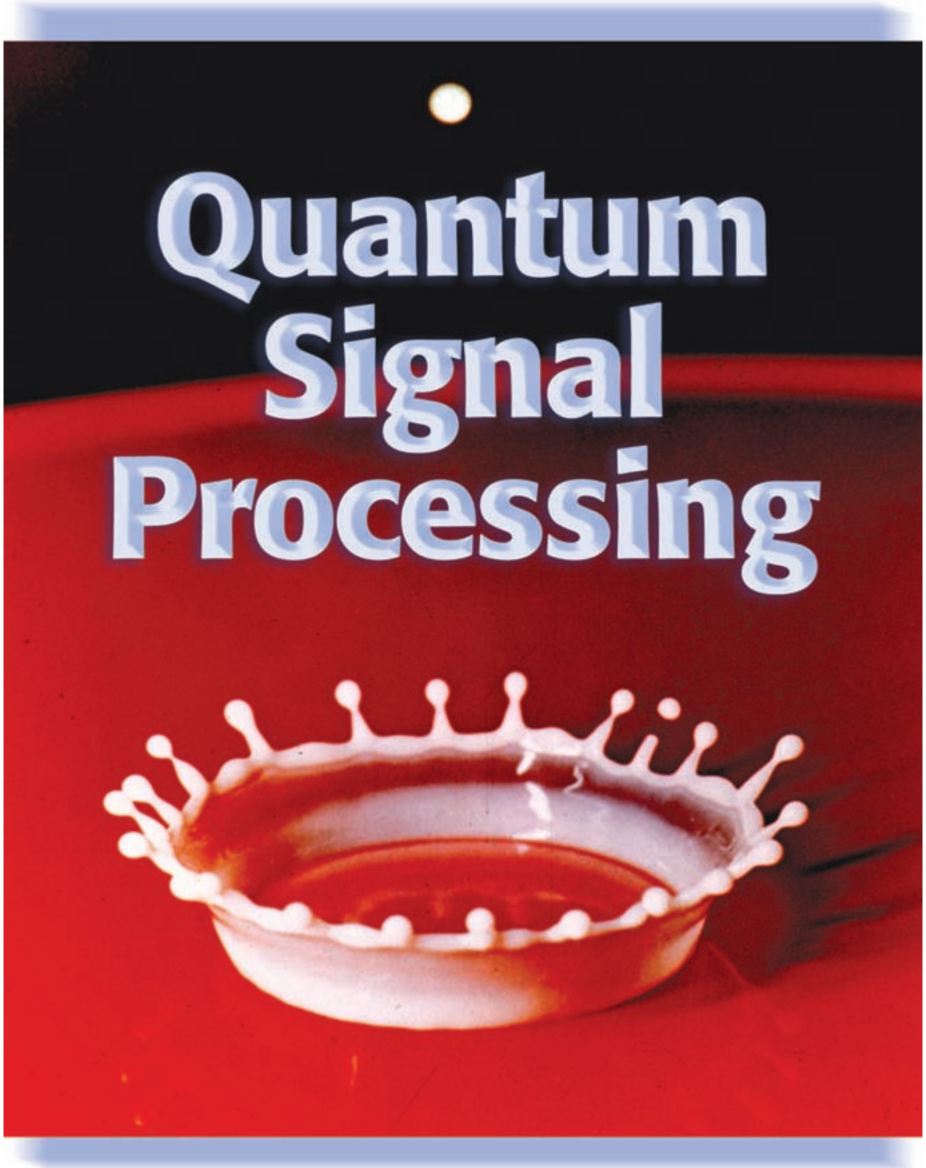


Measurement, Quantization, Consistency

This famous strobe photograph taken by Dr. Harold "Doc" Edgerton at MIT in 1957 captures the transition of a milk drop as it splashes from a single drop into a continuous ring and then coalesces into droplets forming the points of a coronet. Copyright Harold and Ester Edgerton Family Foundation. Reproduced with permission.

Yonina C. Eldar and Alan V. Oppenheim

In this article we present a signal processing framework that we refer to as quantum signal processing (QSP) [1] that is aimed at developing new or modifying existing signal processing algorithms by borrowing from the principles of quantum mechanics and some of its interesting axioms and constraints. However, in contrast to such fields as quantum computing and quantum information theory, it does not inherently depend on the physics associated with quantum mechanics. Consequently, in developing the QSP framework we are free to impose quantum mechanical constraints that we find useful and to avoid those that are not. This framework provides a unifying conceptual structure for a variety of traditional processing techniques and a precise mathematical setting for developing generalizations and extensions of algorithms, leading to a potentially useful paradigm for signal processing with applications in areas including frame theory, quantization and sampling methods, detection, parameter estimation, covariance shaping, and multiuser wireless communication systems.

1053-5888/02/$17.00©2002IEEE

Fig. 1 presents a general overview of the key elements in quantum physics that provide the basis for the QSP framework and an indication of the key results that have so far been developed within this framework. In the remainder of this article, we elaborate on the various elements in this figure.

Overview

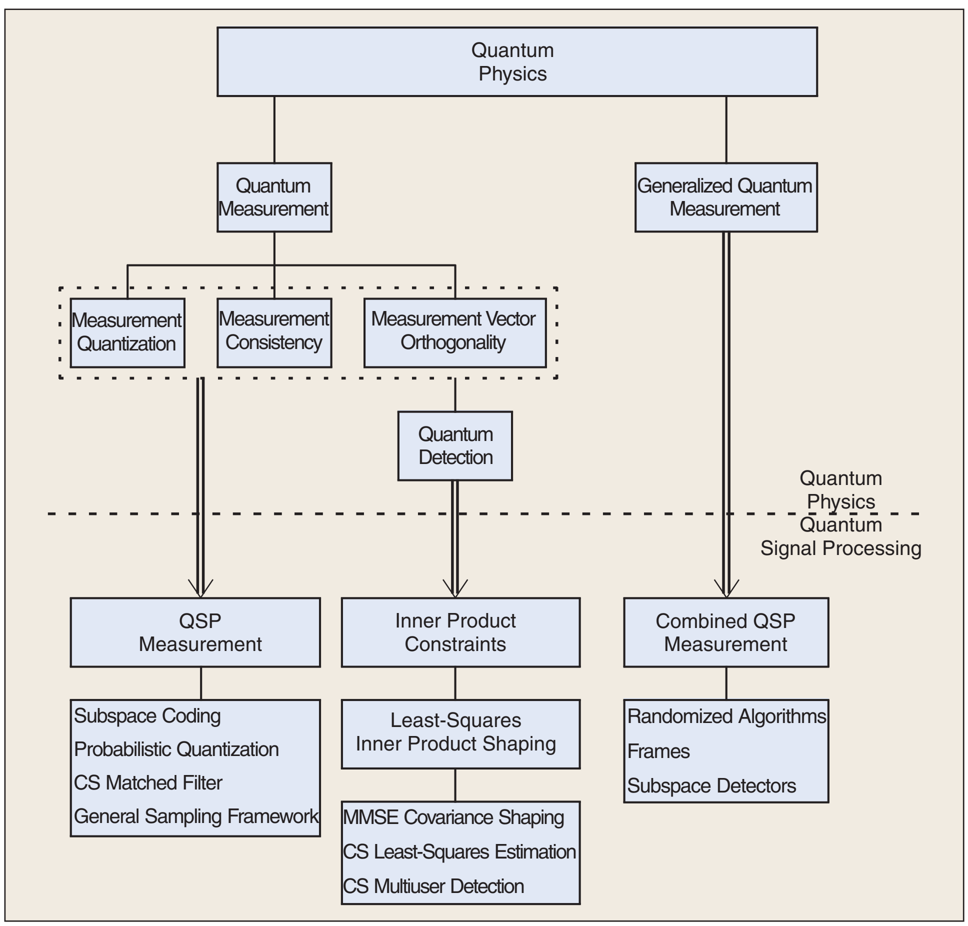
Many new classes of signal processing algorithms have been developed by emulating the behavior of physical systems. There are also many examples in the signal processing literature in which new classes of algorithms have been developed by artificially imposing physical constraints on implementations that are not inherently subject to these constraints. In this article, we survey a new class of algorithms of this type that we refer to as QSP. A fully detailed treatment is contained in [1] as well as in the various journal and conference papers referred to throughout this article. Among the many well-known examples of digital signal processing algorithms that are derived by using physical phenomena and constraints as metaphors are wave digital filters [2]. This class of filters relies on emulating the physical constraints of passivity and energy conservation associated with analog filters to achieve low sensitivity in coefficient variations in digital filters. As another class of examples, the fractal-like aspects of nature [3] and related modeling have inspired interesting signal processing paradigms that are not constrained by the associated physics. These include fractal modulation [4], which emulates the fractal characteristic of nature for communicating over a particular class of unreliable channels, and the various approaches to image compression based on synthetic generation of fractals [5]. Likewise, the chaotic behavior of certain features of nature have inspired new classes of signals for secure communications, remote sensing, and a variety of other signal processing applications [6]-[9]. Other examples of algorithms using physical systems as a simile include solitons [10], genetic algorithms [11], simulated annealing [12], and neural networks [13].

These examples are just several of many that underscore the fact that even in signal processing contexts that are not constrained by the physics, exploiting laws of nature can inspire new methods for algorithm design and may lead to interesting, efficient, and effective processing techniques.

Three fundamental inter-related underlying principles of quantum mechanics that play a major role in QSP as presented here are the concept of a measurement, the principle of measurement consistency, and the principle of quantization of the measurement output. In addition, when using quantum systems in a communication context, other principles arise such as inner product constraints. QSP is based on exploiting these various principles and constraints in the context of signal processing algorithms.

In a broad sense, the terms measurement, measurement consistency, and output quantization are well known in signal processing, although not with the same meaning or precise mathematical interpretation and constraints as in quantum mechanics.

In signal processing, the term measurement can be given a variety of precise or imprecise interpretations. However, as discussed further in this article, in quantum mechanics measurement has a very specific definition and meaning, much of which we carry over to the QSP frame-



▲ *1. QSP framework, its relationship to quantum physics, and key results. In the figure, CS stands for covariance shaping.*

work. Similarly, in signal processing, quantization is traditionally thought of in fairly specific terms. In quantum mechanics, quantization of the measurement output is a fundamental underlying principle, and applying this principle, along with the quantum mechanical notions of measurement and consistency, leads to some potentially intriguing generalizations of quantization as it is typically viewed in signal processing.

Measurement consistency also has a precise meaning in quantum mechanics, specifically that repeated applications of a measurement must yield the same outcome. In signal processing, a similar consistency concept is the basis for a variety of classes of algorithms including signal estimation, interpolation, and quantization methods. Some early examples of consistency as it typically arises in signal processing are the interpolation condition in filter design [14] and the condition for avoiding intersymbol interference in waveforms for pulse amplitude modulation [15]. More recent examples include perfect reconstruction filter banks [16], [17], multiresolution and wavelet approximations [18], [19], and sampling methods in which the perfect reconstruction requirement is replaced by the less stringent consistency requirement [20]-[24]. Here again, viewing measurement consistency in a broader framework motivated by quantum mechanics can lead to some new and interesting signal processing algorithms.

In the next section, we summarize the basic principles of measurement, consistency, and quantization as they relate to quantum mechanics. We also outline the key elements and constraints that are associated with the use of quantum states for communication in the context of what is commonly referred to as the quantum detection problem [25]. We then indicate how these principles and constraints imposed by the physics are emulated and applied in the framework of QSP.

Quantum Systems

In developing the QSP framework, it is convenient to discuss both signal processing and quantum mechanics in a vector space setting. Specifically we consider an arbitrary Hilbert (inner product) space $\mathcal{H}$ with inner product $\langle x, y \rangle$ for any vectors x, y in $\mathcal{H}$. Typically we will refer to elements of $\mathcal{H}$ as vectors or signals interchangeably. In this section we summarize the key elements of physical quantum systems that are emulated in QSP, including those associated with quantum detection.

Measurement

A quantum system in a pure state is characterized by a normalized vector in $\mathcal{H}$. Information about a quantum system is extracted by subjecting the system to a quantum measurement.

A *quantum measurement* is a nonlinear (probabilistic) mapping that in the simplest case can be described in terms of a set of measurement vectors μ_i that span measurement subspaces $\mathcal{S}_i$ in $\mathcal{H}$. The laws of quantum mechanics impose the constraint that the vectors μ_i must be orthonormal and therefore also linearly independent. A measurement of this form is referred to as a *rank-one quantum measurement*. In the more general case, the quantum measurement is described in terms of a set of projection operators P_i onto subspaces $\mathcal{S}_i$ of $\mathcal{H}$, where from the laws of quantum mechanics these projections must form a complete set of orthogonal projections. Such a measurement is referred to as a *subspace quantum measurement*. In quantum mechanics, the outcome of a measurement is inherently probabilistic, with the probabilities of the outcomes of any measurement determined by the vector representing the underlying state of the system at the time of the measurement. The measurement collapses (projects) the state of the quantum system onto a state that is compatible with the measurement outcome. In general the final state of the system is different than the state of the system prior to the measurement.

Measurement consistency is a fundamental postulate of quantum mechanics, i.e., repeated measurements on a system must yield the same outcomes; otherwise we would not be able to confirm the output of a measurement. Therefore the state of the system after a measurement must be such that if we instantaneously remeasure the system in this state, then the final state after this second measurement will be identical to the state after the first measurement.

Quantization of the measurement outcome is a direct consequence of the consistency requirement. Specifically, the consistency requirement leads to a class of states referred to as *determinate states* of the measurement [26]. These are states of the quantum system for which the measurement yields a known outcome with probability one and are the states that lie completely in one of the measurement subspaces $\mathcal{S}_i$. Furthermore, even when the state of the system is not one of the determinate states, after performing the measurement the system is quantized to one of these states, i.e., is certain to be in one of these states, where the probability of being in a particular determinate state is a function of the inner products between the state of the system and the determinate states. See "Quantum Measurement" for more details.

