Research Statement

Huanhuan Yang

My primary research in computational Mathematics focuses on numerical partial differential equations (PDEs), with a particular emphasis on PDE-constrained optimization, reduced order modeling (ROM), and stochastic PDEs. I am also experienced in high performance computing for engineering problems in cardiac electrophysiology, acoustic modeling, ocean modeling, and computational fluid dynamics. I have a broad interest in numerical PDEs, and would like to conduct research continuously in this field.

Projects I have been working on include: (1) developing and analysing variational data assimilation techniques for parameter estimation in electrophysiology; (2) applying reduced-order modeling techniques for cardiac conductivity estimation; (3) developing a data-driven model reduction technique for atrial electrophysiology; (4) solving stochastic optimal control in acoustics with conditional value-at-risk (CVaR) measure and parallel reduced-order modeling; (5) developing fast Centroidal Voronoi tessellation (CVT) grid generation algorithms for ocean modeling. The first two projects were the focus of my dissertation work, the third one is my internship project, and the last two are my postdoctoral research. Below I give a brief summary of the projects and discuss future research plans.

0 Challenges in computational electrocardiology

Computational modeling of healthy and diseased electrophysiology (EC) has a great potential to provide non-invasive, cost effective and personalized assessment of the state of the heart, for improved diagnosis and prognosis of cardiac arrhythmia. In this well-established field, the bidomain model is currently the most physiologically founded description for the dynamics of cardiac electric potentials—the transmembrane potential $u$ and the extracellular potential $u_e$—at the level of cardiac tissue. Its parabolic-elliptic form [9] reads

$$\begin{align}
\beta C_m \frac{\partial u}{\partial t} - \nabla \cdot (\sigma_l \nabla u) - \nabla \cdot (\sigma_i \nabla u_e) + \beta I_{\text{ion}} &= I_{si} \\
-\nabla \cdot (\sigma_i \nabla u) - \nabla \cdot (\sigma_e + \sigma_{se}) \nabla u_e &= I_{si} - I_{se}
\end{align}$$

(1)

where $C_m$ is the membrane capacitance per unit area and $\beta$ denotes the surface-to-volume ratio of the membrane; $\sigma_i$ (resp. $\sigma_e$) is the intracellular (resp. extracellular) conductivity tensor; $I_{si}$ (resp. $I_{se}$) represents the intracellular (resp. extracellular) stimulation current. The total ionic current $I_{\text{ion}}$ is given nonlinearly by a cellular ionic model in a general form

$$\begin{align}
I_{\text{ion}} &= \sum_{i=1}^{M} I_{x_i}(u, w, c) \\
\frac{dw}{dt} + g(u, w, c) &= 0 \\
\frac{dc}{dt} + h(u, w, c) &= 0
\end{align}$$

(2)

where $I_{x_i}$ is the ionic current associated with ion species $x_i$, $w$ is the vector of gating variables and $c$ is the vector of typical ion concentrations. We refer to Fig. 1 (left) for a typical action potential of a cardiac myocyte.

The conductivity tensors can be represented by referring to the cardiac fibers: $\sigma_k(x) = \sigma_{ik} a_i(x) a_k(x)^T + \sigma_{kt} a_t(x) a_t(x)^T + \sigma_{kn} a_n(x) a_n(x)^T$, where $k$ stands for $i$ or $e$, $(a_i, a_k, a_n)$ are orthonormal vectors related to the structure of the myocardium with $a_i$ parallel to the local fibre direction. We may assume that the tissue is axial isotropic (i.e. $\sigma_{kn} = \sigma_{kt}$), as done by groups for conductivity measurement in laboratory experiments.

