Noise-Enhanced Blind Multiple Error Rate Estimators in Wireless Relay Networks

Jen-Yang Liu, *Student Member, IEEE*, and Yu T. Su, *Senior Member, IEEE*

***Abstract*—Data detection or fusion based on output from multiple wireless links often requires channel state information (CSI) about the links' error rate (ER) performance. We consider the scenario that these links include direct source–destination (SD) links and two-hop links that require an intermediate decode-and-forward (DF) node to relay the source signal. Conventional destination-based estimators suffer from slow convergence and are incapable of simultaneously blind estimating all ERs, including, in particular, those of the source–relay (SR) links. They may also require various degrees of CSI about the ERs of the SD and relay–destination (RD) links to remove the ambiguity arising from the insufficient number of links in the network and from that due to the symmetric nature of a cascaded source–relay–destination link's ER as a function of its component SR and RD links' ERs. We propose novel Monte-Carlo-based estimators that overcome all these shortcomings. The estimation process involves injecting noise into the samples received by the destination node to create virtual links and alter link output statistics. We show that the latter scheme exhibits a stochastic resonance effect, i.e., its mean squared estimation error (MSEE) performance is enhanced by injecting proper noise, and there exists an optimal injected noise power level that achieves the maximum improvement. The stochastic resonance effects are analyzed, and numerical examples are provided to display our estimators' MSEE behaviors, as well as to show that the ER performance of the optimal detector using the proposed estimators is almost as good as that with perfect ER information.**

***Index Terms*—Nonlinear detection, nonlinear estimation, relays, stochastic resonance.**

I. Introduction

WE CONSIDER the basic scenario illustrated in Fig. 1, where the destination node (DN) d receives sequences originating from the same source node (SN) s via multiple (L) flat-fading links. These links may include a direct single-hop (SH) source–destination (SD) link and indirect two-hop links, each connecting SN and DN with the help of an intermediate relay node (RN), say, r_k. Such a scenario occurs, for example, in a cooperative communication network (CCN), in which the SD communication is aided by single or multiple relays that act as virtual antennas to allow resource sharing and provide spatial diversity gains [1]. Another popular example is the so-called central estimating officer (CEO) problem associated with a wireless sensor network, where each sensor sends its measurement to the CEO that often does not have direct access to the SN [2]. It is the CEO's responsibility to reliably recover the source information based on the data it has received from various sensors [3].

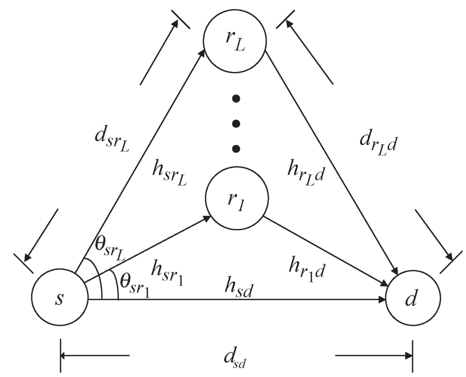


Fig. 1. Wireless multiple-relay network.

For convenience of subsequent discourse, we define a single-relay CCN as one that consists of a source, a relay (or sensor), and a DN only. We refer to the associated SD, source–relay (SR), and relay–destination (RD) links as component links and the indirect source–relay–destination (SRD) link as cascaded link. Although many sensing relay schemes have been proposed, we only consider the decode-and-forward (DF) scheme [1]–[8] for which an RN (sensor) demodulates/decodes the received signal from the SN and remodulates/reencodes the decoded bit stream before retransmitting.

Since a sensor or cooperative RN may erroneously detect or sense its received signal, conventional maximal ratio combining (MRC) or a similar fusion rule is no longer optimal for the DN. In fact, data fusion of various kinds in the presence of imperfect DF relays [6]–[8] and relay selection in a DF-based CCN [9] require some forms of channel state information (CSI). Depending on the modulation used, the required CSI includes short-term CSI (ST-CSI), like instantaneous link gains and signal-to-noise ratios (SNRs), and long-term CSI (LT-CSI), such as average link gains and error rates (ERs) of the component links. The former has intensively been studied in terms of channel estimation, gain control, and carrier recovery loops, whereas the LT-CSI receives much less attention.

Pilot-aided ER estimators are obtainable at the cost of increasing the RNs' computing load and result in bandwidth and power efficiency reductions. The overhead and delay become significant if the true ER is small, the packet size is small,

Manuscript received April 10, 2011; revised September 9, 2011 and November 22, 2011; accepted January 6, 2012. Date of publication January 23, 2012; date of current version March 21, 2012. This work was supported in part by the National Science Council of Taiwan under Grant NSC 96-2221-E-009-076-MY3. The review of this paper was coordinated by Prof. X. Wang.

The authors are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 30056, Taiwan (e-mail: flopyandteet@gmail.com; ytsu@mail.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2012.2185258

or if the number of RNs is large. It is therefore desired that a DN performs all ER estimation tasks blindly. Moreover, blind estimation is mandatory in a sensor network where there is no direct link and the source usually does not or cannot transmit pilot symbols; the fact that it requires less transmission overhead also suits a sensor network's critical need as sensor nodes usually are battery-limited devices.

For multiple-relay networks, the ER estimation problem can be transformed into one of solving a system of nonlinear equations. Each equation describes a relation among the ERs of a pair of links, and the probability that the same bit transmitted through these two independent links is decoded with identical decision. Using all available link pairs and assuming no hidden SR links, Dixit *et al.* [10] converted the problem into a structured eigenvalue task and proposed a modified power method to find the solution. Delmas and Meurisse [11] suggested an expectation-maximization (EM)-based blind ER estimator that outperforms Dixit's estimator by using the method of moments based solution of the nonlinear system as the initial estimate. These novel approaches, however, suffer from some drawbacks. First, the nonlinear system is underdetermined unless we have sufficient relays so that the number of distinct link pair combinations is no smaller than the unknown ERs. Second, even if there are enough RNs, it is not possible to simultaneously estimate all (SR, SD, and RD links) ERs, and LT-CSI is needed to resolve the ambiguity resulting from the fact that the ER of a cascaded SRD link is a symmetric function of the corresponding component links' ERs. Finally, the convergence rate is slow, whence it often takes a long period to obtain a reliable estimate.

It is the purpose of this paper to present novel blind ER estimation schemes that overcome all the foregoing shortcomings. To simply our presentation, we mostly focus on the CCN scenario with the understanding that the proposed schemes can readily be applied in other similar scenarios. As a prelude, we briefly review a unified system model for a multiple-relay wireless network and describe the corresponding maximum likelihood (ML) detector and ER estimator structures in Section II. We begin our discussion with the simplest case of a binary phase-shift keying (BPSK)-based single-relay CCN, assuming the required ST- and LT-CSI's of RD and SD links are all available to the DN, i.e., the only unknown CSI that needs to be estimated is the average ER of the SR link. Even for this case, we show that blind ML ER estimation based on the DN's matched filter outputs requires high computational complexity and storage cost. A simple CSI-aided average-count-based estimator is thus given. We then extend the approach to multiple-relay CCNs with less LT-CSI and obtain the basic nonlinear system (set of equations) for a three-link (two relays plus a direct SD link) CCN and its solution. Some properties of the proposed estimator are given in the same section as well.

The main results are presented in Sections III–VI. In Section III, we discuss the ER ambiguity in a cascaded link and propose a novel approach that creates virtual SD/RD links by either rotating or injecting noise into link output samples to resolve the ambiguity and estimate all ERs without the help of multiple RNs. We show in the same section that the same concept can be applied to binary frequency-shift keying (BFSK)- and differential phase-shift keying (DPSK)-based systems. In Section IV, we first address the convergence rate issue and suggest a simple scheme to improve the virtual-link-aided (VLA) approach using a multitude of virtual links (VLs) to obtain what we call the enhanced VLA (EVLA) estimator. We proceed to propose a more subtle approach that is conceptually similar to importance sampling (IS)-based simulations and is therefore referred to as the IS-inspired (ISI) estimator. The ISI estimator also needs to inject noise into link outputs, but the purpose of noise injecting is not for building VLs but for modifying the output statistic and producing more importance events. Important properties of the proposed estimators and the associated mean squared estimation error (MSEE) performance analysis are given in Section V.

Section VI presents the simulated performance of the proposed schemes and shows that the ML detector using the ERs estimated by our schemes yields performance almost as good as that with perfectly known ERs. Furthermore, a hybrid of the EVLA and ISI (or ISI-VLA) methods is capable of offering significant variance reduction. Both analysis and simulations prove that the ISI estimator exhibits a stochastic resonance effect, i.e., its MSEE performance is improved by injecting noise into the received samples, and there exits an optimal injected noise power that achieves the maximum improvement. Finally, concluding remarks are provided in Section VII.

II. Preliminary

This section begins with descriptions of a generic system model, assumptions, and related parameter definitions. The expressions of the ML data detector and blind ER estimator are then given. The second and third sections review some side-information-aided blind ER estimators for single- and multiple-relay networks. We will frequently refer to these materials in subsequent discussions.

A. System Model, ML Detection, and Blind ER Estimator

We follow the conventional assumption of using a two-phase time division duplex cooperative communication scheme in which the SN in Fig. 1 transmits a sequence of independent identically distributed (i.i.d.) ± 1-valued data $\{x[n]\}$, and all L RNs listen, decode, and re-encode the received message in the first phase. The synchronous samples received by the DN and the kth RN in this phase are

$$y_{sd}[n] = h_{sd}[n]\sqrt{P_s}x[n] + w_{sd}[n] \tag{1.a}$$

$$y_{sr_k}[n] = h_{sr_k}[n]\sqrt{P_s}x[n] + w_{sr_k}[n] \tag{1.b}$$

where P_s is the signal power and the additive noise components, and $w_{sd}[n]$ and $w_{sr_k}[n]$ are independent zero-mean complex white Gaussian random variables with variances σ_d^2 and σ_r^2, respectively. We assume that the complex link gains $h_{ij}[n]$ for the link from node i to node j, where $(i,j) \in \{(s,r_k),(s,d),(r_k,d); k = 1,\ldots,L\}$, and the corresponding noise terms $w_{ij}[n]$ are mutually independent. The RNs send the re-encoded message to the DN in the second phase. Since RNs may detect erroneously, the retransmitted signals are not necessarily equal to $x[n]$. If we denote $\hat{x}_{r_k}[n]$ as the signal sent by the kth relay and $y_{r_k d}[n]$ as the corresponding received sample at the DN in this phase, then

$$y_{r_k d}[n] = h_{r_k d}[n]\sqrt{P_{r_k}}\hat{x}_{r_k}[n] + w_{r_k d} \tag{2}$$

where P_{r_k} is the transmitted signal power of the kth RN, and $w_{r_k d}[n]$ has the same distribution as $w_{sd}[n]$. For frequency-flat fast Rayleigh fading links, $|h_{ij}|^2$ are independent exponentially distributed random variables with variance σ_{ij}^2.

Define the memoryless nonlinearity as

$$f_T(z;\varepsilon) = \ln\left[\frac{\varepsilon + (1-\varepsilon)e^z}{(1-\varepsilon)+\varepsilon e^z}\right], \quad 0 < \varepsilon < 1/2 \tag{3}$$

and for $k = 1, \ldots, L$, the weighting functions

$$q_0(y[n]) = \Re\left\{4h_{sd}^*[n]\sqrt{P_s}y[n]/\sigma_d^2\right\} \tag{4.a}$$

$$q_k(y[n]) = \Re\left\{4h_{r_k d}^*[n]\sqrt{P_{r_k}}y[n]/\sigma_d^2\right\} \tag{4.b}$$

where $\Re\{z\}$ denotes the real part of z. Then, the ML detector for BPSK signals is given by [6]

$$\hat{x}[n] = \operatorname{sgn}\left[q_0(y_{sd}[n]) + \sum_{i=1}^{L} f_T\left(q_k(y_{rd}[n]); e_{sr_k}\right)\right] \tag{5}$$

where $\operatorname{sgn}[z]$ denotes the sign of the real number z, and e_{sr_k} is the ER of the link between the source and the kth RN. Equations (3) and (4a)–(4.b) indicate that besides the instantaneous received complex amplitude-to-noise-power ratio, i.e., $(\sqrt{P_{r_k}}h_{r_k d}[n]/\sigma_d^2)$ and $(\sqrt{P_s}h_{sd}[n]/\sigma_d^2)$, the hidden SR link's ER e_{sr_k} should also be known by the DN for ML detection. As the instantaneous complex link gains $h_{r_k d}[n]$ and $h_{sd}[n]$ are difficult to estimate in a high dynamic wireless environment, noncoherent signals are sometimes preferred for they require no such estimations. Nevertheless, [7] and [8] show that ML noncoherent detections of BFSK and DPSK signals by a DN still need LT-CSI such as ERs for both far-end (SR) and near-end (SD and RD) links or σ_d^2.

For notational brevity, we henceforth omit the subscript k associated with the kth relay r_k unless there is danger of ambiguity. The DN of a single-relay BPSK-based CCN has the samples $\{y_{sd}[n], y_{rd}[n]\}$ of (1.a) and (2) as the sufficient statistics for estimating the BERs of its component links. As an i.i.d. source is assumed, we can easily verify that the probability density function (pdf) of $y_{sd}[n]$ is independent of e_{sr} and so is that of $y_{rd}[n]$. With N coherently received sample pairs $\{(q_0(y_{sd}[i]), q_1(y_{rd}[i]))\}_{i=1}^N \stackrel{def}{=} \{(q_0^{(i)}, q_1^{(i)})\}$, the joint conditional pdf $f(y_{sd}, y_{rd}|\mathrm{I}_{csi})$ of the matched filter outputs y_{sd} and y_{rd} given CSI $\{h_{sd}, h_{rd}, \sigma_d^2, e_{sr}\} = I_{csi}$ and unit transmit powers $P_s = P_r = 1$ is a mixture density, and the ML blind e_{sr} estimator is given by

$$\begin{aligned}\hat{e}_{sr} &= \arg\max_{0\le e_{sr}<0.5} \log f\left(\{y_{sd}[i]\}_{i=1}^N, \{y_{rd}[i]\}_{i=1}^N |\mathrm{I}_{csi}\right)\\ &= \arg\max_{0\le e_{sr}<0.5} \sum_{i=1}^{N} \log\left[\cosh\left(\frac{q_0^{(i)}+q_1^{(i)}}{2}\right)\right.\\ &\qquad \left. -2\sinh\left(\frac{q_0^{(i)}}{2}\right)\sinh\left(\frac{q_1^{(i)}}{2}\right)e_{sr}\right].\end{aligned} \tag{6}$$

The reliability of the ML estimator depends on the sample size N, the true e_{sr}, and two other component links' statistics that, in turn, determine those of $q_0^{(i)}$ and $q_1^{(i)}$. For practical ERs, we usually need large N for the ML estimator to converge. The difficulty in implementing this estimator comes from at least three other concerns: 1) the computing complexity of solving the associated nonconvex optimization problem; 2) no existence of recursive formula for updating the objective function whenever a new received signal pair becomes available; and 3) the large required storage space. These implementation considerations convince us to turn to estimators based on the binary sample sequence $\{\widehat{y}_{rd}[n], \widehat{y}_{sd}[n]\}$ produced by

$$\widehat{y}_{rd}[n] = \operatorname{sgn}\left[q_1(y_{rd}[n])\right], \quad \widehat{y}_{sd}[n] = \operatorname{sgn}\left[q_0(y_{sd}[n])\right]. \tag{7}$$

In addition to their simplicity, an important advantage of such estimators is that they can easily be extended to noncoherent binary modulations, whereas the form of the ML estimator is highly modulation dependent.

