

# Chapter 1

## Reduced Basis Methods for Parameterised Partial Differential Equations

Numerical methods are used to solve PDEs, however for complicated problems the number of parameters required may be quite large, resulting in prohibitive computational cost. The reduced basis method is a means of reducing the number of degrees of freedom in order to reduce costs and make solving the problem feasible. Sometimes we are interested in the field that is the solution to the PDE, perhaps a temperature function. Other times we are interested in a function of this field, perhaps the maximum temperature. This method evaluates the relationship between the input parameters and the output function, hopefully bypassing many of the expensive computations. This is useful in many real-time and multiple query contexts.

### 1.1 Problem Formulation

Consider a PDE with  $P \geq 1$  parameters and let  $\mathcal{D} \subset \mathbb{R}^P$  denote the parameter domain such that  $\vec{\mu} \in \mathcal{D}$  is a particular parameter  $P$ -tuple. These parameters describe properties of the system. In the case of binary coalescence systems there are eight parameters: the mass of each of the compact objects, the spin magnitude, and the spin orientation (2 parameters for each object).

The solutions to the PDE system reside on a manifold defined as

$$\mathcal{M} = \{u(\mu) : \mu \in \mathcal{D}\} \tag{1.1}$$

The PDE can be written

$$F_\mu(u(\mu)) = 0 \quad (1.2)$$

where  $F_\mu$  is a parameter dependent differential operator and  $u(\mu)$  is the solution to the system for parameter  $\mu$ .

Recall (from DG chapter) the dual space  $V^*$  of a vector space  $V$  is the space of (bounded?) linear functionals (aka linear forms)  $g : V \rightarrow \mathbb{R}$  over  $V$ . The norm over  $V^*$  is

$$\|g\|_{V^*} \equiv \sup_{v \in V} \frac{g(v)}{\|v\|_V}$$

The Reisz Representation Theorem tells us that for  $g \in V^* \exists! w_g$  s.t.  $(w_g, v)_V = g(v) \forall v \in V$ . We can then invoke the Cauchy Schwarz inequality to show that  $\|g\|_{V^*} = \|w_g\|_V$ .

For a bilinear form  $b : V \times V \rightarrow \mathbb{R}$  we define the symmetric part  $b_s(w, v) = \frac{1}{2}(b(w, v) + b(v, w))$ , and the skew symmetric part  $b_{ss} = \frac{1}{2}(b(w, v) - b(v, w))$ . It is *coercive* if

$$\alpha = \inf_{w \in V} \frac{b(w, w)}{\|w\|_V^2} > 0 \quad (1.3)$$

Where  $\alpha$  is the *coercivity constant*. It is continuous if

$$\gamma = \sup_{w \in V} \sup_{v \in V} \frac{b(w, v)}{\|w\|_V \|v\|_V} < \infty \quad (1.4)$$

where  $\gamma$  is the *continuity constant*.

$g : V \times \mathcal{D} \rightarrow \mathbb{R}$  is a *parametric linear form* if  $\forall \mu \in \mathcal{D}$ ,  $g(\cdot; \mu) : V \rightarrow \mathbb{R}$  is linear and continuous if  $\forall \mu \in \mathcal{D}$ ,  $g(\cdot; \mu) \in V^*$ . A parametric bilinear form is defined similarly, and is coercive if

$$\alpha(\mu) = \inf_{w \in V} \frac{b(w, w; \mu)}{\|w\|_V^2} > 0 \quad \forall \mu \in \mathcal{D} \quad (1.5)$$

Then we define  $\alpha_0 \equiv \min_{\mu \in \mathcal{D}} \alpha(\mu)$ . A parametric bilinear form is continuous if

$$\gamma(\mu) = \sup_{w \in V} \sup_{v \in V} \frac{b(w, v; \mu)}{\|w\|_V \|v\|_V} < \infty \quad \forall \mu \in \mathcal{D} \quad (1.6)$$

and we define  $\gamma_0 \equiv \max_{\mu \in \mathcal{D}} \gamma(\mu)$ .