To overcome high computational costs associated with the bidomain problem, a simplified monodomain model has been proposed. Its derivation [8] is based upon a proportionality assumption $\sigma_e = \lambda \sigma_i$. Denote $\sigma_m = \frac{1}{1+\lambda} \sigma_i$ and $I_{\text{app}} = \frac{\lambda}{1+\lambda} I_{si} + \frac{1}{1+\lambda} I_{se}$, we have:

$$\beta C_m \frac{\partial u}{\partial t} - \nabla \cdot (\sigma_m \nabla u) + \beta I_{\text{ion}} = I_{\text{app}}.$$

(3)

Although this assumption lacks physiological foundation, the monodomain model has been extensively used in clinic-oriented simulations. Moreover, it is concluded [2] that the discrepancy between the bidomain and monodomain models has order 1% or less in terms of activation time relative error.

Recent computational methods in EC usually suffer from three major limitations that hinder their clinical use: (1) lacking of efficient model personalization strategies; (2) high computational demanding from the EC solver; (3) lacking of good trade-off between the simplification of cellular ionic models and the demand on keeping sufficient biophysical details. In the following, our solutions to above challenges are described.
1 Conductivity estimation by variational data estimation

Electrocardiological models depend strongly on model parameters and in particular on the cardiac conductivities. Unfortunately, it is quite problematic to measure these parameters in vivo and even more so in clinical practice, resulting in no common agreement in the literature. Recent work on variational parameter estimation in electrophysiology utilized either a derivative-free optimization approach or a least-squares approach which usually involves a large number of optimization iterations.

In [16] we consider a variational data assimilation approach and investigate the bidomain inverse conductivity problem (BICP): find $\sigma = (\sigma_{il}, \sigma_{el}, \sigma_{it}, \sigma_{et})$ in an admissible domain $\mathcal{C}_{ad}$ minimizing the misfit functional

$$J(\sigma) = \frac{1}{2} \int_{0}^{T} \int_{\Omega_{obs}} \left( u(\sigma) - u_{\text{meas}} \right)^2 + \left( u_{c}(\sigma) - u_{c,\text{meas}} \right)^2 \, dx \, dt + \frac{\alpha}{2} \mathcal{R}(\sigma)$$

subject to the bidomain equations and a coupled ionic model. Here $u_{\text{meas}}$ and $u_{c,\text{meas}}$ denote the experimental data measured on the observation domain $\Omega_{\text{obs}} \subset \Omega$: $\mathcal{R}$ is a regularization term. The existence of a minimizer of the misfit function is proved with the phenomenological Rogers–McCulloch ionic model [12], that completes the bidomain system. The core of our numerical results is in 3D, on both idealized and real geometries, with the minimal ionic model [3] for more realistic cases—such as an appropriate trade-off between reliability and efficiency. At the best of our knowledge this is the first time variational techniques are used in 3D real geometries to demonstrate the method in cases of clinical interest. We showed the reliability and the stability of the conductivity estimation approach in the presence of noise and with an imperfect knowledge of other model parameters.

We significantly improve the numerical approaches in the literature by resorting to a derivative-based optimization method (see an optimization iteration in Fig. 1 right). The gradient of the misfit functional to minimize is computed by resorting to the adjoint equations of the bidomain system. For the challenge presented by differentiating state-dependent discontinuous terms, we use shape calculus for computing those Gâteaux differentials. Specifically, let $H(\cdot)$ be the standard Heaviside step function intended to model switch-like dynamics in the minimal ionic model (or any other biophysics-based models), the adjoint system of the bidomain equations contains a general differentiation form $\langle b(x)DH(u-c), \varphi \rangle_{L^2}$ with $DH(u-c)$ being the Gâteaux differential of $H(u-c)$, where $c \in \mathbb{R}$ is a constant. If we denote $j_{b,c}(u) = \int_{\Omega} b(x)H(u(x) - c) \, dx$ and its Gâteaux derivative in the direction of $\varphi$ as $D_{\varphi} j_{b,c}(u)$, the differentiation form can be computed through

$$\langle b(x)DH(u-c), \varphi \rangle_{L^2} = \int_{\Omega} b(x)D_{\varphi} H(u-c) \, dx = D_{\varphi} j_{b,c}(u) = \int_{\Gamma_{u(c)}} \frac{b(x)\varphi(x)}{\nabla u(x) \cdot \mathbf{n}} \, d\Gamma$$

where $\mathbf{n}$ denotes the outward unit normal to the domain $\{x : u(x) \geq c\}$ and we assume that $\nabla u(x) \neq 0$ in a neighbourhood of $\Gamma_{u(c)} = \{x : u(x) = c\}$.

