

An Understanding of Partial Differential Equations in Perspective of Stochastic Process, from Concept to Application

Yuan Chen*

November 4, 2020

Abstract

In this report, we discuss an understanding of parabolic Partial Differential Equation (PDE) via stochastic process. We start from a Brownian Motion representation of solution of heat equation. Then, Feynman-Kac Formula is introduced for general parabolic partial differential equations. Based on that, a branch of numerical algorithms using Monte-Carlo Method is generalized to solve PDEs.

1 Introduction

In scientific and engineering area, simulations based on Partial Differential Equations (PDE) are widely used among mechanics, physics and portfolio investment. One way to represent PDE solutions is to take expectation of stochastic process. This idea has raised great attention in recent years and has wide application in solving PDEs.

The organization of this report is: In Section 2, we introduce heat equation and its physics meaning; in Section 3, we raise the Feynman-Kac representation for parabolic PDEs; and in Section 4, we forward one of the applications of Feynman-Kac Formula: Mont-Carlo Method of numerical solving PDEs.

2 Heat Equation

In physics, heat equation is a typical parabolic partial differential equation which has been used to model distribution of quantity such as heat and pollutant in specific (solid or flow) medium [2]. Formally, on domain $\Omega \times [0, T]$, the homogeneous heat equation can be written as:

$$\frac{\partial u}{\partial t} - k\Delta u = 0, \quad \text{on } \Omega \times [0, T] \quad (2.1a)$$

$$u(X, 0) = \phi_0(X), \quad \text{on } \Omega \quad (2.1b)$$

$$u(X, t) = \phi_1(X), \quad \text{on } \partial\Omega \times [0, T] \quad (2.1c)$$

*Email: yuan.chen@gwu.edu. This report serves as a project for STAT 8273, Spring 2020. The author acknowledges Professor Hosam M. Mahmoud and Ph.D Candidate Shuyang Gao for their help through this course.

in which, $u(X, t)$ is the solution satisfying equation 2.1a, initial condition 2.1b and boundary condition 2.1c. k is a constant called conductivity constant of the medium. $\Delta u = \nabla \cdot \nabla u$, in which ∇u is gradient of function u . Explicitly, $\Delta u = \sum_{i=1}^d \frac{\partial^2 u}{\partial x_i^2}$, d is dimension. We assume that $\phi_0(X) \in C^2(\Omega)$ and $\phi_1 = 0$.

Before solving this equation, let's look at the physical meaning of heat equation. For instance, let $u(X, t)$ indicates the temperature of one medium at spatial point X and time t . Then, ∇u , which is the gradient of u , represents direction and rate of the fastest increase of u . The operator $\nabla \cdot f$ is called "divergence" of f , which describes to what extent the neighbor of f diverges from a specific point. Therefore, we can imagine the 'Laplacian' $\Delta u = \nabla \cdot \nabla u$ means 'the extent of direction and rate of u 's fastest increase diverge from a specific point'. On the one hand, if the neighbor points' fastest increasing direction diverges from a point, then we can say this target point has a relatively lower value than the neighbors. Conversely, if the direction of fastest increasing neighbor points converge to a point, then the point is said to have a relatively higher value among the neighbor. In all, the Laplace operator measures how the value of specific point is lower than its neighbors. As shown in Figure 1, the minimum point tends to have a positive (higher) Laplacian, while the maximum point tends to have a negative (lower) Laplacian.