A parametric bilinear form has *affine parameter dependence* if it can be written in the form

$$b(w, v; \mu) = \sum_{q=1}^{Q_b} \Theta_b^q(\mu) b^q(w, v) \quad (1.7)$$

for some finite  $Q_b$ , and where  $\Theta_b^q : \mathcal{D} \rightarrow \mathbb{R}$  are smooth and  $b^q : V_1 \times V_2 \rightarrow \mathbb{R}$  are parameter independent continuous bilinear forms. A coercive affine parametric bilinear form is *parametrically coercive* if  $b_s$  can be written

$$b_s(w, v; \mu) = \sum_{q=1}^{Q_{b_s}} \Theta_{b_s}^q(\mu) b_s^q(w, v)$$

with  $\Theta_{b_s}^q(\mu) > 0 \quad \forall \mu \in \mathcal{D}$  and  $b_s^q(v) \geq 0 \quad \forall v \in V$ .

The *inf-sup stability "constant"* of a parametric bilinear form is defined as

$$\beta(\mu) = \inf_{w \in V_1} \sup_{v \in V_2} \frac{(w, v; \mu)}{\|w\|_{V_1} \|v\|_{V_2}} \geq 0. \quad (1.8)$$

If  $\exists \beta_0 > 0$  s.t.  $\beta(\mu) \geq \beta_0 \quad \forall \mu \in \mathcal{D}$ , then we say  $b$  is *inf-sup stable* over  $V_1 \times V_2$ .

The *supremising operator*  $T_\mu : V_1 \rightarrow V_2$  associated with  $b$  is defined as

$$T_\mu w = \arg \sup_{v \in V_2} \frac{b(w, v; \mu)}{\|v\|_{V_2}} \quad (1.9)$$

Note that  $(T_\mu w, v)_{V_2} = b(w, v; \mu) \quad \forall v \in V_2$ . If  $b$  admits an affine representation, then

$$T_\mu w = \sum_{q=1}^{Q_b} \Theta_b^q(\mu) \mathbb{T}^q w \quad (1.10)$$

where  $\mathbb{T}^q : V_1 \rightarrow V_2$  is given by  $(\mathbb{T}^q w, v)_{V_2} = b^q(w, v) \quad \forall v \in V_2$ .

## 1.2 Parameter Domains and Grids

Consider our parameter domain  $\mathcal{D} \subset \mathbb{R}^P$ , and a parameter value  $\mu \in \mathcal{D}$ . We define  $\mathcal{D}_{box} \subset \mathbb{R}^P$  to be the smallest  $P$  dimensional parallelepiped such that  $\mathcal{D} \subset \mathcal{D}_{box}$ ,  $\mathcal{D} \equiv [\mu_1^{min}, \mu_1^{max}] \times \dots \times [\mu_P^{min}, \mu_P^{max}]$ . If  $\mu^{min} = \min_{p \in 1 \dots P} \mu_p^{min} > 0$  we can perform a logarithmic transformation  $\hat{\mu}_p = \ln \mu_p$ , and  $\hat{\mathcal{D}}, \hat{\mathcal{D}}_{box}$  are the images of  $\mathcal{D}, \mathcal{D}_{box}$  under this transformation.