As an example of a rank-one quantum measurement, suppose that the measurement is defined by two orthonormal measurement vectors μ_1 and μ_2 and the state of the system is given by $x = (1/2)\mu_1 + (\sqrt{3}/2)\mu_2$. As outlined further in "Quantum Measurement," the measurement will project the state x onto one of the measurement vectors μ_1 or μ_2, where the probability of projecting onto μ_1 is $p(1) = |\langle x, \mu_1 \rangle|^2 = 1/4$ and the probability of projecting onto μ_2 is $p(2) = |\langle x, \mu_2 \rangle|^2 = 3/4$. The measurement process is illustrated in Fig. 2.

Quantum Measurement

A standard (von Neumann) measurement in quantum mechanics is defined by a collection of projection operators $\{P_i, i \in \mathcal{I}\}$ onto subspaces $\{\mathcal{S}_i \subseteq \mathcal{H}, i \in \mathcal{I}\}$, where $\mathcal{I}$ denotes an index set and the index $i \in \mathcal{I}$ corresponds to a possible measurement outcome. The laws of quantum mechanics impose the constraint that the operators $\{P_i, i \in \mathcal{I}\}$ form a complete set of orthogonal projections so that with $(\cdot)^*$ denoting the adjoint of the corresponding transformation, for any $i, k \in \mathcal{I}$,

$$P_i = P_i^*; \tag{1}$$

$$P_i^2 = P_i; \tag{2}$$

$$P_i P_k = 0, \quad \text{if} \quad i \neq k; \tag{3}$$

$$\sum_{i \in \mathcal{I}} P_i = I_{\mathcal{H}}. \tag{4}$$

Conditions (3) and (4) imply that the measurement subspaces $\mathcal{S}_i$ are orthogonal and that their direct sum is equal to $\mathcal{H}$.

If the state vector is ϕ, then from the rules of quantum mechanics, the probability of observing the ith outcome is

$$p(i) = \langle P_i \phi, \phi \rangle. \tag{5}$$

Since the state is normalized,

$$\sum_i p(i) = \langle \sum_i P_i \phi, \phi \rangle = \langle \phi, \phi \rangle = 1.$$

In the simplest case, the projection operators are rank-one operators $P_i = \mu_i \mu_i^*$ for some nonzero vectors $\{\mu_i \in \mathcal{H}, i \in \mathcal{I}\}$. We refer to such measurements as *rank-one quantum measurements*. Then (3) and (4) imply that the measurement vectors $\{\mu_i, i \in \mathcal{I}\}$ form an orthonormal basis for $\mathcal{H}$. If the state vector is ϕ, then the probability of observing the ith outcome is

$$p(i) = |\langle \mu_i, \phi \rangle|^2. \tag{6}$$

For a rank-one quantum measurement defined by orthonormal measurement vectors μ_i, it follows from (1) that if $\phi = \mu_i$ for some i, then $p(i) = 1$ and output i is obtained with probability one (w.p. 1). The states $\{\phi = \mu_i\}$ are therefore called the *determinate states* of the measurement. More generally, the determinate states are the states that lie completely in one of the measurement spaces $\mathcal{S}_i$. Indeed, if $\phi \in \mathcal{S}_i$, then $P_i \phi = \phi$ and from (5), $p(i) = 1$.

The quantum measurement can be formulated in terms of a probabilistic mapping between $\mathcal{H}$ and the determinate states. Given a space of states $\mathcal{X}$ and an observation space $\mathcal{Y}$, a *probabilistic mapping* from $\mathcal{X}$ to $\mathcal{Y}$ is a function $f\colon \mathcal{X} \times \mathcal{W} \to \mathcal{Y}$, where $\mathcal{W}$ is the sample space of an auxiliary chance variable $W(x) = \{\mathcal{W}, p_{W|X}(w|x)\}$ with a probability distribution $p_{W|X}(w|x)$ on $\mathcal{W}$ that in general depends on $x \in \mathcal{X}$. Note that a deterministic mapping $f\colon \mathcal{X} \to \mathcal{Y}$ is a special case of a probabilistic mapping in which the auxiliary chance variable is deterministic; i.e., it has one outcome w w.p. 1. In this case the function f is independent of W.

A rank-one quantum measurement corresponding to orthonormal measurement vectors $\{\mu_i \in \mathcal{H}, i \in \mathcal{I}\}$ that span subspaces $\{\mathcal{S}_i \subseteq \mathcal{H}, i \in \mathcal{I}\}$ can be viewed as a probabilistic mapping between $\mathcal{H}$ and the determinate states that is

- 1) a deterministic identity mapping for $\phi \in \mathcal{S}_i$;
- 2) a probabilistic mapping for nondeterminate states that maps ϕ to a normalized vector in the direction of the orthogonal projection $P_i \phi$ for some value $i \in \mathcal{I}$, where $i = f(\{\langle \mu_k, \phi \rangle, k \in \mathcal{I}\}, w_i)$.

Here $f\colon \mathcal{H} \times \mathcal{W} \to \mathcal{I}$ is a probabilistic mapping between elements ϕ of $\mathcal{H}$ and indices $i \in \mathcal{I}$, which depends on a chance variable W with a discrete alphabet $\mathcal{W} = \mathcal{I}$ such that the probability of outcome $w_i \in \mathcal{W}$ depends on the input ϕ only through the inner products $\{\langle \mu_k, \phi \rangle, k \in \mathcal{I}\}$. Specifically, the probability of outcome w_i is $|\langle \mu_i, \phi \rangle|^2$. If w_i is observed, then $f(\{\langle \mu_k, x \rangle, k \in \mathcal{I}\}, w_i) = i$.

A subspace quantum measurement corresponding to a complete set of orthogonal projection operators $\{P_i, i \in \mathcal{I}\}$ onto subspaces $\{\mathcal{S}_i \subseteq \mathcal{H}, i \in \mathcal{I}\}$ can be viewed as a probabilistic mapping between $\mathcal{H}$ and the determinate states that is

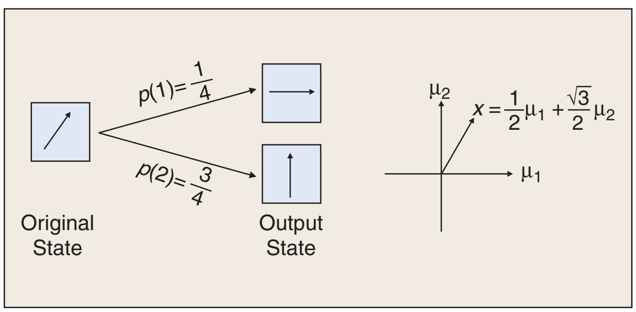
- 1) a deterministic identity mapping for $\phi \in \mathcal{S}_i$;
- 2) a probabilistic mapping for nondeterminate states that maps ϕ to a normalized vector in the direction of the orthogonal projection $P_i \phi$ for some value $i \in \mathcal{I}$, where $i = f(\{\langle P_k \phi, P_k \phi \rangle, k \in \mathcal{I}\}, w_i)$.

Here $f\colon \mathcal{H} \times \mathcal{W} \to \mathcal{I}$ is a probabilistic mapping between elements ϕ of $\mathcal{H}$ and indices $i \in \mathcal{I}$ that depends on a chance variable W with a discrete alphabet $\mathcal{W} = \mathcal{I}$ such that the probability of outcome $w_i \in \mathcal{W}$ depends on the input ϕ only through the inner products $\{\langle P_k \phi, P_k \phi \rangle, k \in \mathcal{I}\}$. Specifically, the probability of outcome w_i is $\langle P_i \phi, P_i \phi \rangle$. If w_i is observed, then $f(\{\langle P_k \phi, P_k \phi \rangle, k \in \mathcal{I}\}, w_i) = i$.

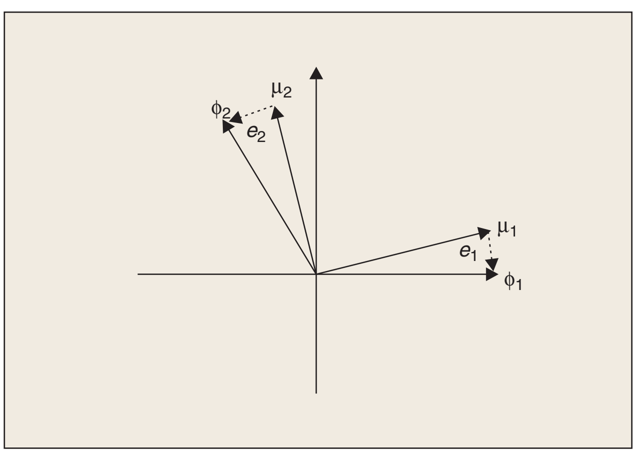
Quantum Detection

The constraints imposed by the physics on a quantum measurement lead to some interesting problems within the framework of quantum mechanics. In particular, an interesting problem that arises when using quantum states for communication is a *quantum detection* problem [25], which we outline here. As we discuss later in a signal processing context, this problem suggests a variety of new or generalized algorithms within the QSP framework based on specific ways of imposing inner product constraints.

In a quantum detection problem a sender conveys classical information to a receiver using a quantum-mechanical channel. The sender represents messages by preparing the quantum channel in a pure quantum state drawn from a collection of known states ϕ_i. The receiver detects the information by subjecting the channel to a quantum measurement with measurement vectors μ_i that are constrained by the physics to be orthogonal. If the states are not orthogonal, then no measurement can distinguish perfectly between them. In the general context of quantum mechanics, a fundamental problem is to construct measurements optimized to distinguish between a set of nonorthogonal pure quantum states. In the context of a communications system, we would like to choose the measurement vectors to minimize the probability of detection error. In this context this problem is commonly referred to as the quantum detection problem.



▲ *2. Illustration of a rank-one quantum measurement.*



▲ *3. Two-dimensional example of the LS measurement. The vectors μ_i are chosen to be orthonormal and to minimize $\sum_i \langle e_i, e_i\rangle = \sum_i \langle \phi_i - \mu_i, \phi_i - \mu_i\rangle$.*

Necessary and sufficient conditions for an optimum measurement minimizing the probability of detection error are known [25], [27]-[29]. However, except in some particular cases [25], [30]-[33], obtaining a closed-form analytical expression for the optimal measurement directly from these conditions is a difficult and unsolved problem.

In [32], an alternative approach is taken based on choosing a squared-error criterion and determining a measurement that minimizes this criterion. Specifically, the measurement vectors μ_i are chosen to be orthogonal and closest in a least-squares (LS) sense to the given set of state vectors ϕ_i so that the vectors μ_i are chosen to minimize the sum of the squared norms of the error vectors $e_i = \mu_i - \phi_i$, as illustrated in Fig. 3. The optimal measurement is referred to as the LS measurement.

The LS measurement problem has a simple closed-form solution with many desirable properties. Its construction is relatively simple; it can be determined directly from the given collection of states; it minimizes the probability of detection error when the states exhibit certain symmetries [32]; it is "pretty good" when the states to be distinguished are equally likely and almost orthogonal [34]; it achieves a probability of error within a factor of two of the optimal probability of error [35]; and it is asymptotically optimal [36].

Thus, in the context of quantum detection the constraints of the physics lead to the interesting problem of choosing an optimal set of orthogonal vectors. Borrowing from quantum detection, a central idea in QSP applications is to impose orthogonality or more general inner product constraints on algorithms and then use the LS measurement and the results derived in the context of quantum detection to design optimal algorithms subject to these constraints.

Quantum Signal Processing

As mentioned previously, the QSP framework draws heavily on the notions of measurement, consistency, and quantization as they relate to quantum systems. Furthermore it borrows from and generalizes the inner product constraint, specifically orthogonality, that quantum physics imposes on measurement vectors. However, the QSP framework is broader and less restrictive than the quantum measurement framework since in designing algorithms we are not constrained by the physical limitations of quantum mechanics.

In quantum mechanics, systems are "processed" by performing measurements on them. In signal processing, signals are processed by applying an algorithm to them. Therefore, to exploit the formalism and rich mathematical structure of quantum physics in the design of algorithms we first draw a parallel between a quantum mechanical measurement and a signal processing algorithm by associating a signal processing measurement with a signal processing algorithm. We then apply the formalism and fundamental principles of quantum measurement to the definition of the QSP measurement. An algorithm is then described by a QSP measurement, with additional input and output mappings if appropriate. This conceptual framework is illustrated schematically in Fig. 4.

The QSP framework is primarily concerned with the design of the QSP measurement, borrowing from the principles, axioms, and constraints of quantum physics and a quantum measurement. As we will show, the QSP measurement depends on a specific set of measurement parameters, so that this framework provides a convenient and useful setting for deriving new algorithms by choosing different measurement parameters, borrowing from the ideas of quantum mechanics. Furthermore, since the QSP measurement is defined to have a mathematical structure similar to a quantum measurement, the mathematical constraints imposed by the physics on the quantum measurement can also be imposed on the QSP measurement leading to some intriguing new signal processing algorithms.

QSP Measurement

The quantum-mechanical principles of measurement, consistency, and quantization lead to the definition of the QSP measurement.

Measurement of a signal in the QSP framework corresponds to applying an algorithm to a signal. As indicated in the previous section and in Fig. 4, one class of algorithms in the QSP framework consists of a QSP measurement (i.e., an algorithm) applied in a signal space that may perhaps be a remapping of the overall input and output signal spaces. For example, if the signal we wish to process is a sequence of scalars, then we may first map the scalar values into vectors in a higher dimensional space, which corresponds to the input mapping in Fig. 4. We then measure the vector representation. The measurement outcome is a signal in the same signal space as the measured signal, which is then mapped to the algorithm output using an output mapping as illustrated in Fig. 4, so that the measurement output represents the output of the algorithm, which in turn may be a signal or any other element. As in quantum mechanics, we require that if we remeasure the outcome signal, then the new outcome will be equal to the original outcome.

