A Proofs

Theorem 1. *For a positive semidefinite matrix L and $\boldsymbol{x} \in [0,1]^N$,*

$$\sum_Y \prod_{i \in Y} x_i \prod_{i \notin Y} (1 - x_i) \det(L_Y) = \det(\operatorname{diag}(\boldsymbol{x})(L - I) + I) . \tag{1}$$

Proof. Assume momentarily that $x_i < 1,\ \forall i$.

$$\begin{aligned}
\sum_Y \prod_{i \in Y} x_i \prod_{i \notin Y} (1 - x_i) \det(L_Y) &= \prod_i (1 - x_i) \sum_Y \prod_{i \in Y} \frac{x_i}{1 - x_i} \det(L_Y) && (2)\\
&= \prod_i (1 - x_i) \sum_Y \det((\operatorname{diag}(\boldsymbol{x}) \operatorname{diag}^{-1}(1 - \boldsymbol{x}) L)_Y) && (3)\\
&= \prod_i (1 - x_i) \det(\operatorname{diag}(\boldsymbol{x}) \operatorname{diag}^{-1}(1 - \boldsymbol{x}) L + I) && (4)\\
&= \det(\operatorname{diag}(\boldsymbol{x}) L + \operatorname{diag}(1 - \boldsymbol{x})) && (5)\\
&= \det(\operatorname{diag}(\boldsymbol{x})(L - I) + I) . && (6)
\end{aligned}$$

The second and fourth equalities follow from the multilinearity of the determinant, and the third follows from DPP normalization. Since Equation (1) is a polynomial in $\boldsymbol{x}$, by continuity, the formula holds when some $x_i = 1$. □

Corollary 2. *For $f(Y) = \log \det(L_Y)$, we have $\tilde{F}(\boldsymbol{x}) = \log \det(\operatorname{diag}(\boldsymbol{x})(L - I) + I)$ and*

$$\frac{\partial}{\partial x_i} \tilde{F}(\boldsymbol{x}) = \operatorname{tr}((\operatorname{diag}(\boldsymbol{x})(L - I) + I)^{-1} (L - I)_i) , \tag{7}$$

where $(L - I)_i$ denotes the matrix obtained by zeroing all except the ith row of $L - I$.

Lemma 3. *When $\boldsymbol{u}, \boldsymbol{v} \geq \mathbf{0}$, we have*

$$\frac{\partial^2}{\partial s \partial t} \tilde{F}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v}) \leq 0 \tag{8}$$

wherever $\mathbf{0} < \boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v} < \mathbf{1}$.

Proof. We begin by rewriting $\tilde{F}$ in a symmetric form:

$$\begin{aligned}
\tilde{F}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v}) &= \log \det(\operatorname{diag}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v})(L - I) + I) && (9)\\
&= \log \det(\operatorname{diag}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v})) + \log \det(L - I + \operatorname{diag}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v})^{-1}) && (10)\\
&= \log \det(D) + \log \det(M) , && (11)
\end{aligned}$$

where $D(s,t) = \operatorname{diag}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v})$ and $M(s,t) = L - I + D^{-1}(s,t)$. Note that $D, M \succ 0$, since $\mathbf{0} < \boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v} < \mathbf{1}$. We have

$$\frac{\partial}{\partial t} \tilde{F}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v}) = \operatorname{tr}(D^{-1}(s,t) \operatorname{diag}(\boldsymbol{v}) - M^{-1}(s,t) D^{-2}(s,t) \operatorname{diag}(\boldsymbol{v})) . \tag{12}$$

Taking the second derivative with respect to s,

$$\begin{aligned}
\frac{\partial^2}{\partial s \partial t} \tilde{F}(\boldsymbol{x} + s\boldsymbol{u} + t\boldsymbol{v}) = \operatorname{tr}(&-D^{-2}(s,t) \operatorname{diag}(\boldsymbol{v}) \operatorname{diag}(\boldsymbol{u}) + 2M^{-1}(s,t) D^{-3}(s,t) \operatorname{diag}(\boldsymbol{v}) \operatorname{diag}(\boldsymbol{u})\\
&- M^{-1}(s,t) D^{-2}(s,t) \operatorname{diag}(\boldsymbol{u}) M^{-1}(s,t) D^{-2}(s,t) \operatorname{diag}(\boldsymbol{v})) . \qquad (13)
\end{aligned}$$

Since diagonal matrices commute and $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, the above is equal to $-\operatorname{tr}(SS^\top) \leq 0$, where

$$S = D^{-1}(s,t) \operatorname{diag}(\sqrt{\boldsymbol{v}}) \operatorname{diag}(\sqrt{\boldsymbol{u}}) - D^{-1}(s,t) \operatorname{diag}(\sqrt{\boldsymbol{v}}) M^{-1}(s,t) \operatorname{diag}(\sqrt{\boldsymbol{u}}) D^{-1}(s,t) . \tag{14}$$

(Note that S is defined since $\boldsymbol{u}, \boldsymbol{v} \geq \mathbf{0}$.) □

