Eigen-inference for multi-source power estimation

Romain Couillet
ST-Ericsson, Supélec
505 Route des Lucioles
06560 Sophia Antipolis, France
Email: romain.couillet@supelec.fr

Jack W. Silverstein
Department of Mathematics
North Carolina State University
Raleigh, North Carolina 27695-8205
Email: jack@math.ncsu.edu

Mérouane Debbah
Alcatel-Lucent Chair, Supélec
Plateau de Moulon, 3 rue Joliot-Curie
91192 Gif sur Yvette, France
Email: merouane.debbah@supelec.fr

***Abstract*—This paper introduces a new method to estimate the power transmitted by multiple signal sources, when the number of sensing devices and the available samples are sufficiently large compared to the number of sources. This work makes use of recent advances in the field of random matrix theory that prove more efficient than previous "moment-based" approaches to the problem of multi-source power detection. Simulations are performed which corroborate the theoretical claims.**

I. Introduction

At a time when radio resources become scarce, the alternative offered by flexible radios [1] is gaining more and more interest. A *flexible* wireless network is a set of opportunistic entities, referred to as the *secondary* network, that benefit from unused spectrum resource to establish communication, with little interference to the established *primary* network. This is performed by letting the secondary devices sense the environment for the presence of active transmissions and exchange the collected information among the secondary network. If the secondary devices can detect the number of primary sources and evaluate the power used by every individual source, their own maximum transmission power (i.e. the maximum transmit power that brings little interference to the primary network) can then be reliably estimated. The detection of the number of neighbors and the estimation of the individual transmit powers is the subject of this work.

The difficulty of estimating transmit powers lies in the little information known *a priori* by the secondary network: the transmitted data and the transmission channels are usually inaccessible. This has motivated much work in the direction of blind detection methods [2], [3]. To solve the harder problem of power inference, it is necessary to assume that the sensed samples are of large dimension compared to the number of active sources.[1] The latter condition allows one to model the channel from the sources to the secondary users, as well as the transmit data and noise, as large random matrices; call them $\mathbf{H}$, $\mathbf{X}$ and $\mathbf{W}$, respectively. Denoting $\mathbf{P}$ a diagonal matrix of the source powers, the detection problem boils down to estimating the entries of $\mathbf{P}$ from the sole knowledge of $\mathbf{Y} = \mathbf{H}\mathbf{P}^{\frac{1}{2}}\mathbf{X}+\mathbf{W}$. Up to this day though, no computationally-cheap *consistent* estimator[2] for the entries of $\mathbf{P}$ has been proposed. Among the existing techniques are convex optimization strategies [4] or moment-based approaches [5], [6]. The latter provide consistent estimators of the moments of the eigenvalue distribution of $\mathbf{P}$ as a function of the moments of the eigenvalue distribution of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$; from those estimates, the entries of $\mathbf{P}$ themselves can be inferred. The moment-based techniques are however expected to perform worse than methods that would fully exploit the eigenvalue distribution of $\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$, and not only the first moments. This problem is addressed in [7] for the *sample covariance matrix model* $\mathbf{Y}' = \mathbf{P}^{\frac{1}{2}}\mathbf{X}$, i.e. the entries of $\mathbf{P}$ are inferred from the full eigenvalue distribution of $\mathbf{X}^{\mathsf{H}}\mathbf{P}\mathbf{X}$.

This work generalizes this result to infer the entries of $\mathbf{P}$ from the observed matrix $\mathbf{Y} = \mathbf{H}\mathbf{P}^{\frac{1}{2}}\mathbf{X} + \mathbf{W}$. The novel estimator proposed here will be shown to have a very compact form, to be computationally inexpensive and to perform better than moment-based approaches. The remainder of this paper is structured as follows: Section II introduces the system model. In Section III, the novel power estimator is derived, part of the technical proofs being left to [8]. Section IV provides simulation results. Section V concludes this work.

Notations: In the following, boldface lower case symbols represent vectors, capital boldface characters denote matrices ($\mathbf{I}_N$ is the size-N identity matrix). The transpose and Hermitian transpose operators are denoted $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$, respectively. We denote by $\mathbb{C}^+$ the set $\{z \in \mathbb{C}, \Im[z] > 0\}$. The symbol '$\xrightarrow{\text{a.s.}}$' denotes almost sure convergence.