Three ways to draw points over  $\mathcal{D}$  are as follows:

*Monte Carlo* sampling to create a uniform distribution of  $m$  points over  $\mathcal{D}$ . To create a linear grid  $G_{[MC; m]}^{lin}$  parameter values are generated by setting

$$\mu_p = \mu_p^{min} + \text{rand}(\mu_p^{max} - \mu_p^{min}) \quad (1.11)$$

where  $\text{rand}$  is a random variable uniformly distributed over  $[0, 1]$ . Any  $\mu = [\mu_1 \dots \mu_P] \notin \mathcal{D}$  is rejected. To create a uniform distribution over the logarithmic space  $\hat{\mathcal{D}}$ , we sample the points as

$$\mu_p = \mu_p^{\min} \exp \left\{ \text{rand} \times \ln \left( \frac{\mu_p^{\max}}{\mu_p^{\min}} \right) \right\}. \quad (1.12)$$

*1D deterministic grids:*  $G_{[z_1, z_2; m]}^{\text{lin}}, G_{[z_1, z_2; m]}^{\text{ln}} : z_1 \in \mathbb{R} > z_2 \in \mathbb{R}, m \in \mathbb{N}$ :

$$\begin{aligned} G_{[z_1, z_2; m]}^{\text{lin}} &= \{z_1 + \frac{i-1}{m-1}(z_2 - z_1), 1 \leq i \leq m\} \\ G_{[z_1, z_2; m]}^{\text{ln}} &= \{z_1 \exp \left( \frac{i-1}{m-1} \ln(z_2/z_1) \right), 1 \leq i \leq m, z_1, z_2 > 0\} \end{aligned} \quad (1.13)$$

*Chebyshev Gauss-Lobatto points*

$$\begin{aligned} G_{[z_1, z_2; m]}^{\text{lin}, \text{Cheb}} &= \{z_1 + \frac{1}{2} \left( 1 - \cos \pi \left( \frac{i-1}{m-1} \right) \right) (z_2 - z_1), 1 \leq i \leq m\} \\ G_{[z_1, z_2; m]}^{\text{ln}, \text{Cheb}} &= \{z_1 \exp \left( \frac{1}{2} \left( 1 - \cos \pi \left( \frac{i-1}{m-1} \right) \ln(z_2/z_1) \right) \right), 1 \leq i \leq m, z_1, z_2 > 0\} \end{aligned} \quad (1.14)$$

1D grids can be combined to create multi-dimensional grids.

### 1.3 Parametric Scalar and vector fields

Consider a parametric scalar or vector field  $w : \Omega \times \mathcal{D} \rightarrow \mathbb{R}^d$  where  $d = 1$  for a scalar field, or the dimension of  $\Omega$  for a vector field. The *parametric derivative*, also known as the *sensitivity derivative*, measures changes in  $w$  as the parameter changes.

$$(D_\sigma w)(\mathbf{x}; \mu) = \frac{\partial^\sigma w}{\partial \mu_1^{\sigma_1} \dots \partial \mu_P^{\sigma_P}} \quad (1.15)$$

where  $\sigma = (\sigma_1, \dots, \sigma_P)$  is an index vector of non negative integers and we denote by  $|\sigma| = \sum_{i=1}^P \sigma_i$  the order of the derivative.  $I^{P,n} = \{\sigma \in \mathbb{N}_0^P \mid |\sigma| \leq n\}$ .

A parametric scalar field, or one component of a parametric vector field is *separable* over  $\Omega$  if for some finite  $M$

$$w(\mathbf{x}; \mu) = \sum_{j=1}^M h^j(\mathbf{x}) g^j(\mu), \quad \forall \mathbf{x} \in \Omega, \mu \in \mathcal{D} \quad (1.16)$$

for  $h^j : \Omega \rightarrow \mathbb{R}$ ,  $g^j : \mathcal{D} \rightarrow \mathbb{R}$ ,  $1 \leq j \leq M$ .  $w$  is  $\mathbf{x}$ -affine separable over  $\Omega$  if each  $h^j(\mathbf{x})$  is affine in  $\mathbf{x}$ .

## 1.4 Problem statement

### 1.4.1 Exact

We consider a parametric PDE over a physical domain  $\Omega \in \mathbb{R}^d$  with boundary  $\partial\Omega$ , where the field variables  $w \in X^e$ , where  $X^e$  is our exact solution space, may be scalar  $w : \Omega \rightarrow \mathbb{R}$  or vector-valued  $w : \Omega \rightarrow \mathbb{R}^d$ . We denote by  $\Gamma_i^D$  boundary measurable segments of  $\partial\Omega$  on which we impose Dirichlet boundary conditions.

