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 Debba
Alcatel-Lucent Chair, Supélec
Plaque 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 active sources.

I.1 Transmitter Power estimation

The detection problem boils down to estimating 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 $(\cdot)^T$ and $(\cdot)^H$, respectively. We denote by $\mathbb{C}^+$ the set $\{z \in \mathbb{C}, \Im[z] > 0\}$. The symbol $\xrightarrow{a.s.}$ 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 $x_k^{(m)} \in \mathbb{C}^{n_k}$, 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 $\sigma^2$ on every sensor; we denote $\sigma_w^{(m)} \in \mathbb{C}^N$ the

1Eq. 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.

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

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 $\mathbf{YY}^H$, 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 $\mathbf{YY}^H$, and not only the first moments. This problem is addressed in [7] for the sample covariance matrix model $\mathbf{Y}' = \mathbf{P}^\frac{1}{2} \mathbf{X} + \mathbf{I}$, i.e. the entries of $\mathbf{P}$ are inferred from the full eigenvalue distribution of $\mathbf{XX}^H$.
receive noise vector where \( w_{k}^{(m)} \) has i.i.d. entries with zero mean and variance 1. At time \( m \), the receiver therefore senses the signal \( y^{(m)} \) defined as

\[
y^{(m)} = \sum_{k=1}^{K} \sqrt{P_k} H_k x_k^{(m)} + \sigma w^{(m)}
\]

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

\[
Y = \sum_{k=1}^{K} \sqrt{P_k} H_k x_k + \sigma W
\]

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

\[
Y = HP^2 X + \sigma W
\]

where \( P \in \mathbb{R}^{n \times n} \) is diagonal with first \( n_1 \) entries \( P_1 \), subsequent \( n_2 \) entries \( P_2 \), and last \( n_K \) entries \( P_K \). \( H = [H_1, \ldots, H_K] \in \mathbb{C}^{N \times n} \) and \( X = [X_1^T, \ldots, X_K^T]^T \in \mathbb{C}^{n \times M} \).

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 \( Y \). That is, for all \( i \), \( \hat{P}_i - P_i \xrightarrow{a.s.} 0 \), 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 \( \frac{1}{M} YY^H \) 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

\[
m(z) = \int \frac{1}{t-z} dF(t)
\]

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

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

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 \( B_N = \frac{1}{M} YY^H \) with eigenvalues \( \lambda_1, \ldots, \lambda_N \). Denote \( \hat{m}(z) \triangleq \frac{1}{N} \sum_{k=1}^{N} \frac{1}{1-z} \) the Stieltjes transform of the eigenvalue distribution \( F_{B_N} \) of \( B_N \). Then, for \( M, N, n \) growing large with limit ratios \( M/N \to c \), \( N/n_k \to c_k \), \( N/n \to c_0 \), \( 0 < c, c_0, c_1, \ldots, c_K < \infty \), for any \( z \in \mathbb{C}^+ \), we have

\[
\hat{m}(z) \xrightarrow{a.s.} m(z)
\]

where \( m(z) \) is defined as

\[
m(z) = cm(z) + (c - 1) \frac{1}{z}
\]

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

\[
\frac{1}{m(z)} = -\sigma^2 + \frac{1}{f(z)} - \sum_{k=1}^{K} \frac{1}{c_k} + P_k f(z)
\]

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

\[
f(z) = (c - 1)m(z) - czm(z)^2
\]

This implies that \( F^{B_N} \) tends weakly and almost surely to a limit \( F \), called the limit spectral density (l.s.d.) of \( B_N \), 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

\[
Y = \left( HP^2 - \sigma I_N \right) \begin{pmatrix} X \\ W \end{pmatrix}
\]

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

\[
Y = \begin{pmatrix} HP^2 - \sigma I_N \\ 0 \end{pmatrix} \begin{pmatrix} X \\ W \end{pmatrix}
\]

We recognize that \( \frac{1}{M} YY^H \) is a sample covariance matrix, with random population covariance matrix \( \left( HP^2 + \sigma^2 I_N \right) \) while \( \left( \tilde{X} \right) \) 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 \( \frac{1}{M} YY^H \) converges almost surely to a limit \( E \) whose Stieltjes transform \( m(z) \) satisfies, for \( z \in \mathbb{C}^+ \),

\[
z = -\frac{1}{m(z)} + \frac{1 + c_0}{c_0} \int \frac{t}{1 + t m(z)} dH(t)
\]

\[
= -\frac{1}{m(z)} + \frac{1 + c_0}{c_0} \frac{1}{m(z)} \left( 1 - \frac{1}{m(z)} \right) \left( \int \frac{dH(t)}{t - (-\frac{1}{m(z)})} \right)
\]