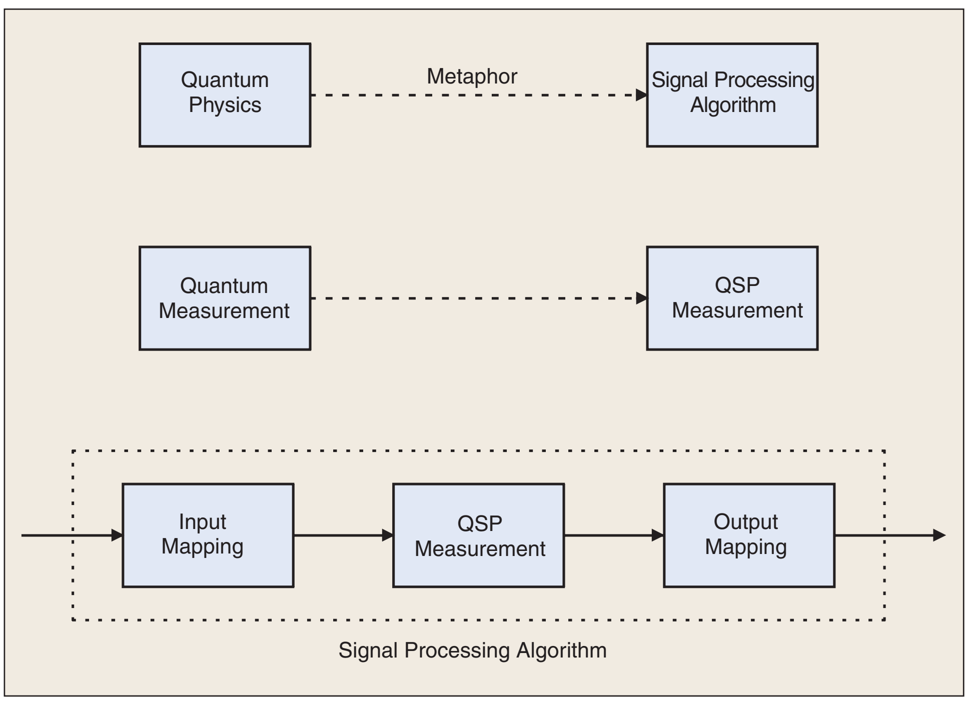
In analogy with the measurement in quantum mechanics, a rank-one QSP measurement (ROM) M on $\mathcal{H}$ is defined by a set of measurement vectors q_i that span one-dimensional subspaces $\mathcal{S}_i$ in $\mathcal{H}$. Since we are not constrained by the physics of quantum mechanics, these vectors are not constrained to be orthonormal nor are they constrained to be linearly independent. Nonetheless, in some applications we will find it useful to impose an orthogonality constraint. A subspace QSP measurement on $\mathcal{H}$ is defined by a set of projection operators E_i onto subspaces $\mathcal{S}_i$ in $\mathcal{H}$. Here again, since we are not constrained by the physics, the projection operators and the subspaces $\mathcal{S}_i$ are not constrained to be orthogonal. The measurement of a signal x is denoted by $M(x)$.

Measurement consistency in our framework is formulated mathematically as

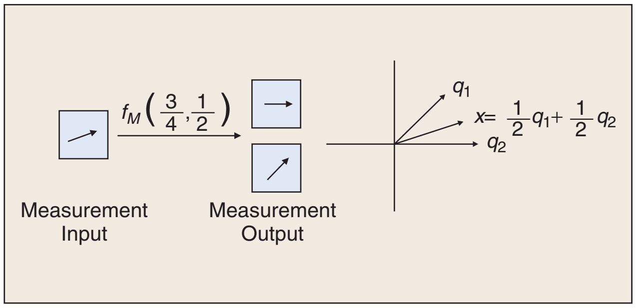
$$M(M(x)) = M(x). \tag{7}$$

Note that by our definition of measurement, if x is a signal in a signal space $\mathcal{H}$ then $M(x)$ is also a signal in $\mathcal{H}$ and can therefore be remeasured.

Quantization of the measurement outcome is imposed by requiring that the outcome signal $M(x)$ is one of a set of signals determined by the measurement M. Specifically, in analogy with the quantum mechanical determinate states we define the set of determinate signals,



▲ 4. *Illustration of a class of algorithms within the QSP framework. In this framework quantum physics is used as a metaphor to design new signal processing algorithms by drawing a parallel between a signal processing algorithm and a quantum mechanical measurement. One class of algorithms is designed by constructing a QSP measurement borrowing from the principles of a quantum measurement, which is then translated into a signal processing algorithm using appropriate input and output mappings.*



▲ 5. Illustration of a ROM.

which are the signals that lie completely in one of the measurement subspaces $\mathcal{S}_i$.

The measurement M is then defined to preserve the two fundamental properties of a quantum measurement.

▲ The measurement outcome is always equal to one of the determinate signals.

▲ For every input signal x to a QSP measurement, (7) is satisfied.

Rank-One QSP Measurements

A ROM M defined by a set of measurement vectors q_i that span one-dimensional measurement subspaces $\mathcal{S}_i$ in $\mathcal{H}$ is in general a nonlinear mapping between $\mathcal{H}$ and the set of determinate signals of M. The measurement subspace $\mathcal{S}_i$ is the one-dimensional subspace that contains all vectors that are multiples of q_i. Therefore, the determinate signals in this case are all signals of the form aq_i for some index i and scalar a. With E_i denoting a projection onto $\mathcal{S}_i$, the measurement is defined such that if x is a determinate signal then $M(x) = E_i x = x$, and otherwise $M(x) = E_i x$ where

$$i = f_M\left(\{\langle x, q_k\rangle\}\right). \tag{8}$$

Here f_M is a mapping between the input signal x and the set of indices that may be probabilistic (see "Quantum Measurement") and that depends on the input x only through the inner products between x and the measurement vectors q_i, which are a subset of the determinate signals. For example, we may choose $f_M(x) = \arg\max_k \langle x, q_k\rangle$. As another example, we may choose $f_M(x) = \arg\max_k (\langle q_k, x\rangle - \langle q_k, q_k\rangle)$. This mapping f_M chooses the vector q_k that minimizes the distance $\|x - q_k\|$ between x and each of the measurement vectors, so that it maps x to the closest vector q_k.

Note that since $E_i x$ is in $\mathcal{S}_i$ for any x, the outcome $M(x)$ is always a determinate signal of M, and since for any determinate signal x, $M(x) = x$, this definition of a measurement satisfies the required properties.

As an example of a ROM, suppose that the measurement input is $x = (1/2)q_1 + (1/2)q_2$ where $q_1 = [1 \quad 0]^T$ and $q_2 = [.5 \quad .5]^T$ with $[\cdot]^T$ denoting the conjugate transpose, are two measurement vectors. Then the measurement output will be either a vector in the direction of q_1 or a vector in the direction of q_2. The particular output chosen is determined by the mapping f_M that depends on the input x only through the inner products $\langle x, q_1\rangle = 3/4$ and $\langle x, q_2\rangle = 1/2$. The measurement process is illustrated in Fig. 5.

Our definition of a ROM is very similar to the definition of a rank-one quantum measurement, with two main differences. First, we allow for an arbitrary mapping f_M in (8); as described in "Quantum Measurement," in quantum mechanics f_M is unique and is the probabilistic mapping in which q_i is chosen with probability $|\langle x, q_i\rangle|^2$. Second, the measurement vectors are not constrained to be orthonormal nor are they constrained to be linearly independent, as in quantum mechanics. The properties of a rank-one quantum measurement and a ROM are summarized in Table 1.

Subspace Measurements

The definition of a subspace QSP measurement parallels that of a ROM and borrows from the definition of a subspace quantum measurement.

A subspace QSP measurement M defined by a set of measurement projections E_i that span measurement subspaces $\mathcal{S}_i$ in $\mathcal{H}$ is a nonlinear mapping between $\mathcal{H}$ and the set of determinate signals of M where if x is a determinate signal then $M(x) = E_i x = x$, and otherwise $M(x) = E_i x$ where

$$i = f_M\left(\{\langle E_k x, E_k x\rangle\}\right). \tag{9}$$

Here f_M is a (possibly probabilistic) mapping between the input signal x and the set of indices that depends on the input x only through the inner products $\{\langle E_k x, E_k x\rangle\}$.

A special case of a subspace QSP measurement is the case in which the measurement is defined by a single projection. Then $M(x) = Ex$ for all x and the subspace QSP measurement reduces to a linear projection operator. We refer to such a measurement as a simple subspace measurement.

Table 1. Comparison Between Rank-One Measurements.

	Quantum Measurement	QSP Measurement
Input x	Vector in $\mathcal{H}$	Vector in $\mathcal{H}$
Measurement vectors q_k	Orthonormal	Any vectors in $\mathcal{H}$
Selection rule	Function of $< x, q_k >$	Function of $< x, q_k >$
	Probabilistic	Deterministic or probabilistic
Output	Multiple of q_k for one value k	Multiple of q_k for one value k

The subspace QSP measurement is very similar to a subspace quantum measurement, with three main differences: we allow for an arbitrary mapping f_M, the measurement projections are not constrained to be orthogonal, and the measurement subspaces are not constrained to be orthogonal.

The properties of a subspace quantum measurement and a subspace QSP measurement are summarized in Table 2.

Algorithm Design in the QSP Framework

Within the QSP framework, the QSP measurement plays a central role in the design of signal processing algorithms. In this framework, signals are processed by either subjecting them to a QSP measurement or by using some of the QSP measurement parameters f_M, q_i, and E_i but not directly applying the measurement, as described in the following sections.

Designing Algorithms Based on QSP Measurements

Algorithm Design

To design an algorithm using a QSP measurement we first identify the measurement vectors q_i in a ROM, or the measurement projection operators E_i in a subspace QSP measurement, that specify the possible measurement outcomes. For example, in a detection scenario the measurement vectors may be equal to the transmitted signals or may represent these signals in a possibly different space. As another example, in a scalar quantizer the measurement vectors may be chosen as a set of vectors that represent the scalar quantization levels. In a subspace QSP measurement, the measurement projection operators may be projections onto a set of subspaces used for signaling. We then embed the measurement vectors (projections) in a Hilbert space $\mathcal{H}$. This basic strategy is illustrated in Fig. 6. If the signal $\tilde{x}$ to be processed does not lie in $\mathcal{H}$, then we first map it into a signal x in $\mathcal{H}$ using a mapping T_x. To obtain the algorithm output we measure the representation x of the signal to be processed. If x is a determinate signal of M, then the measurement outcome is $y = M(x) = x$. Otherwise we approximate x by a determinate signal y using a mapping f_M. If appropriate, the measurement outcome y may be mapped to the algorithm output $\tilde{y}$ using a mapping T_y.

By choosing different input and output mappings T_x and T_y and different measurement parameters f_M, q_i, and E_i, and using the QSP measurement framework of Fig. 6, we can arrive at a variety of new and interesting processing techniques.

Table 2. Comparison Between Subspace Measurements.		
	Quantum Measurement	**QSP Measurement**
Input x	Vector in $\mathcal{H}$	Vector in $\mathcal{H}$
Measurement projections E_k	Orthogonal	Can be nonorthogonal
Measurement subspaces $\mathcal{S}_k$	Orthogonal	Can be nonorthogonal
Selection rule	Function of $< E_k x, E_k x >$	Function of $< E_k x, E_k x >$
	Probabilistic	Deterministic or probabilistic
Output	Multiple of $E_k x$ for one value k	Multiple of $E_k x$ for one value k

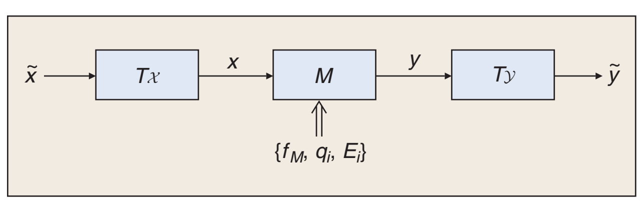
Modifying Known Algorithms

As demonstrated in [1], many traditional detection and processing techniques fit naturally into the framework of Fig. 6. Examples include traditional and dithered quantization, sampling methods, matched-filter detection, and multiuser detection. Once an algorithm is described in the form of a QSP measurement, modifications and extensions of the algorithm can be derived by changing the measurement parameters f_M, q_i, and E_i. Thus, the QSP framework provides a unified conceptual structure for a variety of traditional processing techniques and a precise mathematical setting for generating new, potentially effective, and efficient processing methods by modifying the measurement parameters.

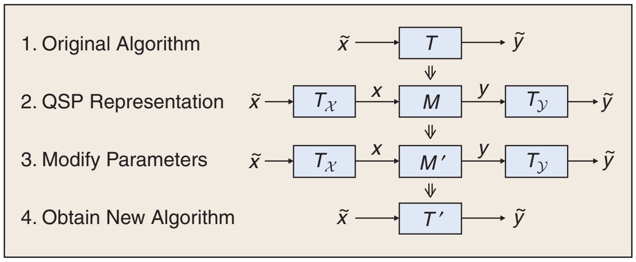
To modify an existing algorithm represented by a mapping T using the QSP framework, we first cast the algorithm as a QSP measurement M, i.e., we choose an input mapping T_x and an output mapping T_y if appropriate, and the measurement parameters f_M, q_i, and E_i. We then systematically change some of these parameters, resulting in a modified measurement M', which can then be translated into a new signal processing algorithm represented by a mapping T'. The modifications we consider result from either imposing some of the additional constraints of quantum mechanics on the measurement parameters of M or from relaxing some of these constraints which we do not have to impose in signal processing. These basic steps are summarized in Fig. 7.

Typical modifications of the parameters that we consider include

- using a probabilistic mapping f_M
- imposing inner product constraints on the measurement vectors q_i



▲ 6. Designing algorithms using a QSP measurement.