Current/Future Work: Follow-up of this work is an experimental validation of the estimation procedure in view of clinical applications. In our future plan, we would like to consider more patient-dependent parameters, like the elevation angle and the transverse angle of the fibers, piecewise-constant conductivities, conductivities in different anatomical structures such as the Purkinje fiber and the bundle of His. After the model personalization step, in collaboration with medical doctors we will carry out applications on patient’s heart geometries, such as optimal pacing maker localization and cardiac resynchronization therapy.

We also intend to pursue the rigorous quantification of uncertainties induced by the presence of noise. In this case, the estimated conductivities will be defined by a probability density function whose moments depend on the noise and the bidomain problem. A statistical estimator of cardiac conductivities will be based on a variational Bayesian approach.

![Diagram](image-url)
2 Reduced-order modeling for cardiac conductivity estimation

The intrinsic complexity of the heart anatomy and the complex interplay among cell, tissue and organ modeling scales make the requirement on computational efficiency of simulation hard to be satisfied. This is even more true for an inverse problem of parameter estimation, as high computational cost arises in many “queries” of forward simulations with different model parameters. Model reduction in electrocardiology is generally challenged by nonlinearity of the models and the exceptional feature like wave-front propagation of their solutions. This prevents “classical” approaches like Proper Orthogonal Decomposition (POD) to be promptly applied. The work in [17] gives a first contribution to model reduction applying to the inverse conductivity problem.

The practical application of POD in electrocardiology only starts from 2011 [1, 4]. In these references, the POD method allows reasonable estimation of cardiac ionic model parameters, however, no systematic study is available on the improvement of efficiency of solving the full nonlinear electrocardiological model. In fact, when reducing a nonlinear problem by projecting onto a low-dimensional space, one critical aspect is to approximate the projected nonlinear terms in a way independent of the full-order model size. This point was not thoroughly addressed in current electrocardiology publications.

In [17] we explore model-order-reduction techniques to fit the estimation procedure into timelines of clinical interest. Specifically we consider the monodomain model and resort to POD techniques to take advantage of an off-line step when solving iteratively the electrocardiological forward model online. In addition, we perform the Discrete Empirical Interpolation Method (DEIM) to tackle the nonlinearity of the model. It is worth mentioning that the conductivity parameter to be estimated considered in this work is more troublesome than other ionic model parameters, since it dominates the speed and direction of fast transient of electrical potential through the cardiac tissue, which is an intrinsic feature of the forward electrocardiology model. This fact thus prevents a successful model reduction via a classical POD procedure. Model reduction procedures need to be specifically customized for the problem, and in particular the construction of the educated basis is a delicate step. Nevertheless, we show in [17] how an appropriate sampling for the basis computation actually leads to significant reduction of the full-order computational cost with a great level of accuracy. We address the sampling required for basis construction based on the novel concept of “Domain of Effectiveness” in the parameter space. A rather small sample set (see Fig. 2) is obtained by sampling the parameter space based on polar coordinates, with refinement in the “small angle–short arc” zone of the sample space utilizing Gaussian nodes. In this way, we manage to use the POD-DEIM reduced-order model with a computational reduction of at least 95% of the full-order conductivity estimation.

![Figure 2: Left: Ten samples (stars in red) generated by a nonuniform sampling on the polar coordinates of conductivity values. The six blue dots are used as test points. (Image 299x325 to 372x404)](image2)

Figure 2: Left: Ten samples (stars in red) generated by a nonuniform sampling on the polar coordinates of conductivity values. The six blue dots are used as test points.