II. System Model

Consider a wireless (primary) network in which K entities are transmitting data. Transmitter $k \in \{1, \ldots, K\}$ has transmission power P_k and is equipped with n_k antennas. We denote $n = \sum_{k=1}^{K} n_k$ the total number of transmit antennas in the primary network. Consider also a secondary network composed of a total of N sensing devices, e.g. a single user embedded with N antennas or N single antenna users; we shall refer to the N sensors collectively as *the receiver*. Denote $\mathbf{H}_k \in \mathbb{C}^{N \times n_k}$ the multiple antenna channel matrix between transmitter k and the receiver. We assume that the entries of $\mathbf{H}_k$ are independent and identically distributed (i.i.d.) with zero mean and variance $1/N$. At time instant m, transmitter k emits signal $\mathbf{x}_k^{(m)} \in \mathbb{C}^{n_k}$, with entries assumed to be i.i.d. of zero mean and variance 1. Assume further that at time instant m the receiver is corrupted by additive white noise of variance σ^2 on every sensor; we denote $\sigma\mathbf{w}^{(m)} \in \mathbb{C}^N$ the

[1] e.g. individual secondary users are equipped with many antennas, or a large number of secondary users, each equipped with few antennas, collect their received data in a central entity.

[2] an estimator $\hat{P}_i$ of the i^{th} entry P_i of $\mathbf{P}$ is said to be *consistent* if $\hat{P}_i - P_i \to 0$ almost surely when the relevant system dimensions grow large.

receive noise vector where $\mathbf{w}_k^{(m)}$ has i.i.d. entries with zero mean and variance 1. At time m, the receiver therefore senses the signal $\mathbf{y}^{(m)}$ defined as

$$\mathbf{y}^{(m)} = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{H}_k\mathbf{x}_k^{(m)} + \sigma\mathbf{w}^{(m)} \tag{1}$$

Assuming the channel fading is constant over M consecutive sampling periods, by concatenating M successive signal realizations into $\mathbf{Y} = [\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(M)}] \in \mathbb{C}^{N\times M}$, we have

$$\mathbf{Y} = \sum_{k=1}^{K} \sqrt{P_k}\mathbf{H}_k\mathbf{X}_k + \sigma\mathbf{W} \tag{2}$$

where $\mathbf{X}_k = [\mathbf{x}_k^{(1)}, \ldots, \mathbf{x}_k^{(M)}] \in \mathbb{C}^{n_k\times M}$, for every k and $\mathbf{W} = [\mathbf{w}^{(1)}, \ldots, \mathbf{w}_M^{(M)}] \in \mathbb{C}^{N\times M}$. This can be further rewritten

$$\mathbf{Y} = \mathbf{H}\mathbf{P}^{\frac{1}{2}}\mathbf{X} + \sigma\mathbf{W} \tag{3}$$

where $\mathbf{P} \in \mathbb{R}^{n\times n}$ is diagonal with first n_1 entries P_1, subsequent n_2 entries P_2, …and last n_K entries P_k, $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_K] \in \mathbb{C}^{N\times n}$ and $\mathbf{X} = [\mathbf{X}_1^{\mathsf{T}}, \ldots, \mathbf{X}_k^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{n\times M}$.

Our objective is to provide an (n, N, M)-*consistent* estimate $\hat{P}_1, \ldots, \hat{P}_K$ of $P_1, \ldots, P_K$, from a single realization of the random matrix $\mathbf{Y}$. That is, for all i, $\hat{P}_i - P_i \xrightarrow{\text{a.s.}} 0$, as n, N and M all grow large. This is the subject of the next section.

III. Multi-source power estimation

We start by analyzing the eigenvalue distribution of $\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ when n, N and M grow large at a similar rate. This is a fundamental prior step to the proper estimation of $P_1, \ldots, P_K$.

A. Spectral analysis

Definition 1: Let F be a distribution function. For $z \in \mathbb{C}^+$, the *Stieltjes transform* $m(z)$ of F is defined as

$$m(z) = \int \frac{1}{t-z} dF(t) \tag{4}$$

For all $a < b \in \mathbb{R}$, we have the *inverse Stieltjes transform* formula

$$F([a,b]) = \frac{1}{\pi} \lim_{y\to 0^+} \int_a^b \Im[m(x+iy)] \tag{5}$$

A consequence of Definition 1 is that studying the distribution function F is equivalent to studying its Stieltjes transform $m(z)$.

In this section, we prove the following theorem

Theorem 1: Let $\mathbf{B}_N = \frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ with eigenvalues $\lambda_1, \ldots, \lambda_N$. Denote $\hat{m}(z) \triangleq \frac{1}{N}\sum_{k=1}^{N} \frac{1}{\lambda_k - z}$ the Stieltjes transform of the eigenvalue distribution $F^{\mathbf{B}_N}$ of $\mathbf{B}_N$. Then, for M, N, n growing large with limit ratios $M/N \to c$, $N/n \to c_1$, $0 < c, c_1 < \infty$, for any $z \in \mathbb{C}^+$, we have

$$\hat{m}(z) \xrightarrow{\text{a.s.}} m(z) \tag{6}$$

where $m(z)$ is defined as

$$m(z) = \frac{M}{N}\underline{m}(z) + \frac{M-N}{N}\frac{1}{z} \tag{7}$$

and $\underline{m}(z)$ is the unique solution with positive imaginary part of the implicit equation

$$\frac{1}{\underline{m}(z)} = -\sigma^2 + \frac{1}{f(z)} - \frac{1}{N}\sum_{k=1}^{K} \frac{n_k P_k}{1 + P_k f(z)} \tag{8}$$

in which we denoted $f(z)$ the function

$$f(z) = \frac{M-N}{N}\underline{m}(z) - \frac{M}{N} z\underline{m}(z)^2 \tag{9}$$

This implies that $F^{\mathbf{B}_N}$ tends weakly and almost surely to a limit F, called the *limit spectral density* (l.s.d.) of $\mathbf{B}_N$, with Stieltjes transform $m(z)$ [9].