As a prelude to the study of simultaneous blind estimation of all component links' ERs, we start with the simpler case of SR link ER estimation, assuming the ST-CSI needed the ERs of either all or some of the remaining component links are available.

B. Side-Information-Aided Blind Single-ER Estimation

Since a cascaded link is composed of two (i.e., SR and RD) binary symmetric links (BSLs) with ERs e_{sr} and e_{rd}, the end-to-end ER e_{srd} is given by $e_{srd} = e_{sr}(1-e_{rd}) + (1-e_{sr})e_{rd} = e_{sr} + e_{rd} - 2e_{sr}e_{rd}$. A single-relay CCN can thus be regarded as the composition of two BSLs connecting the source and the destination. We assume stationary component links with time-invariant ERs and refer to the probability $p = \mathrm{P}_r(\widehat{y}_{sd} = \widehat{y}_{rd})$ as the success matching probability (SMP). Using the identity $p = e_{sd}e_{srd} + (1-e_{sd})(1-e_{srd})$ and the i.i.d. source assumption, we immediately have the following identity that relates various ERs to the SMP between a direct SD link and a cascaded SRD link:

$$e_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - p}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} \stackrel{def}{=} e_{sr}(p). \tag{8}$$

Since the links are assumed to be stationary, $W[i] \stackrel{def}{=} I(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i])$, where $I(\mathbf{E}) = 1$ if the statement $\mathbf{E}$ is true, otherwise it is zero, is Bernoulli distributed with success probability p. Furthermore, the SMP can be estimated by

$$\widehat{p}^{(N)} = \sum_{i=1}^{N} \frac{I\left(\widehat{y}_{sd}[i] = \widehat{y}_{rd}[i]\right)}{N} \tag{9}$$

where the superscript (N) indicates that N sample pairs are used to obtain the estimator. This average-count-based estimator is the sample mean of the Bernoulli process $\{W[i]\}$ and is a uniform minimum variance unbiased estimator if i.i.d. samples are received [12].

Using the sample mean estimator (9) as $\hat{p}$, the method of moments and (8) suggest the estimator

$$\widehat{e}_{sr} = \frac{1 - e_{sd} - e_{rd} + 2e_{sd}e_{rd} - \widehat{p}}{1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}} = e_{sr}(\hat{p}) \tag{10}$$

if both e_{rd} and e_{sd} are known.

As $0 \leq e_{sr} \leq 0.5$, our estimator $\widehat{e}_{sr}$ may have to be modified by the soft limiter

$$\mathcal{J}(\widehat{e}_{sr}) = \min\left[\max(\widehat{e}_{sr}, 0), 0.5\right]. \tag{11}$$

In addition, we can easily derive a recursive relation for $\widehat{p}^{(i)}$ to sequentially estimate p and therefore e_{sr}.

The ER estimator (10) has many desired properties that we summarize in the following two lemmas.

Lemma 1: The estimator $\hat{e}_{sr}$ defined by (10) is 1) unbiased and attains the Cramer–Rao lower bound and 2) a uniformly minimum variance unbiased and ML estimator with variance

$$\mathrm{Var}(\hat{e}_{sr}) = \frac{p(1-p)}{N(1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd})^2} \tag{12}$$

where N is the sample size.

Lemma 2: For any $\varepsilon > 0$, we have

$$\Pr\left(|\widehat{e}_{sr} - e_{sr}| \geq \varepsilon\right) \leq 2\exp\left[-\min\left(\frac{N^2\varepsilon^2 C_1^2}{4p}, \frac{N\varepsilon C_1}{2}\right)\right] \tag{13}$$

where $C_1 = 1 - 2e_{sd} - 2e_{rd} + 4e_{sd}e_{rd}$, N is the sample size, and the soft limiting effect (11) is neglected.

The properties given in Lemma 1 resulted from the fact that $\hat{e}_{sr}$ is a linear function of $\hat{p}$ and the invariance property of an ML estimator. Lemma 2, which is derived from using Chernoff's inequality, implies that the estimator $\widehat{e}_{sr}$ converges to e_{sr} in probability.

C. *Multiple-Relay-Aided Blind Multiple ER Estimation*

When there are L RNs, we have $\binom{L+1}{2}$ combinatorial diversities from pairwise hard-decision matchings. For any (k,l) RN pair $k \neq l$, the random variable $W_{kl} = I(\widehat{y}_{r_k d} = \widehat{y}_{r_l d})$ is Bernoulli distributed with success (matching) probability $p_{kl} = P_r[\widehat{y}_{r_k d} = \widehat{y}_{r_l d}]$, which satisfies the identity

$$p_{kl} = Q_k Q_l + (1 - Q_k)(1 - Q_l) \tag{14}$$

with Q_k being the cascaded link ER given by

$$Q_k = e_{sr_k} + e_{r_k d} - 2e_{sr_k}e_{r_k d} \overset{def}{=} e_{sr_k d}. \tag{15}$$

The preceding equations and (8) imply that p_{sr_k} and p_{kl} are related to the parameter sets $\{e_{sd}, e_{sr_k}, e_{r_k d}\}$ and $\{e_{sr_k}, e_{r_k d}, e_{sr_l}, e_{r_l d}\}$, respectively. Following the approach used for the case $L = 1$, we replace p_{sr_k} and p_{kl} in (8) and (14) by the average sample count (sample mean) estimators

$$\widehat{p}_{sr_k} = \sum_{j=1}^{N} \frac{I\left(\widehat{y}_{sd}[j] = \widehat{y}_{r_k d}[j]\right)}{N}, \quad k=1,\ldots,L \tag{16.a}$$

$$\widehat{p}_{kl} = \sum_{i=1}^{N} \frac{I\left(\widehat{y}_{r_k d}[i] = \widehat{y}_{r_l d}[i]\right)}{N}, \quad 1 \leq k < l \leq L \tag{16.b}$$

to obtain $\binom{L+1}{2}$ equations, where all are of the form similar to (14), involving the unknown ERs $\{Q_i\}$ and e_{sd}.

When the RNs are dedicated stationary nodes and $\{e_{r_k d}\}$ can be reliably estimated, there are only $L+1$ unknown parameters $\{e_{sd}, e_{sr_k}, k = 1, \ldots, L\}$, which can be solved if there are at least $L+1$ independent equations. Since $\binom{L+1}{2} \geq L+1$ whenever $L \geq 2$, the unknown link parameters can be estimated, as long as more than two RNs are available.

For general multiple ER estimation in an L-relay CCN, $L > 2$, we can therefore divide the problem into a sequence of subproblems, each dealing with a smallest two-relay problem. The three-link (two relays plus a direct SD link) CCN is referred to as a basic network in which the link ER is governed by a set of nonlinear equations called a basic (nonlinear) system

$$\begin{bmatrix} 1 - Q_1 - e_{sd} + 2e_{sd}Q_1 \\ 1 - Q_1 - Q_2 + 2Q_1Q_2 \\ 1 - e_{sd} - Q_2 + 2e_{sd}Q_2 \end{bmatrix} = \begin{bmatrix} p_{sr_1} \\ p_{12} \\ p_{sr_2} \end{bmatrix} \approx \begin{bmatrix} \widehat{p}_{sr_1} \\ \widehat{p}_{12} \\ \widehat{p}_{sr_2} \end{bmatrix} \tag{17}$$

where $\widehat{p}_{sr_l}$, $l = 1, 2$, and $\widehat{p}_{12}$ are obtained via (16.a) and (16.b). A similar nonlinear system arose in [10], where the estimations of the ERs e_{sd} and Q_l's were attempted. Unlike our case, there is no cascaded links and hence no need to estimate the ERs of the SR and RD links. It can be shown that the solution to the foregoing basic system gives the basic estimators [13]

$$\widehat{Q}_i = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(2\widehat{p}_{ij} - 1)(2\widehat{p}_{ik} - 1)}{2\widehat{p}_{jk} - 1}}, \quad i, j, k \in \{0, 1, 2\} \tag{18}$$

where $\widehat{Q}_0 = \widehat{e}_{sd}$, $\widehat{p}_{01} = \widehat{p}_{sr_1}$, and $\widehat{p}_{02} = \widehat{p}_{sr_2}$.

The foregoing equation indicates that the presence of multiple RD links enables us to estimate e_{sd} and removes the need for e_{sd} side information, i.e., the relay diversity can be traded for the degree of LT-CSI. To estimate the ERs of the multiple hidden (far-end) SR links, we invoke the relation (15), assuming the ERs of all the RD links are known, to obtain

$$\widehat{e}_{sr_k} = \frac{\widehat{Q}_k - e_{r_k d}}{1 - 2e_{r_k d}}, \quad k = 1, 2. \tag{19}$$

Note that an L relay CCN induces $\binom{L+1}{2}$ basic systems (diversities), where each relay is involved in more than one system so that multiple estimates for a given Q_i may be obtained. Dixit [10] had proposed a complex method to take advantage of this fact and obtained improved ER estimates. On the other hand, [11] shows that the basic estimators given by (18) asymptotically achieve the accuracy achieved by the ML pilot-aided estimator based on the two sequences of hard decision pairs $\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}_{i=1}^{N}$, $l = 1, 2$; for finite N, a better estimate is obtained by maximizing the log-likelihood functions $\Gamma(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}) \overset{def}{=} \log f(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}_{i=1}^{N})$, $l = 1, 2$, which are defined as

$$\Gamma\left(\{\widehat{y}_{sd}[i], \widehat{y}_{r_l d}[i]\}\right) = \prod_{i=1}^{N}\left(e_{sd}^{1-I\left(\widehat{y}_{sd}[i]=x[i]\right)}(1-e_{sd})^{I\left(\widehat{y}_{sd}[i]=x[i]\right)} \times \prod_{l=1}^{2} Q_l^{1-I\left(\widehat{y}_{r_l d}[i]=x[i]\right)}(1-Q_l)^{I\left(\widehat{y}_{r_l d}[i]=x[i]\right)}\right).$$

The derivation of the preceding function is similar to that given in [11, Sec. III] with additional consideration of cascaded link ER Q_l. In [11], an EM-based approach was proposed to obtain blind (unknown $x[i]$) estimates of Q_i, which outperform Dixit's method. However, our numerical experiments conclude that, for both approaches, the performance improvement over the basic estimators is rather limited and do not worth the additional high complexity (see Section VI and Fig. 4).

Before presenting our main results in the following sections, we would like to emphasize that most estimators to be developed are based on some variation or extension of the basic system (17), and their expressions, e.g., (24.a)–(25), (28.a)–(28.d), and (30.a)–(31.b), are derivable from variations or extensions of the basic estimators (18) and (19).

III. Blind Multiple-Error Rate Estimation Using Virtual Links

We first examine the ER ambiguity issue associated with the estimation of a far-end component link's ER and then present a novel solution to resolve this ambiguity. The extension to other binary modulations, i.e., BFSK and DPSK, is discussed at the end of this section.

A. ER Ambiguity in a Cascaded Link

As can be seen from (17), when there are sufficient relays, the resulting equation set leads to formulas for the estimates of e_{sd} and Q_k but not those for e_{sr_k} and e_{r_kd}. This is due to the fact that the ER of an SRD link, as (15) has shown, is a symmetric function of the ERs of the associated component SR and RD links, i.e., there are infinite many (e_{sr_k}, e_{r_kd}) pairs that result in the same Q_k. In fact, the legitimate candidates for the latter two ERs consist of the lower left part of the hyperbola defined by (15), i.e., $(1-2Q_k)/4 = (e_{sr_k} - (1/2))(e_{r_kd} - (1/2))$, that lies within the square $\mathcal{S} \stackrel{def}{=} \{(e_{sr_k}, e_{r_kd}) | 0 < e_{sr_k} < (1/2), 0 < e_{r_kd} < (1/2)\}$. The ambiguity in (15) is resolved in the scenario discussed in the last section by specifying e_{r_kd} so that $\widehat{e}_{sr_k}$ is obtained via (19). Geometrically, this is equivalent to finding the intersection of the hyperbola and the line $e_{r_kd} = e$ within the square $\mathcal{S}$, where e is the true ER of the RD link.

When the LT-CSI e_{r_kd} is not available, we need to find a curve that represents another set of legitimate ER pairs and that has only one intersection point with (15) in $\mathcal{S}$. Since the hyperbola is symmetric with respect to the line $e_{r_kd} = e_{sr_k}$ and we have access to the outputs of the RD and SD links only, finding a curve that has a unique intersection with (15) is possible if an alternate RD link is provided. This can be seen by noting that an RD link with a different average bit SNR $\overline{\gamma}$ yields a different equivalent cascaded link with ER Q'_k and, therefore, a curve of the form $(1-2Q'_k)/4 = (e_{sr_k} - (1/2))(\alpha e_{r_kd} - (1/2))$, where α is such that $0 < \alpha e_{r_kd} \stackrel{def}{=} e_{r'_kd} < (1/2)$.