Corollary 4. *$\tilde{F}(\boldsymbol{x}+t\boldsymbol{v})$ is concave along any direction $\boldsymbol{v} \geq \mathbf{0}$ (equivalently, $\boldsymbol{v} \leq \mathbf{0}$).*

Lemma 5. *If $\boldsymbol{x}$ is a local optimum of $\tilde{F}(\cdot)$, then for any $\boldsymbol{y} \in [0,1]^N$,*

$$2\tilde{F}(\boldsymbol{x}) \geq \tilde{F}(\boldsymbol{x} \vee \boldsymbol{y}) + \tilde{F}(\boldsymbol{x} \wedge \boldsymbol{y}) \;, \tag{15}$$

where $(\boldsymbol{x} \vee \boldsymbol{y})_i = \max(x_i, y_i)$ and $(\boldsymbol{x} \wedge \boldsymbol{y})_i = \min(x_i, y_i)$.

Proof. By definition, $\boldsymbol{x} \vee \boldsymbol{y} - \boldsymbol{x} \geq 0$ and $\boldsymbol{x} \wedge \boldsymbol{y} - \boldsymbol{x} \leq 0$. By Corollary 4 and the first order definition of concavity,

$$\nabla \tilde{F}(\boldsymbol{x})^\top (\boldsymbol{x} \vee \boldsymbol{y} - \boldsymbol{x}) \geq \tilde{F}(\boldsymbol{x} \vee \boldsymbol{y}) - \tilde{F}(\boldsymbol{x}) \tag{16}$$

$$\nabla \tilde{F}(\boldsymbol{x})^\top (\boldsymbol{x} \wedge \boldsymbol{y} - \boldsymbol{x}) \geq \tilde{F}(\boldsymbol{x} \wedge \boldsymbol{y}) - \tilde{F}(\boldsymbol{x}) \;. \tag{17}$$

Adding the two equations gives the desired result, given that $\nabla \tilde{F}(\boldsymbol{x})^\top (\boldsymbol{z} - \boldsymbol{x}) \leq \mathbf{0}$ for any $\boldsymbol{z} \in S$ at a local optimum $\boldsymbol{x}$. □

Let $\mathcal{X}_i \subseteq [0,1]$ be a subset of the unit interval representing $x_i = |\mathcal{X}_i|$, where $|\mathcal{X}_i|$ denotes the measure of $\mathcal{X}_i$. $\tilde{F}^*$ is defined on $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_N)$ by

$$\tilde{F}^*(\mathcal{X}) = \tilde{F}(\boldsymbol{x}), \quad \boldsymbol{x} = (|\mathcal{X}_1|, |\mathcal{X}_2|, \ldots, |\mathcal{X}_N|) \;. \tag{18}$$

Lemma 6. *$\tilde{F}^*$ is submodular.*

Proof. We first show that for $\boldsymbol{x} \leq \boldsymbol{y}$ and $\boldsymbol{a} \geq \mathbf{0}$, $\tilde{F}(\boldsymbol{x}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{x}) \geq \tilde{F}(\boldsymbol{y}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{y})$. By the fundamental theorem of calculus,

$$\tilde{F}(\boldsymbol{x}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{x}) = \int_0^1 \frac{\partial}{\partial t} \tilde{F}(\boldsymbol{x} + t\boldsymbol{a}) \, dt \;, \tag{19}$$

and by a second application,

$$(\tilde{F}(\boldsymbol{y}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{y})) - (\tilde{F}(\boldsymbol{x}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{x})) = \int_0^1 \int_0^1 \frac{\partial^2}{\partial s \partial t} \tilde{F}(\boldsymbol{x} + s(\boldsymbol{y}-\boldsymbol{x}) + t\boldsymbol{a}) \, dt \, ds \;. \tag{20}$$

Since $\boldsymbol{y} - \boldsymbol{x} \geq \mathbf{0}$, Corollary 4 allows us to conclude that the second derivatives are nonpositive.

Now, for $\mathcal{X} \subseteq \mathcal{Y}$ and $\mathcal{A} \cap \mathcal{Y} = \emptyset$ where $\mathcal{X}, \mathcal{Y}, \mathcal{A}$ represent $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{a}$, respectively, we have

$$\tilde{F}^*(\mathcal{X} \cup \mathcal{A}) - \tilde{F}^*(\mathcal{X}) = \tilde{F}(\boldsymbol{x}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{x}) \geq \tilde{F}(\boldsymbol{y}+\boldsymbol{a}) - \tilde{F}(\boldsymbol{y}) = \tilde{F}^*(\mathcal{Y} \cup \mathcal{A}) - \tilde{F}^*(\mathcal{Y}) \;. \tag{21}$$

□

Lemmas 5 and 6 suffice to prove the following theorem, which appears for the multilinear extension in [1], bounding the approximation ratio of our algorithm.

