

# Bayesian Filtering with Wavefunctions

**Lachlan McCalman**

Australian Centre for Field Robotics  
University of Sydney  
New South Wales, Australia  
[l.mccalman@cas.edu.au](mailto:l.mccalman@cas.edu.au)

**Hugh Durrant-Whyte**

Australian Centre for Field Robotics  
University of Sydney  
New South Wales, Australia  
[h.durrant-whyte@cas.edu.au](mailto:h.durrant-whyte@cas.edu.au)

**Abstract** – *This paper describes a general framework for performing Bayesian filtering on probability density functions represented by the modulus squared of a wavefunction in an orthogonal function basis. The objective of this work is to find a sparse representation for high-dimensional non-Gaussian density functions enabling tractable implementations of the general Bayesian filtering problem. The general form of the Bayesian filtering algorithms employing the modulus squared of a wavefunction is presented without specification of a particular basis. A simple 0-th order implementation of this formulation for a bearing only sensor is also demonstrated.*

**Keywords:** orthogonal functions, Bayesian filtering, non-Gaussian estimation

## 1 Introduction

Bayesian filtering is exactly solvable only for particular parametric probability density functions (PDFs) such as Gaussians. In the non-Gaussian case, approximations including particle filters, Parzen density estimates and Gaussian mixture models are most often employed. The problem with these approximations is that they are not sparse and do not scale with the dimension of the state space under consideration making them difficult to employ in large-scale filtering problems.

This paper addresses the fundamental problem of finding a general representation for a PDF which is *sparse* in the chosen basis. That is, it must approximate the PDF with a sufficiently small number of coefficients to allow storage, manipulation and communication. Such a representation would allow large-scale non-Gaussian filtering problems to be addressed in a principled algorithmic manner and enable such methods to be applied in real distributed estimation problems.

Gaussian approximations to a PDF are well known and the Kalman filter provides an optimal, linear solution to the Bayesian filtering equations in this case [9].

The Kalman filter and its non-linear cousins the extended and unscented Kalman filters have been extremely successful in applications where the Gaussian approximation is valid, especially in applications such as tracking, navigation and mapping [2, 12]. Critically, the spacial requirements for storage of a Gaussian function scale at-worst quadratically with dimension. Gaussian models can also be applied in a decentralised estimation context.

A number of Non-Gaussian PDF representations for Bayesian filtering have been developed. These generally use a combination of functions to represent a PDF. Particle filters use Monte-Carlo point sampling estimation methods [8], which is equivalent to a Dirac delta function representation. Filtering algorithms have been developed which are tractable in low-dimensions, but dimensional scaling is particularly poor due to the delta function having zero volume. Parzen density estimates represent PDFs as a tiling of kernel functions [11], usually Gaussian, triangular or raised cosines. Parzen density models are dense and do not scale well with dimension. Gaussian mixture models (GMMs) are similar to Parzen estimates but allow variation in the mean and covariance of the Gaussians being used to tile [1]. In both cases determining the approximation from a known function is computationally expensive, especially if a uniqueness condition is enforced in the case of GMMs. The Bayesian filtering equations have been solved for this class of approximation [1], but in practice, storage requirements scale exponentially with dimension for many functions.

Representing PDFs with an orthogonal function basis can, in principle, overcome these limitations. A choice of basis can be made such that the PDF is sparsely represented even in high-dimensional spaces, and orthogonality allows the use of projection to efficiently compute the basis coefficients from a known function. Some work has already been done in this area; for instance; filtering using  $N$ -D planewaves as an orthogonal function basis [3], and generalised Edgeworth series [4].

Wavelets have also been used [10] to represent PDFs, but the Bayesian filtering equations do not appear to have a closed form solution in this basis.

The work described in this paper starts with the intuition that a PDF can be described in the form of a modulus squared of a wave function representing an orthogonal basis; as is used in quantum mechanics. In principle, it is far easier to find an orthogonal basis for a wave function than a general PDF. Such a basis would provide a unique, sparse and reducible PDF representation allowing efficient filtering and communication of general non-Gaussian estimates. However, while such a representation is relatively easy to find, the key to employing this in Bayesian filtering is to find equivalent expressions for prediction and estimation that apply directly to these wavefunctions. This turns out to be a complex problem.