B. VL Methods

To have an alternate physical link (PL), one can purposely vary the power of the bit stream so that the transmitted sequence is equivalent to one formed by multiplexing two data sources with different powers. If the locations of these two parts in the multiplexed data stream are known, the DN then performs separate comparison and counting based on (16.a) and (16.b). Although such a two-level amplitude modulation makes it possible to solve the e_{sr_k} and e_{r_kd} ambiguity, allocating unequal powers to different parts of the transmitted data stream is often undesirable. This dilemma can be avoided by creating a VL without modifying the existing link.

A VL can be created by rotating the received I–Q vector counterclockwise by an angle θ between 0° and 90°. This is equivalent to introducing an artificial phase offset to the received samples, which are then used as outputs from another link. Since the noise is circular symmetric, the rotation results in an equivalent signal power degradation $\cos^2\theta$ without altering the noise statistic. Such a virtual SNR loss cannot be accomplished by simply multiplying the BPSK matched filter output by a positive constant less than 1.

An alternate method is to add an extra zero-mean white Gaussian noise component to the received in-phase samples. Both schemes give a VL with a smaller $\overline{\gamma}$. The second scheme, i.e., the addition of a perturbation term, incurs no hardware increase but requires the estimation of noise power σ_d^2, which is needed in the subsequent ML detection anyway. As the phase rotation scheme leads to an SNR degradation of magnitude $\cos^2\theta$, the second scheme has to generate i.i.d. zero-mean Gaussian random samples with variance $\sigma_v^2 = \sigma_d^2(1/\cos^2\theta - 1)$ to achieve the same SNR loss. Although both approaches achieve the same effect for BPSK signals, the phase-rotating approach cannot produce a VL for noncoherent systems while the method of inserting extra noise suits both coherent and noncoherent applications. Hence, except for the coherent system discussed in this section, we will adopt the noise injection approach in the following sections.

We use the superscript (v) to indicate that a parameter is associated with a VL, i.e., the kth RD link's synchronous output samples and their rotated (VL) versions are denoted by $y_{r_kd}[n]$ and $y_{r_kd}^{(v)}[n]$ and the corresponding ERs by e_{r_kd} and $e_{r_kd}^{(v)}$. For a BPSK system operating in a flat Rayleigh fading environment, we have [14]

$$P_b^{\text{psk}}(\overline{\gamma}) = \frac{1}{2}\left(1 - \sqrt{\frac{\overline{\gamma}}{1+\overline{\gamma}}}\right) \tag{20}$$

which is equivalent to

$$\overline{\gamma} = \frac{(1-2P_b^{\text{psk}})^2}{1-(1-2P_b^{\text{psk}})^2}. \tag{21}$$

The two ERs are then related by

$$\frac{(1-2e_{r_kd})^2}{1-(1-2e_{r_kd})^2} = \frac{1}{\cos^2\theta}\frac{\left(1-2e_{r_kd}^{(v)}\right)^2}{1-\left(1-2e_{r_kd}^{(v)}\right)^2}. \tag{22}$$

Following a procedure similar to that for solving (17), we can easily show that the nonlinear system, which consists of (15), (22), and the new cascaded link's ER equation

$$Q_k^{(v)} = e_{sr_k} + e_{r_kd}^{(v)} - 2e_{sr_k}e_{r_kd}^{(v)} \tag{23}$$

has the closed-form solution

$$e_{sr_k} = \frac{1-\sqrt{1-4t}}{2}, \quad e_{r_k d} = \frac{Q_k - e_{sr_k}}{1-2e_{sr_k}} \tag{24.a}$$

$$e_{r_k d}^{(v)} = \frac{Q_k^{(v)} - e_{sr_k}}{1-2e_{sr_k}} \tag{24b}$$

where

$$t = \frac{\left(1-2Q_k^{(v)}\right)^2 Q_k(1-Q_k)}{\left(1-2Q_k^{(v)}\right)^2 - \cos^2\theta(1-2Q_k)^2} - \frac{\cos^2\theta(1-2Q_k)^2 Q_k^{(v)}\left(1-Q_k^{(v)}\right)}{\left(1-2Q_k^{(v)}\right)^2 - \cos^2\theta(1-2Q_k)^2}. \tag{25}$$

Based on this solution, we can obtain a complete blind algorithm to estimate the ERs of all component links by using the estimates for Q_k and $Q_k^{(v)}$, which are computed via (18) using another, say, lth ($l \neq k$) relay link; the ER side information is no longer needed. In short, to estimate the triplet $(e_{sd}, e_{sr_k}, e_{r_k d})$ associated with an SD and an SRD linking without the help of CSI, one needs another independent relay. The auxiliary relay requirement can be waived if one creates a virtual SD link to obtain additional combinational diversities. In general, the rotation angle for producing a virtual SD link can be different from that for a virtual RD link. However, we lose no generality by assuming both rotation angles are the same, for example, θ. Denote by $\widehat{p}_{(vs)r}$, $\widehat{p}_{s(vr)}$, and $\widehat{p}_{(vs)(vr)}$ the estimates for the SMPs $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd})$, $\mathrm{P}_r(\widehat{y}_{sd} = \widehat{y}_{rd}^{(v)})$, and $\mathrm{P}_r(\widehat{y}_{sd}^{(v)} = \widehat{y}_{rd}^{(v)})$, respectively, and by $Q = e_{srd}$ and $Q^{(v)} = e_{s(vr)d}$ the ERs for the SRD and SR-plus-virtual relay links. We obtain four nonlinear relations for a single-relay CCN, i.e.,

$$\widehat{p}_{sr} = e_{sd}Q + (1-e_{sd})(1-Q) \tag{26.a}$$

$$\widehat{p}_{(vs)r} = e_{sd}^{(v)}Q + \left(1-e_{sd}^{(v)}\right)(1-Q) \tag{26.b}$$

$$\widehat{p}_{s(vr)} = e_{sd}Q^{(v)} + (1-e_{sd}) \quad \left(1-Q^{(v)}\right) \tag{26.c}$$

$$\widehat{p}_{(vs)(vr)} = e_{sd}^{(v)}Q^{(v)} + \left(1-e_{sd}^{(v)}\right)\left(1-Q^{(v)}\right). \tag{26.d}$$

With the additional PL-VL relation

$$\frac{(1-2e_{sd})^2}{1-(1-2e_{sd})^2} = \frac{1}{\cos^2\theta}\frac{\left(1-2e_{sd}^{(v)}\right)^2}{1-\left(1-2e_{sd}^{(v)}\right)^2} \tag{27}$$

the nonlinear system (26.a)–(27) yields the closed-form estimators

$$\widehat{e}_{sd} = \frac{1}{2}\left[\frac{1-\widehat{p}_{sr}-\widehat{Q}}{1-2\widehat{Q}} + \frac{1-\widehat{p}_{s(rv)}-\widehat{Q}^{(v)}}{1-2\widehat{Q}^{(v)}}\right] \tag{28.a}$$

$$\widehat{Q} = \frac{1-\sqrt{1-4t_1}}{2}, \quad \widehat{Q}^{(v)} = \frac{1-\sqrt{1-4t_2}}{2} \tag{28.b}$$

TABLE I
UNIFIED BLIND NOISE-ENHANCED ER ESTIMATION ALGORITHM

Input:	Received samples, $\mathbf{y}$, noise variance, σ_d^2, scaling factor values, $a_{sd}^{(v)}$, $a_{rd}^{(v)}$, $a_{sd}^{(w)}$, $a_{rd}^{(w)}$ and the number of VLs, n_{vl}.
1:	**for** i=1 to n_{vl} **do**
2:	Add noise-enhanced complex zero-mean Gaussian samples with variances $a_{sd}^{(w)}-1$ and/or $a_{r_i d}^{(w)}-1$ to the received SD- and RD-link output samples.
3:	Create virtual SD and RD links by injecting complex Gaussian noise samples with scaling factors, $a_{sd}^{(v)}$ and $a_{r_i d}^{(v)}$.
4:	Compute SMPs for all physical, virtual and/or noise-enhanced SD-SRD link pairs.
5:	**if** modulation type is BPSK **then**
6:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ through (28.b) and (28.a) with $a_{sd}^{(v)} = a_{rd}^{(v)} = \frac{1}{\cos^2\theta}$.
7:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via (24.a)-(25).
8:	**elseif** modulation type is BFSK or DPSK **then**
9:	Compute $\widehat{Q}$, $\widehat{Q}^{(v)}$, and $\widehat{e}_{sd}$ by using (30.a)-(30.c).
10:	Obtain $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ based on (31.a) and (31.b).
11:	**end if**
12:	**end for**
13:	Convert the estimates of the noise-injected ERs back to those of the original (uncontaminated) ERs via (32) or (33).
14	Take averages of all estimates, if available, to obtain the final estimates $\widehat{e}_{sd}$ and $\widehat{e}_{rd}$.
15	Compute $\widehat{e}_{sr}$ via (10) using the estimates obtained in 14 and $\widehat{p}$ derived from the uncontaminated received samples.
Output:	$\widehat{e}_{sd}$, $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$.

$$t_1 = \frac{\cos^2\theta(2\widehat{p}_{sr}-1)^2\left(\widehat{p}_{(vs)r}-1\right)\widehat{p}_{(vs)r}}{\left(2\widehat{p}_{(vs)r}-1\right)^2 - \cos^2\theta(2\widehat{p}_{sr}-1)^2} - \frac{\left(2\widehat{p}_{(vs)r}-1\right)^2(\widehat{p}_{sr}-1)\widehat{p}_{sr}}{\left(2\widehat{p}_{(vs)r}-1\right)^2 - \cos^2\theta(2\widehat{p}_{sr}-1)^2} \tag{28.c}$$

$$t_2 = \frac{\cos^2\theta\left(2\widehat{p}_{s(vr)}-1\right)^2\left(\widehat{p}_{(vs)(vr)}-1\right)\widehat{p}_{(vs)(vr)}}{\left(2\widehat{p}_{(vs)(vr)}-1\right)^2 - \cos^2\theta\left(2\widehat{p}_{s(vr)}-1\right)^2} - \frac{\left(2\widehat{p}_{(vs)(vr)}-1\right)^2\left(\widehat{p}_{s(vr)}-1\right)\widehat{p}_{s(vr)}}{\left(2\widehat{p}_{(vs)(vr)}-1\right)^2 - \cos^2\theta\left(2\widehat{p}_{s(vr)}-1\right)^2}. \tag{28.d}$$

Estimators $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ can be derived from solving the nonlinear system, which includes (15), (22), and an equation similar to (23). An analytic solution of this nonlinear system is obtained by substituting (28.b) into (25) and then (24.a). As has been mentioned in Section I, we refer to ER estimation algorithms using the approach described in this section as VLA estimators. The corresponding estimation procedure is included in Table I.

Note that the SMP formulas (14) and (26.a)–(26.d) are not valid for the SMP between a PL and its virtual version since their outputs are correlated. Actually, this SMP is the sum of two conditional SMPs defined by (40) and (41), which are derived in Appendix D. Obviously, a system involving these two nonlinear expressions does not easily render a closed-form solution. On the other hand, a VL can provide a new SMP relation similar to (14) with each different PL or its virtual version, and a single-relay CCN can offer two uncorrelated VLs to render a basic system that consists of three independent SMP equations; we thus conclude that by using both virtual RD and SD links, one can estimate all ERs of a single-relay CCN without side information.

C. Blind ER Estimation for BFSK and DPSK Signals

Although we have limited our discussion to BPSK signals so far, such a restriction does not lose any generality as far as the VL concept is concerned. The proposed blind estimation method of the last section can easily be extended to noncoherent binary modulations. Besides using (noncooperative) noncoherent detectors, the DN adds a complex Gaussian perturbation term to each of the received noncoherent sample to generate the corresponding VL with the desired equivalent average SNR.

Since, for the noncoherent case, the definition and estimation of SMPs are the same as those of the BPSK-based system, we have four nonlinear equations similar to (26.a)–(26.d) that relate the SMPs to the corresponding ERs of the connecting SD and cascaded SRD links. The relation between the ER of a cascaded link and its two component links remains the same; we thus obtain two equations similar to (15) and (23). However, as a different modulation type is involved, the equation governing the relation between e_{sd}'s for the physical and the VLs is different from (22), so is that between the two e_{rd}'s. The new relation can be expressed in the generic form

$$F_e(z) = a^{(v)} F_e\left(z^{(v)}\right) \tag{29}$$

where $z = e_{sd}$ or e_{rd}, and as before, the superscript (v) on the right-hand side denotes the corresponding set of parameters for VL. Equation (29) is similar to (22), but the actual expression for $F_e(z)$ depends on the modulation used, and $a^{(v)}$ is a scaling parameter related to the variance of the injected noise (normalized with respect to σ_d^2).

Solving the nonlinear system consisting of SMP equations and $F_e(e_{sd}) = a^{(v)} F_e(e_{sd}^{(v)})$, we obtain

$$\widehat{Q} = \frac{(1-2\widehat{p}_{sr})\left[1v-\widehat{p}_{(vs)r}\right]-a^{(v)}\left[1-2\widehat{p}_{(vs)r}\right](1-2\widehat{p}_{sr})}{(1-2\widehat{p}_{sr})-a^{(v)}\left[1-2\widehat{p}_{(vs)r}\right]} \tag{30.a}$$

$$\widehat{Q}^{(v)} = \frac{\left[1-2\widehat{p}_{s(vr)}\right]\left[1-\widehat{p}_{(vs)(vr)}\right]}{\left[1-2\widehat{p}_{s(vr)}\right]-a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]} - \frac{a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]\left[1-2\widehat{p}_{s(vr)}\right]}{\left[1-2\widehat{p}_{s(vr)}\right]-a^{(v)}\left[1-2\widehat{p}_{(vs)(vr)}\right]} \tag{30.b}$$

$$\widehat{e}_{sd} = \frac{1}{2}\left[\frac{1-\widehat{p}_{sr}-\widehat{Q}}{1-2\widehat{Q}} + \frac{1-\widehat{p}_{s(vr)}-\widehat{Q}^{(v)}}{1-2\widehat{Q}^{(v)}}\right]. \tag{30.c}$$

Similarly, we have

$$\widehat{e}_{sr} = \frac{\widehat{Q}^{(v)} - 2\widehat{Q}\widehat{Q}^{(v)} - a^{(v)}\widehat{Q} + 2a^{(v)}\widehat{Q}\widehat{Q}^{(v)}}{1 - 2\widehat{Q} - a^{(v)} + 2a^{(v)}\widehat{Q}^{(v)}} \tag{31.a}$$

$$\widehat{e}_{rd} = \frac{\widehat{Q} - \widehat{e}_{sr}}{1 - 2\widehat{e}_{sr}}. \tag{31.b}$$

TABLE II
REQUIRED CSI AND SOLUTIONS OF NONLINEAR SYSTEMS UNDER VARIOUS MODULATIONS

	$F_e(x)$	Q	$Q^{(v)}$	e_{sd}	e_{sr}	e_{rd}
BPSK	$\frac{(1-2x)^2}{1-(1-2x)^2}$	(28.b)	(28.b)	(28.a)	(24.a)	(24.a)
BFSK	$\frac{1-2x}{x}$	(30.a)	(30.b)	(30.c)	(31.a)	(31.b)
DPSK	$\frac{1-2x}{2x}$	(30.a)	(30.b)	(30.c)	(31.a)	(31.b)

The explicit forms of $F_e(z)$ for different modulations and the corresponding relations used for computing the ER estimators are listed in Table II.