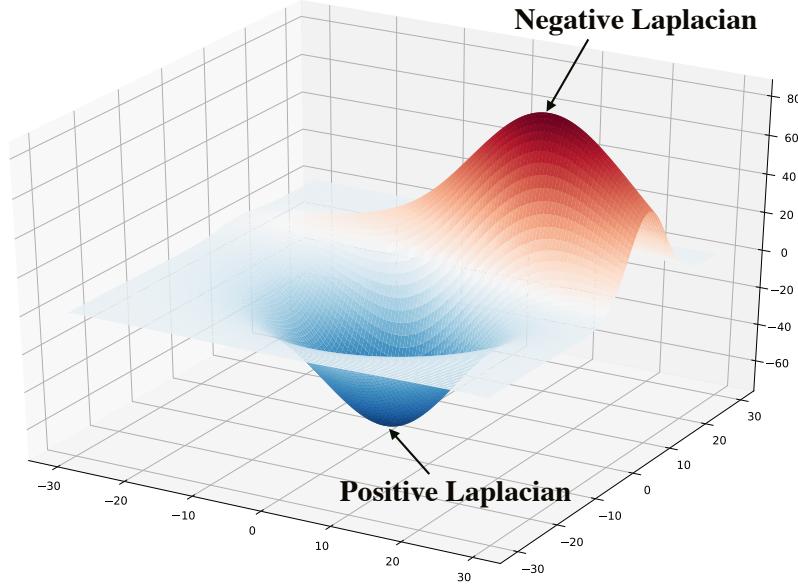


Figure 1: Interpretation of Laplacian

Back to the physics context, by the second law of thermodynamics, heat will flow from bodies with higher temperature to adjacent colder bodies, in proportion to the difference of temperature and to thermal conductivity of the medium [11]. The strength of relationship depends on specific medium, which is represented by constant k (conductivity constant of the medium). This flowing process is described in Equation 2.1a, in which the time variance rate of temperature is related to Laplacian of heat field.

When it comes to deriving exact solution of this equation, we use the technique called ‘separation of variable’. In this report, we won’t introduce much details about this technique. Besides, though there are multiple classical results concerning heat equation, we only introduce uniqueness [6] here for convenience of forwarding the theorems below. k is modified to be 1 under some linear transformation of u .

Theorem 2.1. (*Bounded Domain Uniqueness*) Let ω denote an open bounded set of \mathbb{R}^n , $\Omega = \{(x, t) | x \in \omega, 0 < t < T\}$ is a cylinder in \mathbb{R}^{n+1} . Denote again $\partial'\Omega = \{(x, t) | \text{either } x \in \partial\Omega, 0 \leq t \leq T \text{ or } x \in \omega, t = 0\}$. Let u be continuous in $\bar{\Omega}$, $\frac{\partial u}{\partial t}, \Delta u$ exists and be continuous in Ω . Then u is determined uniquely in $\bar{\Omega}$ by the value $\frac{\partial u}{\partial t} - \Delta u$ in Ω and of u on $\partial'\Omega$.

Theorem 2.2. (*Unbounded Domain Uniqueness*) Let ω be unbounded, and Ω still be defined as above theorem. u satisfies conditions of above theorem. Then u is determined uniquely in $\bar{\Omega}$ by the value $\frac{\partial u}{\partial t} - \Delta u$ in Ω and of u on $\partial'\Omega$ when we restrict ourselves to solutions satisfying:

$$|u(X, t)| \leq M e^{a|x|^2}, \text{ for } 0 < t < T$$

Notice that under the condition of Theorem 2.2, the heat equation itself degenerate to a free-boundary form. Since the domain itself is unbounded, the temperature vanishes at the ‘very end’ of domain naturally. As a result, it looks like the system only has an initial condition. For example, let $\Omega = \mathbb{R}^n$, the system can be written as:

$$\frac{\partial u}{\partial t} - k\Delta u = 0, \quad \text{on } \mathbb{R}^n \times [0, T] \quad (2.2a)$$

$$u(X, 0) = \phi_0(X), \quad \text{on } \mathbb{R}^n \quad (2.2b)$$

In fact, we omit the boundary condition $u(X, t) = 0$ at $\partial\mathbb{R}^n$. This is an important form we will use in this report. With these preparations of heat equation, let’s start our discussion of Brownian Motion.

3 Brownian Motion

Recall the definition of one-dimensional Brownian Motion. Brownian Motion B_t is a real valued process with following properties [3]:

1. If $t_0 < t_1 < \dots < t_n$, then $B(t_0), B(t_1) - B(t_0), \dots, B(t_n) - B(t_{n-1})$ are independent.
2. If $s, t \geq 0$, then

$$\mathbb{P}(B(s+t) - B(s) \in A) = \int_A (2\pi t)^{-1/2} \exp(-x^2/2t) dx$$

3. With probability one, $t \mapsto B_t$ is continuous.

3.1 An Observation

At the beginning of this section, we define an ‘Accelerating Brownian Motion’ which is a time-scaling form of standard Brownian Motion. We define B_t^k by the stochastic process with property 1 and 3 of Brownian Motion, and 2 is modified by: If $s, t \geq 0$, then:

$$\mathbb{P}(B(s+t) - B(s) \in A) = \int_A (4\pi kt)^{-1/2} \exp(-x^2/4kt) dx$$

In this definition, we rescale the Brownian Motion and ‘accelerate’ it (with proper k). Informally, an interpretation of the ‘accelerate’ is shown in Figure 2. In the Figure, when $k = 1$, the right hand side accelerating Brownian Motion acts like a half time-squeezing of the standard Brownian Motion in left hand side.