The accuracy of the POD-DEIM framework has been thoroughly investigated at six test points (blue dots in Fig. 2), which were carefully chosen not too close to the sample points. It was shown that the estimation results using ROM are very satisfying with all test points. In particular, we reconstruct the transmembrane potential corresponding to $\mathbf{\sigma}_{\text{exact}} = [3.2, 0.5]$ and adding 15% uniform noise. Right: reconstruction of the potential computed with the estimated conductivity $\mathbf{\sigma}_{\text{estimated}} = [3.07, 0.425]$.

![Figure 3: Screenshots of $u$ on a real left ventricular geometry reconstructed from SPECT images. The white arrows represent myocardial fiber orientation. Left: synthetic measure of the potential created by simulating with $\mathbf{\sigma}_{\text{exact}} = [3.2, 0.5]$ and adding 15% uniform noise. Right: reconstruction of the potential computed with the estimated conductivity $\mathbf{\sigma}_{\text{estimated}} = [3.07, 0.425]$. (Image 385x325 to 506x402)](image3)

Figure 3: Screenshots of $u$ on a real left ventricular geometry reconstructed from SPECT images. The white arrows represent myocardial fiber orientation. Left: synthetic measure of the potential created by simulating with $\mathbf{\sigma}_{\text{exact}} = [3.2, 0.5]$ and adding 15% uniform noise. Right: reconstruction of the potential computed with the estimated conductivity $\mathbf{\sigma}_{\text{estimated}} = [3.07, 0.425]$.

Current/Future Work: This work opens several interesting challenges to be investigated in future works. In particular, we would like to theoretically quantify the conductivity estimation error caused by reduced-order modeling. An error analysis for some particular optimal control problems has been studied in [5], where the control is on the forcing term or the boundary condition. However, the work on conductivity estimation is still open since the control parameter appears in the differential core of the system. We also plan to extend the sampling strategy proposed here to 3D bidomain inverse conductivity problem.

Another ongoing work is targeting the greedy reduced basis method. After the development of an \textit{a posteriori} error estimator, we need a complete check on its sharpness and computational efficiency. Furthermore, we expect troubles with a single reduced basis, therefore the idea of multiple reduced bases will be verified in the greedy framework.
3 Model order reduction for atrial electrophysiology by statistical learning

In computational EC, it is challenging to keep the cellular model minimally complicated so that it can be solved with little computational effort, but simultaneously make the model sufficiently detailed so that it can reproduce as much clinical data as possible. Recent EC models are capable of describing complex cellular mechanisms, such as detailed intracellular Ca\(^{2+}\) handling. However, these models are computationally demanding due to many coupled ordinary differential equations (ODEs) accounting for different ionic channels. As an example, the Courtemanche–Ramirez–Nattel (CRN) human atrial cell model [5] features 35 static parameters and 21 ODEs, its dominating equation can be stated by the total ionic current \(I_{ion}\) and the stimulus \(I_{stim}\) as

\[
\frac{du}{dt} = -\frac{I_{ion} + I_{stim}}{C_m}.
\]

(6)

Several simplified models have been proposed as computationally efficient surrogates of biophysically detailed models, but they usually lack the capability of describing important physiological properties. Moreover, no simplified model is currently available for human atria-specific cellular electrophysiology.

Motivated by recent progresses in meta-modeling [7], we apply a statistical learning approach in [15] to the reduction of state-of-the-art cellular models used for atria simulation in literature. The reduced model learned keeps the ability to capture the complex dynamics of the original biophysically detailed model, while in very simple form and depending on a smaller number of parameters. This makes the model efficient and suitable for use for large scale simulations at the organ level. To the best of our knowledge, this represents the first example of the application of a model reduction technique based on statistical learning to the multiscale modeling of cardiac electrophysiology.

