

LOCAL LIMIT THEOREMS IN FREE PROBABILITY THEORY

JIUN-CHAU WANG

ABSTRACT. In this paper we study the superconvergence phenomenon in the free central limit theorem for identically distributed, unbounded summands. We prove not only the uniform convergence of the densities to the semicircular density but also their L^p -convergence to the same limit for $p > 1/2$. Moreover, an entropic central limit theorem is obtained as a consequence of the above results.

1. INTRODUCTION

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of freely independent copies of a selfadjoint random variable X . For notational convenience, we assume that the random variable X has mean zero and variance one. Let μ_n be the distribution of the normalized sum $S_n = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$. The aim of this article is to study the convergence properties of the probability measures μ_n in terms of their densities $d\mu_n/dx$.

Convergence of the densities of a sequence of distributions to the density of the limit distribution (assuming the given densities exist) is the object of classical local limit theorems [11]. One such limit theorem concerning the sequence $\{\mu_n\}_{n=1}^{\infty}$ was discovered by Bercovici and Voiculescu in [6], where the random variable X is assumed to be *bounded* (i.e., the law of X has a bounded support.) The limit law in this case is the standard semicircular distribution γ whose density is $\sqrt{4-x^2}/2\pi$ on the interval $[-2, 2]$. This convergence to the semicircle law, referred as the *superconvergence* in [6], is very strong in the sense that the measures μ_n become Lebesgue absolutely continuous after finitely many n 's, the density of μ_n is supported on an interval $[a_n, b_n]$ and analytic on (a_n, b_n) , $\lim_{n \rightarrow \infty} a_n = -2$, $\lim_{n \rightarrow \infty} b_n = 2$, and the sequence $d\mu_n/dx$ converges uniformly to $d\gamma/dx$ as $n \rightarrow \infty$.

On the other hand, the weak convergence of the sequence $\{\mu_n\}_{n=1}^{\infty}$ to the measure γ , also known as the *free central limit theorem*, has been the subject of several investigations. This limit theorem was initially proved by Voiculescu [21] for bounded random variables. Later this result was extended to unbounded variables with finite variance by Maassen [15] (see also [17] for a further development.) An explanation

Date: May 12, 2009; revised on August 10, 2009.

2000 Mathematics Subject Classification. Primary: 46L54; Secondary: 60F05, 60F25.

Key words and phrases. Free central limit theorem; Superconvergence; Local limit theorems; Free entropy.

of the appearance of the semicircle law, both in the free central limit theorem and in the asymptotics of large random matrices, was found by Voiculescu in [23].

Knowing the above results, it is therefore natural to ask whether the superconvergence in the free central limit theorem holds for unbounded random variables. In this paper we show that this is indeed the case and even more, in addition to the uniform convergence and analyticity properties of the densities, we also prove a global central limit theorem in the sense of the L^p -convergence. This further implies an entropic central limit theorem, which is a free analogue of Barron's classical theorem on the convergence to the normal entropy [2]. Note that our results are in sharp contrast with the classical central limit theorem, where the distributions in the central limit process could be all atomic.

This paper is written in terms of free convolution. Let \mathcal{M} be the set of all Borel probability measures on the real line \mathbb{R} . The free convolution of two measures $\mu, \nu \in \mathcal{M}$ is denoted by $\mu \boxplus \nu$. Thus, $\mu \boxplus \nu$ is the probability distribution of $Y + Z$, where Y and Z are free random variables with distributions μ and ν , respectively (see [5] for the details of this construction and the book [28] for a comprehensive introduction to free probability theory.)

Aside from being an interesting subject in itself, the theory of free convolution also plays a significant role in the study of the limit laws for large random matrices. For example, it was shown in [16] that if for each $N \geq 1$, A_N and B_N are independent unitarily invariant $N \times N$ hermitian random matrices such that their spectral distributions converge weakly in probability to μ and ν as $N \rightarrow \infty$ and the sequence $E(\text{tr} |A_N|)$ is bounded (where E denotes the expectation and tr the normalized trace), then the spectral distribution of $A_N + B_N$ converges weakly in probability to the free convolution $\mu \boxplus \nu$ as $N \rightarrow \infty$. Hence, one can use the free probability tools such as the R -transform (see Section 2) to further analyze the limit law of $A_N + B_N$. We refer to [8] for an introduction to some of the most probabilistic aspects of free probability and its connections with random matrices. We now begin by reviewing the analytical machinery needed for the calculation of free convolution.

2. PRELIMINARIES

2.1. Cauchy transforms and free convolution. Denote by $\mathbb{C}^+ = \{z \in \mathbb{C} : \Im z > 0\}$ the complex upper half plane. The *Cauchy transform* of a measure $\mu \in \mathcal{M}$ is

$$G_\mu(z) = \int_{-\infty}^{\infty} \frac{1}{z - x} d\mu(x), \quad z \in \mathbb{C}^+,$$

and its reciprocal $F_\mu = 1/G_\mu$ is an analytic self-mapping of \mathbb{C}^+ . The measure μ can be recovered from G_μ as the weak*-limit of the measures

$$d\mu_y(x) = -\frac{1}{\pi} \Im G_\mu(x + iy) dx$$

as $y \rightarrow 0^+$ (see [10]). If the function $\Im G_\mu(z)$ is continuous at $x \in \mathbb{R}$, then the probability distribution function $\mathcal{F}_\mu(t) = \mu((-\infty, t])$ is differentiable at x and its derivative is given by $\mathcal{F}'_\mu(x) = -\Im G_\mu(x)/\pi$. This inversion formula gives an easy way to extract the density function of a measure from its Cauchy transform.