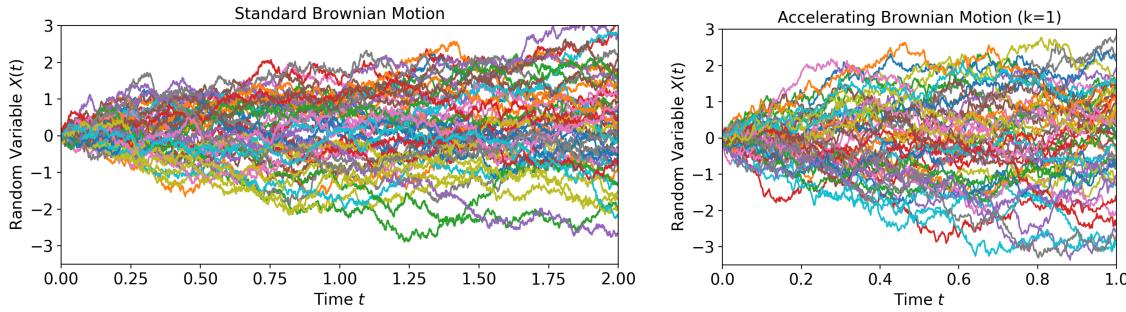


Figure 2: Left: Standard Brownian Motion. Right: Accelerating Brownian Motion with scale 2

Let $u(x, t) = \mathbb{E}(\phi_0(x - B_t^k))$, as you may guess, we can prove this u solves System 2.2.

The first step is to check: $u(x, 0) = \mathbb{E}(\phi_0(x - B_t^k))|_{t=0} = \mathbb{E}(\phi_0(x)) = \phi_0(x)$, so satisfying initial condition.

Then we take derivative on t :

$$\frac{\partial \mathbb{E}(\phi_0(x - B_t^k))}{\partial t} = \frac{\partial}{\partial t} \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi kt}} \exp(-\frac{z^2}{4kt}) \phi_0(x - z) dz \quad (3.1)$$

Since $(4\pi kt)^{-1/2} \exp(-x^2/4kt)$ is bounded uniformly by some M on $[0, T]$ (with continuity extension defined on $t = 0$). So the integrated function is bounded by $M\phi_0(x - z) \in C^2$. By Dominated Convergence Theorem, we can exchange the order of differential symbol and integral symbol.

$$\begin{aligned} \frac{\partial \mathbb{E}(\phi_0(x - B_t^k))}{\partial t} &= \int_{\mathbb{R}} \frac{\partial}{\partial t} \left\{ \frac{1}{\sqrt{4\pi kt}} \exp(-\frac{z^2}{4kt}) \phi_0(x - z) \right\} dz \\ &= \int_{\mathbb{R}} \frac{\partial}{\partial t} \left\{ \frac{1}{\sqrt{4\pi kt}} \exp(-\frac{(x-y)^2}{4kt}) \phi_0(y) \right\} dy \\ &= \frac{1}{\sqrt{4\pi k}} \int_{\mathbb{R}} \exp(-\frac{(x-y)^2}{4kt}) \phi_0(y) \left\{ \frac{(x-y)^2}{4kt^{5/2}} - \frac{1}{2t^{3/2}} \right\} dy \end{aligned} \quad (3.2)$$

With similar strategy and D.C.T. again, we have:

$$\begin{aligned} \frac{\partial^2 \mathbb{E}(\phi_0(x - B_t^k))}{\partial x^2} &= \int_{\mathbb{R}} \frac{\partial^2}{\partial x^2} \left\{ \frac{1}{\sqrt{4\pi k t}} \exp\left(-\frac{(x-y)^2}{4kt}\right) \phi_0(y) \right\} dy \\ &= \frac{1}{\sqrt{4\pi k}} \int_{\mathbb{R}} \exp\left(-\frac{(x-y)^2}{4kt}\right) \phi_0(y) \left\{ \frac{(x-y)^2}{4k^2 t^{5/2}} - \frac{1}{2kt^{3/2}} \right\} dy \end{aligned} \quad (3.3)$$

