39). □

**Proposition 8.5** (Radial block system for the edge problem)**.** *Fix a sign pattern*

$$(\boldsymbol{\epsilon},\boldsymbol{\eta})\in\{\pm1\}^n\times\{\pm1\}^m.$$

*Let* $\mathsf{C}$ *be a domain of the form* (6.41)*, fix*

$$\boldsymbol{z}=(\boldsymbol{t},\boldsymbol{s})\in\mathsf{C},\qquad R_0\leqslant\mathsf{R}(\boldsymbol{z})\leqslant 2R_0,$$

*and let* $\boldsymbol{\omega}$ *be the fixed vector provided by Proposition 8.4 for*

$$\boldsymbol{\sigma}:=(\epsilon_1,\dots,\epsilon_n,\eta_1,\dots,\eta_m).$$

*Define*

$$\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda):=\boldsymbol{z}+(\lambda-R_0)\boldsymbol{\omega},\qquad \lambda\in[R_0,\infty).$$

*Let* $W$ *be a holomorphic solution of the conjugated edge system* (8.12)–(8.13)*. Let* $\mathcal{J}$ *be the set of square-free index pairs*

$$(I,A),\qquad I\subseteq\{1,\dots,n\},\quad A\subseteq\{1,\dots,m\},$$

*and write*

$$\mathcal{J}^*:=\mathcal{J}\setminus\{(\varnothing,\varnothing)\}.$$

*For $(I,A)\in\mathcal{J}$, define the scaled square-free derivatives along $\gamma_{\boldsymbol{z},\boldsymbol{\omega}}$ by*

$$U_{I,A}(\lambda;\boldsymbol{z},\boldsymbol{\omega}) := \lambda^{|I|+|A|}\partial_{t_I}\partial_{s_A}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big).$$

*Write*

$$U=(U_0,U_*),\qquad U_0 := U_{\varnothing,\varnothing}=W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big),$$

*where $U_*$ collects the components indexed by $\mathcal{J}^*$.*

*Set*

$$\zeta := \frac{2}{3}\lambda^{3/2},\qquad \zeta_0 := \frac{2}{3}R_0^{3/2}.$$

*Then there exist matrix-valued functions*

$$B_{01}(\zeta;\boldsymbol{z},\boldsymbol{\omega}),\qquad B_{10}(\zeta;\boldsymbol{z},\boldsymbol{\omega}),\qquad B_{11}(\zeta;\boldsymbol{z},\boldsymbol{\omega}),$$

*and a matrix*

$$\Gamma(\zeta;\boldsymbol{z},\boldsymbol{\omega})$$

*on the non-empty derivative block. The entries of $B_{01},B_{10},B_{11}$ are uniformly bounded for $\zeta\geqslant\zeta_0$, and the restriction of the square-free derivative system to $\gamma_{\boldsymbol{z},\boldsymbol{\omega}}$ has the form*

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}U_0=\frac{1}{\zeta}B_{01}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_*, \tag{8.41}$$

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}U_*=\Gamma(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_*+\frac{1}{\zeta}B_{10}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_0+\frac{1}{\zeta}B_{11}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_*. \tag{8.42}$$

*Order the components of $U_*$ by increasing total derivative order $|I|+|A|$. Then $\Gamma(\zeta;\boldsymbol{z},\boldsymbol{\omega})$ is block lower triangular. More precisely,*

$$\Gamma_{(I,A),(J,B)}(\zeta;\boldsymbol{z},\boldsymbol{\omega})=0 \tag{8.43}$$

*unless either $(I,A)=(J,B)$, or*

$$J\subseteq I,\qquad B\subseteq A,\qquad 1\leqslant |J|+|B|<|I|+|A|.$$

*For $(I,A)\in\mathcal{J}^*$, the corresponding diagonal entry is*

$$\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega}) := \frac{1}{\lambda^{1/2}}\left(-2\sum_{i\in I}\epsilon_i\omega_i\sqrt{t_i+(\lambda-R_0)\omega_i}-\beta\sum_{a\in A}\eta_a\omega_{n+a}\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right). \tag{8.44}$$

*Moreover, whenever $(I,A)\neq(J,B)$ and the entry on the left-hand side of (8.43) is nonzero,*

$$\left|\Gamma_{(I,A),(J,B)}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\right|\leqslant C\ \mathrm{Re}\left[\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega})-\Gamma_{J,B}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\right]. \tag{8.45}$$

*In particular, there exists $c_\delta>0$ such that*

$$\mathrm{Re}\,\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\geqslant c_\delta(2|I|+\beta|A|)\geqslant c_\delta \tag{8.46}$$

*for every $(I,A)\in\mathcal{J}^*$ and every $\zeta\geqslant\zeta_0$, after decreasing $c_\delta$ if necessary.*

*Proof.* We prove the statement in three steps.

*Step 1: Coefficient estimates along the shifted ray.*

By Proposition 8.4, along the shifted ray $\gamma_{\boldsymbol{z},\boldsymbol{\omega}}$ we have

$$\mathsf{R}\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big)\gtrsim\lambda,\qquad \mathrm{Re}\left[-\sigma_j\omega_j\sqrt{z_j+(\lambda-R_0)\omega_j}\right]\gtrsim\lambda^{1/2},\quad \lambda\geqslant R_0$$

Therefore Lemma 8.3 gives, uniformly along the ray,

$$\Omega_i, \Sigma_a, r_i, v_a = \mathrm{O}(\lambda^{-1}), \qquad \partial\Omega_i, \partial\Sigma_a = \mathrm{O}(\lambda^{-2}),$$

and the first-order coefficients in $\mathsf{R}_i^{(t)}$ and $\mathsf{R}_a^{(s)}$ are $\mathrm{O}(\lambda^{-1})$, whereas the scalar coefficients $V_i^{(t)}$ and $V_a^{(s)}$ are $\mathrm{O}(\lambda^{-2})$. More generally, every additional derivative falling on one of these coefficients improves its decay by one further power of $\lambda^{-1}$.

The conjugated equations give the reduction rules

$$\partial_{t_i}^2 W = -2q_i\partial_{t_i} W - \mathsf{R}_i^{(t)} W, \tag{8.47}$$

$$\partial_{s_a}^2 W = -2p_a\partial_{s_a} W - \mathsf{R}_a^{(s)} W. \tag{8.48}$$

We shall use these identities repeatedly to reduce every repeated derivative back to square-free derivatives.

We also record a comparison that will be used for the principal off-diagonal terms. Let $u$ and $v$ be any two coordinates along the shifted ray. Since $|u|,\ |v|,\ |u-v| \gtrsim \lambda$, one has

$$\frac{\lambda|\sqrt{v}|}{|u-v|} \leqslant C|\sqrt{u}|. \tag{8.49}$$

Indeed, if $|v| \leqslant 2|u|$, then

$$\frac{\lambda|\sqrt{v}|}{|u-v|} \leqslant C|\sqrt{u}|$$

because $|u-v| \gtrsim \lambda$. If $|v| > 2|u|$, then $|u-v| \geqslant |v|/2$, and hence

$$\frac{\lambda|\sqrt{v}|}{|u-v|} \leqslant \frac{2\lambda}{|\sqrt{v}|} \leqslant C|\sqrt{u}|,$$

where we used $|u|,|v| \gtrsim \lambda$. The same estimate applies to every pair of $t$- and $s$-coordinates.

*Step 2: The radial system in the $\lambda$-variable.*

First,

$$U_0(\lambda;\boldsymbol{z},\boldsymbol{\omega}) = W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big).$$

Since, along the shifted ray,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \sum_{i=1}^{n} \omega_i\partial_{t_i} + \sum_{a=1}^{m}\omega_{n+a}\partial_{s_a},$$

we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}U_0 = \frac{1}{\lambda}C_{01}(\lambda;\boldsymbol{z},\boldsymbol{\omega})U_*, \tag{8.50}$$

where $C_{01}$ is uniformly bounded.

We next derive the equations for the non-empty square-free derivatives. Fix

$$(I,A) \in \mathcal{J}^*.$$

By the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}U_{I,A} = \frac{|I|+|A|}{\lambda}U_{I,A} + \lambda^{|I|+|A|}\left(\sum_{i=1}^{n}\omega_i\partial_{t_i}\partial_{t_I}\partial_{s_A}W + \sum_{a=1}^{m}\omega_{n+a}\partial_{s_a}\partial_{t_I}\partial_{s_A}W\right)\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big). \tag{8.51}$$

If $i \notin I$, then the derivative remains square-free and

$$\lambda^{|I|+|A|}\partial_{t_i}\partial_{t_I}\partial_{s_A}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) = \frac{1}{\lambda}U_{I\cup\{i\},A}.$$

Similarly, if $a \notin A$, then

$$\lambda^{|I|+|A|}\partial_{s_a}\partial_{t_I}\partial_{s_A}W\big(\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda)\big) = \frac{1}{\lambda}U_{I,A\cup\{a\}}.$$

Thus all terms in (8.51) with $i \notin I$ or $a \notin A$ belong to the $\lambda^{-1}$-part of the radial system.

Suppose now that $i \in I$, and write

$$I' := I \setminus \{i\}.$$

Using (8.47), and observing that $\partial_{t_{I'}}\partial_{s_A}$ does not act on $q_i$, we obtain

$$\begin{aligned}\lambda^{|I|+|A|}\partial_{t_i}^2\partial_{t_{I'}}\partial_{s_A}W &= \lambda^{|I|+|A|}\partial_{t_{I'}}\partial_{s_A}\left(-2q_i(t_i+(\lambda-R_0)\omega_i)\partial_{t_i} - \mathsf{R}_i^{(t)}\right)W \\ &= -2q_i(t_i+(\lambda-R_0)\omega_i)U_{I,A} - \lambda^{|I|+|A|}\partial_{t_{I'}}\partial_{s_A}\big(\mathsf{R}_i^{(t)}W\big),\end{aligned} \tag{8.52}$$

where all functions are evaluated along the shifted ray. Since

$$q_i(t_i+(\lambda-R_0)\omega_i) = \epsilon_i\sqrt{t_i+(\lambda-R_0)\omega_i} + \mathrm{O}(\lambda^{-1}),$$

the first term on the right-hand side of (8.52), after multiplication by $\omega_i$, contributes

$$-2\epsilon_i\omega_i\sqrt{t_i+(\lambda-R_0)\omega_i}\,U_{I,A} + \frac{1}{\lambda}\,\mathrm{O}(1)U_{I,A}.$$

The first term is the $t_i$-part of the diagonal entry in (8.44).

We now analyze the terms arising from $\mathsf{R}_i^{(t)}$. Consider first a first-order summand involving $\partial_{t_j}W$. After applying the Leibniz rule, a typical term is

$$\lambda^{|I|+|A|}\big(\partial_{t_{I_1}}\partial_{s_{A_1}}(\text{coefficient})\big)\partial_{t_{I_2}}\partial_{s_{A_2}}\partial_{t_j}W, \tag{8.53}$$

where

$$I_1 \sqcup I_2 = I', \qquad A_1 \sqcup A_2 = A.$$

By Step 1, the differentiated coefficient in (8.53) is

$$\mathrm{O}\big(\lambda^{-1-|I_1|-|A_1|}\big).$$

If $j \notin I_2$, then the derivative acting on $W$ is already square-free. Passing to the corresponding scaled component gives

$$\lambda^{|I|+|A|}\lambda^{-1-|I_1|-|A_1|}\lambda^{-|I_2|-|A_2|-1} = \lambda^{-1}.$$

Hence this term belongs to the $\lambda^{-1}$-part of the radial system. The same conclusion holds for a first-order summand involving $\partial_{s_b}W$ when $b \notin A_2$.

Suppose instead that $j \in I_2$. Then

$$\partial_{t_{I_2}}\partial_{s_{A_2}}\partial_{t_j}W = \partial_{t_{I_2\setminus\{j\}}}\partial_{s_{A_2}}\partial_{t_j}^2W.$$

Using (8.47), we may either use the leading square-root term in $q_j$, use the remainder $r_j$, or continue the reduction through $-\mathsf{R}_j^{(t)}W$. If we use the leading square-root term, then the resulting square-free derivative is $\partial_{t_{I_2}}\partial_{s_{A_2}}W$, and its coefficient in the $\lambda$-system is bounded by

$$C\lambda^{|I|+|A|}\lambda^{-1-|I_1|-|A_1|}\left|\sqrt{t_j+(\lambda-R_0)\omega_j}\right|\lambda^{-|I_2|-|A_2|} = C\left|\sqrt{t_j+(\lambda-R_0)\omega_j}\right|.$$

Moreover,

$$|I_2| + |A_2| = |I| + |A| - 1 - |I_1| - |A_1| < |I| + |A|.$$

Thus this is a principal off-diagonal contribution from a strictly lower square-free derivative order. After the change of variable $\zeta = 2\lambda^{3/2}/3$, its coefficient acquires the factor $\lambda^{-1/2}$. Such a coefficient need not be uniformly bounded when some coordinates are much larger than the separation scale; it will instead be controlled by the relative estimate below.

If we use $r_j$ rather than the leading square-root part of $q_j$, then $r_j = \mathrm{O}(\lambda^{-1})$, and the resulting term belongs to the $\lambda^{-1}$-part. If we use $-\mathsf{R}_j^{(t)} W$, the reduction may continue. At this stage the remaining square-free derivatives no longer contain $t_j$. Consequently, the term $\Omega_j \partial_{t_j} W$, as well as the $\partial_{t_j}$-part of a pair-interaction operator, cannot create another repeated derivative and therefore terminates the reduction with a $\lambda^{-1}$-term. A scalar term also terminates the reduction and contributes only to the $\lambda^{-1}$-part. The only terms that can continue a principal reduction are pair-interaction terms whose first-order derivative has an index still present among the remaining square-free derivatives. Such a term creates a new repeated derivative, to which the corresponding reduction rule is applied.

The same discussion applies when the first-order summand involves $\partial_{s_b} W$. In particular, a principal reduction chain may pass between the $t$- and $s$-variables, but at every step the next repeated index is already present among the remaining square-free derivatives. Thus no new index can occur in a principal off-diagonal term. If a principal contribution from $U_{J,B}$ appears in the equation for $U_{I,A}$, then

$$J \subseteq I, \qquad B \subseteq A, \qquad |J| + |B| < |I| + |A|.$$

This proves the strict block-lower-triangular structure of the principal matrix.

We next prove the relative estimate for the principal off-diagonal entries. It is enough to consider a chain beginning with the radial $t_i$-derivative; the argument for a chain beginning with an $s_a$-derivative is identical. Fix $i \in I$, and retain the notation $I' = I \setminus \{i\}$. Consider first the pair-interaction term in (8.53) involving

$$\frac{\partial_{t_j}}{t_i - t_j}, \qquad j \in I'.$$

If no derivative falls on its coefficient, then this term produces, up to a fixed constant,

$$\omega_i \lambda^{|I|+|A|} \frac{1}{t_i - t_j + (\lambda - R_0)(\omega_i - \omega_j)} \partial_{t_{I' \setminus \{j\}}} \partial_{s_A} \partial_{t_j}^2 W.$$

Using the leading square-root part of the reduction of $\partial_{t_j}^2 W$, and then passing to the $\zeta$-variable, gives a contribution to $U_{I',A}$ bounded by

$$C|\omega_i| \frac{\lambda^{1/2} \left|\sqrt{t_j + (\lambda - R_0)\omega_j}\right|}{|t_i - t_j + (\lambda - R_0)(\omega_i - \omega_j)|} = C \frac{|\omega_i|}{\lambda^{1/2}} \frac{\lambda \left|\sqrt{t_j + (\lambda - R_0)\omega_j}\right|}{|t_i - t_j + (\lambda - R_0)(\omega_i - \omega_j)|}.$$

By (8.49), this is bounded by

$$C \frac{|\omega_i| \left|\sqrt{t_i + (\lambda - R_0)\omega_i}\right|}{\lambda^{1/2}}.$$

Thus the square-root factor associated with the terminal repeated index has been transferred to the initial index $i$, which is deleted from the final square-free pair.

The same mechanism applies to an arbitrary principal reduction chain. Every additional pair-interaction step contributes one shifted difference denominator and one compensating factor of $\lambda$ coming from the scaled square-free derivatives. Starting from the square-root factor produced at the last reduction and applying (8.49) successively, in the reverse order of the interaction chain, transfers that square-root factor back across all the interaction denominators. Therefore every principal chain beginning with the radial $t_i$-derivative is bounded, in the $\zeta$-system, by

$$C \frac{|\omega_i| \left|\sqrt{t_i + (\lambda - R_0)\omega_i}\right|}{\lambda^{1/2}}.$$

Likewise, every principal chain beginning with the radial $s_a$-derivative is bounded by

$$C\frac{|\omega_{n+a}|\left|\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right|}{\lambda^{1/2}}.$$

Derivatives falling on a pair-interaction coefficient do not change this conclusion. They produce additional inverse powers of the corresponding shifted difference. After passage to the scaled derivatives, each such additional inverse difference is accompanied by one additional factor of $\lambda$. By Proposition 8.4, every shifted difference is bounded below by a constant multiple of $\lambda$, so these additional factors are uniformly bounded. They only change the constant in the preceding estimates.

Since there are only finitely many reduction chains, we conclude that a principal contribution from $U_{J,B}$ to the equation for $U_{I,A}$ is bounded by

$$C\Bigg(\sum_{i\in I\backslash J}\frac{|\omega_i|\left|\sqrt{t_i+(\lambda-R_0)\omega_i}\right|}{\lambda^{1/2}}+\sum_{a\in A\backslash B}\frac{|\omega_{n+a}|\left|\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right|}{\lambda^{1/2}}\Bigg).$$

By the angular estimate (8.39) in Proposition 8.4,

$$\begin{aligned}|\omega_i|\left|\sqrt{t_i+(\lambda-R_0)\omega_i}\right| &\leqslant C\operatorname{Re}\left[-\epsilon_i\omega_i\sqrt{t_i+(\lambda-R_0)\omega_i}\right],\\ |\omega_{n+a}|\left|\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right| &\leqslant C\operatorname{Re}\left[-\eta_a\omega_{n+a}\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right].\end{aligned}$$

All real parts in these displays are nonnegative. Since

$$\begin{aligned}&\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega})-\Gamma_{J,B}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\\ &\quad=\frac{1}{\lambda^{1/2}}\left(-2\sum_{i\in I\backslash J}\epsilon_i\omega_i\sqrt{t_i+(\lambda-R_0)\omega_i}-\beta\sum_{a\in A\backslash B}\eta_a\omega_{n+a}\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right),\end{aligned}$$

the preceding estimate yields

$$\left|\Gamma_{(I,A),(J,B)}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\right|\leqslant C\operatorname{Re}\left[\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega})-\Gamma_{J,B}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\right].$$

This proves the relative estimate for the terms arising from the $t_i$-equations. The proof for the $s_a$-equations is identical.

For completeness, we collect the remaining terms. If a first-order summand does not create a repeated derivative, then the preceding scaling calculation shows that it contributes $\mathrm{O}(\lambda^{-1})$. If a reduction uses one of the remainders $r_i$ or $v_a$, then it again contributes $\mathrm{O}(\lambda^{-1})$. A scalar coefficient is $\mathrm{O}(\lambda^{-2})$; if $|I_1|+|A_1|$ derivatives fall on it, then its contribution after passage to scaled square-free derivatives is bounded by

$$\lambda^{|I|+|A|}\lambda^{-2-|I_1|-|A_1|}\lambda^{-(|I|+|A|-1-|I_1|-|A_1|)}=\lambda^{-1}.$$

The terms involving $\Omega_i\partial_{t_i}$ or $\Sigma_a\partial_{s_a}$ also terminate the reduction and belong to the $\lambda^{-1}$-part, because the remaining square-free derivatives do not contain the currently repeated index. Thus every term which is not part of a principal pair-interaction chain terminating through a leading square-root term is included in the matrices $C_{10}$ and $C_{11}$.

The case $a\in A$ is now identical. Writing $A':=A\backslash\{a\}$, and using (8.48), the leading diagonal contribution is

$$-\beta\eta_a\omega_{n+a}\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\,U_{I,A},$$

up to a term of the form $\lambda^{-1}\,\mathrm{O}(1)U_{I,A}$. The remaining principal terms satisfy the same strict-inclusion and relative estimates, and all other terms belong to the $\lambda^{-1}$-part.

Collecting all contributions in (8.51), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}U_0 = \frac{1}{\lambda}C_{01}(\lambda;\boldsymbol{z},\boldsymbol{\omega})U_*, \tag{8.54}$$

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}U_* = \lambda^{1/2}\Gamma(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_* + \frac{1}{\lambda}C_{10}(\lambda;\boldsymbol{z},\boldsymbol{\omega})U_0 + \frac{1}{\lambda}C_{11}(\lambda;\boldsymbol{z},\boldsymbol{\omega})U_*, \tag{8.55}$$

where the entries of $C_{01}, C_{10}, C_{11}$ are uniformly bounded. Under the ordering by increasing $|I|+|A|$, the matrix $\Gamma$ is block lower triangular, its diagonal entries are given by (8.44), and its off-diagonal entries satisfy (8.45). Notice also that a principal reduction always leaves at least one derivative. Hence no principal term couples $U_0$ into $U_*$; all couplings from $U_0$ into $U_*$ are contained in $\lambda^{-1}C_{10}U_0$.

*Step 3: Change of variable and the diagonal spectral gap.*

Since

$$\zeta = \frac{2}{3}\lambda^{3/2}, \qquad \frac{\mathrm{d}}{\mathrm{d}\zeta} = \lambda^{-1/2}\frac{\mathrm{d}}{\mathrm{d}\lambda}, \qquad \lambda^{-3/2} = \frac{2}{3\zeta},$$

applying $\lambda^{-1/2}\mathrm{d}/\mathrm{d}\lambda$ to (8.54)–(8.55) gives (8.41)–(8.42). The constant factor $2/3$ is absorbed into $B_{01}, B_{10}, B_{11}$. The triangularity property (8.43) is unchanged by this change of variable.