The rest of this section is dedicated to a sketch of the proof of Theorem 1. First remark that (3) can be further simplified into

$$\mathbf{Y} = \begin{pmatrix}\mathbf{H}\mathbf{P}^{\frac{1}{2}} & \sigma\mathbf{I}_N\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{W}\end{pmatrix} \tag{10}$$

Appending $\mathbf{Y} \in \mathbb{C}^{N\times M}$ into the larger matrix $\underline{\mathbf{Y}} \in \mathbb{C}^{(N+n)\times M}$,

$$\underline{\mathbf{Y}} = \begin{pmatrix}\mathbf{H}\mathbf{P}^{\frac{1}{2}} & \sigma\mathbf{I}_N\\ 0 & 0\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{W}\end{pmatrix} \tag{11}$$

We recognize that $\underline{\mathbf{Y}}$ is a *sample covariance matrix*, with random population covariance matrix $\left(\begin{smallmatrix}\mathbf{H}\mathbf{P}^{\frac{1}{2}} & \sigma\mathbf{I}_N\\ 0 & 0\end{smallmatrix}\right)$ while $\left(\begin{smallmatrix}\mathbf{X}\\ \mathbf{W}\end{smallmatrix}\right)$ has independent entries with zero mean and variance 1. Extending (1.4) of [10] to non-deterministic population covariance matrices, we have that the eigenvalue distribution of $\frac{1}{M}\underline{\mathbf{Y}}^{\mathsf{H}}\underline{\mathbf{Y}}$ converges almost surely to a limit $\underline{F}$ whose Stieltjes transform $\underline{m}(z)$ satisfies, for $z \in \mathbb{C}^+$,

$$z = -\frac{1}{\underline{m}(z)} + \frac{N+n}{M}\int \frac{t}{1+t\underline{m}(z)} dH(t) \tag{12}$$

$$= -\frac{1}{\underline{m}(z)} + \frac{N+n}{M\underline{m}(z)}\left(1 - \frac{1}{\underline{m}(z)}\int \frac{1}{t - (-\frac{1}{\underline{m}(z)})} dH(t)\right) \tag{13}$$

where H is the (almost sure) l.s.d. of $\left(\begin{smallmatrix}\mathbf{H}\mathbf{P}\mathbf{H}^{\mathsf{H}} + \sigma^2\mathbf{I}_N & 0\\ 0 & 0\end{smallmatrix}\right)$. The proof of existence of the previous limit distributions is fully argumented in [8]. Now, $\mathbf{P}^{\frac{1}{2}}\mathbf{H}^{\mathsf{H}}$ is itself a sample covariance matrix for which $\mathbf{H}^{\mathsf{H}}$ has i.i.d. entries with zero mean and variance $1/M$; so we have again, for any $w \in \mathbb{C}^+$, [10]

$$w = -\frac{1}{m_1(w)} + \frac{1}{N}\sum_{k=1}^{K} n_k \frac{P_k}{1 + P_k m_1(w)} \tag{14}$$

where $m_1(w)$ is the Stieltjes transform of the l.s.d. of $\mathbf{H}\mathbf{P}\mathbf{H}^{\mathsf{H}}$ evaluated at w. This is also the Stieltjes transform of the l.s.d. of $\mathbf{H}\mathbf{P}\mathbf{H}^{\mathsf{H}} + \sigma^2\mathbf{I}_N$ evaluated at $z = w + \sigma^2$, which, up to n zeros, is distributed as H. More precisely, we have

$$\int \frac{1}{t - (w+\sigma^2)} dH(t) = \frac{N}{N+n} m_1(w) - \frac{n}{N+n}\frac{1}{w} \tag{15}$$

Notice now that the integral in (13) is the left-hand side of (15) evaluated at $w = -1/\underline{m}(z) - \sigma^2$. For $z \in \mathbb{C}^+$, we then have, from (13) and (15),

$$z = -\frac{N}{M}\frac{1}{\underline{m}(z)^2} m_1(-1/\underline{m}(z) - \sigma^2) + \frac{N-M}{M}\frac{1}{\underline{m}(z)} \tag{16}$$