The Cauchy transform G_μ carries certain invertibility properties near the point of infinity, which can be used to calculate the free convolution of measures. To be precise, for each $\theta \in (0, \pi/2)$ and $\beta > 0$, we introduce the cone $\Gamma_\theta = \{z \in \mathbb{C}^+ : \arg z \in (\theta, \pi - \theta)\}$ and the domain

$$D_{\theta, \beta} = \{z \in \mathbb{C}^- : \arg z \in (-\pi + \theta, -\theta); |z| < \beta\},$$

where $\arg z$ is the principal argument of the complex number z and \mathbb{C}^- denotes the complex lower half plane. As shown in [5], we have

$$G_\mu(z) = \frac{1}{z}(1 + o(1)), \quad z \in \mathbb{C}^+,$$

as $z \rightarrow \infty$ *nontangentially* (i.e., $|z| \rightarrow \infty$ but $z \in \Gamma_\theta$ for some angle $\theta > 0$.) This implies that, for every $\theta > 0$, there exists $\beta = \beta(\mu, \theta) > 0$ such that the function G_μ has an analytic inverse G_μ^{-1} (relative to the composition) defined in the set $D_{\theta, \beta}$. Then the *R-transform* of the measure μ is defined as

$$R_\mu(z) = G_\mu^{-1}(z) - \frac{1}{z},$$

and we have

$$R_{\mu_1 \boxplus \mu_2} = R_{\mu_1} + R_{\mu_2},$$

on a domain $D_{\theta, \beta}$ where these three functions are defined. This remarkable property of *R-transform* was first proved by Voiculescu in the case of compactly supported measures [22], then it was extended to the measures with finite variance in [15], and finally to the whole class \mathcal{M} [5].

Note that the *R-transform* R_μ is related with the function φ_μ in [5] via the formula $\varphi_\mu(z) = R_\mu(1/z)$. Thus the properties of the function φ_μ can be translated in terms of R_μ . Among all, we mention the following continuity property with respect to the weak convergence of probability measures (see [5]). The notation δ_x here means the probability measure concentrated at $x \in \mathbb{R}$.

Proposition 2.1. *Let $\{\nu_n\}_{n=1}^\infty \subset \mathcal{M}$ be a sequence. If ν_n converges weakly to δ_0 , then for every $\theta \in (0, \pi/2)$ and $\beta > 0$ there exists $N = N(\theta, \beta) > 0$ such that the*

function R_{ν_n} is defined in the domain $D_{\theta,\beta}$ for $n \geq N$ and $R_{\nu_n} \rightarrow 0$ uniformly on the compact subsets of $D_{\theta,\beta}$ as $n \rightarrow \infty$.

2.2. Free convolution with a semicircle law. For $t > 0$, the centered semicircle distribution of variance t is the probability measure with density

$$d\gamma_t(x) = \frac{1}{2\pi t} \sqrt{4t - x^2} dx$$

on the interval $[-2\sqrt{t}, 2\sqrt{t}]$. Its Cauchy transform is given by

$$G_{\gamma_t}(z) = \frac{z - \sqrt{z^2 - 4t}}{2t}, \quad z \in \mathbb{C}^+,$$

where the branch of the square root on $\mathbb{C}^+ \setminus [0, +\infty)$ is chosen such that $\sqrt{-1} = i$. We use γ to denote the standard semicircle law that has mean zero and variance one. The function G_γ has a continuous extension to $\mathbb{C}^+ \cup \mathbb{R}$, where the extension acts on \mathbb{R} by:

$$\begin{cases} (x - i\sqrt{4 - x^2})/2 & \text{if } |x| \leq 2; \\ (x - \sqrt{x^2 - 4})/2 & \text{if } |x| > 2. \end{cases}$$

In particular, we see that, for each $\delta > 0$, the function G_γ can be continued analytically to $K = \{x + iy : x \in (-2, 2), |y| < \delta\}$ (and beyond, to a whole Riemann surface). This analytic continuation (which we still denote by G_γ) has an explicit formula given by $G_\gamma(z) = (z - i(4 - z^2)^{1/2})/2$, where $(\cdot)^{1/2}$ is the principal branch of the square root function on $\mathbb{C}^+ \setminus (-\infty, 0]$. Note that the function G_{γ_t} also has a similar boundary behavior since $G_{\gamma_t}(z) = t^{-1/2}G_\gamma(t^{-1/2}z)$. We also mention that the R -transform $R_{\gamma_t}(z) = tz$ and the function G_γ satisfies the functional equation:

$$(2.1) \quad G_\gamma(z) + F_\gamma(z) = z, \quad z \in \mathbb{C}^+ \cup K.$$

Proprieties of free convolution by semicircular distributions have been studied thoroughly in [24, 7]. We now review some of these results that are relevant to our approach to the central limit theorem. Fix $t > 0$ and a measure $\nu \in \mathcal{M}$. As shown by Biane [7], the Cauchy transform $G_{\nu \boxplus \gamma_t}$ has a continuous extension to $\mathbb{C}^+ \cup \mathbb{R}$ and an explicit formula for the density of the measure $\nu \boxplus \gamma_t$ can be described as follows.

Define the function $v_t : \mathbb{R} \rightarrow [0, +\infty)$ by

$$v_t(u) = \inf \left\{ v \geq 0 : \int_{-\infty}^{\infty} \frac{1}{(u-x)^2 + v^2} d\nu(x) \leq \frac{1}{t} \right\}.$$

It was proved in [7, Lemma 2] that the function v_t is continuous on \mathbb{R} , analytic on the open set $\{u \in \mathbb{R} : v_t(u) > 0\}$, and for all $u \in \mathbb{R}$

$$(2.2) \quad \int_{-\infty}^{\infty} \frac{1}{(u-x)^2 + v_t(u)^2} d\nu(x) \leq \frac{1}{t}$$

with equality if $v_t(u) > 0$. Let

$$\psi_t(u) = u + t \int_{-\infty}^{\infty} \frac{(u-x)}{(u-x)^2 + v_t(u)^2} d\nu(x), \quad u \in \mathbb{R}.$$

The following formula is proved by Biane [7].