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

\[
w = -\frac{1}{m_1(w)} + \sum_{k=1}^{K} \frac{1}{c_k} + P_k m_1(w)
\]

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

\[
\int \frac{1}{t - (w + \sigma^2)} dH(t) = \frac{c_0}{1 + c_0} m_1(w) - \frac{1}{1 + c_0} \frac{1}{w}
\]

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

\[
z = -\frac{1}{1/c m(z)^2} m_1(-1/m(z) - \sigma^2) - \left( 1 - \frac{1}{c} \right) \frac{1}{m(z)}
\]
where, according to (14), \( f(z) \triangleq m_1(-1/m(z) - \sigma^2) \) satisfies

\[
\frac{1}{m(z)} = -\sigma^2 + \frac{1}{f(z)} - \sum_{k=1}^{K} \frac{1}{c_k} P_k f(z)
\]  

(17)

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

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

\[
m(z) = cm(z) + (c-1) \frac{1}{z}
\]  

(18)

This completes the proof of Theorem 1.

Therefore, the support of the l.s.d. \( F \) of \( \frac{1}{M}YY^H \) can be evaluated as follows: for any \( z \in \mathbb{C}^+ \), \( m(z) \) is given by (7), in which \( m(z) \) 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, where \( P \) has three distinct values \( P_1 = 1, P_2 = 3, P_3 = 10 \) and \( n_1 = n_2 = n_3, N/n = 10, M/N = 10, \sigma^2 = 0.1 \). Empirical test: \( n = 60 \).

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 \( P \hat{X} \).

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 \( P \hat{X} \) [11].

B. Eigen-inference

In this section, we prove our main result.

**Theorem 2:** Let \( B_N \in \mathbb{C}^{N \times N} \) be defined as in Theorem 1, and \( \lambda = (\lambda_1, \ldots, \lambda_N), \lambda_1 < \cdots < \lambda_N, \) be the vector of the ordered eigenvalues of \( B_N \). Further assume that the limiting ratios \( c, c_0 \) and \( P \) 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

\[
\hat{P}_k - P_k \xrightarrow{a.s.} 0
\]  

(19)

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

\[
\hat{P}_k = \frac{NM}{n_k(M-N)} \sum_{i \in N_k} (\eta_i - \mu_i)
\]  

(20)

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

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

\[
P_k = c_k \frac{1}{2\pi i} \oint_{C_k} \sum_{i=1}^{K} \frac{1}{c_i} \frac{P_r - \omega}{P_r - \omega} d\omega
\]  

(21)

for any negatively oriented contour \( C_k \subset \mathbb{C} \), 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 \( C_k \) which is intimately linked to \( m(z) \), the Stieltjes transform of the l.s.d. of \( B_N \). Instead of evaluating \( P_k \), we then evaluate the complex integral resulting from replacing the deterministict \( m(z) \) by the empirical \( \hat{m}(z) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - z} \).

From the convergence \( \hat{m}(z) \xrightarrow{a.s.} m(z) \) 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 \(-1/m_1(x_k^-) < P_k < -1/m_1(x_k^+)\). 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 \( \hat{\Gamma}_k \subset \mathbb{C} \) to be any continuous curve with endpoints \( x_k^- \) and \( x_k^+ \) and with interior points of strictly positive imaginary part. We then define \( \tilde{C}_k \) to be the union of the curve \(-1/m_1(\hat{\Gamma}_k)\) with \( \hat{\Gamma}_k \) oriented from \( x_k^- \) to \( x_k^+ \) and the curve \(-1/m_1(\hat{\Gamma}_k)\) with \( \hat{\Gamma}_k \) 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 \( \tilde{C}_k \), while \( P_i \neq k \), 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 $\Gamma_k$ the surface enclosed in the union of $\Gamma_k$ from $x_k$ to $x_k^+$ and $\Gamma_k^-$ from $x_k^-$ to $x_k$, (21) becomes

$$P_k = i c_k \frac{1}{2\pi i} \oint_{\partial \Omega_k} \left[ c \frac{1}{m_1(w) + (c - 1)} \frac{m_1'(w)}{m_1(w)} dw \right]$$