Through comparison of Computation 3.2 and 3.3, we find $u(x, t) = \mathbb{E}(\phi_0(x - B_t^k))$ solve Equation 2.2a. By uniqueness, solution of 2.1a could be represented as $u(x, t) = \mathbb{E}(\phi_0(x - B_t^k))$. This is an amazing but reasonable result. Informally, think the heat at point (x, t) is owned by a particle, after Δt time passed, the particle moves with a Brownian Motion to $(x + \Delta x, t + \Delta t)$, then the heat is transferred to that point. The solution $\mathbb{E}(\phi_0(x - B_t^k))$ finds the current temperature of point x at time t in the following way: Start Brownian Motion going backward in time from the point (x, t) and run them for time t . Then find the value by initial condition ϕ_0 at ‘backward point $x - B_t^k$ ’ and average them over all Brownian motions. This process is interpreted in Figure 3. Notice that the backward Brownian Motion gets faster when conductivity constant k is larger, which varies with medium Ω . For example, Copper’s constant is 384.1 and water’s is 0.6089. This is consistent with our common sense that heat transfers faster in copper than in water.

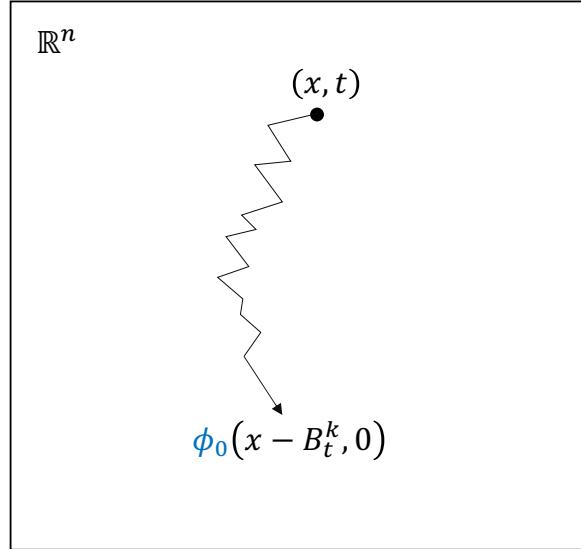


Figure 3: Interpretation of Heat Equation

This representation understands the macroscopic energy transition in terms of the average behavior of microscopic particle motion. In 1905, Albert Einstein claimed probability density of the position x of the Brownian particle at time t should satisfy heat equation (with suitable initial: Dirac delta function, $\phi_0(x) = \delta(x)$). This is why heat equation is also called ‘diffusion equation’ which has also been used to simulate diffusion of pollutants, gas molecules and smoke particles.

3.2 Feynman-Kac Formula

In this section, let's see a more general result [7]. This theorem is named after the physicist Richard Feynman (1918-1988) and mathematician Mark Kac (1914-1984). It claims that solution of parabolic PDE can be represented by expectation of a stochastic process. This stochastic process is also a function of the solution of a Stochastic Differential Equation (SDE).

Theorem 3.1. (1D Feynman-Kac Formula, free boundary) Consider Parabolic PDE:

$$\frac{\partial u}{\partial t} - \mu(x, t)u_x - \frac{1}{2}\sigma^2(x, t)u_{xx} + V(x, t)u = 0, \quad \text{on } \mathbb{R} \times [0, T] \quad (3.4a)$$

$$u(x, 0) = \phi(x), \quad \text{on } \mathbb{R} \quad (3.4b)$$

Under suitable condition of μ, σ, V , the solution can be written as the expectation:

$$u(x, t) = \mathbb{E} \left(\exp \left[- \int_0^t V(X_\tau, t - \tau) d\tau \right] \phi(X_t) \right) \quad (3.5)$$

in which X_t is Ito's Process defined by a Stochastic Differential Equation:

$$\begin{aligned} X_t &= X_0 + \int_0^t \mu(X_\tau, t - \tau) d\tau + \int_0^t \sigma(X_\tau, t - \tau) dB_\tau \\ X_0 &= x \end{aligned}$$