Theorem 2.2. *The function $\psi_t : \mathbb{R} \rightarrow \mathbb{R}$ is a homeomorphism and at the point $x = \psi_t(u)$ we have*

$$G_{\nu \boxplus \gamma_t}(x) = \frac{1}{t} [x - u - iv_t(u)].$$

Thus the density of the measure $\nu \boxplus \gamma_t$ is simply $v_t(u)/\pi t$ at the point $\psi_t(u)$. Biane also showed that the function $G_{\nu \boxplus \gamma_t}(z)$ has an analytic extension to wherever v_t is positive (see Corollary 4 in [7]).

Finally we note for a further reference the following estimate in [7]:

$$(2.3) \quad |G_{\nu \boxplus \gamma_t}(z)| \leq \frac{1}{\sqrt{t}}, \quad z \in \mathbb{C}^+ \cup \mathbb{R}.$$

2.3. Analytic subordination for free convolution powers. For a measure $\mu \in \mathcal{M}$ and a positive integer $n \geq 2$, the n -th free convolution power of μ is defined as

$$\mu^{\boxplus n} = \underbrace{\mu \boxplus \mu \boxplus \cdots \boxplus \mu}_{n \text{ times}}.$$

As shown in [3], the reciprocal Cauchy transform $F_{\mu^{\boxplus n}}$ is subordinated to the function F_μ , that is,

$$(2.4) \quad F_{\mu^{\boxplus n}}(z) = F_\mu(\omega_n(z)), \quad z \in \mathbb{C}^+,$$

where the subordination function ω_n is given by the formula:

$$(2.5) \quad \omega_n(z) = \frac{z}{n} + \frac{n-1}{n} F_{\mu^{\boxplus n}}(z).$$

It was also proved in [3] that the function $F_{\mu^{\boxplus n}}$ has a continuous extension to $\mathbb{C}^+ \cup \mathbb{R}$ and the singular part in the Lebesgue decomposition of the measure $\mu^{\boxplus n}$ is purely atomic. Moreover, the measure $\mu^{\boxplus n}$ has at most one atom and a number α_n is an atom of $\mu^{\boxplus n}$ if and only if $x = \alpha_n/n$ is an atom of μ such that $\mu^{\boxplus n}(\{\alpha_n\}) = n\mu(\{x\}) - n + 1$. Thus, if the measure μ is not a point mass, then the atom of $\mu^{\boxplus n}$ disappears when n is sufficiently large and hence $\mu^{\boxplus n}$ is Lebesgue absolutely continuous.

3. CONVERGENCE OF DENSITIES

Let μ be any probability measure on \mathbb{R} with finite nonzero variance. For simplicity's sake we confine attention to the case of zero mean and unit variance, that

is,

$$\int_{-\infty}^{\infty} x d\mu(x) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} x^2 d\mu(x) = 1.$$

Indeed, the general case can be reduced to this special case by a simple translating and scaling argument. For such a measure μ and a positive integer $n \geq 2$, we introduce

$$\mu_n = \underbrace{D_{\frac{1}{\sqrt{n}}}\mu \boxplus D_{\frac{1}{\sqrt{n}}}\mu \boxplus \cdots \boxplus D_{\frac{1}{\sqrt{n}}}\mu}_{n \text{ times}},$$

where the measure $D_{1/\sqrt{n}}\mu$ is the *dilation* of μ by a factor of $1/\sqrt{n}$, that is,

$$dD_{\frac{1}{\sqrt{n}}}\mu(x) = d\mu(\sqrt{n}x).$$

Note that the measure μ_n can also be realized as the distribution of

$$\frac{1}{\sqrt{n}}(X_1 + X_2 + \cdots + X_n),$$

where X_1, X_2, \dots, X_n are freely independent random variables affiliated to a W^* -probability space, all having distribution μ (see [5]). In other words, we have $\mu_n = D_{1/\sqrt{n}}\mu^{\boxplus n}$. At the level of Cauchy transforms, this means that $G_{\mu_n}(z) = \sqrt{n}G_{\mu^{\boxplus n}}(\sqrt{n}z)$. Meanwhile, by Segal's noncommutative integration theory [19] and the work of Maassen [15], each measure μ_n also has mean zero and variance one. The free central limit theorem [15] states that the sequence μ_n converges weakly to γ as $n \rightarrow \infty$. In this section we will show that the density functions $d\mu_n/dx$ actually converge to $d\gamma/dx$ uniformly on \mathbb{R} as well as in L^p -norms for $p > 1/2$.

To our purposes, we first reinterpret the central limit theorem in terms of free Brownian motion. The following result is essentially the same as Theorem 1.6 in [4]. We provide its proof here for the sake of completeness. Throughout this paper we will use the following notation:

$$t = t(n) = \frac{n}{n+1} \quad \text{and} \quad F_n(z) = F_{\mu_n}(z).$$

Lemma 3.1. *There exists a unique probability measure $\nu \in \mathcal{M}$ such that*

$$(3.1) \quad F_\mu(z) = z - G_\nu(z), \quad z \in \mathbb{C}^+,$$

and, for every $n \geq 1$,

$$(3.2) \quad F_{n+1}(z) = z - G_{\nu_n \boxplus \gamma_t}(z), \quad z \in \mathbb{C}^+ \cup \mathbb{R},$$

where the measure ν_n is given by $d\nu_n(x) = d\nu(\sqrt{n+1}x)$.

Proof. The proof of the representation (3.1) can be found in [15, Proposition 2.2].