IV. Noise-Enhanced Error Rate Estimations

A. Convergence Consideration and a Simple Variance Reduction Method

It is easy to see that, like the estimator for the SMP p defined in Section II, $\widehat{p}_{sr}$, $\widehat{p}_{(vs)r}$, $\widehat{p}_{s(vr)}$, and $\widehat{p}_{(vs)(vr)}$ converge in probability. As the proposed estimators are continuous functions of these estimates, the continuous mapping theorem [15] implies that the estimators $\{\widehat{e}_{sr}, \widehat{e}_{rd}, \widehat{e}_{sd}\}$ converge in probability as well, and their variances depend on those of the SMP estimators. The latter are all derived from the same compare-and-count process, which is similar to that used in simulation-based ER estimations [16]. The main difference is that, for the latter, the desired detector output is known perfectly, and one has complete information and control of the operating average SNR and the link output statistic. In contrast, our scheme can only rely on blind counting without a pilot sequence, and the link statistic is either unavailable or only partially known. Both estimation methods, however, have the same order of convergence rate and require a large number of samples to obtain a reliable estimate if the true ER is small (see Lemma 1 and [16]).

A straightforward approach to improve the convergence performance is to use multiple VLs, i.e., we add $n_{vl} - 1$ virtual RD and/or SD links with the same noise power. Each VL renders a set of new estimates, and the final estimates are obtained by taking average of the n_{vl} estimates. This method is called the EVLA estimator, which yields a reduced variance for a given sample size or equivalently achieves the same variance as that of the original $(n_{vl} = 1)$ estimator with a smaller sample size.

B. ISI Noise-Enhanced Estimator

To further improve the convergence/variance performance, the aforementioned analogy between our method and the simulation-based estimator suggests that we apply a variance reduction method used in the latter approach called IS. The IS method for estimating ER modifies the demodulator output statistic so that it follows a desired probability distribution that makes the important (error) event occur much more often than the original unmodified case does.

The difficulty in applying the IS theory to our scenario, besides the fundamental differences just mentioned, is due to the fact that the estimators, as was shown in (30.a)–(31.b) and other similar equations presented before, are derived from SMPs and, perhaps, other ERs. Complete control of their statistics through dependent variables whose probability distributions are unknown is impossible. For instance, in the case of a BPSK-based CCN, an SMP depends on the inner product of the SD and RD link outputs whose probability distributions depend on, among other parameters, the true ER of the SR link, which needs to be estimated in the first place. In other words, the optimal (variance-minimizing) importance distribution is a function of the parameters whose values we either do not know or want to estimate.

The following observations, however, indicate that a suboptimal importance distribution is obtainable. First, the ultimate parameters of interest are the link ERs, not the pairwise SMPs, and the IS theory says that convergence is faster if the ER to be estimated by simulation is properly increased, which may be realized by simply adjusting the corresponding link output's variance. Second, some ER estimator formulas are functions of other ERs and SMPs; hence, if the estimates of the other ERs can be improved while those for SMPs remain unchanged, e.g., the ER estimator of e_{sr} through (8), we can obtain an improved estimator. Finally, it is reasonable to assume that the link outputs' statistics are partially known, e.g., their noise variances. But even if we are able to partially control the distributions of related parameters, there still exist the problem of weighting the resulting counts, which is needed in a conventional IS-based procedure and can only be done if both the original and modified link output distributions are known.

Our solution that overcomes all these difficulties proceeds as follows. We first add zero-mean complex Gaussian samples with variances N_{sd} and N_{rd} to the received SD and RD link output samples y_{sd} and y_{rd}, respectively. This results in link outputs with larger variances. By solving the nonlinear system associated with the estimated SMPs of the noise-injected links, we obtain the estimates $\{\widetilde{e}_{sd}^{(w)}, \widetilde{e}_{sr}^{(w)}, \widetilde{e}_{rd}^{(w)}\}$, where the superscript (w) is used to signify the fact that the estimates are computed by inserting artificial noises. As the noise injection effectively reduces the average SNR, the scaling relation (29) with $a^{(v)} = a^{(w)} = 1 + N_{sd}/\sigma_d^2$ or $1 + N_{rd}/\sigma_d^2$ enables us to weight and convert the estimates $\{\widetilde{e}_{sd}^{(w)}, \widetilde{e}_{rd}^{(w)}\}$ back to the estimates $\{\widehat{e}_{sd}^{(w)}, \widehat{e}_{rd}^{(w)}\}$ of the true ERs $\{e_{sd}, e_{rd}\}$. For instance, in a noncoherent BFSK or DPSK-based CCN, the relation $(1 - 2e/e) = a^{(w)}(1 - 2e^{(w)}/e^{(w)})$ for $\max\{e, e^{(w)}\} <$ 1/2 suggests that DN uses the conversion rule

$$\widehat{e}^{(w)} = \widetilde{e}^{(w)} / \left(a^{(w)} + 2\widetilde{e}^{(w)} - 2a^{(w)}\widetilde{e}^{(w)}\right) \tag{32}$$

where the subscripts "sd" and "rd" associated with the estimators $\widehat{e}^{(w)}$ and $\widetilde{e}^{(w)}$ are omitted to simplify the expression. Similarly, the conversion rule for a BPSK-based network is

$$\widehat{e}^{(w)} = \frac{1}{2}\left[1 - \sqrt{\frac{a^{(w)}\left(1 - 2\widetilde{e}^{(w)}\right)^2}{1 - \left(1 - 2\widetilde{e}^{(w)}\right)^2 + a^{(w)}\left(1 - 2\widetilde{e}^{(w)}\right)^2}}\right]. \tag{33}$$

The foregoing two conversion rules bypass the need for complete statistics by directly using the ER conversion based only on $a^{(w)}$, i.e., the ratio between the noise-injected and original SNRs (instead of individual SNRs). They also imply that $\widehat{e}^{(w)} < \widetilde{e}^{(w)}$, which has been expected as we have purposely made $e^{(w)}$ larger by injecting noise. If VLs are needed, we have to inject an additional noise term into the noise-injected PLs to create VLs. Hence, the scaling factor is $a^{(v)}$, $a^{(w)}$, or $a^{(v)}a^{(w)}$, depending on whether the link is a VL, a noise-injected PL, or a noise-injected VL. We call the class of estimators based on the preceding concept as ISI-VLA estimator. In the following sections, we show, via both analysis and simulations, that the ISI-VLA estimator does offer significant performance enhancement.

V. Properties and Performance Analysis of the Noise-Enhanced Estimator

For the preceding approach, noise injection is performed to improve the ER estimators and not the SMP p observed at the DN. In fact, it results in a smaller SMP $p^{(w)}$, and if we want to estimate the original p through $p^{(w)}$, we obtain a worse SMP estimate.

Lemma 3: Let p and $p^{(w)}$ be the true SMPs of the original and noise-injected links, and let $\widehat{p}^{(w)}$ and $\widehat{p}$ be the estimates of p with and without the aid of the noise-injected link. Then

$$\mathrm{Var}[\widehat{p}] \leq \mathrm{Var}\left[\widehat{p}^{(w)}\right]. \tag{34}$$

Proof: See Appendix A. ■

As we can only inject noise into samples received by the DN, e.g., y_{sd} and (or) y_{rd}, e_{sr} remains intact and $\widehat{e}_{sr}^{(w)} = \widetilde{e}_{sr}^{(w)}$ if this estimator is obtained by substituting $\widehat{e}_{rd}^{(w)}, \widehat{e}_{sd}^{(w)}$ and $\widehat{p}^{(w)}$ into (8). The preceding lemma suggests that we should replace $\widehat{p}^{(w)}$ by $\widehat{p}$ in the substitution procedure for estimating e_{sr}. As mentioned in the last section, a better estimate for e_{sr} can thus be obtained by using the noise-enhanced estimates $\widehat{e}_{rd}^{(w)}, \widehat{e}_{sd}^{(w)}$ and the original $\widehat{p}$ [see (8)].

The range of appropriate values for the scaling factor $a^{(w)}$ is certainly dependent on the true ERs e and the noise injected ERs $e^{(w)}$. As will be show in Theorem 1 and numerically in the next section, the MSEE performance is improved by injecting proper noise power into the received samples, and there is an optimal injected noise power that achieves the maximum MSEE improvement. This phenomenon is called the stochastic resonance effect, which has been observed in some nonlinear signal processing systems (see [17] and reference therein).

In a BPSK-based single-relay CCN with perfect SD link ($e_{sd} = 0$), when both the average transmitted relay power P_r and the magnitude of the slow-faded RD link gain $|h_{rd}|$ are known, we show in Appendix B that the optimal scaling factor is approximately equal to the RD link output SNR, i.e.,

$$a_{\mathrm{opt}}^{(w)} \approx \frac{P_r|h_{rd}|^2}{\sigma_d^2}. \tag{35}$$

We need the following lemma to derive a closed-form expression of the optimal scaling factor for the more practical case addressed in Theorem 1.

Lemma 4: For a network that consists of three independent (SD or cascaded) flat Rayleigh fading links with ERs e_i, if the ISI-VLA scheme is applied with a common noise-injected ER $e_i^{(w)} = \epsilon$ using the scaling factors $a_i^{(w)}$, $i = 1, 2, 3$, the variance of the noise enhanced estimator $\widehat{e}_i$ using the conversion rule (32) is given by

$$\mathrm{Var}[\widehat{e}_i] \approx \frac{\left(a_i^{(w)}\right)^2}{\left(a_i^{(w)} + 2\epsilon - 2a_i^{(w)}\epsilon\right)^4} \frac{\epsilon - 2\epsilon^2 + 2\epsilon^3 - \epsilon^4}{(2\epsilon - 1)^2 N}. \quad (36)$$

Proof: See Appendix C. ■

In the subsequent discourse, we denote by $\widehat{y}_i$, e_i, and SNR_i the hard decision output, ER, and average SNR of the ith link (direct or cascaded) and by $a_i^{(v)}$ and $a_i^{(w)}$ the associated scaling factor used. To characterize the stochastic resonance effect and the noise enhanced performance, we define the MSEE reduction ratio $\gamma \stackrel{def}{=} \mathrm{MSEE}_{\mathrm{ISI}}/\mathrm{MSEE}_{\mathrm{o}}$, where $\mathrm{MSEE}_{\mathrm{ISI}}$ and $\mathrm{MSEE}_{\mathrm{o}}$ are the MSEEs of the ISI-VLA and VLA estimators with the same sample size. Using the preceding lemma, we obtain the following theorem.

Theorem 1: For a network with three independent flat Rayleigh fading links, the optimal scaling factor under the common noise-injected ER constraint $e_i^{(w)} = \epsilon$, $i = 1, 2, 3$ is approximately equal to

$$a_{i,opt}^{(w)} \approx t_1 \mathrm{SNR}_i \quad (37)$$

where $t_1 = 0.3085$ (DPSK) or 0.15428 (BFSK). The minimum achievable MSEE reduction ratio γ_{min} for $\mathrm{SNR}_i \gg 1$ is given by

$$\gamma_{min} \approx \begin{cases} 9.8277 \frac{\mathrm{SNR}_i^2}{(1+\mathrm{SNR}_i)^3}, & \text{DPSK} \\ 19.655 \frac{\mathrm{SNR}_i^2}{(2+\mathrm{SNR}_i)^3}, & \text{BFSK}. \end{cases} \quad (38)$$

Moreover, noise injection using the optimal scaling factor is beneficial if SNR_i is larger than 3.241 (DPSK) or 6.483 (BFSK).

Proof: See Appendix D. ■

Following a procedure similar to that used in proving Lemma 4 and Theorem 1 and using the relation governing the ER ϵ of a noise-injected BPSK link and the associated scaling factor $\epsilon = (1/2)(1 - \sqrt{(\mathrm{SNR}_i/a_i^{(w)} + \mathrm{SNR}_i)})$, we can prove the following theorem.

Theorem 2: For a three-link BPSK-based network in a flat Rayleigh fading environment, the optimal scaling factors that ensure a common noise-injected ER is $a_i = t_i \mathrm{SNR}_i$, and the MSEE reduction ratio γ for link i is

$$\gamma|_{a_i = t_1 \mathrm{SNR}_i} = \frac{24.68 \mathrm{SNR}_i^2}{3 + 10\mathrm{SNR}_i + 11\mathrm{SNR}_i^2 + 4\mathrm{SNR}_i^3} \quad (39)$$

where $t_1 = (-1 + \sqrt{7}/3)$. Noise injection using the optimal scaling factor is beneficial if $\mathrm{SNR}_i > 1.823$.