## 2 General Treatment

We outline a general treatment of continuous orthogonal bases for Bayesian filtering, which effectively borrows from the mathematical formulation of quantum mechanics. This uses the concept of a Hilbert space, which is an infinite-dimensional generalisation of a vector space. Like a vector space, Hilbert spaces have an inner-product, and contain at least one orthonormal basis. In this case the function space is the complex Lebesgue space  $L_2^{\mathbb{C}}(\mathbb{R}^N)$ , i.e the space of square-integrable complex-valued  $N$ -dimensional functions  $f : \mathbb{R}^N \rightarrow \mathbb{C}$ . The Hilbert space  $\mathcal{H}$  is then defined as this space together with an inner product

$$(\psi, \phi) \triangleq \int_{\mathbf{x} \in \mathbb{R}^N} \psi(\mathbf{x})^* \phi(\mathbf{x}) d\mathbf{x}. \quad (1)$$

Elements of  $\mathcal{H}$  do not yet correspond to probability density functions, which must be non-negative. Hence, we define a PDF  $P(\mathbf{x})$  associated with an element  $\psi$  of  $\mathcal{H}$  as

$$P(\mathbf{x}) = |\psi(\mathbf{x})|^2. \quad (2)$$

Square integrability of  $\psi$  ensures that this function will be normalisable.

In order to develop the Bayesian filtering equations on  $\psi$ , we first establish notation and note a couple of useful results.

**Dirac Notation** Dirac simplified notation in Hilbert spaces by abstracting away from the particular basis or inner product being used [5]. We will use this notation in the remainder of our treatment. In Dirac notation:

$$|\psi\rangle \longrightarrow \psi \quad (3)$$

$$\langle\psi| \longrightarrow \psi^* \quad (4)$$

$$\langle\psi|\phi\rangle \longrightarrow (\psi, \phi) \quad (5)$$

Dirac denoted  $|\psi\rangle \in \mathcal{H}$  as a *ket*,  $\langle\psi| \in \mathcal{H}^*$  as a *bra* (where  $\mathcal{H}^*$  is the conjugate space), and the inner product  $\langle\psi|\phi\rangle$  as a *bracket*.

**Bases and Wavefunctions** Every Hilbert space has at least one orthonormal basis. This is a set of kets  $\{|\xi(\mathbf{x})\rangle\}$  indexed by a continuous parameter  $\mathbf{x}$ . Unless there is confusion we will generally shorten the notation to  $\{|\mathbf{x}\rangle\}$ . The orthonormality condition is:

$$\langle\mathbf{x}|\mathbf{x}'\rangle = \delta(\mathbf{x} - \mathbf{x}') \quad (6)$$

where  $\delta$  is the Dirac delta function.  $\{|\mathbf{x}\rangle\}$  must also span  $\mathcal{H}$ , satisfying

$$|\psi\rangle = \int |\xi(\mathbf{x})\rangle \langle\xi(\mathbf{x})|\psi\rangle d\mathbf{x} \quad (7)$$

for all  $|\psi\rangle \in \mathcal{H}$ .

The function  $\langle\mathbf{x}|\psi\rangle$  gives the coefficients of the basis functions. This is often written as the *wavefunction*

$$\langle\mathbf{x}|\psi\rangle = \psi(\mathbf{x}). \quad (8)$$

**Change of Basis** Let the set of kets  $\{|\mathbf{x}\rangle\}$  and  $\{|\mathbf{r}\rangle\}$  define two orthonormal bases for  $\mathcal{H}$ . The change of basis rules follow from simple projection,

$$\langle\mathbf{r}|\psi\rangle = \langle\mathbf{r}|\left(\int |\mathbf{x}\rangle \langle\mathbf{x}|\psi\rangle d\mathbf{x}\right) \quad (9)$$

$$= \int \langle\mathbf{r}|\mathbf{x}\rangle \langle\mathbf{x}|\psi\rangle d\mathbf{x} \quad (10)$$

or in terms of wavefunctions,

$$\Psi(\mathbf{r}) = \int \langle\mathbf{r}|\mathbf{x}\rangle \psi(\mathbf{x}) d\mathbf{x}. \quad (11)$$

