Information flow in political elections: a stochastic perspective

Santosh Kumar Radha

Department of Physics, Case Western Reserve University,
10900 Euclid Avenue, Cleveland, OH-44106-7079

Often times, a candidate's attractiveness is directly associated with his clear ideologies and opinions on various policies and social issues. Using the ideas of stochastic differential equations and Ornstein-Uhlenbeck Process, we develop a phenomenological model to understand the effect of (un)clearly communicating a candidate's stance on policies to the voting public. We will show that, counter intuitively, there are quantifiable advantages to be vague on one's stance.

Quantitative analysis of political elections have been a staple for many years. Models of election are many and varied, each with it's own focus. There is a vast literature on methods to forecast the elections using tools like fundamental indicators[1], market indicators[2], Bayesian methods[3] and even social media strategies [4] and other general statistical tools [5–7]. Although there is extensive work on modeling of political dynamics between candidates [8–12], most of the study focus on global/macro perspective while either neglecting or including the complex effects of individual components (*e.g.* stances on different policies) as a *mean field* effect.

In this letter, we develop a quantitative phenomenological model to understand and analyze the effect of one such individual component - efficient information flow in an election/voting system. Information flow is a fundamental notion in many areas from physics to quantitative finance. Whenever a decision is being made on a candidate, one typically encounters uncertainties about the candidate's stance on various policies. These uncertainties can occur due to various reasons from candidate's lack of effective communication to spread of deliberate misinformation. While qualitatively, inefficient transfer of information from candidate to public might seem like a negative effect, we show that there are certain cases where these inefficiencies are not only advantages, but also necessary. We use Ornstein-Uhlenbeck equations[13] to model the time dependent information flow between the candidate and public. Use of stochastic theory to model complex elections have also been used by Fenner *et al.* [14] to analyze the polls leading up to the UK 2016 EU referendum. We will start by introducing the model, after which we will explore the properties and effects of various parameters phenomenologically.

Model : Let us denote the stance of the candidate on policy p as μ with his own uncertainty on the stance being σ, where $-\infty < \mu < \infty$ with positive values being in favor for policy and negative values being against. This could model a variety of situations from social issues to ideologies. For example, in case of left vs right, a left leaning moderate candidate might have $\mu \approx 0.5, \sigma \approx 0.5$ with $\mu > 0$ and $\mu < 0$ being left and right respectively. Now we model the public *perceived* stance of the candidate's policy at time t during the election to by a random variable X_t. At the start of election ($t = 0$) (or the start of announcing the candidacy), $X_t|_{t=0} = X_0 = \mathcal{N}(\mu_0, \sigma_0)$ *i.e.* at $t = 0$, the public has a predisposed idea on what the candidate's stance is with a distribution given by X_0. For instance, $X_0 = \mu_0 = 0$ would mean that the public has no idea of what the candidate's stance is. As time flows, the candidate's opinion on p is made clearer though various information transfer processes like public speech, social media posts *etc.* and $E[X_t|_{t\to\infty}] \to \mu$. This system is modeled as a solution to a stochastic differential equation (SDE) given by

$$\mathrm{d}X_t = r(\mu - X_t)\,\mathrm{d}t + \sigma\mathrm{d}W_t, \quad t > 0. \tag{1}$$

where r is the rate at which the candidate makes his stance μ clear with a variance of $\sigma \geq 0$. W_t is a standard Brownian Motion on R. Equation 1 can alternatively be written in terms of stochastic integral form as

$$X_t = \mu\left(1 - e^{-rt}\right) + \sigma e^{-rt}\int_0^t e^{rs}\mathrm{d}W_s + X_0 e^{-rt}, \quad t \geq 0 \tag{2}$$

It is trivial to check that (2) is the unique, strong Markov sollution to (1)[15]. This is Ornstein-Uhlenbeck Process, which is an extension of Brownian motion with friction. The important feature of (1), is that the expectation value at time t is given by

$$E[X_t] = \mu\left(1 - e^{-rt}\right) + e^{-rt}E[X_0], \quad t \geq 0, \tag{3}$$

which gives us the required asymptotic Gaussian behavior $E[X_t|_{t\to\infty}] \to \mu$. These type of processes are often refereed to as *mean-reverting* processes. Although there are different generalizations of Ornstein-Uhlenbeck Process [16], we here choose the vanilla model to introduce the phenomenology.