▲ 7. Using the QSP measurement framework to modify existing signal processing algorithms.

▲ using nonorthogonal (oblique) projections E_i in place of orthogonal projections.

Designing Algorithms Using the Measurement Parameters

Another class of algorithms we develop results from processing a signal with some of the measurement parameters and then imposing quantum mechanical constraints directly on these parameters. For example, we may view any linear processing of a signal as processing with a set of measurement vectors and then imposing inner product constraints on these vectors. Using the ideas of quantum detection we may then design linear algorithms that are optimal subject to these inner product constraints.

To generate new algorithms or modify existing ones, we describe the algorithm as processing by one of the measurement parameters and then modifying these parameters using one of the three modifications outlined at the end of the previous section.

In the remainder of this section we discuss each of these modifications and indicate how they will be applied to the development of new processing methods.

Probabilistic Mappings

The QSP framework naturally gives rise to probabilistic and randomized algorithms by letting f_M be a probabilistic mapping, emulating the quantum measurement. We expand on this idea below in the context of quantization. Further examples are developed in [1]. However, the full potential benefits of probabilistic algorithms in general resulting from the QSP framework remain an interesting area of future study.

Imposing Inner Product Constraints

One of the important elements of quantum mechanics is that the measurement vectors are constrained to be orthonormal. This constraint leads to some interesting problems such as the quantum detection problem described previously. A fundamental problem in quantum mechanics is to construct optimal measurements subject to this constraint that best represent a given set of state vectors. In analogy to quantum mechanics, an important feature of QSP is that of imposing particular types of constraints on algorithms. The QSP framework provides a systematic method for imposing such constraints: the measurement vectors are restricted to have a certain inner product structure, as in quantum mechanics. However, since we are not limited by physical laws, we are not confined to an orthogonality constraint. As part of the QSP framework, we develop methods for choosing a set of measurement vectors that "best" represent the signals of interest and have a specified inner product structure [37], [1]; these methods rely on ideas and results we obtained in the context of quantum detection [32], which unlike QSP are subject to the constraints of quantum physics. Specifically, we construct measurement vectors q_i with a given inner product structure that are closest in an LS sense to a given set of vectors s_i, so that the vectors q_i are chosen to minimize the sum of the squared norms of the error vectors $e_i = q_i - s_i$. These techniques are referred to as LS inner product shaping. Further details on LS inner product shaping are summarized in "Least Squares Inner Product Shaping."

As we show, the concept of LS inner product shaping can be used to develop effective solutions to a variety of problems that result from imposing a deterministic or stochastic inner product constraint on the algorithm and then designing optimal algorithms subject to this constraint. In each of these problems we either describe the algorithm as a QSP measurement and impose an inner product constraint on the corresponding measurement vectors or we consider linear algorithms on which the inner product constraints can be imposed directly. We demonstrate that, even for problems without inherent inner product constraints, imposing such constraints in combination with LS inner product shaping leads to new processing techniques in diverse areas including frame theory, detection, covariance shaping, linear estimation, and multiuser wireless communication, which often exhibit improved performance over traditional methods.

Oblique Projections

In a quantum measurement defined by a set of projection operators, the rules of quantum mechanics impose the constraint that the projections must be orthogonal. In QSP we may explore more general types of measurements defined by projection operators that are not restricted to be orthogonal, i.e., oblique projections [38]-[40].

An oblique projection is a projection operator E satisfying $E^2 = E$ that is not necessarily Hermitian, i.e., the range space and null space of E are not necessarily orthogonal spaces. The notation $E_{\mathcal{US}}$ denotes an oblique projection with range space $\mathcal{U}$ and null space $\mathcal{S}$. If $\mathcal{S} = \mathcal{U}^{\perp}$, then $E_{\mathcal{US}}$ is an orthogonal projection onto $\mathcal{U}$. An oblique projection $E_{\mathcal{US}}$ can be used to decompose x into its components in two disjoint vector spaces $\mathcal{U}$ and $\mathcal{S}$ that are not constrained to be orthogonal, as illustrated in Fig. 8. (Two subspaces are said to be disjoint if the only vector they have in common is the zero vector.)

Oblique projections are used within the QSP framework to develop new classes of frames, effective subspace

Least-Squares Inner Product Shaping

Suppose we are given a set of m vectors $\{s_i, 1 \le i \le m\}$ in a Hilbert space $\mathcal{H}$, with inner product $\langle x, y \rangle$ for any $x, y \in \mathcal{H}$. The vectors s_i span an n-dimensional subspace. If the vectors are linearly independent, then $n = m$; otherwise $n < m$. Our objective is to construct a set of optimal vectors $\{h_i, 1 \le i \le m\}$ with a specified inner product structure, from the given vectors $\{s_i, 1 \le i \le m\}$. Specifically, we seek the vectors h_i that are "closest" to the vectors s_i in the LS sense. Thus, the vectors are chosen to minimize

$$\varepsilon_{LS} = \sum_{i=1}^{m} \langle s_i - h_i, s_i - h_i \rangle, \tag{10}$$

subject to the constraint

$$\langle h_i, h_k \rangle = c^2 r_{ik}, \tag{11}$$

for some $c > 0$ and constants r_{ik}.

We may wish to constrain the constant c in (11) or may choose c such that the LS error ε_{LS} is minimized. Similarly, with $\mathbf{R}$ denoting the matrix with ikth element r_{ik}, we may wish to constrain the elements r_{ik} of $\mathbf{R}$, or we may choose $\mathbf{R}$ to have a specified structure so that the eigenvectors of $\mathbf{R}$ are fixed, but choose the eigenvalues to minimize the LS error.

For example, we may wish to construct a set of orthogonal vectors, so that $\mathbf{R}$ is a diagonal matrix with eigenvector matrix equal to $\mathbf{I}$, but choose the norms of the vectors, i.e., the eigenvalues of $\mathbf{R}$, to minimize the LS error. As another example, we may wish to construct a cyclic set h_i so that $\langle h_i, h_k \rangle$ depends only on $k - i \bmod m$. In this case $\{\langle h_i, h_k \rangle, 1 \le i \le m\}$ is a cyclic permutation of $\{\langle h_1, h_k \rangle, 1 \le k \le m\}$ for all k, and the Gram matrix $H^* H$ is a circulant matrix diagonalized by a DFT matrix, so that the eigenvectors of $\mathbf{R}$ are fixed. (In [37] we show that a vector set has a circulant Gram matrix if and only if the set is cyclic.) We may then wish to specify the values $\{\langle h_1, h_k \rangle, 1 \le k \le m\}$ (possibly up to a scale factor), which corresponds to specifying the eigenvalues of $\mathbf{R}$, or we may choose these values, equivalently the eigenvalues of $\mathbf{R}$, so that the LS error is minimized.

In [1] we consider both the case in which $\mathbf{R}$ is fixed and the case in which the eigenvalues of $\mathbf{R}$ are chosen to minimize the LS error ε_{LS}. As we show, for fixed $\mathbf{R}$ the LS inner product shaping problem has a simple closed form solution; by contrast, if the eigenvalues of $\mathbf{R}$ are not specified, then there is no known analytical solution to the LS inner product shaping problem for arbitrary vectors s_i. An iterative algorithm is developed in [1]. In the simplest case where $\mathbf{R}$ is a specified rank-r matrix, with $r = n = m$, the LS vectors are the "columns" of the (possibly infinite) matrix

$$H = \tilde{\alpha} S(\mathbf{R} S^* S)^{-1/2} \mathbf{R} = \tilde{\alpha} S \mathbf{R} (S^* S \mathbf{R})^{-1/2}, \tag{12}$$

where S is the matrix of columns s_i. If c in (11) is specified then $\tilde{\alpha} = c$, and if c is chosen to minimize the LS error then

$$\tilde{\alpha} = \frac{\mathrm{Tr}\left((S^* S \mathbf{R})^{1/2} \right)}{\mathrm{Tr}(\mathbf{R})}, \tag{13}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace of the corresponding matrix.

If $r = n$ and $\mathcal{N}(S) = \mathcal{N}(\mathbf{R})$ where $\mathcal{N}(\cdot)$ denotes the null space of the corresponding transformation, then

$$H = \tilde{\alpha} S\left((\mathbf{R} S^* S)^{1/2} \right)^{\dagger} \mathbf{R} = \tilde{\alpha} S \mathbf{R} \left((S^* S \mathbf{R})^{1/2} \right)^{\dagger}, \tag{14}$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudoinverse.

detectors, and a general sampling framework for sampling and reconstruction in arbitrary spaces.

Applications of Rank-One Measurements

QSP Quantization

By emulating the quantum measurement, in [1] we develop a probabilistic quantizer and show that it can be used to efficiently implement a dithered quantizer.

In dithered quantization a random signal called a dither signal is added to the input signal prior to quantization [41]-[44]. Dithering techniques have become commonplace in applications in which data is quantized prior to storage or transmission. However, the utility of dithering techniques is limited by the computational complexity associated with generating a random process with an arbitrary joint probability distribution.

A probabilistic quantizer can be used to effectively realize a dither signal with an arbitrary joint probability distribution, while requiring only the generation of one uniform random variable per input. By introducing memory into the probabilistic selection rule we derive a probabilistic quantizer that shapes the quantization noise.

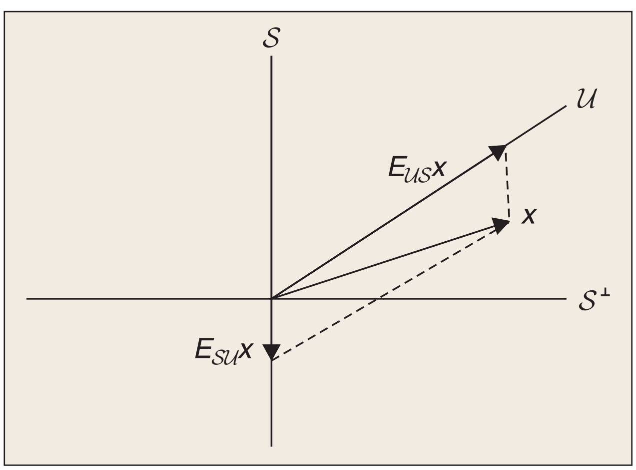
Covariance Shaping Matched Filter Detection

As an example of the type of procedure we may follow in using the concept of LS inner product shaping and optimal QSP measurements to derive new processing methods, in [45] and [46] we consider a generic detection problem where one of a set of signals is transmitted over a noisy channel. When the additive noise is white and Gaussian, it is well known (see, e.g., [15] and [47]) that the receiver that maximizes the probability of correct detection is the matched filter (MF) receiver. If the noise is not Gaussian, then the MF receiver does not necessarily maximize the probability of correct detection. However, it is still used as the receiver of choice in many applications since the optimal detector for non-Gaussian noise is typically nonlinear (see, e.g., [48] and references therein) and depends on the noise distribution, which may not be known.

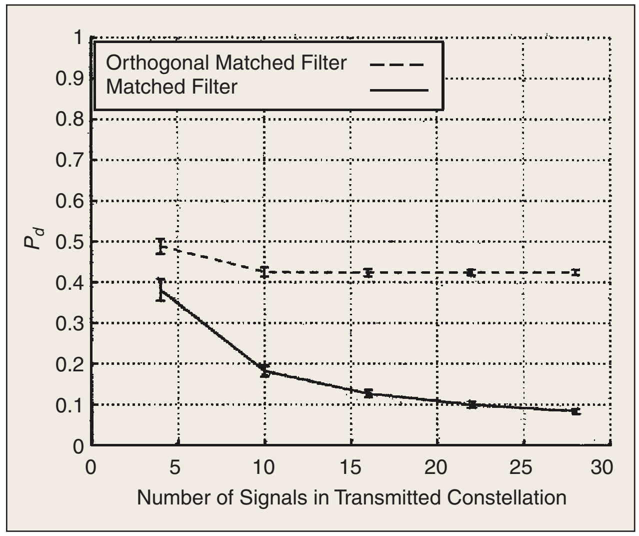
By describing the MF detector as a QSP measurement and imposing an inner product constraint on the measurement vectors, we derive a new class of receivers consisting of a bank of correlators with correlating signals that are matched to a set of signals with a specified inner product structure **R** and are closest in an LS sense to the transmitted signals. These receivers depend only on the transmitted signals, so that they do not require knowledge of the noise distribution or the channel signal-to-noise ratio (SNR). We refer to these receivers as *covariance shaping MF receivers* [1]. In the special case in which $\mathbf{R} = \mathbf{I}$, the receiver consists of a bank of correlators with orthogonal correlating signals that are closest in an LS sense to the transmitted signals and is referred to as the *orthogonal matched filter (OMF) receiver* [45], [46].

Alternatively, we show that the modified receivers can be implemented as an MF demodulator followed by an optimal covariance shaping transformation that optimally shapes the correlation of the outputs of the MF prior to detection. This equivalent representation leads to the concept of minimum mean-squared error (MMSE) covariance shaping, which we consider in its most general form in [1] and [49]. Simulations presented in [45] and [46] show that when the noise is non-Gaussian this approach can lead to improved performance over conventional MF detection in many cases, with only a minor impact in performance when the noise is Gaussian. These results are encouraging since they suggest that if the receiver is designed to operate in different noise environments or in an unknown noise environment, then we may prefer using the modified detectors since for certain non-Gaussian noise distributions these detectors may result in an improvement in performance over an MF detector, without significantly degrading the performance if the noise is Gaussian.