Here  $\langle\mathbf{r}|\mathbf{x}\rangle$  is the transformation function. The inverse transform is then simply

$$\psi(\mathbf{x}) = \int \langle\mathbf{x}|\mathbf{r}\rangle \Psi(\mathbf{r}) d\mathbf{r}. \quad (12)$$

## 3 Bayesian Filtering

We aim to perform Bayesian filtering by representing probability distributions in terms of elements of  $\mathcal{H}$ . The definition  $P(\mathbf{x}) = |\psi(\mathbf{x})|^2$  implies a “natural” basis associated with  $P(\mathbf{x})$ ,

$$|\mathbf{x}\rangle = \delta(\mathbf{x}' - \mathbf{x}). \quad (13)$$

However, this basis is unlikely to be ideal for representing  $P$  sparsely. In general, we will consider the transformation of  $P(\mathbf{x})$  to an arbitrary basis, and the associated wavefunction

$$\mathbf{x} \rightarrow \mathbf{r} : P(\mathbf{x}) \rightarrow \mathcal{P}(\mathbf{r}) \quad (14)$$

$$\mathcal{P}(\mathbf{r}) = |\Psi(\mathbf{r})|^2. \quad (15)$$

Let  $\psi(\mathbf{x})$  be a wavefunction representation of  $P(\mathbf{x})$ , and  $\Psi(\mathbf{r})$  a transformation of that wavefunction into an arbitrary basis. We aim to find the operations, which when performed on  $\psi$  or  $\Psi$ , are equivalent to Bayesian filtering performed on  $P$ . More explicitly, let  $P(\mathbf{x}) = |\psi(\mathbf{x})|^2$  and  $Q(\mathbf{x}) = |\phi(\mathbf{x})|^2$  be probability distributions, and  $F$  be the operation of Bayesian filtering, such that

$$F(P(\mathbf{x})) = Q(\mathbf{x}). \quad (16)$$

Assume under a basis transformation  $|\mathbf{x}\rangle \rightarrow |\mathbf{r}\rangle$  that  $\psi(\mathbf{x}) \rightarrow \Psi(\mathbf{r})$ . We aim to find an operation  $F'_\mathbf{x}$  and more generally  $F'_\mathbf{r}$  such that

$$F'_\mathbf{x}(\psi(\mathbf{x})) = \phi(\mathbf{x}), \quad (17)$$

and

$$F'_\mathbf{r}(\Psi(\mathbf{r})) = \Phi(\mathbf{r}). \quad (18)$$

The Bayesian filtering process consists of 2 steps; the prediction step

$$P(\mathbf{x}_k)^- = \int P(\mathbf{x}_k | \mathbf{x}_{k-1}) P(\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \quad (19)$$

and then the observation update step

$$P(\mathbf{x}_k) = \frac{P(\mathbf{z}_k | \mathbf{x}_k) P(\mathbf{x}_k)^-}{P(\mathbf{z}_k)} \quad (20)$$

$$= \frac{1}{C} P(\mathbf{z}_k | \mathbf{x}_k) P(\mathbf{x}_k)^- \quad (21)$$

for state vector  $\mathbf{x}_k$  and observation  $\mathbf{z}_k$  at time  $k$ . As the denominator is independent of state, we have written it as a normalising factor  $C$ .

These two steps use only 2 operations: multiplication of 2 functions and marginalisation of a function (multiplication by a scalar is a special case of multiplication of 2 functions). So, in order to determine  $F'$ , we must find operations, which when performed on wavefunctions, are equivalent to multiplication and marginalisation on PDFs.

### 3.1 Multiplication

Consider 3 probability distributions:  $P(\mathbf{x})$ ,  $Q(\mathbf{x})$  and  $R(\mathbf{x})$  defined by

$$P(\mathbf{x}) = |\psi(\mathbf{x})|^2 \quad (22)$$

$$Q(\mathbf{x}) = |\phi(\mathbf{x})|^2 \quad (23)$$

$$R(\mathbf{x}) = |\lambda(\mathbf{x})|^2 \quad (24)$$

$$(25)$$

and such that

$$P(\mathbf{x})Q(\mathbf{x}) = R(\mathbf{x}). \quad (26)$$

Then

$$P(\mathbf{x})Q(\mathbf{x}) = |\psi(\mathbf{x})|^2 |\phi(\mathbf{x})|^2 \quad (27)$$

$$= |\psi(\mathbf{x})\phi(\mathbf{x})|^2 \quad (28)$$

and so,

$$\lambda(\mathbf{x}) = \psi(\mathbf{x})\phi(\mathbf{x}). \quad (29)$$

Hence, in the natural basis, multiplication is defined identically for wavefunctions and PDFs.