Finally, Proposition 8.4 gives

$$\operatorname{Re}\left[-\epsilon_i\omega_i\sqrt{t_i+(\lambda-R_0)\omega_i}\right] \geqslant c_\delta\lambda^{1/2}, \quad \operatorname{Re}\left[-\eta_a\omega_{n+a}\sqrt{s_a+(\lambda-R_0)\omega_{n+a}}\right] \geqslant c_\delta\lambda^{1/2}.$$

Substituting these estimates into the diagonal formula (8.44) gives

$$\operatorname{Re}\Gamma_{I,A}(\zeta;\boldsymbol{z},\boldsymbol{\omega}) \geqslant c_\delta(2|I|+\beta|A|).$$

Since $(I,A)\neq(\varnothing,\varnothing)$, the right-hand side is bounded below by a positive constant after decreasing $c_\delta$. This proves (8.46). $\square$

*Proof of the second statement in Theorem 6.5.* Fix a sign pattern $(\boldsymbol{\epsilon},\boldsymbol{\eta})$, and let $\boldsymbol{\omega}$ be the fixed vector from Proposition 8.4, with

$$\boldsymbol{\sigma} = (\epsilon_1,\ldots,\epsilon_n,\eta_1,\ldots,\eta_m).$$

We use this same vector at every point of $\mathsf{C}$. Fix first $\boldsymbol{z}\in\mathsf{C}$, set $R_0 := \mathsf{R}(\boldsymbol{z})$, consider

$$\gamma_{\boldsymbol{z},\boldsymbol{\omega}}(\lambda) := \boldsymbol{z} + (\lambda-R_0)\boldsymbol{\omega}, \qquad \lambda\geqslant R_0.$$

and set

$$\zeta := \frac{2}{3}\lambda^{3/2}, \qquad \zeta_0 := \frac{2}{3}R_0^{3/2}.$$

Along this ray, Proposition 8.5 gives

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}U_0 = \frac{1}{\zeta}B_{01}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_*, \tag{8.56}$$

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}U_* = \Gamma(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_* + \frac{1}{\zeta}B_{10}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_0 + \frac{1}{\zeta}B_{11}(\zeta;\boldsymbol{z},\boldsymbol{\omega})U_*. \tag{8.57}$$

After increasing $M$ if necessary, we may assume that

$$\|B_{01}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\| + \|B_{10}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\| + \|B_{11}(\zeta;\boldsymbol{z},\boldsymbol{\omega})\| \leqslant M, \qquad \zeta\geqslant\zeta_0,$$

and that the constant in (8.45) is at most $M$. Recall that $\Gamma$ is block lower triangular under the ordering by increasing $|I|+|A|$. Its entries need not be uniformly bounded when some coordinates are much larger than the separation scale; the relative estimate (8.45) is used instead.

*Step 1: Exponential decay of the principal backward evolution.*

For $\xi \geqslant \zeta$, let $\mathcal{E}(\zeta, \xi; \boldsymbol{z}, \boldsymbol{\omega})$ be the backward evolution generated by the full triangular matrix $\Gamma$, namely

$$\partial_\zeta \mathcal{E}(\zeta, \xi; \boldsymbol{z}, \boldsymbol{\omega}) = \Gamma(\zeta; \boldsymbol{z}, \boldsymbol{\omega}) \mathcal{E}(\zeta, \xi; \boldsymbol{z}, \boldsymbol{\omega}), \qquad \mathcal{E}(\xi, \xi; \boldsymbol{z}, \boldsymbol{\omega}) = I. \tag{8.58}$$

By (8.46), the backward evolution generated by the diagonal part of $\Gamma$ satisfies

$$\left\| \exp\left( - \int_\zeta^\xi \operatorname{diag} \Gamma(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right) \right\| \leqslant e^{-c_\delta (\xi - \zeta)}. \tag{8.59}$$

By (8.43), every off-diagonal factor strictly increases the total derivative order from its column to its row. Therefore, as in the bulk expansion (7.58), the variation-of-constants expansion terminates after at most $n + m - 1$ off-diagonal factors.

We estimate the terms in this finite expansion. Consider a strictly increasing chain of non-empty square-free pairs, with $1 \leqslant r \leqslant n + m - 1$,

$$(I_0, A_0) \subsetneq (I_1, A_1) \subsetneq \cdots \subsetneq (I_r, A_r),$$

and transition times $\xi \geqslant \tau_1 \geqslant \cdots \geqslant \tau_r \geqslant \zeta$. The diagonal factors appearing in the term associated with this chain are

$$\exp\left( - \int_{\tau_1}^\xi \Gamma_{I_0, A_0}(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right) \prod_{j=1}^{r-1} \exp\left( - \int_{\tau_{j+1}}^{\tau_j} \Gamma_{I_j, A_j}(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right) \exp\left( - \int_\zeta^{\tau_r} \Gamma_{I_r, A_r}(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right).$$

Taking absolute values and using the telescoping identity, we obtain

$$\exp\left( - \int_\zeta^\xi \operatorname{Re} \Gamma_{I_0, A_0}(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right) \prod_{j=1}^{r} \exp\left( - \int_\zeta^{\tau_j} \operatorname{Re} \left[ \Gamma_{I_j, A_j}(u; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(u; \boldsymbol{z}, \boldsymbol{\omega}) \right] \, \mathrm{d}u \right).$$

By (8.45), the off-diagonal factor at $\tau_j$ is bounded by

$$M \operatorname{Re} \left[ \Gamma_{I_j, A_j}(\tau_j; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(\tau_j; \boldsymbol{z}, \boldsymbol{\omega}) \right].$$

The real part in the last display is nonnegative by the angular estimate proved in Proposition 8.4 and the diagonal formula (8.44). Hence, after dropping the ordering restriction on the transition times, the absolute value of the integral associated with the chain is bounded by

$$\begin{aligned} M^r \exp\left( - \int_\zeta^\xi \operatorname{Re} \Gamma_{I_0, A_0}(u; \boldsymbol{z}, \boldsymbol{\omega}) \, \mathrm{d}u \right) \prod_{j=1}^{r} \int_\zeta^\xi \operatorname{Re} \left[ \Gamma_{I_j, A_j}(\tau; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(\tau; \boldsymbol{z}, \boldsymbol{\omega}) \right] \\ \times \exp\left( - \int_\zeta^\tau \operatorname{Re} \left[ \Gamma_{I_j, A_j}(u; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(u; \boldsymbol{z}, \boldsymbol{\omega}) \right] \, \mathrm{d}u \right) \mathrm{d}\tau. \end{aligned}$$

Each integral in the product is at most 1. Indeed, since the real part in question is nonnegative, for every $j = 1, \ldots, r$,

$$\begin{aligned} &\int_\zeta^\xi \operatorname{Re} \left[ \Gamma_{I_j, A_j}(\tau; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(\tau; \boldsymbol{z}, \boldsymbol{\omega}) \right] \\ &\quad \times \exp\left( - \int_\zeta^\tau \operatorname{Re} \left[ \Gamma_{I_j, A_j}(u; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(u; \boldsymbol{z}, \boldsymbol{\omega}) \right] \, \mathrm{d}u \right) \mathrm{d}\tau \\ &= 1 - \exp\left( - \int_\zeta^\xi \operatorname{Re} \left[ \Gamma_{I_j, A_j}(u; \boldsymbol{z}, \boldsymbol{\omega}) - \Gamma_{I_{j-1}, A_{j-1}}(u; \boldsymbol{z}, \boldsymbol{\omega}) \right] \, \mathrm{d}u \right) \leqslant 1. \end{aligned}$$

Moreover, (8.46) gives

$$\exp\left(-\int_\zeta^\xi \operatorname{Re}\Gamma_{I_0,A_0}(u;\boldsymbol{z},\boldsymbol{\omega})\,\mathrm{d}u\right) \leqslant e^{-c_\delta(\xi-\zeta)}.$$

There are only finitely many strictly increasing chains of square-free pairs. Summing the preceding bounds and enlarging $C$, we obtain

$$\|\mathcal{E}(\zeta,\xi;\boldsymbol{z},\boldsymbol{\omega})\| \leqslant Ce^{-c_\delta(\xi-\zeta)}, \qquad \xi \geqslant \zeta \geqslant \zeta_0. \tag{8.60}$$

Thus the full principal backward evolution has uniform exponential decay.

*Step 2: Reduction to a Volterra equation.*

We impose the normalization

$$U_0(\zeta) \to 1, \qquad U_*(\zeta) \to 0, \qquad \zeta \to \infty.$$

Integrating (8.56) from $\zeta$ to infinity gives

$$U_0(\zeta) = 1 + (\mathcal{K}U_*)(\zeta), \qquad (\mathcal{K}U_*)(\zeta) := -\int_\zeta^\infty \frac{1}{\xi} B_{01}(\xi;\boldsymbol{z},\boldsymbol{\omega})U_*(\xi)\,\mathrm{d}\xi. \tag{8.61}$$

Substituting this into (8.57) and using variation of constants with the full triangular propagator $\mathcal{E}$ gives

$$U_* = \mathcal{T}(U_*), \tag{8.62}$$

where

$$(\mathcal{T}U_*)(\zeta) := -\int_\zeta^\infty \mathcal{E}(\zeta,\xi;\boldsymbol{z},\boldsymbol{\omega})\frac{1}{\xi}\Big[B_{10}(\xi;\boldsymbol{z},\boldsymbol{\omega})\big(1+\mathcal{K}U_*(\xi)\big) + B_{11}(\xi;\boldsymbol{z},\boldsymbol{\omega})U_*(\xi)\Big]\,\mathrm{d}\xi. \tag{8.63}$$

*Step 3: The fixed-point space and the basic estimates.*

Let $\mathcal{B}_{\boldsymbol{z}}$ be the Banach space of continuous functions

$$U_* : [\zeta_0,\infty) \longrightarrow \mathbb{C}^{|\mathcal{J}^*|}$$

with norm

$$\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}} := \sup_{\zeta\geqslant\zeta_0} \zeta\|U_*(\zeta)\|_2.$$

If $U_* \in \mathcal{B}_{\boldsymbol{z}}$, then by (8.61),

$$|(\mathcal{K}U_*)(\zeta)| \leqslant M\int_\zeta^\infty \frac{\|U_*(\xi)\|_2}{\xi}\,\mathrm{d}\xi \leqslant M\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}\int_\zeta^\infty \xi^{-2}\,\mathrm{d}\xi = \frac{M}{\zeta}\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}.$$

Hence

$$|(\mathcal{K}U_*)(\zeta)| \leqslant \frac{C}{\zeta}\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{8.64}$$

Using (8.60), (8.63), and (8.64), we have

$$\begin{aligned}
\zeta\|(\mathcal{T}U_*)(\zeta)\|_2 &\leqslant M\zeta\int_\zeta^\infty \|\mathcal{E}(\zeta,\xi;\boldsymbol{z},\boldsymbol{\omega})\|\frac{1}{\xi}\Big(1+|(\mathcal{K}U_*)(\xi)| + \|U_*(\xi)\|_2\Big)\,\mathrm{d}\xi \\
&\leqslant CM\zeta\int_\zeta^\infty e^{-c_\delta(\xi-\zeta)}\Big(\xi^{-1} + \|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}\xi^{-2}\Big)\,\mathrm{d}\xi \\
&\leqslant C_0 + C_1\zeta_0^{-1}\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}.
\end{aligned}$$

Therefore

$$\|\mathcal{T}U_*\|_{\mathcal{B}_{\boldsymbol{z}}} \leqslant C_0 + C_1\zeta_0^{-1}\|U_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{8.65}$$

Similarly, for $U_*, V_* \in \mathcal{B}_{\boldsymbol{z}}$,

$$|(\mathcal{K}U_*)(\zeta) - (\mathcal{K}V_*)(\zeta)| \leqslant \frac{C}{\zeta}\|U_* - V_*\|_{\mathcal{B}_{\boldsymbol{z}}}.$$

Hence

$$\begin{aligned}
\zeta\|(\mathcal{T}U_* - \mathcal{T}V_*)(\zeta)\|_2 &\leqslant M\zeta\int_\zeta^\infty \|\mathcal{E}(\zeta,\xi;\boldsymbol{z},\boldsymbol{\omega})\|\frac{1}{\xi}\Big(|(\mathcal{K}U_*)(\xi) - (\mathcal{K}V_*)(\xi)| + \|U_*(\xi) - V_*(\xi)\|_2\Big)\,\mathrm{d}\xi \\
&\leqslant CM\zeta\int_\zeta^\infty e^{-c_\delta(\xi-\zeta)}\xi^{-2}\,\mathrm{d}\xi\,\|U_* - V_*\|_{\mathcal{B}_{\boldsymbol{z}}} \\
&\leqslant C_2\zeta_0^{-1}\|U_* - V_*\|_{\mathcal{B}_{\boldsymbol{z}}}.
\end{aligned}$$

Thus

$$\|\mathcal{T}U_* - \mathcal{T}V_*\|_{\mathcal{B}_{\boldsymbol{z}}} \leqslant C_2\zeta_0^{-1}\|U_* - V_*\|_{\mathcal{B}_{\boldsymbol{z}}}. \tag{8.66}$$

*Step 4: Existence, asymptotics, and raywise uniqueness.*

If $R_0$, equivalently $\zeta_0$, is sufficiently large, then $C_2\zeta_0^{-1} < 1$, so by (8.66) $\mathcal{T}$ is a contraction on $\mathcal{B}_{\boldsymbol{z}}$. Hence (8.62) admits a unique fixed point $U_* \in \mathcal{B}_{\boldsymbol{z}}$. Defining $U_0$ by (8.61), we obtain a unique solution $(U_0, U_*)$ of (8.56)–(8.57) satisfying

$$U_0(\zeta) \to 1, \qquad U_*(\zeta) \to 0, \qquad \zeta \to \infty.$$

Moreover, (8.64) and (8.65) imply

$$U_0(\zeta) = 1 + \mathrm{O}(\zeta^{-1}), \qquad U_*(\zeta) = \mathrm{O}(\zeta^{-1}), \qquad \zeta \geqslant \zeta_0. \tag{8.67}$$

Now define

$$W(\boldsymbol{z}) := U_0(\zeta_0).$$

Since $\zeta_0 = 2R_0^{3/2}/3$, we have

$$W(\boldsymbol{z}) = 1 + \mathrm{O}(\zeta_0^{-1}) = 1 + \mathrm{O}\big(\mathsf{R}(\boldsymbol{z})^{-3/2}\big).$$

For every $(I, A) \in \mathcal{J}^*$, (8.67) also gives

$$U_{I,A}(\zeta_0) = \mathrm{O}\big(\mathsf{R}(\boldsymbol{z})^{-3/2}\big).$$

At this stage these are estimates for the components of the constructed square-free jet. In Step 5 we verify that this jet is horizontal for the full flat system; its components are then identified with the square-free derivatives of $W$, yielding the desired derivative estimates.

The same argument applied to the difference of two normalized solutions has no constant term in (8.61). Therefore the corresponding Volterra map is a strict contraction with zero forcing, and the difference vanishes. This proves uniqueness along every admissible shifted ray.

*Step 5: Holomorphic dependence and gluing on* $\mathsf{C}$.

Fix $\boldsymbol{z}_0 \in \mathsf{C}$. Choose a fixed number $R_0 > 0$, and shrink to a sufficiently small neighborhood $\mathfrak{U} \subset \mathsf{C}$ of $\boldsymbol{z}_0$, so that

$$R_0 \leqslant \mathsf{R}(\boldsymbol{z}) \leqslant 2R_0, \qquad \boldsymbol{z} \in \mathfrak{U}.$$

For every $\boldsymbol{z} \in \mathfrak{U}$, use the rays

$$\boldsymbol{z} + (\lambda - R_0)\boldsymbol{\omega}, \qquad \lambda \geqslant R_0.$$

The separation and phase estimates of Proposition 8.4 hold uniformly on this family of rays. Since both $R_0$ and $\boldsymbol{\omega}$ are fixed on $\mathfrak{U}$, all coefficients of the Volterra equation depend holomorphically on $\boldsymbol{z}$, and its contraction estimate is uniform. Hence the normalized fixed point depends holomorphically on $\boldsymbol{z}$.

We have so far solved the first-order system only in the radial direction. To check the remaining coordinate directions, fix one coordinate and compare two ways of differentiating the constructed square-free jet: differentiation with respect to the base point, and the corresponding coordinate equation of the full square-free first-order system. Their difference satisfies the homogeneous radial system by flatness. The estimates above hold uniformly for nearby base points, so this difference has zero normalization at infinity. Raywise uniqueness forces it to vanish. Thus the constructed jet is a holomorphic horizontal section of the full flat system. Its empty component is a holomorphic solution $W^{(\mathfrak{U})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ of the conjugated edge system, and

$$W^{(\mathfrak{U})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = 1 + \mathrm{O}\big(\mathsf{R}^{-3/2}\big), \qquad \partial_{\boldsymbol{t}_I}\partial_{\boldsymbol{s}_A} W^{(\mathfrak{U})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \mathrm{O}\big(\mathsf{R}^{-|I|-|A|-3/2}\big).$$

The local solutions agree on overlaps. Indeed, all neighborhoods use the same vector $\boldsymbol{\omega}$. At a point $\boldsymbol{z} \in \mathfrak{U} \cap \mathfrak{V}$, undo the radial scaling and write both constructions on the same geometric ray $\boldsymbol{z} + r\boldsymbol{\omega}$, $r \geqslant 0$. They solve the same radial system and have the same normalization at infinity. Raywise uniqueness therefore gives

$$W^{(\mathfrak{U})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = W^{(\mathfrak{V})}_{\boldsymbol{\epsilon},\boldsymbol{\eta}} \qquad \text{on } \mathfrak{U} \cap \mathfrak{V}.$$

The local solutions glue to a holomorphic function $W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$ on all of $\mathsf{C}$. Setting

$$F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} := \Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}} W_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$$

gives a holomorphic solution of the original edge system satisfying (6.43)–(6.44).

*Step 6: Linear independence.*

For a square-free pair $(I, A)$, the Airy asymptotics, the explicit logarithmic derivatives of $\Phi_{\boldsymbol{\epsilon},\boldsymbol{\eta}}$, and (6.43)–(6.44) give, uniformly on $\mathsf{C}$,

$$\left(\prod_{i\in I}\sqrt{t_i}\prod_{a\in A}\frac{\beta}{2}\sqrt{s_a}\right)^{-1}\Phi^{-1}_{\boldsymbol{\epsilon},\boldsymbol{\eta}}\partial_{\boldsymbol{t}_I}\partial_{\boldsymbol{s}_A}F_{\boldsymbol{\epsilon},\boldsymbol{\eta}} = \prod_{i\in I}\epsilon_i\prod_{a\in A}\eta_a + \mathrm{O}\big(D^{-3/2}\big). \tag{8.68}$$

The leading matrix in (8.68), with rows indexed by $(I, A)$ and columns indexed by $(\boldsymbol{\epsilon}, \boldsymbol{\eta})$, is the character table

$$\left(\prod_{i\in I}\epsilon_i\prod_{a\in A}\eta_a\right)_{(I,A),(\boldsymbol{\epsilon},\boldsymbol{\eta})}$$

of $\{\pm 1\}^{n+m}$. It is invertible. Increasing $D$, if necessary, the exact normalized jet matrix is also invertible. Therefore the $2^{n+m}$ solutions are linearly independent. □

# A Airy functions

We recall some standard facts about the Airy function Ai. For real $x$, it is given by the conditionally convergent integral

$$\mathrm{Ai}(x) = \frac{1}{\pi}\lim_{R\to\infty}\int_0^R \cos\left(\frac{t^3}{3} + xt\right)\,\mathrm{d}t.$$

The function Ai extends to an entire function on $\mathbb{C}$, and it solves the Airy equation

$$\mathrm{Ai}''(w) - w\,\mathrm{Ai}(w) = 0. \tag{A.1}$$

It is characterized by the decay condition

$$\mathrm{Ai}(w) \to 0 \qquad \text{as } w \to +\infty$$

along the positive real axis. All zeros of Ai are real and negative; we write them as

$$0 > \mathfrak{a}_1 > \mathfrak{a}_2 > \mathfrak{a}_3 > \cdots .$$

It is known that the Airy zeros satisfy $\mathfrak{a}_i \sim -(3\pi i/2)^{2/3}$. More precisely, for every $i \in \mathbb{N}$, see for example [32, eqs. (9.9.6) and (9.9.18)],

$$\left|\mathfrak{a}_i + \left(\frac{3\pi i}{2}\right)^{2/3}\right| \lesssim i^{-1/3}. \tag{A.2}$$

We refer to [32, Chapter 9] for further background on Airy functions.

## A.1 Nevanlinna representation

The logarithmic derivative

$$-\frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)}$$

is a particle-generated Nevanlinna function. Its associated configuration is the Airy-zero configuration

$$\Gamma_{\mathrm{Ai}} := \sum_{i=1}^{\infty} \delta_{\mathfrak{a}_i}.$$

More precisely, one has the partial fraction expansion

$$-\frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)} = \sum_{i=1}^{\infty}\left(\frac{1}{\mathfrak{a}_i - w} - \frac{1}{\mathfrak{a}_i}\right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}. \tag{A.3}$$

We shall also use the square-root asymptotic

$$\left|\frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)} + \sqrt{w}\right| \lesssim |w|^{-1}, \tag{A.4}$$

valid for $|w|$ sufficiently large and

$$|\arg w| < \pi - |w|^{-9/7}.$$

The estimates (A.3) and (A.4) are proved in [55, Appendix A.1].

## A.2 Asymptotics for Ai

We recall the standard Airy asymptotic

$$\mathrm{Ai}(w) = \frac{1}{2\sqrt{\pi}} w^{-1/4} \exp\left(-\frac{2}{3}w^{3/2}\right)\left(1 + \mathrm{O}_\delta(|w|^{-3/2})\right), \qquad |w| \to \infty, \tag{A.5}$$

uniformly for $|\arg w| \leqslant \pi - \delta$, see [32, Section 9.7(iv)] and [72, Appendix B]. Here and below, fractional powers such as $\sqrt{w}$, $w^{3/2}$, and $w^{-1/4}$ are taken with the principal branch on $\mathbb{C} \setminus \mathbb{R}_{<0}$.

Let

$$\omega := e^{-2\pi \mathrm{i}/3}.$$

If $\alpha^3 = 1$, then $\mathrm{Ai}(\alpha z)$ also solves the Airy equation (A.1). In particular,

$$\mathrm{i}\omega \, \mathrm{Ai}(\omega z), \qquad -\mathrm{i}\omega^2 \, \mathrm{Ai}(\omega^2 z)$$

are also solutions of (A.1).