Theorem 7. *Let $\tilde{F}(\boldsymbol{x})$ be the softmax extension of a nonnegative submodular function $f(Y) = \log\det(L_Y)$, let $\mathrm{OPT} = \max_{\boldsymbol{x} \in S} \tilde{F}(\boldsymbol{x})$, and let $\boldsymbol{x}$ and $\boldsymbol{z}$ be local optima of $\tilde{F}$ in S and $S \cap \{\boldsymbol{y} \mid \boldsymbol{y} \leq \mathbf{1} - \boldsymbol{x}\}$, respectively. Then*

$$\max(\tilde{F}(\boldsymbol{x}), \tilde{F}(\boldsymbol{z})) \geq \frac{1}{4}\mathrm{OPT} \geq \frac{1}{4} \max_{Y \in S} \log\det(L_Y) \;. \tag{22}$$

We omit the proof since it is unchanged from [1].

Corollary 8. *Algorithm 2 yields a 1/4-approximation to the DPP MAP problem whenever $\log\det(L_Y) \geq 0$ for all Y. In general, the objective value obtained by Algorithm 2 is bounded below by $\frac{1}{4}(OPT - p_0) + p_0$, where $p_0 = \min_Y \log\det(L_Y)$.*

Theorem 9. *If $S = [0,1]^N$, then for any local optimum $\boldsymbol{x}$ of $\tilde{F}$, either $\boldsymbol{x}$ is integral or at least one fractional coordinate x_i can be set to 0 or 1 without lowering the objective.*

Proof. Note that by multilinearity of the determinant, $\det(\operatorname{diag}(\boldsymbol{x})(L - I) + I)$ is linear in each coordinate x_i if all the other coordinates $\boldsymbol{x}_{-i}$ are held fixed. That is, $\tilde{F}(x_i, \boldsymbol{x}_{-i}) = \log(ax_i + b)$, where a and b depend on $\boldsymbol{x}_{-i}$. Suppose that coordinate i is fractional ($0 < x_i < 1$) at a local optimum, then the gradient with respect to x_i must be zero, since the polytope constraint is not active. Since $\frac{\partial \tilde{F}(\boldsymbol{x})}{\partial x_i} = \frac{a}{ax_i + b}$, this is only possible if $a = 0$. Hence setting x_i to 0 or 1 does not affect the objective. □

B Matched summarization example

Figure 1 contains an example set of pairs of statements made by candidates Paul and Romney in the 2012 US Republican primary debates.

	Paul	**Romney**
1	Well, it's a tragedy because this is a consequence of the government being involved in medicine since 1965 ... When the government gets involved in medicine, you don't get better care; you get – cost goes up and it distorts the economy and leads to a crisis ...	... what's wrong with our health care system in America is that government is playing too heavy a role ... 18 percent of our GDP is spent on health care. The next highest nation in the world is 12 percent. ...
2	... Social Security is broke. We spent all the money and it's on its last legs unless we do something ... Now, what I would like to do is to allow all the young people to get out of Social Security and go on their own ...	... Social Security is a responsibility of the federal government, not the state governments, that we're going to have one plan, and we're going to make sure that it's fiscally sound and stable. And I'm absolutely committed to keeping Social Security working ...
3	... I was fighting over a decade to try to explain to people where the housing bubble was coming from. So Freddie Mac is bailed out by the tax payers ... and they're still getting bailed out ...	I look at Fannie and Freddie and just think that obviously they've grown massively beyond the scope that had been envisioned for them originally ... we have to rethink about how we're going to support a growing housing industry.
4	... the people are not ready. We don't have any money. We have too many wars. We – the people want to come home and they certainly don't want a hot war in Iran right now and I – I think that would be the most foolish thing in the world to do right now is take on Iran.	... the president is building roughly nine ships a year. We ought to raise that to 15 ships a year ... We want to show Iran, any action of that nature will be considered an act of war, an act of terror and – and America is going to be keep those sea lanes open.
5	I strongly supported Ronald Reagan ... But in the 1980s, we spent too much, we taxed too much, we built up our deficits, and it was a bad scene ...	you talk about ... the Reagan revolution and the jobs created during the Reagan years and so forth ... But you know what? The free people of America, pursuing their dreams ... those are the people that make America strong, not Washington.
6	... I also resent the fact that illegals come into this country ... I have a strong position on immigration. I don't think that we should give amnesty and they become voters ...	... we let people come across our border illegally or stay here and overstay their visa ... we have to secure our border and crack down on those that bring folks here and hire here illegally.
7	... this administration already has accepted the principle that, when you assume somebody is a terrorist, they can be targeted for assassination, even American citizens ... You don't want to translate our rule of law into a rule of mob rule.	... We have a Constitution and we follow the law ... Our nation was founded on a principal of religious tolerance ... we treat people with respect regardless of their religious persuasion.

Figure 1: An example of the type of sets of pairs our method selects. (The quotes are actually several sentences longer than what is shown here, but for brevity, some information is elided with ellipses.) Qualitatively, this example set is appealing in that it matches our intuition of what maximizing the objective should achieve. That is, this match is both diverse — contains quotes on a variety of topics, like healthcare, taxes, jobs — and at the same time maintains good match quality — the quotes within each pair are frequently on the same topic.

References

[1] C. Chekuri, J. Vondrák, and R. Zenklusen. Submodular Function Maximization via the Multilinear Relaxation and Contention Resolution Schemes. *arXiv:1105.4593*, 2011.