To evaluate the MSEE performance of VLA and ISI-VLA estimators in a CCN, as shown in Appendix C, we need to compute the covariance and matrix $\mathbf{C}$ of the pairwise matching indicators $I(\hat{y}_k[t] = \hat{y}_j[t])$ and the associated Jacobian matrix $\mathbf{J}$. The entries of these two matrices are functions of the (not necessarily pairwise) SMPs, whose expressions are given below.[1]

Lemma 5: For a two-link BPSK-based network, the SMPs $p_{12(v1)}$ that direct PLs 1 and 2 and $a_1^{(v)}$-scaled VL 1 (denoted by $v1$) yield the same hard decision given by

$$p_{12(v1)} = e_2 p_{em}\left(e_1, a_1^{(v)}\right) + (1 - e_2) p_{cm}\left(e_1, a_1^{(v)}\right) \quad (40)$$

where the conditional erroneous matching probability $p_{em}(e_1, a_1^{(v)}) \stackrel{def}{=} \Pr(\widehat{y}_1 = \widehat{y}_1^{(v)} = -s|s)$ and the conditional correct matching probability $p_{cm}(e_1, a_1^{(v)}) \stackrel{def}{=} \Pr(\widehat{y}_1 = \widehat{y}_1^{(v)} = s|s)$, with $s = \pm 1$ being the normalized transmitted BPSK signal, are

$$p_{em}\left(e_1, a_1^{(v)}\right) = \frac{e_1}{2} + \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)} - 1}}\right) - \frac{\tan^{-1}\left(\frac{\left(1-2e_1^{(v)}\right)^{-1}}{\sqrt{a_1^{(v)}-1}}\right)}{\left(1 - 2e_1^{(v)}\right)^{-1}}\right] \quad (41)$$

$$p_{cm}\left(e_1, a_1^{(v)}\right) = 1 - e_1 - e_1^{(v)} + p_{em}\left(e_1, a_1^{(v)}\right). \quad (42)$$

If PL 1 is a cascaded link, the SMP becomes

$$p_{12(v1)} = e_2\left[p_{em}\left(e_1, a_1^{(v)}\right)(1 - e_{sr}) + p_{cm}\left(e_1, a_1^{(v)}\right) e_{sr}\right] + (1-e_2)\left[p_{em}\left(e_1, a_1^{(v)}\right) e_{sr} + p_{cm}\left(e_1, a_1^{(v)}\right)(1 - e_{sr})\right] \quad (43)$$

where e_{sr} is the ER of the hidden component link of PL 1. The SMPs that direct PL 1, cascaded PL 2, and $a_i^{(v)}$-scaled VLs 1 and 2 all yield the same hard decision given by

$$\begin{aligned} p_{12(v1)(v2)} =& p_{em}\left(e_1, a_1^{(v)}\right) \\ &\times \left[p_{em}\left(e_2, a_2^{(v)}\right)(1 - e_{sr}) + p_{cm}\left(e_2, a_2^{(v)}\right) e_{sr}\right] \\ &+ p_{cm}\left(e_1, a_1^{(v)}\right) \\ &\times \left[p_{em}\left(e_2, a_2^{(v)}\right) e_{sr} + p_{cm}\left(e_2, a_2^{(v)}\right)(1 - e_{sr})\right]. \end{aligned} \quad (44)$$

Finally, we have the two joint pairwise SMPs

$$\begin{aligned} \Pr&\left(\widehat{y}_1 = \widehat{y}_2, \quad \widehat{y}_1^{(v)} = \widehat{y}_2^{(v)}\right) \\ &= p_{12(v1)(v2)} + \left[e_1 - p_{em}\left(e_1, a_1^{(v)}\right)\right] \\ &\quad \times \left\{(1 - e_{sr})\left[e_2 - p_{em}\left(e_2, a_2^{(v)}\right)\right] + e_{sr}\left[e_2^{(v)} - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right\} \\ &\quad + \left[e_1^{(v)} - p_{em}\left(e_1, a_1^{(v)}\right)\right] \\ &\quad \times \left\{(1 - e_{sr})\left[e_2^{(v)} - p_{em}\left(e_2, a_2^{(v)}\right)\right] + e_{sr}\left[e_2 - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right\} \end{aligned} \quad (45)$$

[1] Matching probabilities and variance analysis for noise-enhanced estimators are similar. Depending on where the noise-injected links are located, the resulting expressions are obtained by replacing e_i and (or) $e_i^{(v)}$ by $e_i^{(w)}$ and (or) its VL version; the scaling factors are also modified when necessary. This apply to *Lemmas 5* and *6* as well as *Theorems 3* and *4*.

$$\begin{aligned}
\Pr\left(\widehat{y}_1 = \widehat{y}_2, \quad \widehat{y}_1^{(v)} = \widehat{y}_2^{(v)}\right) &\\
= p_{12(v1)(v2)} &\\
+ \left[e_1 - p_{em}\left(e_1, a_1^{(v)}\right)\right] &\\
\times \left\{(1-e_{sr})\left[e_2 - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right. &\\
\left. + e_{sr}\left[e_2 - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right\} &\\
+ \left[e_1^{(v)} - p_{em}\left(e_1, a_1^{(v)}\right)\right] &\\
\times \left\{(1-e_{sr})\left[e_2^{(v)} - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right. &\\
\left. + e_{sr}\left[e_2^{(v)} - p_{em}\left(e_2, a_2^{(v)}\right)\right]\right\}. &\quad (46)
\end{aligned}$$

Proof: See Appendix E. ∎

With the foregoing formulas and the pairwise SMP given by (B.2), we use a procedure similar to that presented in Appendix C to evaluate the covariance matrix of the ER estimators and obtain the following theorem.

Theorem 3: For a two-link BPSK-based network using a virtual R_1D link, the variances of the VLA estimators $\widehat{e}_1$ and $\widehat{e}_2$ are given by

$$\mathrm{Var}\left[\widehat{e}_1 | e_1, e_2, a_1^{(v)}\right] = \frac{b_2^2 p_{12}(1-p_{12}) - 2b_2b_3 + p_{(v1)2}\left(1-p_{(v1)2}\right)}{(1-2e_2)^2(b_2-b_1)^2N} \quad (47)$$

$$\mathrm{Var}\left[\widehat{e}_2 | e_1, e_2, a_1^{(v)}\right] = \frac{b_1^2 p_{12}(1-p_{12}) - 2b_1b_3 + p_{(v1)2}(1-p_{(v1)2})}{(1-2e_1)^2(b_2-b_1)^2N} \quad (48)$$

where p_{12} and $p_{(v1)2}$ are the SMPs for the link pairs (1, 2), and $(v1, 2)$, $b_1 = (a_1^{(v)}/[a_1^{(v)} - (a_1^{(v)} - 1)(1-2e_1)^2]^{3/2})$, $b_2 = (1/[a_1^{(v)} - (a_1^{(v)} - 1)(1-2e_1)^2]^{1/2})$, and $b_3 = p_{12(v1)} - p_{12}p_{2(v1)}$. Moreover, if noise of power $(a_i^{(w)} - 1)\sigma_d^2$ is injected, then the variance of the noise enhanced ISI-VLA (EISI-VLA) estimator $\widehat{e}_i^{(w)}$ is given by

$$\frac{a_i^{(w)}}{\left[1 + (a_i^{(w)} - 1)(1-2e_i^{(w)})^2\right]^3} \mathrm{Var}\left[\widehat{e}_i^{(w)} \,\middle|\, e_1^{(w)}, e_2^{(w)}, a_1^{(v)}\right]. \quad (49)$$

If a virtual R_2D link is used instead, then (47)–(49) should be modified by replacing $a_1^{(v)}$, e_1, $p_{(v1)2}$, and $p_{12(v1)}$ with $a_2^{(v)}$, e_2, $p_{1(v2)}$, and $p_{12(v2)}$, respectively.

Note that the notations used in (47) and (48) imply that the variance of $\widehat{e}_i$ is a function of e_1, e_2, and $a_1^{(v)}$ only. All the other parameters, e.g., b_i's, depend on these three parameters. For the case addressed in Theorem 3, the optimal scaling factors can be obtained by finding the extreme points of (49), a highly nonlinear function of $a_1^{(w)}$ and $a_2^{(w)}$.

The performance analysis of an ISI-VLA estimator for the hidden SR link is more involved. We need the following preliminary result.

Lemma 6: For a single-relay CCN with single virtual SD and RD link, the (i, j)th entry of the covariance matrix $\mathbf{C}$ of the indicator vector $[I(\widehat{y}_1 = \widehat{y}_2)I(\widehat{y}_1^{(v)} = \widehat{y}_2)I(\widehat{y}_1 = \widehat{y}_2^{(v)})I(\widehat{y}_1^{(v)} = \widehat{y}_2^{(v)})]^T$ is given by

$$\mathbf{C}_{ij} = \begin{cases} p_{kl}(1-p_{kl}), & \text{if } k = l, \quad k' = l' \\ \Pr(\widehat{y}_k = \widehat{y}_l, \widehat{y}_{k'} = \widehat{y}_{l'}) - p_{kl}p_{k'l'}, & \text{otherwise} \end{cases} \quad (50)$$

for $i, j = 1, \ldots, 4$, with the mapping $i \to (k, l)$ defined by

$$k = \begin{cases} 1, & \text{if } i \text{ is odd} \\ (v1), & \text{otherwise,} \end{cases} \qquad l = \begin{cases} 2, & i \le 2 \\ (v2), & i \ge 3 \end{cases} \quad (51)$$

and a similar mapping from j to (k', l'). The corresponding inverse Jacobian $\mathbf{J}^{-1}$ is given in (52), shown at the bottom of the page, where $h'(x, a) = (a^{(v)}/[a^{(v)} + (1 - a^{(v)})(1 - 2x)^2]^{3/2})$.

We immediately have the following theorem.

Theorem 4: For a single-relay BPSK-based CCN with $a_{sd}^{(v)}$-scaled virtual SD link and $a_{rd}^{(v)}$-scaled virtual RD link, as described by (26.a)–(27), the variances for the VLA estimators $\widehat{e}_{sr}$, $\widehat{e}_{rd}$, and $\widehat{e}_{sd}$ are given by

$$\mathrm{Var}[\widehat{e}_{sr}] = \frac{\widetilde{\mathbf{C}}_{22}}{N}, \quad \mathrm{Var}[\widehat{e}_{rd}] = \frac{\widetilde{\mathbf{C}}_{33}}{N} \quad (53)$$

$$\mathrm{Var}[\widehat{e}_{sd}] = \frac{\widetilde{\mathbf{C}}_{11} + \widetilde{\mathbf{C}}_{14} + \widetilde{\mathbf{C}}_{41} + \widetilde{\mathbf{C}}_{44}}{4N} \quad (54)$$

where $\widetilde{\mathbf{C}} = \mathbf{JCJ}^T$, and $\widetilde{\mathbf{C}}_{ij}$ denotes the element in the ith row and jth column of $\mathbf{C}$.

Furthermore, the variance of the ISI-VLA estimators $\widehat{e}_{sr}^{(w)}$, $\widehat{e}_{rd}^{(w)}$, and $\widehat{e}_{sd}^{(w)}$ are

$$\mathrm{Var}\left[\widehat{e}_{sr}^{(w)}\right] = \frac{\widetilde{\mathbf{C}}_{22}^{(w)}}{N} \quad (55)$$

$$\mathrm{Var}\left[\widehat{e}_{rd}^{(w)}\right] = \frac{a_{rd}^{(w)}}{\left[1 + \left(a_{rd}^{(w)} - 1\right)\left(1 - 2e_{rd}^{(w)}\right)^2\right]^3} \frac{\widetilde{\mathbf{C}}_{33}}{N} \quad (56)$$

$$\begin{pmatrix} 2e_2 - 1 & (2e_1-1)(1-2e_2) & (2e_1-1)(1-2e_{sr}) & 0 \\ (2e_2-1)h'\left(e_1, a_1^{(v)}\right) & \left(2e_1^{(v)} - 1\right)(1-2e_2) & \left(2e_1^{(v)} - 1\right)(1-2e_{sr}) & 0 \\ 0 & (2e_3-1)\left(1-2e_2^{(v)}\right) & (2e_3-1)(1-2e_{sr})h'\left(e_2, a_2^{(v)}\right) & 2e_2^{(v)} - 1 \\ 0 & \left(2e_3^{(v)} - 1\right)\left(1-2e_2^{(v)}\right) & \left(2e_3^{(v)} - 1\right)(1-2e_{sr})h'\left(e_2, a_2^{(v)}\right) & \left(2e_2^{(v)} - 1\right)h'\left(e_1, a_1^{(v)}\right) \end{pmatrix} \quad (52)$$

$$\mathrm{Var}\left[\widehat{e}_{sd}^{(w)}\right] = \frac{a_{sd}^{(w)}}{\left[1+\left(a_{sd}^{(w)}-1\right)\left(1-2e_{sd}^{(w)}\right)^2\right]^3} \times \frac{\widetilde{\mathbf{C}}_{11}^{(w)}+\widetilde{\mathbf{C}}_{14}^{(w)}+\widetilde{\mathbf{C}}_{41}^{(w)}+\widetilde{\mathbf{C}}_{44}^{(w)}}{4N} \tag{57}$$

where $\widetilde{\mathbf{C}}^{(w)} = [\widetilde{\mathbf{C}}_{ij}^{(w)}] = \mathbf{J}^{(w)}\mathbf{C}^{(w)}(\mathbf{J}^{(w)})^T$, and $\mathbf{J}^{(w)}$ and $\mathbf{C}^{(w)}$ are computed after noise injection into all but the SR link.

We summarize a few remarks regarding the preceding properties, their extensions, and the proposed noise-enhanced estimator in general in the following.

R1: The noise samples play the dual role of a) generating VLs to eliminate the needs for CSI and extra RNs and resolve the symmetric ambiguity and b) altering the statistical property of the received samples.

R2: As the identity, (14), which relates an SMP to the associated ERs, involves two independent links; the three-link network has the special property of offering $\binom{3}{2} = 3$ link pairs such that each link participates in two link-pairs. Such a "uniform participation" is important to guarantee uniform performance, i.e., the MSEE performance for each link is the same if the true ERs are identical. In general, for a network with four or more links, the number of link pairs is larger than the number of independent links, and the performance of an ER estimator for a particular link depends on the number of link pairs it has participated.

R3: Although Theorems 1 and 2 consider a three-link network only, extensions to networks with more independent component links are straightforward, but closed-form expressions for the corresponding optimal scaling factor and noise benefit interval (NBI) can only be determined numerically. Nevertheless, for the special cases considered by both theorems, the minimum achievable MSEE reduction ratio tends to $O(1/\mathrm{SNR}_i)$ at high SNRs.