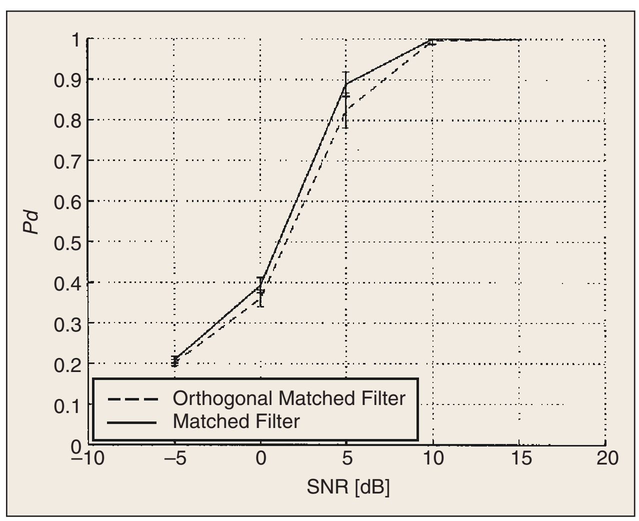
In Fig. 9 we illustrate the performance advantage with one simulation from [45]. In this figure we plot the mean



▲ 8. Decomposition of x into its components in $\mathcal{U}$ and in $\mathcal{S}$ given by $E_{\mathcal{US}}x$ and $E_{\mathcal{SU}}x$, respectively.



▲ 9. Comparison between the OMF and MF detectors in Beta-distributed noise, as a function of the number of signals in the transmitted constellation. The parameters of the distribution are $a = b = 0.1$, and the SNR is 0 dB. The dashed line is the mean P_d using the OMF detector, and the solid line is the mean P_d using the MF detector. The vertical bars indicate the standard deviation of the corresponding P_d.



▲ 10. Comparison between the OMF and MF detectors for transmitted constellations of seven signals in Gaussian noise, as a function of SNR. The dashed line is the mean P_d using the OMF detector, and the solid line is the mean P_d using the MF detector. The vertical bars indicate the standard deviation of the corresponding P_d.

and standard deviation of the probability of correct detection P_d for the OMF detector with $\mathbf{R} = \mathbf{I}$ in Beta-distributed noise, as a function of the number of signals in the transmitted constellation. The dimension m of the signals in the constellation is equal to the number of signals and the samples of the signals are mutually independent zero-mean Gaussian random variables with variance $1/\sqrt{m}$, scaled to have norm one. The vertical bars indicate the standard deviation of P_d. The results in the figure were obtained by generating 500 realizations of signals. For each particular signal realization, we determined the probability of correct detection for the detectors in both types of noise by recording the number of successful detections over 500 noise realizations. From the figure it is evident that the OMF detector outperforms the MF detector, particularly when the probability of correct detection with the MF is marginal. The relative improvement in performance of the OMF detector over the MF detector increases for increasing constellation size. In Fig. 10 we plot the mean of P_d for the OMF detector and the MF detector in Gaussian noise, for transmitted constellations of seven signals in Gaussian noise, as a function of SNR. Again, the vertical bars indicate the standard deviation of P_d.

Optimal Covariance Shaping

Drawing from the quantum detection problem, we can develop new classes of linear algorithms that result from imposing a deterministic or stochastic inner product constraint on the algorithm, i.e., a covariance constraint, and then using the results we obtained in the context of quantum detection to derive optimal algorithms subject to this constraint. In particular, we may extend the concept of LS inner product shaping suggested by the quantum detection framework to develop optimal algorithms that minimize a stochastic mean-squared error (MSE) criterion subject to a covariance constraint.

As an example of this approach, in [1] we exploit the concept of LS inner product shaping to develop a new viewpoint towards whitening and other covariance shaping problems.

Suppose we have a random vector $\mathbf{a}$ that lies in the m-dimensional space $\mathbb{C}^m$ with covariance $\mathbf{C}_a$, and we want to shape the covariance of the vector $\mathbf{a}$ using a shaping transformation $\mathbf{T}$ to obtain the random vector $\mathbf{b} = \mathbf{T}\mathbf{a}$, where the covariance matrix of $\mathbf{b}$ is given by $\mathbf{C}_b = c^2 \mathbf{R}$ for some $c > 0$ and some covariance matrix $\mathbf{R}$. Thus we seek a transformation $\mathbf{T}$ such that

$$\mathbf{C}_b = \mathbf{T}\mathbf{C}_a\mathbf{T}^* = c^2 \mathbf{R}, \tag{15}$$

for some $c > 0$.

Data shaping arises in a variety of contexts in which it is useful to shape the covariance of a data vector either prior to subsequent processing or to control the spectral shape after processing [50], [51]. As is well known, given a covariance matrix $\mathbf{C}_a$, there are many ways to choose a shaping transformation $\mathbf{T}$ satisfying (15). While in some applications certain conditions might be imposed on the transformation such as causality or symmetry, with the exception of the work in [46], [45], [52], [53], [49], and [54], which explicitly relies on the optimality properties developed in [1], there have been no general assertions of optimality for various choices of a linear shaping transformation. In particular, the shaped vector may not be "close" to the original data vector. If this vector undergoes some noninvertible processing or is used as an estimator of some unknown parameters represented by the data, then we may wish to choose the covariance shaping transformation so that the shaped output is close to the original data in some sense.

Building upon the concept of LS inner product shaping, we propose choosing an optimal shaping transformation that results in a shaped vector $\mathbf{b}$ that is as close as possible to the original vector $\mathbf{a}$ in an MSE sense, which we refer to as MMSE covariance shaping. Specifically, among all possible transformations we seek the one that minimizes the total MSE given by

$$\varepsilon_{\mathrm{MSE}} = \sum_{i=1}^{m} E\left((a_i - b_i)^2\right) = E\left((\mathbf{a}-\mathbf{b})^*(\mathbf{a}-\mathbf{b})\right), \tag{16}$$

subject to (15), where a_i and b_i are the ith components of $\mathbf{a}$ and $\mathbf{b}$, respectively.

The solution to the MMSE covariance shaping problem is developed in [1]; further details are summarized in "MMSE Covariance Shaping." This new concept of MMSE shaping can be useful in a variety of signal processing methods that incorporate shaping transformations in which we can imagine using an optimal procedure that shapes the data but at the same time minimizes the distortion to the original data.

Linear Estimation

As another example of an algorithm suggested by the quantum detection framework, where we use the ideas of LS inner product shaping to design an optimal linear algorithm subject to a stochastic inner product constraint, in [54] we derive a new linear estimator for the unknown deterministic parameters in a linear model. The estimator is chosen to minimize an MSE criterion, subject to a constraint on the covariance of the estimator. This new estimator is defined as the *covariance shaping LS (CSLS)* estimator.

Many estimation problems can be represented by the linear model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$, where $\mathbf{H}$ is a known matrix, $\mathbf{x}$ is a vector of unknown deterministic parameters to be estimated, and $\mathbf{w}$ is a random vector with covariance $\mathbf{C}_w$. A common approach to estimating the parameters $\mathbf{x}$ is to restrict the estimator to be linear in the data $\mathbf{y}$ and then find the linear estimate of $\mathbf{x}$ that results in an estimated data vector that is as close as possible to the given data vector $\mathbf{y}$ in a (weighted) LS sense, so that it minimizes the total squared error in the observations [55]-[58]. It is well

MMSE Covariance Shaping

In [1] we show that the MMSE covariance shaping problem can be interpreted as an LS inner product shaping problem, so that the MMSE shaping transformation can be found by applying results derived in that context. In the simplest case in which $\mathbf{C}_a$ and $\mathbf{R}$ are invertible, the optimal shaping transformation is given by

$$\mathbf{T} = \alpha\mathbf{R}(\mathbf{C}_a\mathbf{R})^{-1/2} = \alpha(\mathbf{R}\mathbf{C}_a)^{-1/2}\mathbf{R}, \tag{17}$$

where if c in (15) is specified then $\alpha = c$, and if c is chosen to minimize the MSE then

$$\alpha = \frac{\mathrm{Tr}((\mathbf{C}_a\mathbf{R})^{1/2})}{\mathrm{Tr}(\mathbf{R})}. \tag{18}$$

Here $(\cdot)^{1/2}$ denotes the unique nonnegative-definite square root of the corresponding matrix.

When $\mathbf{R} = \mathbf{I}$ so that $\mathbf{T}$ is a whitening transformation, (17) reduces to

$$\mathbf{T} = \alpha\mathbf{C}_a^{-1/2} \tag{19}$$

and (18) becomes

$$\alpha = \frac{1}{m}\mathrm{Tr}\left(\mathbf{C}_a^{1/2}\right). \tag{20}$$

It is interesting to note that the MMSE whitening transformation has the additional property that it is the unique *symmetric* whitening transformation (up to a factor of ± 1) [103]. It is also proportional to the Mahalanobis transformation, which is frequently used in signal processing applications incorporating whitening (see, e.g., [57], [50], and [51]).

Another interesting case is when the random vector $\mathbf{a}$ is white so that $\mathbf{C}_a = \mathbf{I}$ and the problem is to optimally shape its covariance. The MMSE transformation in this case is

$$\mathbf{T} = \alpha\mathbf{R}^{1/2}, \tag{21}$$

where if c is fixed then $\alpha = c$, and if c is chosen to minimize the MSE then

$$\alpha = \frac{\mathrm{Tr}(\mathbf{R}^{1/2})}{\mathrm{Tr}(\mathbf{R})}. \tag{22}$$

We may also consider a weighted MMSE covariance shaping problem in which the shaping $\mathbf{T}$ is chosen to minimize a weighted MSE. Thus we seek a transformation $\mathbf{T}$ such that $\mathbf{b} = \mathbf{T}\mathbf{A}$ has covariance $\mathbf{C}_b = c^2\mathbf{R}$ for some $c > 0$ and such that

$$\varepsilon^{w}_{\mathrm{MSE}} = E((\mathbf{a}-\mathbf{b})^*\mathbf{A}(\mathbf{a}-\mathbf{b})), \tag{23}$$

is minimized, where $\mathbf{A}$ is some nonnegative-definite Hermitian weighting matrix. In the simplest case in which $\mathbf{C}_a$, $\mathbf{R}$, and $\mathbf{A}$ are all invertible, the weighted MMSE covariance shaping transformation is

$$\mathbf{T} = \alpha(\mathbf{R}\mathbf{A}\mathbf{C}_a\mathbf{A})^{-1/2}\mathbf{R}\mathbf{A}, \tag{24}$$

where if c is specified then $\alpha = c$, and if c is chosen to minimize the weighted MSE then

$$\alpha = \frac{\mathrm{Tr}((\mathbf{R}\mathbf{A}\mathbf{C}_a\mathbf{A})^{1/2})}{\mathrm{Tr}(\mathbf{R}\mathbf{A})}. \tag{25}$$

known that among all possible unbiased linear estimators, the LS estimator minimizes the variance [56]. However, this does not imply that the resulting variance or MSE is small, where the MSE of an estimator is the sum of the variance and the squared norm of the bias. In particular, a difficulty often encountered when using the LS estimator to estimate the parameters $\mathbf{x}$ is that the error in estimating $\mathbf{x}$ can have a large variance and a covariance structure with a very high dynamic range. This is due to the fact that in many cases the data vector $\mathbf{y}$ is not very sensitive to changes in $\mathbf{x}$, so that a large error in estimating $\mathbf{x}$ may translate into a small error in estimating the data vector $\mathbf{y}$, in which case the LS estimate may result in a poor estimate of $\mathbf{x}$. This effect is especially predominant at low to moderate SNR, where the data vector $\mathbf{y}$ is typically affected more by the noise than by changes in $\mathbf{x}$; the exact SNR range will depend on the properties of the model matrix $\mathbf{H}$.

The CSLS estimator, denoted by $\hat{\mathbf{x}}_{\mathrm{CSLS}}$, is a biased estimator directed at improving the performance of the traditional LS estimator at low to moderate SNR by choosing the estimate to minimize the (weighted) total error variance in the observations subject to a constraint on the covariance of the estimation error, so that we control the dynamic range and spectral shape of the covariance of the estimation error. Thus, $\hat{\mathbf{x}}_{\mathrm{CSLS}} = \mathbf{G}\mathbf{y}$ is chosen to minimize

$$\varepsilon_{\mathrm{CSLS}} = E\left((\mathbf{y}'-\mathbf{H}\mathbf{G}\mathbf{y}')^*\mathbf{C}_w^{-1}(\mathbf{y}'-\mathbf{H}\mathbf{G}\mathbf{y}')\right), \tag{26}$$

where $\mathbf{y}' = \mathbf{y} - E(\mathbf{y})$, subject to the constraint that the covariance of the error in the estimate $\hat{\mathbf{x}}_{\mathrm{CSLS}}$, which is equal to the covariance of the estimate $\hat{\mathbf{x}}_{\mathrm{CSLS}}$, is proportional to a given covariance matrix $\mathbf{R}$. Thus $\mathbf{G}$ must satisfy

$$\mathbf{G}\mathbf{C}_w\mathbf{G}^* = c^2\mathbf{R}, \tag{27}$$

Covariance Shaping Least-Squares Estimation

If the model matrix **H** has full column rank, then the covariance shaping LS estimator that minimizes (26) subject to (27) is given by

$$\hat{x}_{CSLS} = \beta(\mathbf{R}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1/2}\mathbf{R}\mathbf{H}^*\mathbf{C}_w^{-1}y, \tag{28}$$

where if c in (27) is specified then $\beta = c$, and if c is chosen to minimize the error then

$$\beta = \frac{\mathrm{Tr}((\mathbf{R}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{1/2})}{\mathrm{Tr}(\mathbf{R}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})}. \tag{29}$$

The CSLS estimator can also be expressed as an LS estimator followed by a weighted MMSE covariance shaping transformation. Specifically, suppose we estimate the parameters x using the LS estimate $\hat{x}_{LS}$. Since $\hat{x}_{LS} = x + \tilde{w}$ where $\tilde{w} = (\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}\mathbf{H}^*\mathbf{C}_w w$, the covariance of the noise component $\tilde{w}$ in $\tilde{x}_{LS}$ is equal to the covariance of $\hat{x}_{LS}$, denoted $\mathbf{C}_{\hat{x}_{LS}}$, which is given by $\mathbf{C}_{\hat{x}_{LS}} = \sigma^2(\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{-1}$. To improve the performance of the LS estimator, we may shape the covariance of the noise component in the estimator $\hat{x}_{LS}$. Thus we seek a transformation **W** such that the covariance matrix of $\hat{x} = \mathbf{W}\hat{x}_{LS}$, denoted by $\mathbf{C}_{\hat{x}}$, satisfies $\mathbf{C}_{\hat{x}} = \mathbf{W}\mathbf{C}_{\hat{x}_{LS}}\mathbf{W}^* = c^2\mathbf{R}$, for some $c > 0$. To minimize the distortion to the estimator $\hat{x}_{LS}$, we choose the transformation **W** that minimizes the weighted MSE

$$E\left((\hat{x}'_{LS} - \mathbf{W}\hat{x}'_{LS})^*\mathbf{C}_{\hat{x}_{LS}}^{-1}(\hat{x}'_{LS} - \mathbf{W}\hat{x}'_{LS})\right). \tag{30}$$

As we show in [54], the resulting estimator $\hat{x}$ is equal to $\hat{x}_{CSLS}$, so that the CSLS estimator can be determined by first finding the LS estimator $\hat{x}_{LS}$ and then optimally shaping its covariance. The CSLS estimator with fixed scaling can also be expressed as a matched correlator estimator followed by MMSE shaping.