In general, the formal solution of (1) is given by $X_t = \mathcal{N}(E[X_t], \mathrm{Var}[X_t])$ where

$$\mathrm{Var}[X_t] = \frac{\sigma^2}{2r}\left(1 - e^{-2rt}\right) + e^{-rt}\mathrm{Var}[X_0] \tag{4}$$

1 shows a numerical simulation of X_t for values of $\mu = 2$ (positive stance) and $\mu = -2$ (negative stance) for 10 different stochastic curves starting with an initial distribution of $\mathcal{N}(0, 2)$ in green and red respectively.

Black lines show the evolution of mean according to (3). As one can see, initial idea of candidate's perspective on policy is quite spread out which then converges to μ as time proceeds.
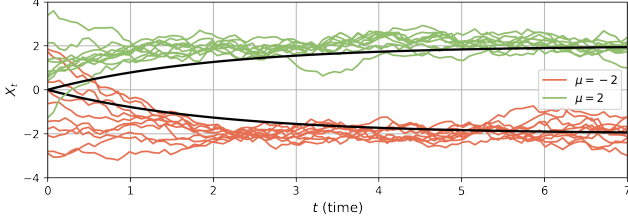

FIG. 1. Sample trajectories of X_t starting from $X_0 \approx \mathcal{N}(0, 2)$ for $\sigma = 0.5$, $r = 1$, (green) $\mu = 2$ and (red) $\mu = -2$

In this model, as seen from (4), $\text{Var}[X_t]$ plays the vital role of how collectively *confused* the public is about the candidate's opinion at any given time t. This is mainly governed by the variable $\frac{\sigma}{2r}$. It is interesting to note that this confusion is not dependent on how strong a negative or positive stance the candidate takes *i.e.* μ.

With these notion in mind, we now propose that the probability that the candidate wins (P_T)

Proposition .1 *Probability (P_T) that the candidate wins maximum votes when the election is at time T is given by Hellinger distance measure $1 - H(X_T, \tilde{X}_p)$ where $\tilde{X}$ is the public collective public opinion on policy p.*

Hellinger distance H is given by[17]

$$H^2(f, g) = \frac{1}{2} \int (\sqrt{f(t)} - \sqrt{g(t)})^2 dt \qquad (5)$$

where f, g are two continuous probability density functions. Because Hellinger distance is a bounded metric on the space of probability distributions, we can directly relate it to the probability of winning. Prop..1 is nothing but a measure of how close the candidate's perceived opinion aligns with actual public onion at any given time, matching of which would dictate the winning.

Based on (3) and 4, we have X_T to be a Gaussian density function and if the public stance on p can be modeled as Gaussian, (5) reduces to

$$P_T = \sqrt{\frac{2\sigma_T \tilde{\sigma}}{\sigma_T^2 + \tilde{\sigma}^2}} \exp\left\{-\frac{1}{4} \frac{(\mu_T - \tilde{\mu})^2}{\sigma_T^2 + \tilde{\sigma}^2}\right\}, \qquad (6)$$

where σ_T, μ_T are the variance and mean of the random variable X_T and public stance $= \mathcal{N}(\tilde{\mu}, \tilde{\sigma})$. Using (6), we can now study the effect of various parameters in the model and its result on the probability of winning.