Moreover, in the upper half-plane $0 < \arg z < \pi$, the solution $\mathrm{i}\omega \, \mathrm{Ai}(\omega z)$ has the pure $e^{2z^{3/2}/3}$ asymptotic

$$\mathrm{i}\omega \, \mathrm{Ai}(\omega z) = \frac{1}{2\sqrt{\pi}} z^{-1/4} \exp\left(\frac{2}{3}z^{3/2}\right)\left(1 + \mathrm{O}(|z|^{-3/2})\right), \qquad |z| \to \infty. \tag{A.6}$$

Similarly, in the lower half-plane $-\pi < \arg z < 0$, $-\mathrm{i}\omega^2 \, \mathrm{Ai}(\omega^2 z)$ has the pure $e^{2z^{3/2}/3}$ asymptotic

$$-\mathrm{i}\omega^2 \, \mathrm{Ai}(\omega^2 z) = \frac{1}{2\sqrt{\pi}} z^{-1/4} \exp\left(\frac{2}{3}z^{3/2}\right)\left(1 + \mathrm{O}(|z|^{-3/2})\right), \qquad |z| \to \infty. \tag{A.7}$$

# B Loop equations of $\beta$-ensembles

Here we recall the notation, estimates, and loop equations for the $\beta$-ensemble from [19]. We use the notation introduced in Section 3.3. As in [19, Section 2.1], $m_V$ satisfies the fixed-point equation

$$m_V(z)^2 + V'(z)m_V(z) + h(z) = 0, \tag{B.1}$$

where

$$h(z) := \int_A^B \frac{V'(\lambda) - V'(z)}{\lambda - z}\, \varrho_V(\lambda)\, \mathrm{d}\lambda.$$

Moreover,

$$2m_V(z) + V'(z) = 2r(z)\sqrt{z-A}\sqrt{z-B}, \tag{B.2}$$

with the usual branch choice. In particular, at the right edge,

$$2m_B + V'(B) = 0. \tag{B.3}$$

We denote the empirical Stieltjes transform as

$$\mathfrak{s}_N(z) := \frac{1}{N}\sum_{j=1}^N \frac{1}{\lambda_j - z}.$$

**Proposition B.1** (Loop equation [19, Proposition A.3])**.** *For any $p \geqslant 1$ and $z, z_2, z_3, \dots, z_p \in \mathbb{C}\setminus\mathbb{R}$,*

$$\begin{aligned} &\mathbb{E}\left[\left(\mathfrak{s}_N(z)^2 + \frac{1}{N}\left(\frac{2}{\beta}-1\right)\mathfrak{s}_N'(z) + \frac{1}{N}\sum_{k=1}^N \frac{V'(\lambda_k)}{\lambda_k - z}\right)\prod_{i=2}^p \mathfrak{s}_N(z_i)\right] \\ &\qquad + \frac{2}{\beta N^2}\mathbb{E}\left[\sum_{j=2}^p \partial_{z_j}\frac{\mathfrak{s}_N(z)-\mathfrak{s}_N(z_j)}{z-z_j}\prod_{\substack{i=2\\ i\neq j}}^p \mathfrak{s}_N(z_i)\right] = 0 \end{aligned} \tag{B.4}$$

Introduce

$$P_N(z) := \mathfrak{s}_N(z)^2 + V'(z)\mathfrak{s}_N(z) + h(z), \quad \Delta_N(z) := \frac{1}{N}\sum_{k=1}^N \frac{V'(\lambda_k)-V'(z)}{\lambda_k - z} - h(z). \tag{B.5}$$

Then (B.4) becomes

$$\mathbb{E}\left[\left(P_N(z) + \frac{1}{N}\left(\frac{2}{\beta}-1\right)\mathfrak{s}_N'(z) + \Delta_N(z)\right)\prod_{i=2}^p \mathfrak{s}_N(z_i)\right] + \frac{2}{\beta N^2}\mathbb{E}\left[\sum_{j=2}^p \partial_{z_j}\frac{\mathfrak{s}_N(z)-\mathfrak{s}_N(z_j)}{z-z_j}\prod_{\substack{i=2\\ i\neq j}}^p \mathfrak{s}_N(z_i)\right] = 0. \tag{B.6}$$

**Proposition B.2** ( [19, Lemma 2.1 and Proposition 3.5])**.** *For any compact set $K \subset \mathbb{C}$, there exists $C > 0$ such that for any $q \geqslant 1$,*

$$\mathbb{E}\left[|\Delta_N(z)|^{2q}\right] \leqslant \frac{(Cq)^{2q}}{N^{2q}}. \tag{B.7}$$

*Moreover, there exist $\eta_0 > 0$ and $C > 0$ such that for any $q \geqslant 1$, any $z = E + \mathrm{i}\eta$ with $0 < \eta \leqslant \eta_0$ and $A - \eta_0 \leqslant E \leqslant B + \eta_0$,*

$$\mathbb{E}\left[|\mathfrak{s}_N(z) - m_V(z)|^q\right] \leqslant \frac{(Cq)^{q/2}}{(N\eta)^q} + \frac{(Cq)^{2q}}{(N\eta)^q}. \tag{B.8}$$

## B.1 Bulk loop equations for $\beta$-ensembles

*Proof of Proposition 3.6.* Set

$$z = E + \frac{w}{N\pi\varrho_E}, \qquad z_j = E + \frac{w_j}{N\pi\varrho_E}. \tag{B.9}$$

By linearity of the loop equation, we may replace $\prod_{j=2}^p \mathfrak{s}_N(z_j)$ by $\prod_{j=2}^p(\mathfrak{s}_N(z_j) + V'(E)/2)$. Starting from the loop equation in the form (B.6), we therefore get

$$\begin{aligned} &\mathbb{E}\left[\left(P_N(z) + \frac{1}{N}\left(\frac{2}{\beta}-1\right)\partial_z \mathfrak{s}_N(z) + \Delta_N(z)\right)\prod_{j=2}^p\left(\mathfrak{s}_N(z_j) + \frac{V'(E)}{2}\right)\right] \\ &\quad + \frac{2}{\beta N^2}\mathbb{E}\left[\sum_{j=2}^p \partial_{z_j}\frac{\mathfrak{s}_N(z)-\mathfrak{s}_N(z_j)}{z-z_j}\prod_{\substack{i=2\\ i\neq j}}^p\left(\mathfrak{s}_N(z_i)+\frac{V'(E)}{2}\right)\right] = 0. \end{aligned} \tag{B.10}$$

We recall from (3.15)

$$\mathfrak{s}_N(z) + \frac{V'(E)}{2} = \pi\varrho_E s_N(w). \tag{B.11}$$

We divide (B.10) by $(\pi\varrho_E)^{p+2}$. First consider the derivative term. Using the relation (B.9) and (B.11) we have

$$\partial_w s_N(w) = \frac{1}{\pi\varrho_E}\frac{1}{N\pi\varrho_E}\partial_z\mathfrak{s}_N(z) = \frac{1}{N(\pi\varrho_E)^2}\partial_z\mathfrak{s}_N(z).$$

Therefore

$$\frac{1}{(\pi\varrho_E)^2}\frac{1}{N}\left(\frac{2}{\beta}-1\right)\partial_z\mathfrak{s}_N(z) = \frac{2-\beta}{\beta}\partial_w s_N(w).$$

Next consider the second loop term. Using the relation (B.9) and (B.11), we get

$$\frac{\mathfrak{s}_N(z)-\mathfrak{s}_N(z_j)}{z-z_j} = N(\pi\varrho_E)^2\frac{s_N(w)-s_N(w_j)}{w-w_j}.$$

We further notice that $\partial_{z_j} = N\pi\varrho_E\,\partial_{w_j}$, and it follows

$$\frac{1}{N^2}\partial_{z_j}\frac{\mathfrak{s}_N(z)-\mathfrak{s}_N(z_j)}{z-z_j} = (\pi\varrho_E)^3\partial_{w_j}\frac{s_N(w)-s_N(w_j)}{w-w_j}.$$

After division by $(\pi\varrho_E)^{p+2}$, and using (B.11)

$$\prod_{\substack{i=2\\ i\neq j}}^p\left(\mathfrak{s}_N(z_i)+\frac{V'(E)}{2}\right) = (\pi\varrho_E)^{p-1}\prod_{\substack{i=2\\ i\neq j}}^p s_N(w_i),$$

the second loop term becomes exactly

$$\partial_{w_j}\frac{s_N(w)-s_N(w_j)}{w-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^p s_N(w_\ell).$$

It remains to identify the leading part of $P_N(z)$. Recall from (B.5)

$$P_N(z) = \mathfrak{s}_N(z)^2 + V'(z)\mathfrak{s}_N(z) + h(z).$$

Using (B.11), we have

$$P_N(z) = (\pi\varrho_E)^2 s_N(w)^2 + \big(V'(z) - V'(E)\big)\pi\varrho_E s_N(w) + h(z) + \frac{V'(E)^2}{4} - \frac{1}{2}V'(z)V'(E). \tag{B.12}$$

At the bulk point $E$, the boundary value of $m_V$ satisfies

$$m_V(E + \mathrm{i}0) = -\frac{1}{2}V'(E) + \mathrm{i}\pi\varrho_E. \tag{B.13}$$

Using the fixed point equation (B.1)

$$m_V(z)^2 + V'(z)m_V(z) + h(z) = 0$$

at $E + \mathrm{i}0$, together with (B.13), we obtain

$$h(E) = \frac{V'(E)^2}{4} + \pi^2\varrho_E^2.$$

Since $V$ and $h$ are analytic near $E$, and $|z - E| = \mathrm{O}_K(1/N)$, we have uniformly for $w \in K$,

$$V'(z) - V'(E) = \mathrm{O}_K(N^{-1}), \qquad h(z) - h(E) = \mathrm{O}_K(N^{-1}).$$

Thus (B.12) gives

$$P_N(z) = (\pi\varrho_E)^2\big(s_N(w)^2 + 1\big) + \mathrm{O}_K(N^{-1})\big(1 + |s_N(w)|\big).$$

Equivalently,

$$\frac{P_N(z)}{(\pi\varrho_E)^2} = s_N(w)^2 + 1 + \mathrm{O}_K(N^{-1})\big(1 + |s_N(w)|\big). \tag{B.14}$$

We now record the moment bounds needed to absorb the error terms. The local law (B.8) gives, for every fixed $q \geqslant 1$,

$$\sup_{w\in K} \mathbb{E}\left[|s_N(w)|^q\right] \leqslant C_{q,K}. \tag{B.15}$$

For points in the lower half-plane, this follows by conjugation. Using (2.2) and (2.4), the derivative terms have uniformly bounded moments as well. For $K \subset \mathbb{C} \setminus \mathbb{R}$,

$$|\partial_w s_N(w)| \leqslant C_K|s_N(w)|, \qquad \left|\partial_{w_j}\frac{s_N(w) - s_N(w_j)}{w - w_j}\right| \leqslant C_K\big(|s_N(w)| + |s_N(w_j)|\big).$$

Finally, the $\Delta_N$-term is negligible. In fact (B.7), (B.15) and Hölder's inequality gives,

$$\sup_{w,w_2,\ldots,w_p\in K}\left|\mathbb{E}\left[\frac{\Delta_N(z)}{(\pi\varrho_E)^2}\prod_{j=2}^{p} s_N(w_j)\right]\right| = \mathrm{O}(N^{-1}) = \mathrm{o}_N(1).$$

Substituting (B.14) into the scaled form of (B.10), and using the preceding moment bounds, we obtain

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\partial_w s_N(w) + s_N(w)^2 + 1\right)\prod_{j=2}^{p} s_N(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s_N(w) - s_N(w_j)}{w - w_j}\prod_{\substack{i=2\\ i\neq j}}^{p} s_N(w_i)\right] = \mathrm{o}_N(1),$$

uniformly for $w, w_2, w_3, \ldots, w_p \in K$. This is (1.4). $\square$

## B.2 Edge loop equations for $\beta$-ensembles

*Proof of Proposition 3.7.* We write

$$z = B + \gamma N^{-2/3} w, \qquad z_j = B + \gamma N^{-2/3} w_j.$$

First we record the deterministic edge expansion of $m_V$. We recall from (B.2) and (B.3)

$$2m_V(z) + V'(z) = 2r(z)\sqrt{z-A}\sqrt{z-B}, \quad 2m_B + V'(B) = 0.$$

We have, uniformly for $w \in K$,

$$m_V(B + \gamma N^{-2/3} w) - m_B = r(B)\sqrt{B-A}\,\gamma^{1/2} N^{-1/3}\sqrt{w} + \mathrm{O}_K(N^{-2/3}).$$

The branch of $\sqrt{w}$ is the principal branch. By the definition of $\gamma$ in (3.17),

$$\gamma^{3/2} r(B)\sqrt{B-A} = 1,$$

and therefore

$$\gamma N^{1/3}\Big(m_V(B+\gamma N^{-2/3}w) - m_B\Big) = \sqrt{w} + \mathrm{O}_K(N^{-1/3}). \tag{B.16}$$

Consequently, we can rewrite (3.18) as

$$\begin{aligned} s_N(w) &= \gamma N^{1/3}\Big(\mathfrak{s}_N(B+\gamma N^{-2/3}w) - m_B\Big), \\ s_N(w) - \sqrt{w} &= \gamma N^{1/3}\Big(\mathfrak{s}_N(B+\gamma N^{-2/3}w) - m_V(B+\gamma N^{-2/3}w)\Big) + \mathrm{O}_K(N^{-1/3}). \end{aligned} \tag{B.17}$$

By the local law (B.8), for every fixed $q \geqslant 1$,

$$\sup_{w\in K} \mathbb{E}\left[|s_N(w)|^q\right] \leqslant C_{q,K}. \tag{B.18}$$

We now apply the loop equation. By linearity, (B.6) remains valid when $\prod_{j=z}^{p} \mathfrak{s}_N(z_j)$ is replaced by the centered product $\prod_{j=2}^{p}\big(\mathfrak{s}_N(z_j) - m_B\big)$. Thus,

$$\begin{aligned} &\mathbb{E}\left[\left(P_N(z) + \frac{1}{N}\left(\frac{2}{\beta}-1\right)\partial_z \mathfrak{s}_N(z) + \Delta_N(z)\right)\prod_{j=2}^{p}\big(\mathfrak{s}_N(z_j) - m_B\big)\right] \\ &\qquad + \frac{2}{\beta N^2}\mathbb{E}\left[\sum_{j=2}^{p} \partial_{z_j}\frac{\mathfrak{s}_N(z) - \mathfrak{s}_N(z_j)}{z - z_j}\prod_{\substack{i=2\\ i\neq j}}^{p}\big(\mathfrak{s}_N(z_i) - m_B\big)\right] = 0. \end{aligned} \tag{B.19}$$

Substituting $z = B + \gamma N^{-2/3} w$ and $z_j = B + \gamma N^{-2/3} w_j$, and multiplying (B.19) by $\big(\gamma N^{1/3}\big)^{p+2}$, we obtain

$$\begin{aligned} &\mathbb{E}\left[\left(\gamma^2 N^{2/3} P_N(z) + \left(\frac{2}{\beta}-1\right)\partial_w s_N(w) + \gamma^2 N^{2/3}\Delta_N(z)\right)\prod_{j=2}^{p} s_N(w_j)\right] \\ &\qquad + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p} \partial_{w_j}\frac{s_N(w) - s_N(w_j)}{w - w_j}\prod_{\substack{i=2\\ i\neq j}}^{p} s_N(w_i)\right] = 0. \end{aligned} \tag{B.20}$$

Indeed, by (B.17)

$$\partial_w s_N(w) = \gamma^2 N^{-1/3}\partial_z \mathfrak{s}_N(z) = \frac{\gamma^2 N^{2/3}}{N}\partial_z \mathfrak{s}_N(z),$$

and the same change of variables gives the second derivative term in (B.20).

It remains to identify the leading part of $\gamma^2 N^{2/3} P_N(z)$. We recall from (B.5) and (B.17)

$$P_N(z) = \mathfrak{s}_N(z)^2 + V'(z)\mathfrak{s}_N(z) + h(z), \quad \mathfrak{s}_N(z) = m_B + \frac{s_N(w)}{\gamma N^{1/3}},$$

then

$$\gamma^2 N^{2/3} P_N(z) = s_N(w)^2 + \gamma N^{1/3}\big(2m_B + V'(z)\big)s_N(w) + \gamma^2 N^{2/3}\big(m_B^2 + V'(z)m_B + h(z)\big). \tag{B.21}$$

By (B.3),

$$2m_B + V'(z) = V'(B + \gamma N^{-2/3} w) - V'(B) = \mathrm{O}_K(N^{-2/3}),$$

and hence the middle term in (B.21) is negligible,

$$\gamma N^{1/3}\big(2m_B + V'(z)\big) = \mathrm{O}_K(N^{-1/3}). \tag{B.22}$$

For the final term in (B.21), use the fixed point equation for $m_V$ in (B.1):

$$m_V(z)^2 + V'(z)m_V(z) + h(z) = 0.$$

Subtracting the expression at $m_B$ gives

$$m_B^2 + V'(z)m_B + h(z) = -\big(2m_B + V'(z)\big)\big(m_V(z) - m_B\big) - \big(m_V(z) - m_B\big)^2.$$

Therefore, using (B.16) and (B.22),

$$\gamma^2 N^{2/3}\big(m_B^2 + V'(z)m_B + h(z)\big) = -w + \mathrm{O}_K(N^{-1/3}). \tag{B.23}$$

Combining (B.21), (B.22), and (B.23), we have

$$\gamma^2 N^{2/3} P_N(z) = s_N(w)^2 - w + \mathrm{O}_K(N^{-1/3})\big(1 + |s_N(w)|\big). \tag{B.24}$$

Using (2.2) and (2.4), the derivative terms have uniformly bounded moments as well. For any compact set $K \subset \mathbb{C} \setminus \mathbb{R}$,

$$|\partial_w s_N(w)| \leqslant C_K |s_N(w)|, \quad \left|\partial_{w_j} \frac{s_N(w) - s_N(w_j)}{w - w_j}\right| \leqslant C_K\big(|s_N(w)| + |s_N(w_j)|\big).$$

Together with (B.18) and Hölder's inequality, this gives uniform moment bounds for all random variables appearing in (B.20).

Finally, the $\Delta_N$-term is negligible. From (B.7),

$$\mathbb{E}\left[|\Delta_N(z)|^{2q}\right] \leqslant \frac{(Cq)^{2q}}{N^{2q}},$$

uniformly for $z = B + \gamma N^{-2/3} w$, $w \in K$. Hence, by Hölder's inequality and (B.18),

$$\sup_{w, w_2, \ldots, w_p \in K} \left| \mathbb{E}\left[\gamma^2 N^{2/3} \Delta_N(z) \prod_{j=2}^{p} s_N(w_j)\right]\right| = \mathrm{O}(N^{-1/3}) = \mathrm{o}_N(1).$$

Substituting (B.24) into (B.20), and using the preceding moment bounds, we obtain, uniformly for $w, w_2, w_3 \ldots, w_p \in K$,

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\partial_w s_N(w) + s_N(w)^2 - w\right)\prod_{j=2}^{p} s_N(w_j)\right] + \frac{2}{\beta}\mathbb{E}\left[\sum_{j=2}^{p} \partial_{w_j} \frac{s_N(w) - s_N(w_j)}{w - w_j} \prod_{\substack{i=2 \\ i \neq j}}^{p} s_N(w_i)\right] = \mathrm{o}_N(1).$$

This is (3.19). $\square$

# C Proof of the edge local law

In this section, we prove Proposition 3.4. The proof is parallel to the bulk case. The main difference from the bulk case is that the edge loop equation is centered around the algebraic relation $s(w)^2 = w$, rather than $s(w)^2 = -1$. Thus the natural quantity to estimate is

$$P(w) := s(w)^2 - w,$$

and one must use a stability estimate to convert moment bounds on $P(w)$ into bounds on $s(w) - \sqrt{w}$.

## C.1 Proof of Equation (3.8) and Equation (3.9)

We start with a deterministic stability estimate.