R4: Theorems 3 and 4 give the MSEE expressions for BPSK-based VLA and ISI-VLA estimators, but we are not able to derive closed-form expressions for the noncoherent modulation-based networks. The optimal injected noise power levels for noncoherent networks with correlated links seem to be mathematically intractable. However, our analysis indicates that a key factor in the MSEE expression is the square of the first derivative of the conversion function (rule) with respect to the scaling factor, which is in the order of $(a_i^{(w)})^{-2}$ for small ERs [see, e.g., (36)]. The increase of $a_i^{(w)}$ reduces this factor's value, but it also impacts the other parameters that might increase the MSEE. For examples, in (36), $a_i^{(w)}$ is fixed by the identical $e_i^{(w)} = \epsilon$ constraint and is not a independent parameter, whereas in (49), $a_i^{(w)}$ affects every parameter on the second rational term. The optimal $a_i^{(w)}$ strikes the best balance between these conflicting effects. Numerical experiments reported in the next section show that, similar to the special cases addressed in Theorems 1 and 2, there is a proper range of injected noise power levels for enhancing the performance with added noise, and an optimal scaling factor (added noise power level) does exist.

R5: Similar to the EVLA scheme, we can add $n_{vl} - 1$ virtual RD and/or SD links to obtain the same number of estimates for $\{\widehat{e}_{sd}\}$ and/or $\{\widehat{e}_{rd}\}$, each with the same reduced variance, and then take the average on the resulting n_{vl} estimators. This sample mean approach guarantees improved performance, but the improvement ratio is bounded by $1/n_{vl}$ due to the correlations among VLs. The resulting multiple-VL algorithm is called the EISI-VLA estimator.

VI. Numerical Results

For convenience of reference, we refer to the ML detector using the ER estimators presented in Section II as the physical-link-only (PLO) detector and that using a VLA estimator as the VLA detector. The ML detector with perfect CSI is called the ideal detector. Let d_{sr_k}, d_{r_kd}, and d_{sd} be the distances of the kth SR, RD, and SD link, and let θ_{sr_k} be the angle between the SD and kth RD links (see Fig. 1). Without loss of generality, we use the normalization $d_{sd} = 10$ so that

$$\begin{aligned} d_{sr_k}^2 &= d_{r_kd}^2 + d_{sd}^2 - 2d_{r_kd}d_{sd}\cos\theta_{sr_k} \\ &= 100 + d_{r_kd}^2 - 20d_{r_kd}\cos\theta_{sr_k}. \end{aligned} \tag{58}$$

We assume the path loss model $\sigma_{ij}^2 \propto d_{ij}^{-\alpha}$ with normalization $\sigma_{sd}^2 = 1$ and $\alpha > 0$. Denote by σ_{ij}^2 the variance of the Rayleigh faded link gain and d_{ij} the distance between nodes i and j, i.e., $(i,j) \in \{(s,r_k),(r_k,d), k=1,\ldots,L\}$. All the simulated performance curves are obtained by sequentially applying the proposed methods, i.e., the estimated ERs are updated sequentially as each new sample becomes available, and the updated estimates are then used for detecting each received bit. As in [7], we define the SH average SNR as the average received SNR for the direct SD link without relaying $\bar{\gamma}_{sd}$. Simulation for a given $\bar{\gamma}_{sd}$ terminates whenever the number of error events in the detector output exceeds 500. We assume that the noise powers at DN and RN are the same, i.e., $\sigma_d^2 = \sigma_r^2$, and use the normalization $P = P_s + \sum_{i=1}^{L} P_{r_i} = 1$ such that $\bar{\gamma}_{sd} = 1/\sigma_d^2$. To reduce the complexity of the ML detector, [6] suggested a piecewise linear function to approximate the nonlinearity (3). As it causes negligible performance degradation with respect to that of the ML detector so long as $e_{sr_k} < (1/2)$, we use the same approximation in our simulation efforts.

The performance of the PLO and VLA detectors for the simplest case, $L = 1$ with BPSK modulation, is illustrated in Fig. 2. For the PLO detector, only e_{sr} is unknown, whereas the VLA detector assumes that the ERs of the other component links are also unavailable and uses a rotation angle $\theta = 45^\circ$, which is equivalent to injecting noise with $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$. The performances of both detectors are found to approach that of the ideal ML detector. We also investigate the effect of correlated fading on the performance of the VLA detector for DPSK signals, and the result is shown in the same figure. Modified Jake's model [18] with normalized Doppler frequency $J = f_dT_s = 0.001$, with f_d and T_s being the Doppler frequency and the sampling period, respectively, is used to generate

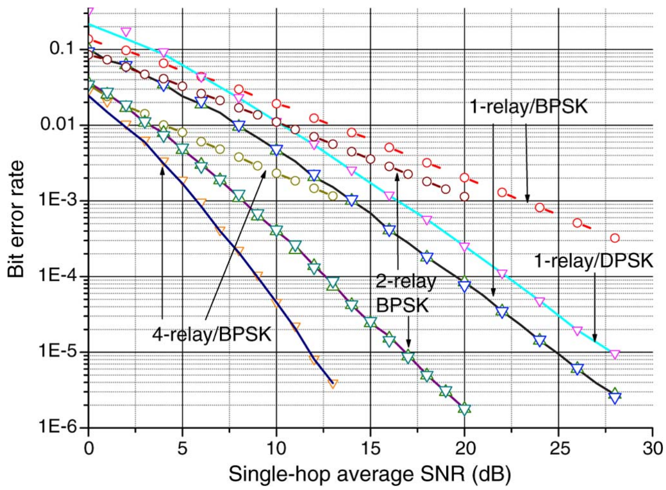


Fig. 2. BER performance of the ML (solid curves), MRC (○), PLO (△), and VLA (∇) detectors. The following system parameter values are used. (i) single-relay system: $P_s = P_r = 0.5P$, $d_{sr}/d_{sd} = 0.8$, $\theta_{sr} = 0°$, and $a^v_{sd} = a^v_{rd} = 2$ (single-relay); (ii) two-relay system: $d_{r_1d} = d_{sr_2} = 7/10$, $\theta_{sr_1} = \theta_{sr_2} = 0°$, $P_s = 0.5P$, $P_{r_1} = P_{r_2} = 0.25P$, and $\theta = 45°$; and (iii) four-relay system: $d_{sd} = 10$, $d_{sr_1} = 5$, $\theta_{sr_1} = 45°$, $d_{sr_2} = 6$, $\theta_{sr_2} = 30°$, $d_{sr_3} = 4$, $\theta_{sr_3} = 60°$, $d_{sr_4} = 5$, $\theta_{sr_1} = 0°$, $P_p = 0.5P$, $P_{r_i} = 0.125P$, for $i = 1, 2, 3, 4$, and $\theta = 30°$.

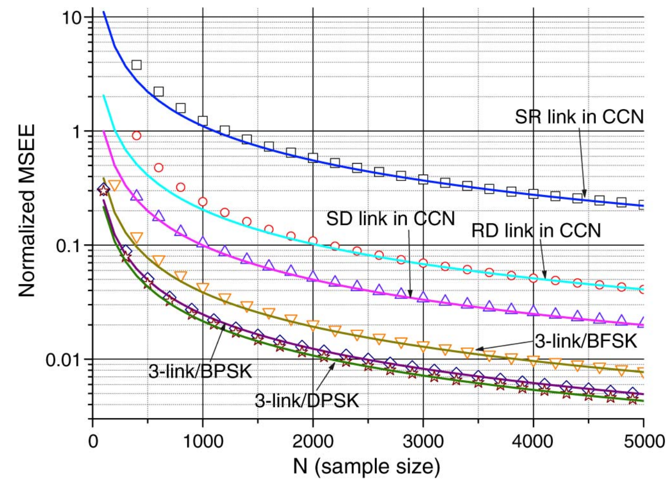


Fig. 3. Normalized MSEE performance of the ISI-VLA scheme for (a) various binary modulated 3-link networks ($e_1 = 0.003$, $e_2 = 0.002$, $e_3 = 0.001$; the injected noise power is such that SH SNR $= 2$ for link 1 and $e_1^{(w)} = e_2^{(w)} = e_3^{(w)}$) and (2) BPSK-based single-relay CCN ($e_{sr} = 0.02922$, $e_{rd} = 0.001988$, $e_{sd} = 0.04356$, $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $a_{sd}^{(w)} = 1$ and $a_{rd}^{(w)} = 30$). For 3-link networks, only the performance of $\widehat{e}_1$ is shown. The analytic predictions (solid curves) for these two scenarios are based on (36) and (55)–(57), respectively.

the component link gains $\{h_{sd}[n]\}$, $\{h_{sr}[n]\}$, and $\{h_{rd}[n]\}$ as a function of sampling epochs. For the DPSK system, we use the noise-injected VLA detector with scaling factors $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$ [see (29)]. Obviously, the performance of the VLA detector is almost the same as that of the ML detector within the range of interest, indicating that the i.i.d. assumption gives accurate ER estimates for moderately correlated fading environments.

Fig. 2 also shows the performance for the cases of two and four RNs. In the two-relay case, we assume that the PLO detector knows e_{r_kd} perfectly. Again, both PLO and VLA detectors yield performance almost identical to that of the ML detector. For the four-relay case, we decompose the problem into four single-relay CCN subproblems, each involving only one SRD and the SD links. It can be seen that at the low SH-SNR region (0–2 dB), the performance of the VLA detector is slightly worse than that of the optimal detector. This is due to fact that the sample size used is not large enough to offer a very reliable BER estimate. Nevertheless, its performance is still superior to that of the MRC detector.

To verify our MSEE analysis, we consider a three-link wireless sensor network in Fig. 3, which shows that, for all three binary modulations considered, the analytic predictions are very close to those obtained by simulations even when the sample size is small, and both give identical results if the sample size is large. A similar performance trend for the ISI-VLA scheme in a BPSK-based single-relay CCN is found in the same figure. The normalized MSEE performance $E[(\widehat{e} - e)^2]/e^2$, where e is the true ER, of the VLA, VLA-EM, and EISI-VLA estimation schemes for a BFSK-based single-relay CCN network is shown in Fig. 4. The VLA-EM scheme refers to a modified version of the EM-based estimator of [11] that did not consider the hidden SR link. The modifications are needed to apply a VL to resolve the ambiguity and replace the normalization factor such that

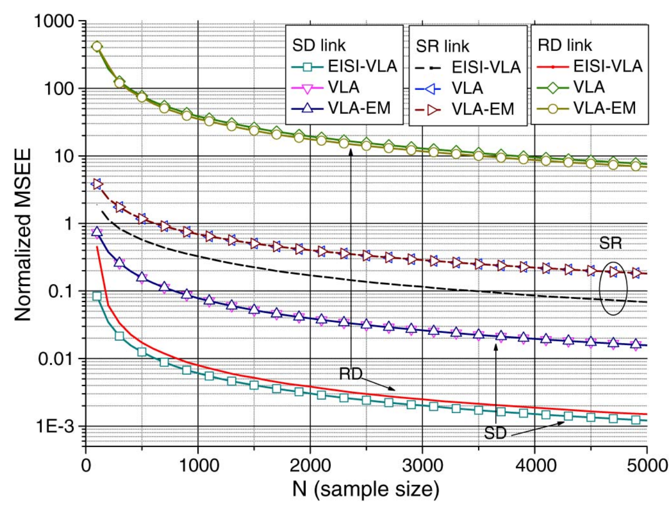


Fig. 4. Normalized MSEE performance of VLA, VLA-EM, and EISI-VLA schemes in a BFSK-based single-relay CCN with $e_{sr} = 0.0127$, $e_{rd} = 5.0711 \times 10^{-5}$, and $e_{sd} = 0.0298$. The other parameter values used are $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$, $e_{sd}^{(w)} = e_{rd}^{(w)} = 0.05$, and $n_{vl} = 30$.

the equation for updating the ER estimate for the cascaded link becomes

$$Q_k^{(i+1)} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y_j}=\widehat{y_k})}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y_j}=\widehat{y_k})}}{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y_j}=\widehat{y_k})}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y_j}=\widehat{y_k})}} + \frac{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{1-I(\widehat{y_j}=\widehat{y_k})}\left(1-Q_j^{(i)}\right)^{I(\widehat{y_j}=\widehat{y_k})}}{\prod_{j=0}^{L}\left(Q_j^{(i)}\right)^{I(\widehat{y_j}=\widehat{y_k})}\left(1-Q_j^{(i)}\right)^{1-I(\widehat{y_j}=\widehat{y_k})}}\right)^{-1} \quad (59)$$

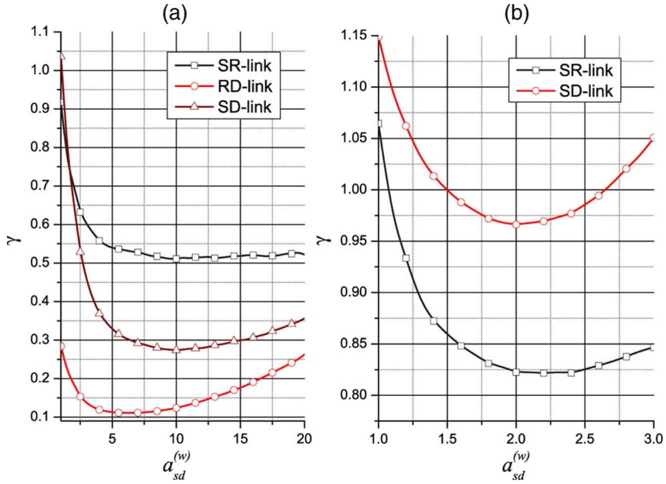


Fig. 5. MSEE reduction ratio (γ) performance of the ISI-VLA estimator with BFSK modulation and $a_{sd}^{(v)} = a_{rd}^{(v)} = 2$. (a) Obtained by assuming $d_{sr} = 5$, SH-SNR $= 25$ dB with the path loss exponent $= 2$ (which leads to $e_{sr} = 0.0016$, $e_{rd} = 0.0016$, $e_{sd} = 0.0062$). (b) Assumes that $d_{sr} = 8$, SH-SNR $= 18$ dB with path loss exponent $= 4$ so that $e_{sr} = 0.0127$, $e_{rd} = 5.0711 \times 10^{-5}$, and $e_{sd} = 0.0298$. The MSEE reduction ratio of the RD link is not shown in part (b) as it is relatively small ($\sim O(10^{-3})$).