Based on (6), as a sanity check, we will first explore the effect of candidates controversial take on policy p, *i.e.*

$|\mu - \tilde{\mu}| >> 0$. Since σ_T is unaffected by μ, we will assign $\alpha_T = 2\sigma_T \tilde{\sigma}$ and $\beta_T = \sigma_T^2 + \tilde{\sigma}^2$, then

$$P_T(\mu)$$
$$= \sqrt{\frac{2\alpha_T}{\beta_T}} \exp\left\{\frac{-1}{4\beta_T} \left(\mu \left(1 - e^{-rT}\right) + e^{-rT} E\left[X_0\right] - \tilde{\mu}\right)^2\right\}, \qquad (7)$$

Which as $T \to \infty$ is a Gaussian distribution with mean $\mu - \tilde{\mu}$, which essentially says that when every other variable $= 1$, the maximum likelihood of winning the election is when the public and candidate have the same stance on p.
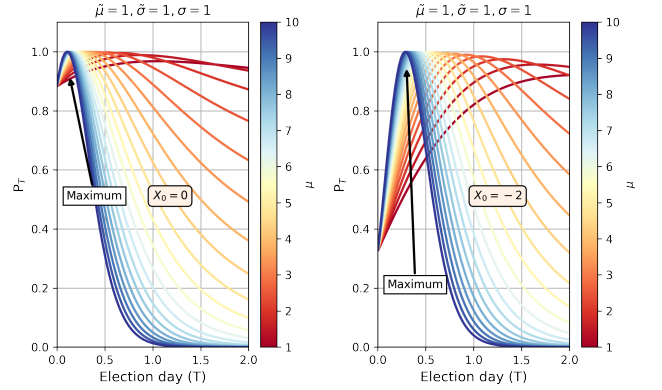

FIG. 2. P_T for $0 < T < 2$ when $\tilde{\mu} = 1, \tilde{\sigma} = 1, \sigma = 1$ and $1 < \mu < 10$ for starting distribution with (a) $X_0 = 0$ (b) $X_0 = -2$

Things get more interesting if we study the likelihood of winning as a function of number of days to election. Figure 2 shows the evolution of P_T as a function of number of days from election for all quantities set to 1, except μ, which varies from 1 to 10. As one can clearly see from (a), for $\tilde{\mu} = 1$ (the majority accepted stance on policy $= 1$) and the candidate's stance is far away from it $\mu = 10$, at $T \to \infty$, the candidate has ≈ 0 chance of winning, as mentioned in previous paragraph. But, when one closely observes the figure, we see that there is an inflation point where the candidate has the maximum likelihood to win. This confirms a very intuitive strategy that, when you have an unpopular opinionated candidate/policy, he is more likely to win when announcing candidacy/stance very close to election as the public cannot yet fully digest his actual stance despite being clear due to lack of time to assimilate. In (b), we start from an already negative perception on the candidates stance *i.e.* $X_0 = -2$, and this inflation point always survives.

Now, we focus on the effect of introducing *confusion* *i.e.* σ, for a given stance. We will focus on $T \to \infty$ limit for clarity as extending to finite T is straightforward. Figure 3 (a) shows the calculated P_T for various μ as a function of σ, with $\tilde{\mu} = 5$ and $\tilde{\sigma} = 1$, since we are taking $T \to \infty$, one should note that the initial starting

point (X_0) does not matter as we are interested in the asymptotic behavior.
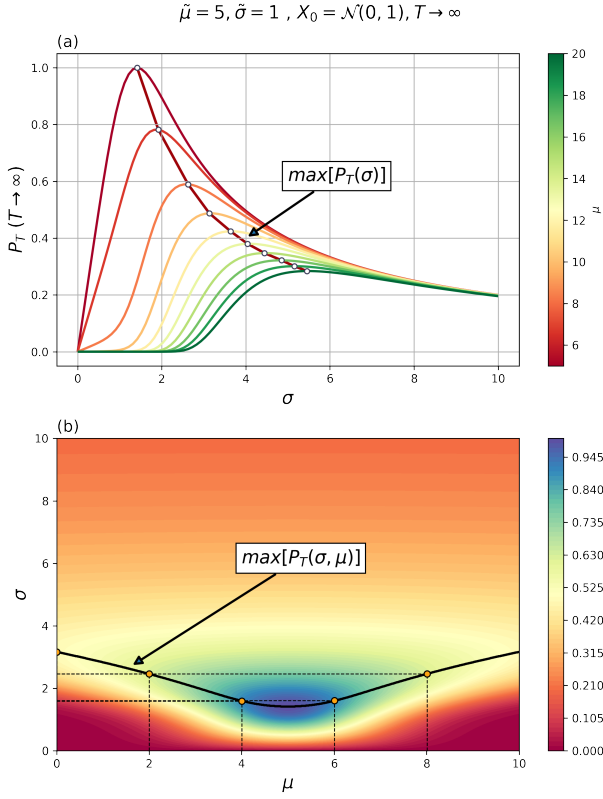