If the noise w in the model $\mathbf{y} = \mathbf{H}x + \mathbf{w}$ is Gaussian with zero-mean and covariance $\mathbf{C}_w$, then the CSLS estimator achieves the CRLB for biased estimators with bias B equal to the bias of the CSLS estimator, which is given by

$$B = \left(\beta(\mathbf{R}\mathbf{H}^*\mathbf{C}_w^{-1}\mathbf{H})^{1/2} - \mathrm{I}\right)x. \tag{31}$$

Thus, from all estimators with bias given by (31) for some β and **R**, the CSLS estimator minimizes the variance.

While it would be desirable to analyze the MSE of the CSLS estimator for more general forms of bias, we cannot directly evaluate the MSE of the CSLS estimator since the bias, and consequently the MSE, depend explicitly on the unknown parameters x. Instead, we may compare the MSE of the CSLS estimator with the MSE of the LS estimator. The analysis in [54] indicates that there are many cases in which the CSLS estimator performs better than the LS estimator in a MSE sense, for all values of the unknown parameters x. Specifically, for a variety of choices of the output covariance **R**, there is a threshold SNR, such that for SNR values below this threshold the CSLS estimator yields a lower MSE than the LS estimator, for all values of x.

If the output covariance is chosen to be equal to $\sigma\mathbf{R}$, where $\mathbf{C}_w = \sigma^2\mathbf{C}$ for some covariance matrix **C**, then it can be shown that there is a threshold SNR value so that for SNR values below this threshold, the MSE of the CSLS estimator is always smaller than the MSE of the LS estimator, regardless of the value of x. Specifically, let $\zeta = \|x\|^2/(\sigma^2 m)$ denote the SNR per dimension. Then with $\mathbf{B} = \mathbf{H}^*\mathbf{C}^{-1}\mathbf{H}$, $\gamma = \arg\max \sigma_k$, and σ_k denoting the eigenvalues of $Q = ((\mathbf{RB})^{1/2} - \mathrm{I})^*((\mathbf{RB})^{1/2} - \mathrm{I})$, the MSE of the CSLS estimator is less than or equal to the MSE of LS estimator for $\zeta \le \zeta_{WC}$, where

$$\zeta_{WC} = \frac{\mathrm{Tr}(\mathbf{B}^{-1}) - \mathrm{Tr}(\mathbf{R})}{\sigma_\gamma}. \tag{32}$$

The bound ζ_{WC} is a worst case bound, since it corresponds to the worst possible choice of parameters, namely when the unknown vector x is in the direction of the eigenvector of Q corresponding to the eigenvalue σ_γ. In practice the CSLS estimator will outperform the LS estimator for higher values of SNR than ζ_{WC}.

Since we have freedom in designing **R**, we may always choose **R** so that $\zeta_{WC} > 0$. In this case we are guaranteed that there is a range of SNR values for which the CSLS estimator leads to a lower MSE than the LS estimator for all choices of the unknown parameters x.

For example, suppose we wish to design an estimator with covariance proportional to some given covariance matrix **Z**, so that $\mathbf{R} = a\mathbf{Z}$ for some $a > 0$. If we choose $a < \mathrm{Tr}(\mathbf{B}^{-1})/\mathrm{Tr}(\mathbf{Z})$, then we are guaranteed that there is an SNR range for which the CSLS estimator will have a lower MSE than the LS estimator for all values of x.

In specific applications it may not be obvious how to choose a particular proportionality factor a. In such cases, we may prefer using the CSLS estimator with optimal scaling. In this case, the scaling is a function of **R** and therefore cannot be chosen arbitrarily, so that in general we can no longer guarantee that there is a positive SNR threshold, i.e., that there is always an SNR range over which the CSLS performs better than the LS estimator. However, as shown in [54], in the special case in which $\mathbf{R} = \mathrm{I}$, there is always such an SNR range. Specifically, with $\{\lambda_k, 1 \le k \le m\}$ denoting the eigenvalues of $\mathbf{B} = \mathbf{H}^*\mathbf{C}^{-1}\mathbf{H}$, and $\alpha = (\sum_{k=1}^m \lambda_k^{1/2})/(\sum_{k=1}^m \lambda_k)$, the MSE of the CSLS estimator is less than or equal to the MSE of LS estimator for $\zeta \le \zeta_{WC}$, where

$$\zeta_{WC} = \frac{(1/m)\sum_{k=1}^m \lambda_k^{-1} - \alpha^2}{\left|\alpha\lambda_\gamma^{1/2} - 1\right|^2}, \tag{33}$$

and $\gamma = \arg\max\left|\alpha\lambda_k^{1/2} - 1\right|^2$.

Here again ζ_{WC} is a worst case bound; in practice the CSLS estimator will outperform the LS estimator for higher values of SNR than ζ_{WC}. Examples presented in [54] indicate that in a variety of applications ζ_{WC} can be pretty large.

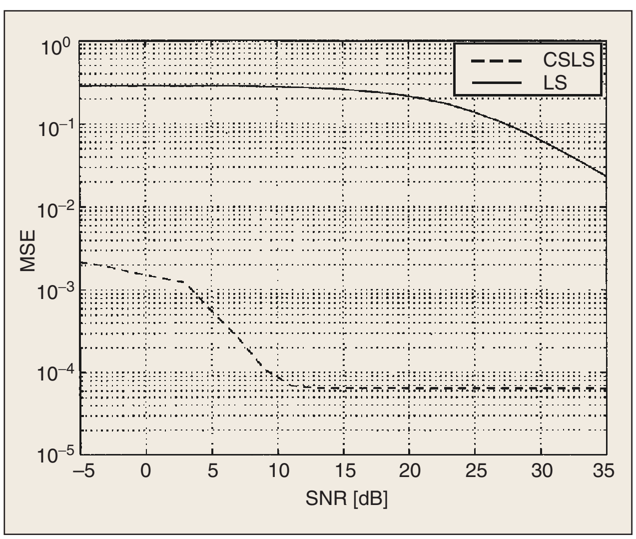
where $c > 0$ is a constant that is either specified or chosen to minimize the error (26).

The CSLS estimator $\hat{\mathbf{x}}_{\text{CSLS}}$ is developed in [1] and [54]. Some of its properties are discussed in "Covariance Shaping Least-Squares Estimation."

Various modifications of the LS estimator under the linear model assumption have been previously proposed in the literature. Among the more prominent alternatives are the ridge estimator [59] (also known as Tikhonov regularization [60]) and the shrunken estimator [61]. In [1] we show that both the ridge estimator and the shrunken estimator can be formulated as CSLS estimators, which allows us to interpret these estimators as the estimators that minimize the total error variance in the observations, from all linear estimators with the same covariance.

As shown in [54], the CSLS estimator has a property analogous to the property of the LS estimator. Specifically, it achieves the Cramer-Rao lower bound (CRLB) for biased estimators [56], [62], [63] when the noise is Gaussian. This implies that for Gaussian noise, there is no linear or nonlinear estimator with a smaller variance, or MSE, and the same bias as the CSLS estimator.

Analysis of the MSE of the CSLS estimator [54] demonstrates that the covariance of the estimation error can be chosen such that over a wide range of SNR, the CSLS estimator results in a lower MSE than the traditional LS estimator, for all values of the unknown parameters. Simulations presented in [1] and [54] strongly suggest that the CSLS estimator can significantly decrease the MSE of the estimation error over the LS estimator for a wide range of SNR values. As an example, in Fig. 11 we plot the MSE in estimating a set of AR parameters in an ARMA model contaminated by white noise, using both the CSLS with $\mathbf{R} = \mathbf{I}$ and the LS estimators from 20 noisy observations of the channel, averaged over 2000 noise realizations, as a function of $-10\log\sigma^2$ where σ^2 is the noise variance. As can be seen from the figure, in this example the CSLS estimator significantly outperforms the LS estimator. In general, the performance advantage using the CSLS estimator will depend on the properties of the model matrix $\mathbf{H}$ and the noise covariance $\mathbf{C}_w$, as indicated by the analysis in [54] and in "Covariance Shaping Least-Squares Estimation."



▲ *11. Mean-squared error in estimating the AR parameters using the LS estimator and the CSLS estimator.*

Multiuser Detection

Based on the concept of CSLS estimation, we propose a new class of linear receivers for synchronous code-division multiple-access (CDMA) systems, which we refer to as the *covariance shaping multiuser (CSMU) receivers*. These receivers depend only on the users' signatures and do not require knowledge of the channel parameters. Nonetheless, over a wide range of these parameters the performance of these receivers can approach the performance of the linear MMSE receiver which is the optimal linear receiver that assumes knowledge of the channel parameters and maximizes the output signal-to-interference ratio. These receivers generalize the recently proposed orthogonal multiuser receiver [52].

Consider the problem of detecting information transmitted by each of the users in an m-user CDMA system. Each user transmits information by modulating a signature sequence. The discrete-time model for the received signal $\mathbf{y}$ is given by [64]

$$\mathbf{y} = \mathbf{SAb} + \mathbf{w}, \quad (34)$$

where $\mathbf{S}$ is the $n \times m$ matrix of columns $\mathbf{s}_k$ and $\mathbf{s}_k$ is the length-n signature vector of the kth user, $\mathbf{A} = \text{diag}(A_1, \ldots, A_m)$ where $A_k > 0$ is the received amplitude of the kth user's signal, $\mathbf{b}$ is a vector of elements b_k where $b_k \in \{1, -1\}$ is the bit transmitted by the kth user, and $\mathbf{w}$ is a white Gaussian noise vector with zero mean and covariance $\mathbf{C}_w = \sigma^2\mathbf{I}$.

Given the received signal $\mathbf{y}$ the problem is to detect the information bits b_k of the different users. One approach consists of estimating the vector $\mathbf{x} = \mathbf{Ab}$ and then detecting the kth symbol as $\hat{b}_k = \text{sgn}(\hat{x}_k)$ where $\hat{x}_k$ is the kth component of $\hat{\mathbf{x}}$.

Estimating $\mathbf{x}$ using an LS estimator results in the well-known decorrelator receiver, first proposed by Lupas and Verdu [65]. Alternatively, we may estimate $\mathbf{x}$ using the CSLS estimator, which results in a new receiver which we refer to as the CSMU receiver. The form of this receiver is discussed in "Covariance Shaping Multiuser Detection."

From the general properties of the CSLS estimator [1], [54], it follows that the CSMU receiver can be implemented as a decorrelator receiver followed by a (weighted) MMSE covariance shaping transformation that optimally shapes the covariance of the output of the decorrelator. The choice of shaping can be tailored to the specific set of signatures. It can also be shown that this receiver is equivalent to an MF receiver followed by an MMSE covariance

Covariance Shaping Multiuser Detection

The CSMU detector results from estimating $x = \mathbf{Ab}$ in (34) using the CSLS estimator, which in the case of linearly independent signature vectors leads to the estimator

$$\hat{x}_{CSLS} = (\mathbf{RS}^*\mathbf{S})^{-1/2}\mathbf{RS}^*\mathbf{y}, \tag{35}$$

where $\mathbf{R}$ is any positive-definite Hermitian matrix. Note that the scaling of $\hat{x}_{CSLS}$ will not effect the detector output and therefore can be chosen arbitrarily. The resulting receiver cross-correlates $\mathbf{y}$ with each of the columns $\mathbf{q}_k$ of $Q = \mathbf{SR}(\mathbf{S}^*\mathbf{SR})^{-1/2}$ to yield the outputs $d_k = \mathbf{q}_k^*\mathbf{y}$. The kth users' bit is then detected as $\hat{b}_k = \text{sgn}(d_k)$.

In the large system limit when both the number of users and the length of the signature vectors goes to infinity with fixed ratio $\beta = n/m$, it can be shown that the SINR γ at the output of the CSMU detector with $\mathbf{R} = \mathbf{I}$ converges to a deterministic limit when random Gaussian signatures and accurate power control are used. Specifically, in [53] we show that

$$\gamma \overset{\text{m.s.}}{\rightarrow} \frac{1}{1 - \frac{\eta_2\left[(\eta_1+\eta_2)E\left(\sqrt{1-\eta_1/\eta_2}\right) - 2\eta_1 K\left(\sqrt{1-\eta_1/\eta_2}\right)\right]^2}{9\pi^2\beta^2(\zeta+1)}} - 1 \tag{36}$$

where $1/\zeta$ is the received SNR,

$$K(k) = \int_0^{\pi/2} \frac{dt}{\sqrt{1-k^2\sin^2 t}} = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{37}$$

$$E(k) = \int_0^{\pi/2} \sqrt{1-k^2\sin^2 t}\,dt = \int_0^1 \sqrt{\frac{1-k^2x^2}{1-x^2}}\,dx \tag{38}$$

are the complete elliptic integrals of the first and second kinds, respectively [104], and

$$\eta_1 = \left(1-\sqrt{\beta}\right)^2 \tag{39}$$

$$\eta_2 = \left(1+\sqrt{\beta}\right)^2. \tag{40}$$

The notation $\overset{\text{m.s.}}{\rightarrow}$ denotes convergence in the mean-squared (L^2) sense [105].

shaping transformation. Finally, the CSMU receiver can also be implemented as a correlation demodulator with correlating signals with inner product matrix $\mathbf{R}$ that are closest in an LS sense to the signature vectors.