In the above theorem, we do integration on a stochastic process with respect to a stochastic process. The rigor definition is called Itô calculus named after mathematician Kiyoshi Itô. Kiyoshi Itô extends the idea of calculus to stochastic process. Here we introduce the basic idea. Follow the spirit of Riemann-Stieltjes integral, the integration of a stochastic process $H_t \in \mathcal{F}_t$ with respect to a Brownian Motion B_t (\mathcal{F}_t is the filtration generated by Brownian motion up to time) on a interval $[0, t]$ is defined as a limit of stochastic process. Let π_n be a partition of $[0, t]$, which is $\pi_n = \{0 = t_0 < t_1 < t_2, \dots < t_n = t\}$, $|\pi_n| = \max_{i=1,2,\dots,n} |t_i - t_{i-1}|$,

$$\int_0^t H_t dB_t = \lim_{|\pi_n| \rightarrow 0} \sum_{i=1}^n H_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}) \quad (3.6)$$

After introducing basic Itô calculus, we take $\mu(x, t) = 0$, $\sigma^2(x, t) = 2k$, $V(x, t) = 0$ in the Theorem 3.1, the problem degenerates to heat equation with coefficient k on free boundary (System 2.2). The form of solution is $u(x, t) = \mathbb{E}(\phi(X_t))$, X_t is defined by $X_t = x - \sqrt{2k}B_t$.

In real-world applications, we often restrict us in a bounded region with specific boundary condition. The Feynman-Kac Formula could apply in this situation as well.

Theorem 3.2. (1D Feynman-Kac Formula, bounded domain) Consider Parabolic PDE:

$$\frac{\partial u}{\partial t} - \mu(x, t)u_x - \frac{1}{2}\sigma^2(x, t)u_{xx} + V(x, t)u = 0, \quad \text{on } \Omega \times [0, T] \quad (3.7a)$$

$$u(x, 0) = \phi_0(x), \quad \text{on } \Omega \quad (3.7b)$$

$$u(x, t) = \phi_1(x, t), \quad \text{on } \partial\Omega \times [0, T] \quad (3.7c)$$

Under suitable condition of μ , σ , V , the solution can be written as the expectation:

$$u(x, t) = \mathbb{E} \left(\exp \left[- \int_0^{\gamma \wedge t} V(X_\tau, t - \tau) d\tau \right] \phi(X_{\gamma \wedge t}, t - \gamma \wedge t) \right) \quad (3.8)$$

in which γ is a stopping time defined as $\inf\{t \geq 0 | X_t \in \mathbb{R} \setminus \Omega\}$, interpreted as the 'hitting time' of X_t to hit the boundary. Symbol $\gamma \wedge t := \min\{\gamma, t\}$. X_t is Ito's Process defined as Theorem 3.1. $\phi(x, t)$ is defined as:

$$\phi(x, t) = \phi_0(x) \mathbb{I}_{\{(x, t) | t=0, x \in \bar{\Omega}\}} + \phi_1(x, t) \mathbb{I}_{\{(x, t) | t>0, x \in \partial\Omega\}}$$

The definition of stopping time here is to say: in a bounded area, if the stochastic path X_t hits the boundary, then its path after that time will not be counted in $u(x, t)$, if X_t always stays in interior of Ω , then the constitution of $u(x, t)$ is the same as the solution of free boundary problem in Theorem 3.1.

Informally, we use Figure 4 to illustrate Feynman-Kac Formula in free boundary and bounded domain condition under a 'heat context'. The left figure is the free boundary condition, the right one is bounded domain condition. In the left hand figure, we still think a particle holds the heat. In order to find value in (x, t) , we do a time reversal with a stochastic process X_t , the value is determined by initial temperature $\phi_0(X_t, 0)$ and affection by $V(X, s)$ along the whole path. The solution is average among all the stochastic paths.

In the bounded area condition (right hand figure), we do time reversal again. However, apart from initial point, the process could be tracked back into a boundary point where time γ is less than t , this is why we define a stopping time γ to represent this time. Apparently, this time is a random variable (actually stopping time). In this condition, the value of (x, t) is determined by boundary condition $\phi_1(X_\gamma, t - \gamma)$. The solution is average among all these two kinds of stochastic paths. X_t in the above two conditions is defined as a stochastic process driven by a Stochastic Differential Equation with information from both a Brownian Motion and PDE itself.