FIG. 3. P_T for $0 < T < 2$ when $\tilde{\mu} = 1, \tilde{\sigma} = 1, \sigma = 1$ and $1 < \mu < 10$ for starting distribution with (a) $X_0 = 0$ (b) $X_0 = -2$

We start with $\mu = 5$ (red curve), where we have the public stance given by $\tilde{\mathcal{N}}(5, 1)$, which has the maximum probability reaching when $\sigma = 1$, giving us that maximum likelihood is reached when $X_T|_{T\to\infty} = \tilde{\mathcal{N}}(5,1)$. As the candidate's stance differs from popular stance *i.e.* $\mu \neq \tilde{\mu}$ (red to green curve), his probability to win reduces. Despite this reduction, one can still tune σ to reach the optimal probability. This essentially dictates that when an unpopular stance is held by a candidate, there is an optimal noise that can be added while communicating to public which maximizes the chances of winning. From (a), one can also see that as $\mu - \tilde{\mu} \to \infty$, the effect of optimizing σ does not create appreciable advantage. Thus, when a polarizing candidate is introduced, the best bet is to be extremely confusing about the stance and very little is to be gained by optimizing σ.

Polar opinions : One important use case of the above formulation is when there is a binary stance by the public on a policy. For instance, there can be scenarios where the public opinion on a policy p is given by $\tilde{x} \in \{-1, 1\}$ while the candidate can have his stance $x \in (-\infty, \infty)$. We now have the random variable $X \in (-\infty, \infty)$, which the public perceive as the candidate's stance, based on

which at election time (T), public makes the choice $\tilde{X}_T$ based on candidate's perceived stance X_T, given by

$$\tilde{X}_T = \begin{cases} +1 & \text{if } X_T > 0, \\ -1 & \text{if } X_T < 0. \end{cases} \tag{8}$$

Equation 12 says that, at the election date, the public makes the decision based on the candidate's positive/negative stance on the policy, irrespective how how strong or weak the stance is. Following the phonological model, the important quantity to calculate is the probability kernel of (1), which is given by

$$\mathcal{P}(t, x, y)$$
$$= \frac{1}{\sqrt{\pi \sigma^2 \left(1 - e^{-2rt}\right)/r}} \exp\left\{ -\frac{\left(y - \mu - (x - \mu)e^{-rt}\right)^2}{\sigma^2 \left(1 - e^{-2rt}\right)/r} \right\}, \tag{9}$$

where $\mathcal{P}$ is the probability kernel for X to reach y from x in time t with μ, σ being the mean and spread of candidate's stance. Using (9), one can calculate the ideal time to start the election[18]. We will illustrate this with an example where we know that the majority favors $+1$ *i.e.* $\tilde{\mu} = +1$ and $\tilde{\sigma} = 0$ (this information can be extracted from public polls *etc.*). In this case, we first need to calculate the probability ($P^\dagger(t, x)$) for a point x to reach $y > 0$ in a time t,

$$P^\dagger(t, x) = \int_0^\infty \mathcal{P}(t, x, y) dy$$
$$= 1 - \frac{1}{2} \operatorname{erfc}\left(-\frac{\sqrt{r}\left(-\mu - (-\mu + x)e^{-rt}\right)}{\sigma\sqrt{1 - e^{-2rt}}} \right) \tag{10}$$