We will focus on the proof of (3.2). Notice that we only need to show that (3.2) holds in an open subset of \mathbb{C}^+ since both functions F_{n+1} and $G_{\nu_n \boxplus \gamma_t}$ extend continuously to $\mathbb{C}^+ \cup \mathbb{R}$, and they are analytic in \mathbb{C}^+ . Let $\omega_{n+1}(z)$ be the subordination function of $F_{\mu^{\boxplus(n+1)}}(z)$ as in (2.4). Then (2.5) and (3.1) imply that

$$\begin{aligned}\omega_{n+1}(z) &= \frac{1}{n+1}z + \frac{n}{n+1}F_{\mu}(\omega_{n+1}(z)) \\ &= \frac{1}{n+1}z + \frac{n}{n+1}\omega_{n+1}(z) - \frac{n}{n+1}G_{\nu}(\omega_{n+1}(z)).\end{aligned}$$

Consequently, we have

$$(3.3) \quad \omega_{n+1}(z) + nG_{\nu}(\omega_{n+1}(z)) = z, \quad z \in \mathbb{C}^+.$$

On the other hand, the additivity of R -transform shows that there exists a large $\beta > 0$ such that the equation

$$G_{\nu \boxplus \gamma_n}(w + nG_{\nu}(w)) = G_{\nu}(w)$$

holds in a truncated cone $\Gamma = \{x + iy \in \mathbb{C}^+ : |x| < y; y > \beta\}$. Combining this with (3.3), we obtain the identity

$$(3.4) \quad G_{\nu \boxplus \gamma_n}(z) = G_{\nu}(\omega_{n+1}(z))$$

in a neighborhood of ∞ ; notice that we have used implicitly a consequence of (2.5), namely, the function $\omega_{n+1}(z) = z(1 + o(1))$ as $z \rightarrow \infty$ nontangentially so that $\omega_{n+1}(z)$ lies in Γ as $z \rightarrow \infty$, $z \in \Gamma$.

Meanwhile, (2.4) and (3.1) show that

$$F_{\mu^{\boxplus(n+1)}}(z) = F_{\mu}(\omega_{n+1}(z)) = \omega_{n+1}(z) - G_{\nu}(\omega_{n+1}(z)), \quad z \in \mathbb{C}^+.$$

Together with (3.3) and (3.4), we deduce that the identity

$$F_{\mu^{\boxplus(n+1)}}(z) = z - (n+1)G_{\nu \boxplus \gamma_n}(z)$$

holds in a neighborhood of ∞ . After dilating by a factor of $1/\sqrt{n+1}$, we conclude that (3.2) holds in an open subset of \mathbb{C}^+ and hence the proof is complete. \square

Following Biane [7], we now parametrize the real line by the homeomorphism ψ_t as in Section 2.2. Setting $x = \psi_t(u)$, we first derive an estimate for the range of x .

Lemma 3.2. *For every $\eta \in (0, 1)$ and $n \geq 1$, the set $\psi_t([-1 + \eta, 1 - \eta])$ is contained in the interval $[-2 + \eta, 2 - \eta]$.*

Proof. By Theorem 2.2, we have $\psi_t(u) = u + t\Re G_{\nu_n \boxplus \gamma_t}(x)$. Moreover, (2.3) implies that

$$|t\Re G_{\nu_n \boxplus \gamma_t}(x)| \leq \sqrt{t} = \sqrt{\frac{n}{n+1}} < 1 - \frac{1}{2(n+1)}.$$

Hence we obtain

$$|\psi_t(u)| < 2 - \eta - \frac{1}{2(n+1)}, \quad u \in [-1 + \eta, 1 - \eta],$$

which implies the desired result. \square

We need one more detail about the boundary behavior of the function $G_{\nu_n \boxplus \gamma_t}(z)$ as $n \rightarrow \infty$. Observe that the sequence ν_n converges weakly to δ_0 .

Lemma 3.3. *For each $\eta \in (0, 1)$ there exists $N = N(\eta) > 0$ such that, for all $n \geq N$, the function $G_{\nu_n \boxplus \gamma_t}(z)$ can be continued analytically to a neighborhood of the interval $[-2 + \eta, 2 - \eta]$. Furthermore, this analytic continuation never vanishes on $[-2 + \eta, 2 - \eta]$.*

Proof. By virtue of Theorem 2.2, it suffices to show that $v_t(u) > 0$ on the interval $[-1 + \eta, 1 - \eta]$ for sufficiently large n . Suppose that, on the contrary, we can find a subsequence $\{n_k\}_{k=1}^\infty \subset \mathbb{N}$ and a sequence $\{u_k\}_{k=1}^\infty$ in $[-1 + \eta, 1 - \eta]$ such that $n_1 < n_2 < \dots$ and $v_{t_k}(u_k) = 0$ for all k 's, where t_k denotes $n_k/(n_k + 1)$. Then the weak convergence of $\{\nu_n\}_{n=1}^\infty$ implies that

$$\nu_{n_k} \left(\left(-\frac{\eta}{2}, \frac{\eta}{2} \right) \right) > 1 - \frac{\eta}{2}$$

for sufficiently large k . Therefore we deduce, from (2.2), that

$$\begin{aligned} \frac{1}{t_k} &> \int_{-\infty}^{\infty} \frac{1}{(u_k - x)^2} d\nu_{n_k}(x) \\ &\geq \int_{|x| < \frac{\eta}{2}} \frac{1}{(u_k - x)^2} d\nu_{n_k}(x) > \frac{2}{2 - \eta} \end{aligned}$$

for large k . This contradicts to the fact that $\lim_{k \rightarrow \infty} t_k = 1$. Hence the function v_t must be positive on $[-1 + \eta, 1 - \eta]$ when n is large, and this completes the proof. \square

We now have all ingredients to proceed with the proof of our main result.

Theorem 3.4. *Suppose that $\mu \in \mathcal{M}$ has mean zero and variance one. Then*

- (1) *the measure μ_n is Lebesgue absolutely continuous for sufficiently large n ;*
- (2) *for each small $\varepsilon > 0$ there exist $\delta > 0$ and $N > 0$ such that the function G_{μ_n} has an analytic continuation h_n to $K = \{x + iy : x \in [-2 + \varepsilon, 2 - \varepsilon], |y| < \delta\}$ whenever $n \geq N$. Moreover, $h_n(z) \rightarrow (z - i(4 - z^2)^{1/2})/2$ uniformly on K as $n \rightarrow \infty$;*
- (3) *the density $d\mu_n/dx$ is continuous for sufficiently large n and $d\mu_n/dx \rightarrow d\gamma/dx$ uniformly on \mathbb{R} as $n \rightarrow \infty$.*

Proof. The statement (1) follows from the regularity results of the measure $\mu^{\boxplus n}$ because μ_n is a dilation of $\mu^{\boxplus n}$.