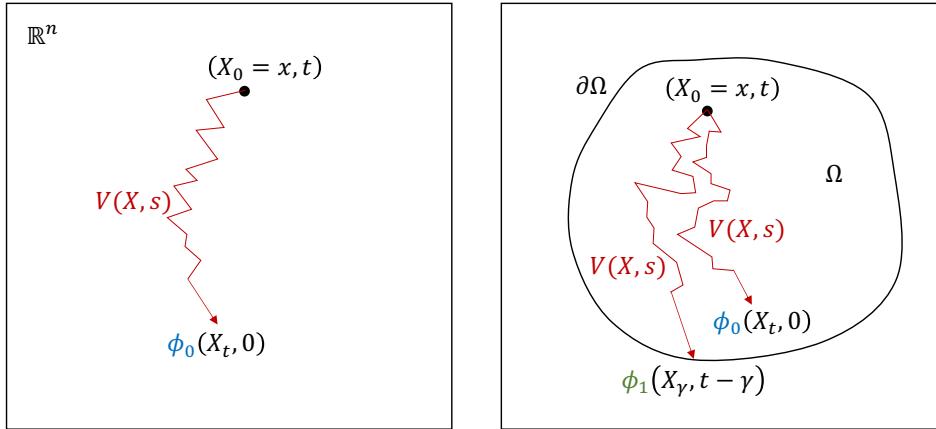


Figure 4: Interpretation of Feynman-Kac Formula

Feynman-Kac Formula establishes a connection between parabolic partial differential equations (PDEs) and stochastic processes. 'It offers a method of solving certain partial differential

equations by simulating random paths of a stochastic process (we will introduce in Section 4). Conversely, an important class of expectations of random processes can be computed by deterministic methods'. This accurate comment is cited from [10].

4 Applications

In this section, we introduce Monte-Carlo method for partial differential equations as an application of Feynman-Kac Formula. Notice that the solution in Theorem 3.1 is the expectation of a function of one stochastic process X_s evaluating at $0 \leq s \leq t$.

$$u(x, t) = \mathbb{E}(g(X_s | 0 \leq s \leq t)) \quad (4.1)$$

in which, $g(X)$ is a known function. By Large Law of Large Number Theory, we have:

$$\mathbb{E}(g(X_t)) \simeq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N g(X_s^n | 0 \leq s \leq t) \quad (4.2)$$

The SLLN states we just need to repeat doing simulation of independent stochastic path X_s . The arithmetic average is the solution of PDE in above section. The main obstacle here is how to simulate sample of X_s , $0 \leq s \leq t$. During preliminary research, we first discretize continuous time interval $[0, t]$ into $\{0 = t_0 < t_1 < t_2, \dots < t_n = t\}$. So the continuous stochastic process X_t is discretized into X_{t_i} , $i = 0, 1, 2, \dots, n$. Then, we generate random samples of standard Brownian Motion B_t at points $t = t_i$, $i = 0, 1, 2, \dots, n$ using Monte-Carlo Simulation. The key step is: X_{t_i} could be generated by a recursive process driven by the Stochastic Differential Equation defined in Theorem 3.1, for $i = 0, 1, 2, \dots, n - 1$:

$$X_{t_{i+1}} \simeq X_{t_i} + \mu(X_{t_i}, t - t_i)(t_{i+1} - t_i) + \sigma(X_{t_i}, t - t_i)(B_{t_{i+1}} - B_{t_i}) \quad (4.3)$$

The algorithm can be written as in the form Algorithm 1.

Algorithm 1: Monte-Carlo Method Simulation for System 3.4

```

 $N \leftarrow 0, R \leftarrow 0, L \leftarrow l;$ 
while  $R$  not converge do
    Sample Standard  $B_t^l$  at level  $l$ ;
    Sample  $X_t^l$  at level  $l$  by 4.3;
     $A \leftarrow g(X_t^l);$ 
     $R \leftarrow \frac{1}{N}[(N - 1)R + A];$ 
     $N \leftarrow N + 1;$ 
end

```

Following the algorithm, a numerical example conducted by [4] solves $\nabla \cdot [\beta \nabla u] = -16$, $\beta = (1, 3)$ with 1000 sample points and $\Delta t = 10^{-4}$. Then they compare the simulation results with results solved by Finite Element Method which is a traditional numerical method for PDEs. The results is shown in the table 1.