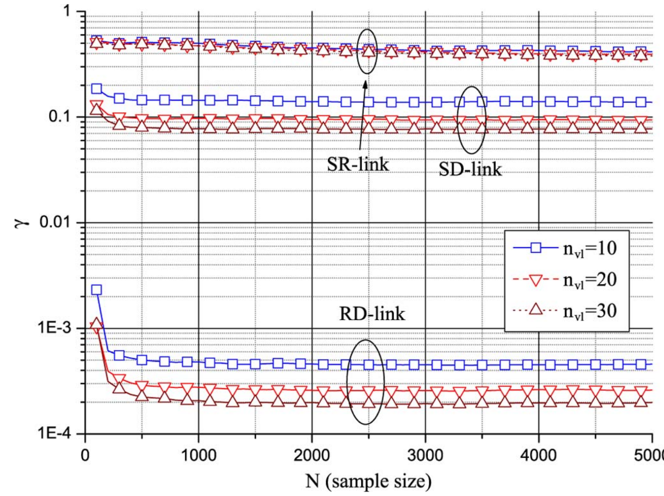


Fig. 6. MSEE reduction ratio behavior of the EISI-VLA estimator for BFSK-based CCN with different n_{vl}. Other system parameter values are the same as those of Fig. 5(b).

where Q_i's are defined in Section II-C with the superscripts denoting the associated iteration number. The ISI method injects additional noise to estimate the ERs of the resulting links and then converting them back to $\widehat{e}_{sr}$ and $\widehat{e}_{rd}$ via the analytic formulas given in Table II. The performance curves clearly demonstrate that the advantage of the VLA-EM scheme against the VLA estimator is negligible, whereas the EISI-VLA scheme far outperforms the other two schemes.

Fig. 5 plots the MSEE reduction ratio as a function of the scaling factor $a_{sd}^{(w)}$, whereas the other scaling factor $a_{rd}^{(w)}$ is chosen such that $e_{rd}^{(w)} = e_{sd}^{(w)}$. These curves reveal that the MSEE performance is improved by injecting proper noise power into the received samples, and there is an optimal injected noise power that achieves the maximum MSEE improvement. This phenomenon is called the stochastic resonance effect, which has been observed in some nonlinear systems (see [17] and references therein). We also notice that the improvement is more impressive when the true ER becomes smaller, which is consistent with what the IS theory has predicted. The NBI, defined as the range of scaling factor values within which the MSEE reduction ratio is less than 1, is a function of the true e_{sd} and e_{rd}. As mentioned before, we are not able to derive closed-form expressions for the optimal scaling factors used in a noncoherent network. Nevertheless, extensive simulations suggest that it is a good strategy to make $e_{sd}^{(w)} = e_{rd}^{(w)} \approx 0.05$ if both e_{sd} and e_{rd} are much smaller than 0.05. As was explained in Sections V and VI, because of the availability of improved estimates for e_{sd} and e_{rd}, the performance of $\widehat{e}_{sr}$ is also improved, although we do not and could not inject noise into samples received at RNs.

Although proper noise injection does improve the convergence rate performance, in some cases such as those shown in Fig. 5, the improvement is not quite as significant as one wishes. The MSEE reduction ratio can be further improved by the EISI-VLA estimator, as shown in Fig. 6, where the simulation conditions are identical to those assumed in Fig. 5(b). As expected, the performance is improved with the increase of n_{vl}, and the improvement is much more impressive when the true ER is small: the required sample size reduction is more than ten times for the SD link and is greater than 8000 times for the RD link when $n_{vl} = 30$. Another benefit of using multiple VLs is that the NBI becomes larger as n_{vl} increases.

VII. Conclusion

In this paper, we have proposed noise-enhanced blind ER estimators for binary modulation-based wireless relay networks. Noise enhancement manifests itself in three aspects. First, noise is added to the received samples to create VLs to remove the CSI requirement and to resolve the ambiguity associated with an underdetermined system and that due to the symmetric nature of a cascaded link. Second, multiple noise-injected VLs are used to reduce the estimation variance and the number of relays needed for estimating ERs. Third, inspired by the IS theory used in computer-simulation-based ER estimation, noise with proper power is inserted to improve the ER estimator's convergence performance. The MSEE performance of some special networks is analyzed, and both analysis and simulations show that the ISI estimator exhibits the so-called stochastic resonance phenomenon that amounts to the effect that injecting noise with a proper power helps improve an estimator's performance, and there exists an optimal injected noise power that offers the best MSEE improvement. Numerical results indicate that the performance of the ML detector using our estimators is very close to that of the ideal ML detector, which knows the SR link's ER perfectly. Moreover, the Monte-Carlo-based ISI approach is capable of bringing about several orders of MSEE reduction.

Appendix A
Proof of Lemma 3

Letting $\widetilde{p}^{(w)}$ be the average count-based estimate of $p^{(w)}$, i.e., the SMP of the noise-injected SD and RD link outputs, we have, from (8), the conversion rule

$$\widehat{p}^{(w)} = \mathcal{D}_o + \frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\widetilde{p}^{(w)} \tag{A.1}$$

where

$$\mathcal{D}_o = \frac{(1-e_{sd}-e_{rd}+2e_{sd}e_{rd})\left(1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}\right)}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}} - \frac{(1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd})\left(1-e_{sd}^{(w)}-e_{rd}^{(w)}+2e_{sd}^{(w)}e_{rd}^{(w)}\right)}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}. \tag{A.2}$$

As $\widehat{p}^{(w)}$ is a linear function of $\widetilde{p}^{(w)}$, the ML estimate of $p^{(w)}$, it is an ML estimator of p. Furthermore, $\widetilde{p}^{(w)}$ is a sample mean estimator; its variance is equal to $\text{var}[\widetilde{p}^{(w)}] = (p^{(w)}(1-p^{(w)})/N)$. Similarly, the variances of $\widehat{p}^{(w)}$ and $\widehat{p}$ are, respectively, given by

$$\text{Var}\left[\widehat{p}^{(w)}\right] = \frac{p^{(w)}\left(1-p^{(w)}\right)}{N}\left(\frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\right)^2$$

$$\text{Var}[\widehat{p}] = \frac{p(1-p)}{N}.$$

Invoking the inequalities $0 \le p < p^{(w)} \le 0.5$ or $1 \ge p > p^{(w)} \ge 0.5$, $e_{sd} \le e_{sd}^{(w)}$, and $e_{rd} \le e_{rd}^{(w)}$, we have $p(1-p) \le p^{(w)}(1-p^{(w)})$ and $((1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}/1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}))^2 \ge 1$. Hence

$$\begin{aligned}\text{Var}[\widehat{p}] &= \frac{p(1-p)}{N} \\ &\le \left(\frac{1-2e_{sd}-2e_{rd}+4e_{sd}e_{rd}}{1-2e_{sd}^{(w)}-2e_{rd}^{(w)}+4e_{sd}^{(w)}e_{rd}^{(w)}}\right)^2 \frac{p^{(w)}(1-p^{(w)})}{N} \\ &= \text{Var}\left[\widehat{p}^{(w)}\right].\end{aligned}$$

In other words, as far as estimating p is concerned, the noise injection method does not help.

Appendix B
Proof of (35)

Following [20], we have the approximation for MSEE reduction ratio

$$\gamma \approx \frac{\int_0^\infty f(y_{rd})dy_{rd}}{\int_0^\infty W(y_{rd})f(y_{rd})dy_{rd}}, \quad W(y_{rd}) \overset{def}{=} \frac{f(y_{rd})}{f^*(y_{rd})} \tag{B.1}$$

where $f^*(y_{rd})$ and $f(y_{rd})$ are Gaussian pdfs with the same mean $\sqrt{P_r|h_{rd}|^2}$ but distinct variances $a_{rd}^{(w)}\sigma_d^2$ and σ_d^2, respectively.

After some calculations, we have

$$\int_0^\infty W(y_{rd})f(y_{rd})dy_{rd} = \frac{a_{rd}^{(w)}}{\sqrt{2a_{rd}^{(w)}-1}} Q\left(\sqrt{\frac{\left(2a_{rd}^{(w)}-1\right)P_r|h_{rd}|^2}{a\sigma_d^2}}\right). \tag{B.2}$$

Since $Q(y) \approx (\exp(-y^2/2)/y\sqrt{2\pi})$, for large y, we obtain

$$\gamma \approx \frac{2a_{rd}^{(w)}-1}{\left(\sqrt{a_{rd}^{(w)}}\right)^3}\exp\left[-\frac{\left(1-a_{rd}^{(w)}\right)P_r|h_{rd}|^2}{2a_{rd}^{(w)}\sigma_d^2}\right]. \tag{B.3}$$

The approximation $2a_{rd}^{(w)}-1 \approx 2a_{rd}^{(w)}$ yields

$$\gamma \approx \frac{2}{\sqrt{a_{rd}^{(w)}}}\exp\left[-\frac{\left(1-a_{rd}^{(w)}\right)P_r|h_{rd}|^2}{2a_{rd}^{(w)}\sigma_d^2}\right] \tag{B.4}$$

which is maximized when $a_{rd}^{(w)} = P_r|h_{rd}|^2/\sigma_d^2$.

Appendix C
Proof of Lemma 4

The analysis presented here follows that of [11] with three major distinctions: 1) We do not use the small ER assumption $e_i^{(w)} \ll 1$; 2) we have equal ER constraint; and 3) we need to consider the ER conversion (32).

Assuming independent links, we can show that the covariance matrix of the pairwise matching indicators $I(\hat{y}_k[t] = \hat{y}_j[t])$ for the noise injected network is

$$\mathbf{C} = \begin{pmatrix} p_{12}(1-p_{12}) & p_{123}-p_{12}p_{13} & p_{123}-p_{12}p_{23} \\ p_{123}-p_{12}p_{13} & p_{13}(1-p_{13}) & p_{123}-p_{13}p_{23} \\ p_{123}-p_{12}p_{23} & p_{123}-p_{13}p_{23} & p_{23}(1p_{23}) \end{pmatrix} \tag{C.1}$$

where

$$p_{kl} = \left(1-e_k^{(w)}\right)\left(1-e_l^{(w)}\right) + e_k^{(w)}e_l^{(w)}$$

$$p_{klm} = \left(1-e_k^{(w)}\right)\left(1-e_l^{(w)}\right)\left(1-e_m^{(w)}\right) + e_k^{(w)}e_l^{(w)}e_m^{(w)}.$$

The three-link network induces the nonlinear system (17) whose solution is given by (18). It is easier to compute the associated inverse Jacobian matrix for such a nonlinear mapping, i.e.,

$$\mathbf{J}^{-1} = \begin{pmatrix} \left(2e_2^{(w)}-1\right) & \left(2e_1^{(w)}-1\right) & 0 \\ \left(2e_3^{(w)}-1\right) & 0 & \left(2e_1^{(w)}-1\right) \\ 0 & \left(2e_3^{(w)}-1\right) & \left(2e_2^{(w)}-1\right) \end{pmatrix}.$$

Using the constraint $e_1^{(w)} = e_2^{(w)} = e_3^{(w)} = \epsilon$, we obtain the Jacobian and covariance matrices as

$$\mathbf{J} = \frac{1}{2(2\epsilon-1)}\begin{pmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} z_1 & z_2 & z_2 \\ z_2 & z_1 & z_2 \\ z_2 & z_2 & z_1 \end{pmatrix}$$

where $z_1 = 2\epsilon - 6\epsilon^2 + 8\epsilon^3 - 4\epsilon^4$, and $z_2 = \epsilon - 5\epsilon^2 + 8\epsilon^3 - 4\epsilon^4$.

The covariance matrix for the estimation error is thus given by

$$\mathbf{JCJ}^T = \frac{1}{4(2\epsilon-1)^2}\begin{pmatrix} 3z_1-2z_2 & 2z_2-z_1 & 2z_2-z_1 \\ 2z_2-z_1 & 3z_1-2z_2 & 2z_2-z_1 \\ 2z_2-z_1 & 2z_2-z_1 & 3z_1-2z_2 \end{pmatrix}.$$

The variance of the estimator $\widehat{\epsilon}$ can be approximated by

$$\mathrm{Var}[\widehat{\epsilon}] \approx \frac{3z_1-2z_2}{4(2\epsilon-1)^2 N} = \frac{4\epsilon-8\epsilon^2+8\epsilon^3-4\epsilon^4}{4(2\epsilon-1)^2 N} \tag{C.2}$$

and the variance of $\widehat{e}_i$ can be approximated by (see [19])

$$\mathrm{Var}[\widehat{e}_i] \approx \left(\frac{dg_i(\epsilon)}{d\epsilon}\right)^2 \tag{C.3}$$

$$\mathrm{Var}[\widehat{\epsilon}] = \frac{\left(a_i^{(w)}\right)^2}{\left(a_i^{(w)}+2\epsilon-2a_i^{(w)}\epsilon\right)^4}\frac{\epsilon-2\epsilon^2+2\epsilon^3-\epsilon^4}{(2\epsilon-1)^2 N} \tag{C.4}$$

where $g_i(x) = x/(a_i^{(w)}+2x-2a_i^{(w)}x)$ is the noncoherent conversion rule.

Appendix D
Proof of Theorem 1

Taking into account the constant noise-injected link ER constraint, we express the average bit ERs for BFSK and DPSK as

$$P_b^{bfsk} = a_i^{(w)}\left(2a_i^{(w)}+\mathrm{SNR}_i\right)^{-1} = \epsilon \tag{D.1}$$

$$P_b^{dpsk} = a_i^{(w)}\left[2\left(a_i^{(w)}+\mathrm{SNR}_i\right)\right]^{-1} = \epsilon. \tag{D.2}$$

Using (D.2) and omitting the superscript (w) for simplicity, we obtain

$$a_i+2\epsilon-2a_i\epsilon = \frac{a_i(\mathrm{SNR}_i+1)}{a_i+\mathrm{SNR}_i} \tag{D.3}$$

which, along with Lemma 4, gives

$$Var(\widehat{e}_i) \approx \frac{a_i^2(a_i+\mathrm{SNR}_i)^4}{(a_i\mathrm{SNR}_i+a_i)^4} \times\left[\frac{3a_i^4+12a_i^3\mathrm{SNR}_i+16a_i^2\mathrm{SNR}_i^2+8a_i\mathrm{SNR}_i^3}{16(a_i+\mathrm{SNR}_i)^4 N}\frac{(a_i+\mathrm{SNR}_i)^2}{\mathrm{SNR}_i^2}\right].$$

The MSEE reduction ratio γ is thus given by (D.4), shown at the bottom of the page. Using the change of variable $q_i = a_i/\mathrm{SNR}_i$, we find that the condition $(\partial\gamma/\partial a_i) = 0$ is equivalent to

$$6q_i^5+27q_i^4+43q_i^3+26q_i^2-4 = 0. \tag{D.5}$$

Since the only positive rational root is $q_i \approx 0.30855316 \equiv t_1$, (D.2) suggests that we inject noise such that

$$\epsilon = \frac{t_1\mathrm{SNR}_i}{2(t_1+1)\mathrm{SNR}_i} = 0.1179. \tag{D.6}$$

Furthermore, the minimum achievable MMSE reduction ratio is given by

$$\begin{aligned}\gamma_{min} &= \gamma|_{a_i=t_1\mathrm{SNR}_i} \\ &= \frac{78.622\mathrm{SNR}^4}{8\mathrm{SNR}_i^5+32\mathrm{SNR}_i^4+52\mathrm{SNR}_i^3+43\mathrm{SNR}_i^2+18\mathrm{SNR}_i+3} \\ &\approx 9.8277\frac{\mathrm{SNR}_i^2}{(1+\mathrm{SNR}_i)^3}.\end{aligned}$$

Solving the equation $\gamma|_{a_i=t_1\mathrm{SNR}_i} = 1$ for SNR_i gives one positive repeated root 3.24092. Since $a_i > 1$, SNR_i must be greater than $1/t_1 = 3.24093$ in order for noise injection to become beneficial.