Let us prove (2). In accordance with the established terminology, we will work with measures indexed by $n + 1$ instead of n . Let $\varepsilon \in (0, 1/10)$ be given. Define the sets $K_\varepsilon = \{x + iy : x \in [-2 + \varepsilon, 2 - \varepsilon], |4y| < \varepsilon\sqrt{\varepsilon}\}$ and $D_{\varepsilon,2} = \{z \in \mathbb{C}^- : \arg z \in (-\pi + \theta, -\theta); |z| < 2\}$, where the angle $\theta = \theta(\varepsilon)$ is chosen so that $4 \sin \theta = \sqrt{\varepsilon(4 - \varepsilon)}$.

We first show that the set $G_\gamma(K_\varepsilon)$ is contained in $D_{\varepsilon,2}$, where $G_\gamma(z) = (z - i(4 - z^2)^{1/2})/2$ is the analytic extension of the Cauchy transform of γ on K_ε . Let $z_0 \in K_\varepsilon$ be given. We write $G_\gamma(z_0) = Re^{i\psi}$, where $\psi = \arg G_\gamma(z_0)$. To prove that $G_\gamma(z_0)$ is in $D_{\varepsilon,2}$, we need to verify that $|\sin \psi| > \sin \theta$ and $R < 2$. Indeed, from the functional equation (2.1), we have

$$\left(R + \frac{1}{R}\right) \cos \psi + i \left(R - \frac{1}{R}\right) \sin \psi = z_0.$$

Using the fact that $|\Re z_0| \leq 2 - \varepsilon$, we get

$$2 |\cos \psi| \leq \left(R + \frac{1}{R}\right) |\cos \psi| \leq 2 - \varepsilon.$$

This implies that $2 |\sin \psi| \geq \sqrt{\varepsilon(4 - \varepsilon)} > 2 \sin \theta$, as desired. On the other hand, we have

$$\frac{1}{4} \varepsilon \sqrt{\varepsilon} > |\Im z_0| = |\sin \psi| \left| R - \frac{1}{R} \right| > \frac{1}{2} \sqrt{\varepsilon} \frac{|R^2 - 1|}{R}.$$

If $R > 1$, then we deduce that $2R^2 - \varepsilon R - 2 < 0$. Therefore, the number R must be bounded from above by the positive x -intercept of the parabola $Y = 2X^2 - \varepsilon X - 2$. This means that R is no more than 1.026 because $\varepsilon < 0.1$.

Recall, from (3.2), that

$$F_{n+1}(z) = z - G_{\nu_n \boxplus \gamma_t}(z), \quad z \in \mathbb{C}^+ \cup \mathbb{R},$$

where the sequence ν_n converges weakly to δ_0 as $n \rightarrow \infty$. We will derive an approximant for the function $G_{\nu_n \boxplus \gamma_t}(z)$ as in (3.7) below. By Proposition 2.1 and Lemma 3.3, there exists $N = N(\varepsilon) > 0$ such that, for $n \geq N$, the following conditions are satisfied:

- (a) $32(n + 1) < \varepsilon\sqrt{\varepsilon}$;
- (b) the R -transform R_{ν_n} is defined in $D_{\varepsilon,2}$ and the sequence R_{ν_n} converges uniformly on $G_\gamma(K_\varepsilon)$ to zero as $n \rightarrow \infty$;

(c)

$$|R_{\nu_n}(w)| < \frac{\varepsilon\sqrt{\varepsilon}}{16}, \quad w \in G_\gamma(K_\varepsilon);$$

- (d) the function $G_{\nu_n \boxplus \gamma_t}(z)$ has an analytical extension (which we still denote by $G_{\nu_n \boxplus \gamma_t}$) to a neighborhood \mathcal{V} of the interval $[-2 + 2\varepsilon, 2 - 2\varepsilon]$ and $G_{\nu_n \boxplus \gamma_t} \neq 0$ in \mathcal{V} .

The additivity of R -transform shows that

$$\begin{aligned} G_{\nu_n \boxplus \gamma_t}^{-1}(w) &= G_{\nu_n}^{-1}(w) + G_{\gamma_t}^{-1}(w) - \frac{1}{w} \\ &= w + \frac{1}{w} - \frac{1}{n+1}w + R_{\nu_n}(w) \end{aligned}$$

wherever the function R_{ν_n} is defined. Hence (b) shows that the function $G_{\nu_n \boxplus \gamma_t}^{-1}$ is defined in $D_{\varepsilon,2}$ for $n \geq N$. Moreover, after replacing w by $G_\gamma(z)$ and making use of (2.1), we conclude that

$$(3.5) \quad f_n(z) = G_{\nu_n \boxplus \gamma_t}^{-1}(G_\gamma(z)) = z + r_n(z), \quad z \in K_\varepsilon,$$

where the sequence

$$r_n(z) = R_{\nu_n}(G_\gamma(z)) - \frac{1}{n+1}G_\gamma(z)$$

converges uniformly on K_ε to zero as $n \rightarrow \infty$ and the estimation

$$|r_n(z)| < \frac{\varepsilon\sqrt{\varepsilon}}{8}$$

holds uniformly for $z \in K_\varepsilon$ and $n \geq N$.