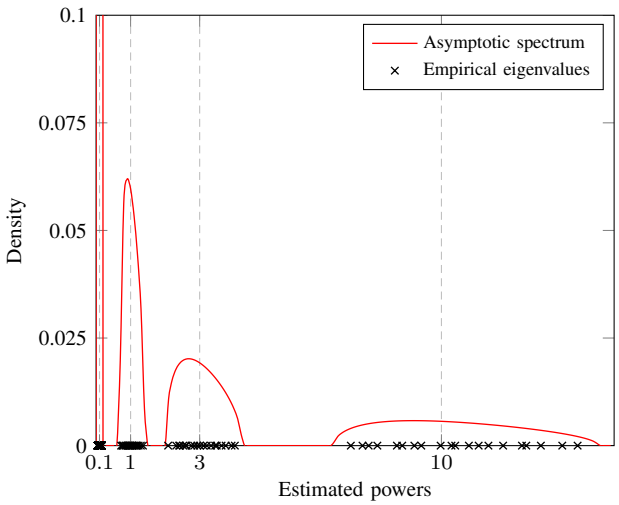

Fig. 1. Empirical and asymptotic eigenvalue distribution of $\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ when $\mathbf{P}$ has three distinct entries $P_1 = 1$, $P_2 = 3$, $P_3 = 10$, $n_1 = n_2 = n_3$, $N/n = 10$, $M/N = 10$, $\sigma^2 = 0.1$. Empirical test: $n = 60$.

where, according to (14), $f(z) \triangleq m_1(-1/\underline{m}(z) - \sigma^2)$ satisfies

$$\frac{1}{\underline{m}(z)} = -\sigma^2 + \frac{1}{f(z)} - \frac{1}{N}\sum_{k=1}^{K} \frac{n_k P_k}{1 + P_k f(z)} \tag{17}$$

Together with (16) (which is equivalent to (9)), this is exactly (8).

Since the eigenvalues of the matrices $\mathbf{Y}$ and $\underline{\mathbf{Y}}$ only differ by n zero eigenvalues, we also have that the Stieltjes transform $m(z)$ of the l.s.d. of $\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ satisfies

$$m(z) = \frac{M}{N}\underline{m}(z) + \frac{M-N}{N}\frac{1}{z} \tag{18}$$

This completes the proof of Theorem 1.

Therefore, the support of the l.s.d. F of $\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ can be evaluated as follows: for any $z \in \mathbb{C}^+$, $m(z)$ is given by (7), in which $\underline{m}(z)$ is the unique solution of (8); the inverse Stieltjes transform formula (5) gives then access to F from $m(z)$, for values of z spanning the set $\{z = x + iy, y > 0\}$, for small y. This is depicted in Figure 1, when $\mathbf{P}$ has three distinct values $P_1 = 1$, $P_2 = 3$, $P_3 = 10$ and $n_1 = n_2 = n_3$, $N/n = 10$, $M/N = 10$, $\sigma^2 = 0.1$.

Two remarks on Figure 1 are of fundamental importance to the following. First, it appears that the asymptotic spectrum is divided in disjoint clusters. This will be in fact true whenever the ratio N/n is sufficiently large; otherwise some of the successive clusters would overlap. An explicit formulation of the condition for the separability of the asymptotic spectrum in clusters is provided in [7] for the simpler matrix model $\mathbf{P}^{\frac{1}{2}}\mathbf{X}$. Secondly, notice that the empirical eigenvalues are here all inside the asymptotic clusters and, most importantly, that the number of those eigenvalues is exactly n_1, n_2 and n_3 for the three clusters corresponding to P_1, P_2 and P_3, respectively. This fact is referred to as *exact separation*; the exact separation for the current model is proven in [8], from an extension of the proof of exact separation for the sample covariance matrix model $\mathbf{P}^{\frac{1}{2}}\mathbf{X}$ [11].

B. Eigen-inference

In this section, we prove our main result,

Theorem 2: Let $\mathbf{B}_N \in \mathbb{C}^{N\times N}$ be defined as in Theorem 1, and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$, $\lambda_1 < \ldots < \lambda_N$, be the vector of the ordered eigenvalues of $\mathbf{B}_N$. Further assume that the limiting ratios c, c_1 and $\mathbf{P}$ are such that the asymptotic cluster separability condition is fulfilled. Then, for $k \in \{1, \ldots, K\}$, as N, n, M grow large, we have

$$\hat{P}_k - P_k \xrightarrow{\text{a.s.}} 0 \tag{19}$$

where the estimate $\hat{P}_k$ is given by

$$\hat{P}_k = \frac{NM}{n_k(M-N)} \sum_{i\in\mathcal{N}_k} (\eta_i - \mu_i) \tag{20}$$

in which $\mathcal{N}_k = \{N - \sum_{i=k}^{K} n_i + 1, \ldots, N - \sum_{i=k+1}^{K} n_i\}$ is the set of indexes matching the cluster corresponding to P_k, $(\eta_1, \ldots, \eta_N)$ are the ordered eigenvalues of the matrix $\mathrm{diag}(\boldsymbol{\lambda}) - \frac{1}{N}\sqrt{\boldsymbol{\lambda}}\sqrt{\boldsymbol{\lambda}}^{\mathsf{T}}$ and $(\mu_1, \ldots, \mu_N)$ are the ordered eigenvalues of the matrix $\mathrm{diag}(\boldsymbol{\lambda}) - \frac{1}{M}\sqrt{\boldsymbol{\lambda}}\sqrt{\boldsymbol{\lambda}}^{\mathsf{T}}$.