In an arbitrary basis

$$\lambda(\mathbf{x}) = \psi(\mathbf{x})\phi(\mathbf{x}) \quad (30)$$

$$\Lambda(\mathbf{r}) = \int \langle \mathbf{r} | \mathbf{x} \rangle \lambda(\mathbf{x}) d\mathbf{x} \quad (31)$$

and so

$$\Lambda(\mathbf{r}) = \int \langle \mathbf{r} | \mathbf{x} \rangle \int \langle \mathbf{x} | \mathbf{r} \rangle \Psi(\mathbf{r}) d\mathbf{r} \int \langle \mathbf{x} | \mathbf{r} \rangle \Phi(\mathbf{r}) d\mathbf{r} d\mathbf{x}. \quad (32)$$

Now, let  $\mathcal{W}$  be the set of wavefunctions  $\{f : \mathbb{R}^N \rightarrow \mathbb{C}\}$ . Then we now define a binary operator  $\star : \mathcal{W} \times \mathcal{W} \rightarrow \mathcal{W}$  by

$$\Psi(\mathbf{r}) \star \Phi(\mathbf{r}) \triangleq \int \langle \mathbf{r} | \mathbf{x} \rangle \int \langle \mathbf{x} | \mathbf{r} \rangle \Psi(\mathbf{r}) d\mathbf{r} \int \langle \mathbf{x} | \mathbf{r} \rangle \Phi(\mathbf{r}) d\mathbf{r} d\mathbf{x}. \quad (33)$$

Clearly then

$$P(\mathbf{x})Q(\mathbf{x}) = R(\mathbf{x}) \leftrightarrow \Psi(\mathbf{r}) \star \Phi(\mathbf{r}) = \Lambda(\mathbf{r}). \quad (34)$$

### 3.2 Marginalisation

Consider 2 PDFs

$$P(\mathbf{x}, \mathbf{y}) = |\psi(\mathbf{x}, \mathbf{y})|^2 \quad (35)$$

$$Q(\mathbf{x}) = |\phi(\mathbf{x})|^2 \quad (36)$$

such that the marginalisation of with respect to  $\mathbf{y}$  of  $P$  is  $Q$ , i.e

$$\int P(\mathbf{x}, \mathbf{y}) d\mathbf{y} = Q(\mathbf{x}). \quad (37)$$

In this case

$$|\phi(\mathbf{x})|^2 = \int |\psi(\mathbf{x}, \mathbf{y})|^2 d\mathbf{y} \quad (38)$$

$$\phi(\mathbf{x}) = \left( \int |\psi(\mathbf{x}, \mathbf{y})|^2 d\mathbf{y} \right)^{\frac{1}{2}} e^{i\theta(\mathbf{x})}. \quad (39)$$

In general  $\theta(\mathbf{x})$  is a free parameter which for the moment we will set to zero. Marginalisation with an arbitrary basis requires substituting the basis transform;

$$\Phi(\mathbf{r}) = \int \langle \mathbf{r}, \mathbf{s} | \mathbf{x}, \mathbf{y} \rangle \left( \int \left| \int \langle \mathbf{x}, \mathbf{y} | \mathbf{r}, \mathbf{s} \rangle \Psi(\mathbf{r}, \mathbf{s}) d\mathbf{r} d\mathbf{s} \right|^2 d\mathbf{y} \right)^{\frac{1}{2}} d\mathbf{x}. \quad (40)$$

Now, if we define a marginalisation function  $M$  by

$$M_y(\Psi(\mathbf{r}, \mathbf{s})) = \int \langle \mathbf{r}, \mathbf{s} | \mathbf{x}, \mathbf{y} \rangle \left( \int \left| \int \langle \mathbf{x}, \mathbf{y} | \mathbf{r}, \mathbf{s} \rangle \Psi(\mathbf{r}, \mathbf{s}) d\mathbf{r} d\mathbf{s} \right|^2 d\mathbf{y} \right)^{\frac{1}{2}} d\mathbf{x} \quad (41)$$

then

$$\int P(\mathbf{x}, \mathbf{y}) d\mathbf{y} = Q(\mathbf{x}) \leftrightarrow M_y(\Psi(\mathbf{r}, \mathbf{s})) = \Phi(\mathbf{r}). \quad (42)$$