**Lemma C.1.** *Let*

$$P(w) := s(w)^2 - w. \tag{C.1}$$

*Then, with the principal branch of $\sqrt{w}$,*

$$\min\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\} \leqslant \frac{|P(w)|}{\sqrt{|w|}} \wedge |P(w)|^{1/2}. \tag{C.2}$$

*Proof.* Since

$$P(w) = \big(s(w) - \sqrt{w}\big)\big(s(w) + \sqrt{w}\big),$$

we have

$$\max\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\} \geqslant \frac{1}{2}\left|(s(w) + \sqrt{w}) - (s(w) - \sqrt{w})\right| = |\sqrt{w}| = \sqrt{|w|}.$$

This gives

$$\min\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\} \leqslant \frac{|P(w)|}{\sqrt{|w|}}.$$

The second bound follows from

$$\min\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\} \leqslant \sqrt{|s(w) - \sqrt{w}|\,|s(w) + \sqrt{w}|} = |P(w)|^{1/2}.$$

□

*Proof of* (3.8) *and* (3.9). Throughout the proof, $C$ denotes a constant depending only on $\beta$, whose value may change from line to line. Recall $P(w)$ from (C.1). By the edge loop equation, exactly as in the bulk case, we have

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q \mathbb{E}\big[\mathrm{Im}[s(w)]^q |s(w)|^q\big], \qquad w = E + \mathrm{i}\eta. \tag{C.3}$$

We first derive a general moment bound for $P(w)$. Since $s$ is a Nevanlinna function, $\mathrm{Im}[s(w)] \geqslant 0$, and hence $\mathrm{Im}[s(w)] \leqslant |s(w)|$. Therefore,

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q \mathbb{E}\big[|s(w)|^{2q}\big]. \tag{C.4}$$

Using $s(w)^2 = w + P(w)$, we get

$$|s(w)|^{2q} = |w + P(w)|^q \leqslant C^q\big(|w|^q + |P(w)|^q\big). \tag{C.5}$$

Thus (C.4) and (C.5) together imply

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q |w|^q + \left(\frac{Cq}{\eta^2}\right)^q \mathbb{E}\big[|P(w)|^q\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q |w|^q + \left(\frac{Cq}{\eta^2}\right)^q \mathbb{E}\big[|P(w)|^{2q}\big]^{1/2}.$$

Solving the above quadratic inequality yields

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^{2q} + \left(\frac{Cq}{\eta^2}\right)^q |w|^q. \tag{C.6}$$

We now prove (3.8). Since $\operatorname{Im}[s(w)] \geqslant 0$, we have

$$|s(w) + \sqrt{w}| \geqslant \operatorname{Im}[s(w) + \sqrt{w}] \geqslant \operatorname{Im}[\sqrt{w}].$$

Consequently,

$$|s(w) - \sqrt{w}| \leqslant |s(w) + \sqrt{w}| + 2|\sqrt{w}| \leqslant C\frac{\sqrt{|w|}}{\operatorname{Im}[\sqrt{w}]}|s(w) + \sqrt{w}|.$$

Therefore the stability estimate (C.2) gives,

$$|s(w) - \sqrt{w}| \leqslant C\frac{\sqrt{|w|}}{\operatorname{Im}[\sqrt{w}]}\min\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\} \leqslant C\frac{\sqrt{|w|}}{\operatorname{Im}[\sqrt{w}]}\left(\frac{|P(w)|}{\sqrt{|w|}} \wedge |P(w)|^{1/2}\right). \tag{C.7}$$

If $|w| \geqslant q/\eta^2$, then by (C.7) and (C.6),

$$\mathbb{E}\big[|s(w) - \sqrt{w}|^{2q}\big] \leqslant \frac{C^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \frac{C^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\left[\left(\frac{Cq}{\eta^2}\right)^{2q} + \left(\frac{Cq}{\eta^2}\right)^q |w|^q\right] \leqslant \frac{(Cq)^q}{\eta^{2q}}\frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}.$$

If $|w| < q/\eta^2$, then again by (C.7) and (C.6),

$$\begin{aligned}\mathbb{E}\big[|s(w) - \sqrt{w}|^{2q}\big] &\leqslant C^q\frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\mathbb{E}\big[|P(w)|^q\big] \leqslant C^q\frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\mathbb{E}\big[|P(w)|^{2q}\big]^{1/2} \\ &\leqslant \frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\left[\left(\frac{Cq}{\eta^2}\right)^q + \left(\frac{Cq}{\eta^2}\right)^{q/2}|w|^{q/2}\right] \leqslant \frac{|w|^q}{(\operatorname{Im}[\sqrt{w}])^{2q}}\frac{(Cq)^q}{\eta^{2q}}.\end{aligned}$$

This proves (3.8).

It remains to prove (3.9). Assume $w = E + \mathrm{i}\eta$ and $E \geqslant \eta$. Then $|w| \asymp E$. Moreover,

$$\left|\operatorname{Im}\left[s(w) - \sqrt{w}\right]\right| \leqslant |s(w) - \sqrt{w}|.$$

Since $\operatorname{Im}[s(w)] \geqslant 0$ and $\operatorname{Im}[\sqrt{w}] \geqslant 0$, we also have

$$\left|\operatorname{Im}\left[s(w) - \sqrt{w}\right]\right| = |\operatorname{Im}[s(w)] - \operatorname{Im}[\sqrt{w}]| \leqslant \operatorname{Im}[s(w)] + \operatorname{Im}[\sqrt{w}] = \operatorname{Im}\left[s(w) + \sqrt{w}\right] \leqslant |s(w) + \sqrt{w}|.$$

Hence, with

$$\delta(w) := \min\{|s(w) - \sqrt{w}|,\ |s(w) + \sqrt{w}|\}, \tag{C.8}$$

we have

$$\left|\operatorname{Im}\left[s(w) - \sqrt{w}\right]\right| \leqslant \delta(w). \tag{C.9}$$

By Lemma C.1,

$$\left|\operatorname{Im}\left[s(w) - \sqrt{w}\right]\right| \leqslant \frac{|P(w)|}{\sqrt{|w|}}. \tag{C.10}$$

We next derive a moment bound for $P(w)$ adapted to the regime $E \geqslant \eta$. In this case,

$$|\sqrt{w}| \leqslant C\sqrt{E}, \qquad \mathrm{Im}[\sqrt{w}] \leqslant C\frac{\eta}{\sqrt{E}}. \tag{C.11}$$

Using (C.8), (C.9) and (C.11), we get

$$\mathrm{Im}[s(w)] \leqslant \mathrm{Im}[\sqrt{w}] + \delta(w) \leqslant C\frac{\eta}{\sqrt{E}} + \delta(w), \quad |s(w)| \leqslant |\sqrt{w}| + \delta(w) \leqslant C\sqrt{E} + \delta(w),$$

because $s(w)$ is within distance $\delta(w)$ of either $\sqrt{w}$ or $-\sqrt{w}$. Therefore,

$$\mathrm{Im}[s(w)]|s(w)| \leqslant C\left(\frac{\eta}{\sqrt{E}} + \delta(w)\right)\left(\sqrt{E} + \delta(w)\right) \leqslant C\big(\eta + \sqrt{E}\,\delta(w) + \delta(w)^2\big).$$

By Lemma C.1 and $|w| \asymp E$,

$$\sqrt{E}\,\delta(w) \leqslant C|P(w)|, \qquad \delta(w)^2 \leqslant |P(w)|.$$

Thus

$$\mathrm{Im}[s(w)]|s(w)| \leqslant C\big(\eta + |P(w)|\big). \tag{C.12}$$

Raising (C.12) to the $q$-th power gives

$$\mathrm{Im}[s(w)]^q|s(w)|^q \leqslant C^q\big(\eta^q + |P(w)|^q\big).$$

Substituting this into (C.3), we obtain

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \left(\frac{Cq}{\eta^2}\right)^q (\eta^q + \mathbb{E}[|P(w)|^q]) \leqslant \left(\frac{Cq}{\eta^2}\right)^q \left(\eta^q + \mathbb{E}[|P(w)|^{2q}]^{1/2}\right).$$

Solving this quadratic inequality yields

$$\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \frac{(Cq)^q}{\eta^q} + \frac{(Cq)^{2q}}{\eta^{4q}}. \tag{C.13}$$

Finally, using $|w| \asymp E$, (C.10), and (C.13), we get

$$\mathbb{E}\left[\left|\mathrm{Im}\left[s(w) - \sqrt{w}\right]\right|^{2q}\right] \leqslant \frac{C}{E^q}\mathbb{E}\big[|P(w)|^{2q}\big] \leqslant \frac{(Cq)^q}{\eta^q E^q} + \frac{(Cq)^{2q}}{\eta^{4q}E^q}.$$

This proves (3.9). □

## C.2 Proof of Equation (3.10)

**Lemma C.2** (Edge counting estimate)**.** *Adopt Assumption 1.10. There exists an almost surely finite random variable $M \geqslant 2$ such that, almost surely,*

$$x_1 \leqslant M, \tag{C.14}$$

*and, for every $a \geqslant 2$,*

$$\left|\#\{x \in P : -a \leqslant x \leqslant M\} - \frac{2}{3\pi}a^{3/2}\right| \leqslant M(\log a)^2. \tag{C.15}$$

*Proof of* (C.14). We first prove that there exists an almost surely finite random integer $M$ such that

$$\mathrm{Im}[s(E+\mathrm{i})] \leqslant \frac{1}{4}, \qquad E \geqslant M-1, \quad E \in \mathbb{Z}. \tag{C.16}$$

Fix $q = 2$. By (3.9), with $\eta = 1$, for all sufficiently large integers $E$,

$$\mathbb{E}\left[\left|\operatorname{Im}\left[s(E+\mathrm{i}) - \sqrt{E+\mathrm{i}}\right]\right|^4\right] \leqslant \frac{C}{E^2}. \tag{C.17}$$

Here we used that the assumptions of (3.9) are satisfied for $\eta = 1$, $q = 2$, and all sufficiently large $E$.

Moreover, for all sufficiently large $E$,

$$\operatorname{Im}[\sqrt{E+\mathrm{i}}] \leqslant \frac{1}{2\sqrt{E}} \leqslant \frac{1}{8}.$$

Therefore,

$$\left\{\operatorname{Im}[s(E+\mathrm{i})] > \frac{1}{4}\right\} \subset \left\{\left|\operatorname{Im}\left[s(E+\mathrm{i}) - \sqrt{E+\mathrm{i}}\right]\right| > \frac{1}{8}\right\}.$$

By Markov's inequality and (C.17),

$$\mathbb{P}\left(\operatorname{Im}[s(E+\mathrm{i})] > \frac{1}{4}\right) \leqslant 8^4 \mathbb{E}\left[\left|\operatorname{Im}\left[s(E+\mathrm{i}) - \sqrt{E+\mathrm{i}}\right]\right|^4\right] \leqslant \frac{C}{E^2}.$$

Consequently,

$$\sum_{E=1}^{\infty} \mathbb{P}\left(\operatorname{Im}[s(E+\mathrm{i})] > \frac{1}{4}\right) < \infty.$$

By the Borel–Cantelli lemma, almost surely only finitely many of the events

$$\left\{\operatorname{Im}[s(E+\mathrm{i})] > \frac{1}{4}\right\}, \qquad E \in \mathbb{Z}_{\geqslant 1},$$

occur. Thus there exists an almost surely finite random integer $M$ such that (C.16) holds.

We now show that this implies boundedness of the top particle. Suppose, for contradiction, that there exists a particle $x \in P$ with $x \geqslant M$. Choose an integer $E \geqslant M - 1$ such that $x \in [E, E+1]$. Evaluating at $z = E + \mathrm{i}$, the contribution of this single particle gives

$$\operatorname{Im}[s(E+\mathrm{i})] = c + \sum_{y \in P} \frac{1}{(y-E)^2 + 1} \geqslant \frac{1}{(x-E)^2 + 1} \geqslant \frac{1}{2},$$

since $c \geqslant 0$ and $0 \leqslant x - E \leqslant 1$. This contradicts (C.16). Therefore no particle can lie in $[M, \infty)$, and hence

$$x_1 \leqslant M.$$

After enlarging $M$, we may also assume $M \geqslant 2$. This proves (C.14). ☐

*Proof of* (C.15). The edge concentration estimates (3.8) and (3.9), together with the net argument, imply that, almost surely, after possibly enlarging $M$, for every dyadic scale $2^n$,

$$\begin{aligned} |s(x+\mathrm{i}y) - \sqrt{x+\mathrm{i}y}| &\leqslant \frac{Mn}{y}, \qquad 2^{-n} \leqslant y \leqslant 2^n, \quad -2^n \leqslant x \leqslant y, \\ |\operatorname{Im}[s(x+\mathrm{i}y) - \sqrt{x+\mathrm{i}y}]| &\leqslant \frac{Mn}{y}, \qquad 0 < y \leqslant 2^n, \quad |x| \leqslant 2^n. \end{aligned} \tag{C.18}$$

This is the edge analogue of the almost-sure rectangle bound in the bulk case Lemma 3.9, and can be proven in the same way. So we omit its proof.

Fix $a \geqslant 2$, and choose $n$ such that

$$2^{n-1} \leqslant a \leqslant 2^n.$$

Choose a smooth nonnegative function $f_+$ such that

$$f_+ = 1 \quad \text{on } [-a, M], \qquad f_+ = 0 \quad \text{outside } [-a - a^{-1/2}, M+1],$$

with

$$|f_+'| \lesssim a^{1/2}, \qquad |f_+''| \lesssim a$$

on the interval $[-a - a^{-1/2}, -a]$, and

$$|f_+'| + |f_+''| \lesssim 1$$

on the interval $[M, M+1]$.

Similarly, choose a smooth nonnegative function $f_-$ such that

$$f_- = 1 \quad \text{on } [-a + a^{-1/2}, M-1], \qquad f_- = 0 \quad \text{outside } [-a, M],$$

with

$$|f_-'| \lesssim a^{1/2}, \qquad |f_-''| \lesssim a$$

on the interval $[-a, -a + a^{-1/2}]$, and

$$|f_-'| + |f_-''| \lesssim 1$$

on the interval $[M-1, M]$. In particular,

$$f_- \leqslant \mathbf{1}_{[-a,M]} \leqslant f_+.$$

We claim that for either $f = f_+$ or $f = f_-$,

$$\left| \sum_{x \in P} f(x) - \frac{1}{\pi} \int_{-\infty}^{0} f(x) \sqrt{-x} \, \mathrm{d}x \right| \leqslant CMn^2. \tag{C.19}$$

We now prove (C.19). Let $\chi \in C_c^\infty(\mathbb{R})$ satisfy

$$\chi = 1 \quad \text{on } [-a, a], \qquad \chi = 0 \quad \text{outside } [-a-1, a+1], \qquad |\chi'| \lesssim 1,$$

and set

$$\widetilde{f}(x + \mathrm{i}y) := (f(x) + \mathrm{i}y f'(x))\chi(y).$$

We apply Lemma 3.11 with

$$s_\varrho(z) = \sqrt{z}, \qquad \varrho(x) = \frac{\sqrt{-x}}{\pi} \mathbf{1}_{x \leqslant 0}.$$

For the $f''$-term near the fixed endpoint $M$, we split the $y$-integral at $\eta = M+1$. For the $f''$-term near the moving endpoint $-a$, we split the $y$-integral at $\eta = a^{-1/2}$. Using (C.18) as input, Lemma 3.11 gives (C.19).

Now apply (C.19) to $f_+$. Since $f_+ \geqslant \mathbf{1}_{[-a,M]}$,

$$\#\{x \in P : -a \leqslant x \leqslant M\} \leqslant \sum_{x \in P} f_+(x).$$

Moreover,

$$\frac{1}{\pi} \int_{-\infty}^{0} f_+(x) \sqrt{-x} \, \mathrm{d}x = \frac{2}{3\pi} a^{3/2} + \mathrm{O}(1),$$

because the smoothing interval near $-a$ has length $a^{-1/2}$, and hence its contribution to the reference integral is $\mathrm{O}(1)$. Therefore

$$\#\{x \in P : -a \leqslant x \leqslant M\} \leqslant \frac{2}{3\pi} a^{3/2} + CMn^2 + \mathrm{O}(1).$$

Similarly, applying (C.19) to $f_-$, and using $f_- \leqslant \mathbf{1}_{[-a,M]}$, gives

$$\#\{x \in P : -a \leqslant x \leqslant M\} \geqslant \frac{2}{3\pi} a^{3/2} - CMn^2 - \mathrm{O}(1).$$

Since $n \asymp \log a$, after enlarging $M$ once more we obtain

$$\left| \#\{x \in P : -a \leqslant x \leqslant M\} - \frac{2}{3\pi} a^{3/2} \right| \leqslant M(\log a)^2. \tag{C.20}$$

This proves (C.15). □

*Proof of* (3.10). Since

$$\#\{x \in P : -a \leqslant x \leqslant M\} \to \infty \qquad \text{as } a \to \infty,$$

the poles may be enumerated in decreasing order as

$$x_1 \geqslant x_2 \geqslant x_3 \geqslant \cdots, \qquad x_j \to -\infty.$$

The estimate (3.10) then follows by rearranging the counting estimate (C.15). □

## C.3 Proof of Equation (3.11)

*Proof of* (3.11). We recall from (A.2),

$$\left| \mathfrak{a}_i + \left( \frac{3\pi i}{2} \right)^{2/3} \right| \lesssim i^{-1/3}. \tag{C.21}$$

Combining (3.10) with (C.21), we obtain

$$|x_i - \mathfrak{a}_i| \leqslant M \frac{(\log(2+i))^2}{i^{1/3}}, \qquad i \geqslant 1. \tag{C.22}$$

Moreover,

$$|x_i| \asymp i^{2/3}, \qquad |\mathfrak{a}_i| \asymp i^{2/3}, \qquad i \geqslant 1. \tag{C.23}$$

We first prove that the right-hand side of (3.11) is well-defined. Let $w$ lie in a compact subset of $\mathbb{C} \setminus \mathbb{R}$. Then, for $i$ large enough, $|x_i - w| \asymp i^{2/3}$, $|\mathfrak{a}_i| \asymp i^{2/3}$. Hence

$$\left| \frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i} \right| = \left| \frac{\mathfrak{a}_i - x_i + w}{(x_i - w)\mathfrak{a}_i} \right| \leqslant C_w \left( \frac{1}{i^{4/3}} + \frac{(\log(2+i))^2}{i^{5/3}} \right).$$

The right-hand side is summable in $i$. Thus the series in (3.11) converges absolutely and locally uniformly on $\mathbb{C} \setminus \mathbb{R}$. Define

$$\widetilde{s}(w) := \sum_{i=1}^{\infty} \left( \frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i} \right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}.$$

We compare $\widetilde{s}$ with the Airy Nevanlinna function. Recall the standard partial fraction expansion (A.3):

$$-\frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)} = \sum_{i=1}^{\infty} \left( \frac{1}{\mathfrak{a}_i - w} - \frac{1}{\mathfrak{a}_i} \right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}. \tag{C.24}$$

Therefore

$$\widetilde{s}(w) + \frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)} = \sum_{i=1}^{\infty} \left( \frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i - w} \right). \tag{C.25}$$

We claim that the right-hand side tends to 0 along the imaginary axis. Indeed, let $w = \mathrm{i}y$, with $y \geqslant 2$, and split the sum into $i \leqslant y^{3/2}$ and $i > y^{3/2}$. For $i \leqslant y^{3/2}$, we have $|x_i - \mathrm{i}y|\,|\mathfrak{a}_i - \mathrm{i}y| \geqslant cy^2$, and therefore, by (C.22),

$$\sum_{i \leqslant y^{3/2}} \left| \frac{1}{x_i - \mathrm{i}y} - \frac{1}{\mathfrak{a}_i - \mathrm{i}y} \right| \leqslant \frac{C}{y^2} \sum_{i \leqslant y^{3/2}} \frac{(\log(2+i))^2}{i^{1/3}} \leqslant \frac{C(\log y)^2}{y}.$$

For $i > y^{3/2}$, using (C.23), we get

$$\sum_{i > y^{3/2}} \left| \frac{1}{x_i - \mathrm{i}y} - \frac{1}{\mathfrak{a}_i - \mathrm{i}y} \right| \leqslant C \sum_{i > y^{3/2}} \frac{(\log(2+i))^2}{i^{5/3}} \leqslant \frac{C(\log y)^2}{y}.$$

Thus

$$\widetilde{s}(\mathrm{i}y) + \frac{\mathrm{Ai}'(\mathrm{i}y)}{\mathrm{Ai}(\mathrm{i}y)} \longrightarrow 0, \qquad y \to \infty. \tag{C.26}$$

By the Airy asymptotic (A.4),

$$-\frac{\mathrm{Ai}'(w)}{\mathrm{Ai}(w)} = \sqrt{w} + \mathrm{O}(|w|^{-1}),$$

uniformly in fixed sectors away from the negative real axis. Together with (C.26), this gives

$$\widetilde{s}(\mathrm{i}y) - \sqrt{\mathrm{i}y} \longrightarrow 0, \qquad y \to \infty. \tag{C.27}$$

On the other hand, (C.18) implies

$$|s(\mathrm{i}y) - \sqrt{\mathrm{i}y}| \leqslant \frac{M \log(2+y)}{y} \longrightarrow 0, \qquad y \to +\infty. \tag{C.28}$$

It follows from (C.27) and (C.28) that

$$s(\mathrm{i}y) - \widetilde{s}(\mathrm{i}y) \longrightarrow 0, \qquad y \to \infty.$$

It remains to identify $s$ and $\widetilde{s}$. Both are Nevanlinna functions with the same representing measure $\sum_i \delta_{x_i}$. Hence their difference is an affine function:

$$s(w) - \widetilde{s}(w) = b + cw, \qquad b \in \mathbb{R}, \quad c \geqslant 0.$$

Since $s(\mathrm{i}y) - \widetilde{s}(\mathrm{i}y) \to 0$ as $y \to \infty$, we must have $c = 0$ and $b = 0$. Therefore $s = \widetilde{s}$, that is,

$$s(w) = \sum_{i \geqslant 1} \left( \frac{1}{x_i - w} - \frac{1}{\mathfrak{a}_i} \right) - \frac{\mathrm{Ai}'(0)}{\mathrm{Ai}(0)}.$$

This proves (3.11). □

# D Explicit solutions for $\beta = 2$ bulk system

We prove Theorem 6.1 in this section. The key idea is to remove the mixed interaction terms by factoring out the mixed Vandermonde product

$$P(\boldsymbol{t}, \boldsymbol{s}) = \prod_{i=1}^{n} \prod_{a=1}^{m} (t_i - s_a).$$

After writing $F = P\,G$, the conjugated equations for $G$ become the ordinary scalar $A_{n+m-1}$-type system in the combined variables

$$x_1 = t_1, \ldots, x_n = t_n, \qquad x_{n+a} = s_a.$$

Thus, when $\beta = 2$, the deformed system is gauge-equivalent, through multiplication by $P(\boldsymbol{t}, \boldsymbol{s})$, to the standard scalar $A$-type CMS system (D.5)–(D.6); see [75].

## D.1 Reduction to the $A$-type CMS system

Let $(\boldsymbol{t}, \boldsymbol{s}) = (t_1, \cdots, t_n, s_1, \cdots, s_m)$, and

$$P(\boldsymbol{t}, \boldsymbol{s}) := \prod_{i=1}^{n} \prod_{a=1}^{m} (t_i - s_a), \qquad F(\boldsymbol{t}, \boldsymbol{s}) = P(\boldsymbol{t}, \boldsymbol{s})\, G(\boldsymbol{t}, \boldsymbol{s}).$$

Introduce

$$A_i := \partial_{t_i} \log P = \sum_{a=1}^{m} \frac{1}{t_i - s_a}, \qquad B_a := \partial_{s_a} \log P = \sum_{i=1}^{n} \frac{1}{s_a - t_i}.$$

Then

$$\partial_{t_i} F = P(\partial_{t_i} + A_i) G, \quad \partial_{t_i}^2 F = P\Big(\partial_{t_i}^2 + 2A_i \partial_{t_i} + \partial_{t_i} A_i + A_i^2\Big) G, \tag{D.1}$$

and similarly

$$\partial_{s_a} F = P(\partial_{s_a} + B_a) G, \quad \partial_{s_a}^2 F = P\Big(\partial_{s_a}^2 + 2B_a \partial_{s_a} + \partial_{s_a} B_a + B_a^2\Big) G. \tag{D.2}$$

**Theorem D.1.** *Assume $\beta = 2$. Then $F = P\,G$ satisfies the original $\beta = 2$ system (5.5) if and only if $G$ satisfies*

$$\left( \partial_{t_i}^2 + 1 + \sum_{j \neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} + \sum_{a=1}^{m} \frac{\partial_{t_i} - \partial_{s_a}}{t_i - s_a} \right) G = 0, \qquad i = 1, \ldots, n, \tag{D.3}$$

*and*

$$\left( \partial_{s_a}^2 + 1 + \sum_{b \neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} + \sum_{i=1}^{n} \frac{\partial_{s_a} - \partial_{t_i}}{s_a - t_i} \right) G = 0, \qquad a = 1, \ldots, m. \tag{D.4}$$

*Equivalently, after setting*

$$x_1 = t_1, \ldots, x_n = t_n, \qquad x_{n+1} = s_1, \ldots, x_{n+m} = s_m,$$

*the function $G$ satisfies*

$$\mathcal{L}_r G = 0, \qquad r = 1, \ldots, n + m, \tag{D.5}$$

*where*

$$\mathcal{L}_r := \partial_{x_r}^2 + 1 + \sum_{u \neq r} \frac{\partial_{x_r} - \partial_{x_u}}{x_r - x_u}. \tag{D.6}$$

*Proof.* We work on the complement of the collision hyperplanes, i.e. (6.4). We first treat the $t_i$-equation; the $s_a$-equation is analogous.