The approach pursued to prove Theorem 2 is based on the original idea of [7]. From the Cauchy integration formula [12],

$$P_k = \frac{n}{n_k}\frac{1}{2\pi i}\oint_{\mathcal{C}_k} \frac{1}{n}\sum_{r=1}^{K} n_r \frac{\omega}{P_r - \omega} d\omega \tag{21}$$

for any negatively oriented contour $\mathcal{C}_k \subset \mathbb{C}$, such that P_k is contained in the surface described by the contour, while for every $i \neq k$, P_i is outside this surface. The strategy is then the following: we first propose a convenient integration contour $\mathcal{C}_k$ which is intimately linked to $m(z)$, the Stieltjes transform of the l.s.d. of $\mathbf{B}_N$. Instead of evaluating P_k, we then evaluate the complex integral resulting from replacing the deterministic $m(z)$ by the empirical $\hat{m}(z) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - z}$. From the convergence $\hat{m}(z) \xrightarrow{\text{a.s.}} m(z)$ proven in Theorem 1, this new integral, denoted $\hat{P}_k$, is shown to be a consistent estimate of P_k in the limit. It then suffices to evaluate $\hat{P}_k$, which is performed by residue calculus [12].

From the spectrum separability condition, we can choose x_k^- and x_k^+ two reals outside the spectrum of H to be such that $-1/m_1(x_k^-) < P_k < -1/m_1(x_k^+)$.[3] The proof of the existence of such x_k^- and x_k^+ is somewhat technical and is not included, due to space limitation (see [8] for details). Define $\bar{\Gamma}_k \subset \mathbb{C}$ to be any continuous curve with endpoints x_k^- and x_k^+ and with interior points of strictly positive imaginary part. We then define $\mathcal{C}_k$ to be the union of the curve $-1/m_1(\bar{\Gamma}_k)$ with $\bar{\Gamma}_k$ oriented from x_k^- to x_k^+ and the curve $-1/m_1(\bar{\Gamma}_k^*)$,[4] with $\bar{\Gamma}_k$ oriented from x_k^+ to x_k^-. Since $-1/m_1(z) \in \mathbb{C}^+$ for $z \in \mathbb{C}^+$ (see e.g. [10]), we verify easily that P_k is included in $\mathcal{C}_k$, while P_i, $i \neq k$, is not, as required.

[3] we implicitly extended here the definition domain of m_1 to all reals outside the support of H; see [8] for an accurate proof.

[4] again here, we implicitly extend the definition of $m_1(z)$ to $z \in \mathbb{C}^-$, which does not represent a major difficulty.

Making the variable change $\omega = -1/m_1(w)$ and denoting Γ_k the surface enclosed in the union of $\bar{\Gamma}_k$ from x_k^- to x_k^+ and $\bar{\Gamma}_k^*$ from x_k^+ to x_k^-, (21) becomes

$$P_k = \frac{n}{n_k}\frac{1}{2\pi i}\oint_{\partial\Gamma_k}\left(\frac{N}{n}wm_1(w)+\frac{N-n}{n}\right)\frac{m_1'(w)}{m_1(w)^2}dw \tag{22}$$

From Theorem 1, we can then rewrite (22) as a function of $\underline{m}(z)$. Thanks to the variable change $w = -1/\underline{m}(z) - \sigma^2$,

$$P_k = \frac{n}{n_k}\frac{1}{2\pi i}\oint_{\partial\Omega_k}\left[\frac{N}{n}\left(1+\sigma^2\underline{m}(z)\right)+\frac{N-n}{n}\frac{1}{zm(z)}\right] \times\left[-\frac{1}{z\underline{m}(z)}-\frac{\underline{m}'(z)}{\underline{m}(z)^2}-\frac{m'(z)}{m(z)\underline{m}(z)}\right]dz \tag{23}$$

where Ω_k is the surface described by $-1/\underline{m}(\Gamma_k) - \sigma^2$. Similarly as we show that Γ_k encloses P_k and none of the P_i's, $i \neq k$, we show in [8] that Ω_k contains the asymptotic spectrum cluster corresponding to P_k but none of the clusters corresponding to P_i, $i \neq k$.