We construct the exact solution space  $X^e$  as the Cartesian product of scalar spaces  $Y_i^e \equiv v \in H^1(\Omega) | v|_{\Gamma_i^D} = 0$  where  $1 \leq i \leq d$  if the field variable is vector-valued, or simply  $i = 1$  if the field variable is scalar, in which case the subscript can be omitted. An element  $w \in X^e$  is denoted  $w = (w_1, \dots, w_d)$ . We have a number of options for the inner product and norm with which to equip  $X^e$ .

**Energy norm** (for coercive and continuous  $a$ ):

$$((w, v)) = a(w, v), \quad \forall w, v \in X^e \quad (1.17)$$

$$|||w||| = \sqrt{a(w, w)}, \quad \forall w \in X^e \quad (1.18)$$

We can then define the parametric field variable as  $u : \Omega \times \mathcal{D} \rightarrow \mathbb{R}$  where for a given  $\mu \in \mathcal{D}$ ,  $u(\cdot; \mu) \in X^e$ .

### Parametric weak form

Suppose we are given parametric linear forms  $f$  and  $l$  that are bounded over  $X^e$ , and an affine parametric bilinear form  $a$  that is inf-sum stable and continuous over  $X^e$ , and which we assume is parametrically coercive. Then the problem is given  $\mu \in \mathcal{D}$  we find  $u^e(\cdot; \mu) \in X^e$  such that

$$a(u^e(\mu), v; \mu) = f(v; \mu), \quad \forall v \in X^e \quad (1.19)$$

and evaluate our output of interest

$$s^e(\mu) = l(u^e(\mu); \mu). \quad (1.20)$$

A problem is *compliant* if

1.  $l(\cdot; \mu) = f(\cdot; \mu), \forall \mu \in \mathcal{D}$  that is the output functional and "load/source are identical, and
2.  $a$  is symmetric.

### 1.4.2 “Truth” Approximation

We introduce a family of *conforming* approximation spaces  $X^{\mathcal{N}} \subset X^e$  of dimension  $\mathcal{N}$ . As in the exact case we form  $X^{\mathcal{N}}$  as a Cartesian product of scalar approximations  $Y_i^{\mathcal{N}_i} \subset Y_i^e$  of dimension  $\mathcal{N}_i$  such that  $X^{\mathcal{N}} = Y_i^{\mathcal{N}_1} \times \cdots \times Y_i^{\mathcal{N}_d}$ ,  $\mathcal{N} = \sum_i \mathcal{N}_i$ . We choose a set of basis functions  $\psi_k^{\mathcal{N}} \in X^{\mathcal{N}}$ ,  $1 \leq k \leq \mathcal{N}$ . The inner products and norms are inherited from the exact case. We require that the family of truth subspaces  $X^{\mathcal{N}}$  satisfies the approximation condition that for any  $\epsilon > 0$ ,  $\exists \mathcal{N}$  such that the error in the best fit to  $u(\mu)$  in  $X^{\mathcal{N}}$  is less than or equal to  $\epsilon$  for all  $\mu$  in  $\mathcal{D}$ :

$$\max_{\mu \in \mathcal{D}} \inf_{w \in X^{\mathcal{N}}} \|u(\mu) - w\|_{X^e} \rightarrow 0 \text{ as } \mathcal{N} \rightarrow \infty \quad (1.21)$$