In the special case in which $\mathbf{R} = \mathbf{I}$, the shaping transformation is a whitening transformation, and the CSMU receiver is equivalent to a decorrelator receiver followed by MMSE whitening. This receiver has been referred to as the orthogonal multiuser receiver [22], [53]. By allowing for other choices of $\mathbf{R}$ we can improve the performance over both the decorrelator and the orthogonal multiuser receiver for a wide range of channel parameters.

To demonstrate the performance advantage in using the CSLS estimator, we consider the case in which the signature vectors are chosen as pseudonoise sequences corresponding to maximal-length, shift-register sequences [64], [66], so that

$$\mathbf{s}_k^*\mathbf{s}_l = \begin{cases} 1, & l = k; \\ -1/n, & l \neq k. \end{cases} \tag{41}$$

The shaping $\mathbf{R}$ is chosen as a circulant matrix with parameter ρ:

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & & \ddots & & \vdots \\ \rho & \rho & \cdots & \rho & 1 \end{bmatrix}. \tag{42}$$

Fig. 12 compares the theoretical probability of bit error of the CSMU receiver in the case of ten users with $\rho = 0.35$ and with accurate power control so that $A_m = 1$ for all m. The corresponding curves for the decorrelator, MF, and linear MMSE receivers are plotted for comparison. The MMSE receiver is the linear receiver that maximizes the signal-to-interference ratio and, unlike the decorrelator, MF, and CSMU receivers, requires knowledge of the channel parameters. We see that the CSMU receiver performs better than the decorrelator and the MF and performs similarly to the linear MMSE receiver, even though it does not rely on knowledge of the channel parameters.

In [67] we develop methods to analyze the output signal-to-interference+noise ratio (SINR) of the CSMU receiver in the large system limit. We show that the SINR converges to a deterministic limit and compare this limit to the known SINR limits for the decorrelator, MF, and linear MMSE receivers [67]-[69]. The analysis suggests that this modified receiver can lead to improved performance over the decorrelator and MF receiver and can approach the performance of the linear MMSE receiver over a wide range of channel parameters without requiring knowledge of these parameters.

Applications of Subspace Measurements

Simple Subspace Measurements

A simple subspace QSP measurement is equivalent to a linear projection operator. Numerous signal processing and detection algorithms based on orthogonal projections have been developed. Algorithms based on oblique projections have received much less attention in the signal processing literature. Recently, oblique projections have been applied to various detection problems [70], [71], to signal and parameter estimation [72], to computation of wavelet transforms [73], and to the formulation of consistent interpolation and sampling methods [20], [74].

In [20] the authors develop consistent reconstruction algorithms for sampled signals, in which the reconstructed signal is in general not equal to the original signal but nonetheless yields the same samples. Using a simple subspace QSP measurement corresponding to an oblique projection operator, in [23] and [24] the results of [20] are extended to a broader framework that can be applied to arbitrary subspaces of an arbitrary Hilbert space, as well as arbitrary input signals. The algorithms developed yield perfect reconstruction for signals in a subspace of $\mathcal{H}$, and consistent reconstruction for arbitrary signals, so that this framework includes the more restrictive perfect reconstruction theories as special cases. This framework leads to some new sampling theorems and can also be used to construct signals from a certain class with prescribed properties [75]. For example, we can use this framework to construct a finite-length signal with specified low-pass coefficients or an odd-symmetric signal with specified local averages.

Subspace Coding and Decoding

Subspace measurements also lead to interesting and potentially useful coding and decoding methods for communication-based applications over a variety of channel models. In particular, in [1] a subspace approach for transmitting information over a noisy channel is proposed, in which the information is encoded in disjoint subspaces. To detect the information, we design a receiver based on a subspace QSP measurement and show that for a certain class of channel models this receiver implements a generalized likelihood ratio test. Although the discussion constitutes a rather preliminary exploration of such coding techniques, it represents an interesting and potentially useful model for communication in many contexts. In particular, decoding methods suggested by the QSP framework may prove useful in the context of recent advances in multiple-antenna coding techniques [76], [77].

Combined Measurements

An interesting class of measurements in quantum mechanics results from restricting measurements to a subspace in which the quantum system is known a priori to lie. This leads to the notion of generalized measurements, or positive operator-valued measures (POVMs) [78], [79]. It can be shown that a generalized measurement on a quantum system can be implemented by performing a standard measurement on a larger system. Alternatively, we can view a generalized quantum measurement as a combination of a standard measurement followed by an orthogonal projection onto a lower space.

Drawing from the quantum mechanical POVM, as part of the QSP framework we also consider combined QSP measurements. The QSP analogue of a quantum POVM is a ROM followed by a simple subspace QSP measurement corresponding to an orthogonal projection operator. Since the QSP framework does not depend on the physics associated with quantum mechanics, we may extend the notion of a (physically realizable) POVM to include other forms of combined QSP measurements, where we perform any two measurements successively. As developed further in [1], such measurements lead to a variety of extensions and rich insights into frames, to new classes of frames, and to the concept of oblique frame expansions. This framework also leads to subspace MF detectors and randomized algorithms for improving worst-case performance.

Combined Measurements and Tight Frames

Emulating the quantum POVM leads to combined measurements where a ROM is followed by an orthogonal projection onto a subspace $\mathcal{U}$. Such measurements are characterized by an effective set of measurement vectors. In [80] we show that the family of possible effective measurement vectors in $\mathcal{U}$ is equal to the family of rank-one POVMs on $\mathcal{U}$ and is precisely the family of (normalized) tight frames for $\mathcal{U}$.

Frames are generalizations of bases which lead to redundant signal expansions [81], [82]. A *frame* for a Hilbert space $\mathcal{U}$ is a set of not necessarily linearly independent vectors that spans $\mathcal{U}$ and has some additional properties. Frames were first introduced by Duffin and Schaeffer [81] in the context of nonharmonic Fourier series and play an important role in the theory of nonuniform sampling [81]-[83]. Recent interest in frames has been motivated in part by their utility in analyzing wavelet expansions [84], [85]. A *tight frame* is a special case of a frame for which the reconstruction formula is particularly simple and is reminiscent of an orthogonal basis expansion, even though the frame vectors in the expansion are linearly dependent.

Exploiting the equivalence between tight frames and quantum POVMs, we develop frame-theoretic analogues of various quantum-mechanical concepts and results [80]. In particular, motivated by the construction of optimal LS quantum measurements [32], we consider the problem of constructing optimal LS tight frames for a subspace $\mathcal{U}$ from a given set of vectors that span $\mathcal{U}$.

The problem of frame design has received relatively little attention in the frame literature. A popular frame construction from a given set of vectors is the canonical frame [86]-[89], first proposed in the context of wavelets in [90]. The canonical frame is relatively simple to construct, can be determined directly from the given vectors, and plays an important role in wavelet theory [18], [91], [92]. In [80] we show that the canonical frame vectors are proportional to the LS frame vectors.

This relationship between combined measurements and frames suggests an alternative definition of frames in terms of projections of a set of linearly independent signals in a larger space. This perspective provides additional insights into frames and suggests a systematic approach for generating new classes of frames by changing the properties of the signals or changing the properties of the projection.

Geometrically Uniform Frames

In the context of a single QSP measurement, we have seen that imposing inner product constraints on the measurement vectors of a ROM leads to interesting new processing techniques. Similarly, in the context of combined measurements imposing such constraints leads to the definition of *geometrically uniform frames* [93]. This class of frames is highly structured, resulting in nice computational properties, and possesses strong symmetries that may be advantageous in a variety of applications such as channel coding [94]-[96] and multiple description source coding [97]. Further results regarding these frames are developed in [93].

Consistent Sampling and Oblique Dual Frame Vectors

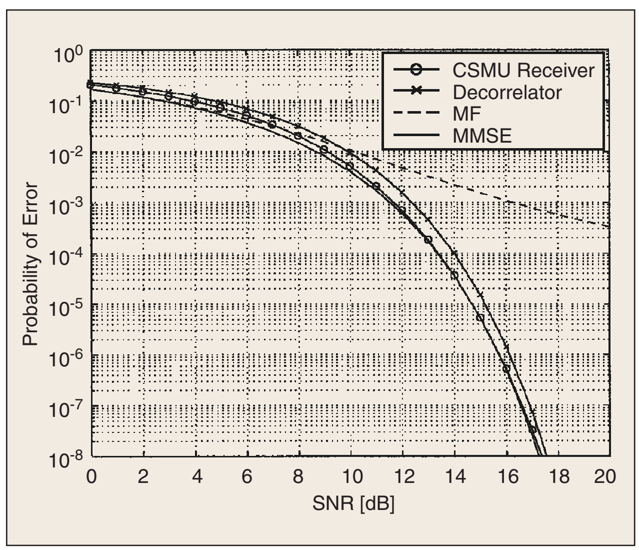
We also explore extensions of frames that result from choosing an oblique projection operator onto $\mathcal{U}$. In this case the the measurement is described in terms of two sets of effective measurement vectors, where the first set forms a frame for $\mathcal{U}$ and the second set forms what we define as an *oblique dual frame*. The frame operator corresponding to these vectors is referred to as an *oblique dual frame operator* and is a generalization of the well-known dual frame operator [86]. As we show in [23], these frame vectors have properties that are very similar to those of the conventional dual frame vectors. However, in contrast with the dual frame vectors, they are not constrained to lie in the same space as the original frame vectors. Thus, using oblique dual frame vectors we can extend the notion of a frame expansion to include redundant expansions in which the analysis frame vectors and the synthesis frame vectors lie in different spaces.

Based on the concept of oblique dual frame vectors, in [23] we develop *redundant* consistent sampling procedures with (almost) arbitrary sampling and reconstruction spaces. By allowing for arbitrary spaces, the sampling and reconstruction algorithms can be greatly simplified in many cases with only a minor increase in approximation error [20], [21], [98]-[101]. By using oblique dual frame vectors, we can further simplify the sampling and reconstruction processes while still retaining the flexibility of choosing the spaces almost arbitrarily, due to the extra degrees of freedom offered by the use of frames that allow us to construct frames with prescribed properties [84], [102]. Furthermore, if the measurements are quantized prior to reconstruction, then as we show the average power of the reconstruction, error using this redundant procedure can be reduced by as much as the redundancy of the frame in comparison with the nonredundant procedure.

Three principles of quantum mechanics that play a major role in QSP are the concept of a measurement, the principle of measurement consistency, and the principle of quantization of the measurement output.

Acknowledgments

We are extremely grateful to Prof. G.D. Forney for his many thoughtful comments and valuable suggestions that greatly influenced the formulation of the QSP framework and many of the specific results that follow from it. This work has been supported in part by the Army Research Laboratory (ARL) Collaborative Technology Alliance through BAE Systems, Inc. subcontract RK7854; BAE Systems Cooperative Agreement RP6891 under Army Research Laboratory Grant DAAD19-01-2-0008; BAE Systems, Inc. through Contract RN5292; Texas Instruments through the TI Lead-



▲ *12. Probability of bit error with ten users,* $\rho = 0.35$*, and accurate power control, as a function of SNR.*

ership University Consortium; and by IBM through an IBM Fellowship.

Yonina C. Eldar received the B.Sc. degree in physics (summa cum laude) in 1995 and the B.Sc. degree in electrical engineering (summa cum laude) in 1996 both from Tel-Aviv University (TAU), Tel-Aviv, Israel, and the Ph.D. degree in electrical engineering and computer science in 2001 from the Massachusetts Institute of Technology (MIT), Cambridge. From January 2002 to July 2002 she was a postdoctoral fellow at the Digital Signal Processing Group at MIT. She is currently a senior lecturer in the Department of Electrical Engineering at the Technion-Israel Institute of Technology, Haifa. Her current research interests are in the general areas of quantum detection and signal processing. In 1998 she held the Rosenblith Fellowship for study in electrical engineering at MIT, and in 2000 she held an IBM Research Fellowship.

Alan V. Oppenheim received the S.B. and S.M. degrees in 1961 and the Sc.D. degree in 1964, all in electrical engineering, from the Massachusetts Institute of Technology (MIT). He is also the recipient of an honorary doctorate from Tel Aviv University in 1995. In 1964, he joined the faculty at MIT, where he currently is Ford Professor of Engineering and a MacVicar Faculty Fellow. His research interests are in the general area of signal processing and its applications. He is coauthor of the textbooks *Discrete-Time Signal Processing* and *Signals and Systems*. He is also editor of several advanced books on signal processing. He is a member of the National Academy of Engineering, a Fellow of the IEEE, and a member of Sigma Xi and Eta Kappa Nu. He has been a Guggenheim Fellow and a Sackler Fellow at Tel Aviv University. He has also received a number of awards for outstanding research and teaching including the IEEE Education Medal, the IEEE Centennial Award, the Society Award, the Technical Achievement Award, and the Senior Award of the IEEE Society on Acoustics, Speech and Signal Processing. He has also received a number of awards at MIT for excellence in teaching, including the Bose Award and the Everett Moore Baker Award.