Instead of going further with (23), define $\hat{P}_k$, the "empirical counterpart" of P_k, as

$$\hat{P}_k = \frac{n}{n_k}\frac{1}{2\pi i}\oint_{\partial\Omega_k}\left[\frac{N}{n}\left(1+\sigma^2\underline{\hat{m}}(z)\right)+\frac{N-n}{n}\frac{1}{z\hat{m}(z)}\right] \times\left[-\frac{1}{z\underline{\hat{m}}(z)}-\frac{\underline{\hat{m}}'(z)}{\underline{\hat{m}}(z)^2}-\frac{\hat{m}'(z)}{\hat{m}(z)\underline{\hat{m}}(z)}\right]dz \tag{24}$$

where we recall $\hat{m}(z) \triangleq \frac{1}{N}\sum_{i=1}^N \frac{1}{\lambda_i - z}$ is the eigenvalue distribution of $\mathbf{B}_N = \frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$ and $\hat{m}(z) = \frac{1}{N}\sum_{i=1}^N \frac{1}{\lambda_i - z}$ is that of $\frac{1}{M}\underline{\mathbf{Y}}^{\mathsf{H}}\underline{\mathbf{Y}}$.

The integrand can then be expanded into 9 terms, for which residue calculus [12] can easily be performed. Denote first $\eta_1, \ldots, \eta_N$ the N real roots of $\hat{m}(z) = 0$ and $\mu_1, \ldots, \mu_N$ the N real roots of $\underline{\hat{m}}(z) = 0$. We identify three sets of possible poles for the 9 aforementioned terms: (i) the set $\{\lambda_1, \ldots, \lambda_N\} \cap \Omega_k$, (ii) the set $\{\eta_1, \ldots, \eta_N\} \cap \Omega_k$ and (iii) the set $\{\mu_1, \ldots, \mu_N\} \cap \Omega_k$. The full calculus, detailed in [8], leads to

$$\hat{P}_k = \frac{NM}{n_k(M-N)}\Big[\sum_{\substack{1\leq i\leq N\\ \eta_i\in\Omega_k}}\eta_i - \sum_{\substack{1\leq i\leq N\\ \mu_i\in\Omega_k}}\mu_i\Big] \tag{25}$$

From [11], we know that for N sufficiently large, with probability one, there will be no eigenvalue of $\mathbf{B}_N$ outside the support of F and the number of eigenvalues inside the cluster corresponding to P_k is exactly n_k. Supposing N large, since Ω_k encloses the cluster corresponding to P_k only, $\{\lambda_1, \ldots, \lambda_N\} \cap \Omega_k = \{\lambda_i, i \in \mathcal{N}_k\}$. Also, for any $i \in \{1, \ldots, N\}$, it is easy to see from its definition that $\hat{m}(z) \to \infty$ when $z \to \lambda_i$ from below and $\hat{m}(z) \to -\infty$ when $z \to \lambda_i$ from above. Therefore $\hat{m}(z) = 0$ has at least one solution in each interval $(\lambda_i, \lambda_{i+1})$, hence $\mu_1 < \lambda_1 < \mu_2 < \ldots < \mu_N < \lambda_N$. This implies that, if Ω_k contains $\lambda_{i_0}, \ldots, \lambda_{i_0+(n_k-1)}$, then it also contains $\{\mu_{i_0+1}, \ldots, \mu_{i_0+(n_k-1)}\}$. In [8], we show that in fact, for N large, Ω_k contains exactly $\mu_{i_0}, \ldots, \mu_{i_0+(n_k-1)}$ and no other μ_i. The same result holds for $\eta_{i_0}, \ldots, \eta_{i_0+(n_k-1)}$. This therefore leads to the expression of $\hat{P}_k$ given in (20).

Now, we know from [9] that $\hat{m}(z) \xrightarrow{\text{a.s.}} m(z)$ and $\underline{\hat{m}}(z) \xrightarrow{\text{a.s.}} \underline{m}(z)$ as $N \to \infty$. Observing that the integrand in (24) is uniformly bounded on the compact $\partial\Omega_k$, the dominated convergence theorem [13] ensures $\hat{P}_k \xrightarrow{\text{a.s.}} P_k$.

It now remains to show that the η_i's and the μ_i's are the eigenvalues of $\mathrm{diag}(\boldsymbol{\lambda}) - \frac{1}{N}\sqrt{\boldsymbol{\lambda}}\sqrt{\boldsymbol{\lambda}}^{\mathsf{T}}$ and $\mathrm{diag}(\boldsymbol{\lambda}) - \frac{1}{M}\sqrt{\boldsymbol{\lambda}}\sqrt{\boldsymbol{\lambda}}^{\mathsf{T}}$ respectively. For this, we need the following lemma, proven in parallel in [8] and [14],