The family of approximations to the exact problems is given by finding  $u^{\mathcal{N}}(\mu) \in X^{\mathcal{N}}$  such that

$$a(u^{\mathcal{N}}(\mu), v; \mu) = f(v; \mu) \quad \forall v \in X^{\mathcal{N}} \quad (1.22)$$

and evaluating

$$s^{\mathcal{N}}(\mu) = f(u^{\mathcal{N}}(\mu); \mu) \quad (1.23)$$

for a given  $\mu \in \mathcal{D}$ . This is a standard Galerkin projection. We implicitly assume that we can represent the exact geometry of the domain, and that quadratures are exact. From 1.21 we have that as  $\mathcal{N} \rightarrow \infty$ ,  $u^{\mathcal{N}}(\cdot; \mu) \rightarrow u^e(\cdot; \mu)$ . We define the error

$$\epsilon^{\mathcal{N}} = \max_{\mu \in \mathcal{D}} \|u(\cdot; \mu) - u^{\mathcal{N}}(\cdot; \mu)\|_{X^e} \quad (1.24)$$

then  $\epsilon^{\mathcal{N}} \rightarrow 0$  as  $\mathcal{N} \rightarrow \infty$ .

If we expand the solution  $u^{\mathcal{N}}(\mu)$  in our basis  $\psi_k^{\mathcal{N}}$  as

$$u^{\mathcal{N}}(\mathbf{x}; \mu) = \sum_{j=1}^{\mathcal{N}} u_j^{\mathcal{N}}(\mu) \psi_j^{\mathcal{N}}(\mathbf{x}), \quad (1.25)$$

then

$$\underline{u}^{\mathcal{N}}(\mu) \equiv [u_1^{\mathcal{N}} \cdots u_{\mathcal{N}}^{\mathcal{N}}]^T \in \mathbb{R}^{\mathcal{N}} \quad (1.26)$$

and the problem can be written

$$\underline{A}^{\mathcal{N}}(\mu) \underline{u}^{\mathcal{N}}(\mu) = \underline{F}^{\mathcal{N}}(\mu) \quad (1.27)$$

where the stiffness matrix  $\underline{A}^{\mathcal{N}}(\mu) \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$  has elements given by

$$A_{ij}^{\mathcal{N}}(\mu) = a(\psi_j^{\mathcal{N}}, \psi_i^{\mathcal{N}}; \mu) \quad (1.28)$$

and the components of the load/source vector (and in the compliant case, output vector), are given by

$$F_i^{\mathcal{N}}(\mu) = f(\psi_i^{\mathcal{N}}; \mu). \quad (1.29)$$

Invoking the affine properties of  $f$  and  $a$ , we can write these as follows:

$$\begin{aligned} \underline{A}^{\mathcal{N}}(\mu) &= \sum_{q=1}^{Q_a} \Theta_a^q(\mu) \mathbb{A}^{\mathcal{N}_q}, \\ \mathbb{A}_{ij}^{\mathcal{N}_q} &= a^q(\psi_j^{\mathcal{N}}, \psi_i^{\mathcal{N}}) \\ 1 \leq i, j \leq \mathcal{N}, \quad 1 \leq q \leq Q_a \end{aligned} \quad (1.30)$$

and

$$\begin{aligned} \underline{F}^{\mathcal{N}}(\mu) &= \sum_{q=1}^{Q_f} \Theta_f^q(\mu) \mathbb{F}^{\mathcal{N}_q}, \\ \mathbb{F}_i^{\mathcal{N}_q} &= f^q(\psi_i^{\mathcal{N}}), \\ 1 \leq i \leq \mathcal{N}, \quad 1 \leq q \leq Q_f \end{aligned} \quad (1.31)$$

Note that  $\mathbb{F}^{\mathcal{N}_q}$  and  $\mathbb{A}^{\mathcal{N}_q}$  are parameter independent. We also introduce another parameter independent matrix  $\mathbb{X}^{\mathcal{N}}$  :

$$\mathbb{X}_{ij}^{\mathcal{N}} = (\psi_j^{\mathcal{N}}, \psi_i^{\mathcal{N}})_{X^{\mathcal{N}}}, \quad 1 \leq i, j \leq \mathcal{N} \quad (1.32)$$

so that for any two members  $w, v \in X^{\mathcal{N}}$  written in terms of the  $\psi^{\mathcal{N}}$  basis, the  $X^{\mathcal{N}}$ - inner product can be written

$$(w, v)_{X^{\mathcal{N}}} = \underline{w}^T \mathbb{X}^{\mathcal{N}} \underline{v} \quad (1.33)$$