Substituting $F = PG$ into the $t_i$-equation of the original $\beta = 2$ system (5.5) and dividing by $P$, using (D.1) and (D.2), gives

$$0 = \Big(\partial_{t_i}^2 + 2A_i \partial_{t_i} + \partial_{t_i} A_i + A_i^2 + 1\Big) G + \sum_{j \neq i} \frac{(\partial_{t_i} + A_i) - (\partial_{t_j} + A_j)}{t_i - t_j} G - \sum_{a=1}^{m} \frac{(\partial_{t_i} + A_i) + (\partial_{s_a} + B_a)}{t_i - s_a} G.$$

Collecting the first-order terms involving the mixed variables, we find

$$2A_i \partial_{t_i} - \sum_{a=1}^{m} \frac{\partial_{t_i} + \partial_{s_a}}{t_i - s_a} = \sum_{a=1}^{m} \frac{\partial_{t_i} - \partial_{s_a}}{t_i - s_a},$$

since $A_i = \sum_a (t_i - s_a)^{-1}$. Therefore

$$0 = \left( \partial_{t_i}^2 + 1 + \sum_{j \neq i} \frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} + \sum_{a=1}^{m} \frac{\partial_{t_i} - \partial_{s_a}}{t_i - s_a} \right) G + Z_i G,$$

where

$$Z_i = \partial_{t_i} A_i + A_i^2 + \sum_{j \neq i} \frac{A_i - A_j}{t_i - t_j} - \sum_{a=1}^{m} \frac{A_i + B_a}{t_i - s_a}.$$

It remains to show that $Z_i = 0$. Set

$$c_{ia} := \frac{1}{t_i - s_a}.$$

Then

$$A_i = \sum_{a=1}^{m} c_{ia}, \qquad B_a = -\sum_{i=1}^{n} c_{ia}.$$

We compute the three contributions to $Z_i$. First,

$$\partial_{t_i} A_i + A_i^2 = -\sum_a c_{ia}^2 + \left( \sum_a c_{ia} \right)^2 = 2 \sum_{a<b} c_{ia} c_{ib}.$$

Second,

$$\frac{A_i - A_j}{t_i - t_j} = \sum_{a=1}^{m} \frac{c_{ia} - c_{ja}}{t_i - t_j} = -\sum_{a=1}^{m} c_{ia} c_{ja},$$

and hence

$$\sum_{j \neq i} \frac{A_i - A_j}{t_i - t_j} = -\sum_{j \neq i} \sum_{a=1}^{m} c_{ia} c_{ja}.$$

Third, since

$$A_i + B_a = \sum_{b \neq a} c_{ib} - \sum_{j \neq i} c_{ja},$$

we have

$$\sum_{a=1}^{m} \frac{A_i + B_a}{t_i - s_a} = \sum_{a=1}^{m} \sum_{b \neq a} c_{ia} c_{ib} - \sum_{a=1}^{m} \sum_{j \neq i} c_{ia} c_{ja} = 2 \sum_{a<b} c_{ia} c_{ib} - \sum_{j \neq i} \sum_{a=1}^{m} c_{ia} c_{ja}.$$

Combining these identities gives $Z_i = 0$, and therefore (D.3) follows.

The $s_a$-equation is handled in the same way. Substituting $F = PG$ into the $s_a$-equation and collecting first-order terms gives

$$0 = \left( \partial_{s_a}^2 + 1 + \sum_{b \neq a} \frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} + \sum_{i=1}^{n} \frac{\partial_{s_a} - \partial_{t_i}}{s_a - t_i} \right) G + W_a G,$$

where

$$W_a = \partial_{s_a} B_a + B_a^2 + \sum_{b \neq a} \frac{B_a - B_b}{s_a - s_b} - \sum_{i=1}^{n} \frac{B_a + A_i}{s_a - t_i}.$$

The same algebra as above gives $W_a = 0$. Hence (D.4) follows.

Finally, the unified form (D.5) is immediate after the identification

$$x_1 = t_1, \ldots, x_n = t_n, \qquad x_{n+1} = s_1, \ldots, x_{n+m} = s_m.$$

This proves the theorem. $\square$

## D.2 Proof of Theorem 6.1

*Proof of Theorem 6.1.* Set

$$x_1 = t_1, \dots, x_n = t_n, \qquad x_{n+1} = s_1, \dots, x_{n+m} = s_m,$$

and write

$$\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{n+m}) := (\epsilon_1, \dots, \epsilon_n, \eta_1, \dots, \eta_m) \in \{\pm 1\}^{n+m}.$$

Let

$$P(\boldsymbol{t}, \boldsymbol{s}) := \prod_{i=1}^{n} \prod_{a=1}^{m} (t_i - s_a).$$

By Theorem D.1, a function of the form

$$F(\boldsymbol{t}, \boldsymbol{s}) = P(\boldsymbol{t}, \boldsymbol{s})\, G(\boldsymbol{x})$$

solves the original $\beta = 2$ system if and only if $G$ solves the reduced system

$$\mathcal{L}_r G = 0, \qquad r = 1, \dots, n+m, \tag{D.7}$$

where

$$\mathcal{L}_r := \partial_{x_r}^2 + 1 + \sum_{u \neq r} \frac{\partial_{x_r} - \partial_{x_u}}{x_r - x_u}.$$

*Step 1. Explicit solutions of the reduced system.* For $\boldsymbol{\sigma} \in \{\pm 1\}^{n+m}$, define

$$\Delta_{\boldsymbol{\sigma}}(\boldsymbol{x}) := \prod_{1 \leqslant r < u \leqslant n+m} (x_r - x_u)^{(\sigma_r \sigma_u - 1)/2}. \tag{D.8}$$

Since $(\sigma_r \sigma_u - 1)/2 \in \{0, -1\}$, the function $\Delta_{\boldsymbol{\sigma}}$ is rational. Set

$$G_{\boldsymbol{\sigma}}(\boldsymbol{x}) := \exp\left( \mathrm{i} \sum_{r=1}^{n+m} \sigma_r x_r \right) \Delta_{\boldsymbol{\sigma}}(\boldsymbol{x}). \tag{D.9}$$

We claim that

$$\mathcal{L}_r G_{\boldsymbol{\sigma}} = 0, \qquad r = 1, \dots, n+m.$$

Fix $r$, and introduce

$$J_r := \{u \neq r : \sigma_u \neq \sigma_r\}, \qquad K_r := \{u \neq r : \sigma_u = \sigma_r\}.$$

Then

$$\psi_r := \partial_{x_r} \log \Delta_{\boldsymbol{\sigma}} = -\sum_{u \in J_r} \frac{1}{x_r - x_u}.$$

Hence

$$\frac{\partial_{x_r} G_{\boldsymbol{\sigma}}}{G_{\boldsymbol{\sigma}}} = \mathrm{i}\sigma_r + \psi_r, \qquad \frac{\partial_{x_r}^2 G_{\boldsymbol{\sigma}}}{G_{\boldsymbol{\sigma}}} = (\mathrm{i}\sigma_r + \psi_r)^2 + \partial_{x_r} \psi_r.$$

Therefore

$$\frac{\mathcal{L}_r G_{\boldsymbol{\sigma}}}{G_{\boldsymbol{\sigma}}} = (\mathrm{i}\sigma_r)^2 + 1 + \partial_{x_r}\psi_r + \psi_r^2 + 2\mathrm{i}\sigma_r \psi_r + \sum_{u \neq r} \frac{\mathrm{i}(\sigma_r - \sigma_u) + \psi_r - \psi_u}{x_r - x_u}. \tag{D.10}$$

Since $(\mathrm{i}\sigma_r)^2 + 1 = 0$, it remains to check that the rational terms cancel.

First, the exponential contributions cancel:

$$2\mathrm{i}\sigma_r\psi_r + \sum_{u\in J_r}\frac{\mathrm{i}(\sigma_r-\sigma_u)}{x_r-x_u} = -2\mathrm{i}\sigma_r\sum_{u\in J_r}\frac{1}{x_r-x_u} + 2\mathrm{i}\sigma_r\sum_{u\in J_r}\frac{1}{x_r-x_u} = 0.$$

For the remaining terms, we compute

$$\partial_{x_r}\psi_r + \psi_r^2 = 2\sum_{a\in J_r}\frac{1}{(x_r-x_a)^2} + \sum_{\substack{a,b\in J_r\\ a\neq b}}\frac{1}{(x_r-x_a)(x_r-x_b)}.$$

Moreover,

$$\sum_{u\in K_r}\frac{\psi_r-\psi_u}{x_r-x_u} = \sum_{u\in K_r}\sum_{a\in J_r}\frac{1}{(x_r-x_a)(x_u-x_a)},$$

while

$$\sum_{a\in J_r}\frac{\psi_r-\psi_a}{x_r-x_a} = -2\sum_{a\in J_r}\frac{1}{(x_r-x_a)^2} - \sum_{\substack{a,b\in J_r\\ a\neq b}}\frac{1}{(x_r-x_a)(x_r-x_b)} - \sum_{a\in J_r}\sum_{u\in K_r}\frac{1}{(x_r-x_a)(x_u-x_a)}.$$

These three expressions cancel term by term. Thus

$$\mathcal{L}_r G_{\boldsymbol{\sigma}} = 0, \qquad r = 1,\dots,n+m.$$

*Step 2. Lifting back to the original system.* Define

$$F_{\boldsymbol{\sigma}}(\boldsymbol{t},\boldsymbol{s}) := P(\boldsymbol{t},\boldsymbol{s})\,G_{\boldsymbol{\sigma}}(\boldsymbol{t},\boldsymbol{s}).$$

Since $G_{\boldsymbol{\sigma}}$ solves the reduced system (D.7), Theorem D.1 implies that $F_{\boldsymbol{\sigma}}$ solves the original $\beta=2$ system.

We now identify $F_{\boldsymbol{\sigma}}$. By definition,

$$\Delta_{\boldsymbol{\sigma}}(x) = \prod_{1\leqslant i<j\leqslant n}(t_i-t_j)^{(\epsilon_i\epsilon_j-1)/2}\prod_{1\leqslant a<b\leqslant m}(s_a-s_b)^{(\eta_a\eta_b-1)/2}\prod_{i=1}^{n}\prod_{a=1}^{m}(t_i-s_a)^{(\epsilon_i\eta_a-1)/2}.$$

Multiplication by

$$P(t,s) = \prod_{i=1}^{n}\prod_{a=1}^{m}(t_i-s_a)$$

changes the mixed exponent from $(\epsilon_i\eta_a-1)/2$ to $(1+\epsilon_i\eta_a)/2$. Therefore

$$\begin{aligned}F_{\boldsymbol{\sigma}}(\boldsymbol{t},\boldsymbol{s}) = \exp\left(\mathrm{i}\sum_{i=1}^{n}\epsilon_i t_i + \mathrm{i}\sum_{a=1}^{m}\eta_a s_a\right)&\prod_{1\leqslant i<j\leqslant n}(t_i-t_j)^{(\epsilon_i\epsilon_j-1)/2}\\ \times\prod_{1\leqslant a<b\leqslant m}(s_a-s_b)^{(\eta_a\eta_b-1)/2}&\prod_{i=1}^{n}\prod_{a=1}^{m}(t_i-s_a)^{(1+\epsilon_i\eta_a)/2}.\end{aligned}$$

This is exactly (6.1). Since all exponents are integers, these solutions are single-valued.

*Step 3. Linear independence.* Choose $\boldsymbol{v} = (v_1,\cdots,v_{n+m})\in\mathbb{C}^{n+m}$ such that

$$v_r \neq v_u \quad (r\neq u),$$

and

$$\mathrm{Re}[\mathrm{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{v}] \neq \mathrm{Re}[\mathrm{i}\,\boldsymbol{\sigma}'\cdot\boldsymbol{v}] \qquad \text{for } \boldsymbol{\sigma}\neq\boldsymbol{\sigma}'.$$

Such a choice is possible after excluding finitely many proper real hyperplanes.

Along the ray $\boldsymbol{x} = R\boldsymbol{v}$, as $R \to +\infty$, we have

$$G_{\boldsymbol{\sigma}}(R\boldsymbol{v}) = c_{\boldsymbol{\sigma}} R^{-N_+(\boldsymbol{\sigma})N_-(\boldsymbol{\sigma})} \exp(\mathrm{i}R\,\boldsymbol{\sigma}\cdot\boldsymbol{v}) \left(1 + \mathrm{O}(R^{-1})\right),$$

where

$$c_{\boldsymbol{\sigma}} = \prod_{1\leqslant r<u\leqslant n+m} (v_r - v_u)^{(\sigma_r\sigma_u - 1)/2} \neq 0, \quad N_\pm(\boldsymbol{\sigma}) := \#\{r : \sigma_r = \pm 1\}.$$

The numbers $\mathrm{Re}[\mathrm{i}\,\boldsymbol{\sigma}\cdot\boldsymbol{v}]$ are pairwise distinct, so the exponentials $\exp(\mathrm{i}R\,\boldsymbol{\sigma}\cdot\boldsymbol{v})$ have distinct growth rates along this ray. Hence no nontrivial linear combination of the $G_{\boldsymbol{\sigma}}$ can vanish identically.

Therefore the functions $G_{\boldsymbol{\sigma}}$ are linearly independent. Since $F_{\boldsymbol{\sigma}} = P\,G_{\boldsymbol{\sigma}}$ and $P$ is not identically zero, the functions $F_{\boldsymbol{\sigma}}$ are also linearly independent. This proves the theorem. ☐

# E Explicit solutions for $\beta = 2$ edge system

## E.1 Reduction to the edge $A$-type CMS system

Let

$$(\boldsymbol{t}, \boldsymbol{s}) = (t_1, \ldots, t_n, s_1, \ldots, s_m),$$

and define

$$P(\boldsymbol{t}, \boldsymbol{s}) := \prod_{i=1}^{n}\prod_{a=1}^{m}(t_i - s_a), \qquad F(\boldsymbol{t}, \boldsymbol{s}) = P(\boldsymbol{t}, \boldsymbol{s})\,G(\boldsymbol{t}, \boldsymbol{s}).$$

**Theorem E.1** (Reduction to the scalar edge $A_{n+m-1}$-system)**.** *Assume $\beta = 2$. Then $F = P\,G$ satisfies the original edge system (5.20) if and only if $G$ satisfies*

$$\left(\partial_{t_i}^2 - t_i + \sum_{j\neq i}\frac{\partial_{t_i} - \partial_{t_j}}{t_i - t_j} + \sum_{a=1}^{m}\frac{\partial_{t_i} - \partial_{s_a}}{t_i - s_a}\right) G = 0, \qquad i = 1, \ldots, n, \tag{E.1}$$

*and*

$$\left(\partial_{s_a}^2 - s_a + \sum_{b\neq a}\frac{\partial_{s_a} - \partial_{s_b}}{s_a - s_b} + \sum_{i=1}^{n}\frac{\partial_{s_a} - \partial_{t_i}}{s_a - t_i}\right) G = 0, \qquad a = 1, \ldots, m. \tag{E.2}$$

*Equivalently, after setting*

$$x_1 = t_1, \ldots, x_n = t_n, \qquad x_{n+1} = s_1, \ldots, x_{n+m} = s_m,$$

*the function $G$ satisfies*

$$\mathsf{L}_r G = 0, \qquad r = 1, \ldots, n+m, \tag{E.3}$$

*where*

$$\mathsf{L}_r := \partial_{x_r}^2 - x_r + \sum_{u\neq r}\frac{\partial_{x_r} - \partial_{x_u}}{x_r - x_u}. \tag{E.4}$$

*Proof.* This is the same conjugation as in the bulk $\beta = 2$ case. Indeed, after writing $F = P\,G$, the logarithmic derivatives

$$A_i := \partial_{t_i}\log P = \sum_{a=1}^{m}\frac{1}{t_i - s_a}, \qquad B_a := \partial_{s_a}\log P = \sum_{i=1}^{n}\frac{1}{s_a - t_i}$$

are exactly the same as in the proof of Theorem D.1. Hence the first-order terms conjugate in the same way, and the same $c_{ia}$-algebra shows that all zero-order terms coming from $P$ cancel.

The only difference from the bulk computation is that the constant term $+1$ is replaced by the linear potential $-t_i$ or $-s_a$. These terms are unaffected by conjugation with $P$. Therefore the bulk proof carries over verbatim and yields (E.1) and (E.2). The unified form (E.3) is then immediate from the identification of the $x$-variables. $\square$

## E.2 Proof of Theorem 6.4

The proof is parallel in spirit to the bulk case. The reduced scalar equation is now the edge $A$-type Calogero system (E.3), so the role of the plane waves is played by Airy solutions and a Slater determinant.

*Proof of Theorem 6.4.* We write

$$\boldsymbol{x} = (x_1, \dots, x_{n+m}), \qquad x_1 = t_1, \dots, x_n = t_n, \qquad x_{n+1} = s_1, \dots, x_{n+m} = s_m.$$

Let

$$\Delta(\boldsymbol{x}) := \prod_{1 \leqslant r < u \leqslant n+m} (x_r - x_u).$$

For each sign pattern

$$\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{n+m}) \in \{\pm 1\}^{n+m},$$

choose Airy solutions $\phi_{\sigma_r}$ (recall from (6.34)) of

$$\phi''_{\sigma_r}(x) - x\,\phi_{\sigma_r}(x) = 0.$$

Define

$$D_{\boldsymbol{\sigma}}(\boldsymbol{x}) := \det\big[\partial_x^{j-1}\phi_{\sigma_r}(x_r)\big]_{1 \leqslant r, j \leqslant n+m},$$

and

$$G_{\boldsymbol{\sigma}}(\boldsymbol{x}) := \frac{D_{\boldsymbol{\sigma}}(\boldsymbol{x})}{\Delta(\boldsymbol{x})}. \tag{E.5}$$

The above determinant quotient, with an arbitrary row-by-row choice of one-variable Airy solutions, gives local solutions of (E.3) on the configuration space $x_r \neq x_u$. The special symmetric solution obtained by taking $\phi_1 = \cdots = \phi_{n+m} = \mathrm{Ai}$ is regular across the diagonals and agrees, up to normalization, with the standard $\beta = 2$ multivariate Airy function, equivalently the Kontsevich matrix Airy function; see [30, 61].

*Step 1. Solving the reduced system.* We first prove that

$$\mathsf{L}_r G_{\boldsymbol{\sigma}} = 0, \qquad r = 1, \dots, n+m.$$

Equivalently,

$$\left(\partial_{x_r}^2 - x_r + \sum_{u \neq r} \frac{\partial_{x_r} - \partial_{x_u}}{x_r - x_u}\right) G_{\boldsymbol{\sigma}} = 0. \tag{E.6}$$

Let

$$\mathfrak{X} := \{\boldsymbol{x} \in \mathbb{C}^{n+m} : \ x_r \neq x_u \text{ for } r \neq u\}.$$

We prove the identity on $\mathfrak{X}$, where the operator is well defined.