The uniform bound of r_n and (3.5) imply that the rectangle $\tilde{K}_\varepsilon = \{x + iy : x \in [-2 + 2\varepsilon, 2 - 2\varepsilon], |y| < \varepsilon\sqrt{\varepsilon}\}$ is contained in the set $f_n(K_\varepsilon)$. Then Rouché's theorem shows that each function f_n has an analytic inverse f_n^{-1} defined in \tilde{K}_ε . Moreover, by inverting (3.5), we have

$$f_n^{-1}(z) = z + O(|r_n(z)|)$$

for every $z \in \tilde{K}_\varepsilon$ and $n \geq N$. Meanwhile, (d) shows that the composition $G_\gamma^{-1} \circ G_{\nu_n \boxplus \gamma_t}$ is defined and analytic in \mathcal{V} , and hence it coincides with the function f_n^{-1} on the interval $[-2 + 2\varepsilon, 2 - 2\varepsilon]$. Therefore we conclude that there exist constants $Q > 0$ and $\delta = \delta(\varepsilon) \in (0, \varepsilon\sqrt{\varepsilon}/8)$, both independent of n , such that

$$(3.6) \quad G_\gamma^{-1}(G_{\nu_n \boxplus \gamma_t}(z)) = f_n^{-1}(z) = z + g_n(z)$$

for every z in the rectangle $K = \{x + iy : x \in [-2 + 2\varepsilon, 2 - 2\varepsilon], |y| < \delta\}$, where $|g_n| \leq Q|r_n|$ on K . Note that the Cauchy estimate shows that $|G_\gamma'(z)| \leq \varepsilon^{-1}$ on K . Therefore, by applying G_γ on (3.6), we get

$$(3.7) \quad G_{\nu_n \boxplus \gamma_t}(z) = G_\gamma(z + g_n(z)) = G_\gamma(z) + l_n(z), \quad z \in K, n \geq N,$$

where $|l_n(z)| \leq Q \sup_{z \in K} |G_\gamma'(z)| |r_n(z)| \leq Q\varepsilon^{-1} |r_n(z)|$ on K .

By (3.2), (3.7), and (2.1), the function G_{n+1} can be continued meromorphically to the set K . Moreover, if h_{n+1} denotes this continuation, then we have

$$h_{n+1}(z) = \frac{G_\gamma(z)}{1 - l_n(z)G_\gamma(z)}, \quad z \in K, n \geq N.$$

Therefore, after letting n be large enough so that $|l_n(z)G_\gamma(z)| < 1$, we conclude that the function h_{n+1} is the analytic continuation of G_{n+1} to K , and the uniform convergence property in the statement (2) follows because $l_n \rightarrow 0$ uniformly on K as $n \rightarrow \infty$. The proof of (2) is complete.

We next prove (3). The proof of (2) shows that, for large n , we have

$$|G_{n+1}(x) - G_\gamma(x)| < \varepsilon, \quad x \in [-2 + 2\varepsilon, 2 - 2\varepsilon].$$

Hence, by the inversion formula, the density function $d\mu_{n+1}/dx$ is uniformly close to $d\gamma/dx$ on $[-2 + 2\varepsilon, 2 - 2\varepsilon]$ for sufficiently large n . Combining with (3.2) and (2.3), we then have

$$\Im G_{n+1}(x) = \frac{\Im G_{\nu_n \boxplus \gamma_t}(x)}{[x - \Re G_{\nu_n \boxplus \gamma_t}(x)]^2 + [\Im G_{\nu_n \boxplus \gamma_t}(x)]^2}, \quad x \in \mathbb{R},$$

for large n . Hence we see that the function $d\mu_{n+1}/dx$ is continuous for such n .

In order to finish the proof, we need to estimate $d\mu_{n+1}/dx$ on the set $\mathbb{R} \setminus [-2 + 2\varepsilon, 2 - 2\varepsilon]$. Let us take $x = \psi_t(u)$ and $v_t(u) = -t\Im G_{\nu_n \boxplus \gamma_t}(x)$ as in Theorem 2.2. Then Lemma 3.2 and (2.3) imply that $|u| > 1 - 2\varepsilon$ and

$$|x - \Re G_{\nu_n \boxplus \gamma_t}(x)| \geq |x| - \frac{7}{6}, \quad |x| > 2 - 2\varepsilon, \quad n \geq 3.$$

Hence we obtain the bound:

$$(3.8) \quad |\Im G_{n+1}(x)| \leq \frac{72v_t(u)}{(|x| + 1)^2}, \quad |u| > 1 - 2\varepsilon, \quad |x| > 2 - 2\varepsilon,$$

for every $n \geq 3$.

By the inversion formula, it remains to show that v_t is uniformly small on $\{u : |u| > 1 - 2\varepsilon\}$ as $n \rightarrow \infty$. This can be done as follows. For each $n \geq 1$, we introduce the probability tail-sum $k_n = 1 - \nu_n((-\varepsilon, \varepsilon))$. From (2.2), we have, for $v_t(u) > 0$ and $|u| > 1 - 2\varepsilon$, that

$$\begin{aligned} \frac{1}{t} &= \int_{-\infty}^{\infty} \frac{1}{(u-x)^2 + v_t(u)^2} d\nu_n(x) \\ &\leq \int_{|x| < \varepsilon} \frac{1}{(1-3\varepsilon)^2 + v_t(u)^2} d\nu_n(x) + \int_{|x| \geq \varepsilon} \frac{1}{v_t(u)^2} d\nu_n(x) \\ &= \frac{1 - k_n}{(1-3\varepsilon)^2 + v_t(u)^2} + \frac{k_n}{v_t(u)^2}. \end{aligned}$$

Therefore we see that the number $v_t(u)^2$ is bounded above by the positive x -intercept of the parabola $Y = X^2 - bX - c$, where $b = t - (1 - 3\varepsilon)^2$ and $c = tk_n(1 - 3\varepsilon)^2 \geq 0$. In other words, we have

$$v_t(u)^2 \leq \frac{b + \sqrt{b^2 + 4c}}{2} \leq |b| + \sqrt{k_n}, \quad |u| \geq 1 - 2\varepsilon, \quad n \geq 1.$$

Hence the desired conclusion for v_t follows from the facts that $\lim_{n \rightarrow \infty} k_n = 0$ and $\lim_{n \rightarrow \infty} b = 6\varepsilon - 9\varepsilon^2$. Since ε is arbitrary, this completes the proof of (3). \square

Remark. Observe first that, in Theorem 3.4, the function $(\overline{h_n(\bar{z})} - h_n(z))/2\pi i$ is a complex analytic extension of $d\mu_n/dx$ to K for $n \geq N$. Hence, for any $\varepsilon > 0$ and $k \geq 1$, Theorem 3.4 actually implies that the k -th derivatives $d^k \mu_n/dx^k$ converge to $d^k \gamma/dx^k$ uniformly on $[-2 + \varepsilon, 2 - \varepsilon]$ as $n \rightarrow \infty$.