Employing a similar approach for a BFSK-based network, we conclude that

$$\begin{aligned}\gamma_{min} &= \gamma|_{a_i=t_1'\mathrm{SNR}_i} \\ &= \frac{19.655\mathrm{SNR}_i^4}{\mathrm{SNR}_i^5+8\mathrm{SNR}_i^4+26\mathrm{SNR}_i^3+43\mathrm{SNR}_i^2+36\mathrm{SNR}_i+12} \\ &\approx 19.655\frac{\mathrm{SNR}_i^2}{(2+\mathrm{SNR}_i)^3}\end{aligned} \tag{D.7}$$

where $t_1' = 0.15427658$, and noise injection is beneficial only if $\mathrm{SNR}_i \geq 6.4828$.

Appendix E
Proof of Lemma 5

We begin with the simpler case where the network only consists of PLs 1 and 2 and VL 1, whose outputs are y_1, y_2, and $y_1^{(v)}$. The probability that two PLs and the VL all give identical decisions can be decomposed as

$$\Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}\right) = \frac{1}{2}\left[\Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=1\right) + \Pr\left(\widehat{y}_2 = \widehat{y}_1 = \widehat{y}_1^{(v)}|x=-1\right)\right] \stackrel{def}{=} p_{12(v1)}. \tag{E.1}$$

$$\gamma = \frac{1}{a_i^2}\left[\frac{3a_i^6+18a_i^5\mathrm{SNR}_i+43a_i^4\mathrm{SNR}_i^2+52a_i^3\mathrm{SNR}_i^3+32a_i^2\mathrm{SNR}_i^4+8a_i\mathrm{SNR}_i^5}{(1+\mathrm{SNR}_i)^2\left(3+12\mathrm{SNR}_i+16\mathrm{SNR}_i^2+8\mathrm{SNR}_i^3\right)}\right] \tag{D.4}$$

The binary symmetric nature of both PLs gives

$$\begin{aligned}\Pr&\left(\widehat{y}_2=\widehat{y}_1=\widehat{y}_1^{(v)}|x=-1\right)\\&=\Pr\left(\widehat{y}_2=\widehat{y}_1=\widehat{y}_1^{(v)}|x=1\right)\\&=\Pr(\widehat{y}_2=1|x=1)\Pr\left(\widehat{y}_1=\widehat{y}_1^{(v)}=1|x=1\right)\\&\quad+\Pr(\widehat{y}_2=-1|x=1)\Pr\left(\widehat{y}_1=\widehat{y}_1^{(v)}=-1|x=1\right)\\&\stackrel{\text{def}}{=}(1-Q_2)p_{sm}+Q_2p_{em}.\end{aligned}\tag{E.2}$$

Based on the normalized model for link 1, i.e., $y_1=hx+w$, where $x\in\{\pm1\}$, h is Raleigh distributed, and w is a zero-mean Gaussian random variable with variance $\text{var}(w)=N_0/2=1/2\text{SNR}_1$, we obtain

$$\begin{aligned}p_{em}&=\Pr\left(\widehat{y}_1=\widehat{y}_1^{(v)}=-1|x=1\right)\\&=\Pr\left(\widehat{y}_1=\widehat{y}_1^{(v)}=1|x=-1\right)\\&=\int_h\Pr\left(-h+w>0,-h+w+w_v>0|x=-1,h\right)\\&\quad\times f(h)dh\\&=\int_h\Pr\left(n>h\sqrt{\frac{2}{N_0}},m>h\sqrt{\frac{2}{a_1^{(v)}N_0}}\,\middle|\,x=-1,h\right)\\&\quad\times f(h)dh\end{aligned}\tag{E.3}$$

where $m=(w+w_v)/\sqrt{(a_1^{(v)}N_0)/2}$, $n=w/\sqrt{N_0/2}$, w_v is a zero-mean real Gaussian random variable with variance $(a_1^{(v)}-1)N_0/2$, and $E[nm]=1/\sqrt{a_1^{(v)}}$.

The first integrand of (E.3) can be expressed as a standard bivariate Gaussian distribution function $Q(x,y;\rho)$, which, in turn, yields the Craig form as [14, eq. (4.17)]

$$\begin{aligned}\Pr&\left[n>h\sqrt{\frac{2}{N_0}},m>h\sqrt{\frac{2}{a_1^{(v)}N_0}}\,\middle|\,x=-1,h\right]\\&=Q\left(h\sqrt{\frac{2}{N_0}},h\sqrt{\frac{2}{a_1^{(v)}N_0}};\rho\right)\\&=\frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{\sqrt{a_1^{(v)}-1}}{1-\rho\sqrt{a_1^{(v)}}}\right)}\exp\left(-\frac{2h^2}{2N_0\sin^2\Phi}\right)d\Phi\\&\quad+\frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)}-1}}\right)}\exp\left(-\frac{2\rho^2h^2}{2N_0\sin^2\Phi}\right)d\Phi\end{aligned}\tag{E.4}$$

where $\rho=(1/\sqrt{a_1^{(v)}})$ is the correlation coefficient.

Using the method described in [14, ch. 5] and the identity [14, eq. (5.A.11)]

$$\int\left(1+\frac{c}{\sin^2\Phi}\right)^{-1}d\Phi=\Phi-\sqrt{\frac{c}{c+1}}\tan^{-1}\left[\frac{\tan\Phi}{\sqrt{\frac{c}{c+1}}}\right]$$

we obtain

$$\begin{aligned}\int_h&Q\left(h\sqrt{\frac{2}{N_0}},h\sqrt{\frac{2}{a_1^{(v)}N_0}},\rho\right)f(h)dh\\&=\frac{1}{2\pi}\int_0^{\pi/2}\left(1+\frac{1}{N_0\sin^2\Phi}\right)^{-1}d\Phi\\&\quad+\frac{1}{2\pi}\int_0^{\tan^{-1}\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)}\left(1+\frac{\rho^2}{N_0\sin^2\Phi}\right)^{-1}d\Phi\\&=\frac{1}{4}\left(1-\sqrt{\frac{\text{SNR}_1}{1+\text{SNR}_1}}\right)\\&\quad+\frac{1}{2\pi}\left[\tan^{-1}\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)\right.\\&\quad\left.-\sqrt{\frac{\rho^2\text{SNR}_1}{1+\rho^2\text{SNR}_1}}\tan^{-1}\left(\frac{\rho\sqrt{1+\rho^2\text{SNR}_1}}{\sqrt{(1-\rho^2)\rho^2\text{SNR}_1}}\right)\right]\\&=\frac{e_1}{2}+\frac{1}{2\pi}\\&\quad\times\left[\tan^{-1}\left(\frac{1}{\sqrt{a_1^{(v)}-1}}\right)-\left(1-2e_1^{(v)}\right)\right.\\&\quad\left.\times\tan^{-1}\left[\frac{\left(1-2e_1^{(v)}\right)^{-1}}{\sqrt{(a_1^{(v)}-1)}}\right]\right]\stackrel{def}{=}p_{em}\left(e_1,a_1^{(v)}\right).\end{aligned}\tag{E.5}$$

Invoking the relation [14, eq. (6.42)]

$$Q(-x,-y;\rho)=1-Q(x)-Q(y)+Q(x,y;\rho)\quad x,y\geq0$$

and (E.5), we express the conditional correct (pairwise) SMP as

$$\begin{aligned}\Pr&\left(\widehat{y}_1=\widehat{y}_1^{(v)}=1|x=1\right)=\Pr\left(\widehat{y}_1=\widehat{y}_1^{(v)}=-1|x=-1\right)\\&=\int_h\Pr(w>-h,m>-h|x=-1,h)f(h)dh\\&=1-e_1-e_1^{(v)}+p_{em}\left(e_1,a_1^{(v)}\right)\stackrel{def}{=}p_{cm}\left(e_1,a_1^{(v)}\right).\end{aligned}\tag{E.6}$$

Summarizing (E.1)–(E.6), we then obtain

$$p_{12(v1)}=e_2p_{em}\left(e_1,a_1^{(v)}\right)+(1-e_2)p_{cm}\left(e_1,a_1^{(v)}\right)\tag{E.7}$$

which is (40) in the main text. The other probabilities (43)–(46) can similarly be derived with the aid of the following two identities [14, eq. (6.42)]:

$$Q(x,y,\rho)=Q(x)-Q(x,-y,-\rho),\quad x\geq 0, y<0 \tag{E.8}$$

$$Q(x,y,\rho)=Q(y)-Q(-x,y,-\rho),\quad x<0, y\geq 0. \tag{E.9}$$

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] I. Olabarrieta and J. Del Ser, "Enhanced sensing error probability estimation for iterative data fusion in the low SNR regime," in *Proc. Int. ITG WSA*, Bremen, Germany, 2010, pp. 270–274.

[3] J. Del Ser, J. Garcia-Frias, and P. M. Crespo, "Iterative concatenated zigzag decoding and blind data fusion of correlated sensors," in *Proc. ICUMT*, St. Petersburg, Russia, 2009, pp. 1–6.

[4] A. Razi, F. Afghah, and A. Abedi, "Binary source estimation using a two-tiered wireless sensor network," *IEEE Commun. Lett.*, vol. 15, no. 4, pp. 449–451, Apr. 2011.

[5] F. Daneshgaran, M. Laddomada, and M. Mondin, "Iterative joint channel decoding of correlated sources," *IEEE Inf. Theory*, vol. 51, no. 7, pp. 2721–2731, Jul. 2005.

[6] J. N. Laneman and G. W. Wornell, "Energy-efficient antenna sharing and relaying for wireless networks," in *Proc. IEEE WCNC*, Chicago, IL, Sep. 2000, pp. 7–12.

[7] D. Chen and J. N. Laneman, "Modulation and demodulation for cooperative diversity in wireless systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1785–1794, Jul. 2006.

[8] Q. Zhao and H. Li, "Differential modulation for cooperative wireless systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2273–2283, May 2007.

[9] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich, "Link-failure probabilities for practical cooperative relay networks," in *Proc. IEEE Veh. Tech. Conf. Spring*, Stockholm, Sweden, 2005, pp. 1489–1493.

[10] A. Dixit, S. C. Douglas, and G. C. Orsak, "Blind estimation of channel BERs in a multi-receiver network," in *Proc. IEEE ICASSP*, Montreal, QC, Canada, May 17–21, 2004, vol. 2, pp. II-305–II-308.

[11] J. P. Delmas and Y. Meurisse, "Performance analysis of optimal blind fusion of bits," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1477–1485, Apr. 2007.

[12] E. L. Lehmann and G. Casella, *Theory of Point Estimation,* 2nd ed. New York: Springer-Verlag, 1998.

[13] A. Dixit and G. C. Orsak, "Asymptotically optimal blind fusion of bit estimates," in *Proc. IEEE Signal Process. Workshop*, Callaway Gardens, GA, Oct. 13–16, 2002, pp. 377–382.

[14] M. K. Simon and M. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis,* 2nd ed. Hoboken, NJ: Wiley, 2004.

[15] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

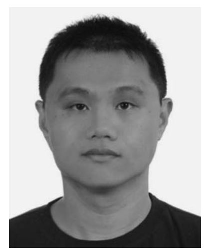
[16] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: A review of importance sampling techniques in communications systems," *IEEE J. Sel. Areas Commun.*, vol. 15, no. 4, pp. 597–613, May 1997.

[17] H. Chen, P. K. Varshney, and J. H. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.

[18] Y. B. Li and Y. L. Guan, "Modified Jakes' model for simulating multiple uncorrelated fading waveforms," in *Proc. 51st IEEE Veh. Technol. Conf.*, Tokyo, Japan, May 2000, pp. 1819–1822.

[19] G. Casella and R. L. Berger, *Statistical Inference,* 2nd ed. Pacific Grove, CA: Duxbury, 2002.

[20] K. S. Shanmugam and P. Balaban, "A modified Monte-Carlo simulation techniques for the evaluation of error rate in digital communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1916–1924, Nov. 1980.



Jen-Yang Liu (S'07) was born in Taipei, Taiwan, in 1981. He received the B.S. degree (with the highest distinction) in electrical engineering from the National Sun Yat-Sen University, Kaushiung, Taiwan, in 2002 and the M.S. degree in communications engineering from the National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 2004, where he is currently working toward the doctoral degree.

Among his research topics are communication signal processing (including stochastic resonance) and performance analysis of communication systems.

Yu T. Su (SM'06) received the B.S.E.E. degree from Tatung Institute of Technology, Taipei, Taiwan, in 1974 and the Ph.D. degree from the University of Southern California, Los Angeles, in 1983.

From 1983 to 1989, he was with LinCom Corporation, Los Angeles, where his last position was Corporate Scientist. Since September 1989, he has been with the National Chiao Tung University, Hsinchu, Taiwan, where he is currently a Professor with the Department of Electrical Engineering. He was an Associative Dean of the College of Electrical and Computer Engineering from 2004 to 2007 and Head of the Communications Engineering Department from 2001 to 2003. He is also affiliated with the Microelectronic and Information Systems Research Center of the same university and served as a Deputy Director during 1997–2000. From 2005 to 2008, he was the Area Coordinator of the Taiwan National Science Council's Telecommunications Program. His main research interests include communication theory and statistical signal processing.