To be specific, in [15], we focus on the CRN atrial cell model. We first use the Principal Component Analysis (PCA) to reduce the dimension of the action potential (AP) manifold, to which the state variable \(u\) in (6) belongs. In model construction, various AP quantities (e.g. action potential duration) were accurately regressed using the projection pursuit regression (PPR) method. The embedding coordinates of the action potential in its reduced space were then regressed by PPR or multivariate adaptive regression spline (MARS) using above quantities as extra parameters. Finally, the registration of the regression cellular model into tissue-level EC modeling is after each upstroke: at time \(t^i = t_{upstroke} + i\Delta t\), the current is computed as

\[
-\frac{I_{ion} + I_{stim}}{C_m}(x, t^i) = \frac{u_{ref}(i) - u(x, t^{i-1})}{\Delta t},
\]

where \(u_{ref}\) is from the regression model, \(i \in \mathbb{N}\) corresponds to the selected time snapshot of \(u_{ref}\).

As reported in [15], AP manifold dimension can be reduced to 15 despite being the output of a nonlinear system. Our regression model demonstrates the ability of capturing the physiological complexity of cardiac AP (Fig. 4, Left). Its accuracy is guarantee even with sample parameters having standard deviation 0.3. Most significantly, the application of this regression model to tissue-level EC modeling (Fig. 4, Right) dramatically improves the computational efficiency: it decreases the computational time up to two orders of magnitude as compared to using the original non-reduced model and enables almost real-time computations (order of seconds for computing a heart cycle on a standard workstation).

The work on the modeling of atrial electrophysiology also produces a publication of US patent [14].

**Current/Future Work:** There is plenty of room for improving the proposed method. In future work, we will focus on more precise ways for monitoring AP upstroke, such as the use of the eikonal equation for the depolarization time, or a regression approach on the sodium ionic channel which controls the AP depolarization phase. We also intend to study the AP restitution properties with a more sophisticated control of the action potential duration alternans, with the aim of modeling complex pathological patterns such as atrial fibrillation. This statistical learning framework can be trivially extended to other cardiac cells for whole heart modeling.
4 Stochastic optimal control in acoustics

Simulation-based optimization of acoustic liner design in a turbofan engine nacelle for noise reduction purposes can dramatically reduce the cost and time needed for experimental designs. The acoustic model for sound propagation governed by the Helmholtz equation contain coefficients, such as the acoustic wavenumber, that are not exactly known due to incomplete knowledge or an inherent variability in the system. These uncertainties, inevitable in the design process, should be introduced into the model by treating those parameters as random variables. Optimization of the resulting stochastic system would be more complex than the deterministic one, but its accommodation to model uncertainties provides a more robust and realistic tool in practice.

In [13], we take into account uncertainties on the acoustic wavenumber due to variability in the weather, and on the fan noise source due to incomplete knowledge. We formulate the optimization on the conditional value-at-risk (CVaR) measure [11], which quantifies the conditional expectation of the sound energy provided that the sound is above a certain threshold.

To be specific, let \( X(\vartheta) \) be a general cost function with uncertainties denoted by the random vector \( \vartheta : \Omega \rightarrow \Lambda \). We denote the probability density function of \( \vartheta \) as \( \rho(\vartheta) \) and the distribution function of \( X \) as \( \Psi(\alpha) \). At a specified confidence level \( \beta \in (0,1) \), the corresponding value-at-risk (VaR\(_\beta\)) of \( X \) is defined as the \( \beta \)-quantile, that is, \( \text{VaR}_{\beta}[X] = \min\{\alpha \in \mathbb{R} : \Psi(\alpha) \geq \beta\} \). Using this concept, at probability level \( \beta \), the conditional value-at-risk CVaR\(_\beta\) is defined as the conditional expectation

\[
\text{CVaR}_{\beta}[X] = \mathbb{E}\left[X | X \geq \text{VaR}_{\beta}[X]\right] = \frac{1}{1-\beta} \int_{\{\vartheta : X(\vartheta) \geq \text{VaR}_{\beta}[X]\}} X(\vartheta) \rho(\vartheta) d\vartheta.
\]