Since the Airy equation has a two-dimensional space of solutions spanned by Airy contour branches, and since $D_{\boldsymbol{\sigma}}$ is multilinear in its rows, it suffices to prove the claim when each $\phi_{\sigma_a}$ is represented by a single admissible Airy contour integral. Thus, for each $a = 1, \dots, n+m$, choose an admissible Airy contour $\Gamma_{\sigma_a}$ such that

$$\phi_{\sigma_a}(x) = \int_{\Gamma_{\sigma_a}} e^{z^3/3 - xz}\,\mathrm{d}z.$$

Then

$$\partial_x^{j-1}\phi_{\sigma_a}(x_a) = (-1)^{j-1}\int_{\Gamma_{\sigma_a}} z^{j-1}e^{z^3/3-x_a z}\,\mathrm{d}z.$$

By multilinearity of the determinant in the rows,

$$\begin{aligned} D_{\boldsymbol{\sigma}}(\boldsymbol{x}) &= (-1)^{\frac{(n+m)(n+m-1)}{2}}\int_{\Gamma_{\sigma_1}\times\cdots\times\Gamma_{\sigma_{n+m}}} \det[z_a^{j-1}]_{a,j=1}^{n+m}\exp\left(\sum_{a=1}^{n+m}\left(\frac{z_a^3}{3}-x_a z_a\right)\right)\mathrm{d}\boldsymbol{z} \\ &= c_{n+m}\int_{\Gamma_{\boldsymbol{\sigma}}}\Delta(\boldsymbol{z})e^{S(\boldsymbol{z};\boldsymbol{x})}\,\mathrm{d}\boldsymbol{z}, \end{aligned} \tag{E.7}$$

where

$$c_{n+m} := (-1)^{\frac{(n+m)(n+m-1)}{2}}, \qquad \Gamma_{\boldsymbol{\sigma}} := \Gamma_{\sigma_1}\times\cdots\times\Gamma_{\sigma_{n+m}},$$

$$S(\boldsymbol{z};\boldsymbol{x}) := \sum_{a=1}^{n+m}\left(\frac{z_a^3}{3}-x_a z_a\right), \qquad \mathrm{d}\boldsymbol{z} := \mathrm{d}z_1\cdots\mathrm{d}z_{n+m}.$$

Hence, for $\boldsymbol{x}\in\mathfrak{X}$,

$$G_{\boldsymbol{\sigma}}(\boldsymbol{x}) = c_{n+m}\int_{\Gamma_{\boldsymbol{\sigma}}} K(\boldsymbol{x},\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}, \qquad K(\boldsymbol{x},\boldsymbol{z}) := \frac{\Delta(\boldsymbol{z})}{\Delta(\boldsymbol{x})}e^{S(\boldsymbol{z};\boldsymbol{x})}. \tag{E.8}$$

Fix $r\in\{1,\dots,n+m\}$, and set

$$\mathsf{L}_r := \partial_{x_r}^2 - x_r + \sum_{u\neq r}\frac{\partial_{x_r}-\partial_{x_u}}{x_r-x_u}.$$

For $r=1,\dots,n+m$, write

$$A_r(\boldsymbol{x}) := \partial_{x_r}\log\Delta(\boldsymbol{x}) = \sum_{u\neq r}\frac{1}{x_r-x_u}.$$

Then

$$\partial_{x_r}K(\boldsymbol{x},\boldsymbol{z}) = -(z_r + A_r(\boldsymbol{x}))K(\boldsymbol{x},\boldsymbol{z}).$$

Therefore,

$$\partial_{x_r}^2 K = \left((z_r+A_r)^2 - \partial_{x_r}A_r\right)K,$$

and hence

$$\begin{aligned} \mathsf{L}_r K &= \left(z_r^2 - x_r + 2z_r A_r + A_r^2 - \partial_{x_r}A_r + \sum_{u\neq r}\frac{z_u-z_r}{x_r-x_u} + \sum_{u\neq r}\frac{A_u-A_r}{x_r-x_u}\right)K \\ &= \left(z_r^2 - x_r + \sum_{u\neq r}\frac{z_r+z_u}{x_r-x_u} + C_r(\boldsymbol{x})\right)K, \end{aligned} \tag{E.9}$$

where

$$C_r(\boldsymbol{x}) := A_r^2 - \partial_{x_r}A_r + \sum_{u\neq r}\frac{A_u-A_r}{x_r-x_u}.$$

We now simplify $C_r(\boldsymbol{x})$. Write

$$d_u := x_r - x_u, \qquad u\neq r.$$

Then

$$A_r = \sum_{u\neq r}\frac{1}{d_u}, \qquad -\partial_{x_r}A_r = \sum_{u\neq r}\frac{1}{d_u^2},$$

and therefore

$$A_r^2 - \partial_{x_r} A_r = 2\sum_{u\neq r} \frac{1}{d_u^2} + 2\sum_{\substack{u<v\\u,v\neq r}} \frac{1}{d_u d_v}. \tag{E.10}$$

On the other hand, for $u \neq r$,

$$A_u = -\frac{1}{d_u} + \sum_{\substack{v\neq r\\v\neq u}} \frac{1}{x_u - x_v}.$$

Thus

$$\frac{A_u - A_r}{d_u} = -\frac{2}{d_u^2} + \sum_{\substack{v\neq r\\v\neq u}} \frac{1}{d_u(x_u - x_v)} - \sum_{\substack{v\neq r\\v\neq u}} \frac{1}{d_u d_v}.$$

Summing over $u \neq r$ and pairing the ordered pairs $(u,v)$ and $(v,u)$, we get

$$\begin{aligned}
\sum_{u\neq r} \frac{A_u - A_r}{d_u} &= -2\sum_{u\neq r} \frac{1}{d_u^2} + \sum_{\substack{u<v\\u,v\neq r}} \left(\frac{1}{d_u(x_u - x_v)} + \frac{1}{d_v(x_v - x_u)}\right) - 2\sum_{\substack{u<v\\u,v\neq r}} \frac{1}{d_u d_v} \\
&= -2\sum_{u\neq r} \frac{1}{d_u^2} + \sum_{\substack{u<v\\u,v\neq r}} \frac{1}{x_u - x_v}\left(\frac{1}{d_u} - \frac{1}{d_v}\right) - 2\sum_{\substack{u<v\\u,v\neq r}} \frac{1}{d_u d_v} \\
&= -2\sum_{u\neq r} \frac{1}{d_u^2} - \sum_{\substack{u<v\\u,v\neq r}} \frac{1}{d_u d_v}.
\end{aligned} \tag{E.11}$$

In the last step we used

$$\frac{1}{x_u - x_v}\left(\frac{1}{d_u} - \frac{1}{d_v}\right) = \frac{1}{d_u d_v}.$$

Combining (E.10) and (E.11), we find

$$C_r(\boldsymbol{x}) = \sum_{\substack{u<v\\u,v\neq r}} \frac{1}{(x_r - x_u)(x_r - x_v)}.$$

Thus (E.9) becomes

$$\mathsf{L}_r K(\boldsymbol{x},\boldsymbol{z}) = B_r(\boldsymbol{x},\boldsymbol{z}) K(\boldsymbol{x},\boldsymbol{z}), \tag{E.12}$$

where

$$B_r(\boldsymbol{x},\boldsymbol{z}) := z_r^2 - x_r + \sum_{u\neq r} \frac{z_r + z_u}{x_r - x_u} + \sum_{\substack{u<v\\u,v\neq r}} \frac{1}{(x_r - x_u)(x_r - x_v)}. \tag{E.13}$$

Therefore, by (E.8),

$$\mathsf{L}_r G_{\boldsymbol{\sigma}}(\boldsymbol{x}) = \frac{c_{n+m}}{\Delta(\boldsymbol{x})} \int_{\Gamma_{\boldsymbol{\sigma}}} B_r(\boldsymbol{x},\boldsymbol{z}) \Delta(\boldsymbol{z}) e^{S(\boldsymbol{z};\boldsymbol{x})} \, \mathrm{d}\boldsymbol{z}.$$

It remains to prove that

$$\int_{\Gamma_{\boldsymbol{\sigma}}} B_r(\boldsymbol{x},\boldsymbol{z}) \Delta(\boldsymbol{z}) e^{S(\boldsymbol{z};\boldsymbol{x})} \, \mathrm{d}\boldsymbol{z} = 0. \tag{E.14}$$

*Step 1a. A family of integration-by-parts identities.* Set

$$\Phi(\boldsymbol{z};\boldsymbol{x}) := \Delta(\boldsymbol{z}) e^{S(\boldsymbol{z};\boldsymbol{x})}, \qquad U_r := \{1,\dots,n+m\} \setminus \{r\}.$$

For each subset $I \subseteq U_r$, define

$$J_I := I \cup \{r\}, \qquad \Delta_{J_I}(\boldsymbol{z}) := \prod_{\substack{a<b \\ a,b \in J_I}} (z_a - z_b), \qquad \Omega_I(\boldsymbol{z};\boldsymbol{x}) := \frac{\Phi(\boldsymbol{z};\boldsymbol{x})}{\Delta_{J_I}(\boldsymbol{z})}.$$

Since $\Delta_{J_I}(\boldsymbol{z})$ divides $\Delta(\boldsymbol{z})$, the function $\Omega_I$ is entire in $\boldsymbol{z}$.

If $J_I = \{j_0 < j_1 < \cdots < j_m\}$, define

$$\mathcal{D}_I := \sum_{\ell=0}^{m} (-1)^\ell \Delta_{J_I \setminus \{j_\ell\}}(\boldsymbol{z})\, \partial_{z_{j_\ell}}. \tag{E.15}$$

For each $\ell$, the coefficient $\Delta_{J_I \setminus \{j_\ell\}}(\boldsymbol{z})$ is independent of $z_{j_\ell}$. Moreover, $\Omega_I$ is a polynomial in $\boldsymbol{z}$ times $e^{S(\boldsymbol{z};\boldsymbol{x})}$, and it decays exponentially at the ends of the admissible Airy contours in the variable being differentiated. Therefore integration by parts in $z_{j_\ell}$ yields

$$\int_{\Gamma_{\boldsymbol{\sigma}}} \mathcal{D}_I \Omega_I \, \mathrm{d}\boldsymbol{z} = 0, \qquad I \subseteq U_r. \tag{E.16}$$

We now introduce

$$c_I(\boldsymbol{x}) := \prod_{u \in I} \frac{1}{x_r - x_u}, \qquad c_\varnothing(\boldsymbol{x}) := 1.$$

The key identity is

$$\sum_{I \subseteq U_r} c_I(\boldsymbol{x})\, \mathcal{D}_I \Omega_I = B_r(\boldsymbol{x}, \boldsymbol{z})\, \Phi(\boldsymbol{z};\boldsymbol{x}). \tag{E.17}$$

Once (E.17) is proved, integrating it over $\Gamma_{\boldsymbol{\sigma}}$ and using (E.16) gives (E.14). Then (E.12) implies $\mathsf{L}_r G_{\boldsymbol{\sigma}} = 0$.

*Step 1b. Proof of the key identity.* For a finite set $J \subseteq \{1, \ldots, n+m\}$ and $a \in J$, define

$$w_a^J(\boldsymbol{z}) := \frac{1}{\prod_{b \in J \setminus \{a\}} (z_a - z_b)}.$$

If $J = \{j_0 < \cdots < j_m\}$ and $a = j_\ell$, then a direct sign check gives

$$(-1)^\ell \frac{\Delta_{J \setminus \{a\}}(\boldsymbol{z})}{\Delta_J(\boldsymbol{z})} = w_a^J(\boldsymbol{z}). \tag{E.18}$$

Since

$$\frac{\partial_{z_a} \Omega_I}{\Omega_I} = z_a^2 - x_a + \sum_{v \notin J_I} \frac{1}{z_a - z_v}, \qquad a \in J_I,$$

it follows from (E.15) and (E.18) that

$$\frac{\mathcal{D}_I \Omega_I}{\Phi} = \sum_{a \in J_I} w_a^{J_I}(\boldsymbol{z}) \left( z_a^2 - x_a + \sum_{v \notin J_I} \frac{1}{z_a - z_v} \right). \tag{E.19}$$

We use two elementary identities.

*(i) Lagrange interpolation identity.* For every finite set $J$,

$$\sum_{a \in J} w_a^J(\boldsymbol{z})\, z_a^m = \begin{cases} 0, & 0 \leqslant m \leqslant |J| - 2, \\ 1, & m = |J| - 1. \end{cases} \tag{E.20}$$

This is the standard Lagrange interpolation formula applied to $t^m$.

*(ii) Partial fraction identity.* If $v \notin J$, then

$$\sum_{a\in J} \frac{w_a^J(\boldsymbol{z})}{z_a - z_v} = -\frac{1}{\prod_{a\in J}(z_v - z_a)} = -w_v^{J\cup\{v\}}(\boldsymbol{z}). \tag{E.21}$$

This follows from the partial fraction decomposition of

$$\frac{1}{\prod_{a\in J}(t - z_a)} = \sum_{a\in J} \frac{w_a^J(\boldsymbol{z})}{t - z_a},$$

evaluated at $t = z_v$.

We now sum (E.19) over all $I \subseteq U_r$ with coefficients $c_I(\boldsymbol{x})$.

First, consider the terms involving $z_a^2$. By (E.20), only the cases $|J_I| = 1, 2, 3$ contribute:

$$\sum_{a\in J_I} w_a^{J_I}(\boldsymbol{z}) z_a^2 = \begin{cases} z_r^2, & I = \varnothing, \\ z_r + z_u, & I = \{u\}, \\ 1, & I = \{u, v\},\ u < v,\ u, v \neq r, \\ 0, & |I| \geqslant 3. \end{cases}$$

Hence the total contribution of the $z_a^2$-terms is

$$z_r^2 + \sum_{u\neq r} \frac{z_r + z_u}{x_r - x_u} + \sum_{\substack{u<v\\ u,v\neq r}} \frac{1}{(x_r - x_u)(x_r - x_v)}. \tag{E.22}$$

It remains to treat the $-x_a$-terms and the terms containing $(z_a - z_v)^{-1}$. These cancel telescopically.

Suppose $I \neq \varnothing$. Since $|J_I| \geqslant 2$, identity (E.20) with $m = 0$ gives

$$\sum_{a\in J_I} w_a^{J_I}(\boldsymbol{z}) = 0.$$

Therefore

$$\begin{aligned} -c_I(\boldsymbol{x}) \sum_{a\in J_I} w_a^{J_I}(\boldsymbol{z}) x_a &= -c_I(\boldsymbol{x}) \sum_{b\in I} w_b^{J_I}(\boldsymbol{z})(x_b - x_r) \\ &= \sum_{b\in I} c_{I\setminus\{b\}}(\boldsymbol{x}) w_b^{J_I}(\boldsymbol{z}), \end{aligned} \tag{E.23}$$

because

$$-c_I(\boldsymbol{x})(x_b - x_r) = \prod_{u\in I\setminus\{b\}} \frac{1}{x_r - x_u} = c_{I\setminus\{b\}}(\boldsymbol{x}).$$

Now fix $b \in I$. Consider the contribution of the term $v = b$ in the last sum of (E.19), but for the smaller subset $I \setminus \{b\}$. Since

$$J_{I\setminus\{b\}} \cup \{b\} = J_I,$$

the partial fraction identity (E.21) gives

$$c_{I\setminus\{b\}}(\boldsymbol{x}) \sum_{a\in J_{I\setminus\{b\}}} \frac{w_a^{J_{I\setminus\{b\}}}(\boldsymbol{z})}{z_a - z_b} = -\, c_{I\setminus\{b\}}(\boldsymbol{x}) w_b^{J_I}(\boldsymbol{z}). \tag{E.24}$$

This cancels exactly with the corresponding summand in (E.23). Thus every $-x_a$-term with $I \neq \varnothing$ is cancelled by a unique term coming from the $(z_a - z_v)^{-1}$-part for a smaller subset.

The only unmatched $x$-term is the one coming from $I = \varnothing$, namely

$$-\sum_{a\in\{r\}} w_a^{\{r\}}(\boldsymbol{z})x_a = -x_r.$$

Therefore, after all cancellations, the total contribution of the $-x_a$-terms together with the $(z_a - z_v)^{-1}$-terms is exactly $-x_r$. Combining this with (E.22), we obtain

$$\sum_{I\subseteq U_r} c_I(\boldsymbol{x})\frac{\mathcal{D}_I\Omega_I}{\Phi} = z_r^2 - x_r + \sum_{u\neq r}\frac{z_r+z_u}{x_r-x_u} + \sum_{\substack{u<v\\ u,v\neq r}}\frac{1}{(x_r-x_u)(x_r-x_v)}.$$

By the definition of $B_r(\boldsymbol{x},\boldsymbol{z})$ in (E.13), this proves (E.17).

Integrating (E.17) over $\Gamma_{\boldsymbol{\sigma}}$ and using (E.16) give

$$\int_{\Gamma_{\boldsymbol{\sigma}}} B_r(\boldsymbol{x},\boldsymbol{z})\Phi(\boldsymbol{z};\boldsymbol{x})\,\mathrm{d}\boldsymbol{z} = 0.$$

Thus $\mathsf{L}_r G_{\boldsymbol{\sigma}} = 0$, which proves (E.6).

*Step 2. Lifting back to the original system.* By Theorem E.1,

$$F_{\boldsymbol{\sigma}} = P\,G_{\boldsymbol{\sigma}}$$

solves the original edge system (5.20). Since

$$\Delta(\boldsymbol{x}) = \Delta(\boldsymbol{t})\Delta(\boldsymbol{s})P(\boldsymbol{t},\boldsymbol{s}),$$

formula (6.37) follows directly.

*Step 3. Linear independence.* In the sector $0 < |\arg x| < \pi - \delta$,

$$\partial_x^{j-1}\phi_\pm(x) \sim (\pm\sqrt{x})^{j-1}\phi_\pm(x).$$

Hence, along any admissible ray in this sector,

$$D_{\boldsymbol{\sigma}}(\boldsymbol{x}) \sim \left(\prod_{r=1}^{n+m}\phi_{\sigma_r}(x_r)\right)\det\big[(\sigma_r\sqrt{x_r})^{j-1}\big]_{1\leqslant r,j\leqslant n+m} = \left(\prod_{r=1}^{n+m}\phi_{\sigma_r}(x_r)\right)\Delta(\sigma_1\sqrt{x_1},\ldots,\sigma_{n+m}\sqrt{x_{n+m}}).$$

Therefore

$$F_{\boldsymbol{\sigma}}(\boldsymbol{t},\boldsymbol{s}) \sim \frac{\Delta(\sigma_1\sqrt{x_1},\ldots,\sigma_{n+m}\sqrt{x_{n+m}})}{\Delta(\boldsymbol{t})\Delta(\boldsymbol{s})}\prod_{r=1}^{n+m}\phi_{\sigma_r}(x_r).$$

Distinct sign vectors $\boldsymbol{\sigma}$ produce distinct leading exponential factors

$$\exp\left(\sum_{r=1}^{n+m}\sigma_r\frac{2}{3}x_r^{3/2}\right).$$

Thus the same large-ray argument as in the bulk proof shows that the functions $\{F_{\boldsymbol{\sigma}}\}$ are linearly independent. $\square$

# F The equilibrium BBGKY hierarchy and the loop equations

In this appendix we explain that the equilibrium Bogoliubov–Born–Green–Kirkwood–Yvon equations for the one-dimensional logarithmic interaction imply the loop equations (1.2), and therefore the particle system is the $\mathrm{Sine}_\beta$ ensemble. This provides one example of a point process characterized by the BBGKY hierarchy in the case of long-range and singular interactions. For the proof, we assume a priori mild estimates on the point process, see (F.4).

The definition and analysis of the BBGKY hierarchy for particle systems are delicate for long-range interactions. We first discuss a general setting and then specialize to the log interaction in dimension 1. Consider a simple point process

$$\mu = \sum_{i\in I} \delta_{x_i},\ x_i \in X = \mathbb{R}^d,$$

and its correlation functions defined as the measures $(\rho_p)_{p\geqslant 1}$ such that

$$\mathbb{E}\Bigg[\sum_{i_1,\dots,i_p\in I}^{*} f(x_{i_1},\dots,x_{i_p})\Bigg] = \int_{X^p} f(q_1,\dots,q_p)\,\mathrm{d}\rho_p(q_1,\dots,q_p),$$

for all suitable test functions $f$, where $\sum^*$ means the summation is over all ordered $p$-tuples of distinct indices. Assume $\mu$ is the local limit of a finite interacting particle system with the same charges and symmetric, singular, long-range interaction potential $\varphi$ (e.g. consider either Coulombic $\varphi = (-\Delta)^{-1}$ or logarithmic $\varphi = -\log$ below), at inverse temperature $\beta$: defining

$$\mathbb{P}_N(y_1,\dots,y_N) = \frac{1}{Z_{N,\beta}} e^{-\beta N^{\alpha}\left(\frac12\sum_{i\neq j}\varphi(y_i-y_j) - N\sum_{k=1}^N V(y_k)\right)}\mathrm{d}y_1\dots\mathrm{d}y_N \tag{F.1}$$

with confining $V$ (e.g. $V(y) = |y|^2$) we assume that the following convergence holds in distribution with respect to the weak topology, as $N\to\infty$ (we only consider convergence around the energy level $E=0$ to simplify the discussion):

$$\sum_{i=1}^N \delta_{N^{1/d}y_i} \to \mu \quad \text{as } N\to\infty.$$

Depending on $\alpha\in\mathbb{R}$, the point process $\mu$ is expected to be either Poisson, crystallized, or a non-trivial indermediate depending on $\beta$. When $\mu$ is non-trivial, one expects the following Yvon–Born–Green or YBG hierarchy (which corresponds to the BBGKY hierarchy at equilibrium) for the correlation functions of $\mu$, see [67, Equation (2.10)]: assuming some decay of correlations (called a clustering condition in [67]), we have for distinct points $q_1,\dots,q_p$,

$$\begin{aligned}\frac{1}{\beta}\nabla_{q_1}\rho_p(q_1,\dots,q_p) + \big(\sum_{j=2}^p \nabla\varphi(q_1-q_j)\big)\rho_p(q_1,\dots,q_p)&\\ + \int_X \nabla\varphi(q_1-y)\,(\rho_{p+1}(q_1,\dots,q_p,y) - \rho_p(q_1,\dots,q_p)\rho_1(y))\,\mathrm{d}y &= 0.\end{aligned} \tag{F.2}$$

For singular $\varphi$, this pointwise identity (F.2) holds only away from the collision diagonals. It does not by itself imply the following weak formulation, as this would require *a priori* information on the behavior of $\rho_p$ near collisions. We therefore adopt the following distributional BBGKY hierarchy as the definition.