Figure 4(a) shows $P^\dagger(t, x)$ for $\mu = \sigma = r = 1$. This is the probability that a person who thinks that the candidate's stance is x would choose $\tilde{X}_t = 1$, qt time t. Thus under the assumption that majority favors $+1$, this would guarantee that the person is voting for the candidate. Using $P^\dagger(t, x)$, we can now calculate exactly how long to wait for the election for various starting points. For example, if we know that at this given moment $t = 0$, the public perceives the stance of candidate to be $f(x)$ where $f(x) := \mathcal{N}(\mu_0, \sigma_0)$, then the probability that they reach $+1$ at time t is given by

$$P'(t) = \int_{-\infty}^{+\infty} f(x) P^\dagger(t, x) dx. \tag{11}$$

This is easy to Figure 4(a), as we integrate out x, sampling from the points where we start *i.e.* $\mathcal{N}(\mu_0, \sigma_0)$ (shown in black line (a)). We show $P'(t)$ in (b) for various starting point μ_0. It is clear that when the public perceives the candidate to be on the negative side, there is a huge advantage to wait (assuming the public stance
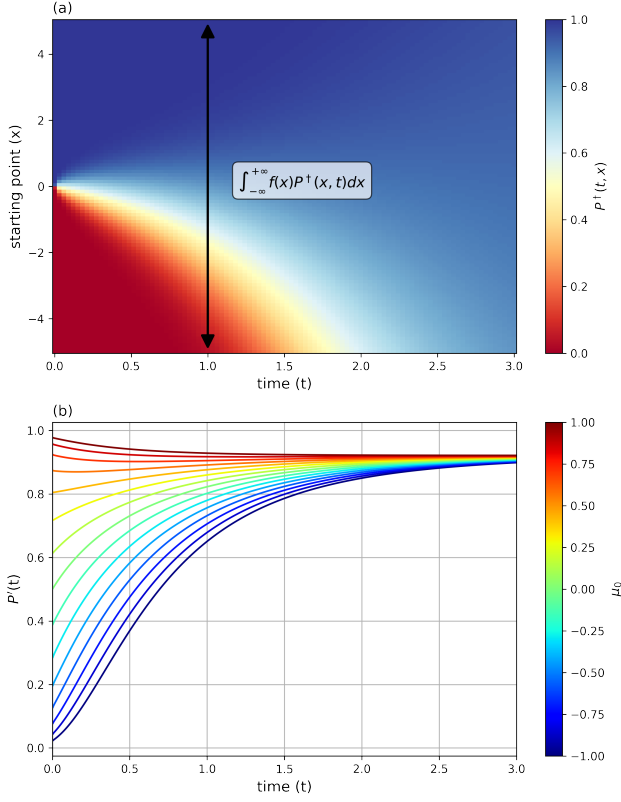
FIG. 4. P_T for $0 < T < 2$ when $\tilde{\mu} = 1, \tilde{\sigma} = 1, \sigma = 1$ and $1 < \mu < 10$ for starting distribution with (a) $X_0 = 0$ (b) $X_0 = -2$

on policy is positive $\tilde{\mu} = +1$) while the inverse is true if they perceive him to be on the positive side. Although we discussed the case for majority favors $+1$, this can easily be extended to -1.

Finally, (1) can be used as a base to model and study other intricate effects. For instance one can study effect of sudden impact full negative/positive news by simply adding terms to (1) to include jump process which can be analyzed numerically. Recent work by Brody *et al.* [19] showed that *Fake news* can be modeled by adding a noise with non-zero drift.