Lemma 1: Let $\mathbf{A} \in \mathbb{R}^{N\times N}$ be diagonal with entries $\lambda_1, \ldots, \lambda_N$, and let $\mathbf{y} \in \mathbb{R}^N$. Then the eigenvalues of $\mathbf{A} - \mathbf{y}\mathbf{y}^{\mathsf{H}}$ are the N real solutions of the following equation in x,

$$\sum_{i=1}^N \frac{y_i^2}{\lambda_i - x} = 1 \tag{26}$$

Applying Lemma 1 to $\mathbf{A} = \mathrm{diag}\,\boldsymbol{\lambda}$ and $\mathbf{y} = \sqrt{\frac{1}{N}\boldsymbol{\lambda}}$, we find that the eigenvalues of $\mathrm{diag}(\boldsymbol{\lambda}) - \frac{1}{N}\sqrt{\boldsymbol{\lambda}}\sqrt{\boldsymbol{\lambda}}^{\mathsf{T}}$ are the solutions of

$$\sum_{i=1}^N \frac{\frac{1}{N}\lambda_i}{\lambda_i - x} = 1 \tag{27}$$

which is equivalent to $\hat{m}(x) = 0$ and whose solutions are by definition $\eta_1, \ldots, \eta_N$. The same argument applies similarly to $\mu_1, \ldots, \mu_N$.

C. Discussion

Theorem 2 states that, under spectrum separability condition, when $n_1, \ldots, n_K$ are known *a priori* to the receiver, then $\hat{P}_1, \ldots, \hat{P}_K$ are consistent estimators for $P_1, \ldots, P_K$. Now, in practice, it is rare that $n_1, \ldots, n_K$ and even K are *a priori* known to the receiver. However, if separability is assumed, i.e. the ratio n/N is sufficiently small, then one can estimate simultaneously $K, n_1, \ldots, n_K$ and $P_1, \ldots, P_K$. This is performed by (i) "visually" determining the clusters of the empirical eigenvalues of $\mathbf{B}_N$ (every jump of eigenvalue characterizes a new cluster), which determines K, (ii) counting the number of eigenvalues in each cluster to determine the multiplicities $n_1, \ldots, n_K$ and (iii) evaluating $\hat{P}_1, \ldots, \hat{P}_K$ from Theorem 2.

However, in practical applications, it is obviously impossible to ensure the cluster separability condition. If the condition is not met, say the empirical eigenvalues corresponding to p values $P_i, \ldots, P_{i+(p-1)}$ are merged into a single cluster, then applying the method described above leads to an estimator of the mean $\frac{1}{n}\sum_{k=0}^{p-1} n_k P_k$ (since the contour of integration encloses all the values), instead of an estimator of the individual values. In this case, the receiver can therefore only say that a given estimate $\hat{P}_k$ obtained from Theorem 2 corresponds either to a single transmit source with dimension n_k or to multiple transmit sources with average transmit power well approximated by $\hat{P}_k$ of cumulated dimension n_k; for practical blind detection purposes, this might be good enough. Further investigation is being currently carried out to alleviate the cluster separability constraint.

IV. Simulations

We provide hereafter simulation results for the model presented in Figure 1, i.e. $K = 3$, $P_1 = 1$, $P_2 = 3$, $P_3 = 10$, $n_1/n = n_2/n = n_3/n = 1/10$ and $n/N = N/M = 1/10$. The SNR is 10 dB. In Figure 2, we present results both for $n = 60$ and $n = 6$. These are compared against a classical moment-based approach. The latter consists in computing the first three moments of the empirical eigenvalue distribution of $\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}$, i.e. $\frac{1}{N}\operatorname{tr}\left(\frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\right)^k$, for $k = 1, 2, 3$, from which the *deconvolved* asymptotic moments $\frac{1}{3}(P_1^k + P_2^k + P_3^k)$ of $\mathbf{P}$ can be evaluated, see e.g. [15]; these moments can then be inverted using Newton-Girard polynomial formulas [16] to obtain estimates of P_1, P_2, P_3. We observe a particularly accurate fit between the empirical estimates $\hat{P}_1, \hat{P}_2, \hat{P}_3$ and the true P_1, P_2, P_3 in the case $n = 60$, while the case $n = 6$ at least demonstrates the robustness (and the apparent unbiasedness) of this Stieltjes transform based technique. In comparison, the moment-based approach estimates are slightly less accurate, and more biased for small n.[5] Other approaches than Newton-Girard inversion, such as minimum mean square error estimates (assuming asymptotic Gaussian behaviour of the moments) [5] would give better results but at an extremely higher computational price.

V. Conclusion

In this paper, a blind source power estimator was derived. Under the assumptions that the ratio between the number of signal sources and the number of sensors is small and the source transmit powers are sufficiently distinct from one another, we derived a method to infer the number of antennas of each source as well as its approximate transmit power. Simulations show that the performance achieved by this novel approach is extremely satisfactory compared to classical moment-based approaches, is computationally inexpensive and is particularly robust to small system dimensions.