Point	Finite element	Monte Carlo	Difference (%)
(0.25, 0.25)	1.788	1.802	0.780
(0.50, 0.25)	6.350	6.249	-1.59
(0.75, 0.25)	23.87	23.98	0.440
(0.25, 0.50)	2.480	2.406	-2.97
(0.50, 0.50)	8.940	8.477	-5.18
(0.75, 0.50)	32.43	33.29	2.660

Table 1: Finite element and proposed Monte Carlo solutions

The table shows that the bias of Mont Carlo Method is acceptable. This new method which utilizes probabilistic method to solve PDEs has its advantage in efficiency, which are also pointed out in [8]: (1) we can get the solution for points we are interested in without working on the whole field. In contrast, traditional numerical methods such as Finite Element aforementioned must solve the whole field in one run. (2) It can overcome the ‘curse of dimension’. Here, the ‘curse’ means the number of mesh increases exponentially with dimension d , because traditional numerical methods rely on meshes generated to solve PDE. For large d , it is impossible to simulate large systems. However, the computational cost of Mont-Carlo Method only increases linearly with the dimension of X_t . While this branch of so-called non-mesh Mont-Carlo Method is a relative young method, it has received attention in wide application areas. For instance, [1] simulates 3D heat transition in a multi-layered composite material component using Mont-Carlo Method shown in Figure 5.

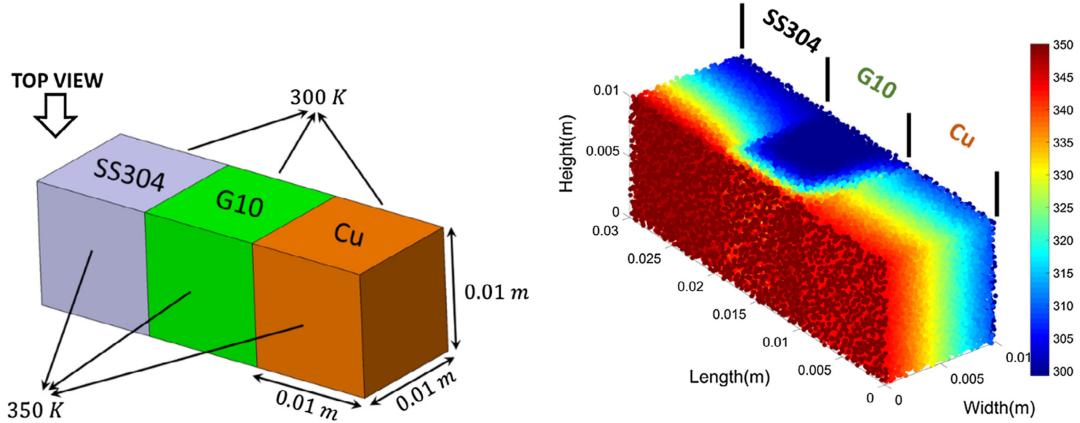


Figure 5: Left: A composite material component, Right: Heat map [1]

In recent years, using probabilistic methods to solve PDEs is becoming a research hotspot coupled with rising personal computers and development of artificial intelligence. Recently, Deep BSDE Solver developed by [9, 5] formulate parabolic PDE into BSDE (Backward Stochastic Equation) through non-linear Feynman-Kac formula. The authors represented the solution by a specific function (involving gradient of u) of Brown motion which can be simulated by Mont Carlo Method. The key step of the method is to approximate the specific function by a Deep

Neural Network and obtain the parameters using loss function defined by boundary condition.

References

- [1] Reza Bahadori, Hector Gutierrez, Shashikant Manikonda, and Rainer Meinke. A mesh-free monte-carlo method for simulation of three-dimensional transient heat conduction in a composite layered material with temperature dependent thermal properties. *International Journal of Heat and Mass Transfer*, 119:533–541, 2018.
- [2] John Rozier Cannon. *The one-dimensional heat equation*. Number 23. Cambridge University Press, 1984.
- [3] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.
- [4] M Grigoriu. A monte carlo solution of heat conduction and poisson equations. *J. Heat Transfer*, 122(1):40–45, 2000.
- [5] Jiequn Han, Arnulf Jentzen, and E Weinan. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.
- [6] Fritz John. *Partial Differential Equations*. Applied Mathematical Sciences 1. Springer US, 3 edition, 1978.
- [7] Huyêñ Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.
- [8] Fq Takens. Lecture notes in mathematics. *by DA Rand and L.-S. Young Springer, Berlin*, 898:366, 1981.
- [9] E Weinan, Jiequn Han, and Arnulf Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4):349–380, 2017.
- [10] Wikipedia contributors. Feynman–kac formula, 2020. [Online; accessed 28-April-2020].
- [11] Wikipedia contributors. Heat equation, 2020. [Online; accessed 11-April-2020].