For a family of approximation spaces, we define the "truth approximation space"  $X^{\mathcal{N}}$  to correspond to  $\mathcal{N} = \mathcal{N}_t$ , which is typically quite large, such that  $\epsilon^{\mathcal{N}_t}$  is smaller (in the worst case scenario) than some error threshold. We can then consider the parametrically induced manifold  $\mathcal{M} \equiv \{u^{\mathcal{N}_t}(\mu) | \mu \in \mathcal{D}\}$ , which is typically of comparably low dimension. We seek to represent the solution  $u^{\mathcal{N}}(\cdot; \mu)$  in terms of elements of the space  $\text{span}\{\mathcal{M}^{\mathcal{N}_t}\}$ .

## 1.5 Reduced Basis Approximation

### 1.5.1 Lagrangian approach

For  $\mu \in \mathcal{D}$ , approximate  $u^{\mathcal{N}_t} = u^{\mathcal{N}_t}(\mathbf{x}; \mu)$  by a linear combination of  $N$  precomputed “snapshots” on  $\mathcal{M}^{\mathcal{N}_t}$ :  $u^1, \dots, u^N$ . We specify the maximum dimension of the reduced basis space  $N_{\max}$ . We introduce a set of parameter points  $\mu^n \in \mathcal{D}$  and define our RB samples sets

$$S_N = \{\mu_1, \dots, \mu_N\} \quad (1.34)$$

Note that these sets are nested  $S_1 \subset S_2 \subset \dots \subset S_{N_{\max}} \subset \mathcal{D}$ . The “snapshots” are then

$$u^n \equiv u(\cdot; \mu^n) \quad (1.35)$$

in terms of which we define our Lagrange RB spaces:

$$W_N = \text{span}\{u^n, 1 \leq n \leq N\} \quad 1 \leq N \leq N_{\max} \quad (1.36)$$

These spaces are hierarchical. It is also possible to define the RB spaces in other ways, for example the Taylor RB spaces are generated by both the field  $u^n$  and the field sensitivity derivatives- the derivative of  $u(\cdot; \mu^n)$  with respect to  $\mu$ - at a particular point in  $\mathcal{D}$ , while the Hermite RB spaces combine both of these and consider the field and sensitivity derivatives at several points in  $\mathcal{D}$ .

The snapshots can be written in terms of the finite element basis function  $\psi_i^{\mathcal{N}_t}$  as per equation 1.25. We can then construct an orthogonal basis for the Lagrange RB space using Gram Schmidt orthogonalisation with  $\mathbb{X}$  the inner product matrix defined in 1.32.:

We can now define “basis” matrices  $\mathbb{Z}_N^{\mathcal{N}_t} \in \mathbb{R}^{\mathcal{N}_t \times N}$  as

$$\mathbb{Z}_{Njn}^{\mathcal{N}_t} = \hat{u}_j^n \quad (1.37)$$

That is the  $n$ th column of  $\mathbb{Z}_{N_{\max}}^{\mathcal{N}_t}$  contains the vector of truth approximation finite element basis coefficients associated with the  $n$ th RB basis function. Note that  $\mathbb{Z}_n^{\mathcal{N}_t}$  is a principal submatrix of  $\mathbb{Z}_{n+1}$ . We can express the orthogonality condition as

$$\mathbb{Z}_{N_{\max}}^T \mathbb{X} \mathbb{Z}_{N_{\max}} = \mathbb{I}_{N_{\max}} \quad (1.38)$$

We can now express our problem in the RB approximation space. We look for  $u_{X_N} \in X_N$  such that

$$a(u_{X_N}, v; \mu) = f(v; \mu), \quad \forall v \in X_N \quad (1.39)$$



and evaluate

$$s_{X_N}(\mu) = f(u_{X_N}(\mu); \mu) \quad (1.40)$$

## 1.6 Reduced basis for binary coalescences

Binary coalescence systems are described by eight parameters: the mass of each of the compact objects, the spin magnitude, and the spin orientation (2 parameters for each object).