Secondly, if the measure μ has a bounded support, then so does the measure μ_n . In this case, the equation (3.2) shows that the support of the measures μ_n can not deviate too much from the interval $[-2, 2]$. This property has been noted in [6], where it was shown that, for any $\varepsilon > 0$ there exists $N > 0$ such that the support of μ_n is contained in $[-2 - \varepsilon, 2 + \varepsilon]$ for $n \geq N$. An earlier result of Voiculescu [22, Lemmas 3.1 and 3.2] provides a precise estimate for the support of a free convolution of compactly supported measures (see also [14]). In particular, Voiculescu's result shows that the support of μ_n is contained in the interval $[-2 - (L/\sqrt{n}), 2 + (L/\sqrt{n})]$ for $n \geq 1$, where $L = \sup\{|x| : x \in \text{supp}(\mu)\}$.

We conclude this section with a global form of the central limit theorem.

Theorem 3.5. *Let μ and μ_n be as in Theorem 3.4. Then we have*

$$\lim_{n \rightarrow \infty} \left\| \frac{d\mu_n}{dx} - \frac{d\gamma}{dx} \right\|_{L^p(\mathbb{R})} = 0$$

for every $p > 1/2$.

Proof. Fix $p > 1/2$. Let $\varepsilon \in (0, 1/10)$ be given. From the inversion formula, we have

$$\begin{aligned} \left\| \frac{d\mu_{n+1}}{dx} - \frac{d\gamma}{dx} \right\|_{L^p(\mathbb{R})}^p &= \frac{1}{\pi^p} \int_{-2+2\varepsilon}^{2-2\varepsilon} |\Im G_{n+1}(x) - \Im G_\gamma(x)|^p dx \\ &\quad + \frac{1}{\pi^p} \int_{|x| > 2-2\varepsilon} |\Im G_{n+1}(x) - \Im G_\gamma(x)|^p dx. \end{aligned}$$

Then Theorem 3.4 (2) shows that within the interval $[-2 + 2\varepsilon, 2 - 2\varepsilon]$ the integrand tends uniformly to zero as $n \rightarrow \infty$ and so the contribution of $[-2 + 2\varepsilon, 2 - 2\varepsilon]$ tends to zero. On the other hand, (3.8) implies, for $n \geq 3$, that

$$\begin{aligned} \int_{|x| > 2-2\varepsilon} |\Im G_{n+1}(x) - \Im G_\gamma(x)|^p dx &\leq \int_{|x| > 2-2\varepsilon} 2^p (|\Im G_{n+1}(x)|^p + |\Im G_\gamma(x)|^p) dx \\ &\leq 144^p \left(\sqrt{|b|} + \sqrt[4]{k_n} \right)^p \int_{|x| > 2-2\varepsilon} \frac{1}{(|x| + 1)^{2p}} dx \\ &\quad + \int_{2 \geq |x| > 2-2\varepsilon} (4 - x^2)^{\frac{p}{2}} dx, \end{aligned}$$

where b and k_n are as in the proof of Theorem 3.4. Then the proof is completed by taking $n \rightarrow \infty$, then $\varepsilon \rightarrow 0^+$. \square

4. ENTROPIC CENTRAL LIMIT THEOREM

In this section we study the convergence in the free central limit theorem in terms of entropy. Our motivation is a result of Barron, which we now review as follows.

The classical entropy of a measure $\rho \in \mathcal{M}$ with density f is defined as $H(\rho) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx$, provided the positive part of the integral is finite. Thus we have $H(\rho) \in [-\infty, \infty)$. It is well-known that the standard Gaussian distribution G has the largest entropy among all probability measures on \mathbb{R} with variance one. Let μ , $D_{1/\sqrt{n}}\mu$ and μ_n be as in Section 3, and further let ρ_n be the classical convolution:

$$\rho_n = \underbrace{D_{\frac{1}{\sqrt{n}}}\mu * D_{\frac{1}{\sqrt{n}}}\mu * \cdots * D_{\frac{1}{\sqrt{n}}}\mu}_{n \text{ times}}.$$

Barron [2] showed that if the entropy $H(\rho_n)$ ever becomes different from $-\infty$, then the sequence $H(\rho_n)$ converges to $H(G) = (1/2) \log 2\pi e$. A detailed exposition of the connections between the entropy and central limit theorems can be found in [13].

Let ν be a probability measure on \mathbb{R} . The quantity

$$\chi(\nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \log |x - y| d\nu(x) d\nu(y) + \frac{3}{4} + \frac{1}{2} \log 2\pi,$$

called *free entropy*, was discovered by Voiculescu in [24] to be a good substitute for entropy in free probability theory. Remarkably, free entropy χ behaves like the classical entropy H in many instances. For example, the free entropy is maximized by the standard semicircular law γ (with the value $\chi(\gamma) = (1/2) \log 2\pi e$) among all probability measures with variance one [25, 12] and both $\chi(\mu_n)$ and $H(\rho_n)$ are monotonically increasing [20, 18, 1]. We also refer to [27] for a survey of free entropy. Our goal here is to prove a free analogue of Barron's result.