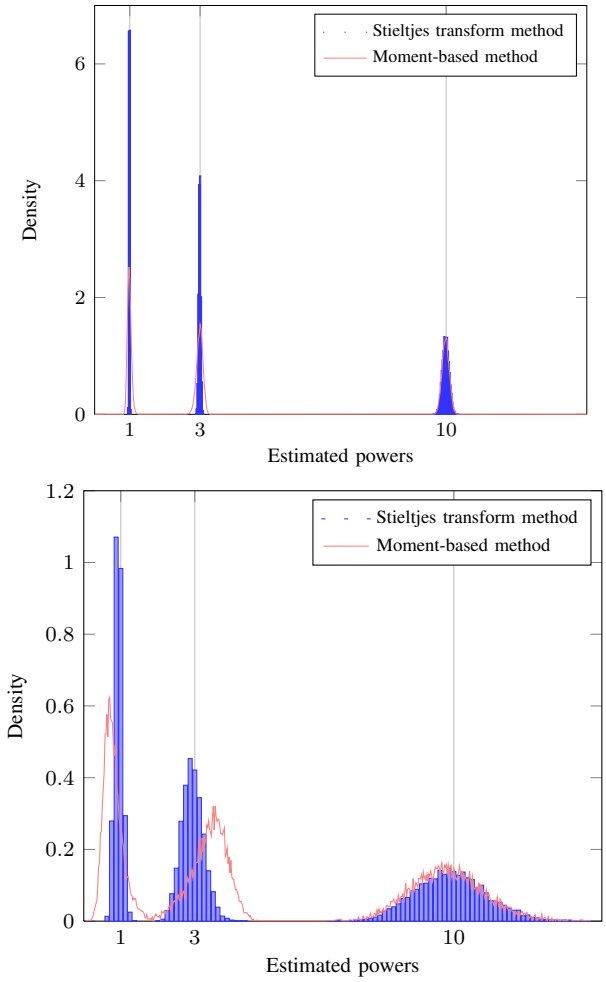

Fig. 2. Multi-source power estimation, for $K = 3$, $P_1 = 1, P_2 = 3, P_3 = 10$, $n_1/n = n_2/n = n_3/n = 1/3$,$n/N = N/M = 1/10$, SNR $= 10$ dB, for $10,000$ simulation runs; Top $n = 60$, bottom $n = 6$.

VI. Acknowledgment

Silverstein's work is supported by the U.S. Army Research Office, Grant W911NF-09-1-0266. Debbah's work is supported by the European Commission, FP7 Network of Excellence in Wireless Communications NEWCOM++.

References

[1] J. Mitola, and Jr GQ Maguire, "Cognitive radio: making software radios more personal", Personal Communications, IEEE [see also IEEE Wireless Communications] 6(4), 13-18, 1999.

[2] H. Urkowitz, "Energy detection of unknown deterministic signals," Proc. of the IEEE, vol. 55, no. 4, pp. 523-531, Apr. 1967

[3] V. I. Kostylev, "Energy detection of a signal with Random Amplitude", Proc IEEE Int. Conf. on Communications (ICC'02). New York City, pp. 1606-1610, May 2002.

[4] N. El Karoui, "Spectrum estimation for large dimensional covariance matrices using random matrix theory," Annals of Statistics, vol. 36, no. 6, pp. 2757-2790, 2008.

[5] N. R. Rao, J. A. Mingo, R. Speicher, A. Edelman, "Statistical eigen-inference from large Wishart matrices," Annals of Statistics, vol. 36, no. 6, pp. 2850-2885, 2008.

[5] note also that sometimes the Newton-Girard formulas returns purely complex roots, which were discarded here.

[6] R. Couillet, M. Debbah, "Free deconvolution for OFDM multicell SNR detection", Proceedings of IEEE PIMRC 2008, Cannes, France.

[7] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," IEEE trans. on Information Theory, vol. 54, no. 11, pp. 5113-5129, 2008.

[8] R. Couillet, J. Silverstein, M. Debbah, "Limiting results on statistical eigen-Inference for wireless communication applications," *to be submitted to IEEE trans. on Information Theory*.

[9] J. W. Silverstein, Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," Journal of Multivariate Analysis, vol. 54, no. 2, pp. 175-192, 1995.

[10] S. Choi, J. W. Silverstein, "Analysis of the limiting spectral distribution of large dimensional random matrices" Journal of Multivariate Analysis, vol 54, no. 2, pp. 295-309, 1995.

[11] J. W. Silverstein, Z. Bai, "Exact separation of eigenvalues of large dimensional sample covariance matrices" Annals of Probability, vol. 27, no. 3, pp. 1536-1555, 1999.

[12] W. Rudin, *Real and complex analysis*, McGraw-Hill, 2006.

[13] P. Billingsley, "Probability and measure," Wiley New York, 2008.

[14] D. Gregoratti, X. Mestre, "Random DS/CDMA for the amplify and forward relay channel," IEEE Trans. on Wireless Communications, vol. 8, no. 2, pp. 1017-1027, 2009.

[15] Ø. Ryan, M. Debbah, "Free Deconvolution for Signal Processing Applications," Proceedings of IEEE ISIT, pp. 1846-1850, 2007.

[16] R. Seroul, D. O'Shea, "Programming for Mathematicians", Springer, 2000