### 3.3 Wavefunction Filtering

The two equivalent operations for multiplication and marginalisation allow us to define Bayesian filtering on wavefunctions. Let  $P(\mathbf{x}_k)$  be a posterior distribution for  $\mathbf{x}$  at timestep  $k$ , and  $\Psi(\mathbf{x}'_k)$  be the corresponding wavefunction in some arbitrary basis  $\{\mathbf{x}'\}$ . Similarly, let  $\Psi(\mathbf{x}'_k | \mathbf{x}'_{k-1})$  be the wavefunction of the state transition function in basis  $\{\mathbf{x}'\}$ , and  $\Psi(\mathbf{z}'_k | \mathbf{x}'_k)$  be the observation model wavefunction. Then the prediction step on  $\Psi(\mathbf{x}'_k)$  is given by

$$\Psi(\mathbf{x}'_k)^- = M_{\mathbf{x}_{k-1}}(\Psi(\mathbf{x}'_k | \mathbf{x}'_{k-1}) \star \Psi(\mathbf{x}'_{k-1})) \quad (43)$$

and the observation update step by

$$\Psi(\mathbf{x}'_k) = \frac{1}{C} \Psi(\mathbf{z}'_k | \mathbf{x}'_k) \star \Psi(\mathbf{x}'_k)^-. \quad (44)$$

Clearly, these equations are intractable in the general case; both contain nested integrals over all  $\mathbb{R}^N$ . A specific function basis must be used which allows simplification of these integrals. Though such a basis has not yet been found, there are many candidates in the literature; Hermite, Laguerre or Jacobi polynomials for instance [6, 7].

## 4 Example: Planewave (Fourier) Basis

For illustrative purposes we will now demonstrate the above results in the  $N$ -d planewave or Fourier basis. Let the planewave basis be denoted by  $\{|\boldsymbol{\omega}\rangle\}$ . Using the standard definition of an  $N$ -d Fourier series, we can write the projection

$$\langle \mathbf{x} | \boldsymbol{\omega} \rangle = e^{-2\pi i \mathbf{x} \cdot \boldsymbol{\omega}} \quad (45)$$

which then allows us to derive the relationship between the wavefunctions

$$\psi(\mathbf{x}) = \langle \mathbf{x} | \left( \int |\boldsymbol{\omega}\rangle \langle \boldsymbol{\omega}| \right) | \psi \rangle \quad (46)$$

$$= \int e^{-2\pi i \mathbf{x} \cdot \boldsymbol{\omega}} \Psi(\boldsymbol{\omega}) \quad (47)$$

$$= \mathcal{F}[\psi(\mathbf{x})](\boldsymbol{\omega}), \quad (48)$$

where  $\mathcal{F}[\cdot]$  is the Fourier transform. The equations for multiplication and marginalisation then become

$$\Psi(\boldsymbol{\omega}) \star \Phi(\boldsymbol{\omega}) = \Psi(\boldsymbol{\omega}) * \Phi(\boldsymbol{\omega}). \quad (49)$$

where  $*$  is the convolution operator.

Substituting into the general form of the marginalisation equation yields

$$M_y(\Psi(\boldsymbol{\omega}, \boldsymbol{\theta})) = \int e^{2\pi i \mathbf{x} \cdot \boldsymbol{\omega}} \left( \int \left| \int e^{-2\pi i (\mathbf{x} \cdot \boldsymbol{\omega} + \mathbf{y} \cdot \boldsymbol{\theta})} \Psi(\boldsymbol{\omega}, \boldsymbol{\theta}) d\boldsymbol{\omega} d\boldsymbol{\theta} \right|^2 d\mathbf{y} \right)^{\frac{1}{2}} d\mathbf{x}. \quad (50)$$

Marginalisation in this treatment differs from [3] due to the modulus squared of the wavefunction being used to represent the PDF, rather than the wavefunction itself. Unfortunately, in this case it means that marginalisation cannot be simplified to avoid a co-ordinate change into the  $\delta$  basis, thereby destroying the sparsity of the representation. The essence of the problem is that squaring does not commute with the Fourier transform. Therefore, the planewave basis is not suitable for an implementation of Bayesian filtering using this treatment.