The key ingredient in the proof of our result is a free logarithmic Sobolev inequality first proved by Voiculescu [26], which requires a further assumption on the regularity of the measure ν . Suppose that the measure ν has a density p in $L^3(\mathbb{R})$, and let us assume the support of ν is bounded for the moment. Then the *free Fisher information* of the measure ν is

$$\Phi(\nu) = \frac{4\pi^2}{3} \int_{-\infty}^{\infty} p(x)^3 dx,$$

see [26, Proposition 7.9], where the following inequality is proved:

$$(4.1) \quad \chi(\nu) \geq \frac{1}{2} \log \left(\frac{2\pi e}{\Phi(\nu)} \right).$$

Using random matrix approximations, the inequality (4.1) is further shown to hold for measures with finite variance, not necessarily having a bounded support (see [9]). Thus we can use this generalized version of (4.1) to obtain the following:

Theorem 4.1. *Let μ and μ_n be as in Theorem 3.4. Then the free entropy $\chi(\mu_n)$ converges to the semicircular entropy $(1/2)\log 2\pi e$.*

Proof. Notice first that $\Phi(\gamma) = 1$. Theorem 3.5 shows that the free Fisher information $\Phi(\mu_n)$ converges to $\Phi(\gamma) = 1$ as $n \rightarrow \infty$. Since μ_n has variance one, we conclude, from (4.1), that $\chi(\mu_n) > -\infty$ and

$$\chi(\gamma) \geq \chi(\mu_n) \geq \chi(\gamma) - \frac{1}{2} \log(\Phi(\mu_n))$$

for sufficiently large n . The result follows immediately. \square

Finally, we conclude this paper with the following remark. A version of Theorem 4.1 for bounded variables satisfying more restrictive conditions was noted in [13]. Here we only require the variance to be one, which is the most general condition to have the free central limit theorem (see [17]). Actually, Theorem 4.1 can be extended to any probability measure with finite nonzero variance via a standard translating and scaling procedure by observing that the free entropy χ is invariant under the translation and

$$\chi(D_a\nu) = \chi(\nu) + \log a, \quad a > 0.$$

Acknowledgement. The author would like to thank Professors Hari Bercovici and Roland Speicher for their encouragement and valuable help during the course of the investigation. He is also grateful to Professor Michael Anshelevich for a reference, and to a referee for helpful comments on the preliminary version of this work.

REFERENCES

- [1] S. Artstein, K. Ball, F. Barthe, and A. Naor, *Solution of Shannon's problem on monotonicity of entropy*, J. Amer. Math. Soc. **17** (2004) 975-982.
- [2] A. Barron, *Entropy and the central limit theorem*, Ann. Probab. **14** (1986), 336-342.
- [3] S. Belinschi and H. Bercovici, *Atoms and regularity for measures in a partially defined free convolution semigroup*, Math. Z. **248** (2004), no. 4, 665-674.
- [4] S. Belinschi and A. Nica, *On a remarkable semigroup of homomorphisms with respect to free multiplicative convolution*, Indiana Univ. Math. J. **57** (2008), no. 4, 1679-1713.
- [5] H. Bercovici and D. Voiculescu, *Free convolution of measures with unbounded support*, Indiana Univ. Math. J. **42** (1993), no. 3, 733-773.
- [6] ———, *Superconvergence to the central limit and failure of the Cramér theorem for free random variables*, Probab. Theory Relat. Fields **103** (1995), no. 2, 215-222.
- [7] Ph. Biane, *On the free convolution with a semi-circular distribution*, Indiana Univ. Math. J. **46** (1997), no. 3, 705-718.

-
- [8] ———, *Free probability for probabilists*, QP-PQ, Vol. XI, 55-71, World Sci. Publ. 2003.
- [9] ———, *Logarithmic Sobolev inequalities, matrix models and free entropy*, Acta Math. Sin. (Engl. Ser.) **19** (2003), no. 3, 497-506.
- [10] W. F. Donoghue, Jr., *Monotone matrix functions and analytic continuation*, Die Grundlehren der mathematischen Wissenschaften, Band 207, Springer-Verlag, 1974.
- [11] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley, Reading, MA, 1968.
- [12] F. Hiai and D. Petz, *Maximizing free entropy*, Acta Math. Hungar. **80** (1998), no. 4, 335-356.
- [13] O. Johnson, *Information theory and the central limit theorem*, Imperial College Press, London, UK, 2004.
- [14] V. Kargin, *On superconvergence of sums of free random variables*, Ann. Probab. **35** (2007), no. 5, 1931-1949.
- [15] H. Maassen, *Addition of freely independent random variables*, J. Funct. Anal. **106** (1992), no. 2, 409-438.
- [16] L. Pastur and V. Vasilchuk, *On the law of addition of random matrices*, Comm. Math. Phys. **214** (2000), no. 2, 249-286.
- [17] V. Pata, *The central limit theorem for free additive convolution*, J. Funct. Anal. **140** (1996), no. 2, 359-380.
- [18] H. Schultz, *Semicircularity, Gaussianity and monotonicity of entropy*, J. Operator Theory **60** (2008), no. 1, 125-136.
- [19] I. E. Segal, *A non-commutative extension of abstract integration*, Ann. of Math. **57** (1953), 401-457.
- [20] D. Shlyakhtenko, *A free analogue of Shannon's problem on monotonicity of entropy*, Adv. Math. **208** (2007), no. 2, 824-833.
- [21] D. Voiculescu, *Symmetries of some reduced free product C^* -algebras*, Operator Algebras and Their Connections with Topology and Ergodic Theory (Buşteni, 1983). Lecture Notes in Math. **1132**, Springer, Berlin, 1985, 556-588.
- [22] ———, *Addition of certain noncommuting random variables*, J. Funct. Anal. **66** (1986), no. 3, 323-346.
- [23] ———, *Limit laws for random matrices and free products*, Invent. Math. **104** (1991), 201-220.
- [24] ———, *The analogues of entropy and of Fisher's information measure in free probability theory. I*, Comm. Math. Phys. **155** (1993), no. 1, 71-92.
- [25] ———, *The analogues of entropy and of Fisher's information measure in free probability theory. IV. Maximum entropy and freeness*, Free Probability Theory, 293-302, Fields Inst. Commun. **12**, AMS, 1997.
- [26] ———, *The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms*, Invent. Math. **132** (1998), no. 1, 189-227.
- [27] ———, *Free entropy*, Bull. London Math. Soc. **34** (2002), no. 3, 257-278.
- [28] D. Voiculescu, K. Dykema, and A. Nica, *Free random variables*. CRM Monograph series, 1. Amer. Math. Soc., Providence, RI, 1992.

DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN'S UNIVERSITY, KINGSTON, ONTARIO, K7L 3N6, CANADA

E-mail address: jiuwang@mast.queensu.ca