Search via matched filtering- comparing possible waveform templates to actual data from detector. For low mass systems the inspiral part of the coalescence is at detectable frequencies. At higher masses, the merger part is (this requires NR techniques). These are expensive, therefore we search for the most relevant points in the parameter space. Searching is also expensive AND time critical so that events can be matched with EM astronomy.

GW is dependent on time (or frequency) and  $P$  parameters  $\vec{\mu} = \mu_1, \dots, \mu_P$ . Denote a GW by  $h_{\vec{\mu}}$ .  $\mathcal{H}$  is the space of all normalised GWs for a source. This space is non-linear, but can be represented by a linear space to arbitrarily high accuracy.

Approximate  $\mathcal{H}$  with best linear combinations of members  $\Psi_i \equiv h_{\vec{\mu}_i}$  of a catalogue  $C_N = \{\Psi_i\}_{i=1}^N$ . The reduced basis space is  $W_N = \text{span}(C_N)$ .

Optimal method: Choose waveforms to make this catalog so that the error in representing  $\mathcal{H}$  with  $W_N$  is minimised over the choice of  $N$  catalogue members. Optimal error given by Kolmogorov  $N$  width:

$$d_N(\mathcal{H}) = \min_{C_N} \max_{\vec{\mu}} \min_{u \in W_N} \|u - h_{\vec{\mu}}\|$$

Where the norm is calculated from the complex inner product such that for two waveforms  $F$  and  $G$  in Fourier space

$$\langle F, G \rangle \equiv \int_{f_L}^{f_U} \frac{F^*(f)G(f)}{S_n(f)} df$$

where  $S_n(f)$  is the power spectral density of the detector.

In reality, this approach is too computationally expensive. Use greedy method.

## 1.7 Greedy Algorithm

A greedy algorithm solves a problem by making the locally optimal choice at each stage in the hope that this will lead to a global optimum.

Choose a waveform with arbitrary parameter value. Basis vector  $e_1 = h_{\vec{\mu}_1}$  is identified with this waveform, and the first catalogue  $C_1 = \Psi_1 = h_{\vec{\mu}_1}$ . Subsequent catalogues are built by adding another waveform by finding

$$\vec{\mu}_j = \arg \min_{\mu} \|h_{\mu} - P_{j-1}(h_{\mu})\| \quad (1.41)$$

where  $P_N(h_{\mu}) = \sum_{i=1}^N e_i \langle e_i, h_{\mu} \rangle$  is the representation of  $h_{\mu}$  using the catalogue  $C_N$  obtained by finding orthogonal projection of  $h_{\mu}$  onto  $W_N$ . Note that this creates a hierarchical sequence of catalogues  $C_1 \subset C_2 \dots$

The *greedy error*  $\epsilon_N$  is the maximum error for a catalog  $C_N$ , given by

$$\epsilon_N \equiv \max_{\mu} \|h_{\vec{\mu}} - P_N(h_{\mu})\| \quad (1.42)$$

It can be shown that if the decay of the  $N$ -width can be bounded by an exponential,  $d_N(\mathcal{H}) \leq A \exp^{-cN^{\alpha}}$  for some real  $c$  and  $\alpha$ , then  $\epsilon_N \leq \tilde{A} \exp^{-dN^{\beta}}$  where  $d$  and  $\beta$  depend on  $c$  and  $\alpha$ .