**Definition F.1** (Equilibrium BBGKY)**.** *We say that $\mu$ satisfies the equilibrium BBGKY hierarchy for the interaction $\varphi$ if for any $p\geqslant 1$ and any test function $\Phi\in C_c^\infty(X^p)$, we have*

$$\mathbb{E}\sum_{i_1,\dots,i_p\in I}\left[\frac{1}{\beta}\nabla_{x_{i_1}}\Phi(x_{i_1},\dots,x_{i_p}) - \Phi(x_{i_1},\dots,x_{i_p})\int\nabla\varphi(x_{i_1}-u)(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(u)\right] = 0, \tag{F.3}$$

*where the indices $i_1,\dots,i_p$ are not required to be distinct and we denote $\mu_i = \mu - \delta_{x_i}$.*

**Remark F.2.** *For symmetric interactions, constant density of states $\rho_1$, and $\nabla\varphi \in \mathrm{L}^1$ we have $\int \nabla\varphi(q-y)\rho_1(y) = 0$, so $\rho_1$ above and the $\rho_p\rho_1$ term in (F.2) can be omitted. However, in our long-range and singular context $\nabla\varphi \notin \mathrm{L}^1$ and this counter-term is needed.*

**Remark F.3.** *The above definition is natural: integrating by parts with respect to the measure (F.1) gives, for any $F \in C_c^\infty(X^p)$*

$$\mathbb{E}_N\left[\frac{1}{\beta}\nabla_{y_1}F - F\Big(\sum_{k\neq 1}\nabla\varphi(y_1-y_k) + N\nabla V(y_1)\Big)\right] = 0.$$

*For example, picking $V(y) = |y|^2$ and $F(y) = \Phi(N^{1/d}y)$, the contribution from $V$ can be shown to be negligible as $N \to \infty$. Defining the equilibrium density through $N^{-1}\sum\delta_{y_i} \to \rho_{\rm eq}$, up to a negligible error one can replace $\sum_{k\neq 1}\nabla\varphi(y_1-y_k)$ with $\sum_{k\neq 1}\nabla\varphi(y_1-y_k) - N\int\nabla\varphi(y_1-u)\mathrm{d}\rho_{\rm eq}(u)$. Under some mild assumption on the discrepancy of the point process, dominated convergence can then be justified and one obtains (F.3) with constant $\rho_1$.*

We now show that the distributional formulation Definition F.1 implies the pointwise BBGKY equation (F.2) away from the collision diagonals. To see this, consider distinct $q_i$'s and $\Phi(x_1,\dots,x_p) = \prod_{i=1}^p \xi_\varepsilon^i(x_i)$ with $\xi_\varepsilon^i(x) = \varepsilon^{-d}\xi((x-q_i)/\varepsilon)$, $\xi$ nonnegative, smooth, supported on $|x| < 1$ and $\int\xi = 1$. When $\varepsilon < \min_{i\neq j}|q_i - q_j|/2$, the supports of $\xi_\varepsilon^1,\dots,\xi_\varepsilon^p$ are pairwise disjoint. Applying Definition F.1 with this test function, we observe that, since the supports of the $\xi_\varepsilon^a$'s are pairwise disjoint, every term with repeated indices vanishes. Indeed, if $i_a = i_b$ for some $a \neq b$, then the factor

$$\xi_\varepsilon^a(x_{i_a})\xi_\varepsilon^b(x_{i_b}) = \xi_\varepsilon^a(x_{i_a})\xi_\varepsilon^b(x_{i_a})$$

is zero. The same argument applies to the derivative term, since $\operatorname{supp}\nabla\xi_\varepsilon^i \subset \operatorname{supp}\xi_\varepsilon^i$. Therefore, for this particular choice of $\Phi$, the contribution of the unrestricted sum is equal to the contribution of the distinct-index sum. Hence Definition F.1 gives

$$\mathbb{E}\sum_{i_1,\dots,i_p\in I}^{*}\Bigg[\frac{1}{\beta}(\nabla_{x_{i_1}}\xi_\varepsilon^1(x_{i_1}))\prod_{j=2}^p\xi_\varepsilon^j(x_{i_j}) - \prod_{j=1}^p\xi_\varepsilon^j(x_{i_j})\sum_{k=2}^p\nabla\varphi(x_{i_1}-x_{i_k}) \\ - \prod_{j=1}^p\xi_\varepsilon^j(x_{i_j})\Bigg(\sum_{\ell\in I\setminus\{i_1,\dots,i_p\}}\nabla\varphi(x_{i_1}-x_\ell) - \int\nabla\varphi(x_{i_1}-u)\mathrm{d}\rho_1(u)\Bigg)\Bigg] = 0,$$

so

$$\int_{X^p}\frac{1}{\beta}(\nabla_{r_1}\xi_\varepsilon^1(r_1))\prod_{j=2}^p\xi_\varepsilon^j(r_j)\rho_p(r_1,\dots,r_p)\mathrm{d}r_1\dots\mathrm{d}r_p - \int_{X^p}\prod_{j=1}^p\xi_\varepsilon^j(r_j)\sum_{k=2}^p\nabla\varphi(r_1-r_k)\rho_p(r_1,\dots,r_p)\mathrm{d}r_1\dots\mathrm{d}r_p \\ - \int_{X^{p+1}}\prod_{j=1}^p\xi_\varepsilon^j(r_j)\nabla\varphi(r_1-x)\big(\rho_{p+1}(r_1,\dots,r_p,x) - \rho_p(r_1,\dots,r_p)\rho_1(x)\big)\mathrm{d}r_1\dots\mathrm{d}r_p\mathrm{d}x = 0.$$

Integrating by parts in $r_1$, taking $\varepsilon \to 0$, and using that $\prod_{j=1}^p\xi_\varepsilon^j$ converges to a Dirac mass at $(q_1,\dots,q_p)$, we obtain (F.2).

We now consider the case $d = 1$, $\varphi(x) = -\log|x|$, and

$$\nabla\varphi(x-y) = -\frac{1}{x-y}.$$

We also assume that the density of states is constant, with normalization $\rho_1 \equiv \pi^{-1}$. Finally, we will need a very weak control on the eigenvalue counts, which can possibly be weakened (we have not tried to optimize this assumption): there exists $\kappa > 0$ such that, for any $p \geqslant 1$, the following holds. There exists $C_p > 0$ such that, for any interval $I$ with $|I| \geqslant 1$,

$$\mathbb{E}\big[|\mu(I) - |I|/\pi|^p\big]^{1/p} \leqslant C_p |I|^{1-\kappa}. \tag{F.4}$$

Under the above assumption, the following properties are elementary to prove: for any $w \notin \mathbb{R}$,

$$s(w) = \mathrm{P.V.} \sum_i \frac{1}{x_i - w}$$

is a.s. well-defined and belongs to $L^p$ for every $p \geqslant 1$, and it is a Nevanlinna function with parameters $c = 0$, $b = \mathrm{P.V.} \sum_i x_i/(x_i^2+1)$. The main result of this section is the following.

**Proposition F.4.** *Fix $\beta > 0$ rational. Let $\mu$ satisfy the above assumption (F.4) and the equilibrium BBGKY hierarchy (F.3). Then $s$ satisfies the loop equations (1.2), and $\mu$ is the* $\mathrm{Sine}_\beta$ *point process.*

*Proof.* Let $\varepsilon > 0$ be small enough, for example $\varepsilon = \kappa/10$, where $\kappa$ is defined in (F.4). Let $R > 0$, which will eventually tend to infinity, and let $\chi_a : \mathbb{R} \to \mathbb{R}_+$ be smooth and even, with

$$\chi_a = 1 \quad \text{on } [-a, a], \qquad \chi_a = 0 \quad \text{on } [-(a + R^{1-\varepsilon}), a + R^{1-\varepsilon}]^{\complement},$$

and $\|\chi_a'\|_\infty \leqslant R^{-1+\varepsilon}$. We will consider $a = R$ and $a = R + 10R^{1-\varepsilon}$ in the proof.

We write $Y = Y(R)$ for a random variable, changing from line to line, which is uniformly bounded in $L^q$, for every fixed $q \geqslant 1$.

We introduce

$$s_R(w) = \sum_i \frac{\chi_R(x_i)}{x_i - w}, \quad \Phi(x_1, \dots, x_p) = \frac{\chi_R(x_1)}{x_1 - w_1} \cdots \frac{\chi_R(x_p)}{x_p - w_p}. \tag{F.5}$$

Then (F.3) gives

$$\mathbb{E}\left[\frac{1}{\beta} \sum_{i_1,\dots,i_p} \partial_{x_{i_1}} \left(\frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} \prod_{j=2}^p \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j}\right) + \sum_{i_1,\dots,i_p} \prod_{j=1}^p \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j} \int \frac{1}{x_{i_1} - x} (\mathrm{d}\mu_{i_1} - \mathrm{d}\rho_1)(x)\right] = 0. \tag{F.6}$$

We first consider the $\partial_{x_{i_1}}$ term:

$$\begin{aligned}\partial_{x_{i_1}} \left(\frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} \prod_{j=2}^p \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j}\right) &= \left(-\partial_{w_1} \frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} + \frac{\chi_R'(x_{i_1})}{x_{i_1} - w_1}\right) \prod_{j=2}^p \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j} \\ &\quad - \frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} \sum_{k=2}^p \mathbb{1}_{i_k = i_1} \left(\frac{\chi_R(x_{i_k})}{(x_{i_k} - w_k)^2} - \frac{\chi_R'(x_{i_k})}{x_{i_k} - w_k}\right) \prod_{\substack{2 \leqslant j \leqslant p \\ j \neq k}} \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j},\end{aligned}$$

so that

$$\begin{aligned}\sum_{i_1,\dots,i_p} \partial_{x_{i_1}} \left(\frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} \prod_{j=2}^p \frac{\chi_R(x_{i_j})}{x_{i_j} - w_j}\right) &= -\partial_{w_1} s_R(w_1) \prod_{j=2}^p s_R(w_j) - \sum_{k=2}^p \sum_{i_1} \frac{\chi_R(x_{i_1})^2}{(x_{i_1} - w_1)(x_{i_1} - w_k)^2} \prod_{\substack{2 \leqslant j \leqslant p \\ j \neq k}} s_R(w_j) \\ &\quad + \sum_{i_1} \frac{\chi_R'(x_{i_1})}{x_{i_1} - w_1} \prod_{j=2}^p s_R(w_j) + \sum_{k=2}^p \sum_{i_1} \frac{\chi_R(x_{i_1})}{x_{i_1} - w_1} \frac{\chi_R'(x_{i_1})}{x_{i_1} - w_k} \prod_{\substack{2 \leqslant j \leqslant p \\ j \neq k}} s_R(w_j).\end{aligned}$$

From (F.4), all terms above involving $\chi'_R$ are of order $Y/R$. For example,

$$\left|\sum_{i_1}\frac{\chi'_R(x_{i_1})}{x_{i_1}-w_1}\right| \leqslant \|\chi'_R\|_\infty \sum_{|x_{i_1}|\in[R,R+R^{1-\varepsilon}]}\frac{1}{|x_{i_1}-w_1|} \lesssim R^{-1+\varepsilon}\frac{\#\{|x_i|\in[R,R+R^{1-\varepsilon}]\}}{R} \leqslant \frac{Y}{R}.$$

Note that

$$\sum_i \frac{\chi_R(x_i)^2}{(x_i-w_1)(x_i-w_k)^2} = \partial_{w_k}\frac{s_R(w_1)-s_R(w_k)}{w_1-w_k} + \sum_i \frac{\chi_R(x_i)^2-\chi_R(x_i)}{(x_i-w_1)(x_i-w_k)^2}$$
$$= \partial_{w_k}\frac{s_R(w_1)-s_R(w_k)}{w_1-w_k} + \mathrm{O}\left(\frac{\#\{|x_i|\in[R,R+R^{1-\varepsilon}]\}}{R^3}\right) = \partial_{w_k}\frac{s_R(w_1)-s_R(w_k)}{w_1-w_k} + \frac{Y}{R^{2+\varepsilon}}.$$

We obtain

$$\begin{aligned}&\sum_{i_1,\dots,i_p}\partial_{x_{i_1}}\left(\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\prod_{j=2}^p\frac{\chi_R(x_{i_j})}{x_{i_j}-w_j}\right)\\ &= -\partial_{w_1}s_R(w_1)\prod_{j=2}^p s_R(w_j) - \sum_{k=2}^p \partial_{w_k}\frac{s_R(w_1)-s_R(w_k)}{w_1-w_k}\prod_{\substack{2\leqslant j\leqslant p\\ j\neq k}} s_R(w_j) + \frac{Y}{R}.\end{aligned} \tag{F.7}$$

For the remaining term in (F.6), using the symmetry under exchange of $i_1$ and $k$,

$$\sum_{i_1\neq k}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\frac{\chi_R(x_k)}{x_{i_1}-x_k} = \frac{1}{2}\sum_{i_1\neq k}\frac{\chi_R(x_{i_1})\chi_R(x_k)}{x_{i_1}-x_k}\left(\frac{1}{x_{i_1}-w_1}-\frac{1}{x_k-w_1}\right)$$
$$= -\frac{1}{2}\sum_{i_1\neq k}\frac{\chi_R(x_{i_1})\chi_R(x_k)}{(x_{i_1}-w_1)(x_k-w_1)} = -\frac{1}{2}\left(s_R(w_1)^2-\partial_{w_1}s_R(w_1)\right) + \frac{Y}{R^{1+\varepsilon}},$$

so that

$$\begin{aligned}\sum_{i_1,\dots,i_p}\prod_{j=1}^p\frac{\chi_R(x_{i_j})}{x_{i_j}-w_j}\int\frac{1}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x) &= -\frac{1}{2}\left(s_R(w_1)^2-\partial_{w_1}s_R(w_1)\right)\prod_{j=2}^p s_R(w_j)\\ &+ \sum_{i_1,\dots,i_p}\prod_{j=1}^p\frac{\chi_R(x_{i_j})}{x_{i_j}-w_j}\int\frac{1}{x_{i_1}-x}\big((1-\chi_R)\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1\big)(x) + \frac{Y}{R}.\end{aligned} \tag{F.8}$$

For $|x_{i_1}|<R-R^{1-\varepsilon/2}$, the point $x_{i_1}$ is far from the transition region of $\chi_R$. Hence

$$\int\frac{\chi_R(x)}{x_{i_1}-x}\,\mathrm{d}x = \mathrm{P.V.}\int_{-R}^{R}\frac{\mathrm{d}x}{x_{i_1}-x} + \mathrm{O}(R^{-\varepsilon/2}) = \log\frac{R+x_{i_1}}{R-x_{i_1}} + \mathrm{O}(R^{-\varepsilon/2}).$$

The remaining particles, satisfying $R-R^{1-\varepsilon/2}<|x_{i_1}|<R+R^{1-\varepsilon}$, lie in a boundary layer. Their contribution is negligible. Indeed, by (F.4), the number of such particles is $YR^{1-\varepsilon/2}$, while $|\chi_R(x_{i_1})/(x_{i_1}-w_1)|\lesssim 1/R$ in this region, and the integral against $\chi_R(x)\mathrm{d}x/(x_{i_1}-x)$ is at most logarithmic. Thus the total boundary-layer contribution is $Y/R^{\varepsilon/4}$.

Combining the interior and boundary estimates gives

$$\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int\frac{\chi_R(x)}{x_{i_1}-x}\,\mathrm{d}\rho_1(x) = \frac{1}{\pi}\sum_{|x_{i_1}|\leqslant R-R^{1-\varepsilon/2}}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\log\frac{R+x_{i_1}}{R-x_{i_1}} + \frac{Y}{R^{\varepsilon/4}}.$$

We now replace the particle sum by the corresponding density integral, using the weak rigidity estimate (F.4). This gives

$$\frac{1}{\pi}\sum_{|x_{i_1}|\leqslant R-R^{1-\varepsilon/2}}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\log\frac{R+x_{i_1}}{R-x_{i_1}}=\frac{1}{\pi^2}\mathrm{P.V.}\int_{-R}^{R}\frac{1}{x-w_1}\log\frac{R+x}{R-x}\,\mathrm{d}x+\frac{Y}{R^{\varepsilon/4}}.$$

After the change of variables $x=Ru$, the deterministic integral becomes

$$\frac{1}{\pi^2}\int_{-1}^{1}\frac{1}{u-w_1/R}\log\frac{1+u}{1-u}\,\mathrm{d}u.$$

Letting $R\to\infty$, this converges to

$$\frac{1}{\pi^2}\int_{-1}^{1}\frac{1}{u}\log\frac{1+u}{1-u}\,\mathrm{d}u=\frac{1}{2}.$$

Consequently,

$$\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int\frac{\chi_R(x)}{x_{i_1}-x}\,\mathrm{d}\rho_1(x)=\frac{1}{2}+\frac{Y}{R^{\varepsilon/4}}.$$

Substituting the above equation and the long-range cancellation Lemma F.5 below into (F.8) gives

$$\sum_{i_1,\dots,i_p}\prod_{j=1}^{p}\frac{\chi_R(x_{i_j})}{x_{i_j}-w_j}\int\frac{1}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x)=\left(-\frac{1}{2}\left(s_R(w_1)^2-\partial_{w_1}s_R(w_1)\right)-\frac{1}{2}\right)\prod_{j=2}^{p}s_R(w_j)+\frac{Y}{R^{\varepsilon/4}}. \tag{F.9}$$

From the estimates (F.7) and (F.9), (F.6) can be written as

$$\mathbb{E}\left[\left(\frac{2-\beta}{\beta}\,\partial_{w_1}s_R(w_1)+s_R(w_1)^2+1\right)\prod_{j=2}^{p}s_R(w_j)\right]+\frac{2}{\beta}\,\mathbb{E}\left[\sum_{j=2}^{p}\partial_{w_j}\frac{s_R(w_1)-s_R(w_j)}{w_1-w_j}\prod_{\substack{\ell=2\\ \ell\neq j}}^{p}s_R(w_\ell)\right]=\mathrm{O}(R^{-\varepsilon/4}).$$

Taking $R\to\infty$ gives the loop equation (1.2) by dominated convergence. Since the principal value $s$ is a Nevanlinna function, we can apply Theorem 1.5, and $\mu$ is the $\mathrm{Sine}_\beta$ point process. □

**Lemma F.5.** *We have*

$$\sum_i\frac{\chi_R(x_i)}{x_i-w_1}\int\frac{1-\chi_R(x)}{x_i-x}(\mathrm{d}\mu_i-\mathrm{d}\rho_1)(x)=\frac{Y}{R^{\varepsilon/2}}, \tag{F.10}$$

*where $Y=Y(R)$ is uniformly bounded in $L^p$ for every fixed $p\geqslant 1$.*

*Proof.* Let $R'=R+10R^{1-\varepsilon}$. We decompose (F.10) as

$$\begin{aligned}(\text{F.10})=&\sum_i\frac{\chi_R(x_i)}{x_i-w_1}\int\frac{1-\chi_{R'}(x)}{x_i-x}(\mathrm{d}\mu_i-\mathrm{d}\rho_1)(x)\\&+\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int\frac{(\chi_{R'}-\chi_R)(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x).\end{aligned} \tag{F.11}$$

We start with the first term on the right-hand side of (F.11). Note that, for any $f$ defined on $[0,\infty)$ such that $f(0)=0$, $|f(u)|\lesssim u^{-1}$, and $|f'(u)|\lesssim u^{-2}$ as $u\to\infty$, we have

$$\int_0^\infty f(x)(\mathrm{d}\mu-\mathrm{d}\rho_1)(x)=-\int_0^\infty f'(x)(\mu[0,x]-x/\pi)\mathrm{d}x.$$

For $f(u)=\frac{1-\chi_{R'}(u)}{x_{i_1}-u}$, for any $|x_{i_1}|<R+R^{1-\varepsilon}=:b$ and $u>R+5R^{1-\varepsilon}$, we have

$$|f'(u)|\leqslant \frac{1-\chi_{R'}(u)}{|x_{i_1}-u|^2}+\frac{|\chi'_{R'}(u)|}{|x_{i_1}-u|}\leqslant \frac{1-\chi_{R'}(u)}{|b-u|^2}+\frac{|\chi'_{R'}(u)|}{|b-u|}=:g(u).$$

Denoting $\Delta(u)=|\mu[0,u]-u/\pi|$, we have

$$\left|\int_0^\infty \frac{1-\chi_{R'}(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x)\right|\leqslant \int_0^\infty g(u)\Delta(u)\mathrm{d}u=:Y.$$

Moreover, for any positive $\ell(u)$,

$$\begin{aligned}|Y|^p&\leqslant \int_{[0,\infty)^p} g(u_1)\dots g(u_p)\Delta(u_1)\dots\Delta(u_p)\mathrm{d}u_1\dots\mathrm{d}u_p\\ &\lesssim \int_{[0,\infty)^p} g(u_1)\dots g(u_p)\ell(u_1)\dots\ell(u_p)\left(\frac{\Delta(u_1)^p}{\ell(u_1)^p}+\dots+\frac{\Delta(u_p)^p}{\ell(u_p)^p}\right)\mathrm{d}u_1\dots\mathrm{d}u_p.\end{aligned}$$

Choosing $\ell(u)=|b-u|^{1-\frac{\kappa}{2}}$, and using the weak local law estimate (F.4), we obtain

$$\mathbb{E}\left[|Y|^p\right]\lesssim \frac{1}{R^{p(1-\varepsilon)\frac{\kappa}{2}}}.$$

Thus, for some $Y$ uniformly bounded in $L^p$, $p\geqslant 1$, we have

$$\left|\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int_0^\infty \frac{1-\chi_{R'}(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x)\right|\lesssim \frac{Y}{R^{\kappa/4}}\sum_{|x_{i_1}|<2R}\frac{1}{|x_{i_1}-w_1|}.$$

The sum on the right-hand side is at most $Y\log R$, and the same estimate holds over $(-\infty,0]$. We have therefore proved

$$\left|\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int_{-\infty}^\infty \frac{1-\chi_{R'}(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x)\right|\lesssim \frac{Y}{R^{\kappa/8}}.\tag{F.12}$$

Next we bound the second term on the right-hand side of (F.11),

$$\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int\frac{(\chi_{R'}-\chi_R)(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x),$$

by considering the $\mu_{i_1}$ and $\rho_1$ contributions separately, starting with $\mu_{i_1}$. Denoting $h=\chi_{R'}-\chi_R$, we notice that $h(x_j)\neq 0$ only if $|x_j|\in[R,R+20R^{1-\varepsilon}]$. In the following we restrict to the case $x_j\in[R,R+20R^{1-\varepsilon}]$; the case $x_j<0$ can be bounded in the same way.

We discuss the two cases $|x_{i_1}-x_j|<100R^{1-\varepsilon}$ and $|x_{i_1}-x_j|\geqslant 100R^{1-\varepsilon}$ separately.

For the sum over $|x_{i_1}-x_j|<100R^{1-\varepsilon}$, by symmetrization we have

$$\begin{aligned}\left|\sum_{\substack{i_1\neq j\\|x_{i_1}-x_j|<100R^{1-\varepsilon}}}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\frac{h(x_j)}{x_{i_1}-x_j}\right|&\lesssim \sum_{\substack{i_1\neq j\\|x_{i_1}-x_j|<100R^{1-\varepsilon}}}\left|\frac{\chi_R(x_{i_1})h(x_j)}{x_{i_1}-w_1}-\frac{\chi_R(x_j)h(x_{i_1})}{x_j-w_1}\right|\frac{1}{|x_{i_1}-x_j|}\\ &\lesssim \frac{\|h'\|_\infty+\|\chi'_{R'}\|_\infty+R^{-1}}{R}\cdot\left(\#\{x_i\in[R-1000R^{1-\varepsilon},R+1000R^{1-\varepsilon}]\}\right)^2\lesssim\frac{Y}{R^{\varepsilon}},\end{aligned}$$

where in the second line we used that $|x_{i_1}-w_1|,\ |x_j-w_1|\asymp R$.