For illustrative purposes, figure 1 shows 2 PDFs and their wavefunction representations in the planewave basis; an approximate Dirac delta function, and an ‘‘egg-carton’’ planewave function given by  $P(x, y) = \sin(kx) \sin(ky)$ . The functions were initially defined on a  $100 \times 100$  grid. The need to choose a basis suitable for the particular PDFs being represented is apparent: The delta function representation is dense in the planewave basis, while the planewave PDF is of course defined by a single co-efficient in the planewave basis.

## 5 Implementation

A target tracker with a fixed bearing-only sensor was implemented using wavefunctions in the natural basis  $\{\delta(\mathbf{x} - \mathbf{x}_0)\}$ . This is effectively equivalent to a grid-based representation of the PDFs, and hence would not scale to higher dimensional problems. The state vector is a simple  $(x, y)$  position coordinate, while the target motion model is

$$\mathbf{x}' = A\mathbf{x} + \mathbf{q} \quad (51)$$

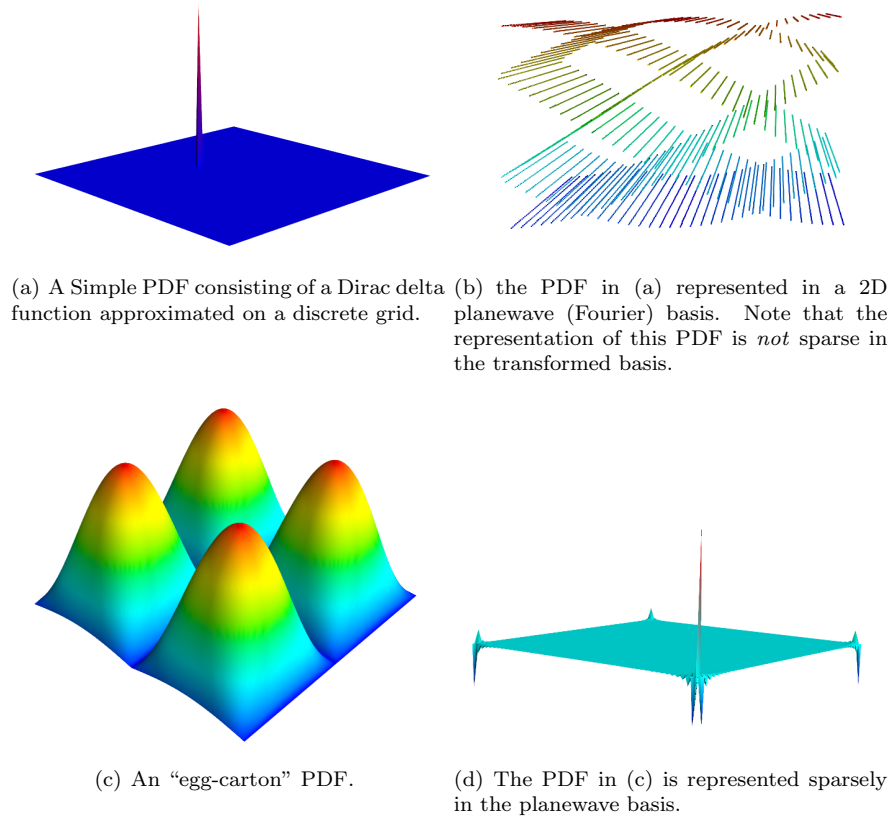


Figure 1: Two different probability density functions, and their (real) wavefunction representations in the planewave basis.