Even though all previous discussions pertained to single policy, one can easily generalize this model to accommodate multiple policy by increasing the dimensionality of the random variable. This is important because public decide between candidates based on number of different policies that are both quantifiable and non quantifiable. Subconsciously, these stances are mapped into an overall score, and the public at the end votes for the candidate with the highest score. Similar to (1), one can define the $n-$dimensional OU process with $\mathbf{X_t} = (X_t^1, X_t^2, \ldots, X_t^n)$ which are random variables for policies $\mathbf{p} = (p^1, p^2, \ldots, p^n)$ satisfying

$$\mathrm{d}\mathbf{X}_t^i = \sum_k R^{ik} \left(M^k - \mathbf{X}_t^k \right) \mathrm{d}t + \sum_l \Sigma^{il} \mathrm{d}\mathbf{W}_t^l, \quad t > 0,$$

(12)

where R and Σ are $n \times n$ matrices and M is a $n \times 1$ matrix while $\mathbf{W_t}$ is a vector of n independent Brownian motions and repeated. Formal solution of (12) is given by time dependent Gaussian vector[20]

Summery : In this work, we have developed a quantitative phenomenological model to understand the effect of uncertainty in information flow in elections. This offers a initial model upon which complexities can be adorned to quantitatively analyze the effects of isolated variables, which helps in quantitative strategizing of election campaigns. Further, we showed the existence of situations where ineffective information flow can be advantageous. Finally, we noted that this formalism, although developed for single policy study, can be extended to multiple policies and is still analytically tractable.

[1] P. Hummel and D. Rothschild, Fundamental models for forecasting elections, ResearchDMR.com/HummelRothschild_FundamentalModel (2013).

[2] J. Berg, R. Forsythe, F. Nelson, and T. Rietz, Results from a dozen years of election futures markets research, Handbook of experimental economics results **1**, 742 (2008).

[3] D. A. Linzer, Dynamic bayesian forecasting of presidential elections in the states, Journal of the American Statistical Association **108**, 124 (2013).

[4] D. Gayo-Avello, No, you cannot predict elections with twitter, IEEE Internet Computing **16**, 91 (2012).

[5] P. Hummel and D. Rothschild, Fundamental models for forecasting elections at the state level, Electoral Studies **35**, 123 (2014).

[6] C. Klarner, Forecasting the 2008 us house, senate and presidential elections at the district and state level, PS: Political Science and Politics **41**, 723 (2008).

[7] B. E. Lauderdale and D. Linzer, Under-performing, over-performing, or just performing? the limitations of fundamentals-based presidential election forecasting, International Journal of Forecasting **31**, 965 (2015).

[8] L. Böttcher, H. J. Herrmann, and H. Gersbach, Clout, activists and budget: The road to presidency, PloS one **13**, e0193199 (2018).

[9] D. Braha and M. A. De Aguiar, Voting contagion: Modeling and analysis of a century of us presidential elections, PloS one **12**, e0177970 (2017).

[10] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz, Is the voter model a model for voters?, Physical review letters **112**, 158701 (2014).

[11] S. Galam, The dynamics of minority opinions in democratic debate, Physica A: Statistical Mechanics and its Applications **336**, 56 (2004).

[12] S. K. Radha, Stochastic differential theory of cricket (2019), arXiv:1908.07372 [physics.soc-ph].

[13] G. E. Uhlenbeck and L. S. Ornstein, On the theory of the brownian motion, Phys. Rev. **36**, 823 (1930).

[14] T. Fenner, M. Levene, and G. Loizou, A stochastic differential equation approach to the analysis of the uk 2016 eu referendum polls, Journal of Physics Communications **2**, 055022 (2018).

[15] P. E. Protter, Stochastic differential equations, in *Stochastic integration and differential equations* (Springer, 2005) pp. 249–361.

[16] A. K. Dixit, R. K. Dixit, and R. S. Pindyck, *Investment under uncertainty* (Princeton university press, 1994).

[17] E. Hellinger, Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen., Journal für die reine und angewandte Mathematik (Crelles Journal) **1909**, 210 (1909).

[18] Time to start the election is the same as time till the election after announcing candidacy.

[19] D. C. Brody and D. M. Meier, How to model fake news, arXiv preprint arXiv:1809.00964 (2018).

[20] P. Vatiwutipong and N. Phewchean, Alternative way to derive the distribution of the multivariate ornstein–uhlenbeck process, Advances in Difference Equations **2019**, 1 (2019).