For the sum over $|x_{i_1}-x_j|\geqslant 100R^{1-\varepsilon}$, if $\chi_R(x_{i_1})\neq 0$ and $h(x_j)\neq 0$, then $|x_{i_1}-x_j|\gtrsim |x_{i_1}-R|$. Thus we also have

$$\left|\sum_{\substack{i_1\neq j\\ |x_{i_1}-x_j|\geqslant 100R^{1-\varepsilon}}}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\frac{h(x_j)}{x_{i_1}-x_j}\right|\lesssim \#\{x_j\in[R,R+20R^{1-\varepsilon}]\}\cdot\sum_{|x_i|<R-10R^{1-\varepsilon}}\frac{1}{|x_i-w_1|\cdot|x_i-R|}$$

$$\lesssim Y\,R^{1-\varepsilon}\left(\frac{1}{R}\sum_{|x_i|<R/2}\frac{1}{|x_i-w_1|}+\frac{1}{R}\sum_{|x_i|\in[R/2,R-10R^{1-\varepsilon}]}\frac{1}{|x_i-R|}\right)\leqslant\frac{Y}{R^{\varepsilon/2}}.$$

Together with similar, easier estimates for the contribution of $\rho_1$, we have obtained

$$\left|\sum_{i_1}\frac{\chi_R(x_{i_1})}{x_{i_1}-w_1}\int\frac{(\chi_{R'}-\chi_R)(x)}{x_{i_1}-x}(\mathrm{d}\mu_{i_1}-\mathrm{d}\rho_1)(x)\right|\lesssim\frac{Y}{R^{\varepsilon/2}}.$$

This estimate, combined with (F.12) and the choice $\varepsilon\leqslant\kappa/4$, concludes the proof. $\square$

Finally, we explain how the loop equations (1.2) allow us to obtain the equilibrium hierarchy (F.2). The reasoning below can be made fully rigorous by testing against local test functions and passing to limits, but we focus only on the algebraic aspects. For a function $F$ analytic off the real axis, define the jump operator

$$\Delta_E F:=\frac{1}{2\pi\mathrm{i}}\left(F(E+\mathrm{i}0)-F(E-\mathrm{i}0)\right).$$

**Lemma F.6.** *Let*

$$s(z)=\mathrm{P.V.}\sum_i\frac{1}{x_i-z},\qquad \mu=\sum_i\delta_{x_i},\qquad H(E)=\mathrm{P.V.}\sum_i\frac{1}{x_i-E}.$$

*Then, in the sense of distributions,*

$$\Delta_E s=\mu(E),\quad \Delta_E s'=\partial_E\mu(E),\quad \Delta_E(s^2)=2\mu(E)H(E)+\partial_E\mu(E). \tag{F.13}$$

*Here $\mu(E)H(E)$ is understood with the pole at $E=x_i$ omitted, namely*

$$\mu(E)H(E)=\sum_i\left(\mathrm{P.V.}\sum_{j\neq i}\frac{1}{x_j-x_i}\right)\delta_{x_i}(E).$$

*Moreover, for $E\neq E'$,*

$$\Delta_E\Delta_{E'}\partial_{w'}\frac{s(w)-s(w')}{w-w'}=0, \tag{F.14}$$

*where $\Delta_E$ acts on the variable $w$, and $\Delta_{E'}$ acts on the variable $w'$.*

*Proof of Lemma F.6.* The identity $\Delta_E s=\mu(E)$ follows directly from the Sokhotski–Plemelj formula. Indeed,

$$s(E\pm\mathrm{i}0)=H(E)\pm\mathrm{i}\pi\mu(E),$$

and therefore

$$\Delta_E s=\frac{1}{2\pi\mathrm{i}}\left(s(E+\mathrm{i}0)-s(E-\mathrm{i}0)\right)=\mu(E).$$

We also have, in the distributional sense,

$$\Delta_E \frac{1}{(x_i - z)^2} = \partial_E \delta_{x_i}(E).$$

Summing over $i$ gives

$$\Delta_E \sum_i \frac{1}{(x_i - z)^2} = \Delta_E s' = \sum_i \partial_E \delta_{x_i}(E) = \partial_E \mu(E). \tag{F.15}$$

We next prove the identity for $\Delta_E(s^2)$. Split $s^2$ into its off-diagonal and diagonal parts:

$$s(z)^2 = \sum_{i \neq j} \frac{1}{(x_i - z)(x_j - z)} + \sum_i \frac{1}{(x_i - z)^2}.$$

Consider first the off-diagonal terms. For $i \neq j$, the function

$$\frac{1}{(x_i - z)(x_j - z)}$$

has simple poles at $x_i$ and $x_j$. Hence its jump is

$$\Delta_E \frac{1}{(x_i - z)(x_j - z)} = \frac{\delta_{x_i}(E)}{x_j - x_i} + \frac{\delta_{x_j}(E)}{x_i - x_j}.$$

Summing over $i \neq j$, we obtain

$$\Delta_E \sum_{i \neq j} \frac{1}{(x_i - z)(x_j - z)} = 2 \sum_i \left( \sum_{j \neq i} \frac{1}{x_j - x_i} \right) \delta_{x_i}(E) = 2\mu(E)H(E).$$

The diagonal contribution is given by (F.15). Combining the off-diagonal and diagonal parts proves

$$\Delta_E(s^2) = 2\mu(E)H(E) + \partial_E \mu(E).$$

Finally, we prove (F.14). We note that

$$\partial_{w'} \frac{s(w) - s(w')}{w - w'} = \sum_i \frac{1}{(x_i - w)(x_i - w')^2}.$$

Taking the jump first in $w'$ and then in $w$ gives

$$\Delta_E \Delta_{E'} \partial_{w'} \frac{s(w) - s(w')}{w - w'} = \sum_i \delta_{x_i}(E)\, \partial_{E'} \delta_{x_i}(E').$$

This distribution is supported on the diagonal $E = E'$. Hence it vanishes on the region $E \neq E'$, which proves (F.14). $\square$

We now start from the loop equation (1.2)

$$\mathbb{E}\left[ \left( \frac{2-\beta}{\beta} \partial_{w_1} s(w_1) + s(w_1)^2 + 1 \right) \prod_{j=2}^{p} s(w_j) \right] + \frac{2}{\beta} \sum_{j=2}^{p} \mathbb{E}\left[ \partial_{w_j} \left( \frac{s(w_1) - s(w_j)}{w_1 - w_j} \right) \prod_{\substack{\ell=2 \\ \ell \neq j}}^{p} s(w_\ell) \right] = 0.$$

We now apply

$$\Delta_{E_1} \cdots \Delta_{E_p}$$

with $w_a = E_a \pm \mathrm{i}0$, where the $E_a$'s are distinct. Lemma F.6 gives, for the $\partial_{w_1} s$ term,

$$\Delta_{E_1} \cdots \Delta_{E_p} \mathbb{E}\left[\partial_{w_1} s(w_1) \prod_{j=2}^{p} s(w_j)\right] = \partial_{E_1} \rho_p(E_1, \dots, E_p);$$

for the quadratic term,

$$\Delta_{E_1} \cdots \Delta_{E_p} \mathbb{E}\left[s(w_1)^2 \prod_{j=2}^{p} s(w_j)\right] = 2\,\mathrm{P.V.}\int_{\mathbb{R}} \frac{\rho_{p+1}(E_1, x, E_2, \dots, E_p)}{x - E_1} \mathrm{d}x + 2\sum_{j=2}^{p} \frac{\rho_p(E_1, \dots, E_p)}{E_j - E_1} + \partial_{E_1} \rho_p(E_1, \dots, E_p).$$

The analytic term $+1$ has no jump in $E_1$ and hence disappears. When the $E_a$'s are distinct, the jumps of the last term also vanish:

$$\Delta_{E_1} \cdots \Delta_{E_p} \mathbb{E}\left[\partial_{w_j}\left(\frac{s(w_1) - s(w_j)}{w_1 - w_j}\right) \prod_{\substack{\ell=2 \\ \ell \neq j}}^{p} s(w_\ell)\right] = 0. \tag{F.16}$$

Note that this term would be important in the equilibrium BBGKY hierarchy that includes contact terms, but it does not contribute to deriving (F.2).

Therefore, away from coincident external points, we have proved

$$\frac{1}{\beta}\partial_{E_1}\rho_p + \mathrm{P.V.}\int \frac{\rho_{p+1}(x, E_1, E_2, \dots, E_p)}{x - E_1}\,\mathrm{d}x + \sum_{j=2}^{p} \frac{\rho_p(E_1, \dots, E_p)}{E_j - E_1} = 0.$$

This is the equilibrium BBGKY hierarchy for the one-dimensional logarithmic gas as stated in (F.2); the background term $\rho_p \rho_1$ vanishes after integration in the principal-value sense.

# References


[1] M. Aizenman and S. Warzel. On the ubiquity of the Cauchy distribution in spectral problems. *Probab. Theory Related Fields*, 163(1-2):61–87, 2015.

[2] G. Akemann, J. Baik, and P. Di Francesco. *The Oxford handbook of random matrix theory.* Oxford University Press, 2011.

[3] G. Akemann and Y. Fyodorov. Universal random matrix correlations of ratios of characteristic polynomials at the spectral edges. *Nuclear Physics B*, 664(3):457–476, 2003.

[4] F. Atai, M. Hallnäs, and E. Langmann. Orthogonality of super-Jack polynomials and a Hilbert space interpretation of deformed Calogero–Moser–Sutherland operators. *Bulletin of the London Mathematical Society*, 51(2):353–370, 2019.

[5] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[6] J. Baik, P. Deift, and E. Strahov. Products and ratios of characteristic polynomials of random Hermitian matrices. *Journal of Mathematical Physics*, 44(8):3657–3670, 2003.

[7] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau. Bulk eigenvalue statistics for random regular graphs. *The Annals of Probability*, 45(6A):3626 – 3663, 2017.

[8] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau. Edge rigidity and universality of random regular graphs of intermediate degree. *Geometric and Functional Analysis*, 30(3):693–769, 2020.

[9] R. Bauerschmidt, A. Knowles, and H.-T. Yau. Local semicircle law for random regular graphs. *Communications on Pure and Applied Mathematics*, 70(10):1898–1960, 2017.

[10] A. F. Beardon. *The geometry of discrete groups*. Springer Science & Business Media, 2012.

[11] F. Bekerman, A. Figalli, and A. Guionnet. Transport maps for $\beta$-matrix models and universality. *Comm. Math. Phys.*, 338(2):589–619, 2015.

[12] O. Bohigas, M.-J. Giannoni, and C. Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Physical Review Letters*, 52(1):1, 1984.

[13] A. Borodin and E. Strahov. Averages of characteristic polynomials in random matrix theory. *Communications on Pure and Applied Mathematics*, 59(2):161–253, 2006.

[14] G. Borot and A. Guionnet. Asymptotic expansion of $\beta$ matrix models in the one-cut regime. *Comm. Math. Phys.*, 317(2):447–483, 2013.

[15] G. Borot, A. Guionnet, and K. K. Kozlowski. Large-N asymptotic expansion for mean field models with Coulomb gas interaction. *International Mathematics Research Notices*, 2015(20):10451–10524, 2015.

[16] P. Bourgade, L. Erdős, and H.-T. Yau. Edge universality of beta ensembles. *Comm. Math. Phys.*, 332(1):261–353, 2014.

[17] P. Bourgade, L. Erdős, and H.-T. Yau. Universality of general $\beta$-ensembles. *Duke Math. J.*, 163(6):1127–1190, 2014.

[18] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. *Communications on Pure and Applied Mathematics*, 69(10):1815–1881, 2016.

[19] P. Bourgade, K. Mody, and M. Pain. Optimal local law and central limit theorem for $\beta$-ensembles. *Comm. Math. Phys.*, 390(3):1017–1079, 2022.

[20] P. Bourgade, M. Sodin, and O. Zeitouni. The Paley-Wiener excess and a local law for the $\mathrm{Sine}_\beta$ process. *in preparation*.

[21] P. Bourgade and H.-T. Yau. The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.*, 350(1):231–278, 2017.

[22] O. Chalykh, M. Feigin, and A. Veselov. New integrable generalizations of Calogero–Moser quantum problem. *Journal of Mathematical Physics*, 39(2):695–703, 1998.

[23] G. Cipolloni, L. Erdős, and D. Schröder. Eigenstate thermalization hypothesis for Wigner matrices. *Comm. Math. Phys.*, 388(2):1005–1048, 2021.

[24] J. Conrey, D. Farmer, and M. Zirnbauer. Howe pairs, supersymmetry, and ratios of random characteristic polynomials for the unitary groups U (N). *arXiv preprint math-ph/0511024*, 2005.

[25] I. Corwin. The Kardar–Parisi–Zhang equation and universality class. *Random matrices: Theory and applications*, 1(01):1130001, 2012.

[26] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach: A Riemann-Hilbert Approach*, volume 3. American Mathematical Soc., 2000.

[27] P. Deift and D. Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18. American Mathematical Soc., 2009.

[28] D. Dereudre, A. Hardy, T. Leblé, and M. Maïda. DLR equations and rigidity for the sine-beta process. *Comm. Pure Appl. Math.*, 74(1):172–222, 2021.

[29] P. Desrosiers. Duality in random matrix ensembles for all $\beta$. *Nuclear Physics B*, 817(3):224–251, 2009.

[30] P. Desrosiers and D.-Z. Liu. Asymptotics for products of characteristic polynomials in classical $\beta$-ensembles. *Constructive Approximation*, 39(2):273–322, 2014.

[31] P. Desrosiers and D.-Z. Liu. Selberg integrals, super-hypergeometric functions and applications to $\beta$-ensembles of random matrices. *Random Matrices: Theory and Applications*, 4(02):1550007, 2015.

[32] *NIST Digital Library of Mathematical Functions*. https://dlmf.nist.gov/, Release 1.2.1 of 2024-06-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[33] M. Erbar, M. Huesmann, and T. Leblé. The one-dimensional log-gas free energy has a unique minimizer. *Communications on Pure and Applied Mathematics*, 74(3):615–675, 2021.

[34] L. Erdős, A. Knowles, H.-T. Yau, J. Yin, et al. The local semicircle law for a general class of random matrices. *Electron. J. Probab*, 18(59):1–58, 2013.

[35] L. Erdős, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau. Bulk universality for Wigner matrices. *Communications on Pure and Applied Mathematics*, 63(7):895–925, 2010.

[36] L. Erdős, B. Schlein, and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009.

[37] L. Erdős, B. Schlein, and H.-T. Yau. Universality of random matrices and local relaxation flow. *Invent. Math.*, 185(1):75–119, 2011.

[38] L. Erdős, B. Schlein, H.-T. Yau, and J. Yin. The local relaxation flow approach to universality of the local statistics for random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(1):1–46, 2012.

[39] L. Erdős and H.-T. Yau. Gap universality of generalized Wigner and $\beta$-ensembles. *J. Eur. Math. Soc. (JEMS)*, 17(8):1927–2036, 2015.

[40] L. Erdős and H.-T. Yau. *A dynamical approach to random matrix theory*, volume 28. American Mathematical Soc., 2017.

[41] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probability Theory and Related Fields*, 154(1):341–407, 2012.

[42] L. Erdős, H.-T. Yau, and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.*, 229(3):1435–1515, 2012.

[43] B. Eynard and N. Orantin. Invariants of algebraic curves and topological expansion. *Commun. Number Theory Phys.*, 1(2):347–452, 2007.

[44] A. Figalli and A. Guionnet. Universality in several-matrix models via approximate transport maps. *Acta Math.*, 217(1):81–176, 2016.

[45] P. J. Forrester. Dualities in random matrix theory. *arXiv preprint arXiv:2501.07144*, 2025.

[46] G. Gallavotti and E. Verboven. On the classical KMS boundary condition. *Il Nuovo Cimento*, 28(12):274–286, 1975.

[47] G. Genovese and S. Simonella. On the stationary BBGKY hierarchy for equilibrium states. *J. Stat. Phys.*, 148(1):89–112, 2012.

[48] A. Guionnet. *Asymptotics of random matrices and related models: the uses of Dyson-Schwinger equations*, volume 130. American Mathematical Soc., 2019.

[49] Z. Ha. Exact dynamical correlation functions of Calogero-Sutherland model and one-dimensional fractional statistics. *Physical Review Letters*, 73(12):1574, 1994.

[50] M. Hallnäs. New orthogonality relations for super-Jack polynomials and an associated Lassalle–Nekrasov Correspondence. *Constructive Approximation*, 59(1):113–142, 2024.

[51] Y. He, A. Knowles, and R. Rosenthal. Isotropic self-consistent equations for mean-field random matrices. *Probability Theory and Related Fields*, 171(1):203–249, 2018.

[52] J. Huang, T. McKenzie, and H.-T. Yau. Ramanujan property and edge universality of random regular graphs. *arXiv preprint arXiv:2412.20263*, 2024.

[53] J. Huang and H.-T. Yau. Spectrum of random d-regular graphs up to the edge. *Communications on Pure and Applied Mathematics*, 77(3):1635–1723, 2024.

[54] J. Huang and H.-T. Yau. Lecture notes on edge universality for random regular graphs. *arXiv preprint arXiv:2602.00975*, 2026.

[55] J. Huang and L. Zhang. A convergence framework for $\mathrm{Airy}_\beta$ line ensemble via pole evolution. *Adv. Math.*, 498, 2026.

[56] D. Jakobson, S. D. Miller, I. Rivin, and Z. Rudnick. Eigenvalue spacings for regular graphs. In *Emerging Applications of Number Theory*, pages 317–327. Springer, 1999.

[57] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Mathematical Journal*, 91(1):151 – 204, 1998.

[58] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[59] K. Johansson. Universality of the Local Spacing Distribution in Certain Ensembles of Hermitian Wigner Matrices. *Comm. Math. Phys.*, 215(3):683–705, 2001.

[60] S. Kerov, A. Okounkov, and G. Olshanski. The boundary of the Young graph with Jack edge multiplicities. *International Mathematics Research Notices*, 1998(4):173–199, 1998.

[61] M. Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Communications in mathematical physics*, 147(1):1–23, 1992.

[62] S. G. Krantz. *Function theory of several complex variables*, volume 340. American Mathematical Soc., 2001.

[63] B. Landon, P. Sosoe, and H.-T. Yau. Fixed energy universality of Dyson Brownian motion. *Advances in Mathematics*, 346:1137–1332, 2019.

[64] T. Leblé and S. Serfaty. Large deviation principle for empirical fields of log and Riesz gases. *Inventiones mathematicae*, 210(3):645–757, 2017.

[65] F. Lesage, V. Pasquier, and D. Serban. Dynamical correlation functions in the Calogero-Sutherland model. *Nuclear Physics B*, 435(3):585–603, 1995.

[66] Y. M. Makeenko and A. A. Migdal. Exact equation for the loop average in multicolor QCD. *Physics Letters B*, 88(1-2):135–137, 1979.

[67] P. A. Martin. Sum rules in charged fluids. *Rev. Modern Phys.*, 60(4):1075–1127, 1988.

[68] M. L. Mehta. *Random matrices*, volume 142. Elsevier, 2004.

[69] A. A. Migdal. Loop equations and $1/N$ expansion. *Physics Reports*, 102(4):199–290, 1983.

[70] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Proc. Symp. Pure Math*, volume 24, 1973.

[71] C. B. Morrey, Jr. On the derivation of the equations of hydrodynamics from statistical mechanics. *Comm. Pure Appl. Math.*, 8:279–326, 1955.

[72] G. Nemes. Error bounds for the large-argument asymptotic expansions of the Hankel and Bessel functions. *Acta Applicandae Mathematicae*, 150:141–177, 2017.

[73] A. M. Odlyzko. On the distribution of spacings between zeros of the zeta function. *Mathematics of Computation*, 48(177):273–308, 1987.

[74] A. Okounkov. On n-point correlations in the log-gas at rational temperature. *arXiv preprint hep-th/9702001*, 1997.

[75] M. Olshanetsky and A. Perelomov. Quantum integrable systems related to Lie algebras. *Physics Reports*, 94(6):313–404, 1983.

[76] M. Prähofer and H. Spohn. Scale invariance of the PNG droplet and the Airy process. *Journal of statistical physics*, 108(5):1071–1106, 2002.

[77] J. Quastel and H. Spohn. The one-dimensional KPZ equation and its universality class. *Journal of Statistical Physics*, 160(4):965–984, 2015.

[78] J. Ramirez, B. Rider, and B. Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *Journal of the American Mathematical Society*, 24(4):919–944, 2011.

[79] M. Saito, B. Sturmfels, and N. Takayama. *Gröbner deformations of hypergeometric differential equations.* Springer Science & Business Media, 2013.

[80] D. Serban, F. Lesage, and V. Pasquier. Single-particle Green function in the Calogero-Sutherland model for rational couplings $\beta$= p/q. *Nuclear Physics B*, 466(3):499–512, 1996.

[81] A. Sergeev. Superanalogs of the Calogero operators and Jack polynomials. *Journal of Nonlinear Mathematical Physics*, 8(1):59–64, 2001.

[82] A. Sergeev and A. Veselov. Deformed quantum Calogero-Moser problems and Lie superalgebras. *Comm. Math. Phys.*, 245(2):249–278, 2004.

[83] A. N. Sergeev and A. P. Veselov. Generalised discriminants, deformed Calogero–Moser–Sutherland operators and super-Jack polynomials. *Advances in Mathematics*, 192(2):341–375, 2005.

[84] A. N. Sergeev and A. P. Veselov. Dunkl operators at infinity and Calogero–Moser systems. *International Mathematics Research Notices*, 2015(21):10959–10986, 2015.

[85] M. Shcherbina. Fluctuations of linear eigenvalue statistics of $\beta$ matrix models in the multi-cut regime. *Journal of Statistical Physics*, 151(6):1004–1034, 2013.

[86] M. Shcherbina. Change of variables as a method to study general $\beta$-models: bulk universality. *J. Math. Phys.*, 55(4):043504, 23, 2014.

[87] G. 't Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, pages 461–473, 1974.

[88] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010.

[89] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, 206(1):127–204, 2011.

[90] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Inventiones mathematicae*, 177(3):463–508, 2009.

[91] H.-T. Yau. A journey of random matrices through Wigner universality, Anderson delocalization, and beyond. *ICM proceedings*, 2026.