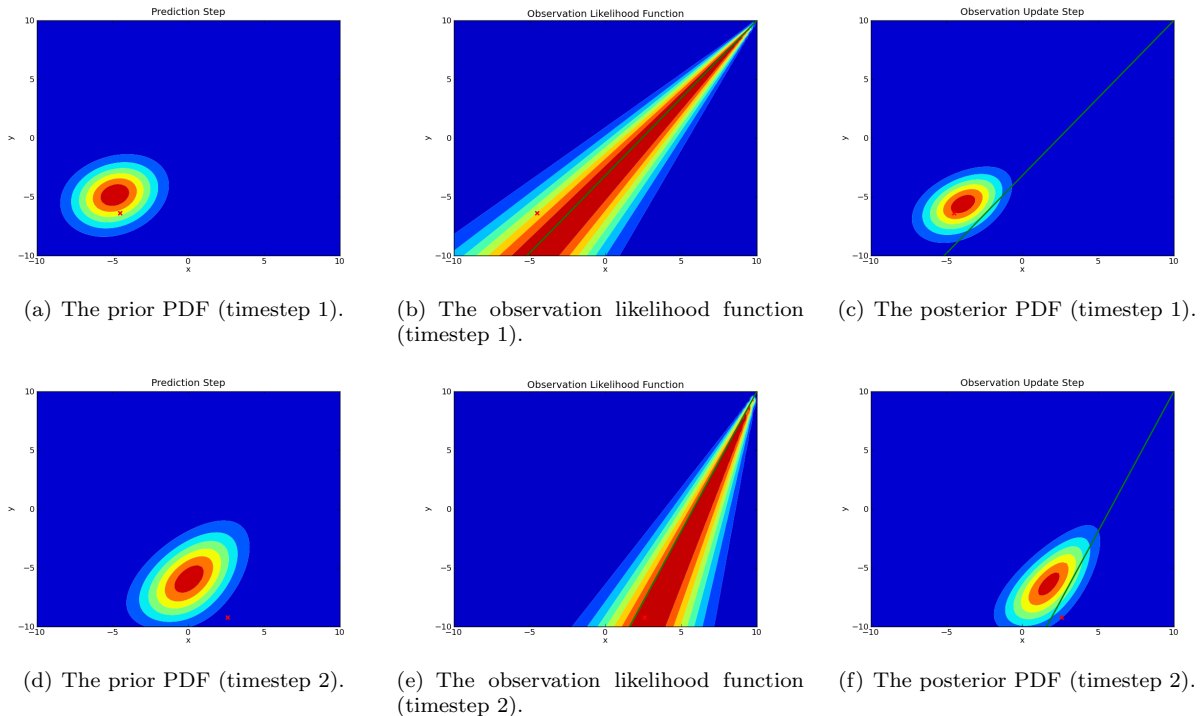


Figure 2: Two timesteps of bearing-only tracking of a moving target from a fixed sensor. The true position of the particle is denoted with a red cross, and the actual bearing measurement with a line.

where  $A$  is a 2x2 matrix,  $v$  is a translation and  $Q$  is a normally-distributed random vector of mean  $v$ .

The observation model is a device periodically taking normally distributed bearing measurements centered about the true location of the target.

The implementation stores the state wavefunction  $\psi(\mathbf{x}_k)$  as an  $N \times N$  array, and pre-calculates the prediction model wavefunction  $\phi(\mathbf{x}_{k+1}|\mathbf{x}_k)$  in an  $N^4$  array.

The prediction step of the filter involves multiplying the two arrays  $\psi(\mathbf{x}_k)$  and  $\phi(\mathbf{x}_{k+1}|\mathbf{x}_k)$  then performing a wavefunction marginalisation over  $\mathbf{x}_k$ ;

$$\psi_{i,j}^+ = \left( \sum_{l=0}^N \sum_{m=0}^N |\psi_{i,j} \phi_{i,j,l,m}|^2 \right)^{\frac{1}{2}}. \quad (52)$$

The observation update step then involves multiplying  $\psi^+$  by the likelihood wavefunction  $\Lambda(\mathbf{x})$  (which in this implementation is not stored in memory but called element-wise), then normalising such that

$$\psi(\mathbf{x}_{k+1})_{i,j} = \frac{1}{C} \Lambda(x_i, y_j) \psi_{i,j}^+, \quad (53)$$

where

$$C = \left( \sum_{i,j} |\Lambda(x_i, y_j) \psi_{i,j}^+|^2 \right)^{\frac{1}{2}}. \quad (54)$$

A pseudo-code description of the algorithm is as follows. Here a “wf” preceding the variable name indicates a wavefunction. The prediction model wavefunction `wf_predict_model` is first calculated from the prediction model equation, which in this basis simply means taking the square root. Next the initial value of the state wavefunction `wf_state` is calculated from the addition of noise to the true state. The main loop then begins. The true state is updated from the transition matrix and the process noise, then the joint wavefunction `wf_joint` is calculated by multiplying the prediction model wavefunction and the state wavefunction.