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

$$P_k = c_k \frac{1}{2\pi i} \oint_{\partial \Omega_k} \left[ c \left( 1 + \sigma^2 m(z) \right) \right] \times \left[ -\frac{1}{z m(z)} - \frac{m'(z)}{m(z)} - \frac{m''(z)}{m(z)^2} \right] dz$$

where $\Omega_k$ is the surface described by $-1/m(\Gamma_k) - \sigma^2$. Similarly as we show that $\Gamma_k$ encloses $P_k$ and none of the $P_i$, $i \neq k$, we show in [8] that $\Omega_k$ contains the 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

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

where we recall $\hat{m}(z) \triangleq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z_i - z}$ is the eigenvalue distribution of $\mathbf{B}_N = \frac{1}{n^2} \mathbf{Y} \mathbf{Y}^H$ and $\hat{m}(z) = \frac{1}{n} \sum_{i=1}^{N} \frac{1}{z_i - z}$ is that of $\frac{1}{n} \mathbf{Y} \mathbf{Y}^H$.

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 $\hat{m}(z) = 0$. 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

$$\hat{P}_k = \frac{n M}{n_k (M - N)} \left[ \sum_{1 \leq i < N}^{N} \eta_i - \sum_{1 \leq i < N}^{N} \mu_i \right]$$

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 $\Omega_k$ encloses the cluster corresponding to $P_k$ only, $\{\lambda_1, \ldots, \lambda_N\} \cap \Omega_k \Rightarrow \{\lambda_i, i \in \mathbb{N} \}$. Also, for any $i \in \{1, \ldots, N\}$, it is easy to see from its definition that $\hat{m}(z) \rightarrow \infty$ when $z \rightarrow \lambda_i$ from below and $\hat{m}(z) \rightarrow -\infty$ when $z \rightarrow \lambda_i$ 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 $\Omega_k$ contains $\lambda_1, \ldots, \lambda_{i_0+(n_k-1)}$, 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, $\Omega_k$ contains exactly $\mu_{i_0}, \ldots, \mu_{i_0+(n_k-1)}$ and no other $\mu_i$. The same result holds for $\eta_0, \ldots, \eta_{i_0+(n_k-1)}$. This therefore leads to the expression of $P_k$ given in (20).

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

It now remains to show that the $\eta_i$’s and the $\mu_i$’s are the eigenvalues of $\text{diag}(\lambda_i - \frac{1}{n} \sqrt{\lambda_i \lambda_i^T})$ and $\text{diag}(\lambda_i - \frac{1}{n} \sqrt{\lambda_i \lambda_i^T})$ respectively. For this, we need the following lemma, proven in parallel in [8] and [14].

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

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

Applying Lemma 1 to $A = \text{diag} \lambda$ and $y = \sqrt{n} \lambda$, we find that the eigenvalues of $\text{diag}(\lambda_i - \frac{1}{n} \sqrt{\lambda_i \lambda_i^T})$ are the solutions of

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

which is equivalent to $\hat{m}(z) = 0$ 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 $P_1, \ldots, P_K$ 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 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 $P_1, \ldots, P_K$ 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 $P_1, \ldots, P_{i+(p-1)}$ are merged into a single cluster, then applying the method described above leads to an estimator of the mean $\frac{1}{n} \sum_{k=1}^{p} n_k P_k$ (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. In [8], a characterization of the conditions to ensure cluster separability is documented. Further investigation is being currently carried out to alleviate the cluster separability constraint.
IV. Simulations

We now provide simulation results for the model presented in Figure 1, i.e. $K = 3$, $(P_1, P_2, P_3) = (1, 3, 10)$, $n_1/n = n_2/n = n_3/n = 1/10$ and $n/N = N/M = 1/10$. The SNR is 10 dB. In Figure 2, we present results 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 eigenvalue distribution of $\frac{1}{N} YY^H$, i.e. $\frac{1}{N} \text{tr} (\frac{1}{M} YY^H)^k$, for $k = 1, 2, 3$, from which $\frac{1}{3}(P_1^k + P_2^k + P_3^k)$ of $P$ can be evaluated for these $k$, see e.g. [15]; these values 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 $n$. 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.

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