References

[1] Y.C. Eldar, "Quantum signal processing," Ph.D. dissertation, Mass. Inst. Technol., Dec. 2001. Available: http://allegro.mit.edu/dspg/publications/TechRep/index.html

[2] A. Fettweis, "Wave digital filters: Theory and practice," *Proc. IEEE*, vol. 74, pp. 270-316, Feb. 1986.

[3] B.B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1982.

[4] G.W. Wornell and A.V. Oppenheim, "Wavelet-based representations for a class of self similar signals with application to fractal modulation," *IEEE Trans. Inform. Theory*, vol. 38, pp. 785-800, Mar. 1992.

[5] M.F. Barnsley and A.D. Sloan, "A better way to compress images," *BYTE*, vol. 1, no. 12, pp. 215-223, Jan. 1988.

[6] T.L. Carroll and L.M. Pecora, "Synchronizing chaotic circuits," *IEEE Trans. Circuits Syst*., vol. 38, pp. 453-456, Apr. 1991.

[7] L.M. Pecora and T.L. Carroll, "Driving systems with chaotic signals," *Phys. Rev. A*, vol. 44, pp. 2374-2383, Aug. 1991.

[8] K.M. Cuomo, A.V. Oppenheim, and S.H. Strogatz, "Synchronization of Lorenz-based chaotic circuits with applications to communications," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 626-639, Oct. 1993.

[9] "Special issue on applications of chaos in modern communications systems," *IEEE Trans. Circuits Syst. I*, vol. 48, Dec. 2001.

[10] A.C. Singer, A.V. Oppenheim, and G.W. Wornell, "Detection and estimation of multiplexed soliton signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 2768-2782, Oct. 1999.

[11] K.S. Tang, K.F. Man, S. Kwong, and Q. He, "Genetic algorithms and their applications," *IEEE Signal Processing Mag*., pp. 22-37, Nov. 1996.

[12] M. Pirlot and R.V.V. Vidal, "Simulated annealing: A tutorial," *Contr. Cybern.,* vol. 25, no. 1, pp. 9-31, 1996.

[13] S.S. Haykin, *Neural Networks: A Comprehensive Foundation,* 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[14] H.W. Schüssler and P. Steffen, "Some advanced topics in filter design," in *Advanced Topics in Signal Processing*, J.S. Lim and A.V. Oppenheim, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[15] J.G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[16] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[17] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge, 1996.

[18] S.G. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 674-693, 1989.

[19] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[20] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, pp. 2915-2925, Nov. 1994.

[21] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, pp. 569-587, Apr. 2000.

[22] Y.C. Eldar and A.V. Oppenheim, "Nonredundant and redundant sampling with arbitrary sampling and reconstruction spaces," in *Proc. 2001 Workshop on Sampling Theory and Applicatations*, SampTA'01, May 2001, pp. 229-234.

[23] Y.C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *J. Fourier Analys. Appl.*, to be published.

[24] Y.C. Eldar, "Sampling without input constraints: Consistent reconstruction in arbitrary spaces," *Sampling, Wavelets and Tomography,* A. Zayed and J.J. Benedetto, Eds., to be published.

[25] C.W. Helstrom, *Quantum Detection and Estimation Theory*. New York: Academic, 1976.

[26] D.G. Griffiths, *Introduction to Quantum Mechanics*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[27] A.S. Holevo, "Statistical decisions in quantum theory," *J. Multivar. Anal*., vol. 3, pp. 337-394, Dec. 1973.

[28] H.P. Yuen, R.S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 125-134, Mar. 1975.

[29] Y.C. Eldar, A. Megretski, and G.C. Verghese, "Designing optimal quantum detectors via semidefinite programming," *IEEE Trans. Inform. Theory*, to be published.

[30] M. Charbit, C. Bendjaballah, and C.W. Helstrom, "Cutoff rate for the -ary PSK modulation channel with optimal quantum detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1131-1133, Sept. 1989.

[31] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, "Optimum measurements for discrimination among symmetric quantum states and parameter estimation," *Int. J. Theor. Phys.*, vol. 36, pp. 1269-1288, 1997.

[32] Y.C. Eldar and G.D. Forney, Jr., "On quantum detection and the square-root measurement," *IEEE Trans. Inform. Theory*, vol. 47, pp. 858-872, Mar. 2001.

[33] Y.C. Eldar, A. Megretski, and G.C. Verghese, "Optimal detection of symmetric mixed quantum states," in preparation.

[34] P. Hausladen and W.K. Wootters, "A 'pretty good' measurement for distinguishing quantum states," *J. Mod. Opt.*, vol. 41, pp. 2385-2390, 1994.

[35] H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," [Online]. Available: http://xxx.lanl.gov/abs/quant-ph/0004088

[36] P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, and W.K. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A*, vol. 54, pp. 1869-1876, Sept. 1996.

[37] Y.C. Eldar, "Least-squares inner product shaping," *Linear Algebra Applicat.,* vol. 348, pp. 153-174, May 2002.

[38] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[39] A. Aldroubi, "Oblique projections in atomic spaces," *Proc. Amer. Math. Soc*., vol. 124, no. 7, pp. 2051-2060, 1996.

[40] S. Kayalar and H.L. Weinert, "Oblique projections: Formulas, algorithms, and error bounds," *Math. Contr. Signals Syst.*, vol. 2, no. 1, pp. 33-45, 1989.

[41] S.P. Lipshitz, R.A. Wannamaker, and J. Vanderkooy, "Quantization and dither: A theoretical survey," *J. Audio Eng. Soc.*, vol. 40, no. 5, pp. 355-375, May 1992.

[42] R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, and J.N. Wright, "A theory of nonsubtractive dither," *IEEE Trans. Signal Processing*, vol. 48, pp. 499-516, Feb. 2000.

[43] R.A. Wannamaker, "The theory of dithered quantization," Ph.D. dissertation, Univ. Waterloo, Waterloo, Canada, 1997.

[44] R.M. Gray and T.G. Stockham, "Dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 805-812, May 1993.

[45] Y.C. Eldar, A.V. Oppenheim, and D. Egnor, "Orthogonal and projected orthogonal matched filter detection," *Signal Processing*, June 2002, submitted.

[46] Y.C. Eldar and A.V. Oppenheim, "Orthogonal matched filter detection," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (ICASSP-2001), Salt Lake City, UT, 2001, vol. 5, pp. 2837-2840.

[47] L.L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

[48] T. Kailath and V. Poor, "Detection of stochastic processes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2230-2259, Oct. 1998.

[49] Y.C. Eldar and A.V. Oppenheim, "MMSE whitening and subspace whitening," submitted for publication.

[50] E.M. Friel and K.M. Pasala, "Direction finding with compensation for a near field scatterer," in Proc. *Int. Symp. Ant. Propagat. Soc.*, 1995, pp. 106-109.

[51] R.J. Piechocki, N. Canagarajah, and J.P. McGeehan, "Improving the direction-of-arrival resolution via double code filtering in WCDMA," in *Proc. First Int. Conf. 3G Mobile Commun. Technol.*, Mar. 2000, pp. 204-207.

[52] Y.C. Eldar and A.V. Oppenheim, "Orthogonal multiuser detection," *Signal Processing*, vol. 82, pp. 321-325, 2002.

[53] Y.C. Eldar and A.M. Chan, "An optimal whitening approach to linear multiuser detection," submitted for publication.

[54] Y. C. Eldar and A. V. Oppenheim, "Covariance shaping least-squares estimation," *IEEE Trans. Signal Processing*, to be published.

[55] T. Kailath, *Lectures on Linear Least-Squares Estimation*. Wein, NY: Springer, 1976.

[56] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[57] C.W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[58] J.A. Cadzow, "Signal processing via least squares modeling," *IEEE ASSP Mag*., pp. 12-31, Oct. 1990.

[59] A.E. Hoerl and R.W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, pp. 55-67, Feb. 1970.

[60] A.N. Tikhonov and V.Y. Arsenin, *Solution of Ill-Posed Problems*. Washington, DC: Winston, 1977.

[61] L.S. Mayer and T.A. Willke, "On biased estimation in linear models," *Technometrics*, vol. 15, pp. 497-508, Aug. 1973.

[62] C.R. Rao, "Minimum variance and the estimation of several parameters," *Proc. Cambridge Phil. Soc.*, pp. 280-283, 1946.

[63] H.L. Van Trees*, Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.

[64] S. Verdu, *Multiuser Detection*, Cambridge, U.K.: Cambridge Univ. Press, 1998.

[65] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123-136, Jan. 1989.

[66] S. Ulukus and R.D. Yates, "Optimum multiuser detection is tractable for synchronous CDMA systems using *m*-sequences," *IEEE Commun. Lett.*, vol. 2, pp. 89-91, Apr. 1998.

[67] Y.C. Eldar and A.M. Chan, "On the asymptotic large system performance of the decorrelator," *IEEE Trans. Inform. Theory,* to be published.

[68] D.N.C. Tse and S.V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641-657, Mar. 1999.

[69] S. Verdu and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622-640, Mar. 1999.

[70] R.T. Behrens and L.L. Scharf, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Processing*, vol. 42, pp. 1413-1424, June 1994.

[71] Y.C. Eldar, "On geometrical properties of the decorrelator," *IEEE Commun. Lett.*, vol. 6, pp. 16-18, Jan. 2002.

[72] R.T. Behrens and L.L. Scharf, "Parameter estimation in the presence of low-rank noise," in *Proc. 22nd Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, Nov. 1988.

[73] M.J. Vrhel, C. Lee, and M. Unser, "Rapid computation of the continuous wavelet transform by oblique projections," *IEEE Trans. Signal Processing*, vol. 45, pp. 891-900, Apr. 1997.

[74] C. Lee, M. Eden, and M. Unser, "High-quality image resizing using oblique projection operators," *IEEE Trans. Image Processing*, vol. 7, pp. 670-692, May 1998.

[75] Y.C. Eldar, "Constructing signals with prescribed properties," *IEEE Signal Processing Lett.*, to be published.

[76] B.M. Hochwald and W. Sweldens, "Space-time modulation for unknown fading," in *Digital Wireless Communication,* R.M. Rao, S.A. Dianat, and M.D. Zoltowski, Eds., SPIE Proc., vol. 3708, June 1999.

[77] B.M. Hochwald and T.L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543-564, 2000.

[78] M.A. Neumark, "On a representation of additive operator set functions," *C. R. (Doklady) Acad. Sci. URSS* (N.S.), vol. 41, pp. 359-361, 1943.

[79] A. Peres, "Neumark's theorem and quantum inseparability," *Found. Phys.*, vol. 20, no. 12, pp. 1441-1453, 1990.

[80] Y.C. Eldar and G.D. Forney, Jr., "Optimal tight frames and quantum measurement," *IEEE Trans. Inform. Theory*, vol. 48, pp. 599-610, Mar. 2002.

[81] R.J. Duffin and A.C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Amer. Math. Soc.*, vol. 72, pp. 314-366, 1952.

[82] R.M. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic, 1980.

[83] J.J. Benedetto, "Irregular sampling and frames," in *Wavelets—A Tutorial in Theory and Applications*, C.K. Chui, Ed. Boca Raton, FL: CRC, 1992, pp. 445-507.

[84] C.E. Heil and D.F. Walnut, "Continuous and discrete wavelet transforms," *SIAM Rev.*, vol. 31, no. 4, pp. 628-666, Dec. 1989.

[85] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961-1005, Sept. 1990.

[86] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[87] H. Bölcskei, "Oversampled filter banks and predictive subband coders," Ph.D. dissertation, Vienna Univ. Technol., Nov. 1997.

[88] H. Bölcskei and A.J.E.M. Janssen, "Gabor frames, unimodularity, and window decay," *J. Fourier Analys. Applicat.*, vol. 6, no. 3, pp. 255-276, 2000.

[89] A.J.E.M. Janssen and H. Bölcskei, "Equivalence of two methods for constructing tight Gabor frames," *IEEE Signal Processing Lett.*, vol. 7, pp. 79-82, Apr. 2000.

[90] Y. Meyer, "Ondelettes et functions splines," Dec. 1986, Seminaire EDP, Ecole Polytechn., Paris, France.

[91] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.,* pp. 909-996, 1988.

[92] M. Unser and A. Aldroubi, "Families of multiresolution and wavelet spaces with optimal properties," *Numer. Funct. Anal. Optimiz.*, vol. 14, pp. 417-446, 1993.

[93] Y.C. Eldar and H. Bölcskei, "Geometrically uniform frames," *IEEE Trans. Inform. Theory*, to be published.

[94] G.D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241-1260, Sept. 1991.

[95] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, pp. 55-67, 1982.

[96] D. Raphaeli, "On multidimensional coded modulations having uniform error property for generalized decoding and flat-fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 34-40, 1998.

[97] V.K. Goyal, J. Kovavcevic, and J.A. Kelner, "Quantized frame expansions with erasures," *Appl. Comp. Harmonic Anal.*, vol. 10, pp. 203-233, May 2001.

[98] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, pp. 22-38, Nov. 1999.

[99] M. Unser and J. Zerubia, "Generalized sampling: Stability and performance analysis," *IEEE Trans. Signal Processing*, vol. 45, pp. 2941-2950, Dec. 1997.

[100] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors," *IEEE Trans. Signal Processing*, vol. 47, pp. 2783-2795, Oct. 1999.

[101] T. Blu and M. Unser, "Quantitative Fourier analysis of approximation techniques: Part II—Wavelets," *IEEE Trans. Signal Processing*, vol. 47, pp. 2796-2806, Oct. 1999.

[102] A. Aldroubi, "Portraits of frames," *Proc. Amer. Math. Soc.*, vol. 123, pp. 1661-1668, 1995.

[103] J.J. Atick and A.N. Redlich, "Convergent algorithm for sensory receptive filed development," *Neurak Comp.*, vol. 5, 1993.

[104] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products*. New York: Academic, 1980.

[105] K.L. Chung, *A Course in Probability Theory*. San Diego, CA: Academic, 2001.