The predicted state wavefunction `wf_predict` is found by squaring, marginalising away the old state, and square-rooting. Next a simulated measurement is taken based on the true state, and the likelihood wavefunction `wf_likelihood` is calculated by square-rooting the measurement likelihood. Finally the state wavefunction is updated by multiplying the prediction wavefunction by the likelihood wavefunction.

Clearly, the natural basis is totally unsuitable for a real implementation; this pseudo-code is equivalent to a grid-based discrete solution to the Bayesian filtering equations, with a superfluous square-root and square operations that ultimately negate each-other. It mainly serves to explicitly set out how the algorithm *would* look if a basis could be found which allowed for tractable marginalisation and multiplication of wavefunctions.

#### Bearing-Only Tracking Pseudo-code

```
wf_predict_model = sqrt(prediction_model)
wf_state = initial_true_state + noise
LOOP
  true_state = true_state * transition + noise
  wf_joint = wf_predict_model * wf_state
  pdf_joint = wf_joint^2
  pdf_marginalised = sum_last_axes(pdf_joint)
  wf_predict = sqrt(pdf_marginalised)
  measurement = bearing_measurement(true_state)
  wf_likelihood = sqrt(likelihood(measurement))
  wf_state = wf_predict * wf_likelihood
END
```

## 5.1 Simulation Parameters

A small simulation was run with  $N = 100$  points, over the range  $[-10, 10] \times [-10, 10]$ . The number of points is constrained by the prediction model, which being 4-dimensional in this case requires  $N^4 = 10^8$  points. The bearing noise covariance was set at 0.1, while the process noise mean was 0.8 and the covariance 1.0 in  $x$  and  $y$  with no cross-correlation.

## 6 Conclusion

This paper outlines the general form of the equations for Bayesian filtering with wavefunctions. The critical question which must be answered is, is it possible to find an orthonormal function basis i) which sparsely represents PDFs in a given application, and ii) for which the wavefunction filtering equations can actually be solved. Future work would aim to first find representations satisfying ii), and then compare these representations experimentally, measuring how sparsely they are able to approximate PDFs found in real-world Bayesian filtering applications.

## References

- [1] D. Alspach and H. Sorenson. Nonlinear Bayesian Estimation Using Gaussian Sum Approximations. *IEEE transactions on automatic control*, 17(4):439–448, 1972.
- [2] Y. Bar-Shalom and T.E. Fortmann. *Tracking and data Association*. Academic Press Professional, Inc., 1987.
- [3] D. Brunn, F. Sawo, and U. D. Hanebeck. Nonlinear Multidimensional Bayesian Estimation with Fourier Densities. *Proceedings of the 45th IEEE Conference on Decision and Control*, (1):1303–1308, 2006.
- [4] S. Challa, Y. Bar-Shalom, and V. Krishnamurthy. Nonlinear Filtering via Generalized Edgeworth Series and Gauss-Hermite Quadrature. *IEEE Transactions on Signal Processing*, 48(6):1816–1820, 2000.

- [5] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 1981.
- [6] I. Dumitriu, A. Edelman, and G. Shuman. Mops: Multivariate orthogonal polynomials (symbolically). *Journal of Symbolic Computation*, 42(6):587–620, 2007.
- [7] C.F. Dunkl and Y. Xu. *Orthogonal Polynomials of Several Variables*. Cambridge University Press, 2001.
- [8] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEEE Proceedings-F*, 140(2):107–113, 1993.
- [9] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems 1. *Transactions of the ASME Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- [10] A. Pinheiro and B. Vidakovic. Estimating The Square Root of a Density via Compactly Supported Wavelets. *Computational Statistics and Data Analysis*, 25(4):399–416, 1997.
- [11] B.W. Silverman. *Density estimation for statistics and data analysis*. Chapman & Hall/CRC, 1986.
- [12] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, September 2005.