It measures the conditional mean value of the cost above the amount VaR\(_\beta\)[X]. The second equality results from the probability \( P\left[X \geq \text{VaR}_{\beta}[X]\right] = 1 - \beta \). In [11], it is proved that \( \text{CVaR}_{\beta}[X] \) can be characterized in terms of \( \text{CVaR}_{\beta}[X] = \min_{\alpha \in \mathbb{R}} F_\beta(\alpha; X) \), where \( F_\beta(\alpha; X) = \alpha + \frac{1}{1-\beta} \int_{\Lambda} \left[X(\vartheta) - \alpha\right]^+ \rho(\vartheta) d\vartheta \) with \([x]^+ = \max\{x, 0\}\).

Based on this concept, the stochastic optimization problem we propose is to solve

\[
\min_{\xi, \alpha} \left\{ \frac{1}{2} \left[ \alpha + \frac{1}{1-\beta} \int_{\Lambda} \left[ \frac{1}{|p|} \int_{D} |p(x, \vartheta; \xi)|^2 dx - \alpha \right]^+ \rho(\vartheta) d\vartheta \right] + \frac{\gamma}{2} |\xi|^2 \right\},
\]

where \( \xi = \xi_r + i\xi_i \in \mathbb{C} \) is the impedance factor of the acoustic liner whose real part \( \xi_r \) represents resistance and the imaginary part \( \xi_i \) reactance; \( p(x, \vartheta; \xi) \) denotes the complex-valued acoustic pressure at \( x \) under uncertainty \( \vartheta \). The constant \( \gamma_p \) is chosen to scale the energy of the acoustic potential, and \( \gamma \) is a regularization coefficient.

The optimization based on CVaR measure is expected to determine optimal impedance factor that are robust to uncertainty. Solving the stochastic optimization problem would facilitate the optimal acoustic liner design with different significance levels.

A parallel reduced-order modeling framework is developed that dramatically improves the computational efficiency of the stochastic optimization solver for a realistic nacelle geometry. Whereas the computation of the CVaR measure of the full-order Helmholtz solutions is forbidding, the reduced stochastic optimization problem can be solved within 500 seconds. Numerical experiments based on minimizing the CVaR measure indicate: with 95\% certainty the acoustic noise energy can be optimally controlled within 48.66\% of the noise level associated with the hard-wall condition without acoustic liner. A typical example on the acoustic noise reduction is shown in Fig. 5. In addition, well posedness and finite element error analyses of the state system and optimization problem are provided.

**Current/Future Work:** The limitation of the work lies in the lack of an appropriate acoustic liner model that connects the design feature with the impedance factor. This will be a topic of our future work.

The airflow inside/outside the engine nacelle also has an impact on sound radiation. A coupling of acoustics with aerodynamics, that is the integration of computational fluid dynamics (CFD) results such as flow velocity, pressure and temperature with acoustic simulation, will increase the reliability of the model. Its computational challenges draw our attention for future work.

5 Fast CVT grid generation for ocean modeling

This is an ongoing research project, in collaboration with members from Florida State University, University of South Carolina and Los Alamos National Lab.
Atmospheric and oceanic modeling is a fascinating field for weather forecasting, since the ocean covers about two-thirds of the Earth. Ocean modeling for different phenomena, such as tides, eddies, and large-scale ocean currents, contains turbulent fluid dynamics in active scales ranging from the global down to order of kilometers horizontally and meters vertically. It is computationally very challenging, and in the project we consider at least two aspects. (1) In pairing with a previously developed finite-volume method [10], current centroidal Voronoi tessellation (CVT) algorithms should be largely accelerated for generating multi-resolution Voronoi diagrams and Delaunay triangulations. In particular, the best initialization technique and an effective multi-level method based on quasi-Newton optimization are still missing. Our work is ongoing. (2) Important space and time scale interactions occur over the entire spectrum as the ocean interacts with the land, the sea-ice, and the atmosphere systems. The coupling of these local systems for global climate modeling calls for effective numerical schemes to preserve stability.

References


